CORC REPORT 2002-01

Subset Algebra Lift Operators for 0-1 Integer Programming (Extended version)*

Daniel Bienstock and Mark Zuckerberg
Columbia University
New York, NY 10027

July, 2002
version 2003-Dec-10

Abstract

We extend the Sherali-Adams, Lovász-Schrijver, Balas-Ceria-Cornuéjols and Lasserre lift-and-project methods for 0-1 optimization by considering liftings to subset algebras. Our methods yield polynomial-time algorithms for solving a relaxation of a set-covering problem at least as strong as that given by the set of all valid inequalities with small coefficients, and, more generally, all valid inequalities where the right-hand side is not very large relative to the positive coefficients in the left-hand side. Applied to generalizations of vertex-packing problems, our methods yield, in polynomial time, relaxations that have unbounded rank using for example the N_+ operator.

1 Introduction

Consider a 0-1 integer programming problem

$$\min\{c^T x \,:\, x \in \mathcal{F}\},$$

where

$$\mathcal{F} \;=\; \{\, x \in \{0,1\}^n \;:\; Ax \geq b \,\}. \tag{1}$$

The procedures in [SA90], [LS91], [L01b] and [BCC93] solve this problem by iteratively strengthening its continuous relaxation, until, after at most n iterations, the convex hull of $\mathcal{F}$ is obtained. This bound on the number of iterations is tight ([CD01], [L01], also see [GT01], and [CL01] for related topics). Nevertheless, a question of theoretical and practical interest is whether it is possible to modify the procedures so that the earlier iterations produce stronger relaxations.

As shown in [L01], the methods used in [SA90], [LS91], [L01b], and [BCC93] can be viewed as relying on a common paradigm: that of "lifting" a point in $\{0,1\}^n$ to an appropriate zeta-vector of the subset lattice of $\{1, 2, \cdots, n\}$.

In this paper we introduce operators that lift instead to the (much larger) subset *algebra* of $\{0,1\}^n$. One example of a result which is derived using our operators is the following:

Theorem 1.1 *Let $k \geq 1$ be a fixed integer. Consider a set-covering problem*

$$\min\{c^T x \,:\, Ax \geq e,\; x \in \{0,1\}^n\,\},$$

where A is 0-1 and e is a vector of 1s. Let V_k denote the set of inequalities $a^T x \geq a_0$ which are valid for $\{x \in \{0,1\}^n \,:\, Ax \geq e\}$ and for which $a_j \in \{0, 1, 2, \cdots, k\}$, $0 \leq j \leq n$. Let $\mathcal{V}_k$ denote the set of points

*This research was partially funded by NSF awards 29132-5538 (CORC) and DMI-0200221.

in $\{0,1\}^n$ *that satisfy all inequalities in* V_k*. Then there is an algorithm of complexity polynomial in* n*, for solving*

$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & x \in \mathcal{R}_k \end{aligned}$$

where $\mathcal{R}_k$ *is a certain polyhedron satisfying*

$$\{x \in \{0,1\}^n \,:\, Ax \geq e\} \subseteq \mathcal{R}_k \subseteq conv(\mathcal{V}_k).$$

In other words, $\mathcal{R}_k$ is a polyhedral relaxation of the set-covering problem, all of whose points satisfy all inequalities in the set V_k.

Theorem 1.1 is a special case of a more general result. Given an inequality $a^T x \geq a_0$ with indices ordered so that $0 < a_1 \leq a_2 \leq \cdots \leq a_J$ and $a_j = 0$ for $j > J$, its *pitch* is the minimum t such that $\sum_{j=1}^t a_j \geq a_0$. Then we have:

Theorem 1.2 *Let* $k \geq 1$ *be a fixed integer. Consider a set-covering problem*

$$\min\{c^T x \,:\, Ax \geq e,\ x \in \{0,1\}^n\},$$

where A *is an* $m \times n$*, 0-1 matrix and* e *is the vector of* m 1*s. Let* P_k *denote the set of all valid inequalities for* $\{x \in \{0,1\}^n \,:\, Ax \geq e\}$ *of pitch* $\leq k$*. Then there a positive integer* $g(k)$*, a polytope* $\mathcal{Q}_k \subseteq R^n$ *and a polytope* $\bar{\mathcal{Q}}_k \subseteq R^{(m+n)^{g(k)}}$ *satisfying:*

- *(a)* $\{x \in \{0,1\}^n \,:\, Ax \geq e\} \subseteq \mathcal{Q}_k$*,*
- *(b)* $a^T x \geq a_0, \quad \forall x \in \mathcal{Q}_k$ *and* $\forall (a, a_0) \in P_k$*,*
- *(c)* $\mathcal{Q}_k$ *is the projection to* R^n *of* $\bar{\mathcal{Q}}_k$*,*
- *(d)* $\bar{\mathcal{Q}}_k$ *can be described by a system of at most* $(m+n)^{g(k)}$ *linear constraints, with integral coefficients of absolute value at most* k*. This system can be computed in time polynomial in* n *and* m *for fixed* k*.*

Thus, Theorem 1.2 introduces a natural hierarchy among the valid inequalities for set-covering problems, and shows that any fixed level in the hierarchy can be satisfied in polynomial time. How strong are the pitch $\leq k$ inequalities? First, note that the hierarchy is "complete" in the sense that any facet-defining inequality has pitch $\leq n$. Further, one can produce examples of set-covering problems with exponentially many facets with coefficients $0, 1, 2$ (Balas and Ng have completely characterized the set of facet-defining inequalities with coefficients $0, 1, 2$ [BN89]). Also, there are examples of Gomory inequalities of rank greater than one with coefficients ≤ 3, and we conjecture that for each fixed integer $k' > 0$ there is a $k = k(k')$ such that there are examples of set-covering problems where some of the inequalities of pitch $\leq k$ have (fractional) Gomory rank $\geq k'$.

At the same time, we show that given a set-covering problem with a full-circulant constraint matrix ($\sum_{j \neq i} x_j \geq 1$ for each $1 \leq i \leq n$) the valid inequality $\sum_j x_j \geq 2$ (a constraint of pitch 2) has rank at least $n-3$ for a lifting operator stronger than the Sherali-Adams and the N_+ procedures combined. In other words (if somewhat incompletely) lifting to a polynomially larger space yields an operator that is exponentially stronger. These results, and others on set-covering problems, are presented in Sections 3.3. and 4.4.

The hierarchy of algorithms that we present (for fixed k we call our level k procedure the Σ^k*-algorithm*, and we define these for $k \geq 2$) do not require positive-semidefiniteness in order to achieve Theorem 1.2. On the other hand, one of the key results in [LS91] is that positive-semidefiniteness allows the N_+ operator to guarantee that classical inequalities, for example clique inequalities in vertex-packing problems, are achieved with rank 1. In Section 4.5 we consider set packing problems where, given a graph G, and given a family of pairwise disjoint index sets S_i, $i \in V(G)$, for every edge $\{i, j\}$ we have the constraint

$$\sum_{k \in S_i} x_k + \sum_{k \in S_j} x_k \quad \leq \quad |S_i| + |S_j| - 1. \tag{2}$$

When $|S_i| = 1$ for each i we obtain precisely the vertex-packing polyhedron of G. Thus, one can think of generalizations of several classical valid inequalities (clique, odd-hole, odd-antihole, odd-wheel). For example, when G contains a clique with vertex set K the resulting *set-clique* constraint is:

$$\sum_{i \in K} \sum_{k \in S_i} x_k \le \sum_{i \in K} |S_i| - |K| + 1.$$

One can similarly define set-odd-hole, set-odd-antihole and set-odd-wheel inequalities. In Section 4.5 we show:

Theorem 1.3 *Suppose we have a formulation that contains inequalities (2) corresponding to a clique K (with pairwise disjoint sets S_i). Then the Σ^2-algorithm generates a vector that satisfies the set-clique constraint. On the other hand, set-clique inequalities have unbounded N_+- (and Sherali-Adams) rank. Similarly, the Σ^2-algorithm generates a vector that satisfies all set-odd-hole, set-odd-antihole and set-odd-wheel constraints.*

Finally, section 4.6 presents further comparisons between our algorithms and the N_+ procedure.

The development of our algorithms starts in Section 3.

2 Background

Let $\mathcal{F} \subseteq \{0,1\}^n$ be as before. Lovász and Schrijver [LS91] introduced the following general paradigm for the problem of separating over $conv(\mathcal{F})$.

Let $N \gg n$, and suppose we have a function that maps ("lifts") each $v \in \mathcal{F}$ into $z = z(v) \in \{0,1\}^N$ with $z_j = v_j$, $1 \le j \le n$. Let $\hat{\mathcal{F}} \subseteq \{0,1\}^N$ denote the image of $\mathcal{F}$ under this mapping. Then given $x \in R^n_+$, the question of whether $x \in conv(\mathcal{F})$ is equivalent to answering whether there exists $y \in conv(\hat{\mathcal{F}})$ such that $y_i = x_i$, $1 \le i \le n$.

This second membership question may be easier to answer than the original one because the vectors $z(v)$ reveal information about $\mathcal{F}$ in a more explicit way. As pointed out in [LS91], this basic idea was implicit in earlier work on specific combinatorial problems, see [BP83], [B93], [BP89], and others. Furthermore, as it turns out, the much earlier work of Balas on disjunctive programming [B75], [B79] provides a common underlying viewpoint for much of this work.

Writing $E_n = \{1, 2, \cdots, n\}$, the concrete application of this idea in [LS91] is as follows. We map each $v \in \{0,1\}^n$ into $\hat{v} \in \{0,1\}^{2^n}$, where

(i) the entries of $\hat{v}$ are indexed by subsets of E_n, and

(ii) For $S \subseteq \{1, \ldots, n\}$, $\hat{v}_S = 1$ iff $v_j = 1$ for all $j \in S$.

Clearly, for $1 \le j \le n$, $\hat{v}_{\{j\}} = v_j$, so each $v \in \{0,1\}^n$ is mapped into a distinct column of the zeta matrix $\mathcal{Z}$ of the subset lattice L of E_n (see [R64]). For simplicity, we will forgo the standard lattice-theoretic notation ($\le$, $\vee$, $\wedge$) and use the corresponding set-theoretic operators instead ($\subseteq$, $\cap$ and $\cup$), and identify elements of the lattice with subsets of E_n.

For completeness, we state the definition of $\mathcal{Z}$: it has a row and a column for each element of L (i.e., each subset of E_n), and for a given $p \in L$ its corresponding column ζ^p is defined by:

$$\zeta^p_q = \begin{cases} 1 & \text{if } q \subseteq p, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Note that each column of $\mathcal{Z}$ contains an entry (of value 1) for the empty set $\emptyset$; we will assume that this is the *zeroth* coordinate and usually indicate it as such, but may sometimes use $\emptyset$ instead. Assuming an appropriate ordering of columns, $\mathcal{Z}$ is upper triangular with 1s along the main diagonal, and therefore invertible. Its inverse, $\mathcal{M}$, is called the Möbius matrix of the lattice.

Let $y \in R^L$. Then since $\mathcal{Z}$ is invertible we can write $y = \sum_r \lambda_r \zeta^r$ for unique reals λ_r, $r \in L$. Thus, since $\mathcal{M} = \mathcal{Z}^{-1}$, we have $\lambda_r = m_r y$, where for $p \in L$ we denote by m_p the corresponding row of $\mathcal{M}$. In other words, we can completely characterize $conv(\hat{\mathcal{F}})$ as follows:

$$conv(\hat{\mathcal{F}}) \;=\; \{y \in R^L \,:\, m_p y \geq 0 \;\; \forall p \in \hat{\mathcal{F}};\;\; m_p y = 0 \;\; \forall p \notin \hat{\mathcal{F}};\;\; e^T \mathcal{M} y = 1\}. \tag{4}$$

We can summarize what we know so far by considering the following questions, given $y \in R^L$:

(A) Is y a linear combination of the columns of $\mathcal{Z}$? This is always true because $\mathcal{Z}$ is invertible.

(B) Is y an affine combination of the columns of $\mathcal{Z}$? This requires $y_\emptyset = 1$.

(C) Is y a convex combination of the columns of $\mathcal{Z}$? This requires $y_\emptyset = 1$ and $\lambda = My \geq 0$.

(D) Is $y \in conv(\hat{\mathcal{F}})$? This requires $y_\emptyset = 1$, $My \geq 0$, and $m_r y = 0$ for all $r \notin \hat{\mathcal{F}}$.

We stress that (C) is already a nontrivial requirement (since $\mathcal{Z}$ is invertible and therefore λ is unique).

Even though condition (4) completely determines $conv(\hat{\mathcal{F}})$, it is algorithmically cumbersome – it requires that we handle exponentially large matrices and vectors. [LS91], [SA90], [L01b] provide methods to approximate this condition while only considering lower dimensional lattice elements. The procedure in [BCC93], which has achieved some computational success [BCC96], is related to earlier work by Balas (see [B75]) and can be viewed as a simplified version of the procedures in [LS91] and [SA90] – it relies on convexifying one variable at a time. Here we outline the approach in [LS91].

Suppose $y = \sum_{r \in L} \lambda_r \zeta^r$. Consider the $2^n \times 2^n$-matrix W^y defined by

$$W^y \;=\; \sum_{r \in L} \lambda_r \zeta^r (\zeta^r)^T. \tag{5}$$

We have:

$$\lambda \geq 0 \quad iff \quad W^y \succeq 0. \tag{6}$$

This fact is clear in one direction; for the other implication note that $\lambda_r = m_r^T W^y m_r$ for all r.

Hence, the condition $\lambda \geq 0$ may be approximated by requiring that some "small" (e.g., polynomial-sized) minor of W^y be positive-semidefinite. Thus, it is of interest to approximate small minors of W^y without generating W^y itself.

The approach in [LS91] approximates the $(n+1) \times (n+1)$ leading minor of W^y, as follows. Given $\bar{x} \in R^n_+$, if $\bar{x} \in conv(\mathcal{F})$ then by (5) we can lift $\bar{x}$ to an $(n+1) \times (n+1)$-matrix $M^{\bar{x}}$ with rows and columns indexed by singletons and the empty set, such that:

(a) $M^{\bar{x}} \succeq 0$,

(b) $M^{\bar{x}}$ is symmetric,

(c) The zeroth row of $M^{\bar{x}}$ is equal to its diagonal, and

(d) The zeroth row of $M^{\bar{x}}$ is $(1, \;\; \bar{x}^T)$.

Even though when $\bar{x} \in conv(\mathcal{F})$ such a lifting exists, it is not necessarily the case that any matrix $M^{\bar{x}}$ satisfying (a)-(d) is a minor of a matrix the form W^y for some lifting $y \in R^L$ of $\bar{x}$. In fact, there are other structural properties that any such W^y has to satisfy which can also be required of $M^{\bar{x}}$.

Consider one of the constraints $a_i^T x \geq b_i$ in the definition of $\mathcal{F}$. Suppose $\bar{x} \in conv(\mathcal{F})$, and let w be the k^{th} column of $M^{\bar{x}}$, $0 \leq k \leq n$. Then it is easy to see that w satisfies the (homogenized) constraint

$$(-b_i, a_i^T) w \;=\; \sum_{j=1}^{n} a_{ij} w_j - b_i w_0 \geq 0. \tag{7}$$

This is clear if $k = 0$ by (d) and (b), and for $k \geq 1$ use the fact that the k^{th} column in each of the terms in (5) satisfy the constraint (7).

Further, using basic properties of lattices and the Möbius matrix one can also show that the vector obtained by subtracting any column of $M^{\bar{x}}$ from the zeroth column also satisfies each homogenized constraint. This can also be directly obtained [Z03] by expanding the formulation to include the columns $x'_j = 1 - x_j$ $(1 \le j \le n)$ and studying the corresponding W and M matrices.

We summarize these facts as

(e) Let w^k indicate the k^{th} column of $M^{\bar{x}}$. Then $(-b, A)w^k \ge 0$ for $0 \le k \le n$, and $(-b, A)(w^0 - w^k) \ge 0$ for $1 \le k \le n$.

In [LS91] the lifting $\bar{x} \to M^{\bar{x}}$ required to satisfy conditions (b)-(e) is denoted by M. If, in addition, we require (a) then the lifting is denoted by M_+. Or, more precisely, we may think of M (or M_+) as describing operators: if we start with

$$Q = \{x \in [0,1]^n \ : \ Ax \ge b\}$$

then we can define $N(Q)$ (resp., $N_+(Q)$) as that subset of Q for which a lifting M^x exists satisfying (b)-(e) (resp., (a)-(e)). Clearly $N_+(Q) \subseteq N(Q) \subseteq Q$, and both $N_+(Q)$ and $N(Q)$ are convex sets (a polytope in the second case). As shown in [LS91] after iterating n times we have $N^n(Q) = conv(\mathcal{F})$. In fact, the operator in [BCC93], which is weaker than N, also requires at most n iterations.

While these operators all require, in the worst case, the same number of iterations, it is clearly important to study their relative strength, i.e., how comparatively tight a relaxation they produce. In this regard, there is an additional critical property that is satisfied by the matrix W^y (see [LS91] for references):

$$W^y_{p,q} \quad = \quad y_{p \cup q}, \quad \forall \ p, \ q \in L. \tag{8}$$

Thus, *every* entry of W^y can be found in its zeroth row (or column), and, in general, there are nontrivial relationships between the entries appearing in any minor of W^y of size greater than $n+1$. In general, by approximating W^y with a minor restricted to lattice elements of cardinality $\le k$ we are able to make some statements about coordinates of y corresponding to lattice elements of cardinality $k+1$ or larger.

The procedures in [SA90], [L01b] take advantage of this fact and introduce some further ideas which we discuss next. [L01] has shown how to cast these methods in the general framework we have been using, although originally they were presented quite differently. In addition, they apply to more general problems than linear integer programs, but here we will restrict attention to the linear case.

As shown in [L01] there is a common underlying theme to the algorithms in [SA90] and [L01b]. Let $a_i^T x \ge b_i$ be once more one of the inequalities in $Ax \ge b$. Define $\bar{a}_i \in R^L$ by $\bar{a}_i^T \ = \ (-b_i, a_i^T, 0, 0, \cdots, 0)$ where the number of appended 0s equals $2^n - n - 1$, i.e., we append to a_i a zero for each element of L of cardinality greater than 1. Suppose again that $y = \sum_{r \in L} \lambda_r \zeta^r$. Then

$$\bar{a}_i * y \quad \doteq \quad W^y \bar{a}_i \tag{9}$$

satisfies

$$\bar{a}_i * y \quad = \quad \sum_{r \in L} \lambda_r \zeta^r (\zeta^r)^T \bar{a}_i. \tag{10}$$

If, in addition, y is a lifting of $\bar{x} \in conv(\mathcal{F})$, then the sum in (10) can be restricted to elements of $\hat{\mathcal{F}}$, and each such term r satisfies $(\zeta^r)^T \hat{a}_i \ge 0$ (note: a similar idea is implicit in the $N(K, K)$ operator in [LS91]). Consequently, $\bar{a}_i * y$ is a nonnegative linear combination of columns of the zeta matrix. We may summarize this fact as another condition to be satisfied by x:

(f) $W^{\bar{a}_i * y} \succeq 0$ for each constraint $a_i^T x \ge b_i$ in $Ax \ge b$.

Of course, (f) involves y, not x, but notice that by definition of $\bar{a}_i$, we only need the first $n+1$ columns of W^y in order to compute $\bar{a}_i * y$. Hence (through another application of (8)) condition (f) may be approximated by requiring positive semidefiniteness of appropriate minors of $W^{\bar{a}_i * y}$.

As shown in [L01], round $t \ge 1$ of the Sherali-Adams procedure requires that for each $U \subseteq E_n$ with $|U| \le \min\{t+1, n\}$, the minor of W^y corresponding to the set of rows and columns arising from all subsets of U be positive-semidefinite, and that for each $U \subseteq E_n$ with $|U| \le t$, the minor of $W^{\bar{a}_i * y}$ corresponding to the set of rows and columns arising from all subsets of U be positive-semidefinite. In contrast, round t of the Lasserre procedure requires the stronger condition that the minor of W^y corresponding to the set of rows and columns arising from *all* subsets of E_n of cardinality $\leq \min\{t+1, n\}$ be positive-semidefinite, and that the same holds for the minor of $W^{\bar{a}_i * y}$ arising from all subsets of cardinality $\leq t$. Further, the Lasserre operator is stronger than the Sherali-Adams operator and than the N_+ operator, whereas the Sherali-Adams operator is stronger than the N operator (but not N_+). It seems likely that the Lasserre operator is in fact far stronger than the Sherali-Adams operator.

2.1 Probability measures

[LS91] introduces an additional, very useful idea. In order to describe this idea we need to review the definition of a probability measure. The definition we present below is slightly imprecise and economizes on notation; see [F66] for formal details.

Definition 2.1 *Let $\mathcal{W}$ be a set. A probability measure is a function $\Upsilon : 2^{\mathcal{W}} \rightarrow R$ satisfying the following properties:*

(i) $\Upsilon(A) \geq 0$ *for all* $A \subseteq \mathcal{W}$*,*

(ii) $\Upsilon(\mathcal{W}) = 1$*, and*

(iii) For all disjoint subsets A*,* B *of* $\mathcal{W}$*,* $\Upsilon(A \cup B) = \Upsilon(A) + \Upsilon(B)$*.*

Note that (i)-(iii) imply that $\Upsilon(\emptyset) = 0$, and that Υ is nondecreasing.

The following result is stated in [LS91] (p. 186, "Remark"). This topic has been studied in some detail in [DL97]. In particular, it appears that ideas in this direction existed prior to the Lovász-Schrijver paper [LS91]. A proof of the "Remark" using the viewpoint in this paper appears in [Z03].

Theorem 2.2 *Let L be the subset lattice of E_n. Suppose $z \in R^L$. Then z is a convex combination of the columns of $\mathcal{Z}$ iff there exists*

(a) a probability measure Υ on some (abstract) set $\mathcal{W}$, and

(b) a family of subsets $\{I_j \subseteq \mathcal{W} : 1 \leq j \leq n\}$,

such that

$$\forall r \subseteq E_n, \quad \Upsilon(\bigcap_{j \in r} I_j) = z_r. \tag{11}$$

When the conditions of the theorem apply to $z \in R^L$ we will say that z is *measure consistent*. There are several useful observations to be made here. First, the set $\mathcal{W}$ is abstract – we are free to choose it (and the probability measure Υ) as convenient. For example, we might choose $\mathcal{W} = R^1$.

Second, consider again the set $\mathcal{F} \subseteq \{0,1\}^n$ and its lifting $\hat{\mathcal{F}} \subseteq R^L$ as we have been discussing above. Given $\bar{x} \in R^n$, suppose indeed we can lift it to a $y \in R^L$ that is measure consistent. Then the measure Υ and sets I_J produced by Theorem 2.2 (applied to y) satisfy $\Upsilon(I_j) = \bar{x}_j$, $1 \leq j \leq n$.

Hence, we may think of I_j as representing the hyperplane $\{v \in R^n : v_j = 1\}$, or, more precisely, the intersection of this hyperplane with $conv(\mathcal{F})$. In other words, if $\bar{x} \in R^n$ we may think of $\bar{x}_j$ $(1 \leq j \leq n)$ as stating the probability of being in the set I_j.

This point can be pursued further. Suppose we are trying to construct a function Υ so as to prove that y is measure consistent. In addition to being a probability measure, the only structural condition to be satisfied by Υ is (11). Even though this condition appears simple, a large number of additional conditions are implied by it. In fact, there is a condition that can be stated for each element of the algebra generated by the sets I_j [F66]. This is the starting point for our work.

3 New results

3.1 Preliminaries

Roughly speaking, the lifting operators we will describe below lift any point of $\mathcal{F}$ to a zeta-vector of the subset algebra of $\{0,1\}^n$, viewed as a lattice. A completely formal definition of a subset algebra is beyond the scope of this paper (see [C74]), but the following should suffice.

Let S be a finite set, and suppose $\{A_j,\, 1 \le j \le k\}$ is a collection of subsets of S. Let $\bar{A}_j = S - A_j$ for each j.

Definition 3.1 *The subset algebra $\Sigma = \Sigma(A_1, \cdots, A_k)$ generated by $A_1, \cdots, A_k$ is the set of all subsets of S that can be obtained from set-theoretic expressions involving the $A_1, \cdots, A_k$ and $\bar{A}_1, \cdots, \bar{A}_k$.*

Thus, expressions of the form $A_1 \cup (A_2 \cap \bar{A}_3))$, etc, are in $\Sigma(A_1, \cdots, A_k)$.

Note that $\Sigma(A_1, \cdots, A_k)$ can be viewed as contained in 2^S.

Definition 3.2 *Let $J \subseteq \{1, 2, \cdots k\}$. The subset*

$$(\bigcap_{j \in J} A_j) \cap (\bigcap_{j \notin J} \bar{A}_j)$$

is called an atom (also called a Boolean function, or a complete product).

It can be shown that every element of Σ can be written as a (finite) union of atoms. Thus, Σ has at most 2^{2^k} distinct elements.

Example 3.3 *Consider $S = \{0,1\}^n$. For each $1 \le j \le n$, let $H_j \doteq \{x \in \{0,1\}^n \,:\, x_j = 1\}$. Then the subset algebra generated by the H_j is exactly the set of all subsets of $\{0,1\}^n$. Each atom corresponds to a distinct point in $\{0,1\}^n$, and there are 2^{2^n} elements in the algebra.*

Σ is also a lattice. In our case we will use the reverse of the inclusion order; i.e., we will declare $b \le a$ when $a \subseteq b$. That we have a lattice follows since Σ is closed under unions and intersections, and hence $\vee$ and $\wedge$ are well-defined. For convenience, we will also denote this lattice by Σ. To avoid confusion with the standard subset lattice, we will use ξ to denote a zeta-vector of this lattice. Thus, if $p \in \Sigma$, the ξ-vector for p is the vector $\xi \in \{0,1\}^\Sigma$ defined by

$$\xi^p_q = \begin{cases} 1 & \text{if } p \subseteq q, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

where we use "$\subseteq$" whenever the underlying subsets of S satisfy the relation. In our development of lift operators we will work with (slight) generalizations of this type of lattice, as indicated next.

As a final note, when Σ has a member for each subset of S, then, as a lattice, Σ is isomorphic to the subset lattice of S (with the order relationship reversed). Even in such a case, however, it is useful to view the lattice as generated by the A_j.

3.2 Lifting to an algebra

We now return to the feasible region (1) of a 0-1 integer program, repeated here for convenience:

$$\mathcal{F} = \{\, x \in \{0,1\}^n \;:\; Ax \ge b \,\}.$$

As before, we will use the following notation. For $1 \le j \le n$ denote $H_j \doteq \{x \in \{0,1\}^n \,:\, x_j = 1\}$, and

$$\mathrm{Y}_j \,=\, \mathcal{F} \cap H_j, \qquad \mathrm{N}_j \,=\, \bar{\mathrm{Y}}_j \,=\, \mathcal{F} - \mathrm{Y}_j. \tag{13}$$

In other words, Y_j (resp., N_j) is the subset of $\mathcal{F}$ with $x_j = 1$ (resp., $x_j = 0$). Let Σ denote the subset algebra $\Sigma(\mathrm{Y}_1, \cdots, \mathrm{Y}_n)$, Finally, if $z \in R^{\Sigma}$, its α-entry (for $\alpha \in \Sigma$) is indicated by $z[\alpha]$, and for a matrix $B \in R^{\Sigma \times \Sigma}$ its α, β-entry is indicated by $B[\alpha, \beta]$.

Our lifting will map points in $\mathcal{F}$ to zeta-vectors arising from Σ. Formally, this is done as follows. Given $x \in \mathcal{F}$, denote by $\alpha(x) \in \Sigma$ the atom

$$(\bigcap_{j\,:\,x_j=1} \mathrm{Y}_j) \cap (\bigcap_{j\,:\,x_j=0} \mathrm{N}_j). \tag{14}$$

Note that the set-theoretic value of $\alpha(x)$ is precisely x, i.e., $\alpha(x)$ belongs in the equivalence class defined by x. Our lifting maps

$$x \quad \rightarrow \quad \xi^{\alpha(x)}. \tag{15}$$

Example 3.4 *Suppose $n = 5$. Given $v = (1, 1, 1, 0, 0) \in \{0, 1\}^5$, its lifting $\check{v} \in \{0, 1\}^{2^{32}}$ satisfies*

$$\begin{aligned} \check{v}[(Y_1 \cap Y_2) \cup Y_5] &= 1, \\ \check{v}[Y_3 \cap Y_4] &= 0, \\ \check{v}[Y_3 \cap (Y_4 \cup N_5)] &= 1, \end{aligned}$$

among other conditions.

An important point to notice is that only a (very small) subset of all zeta-vectors are images under this lifting – in this, our approach differs from those outlined in Section 1. Further, even though the dimension of the image space is very large (doubly exponential in n) in our algorithms we will only consider polynomially many coordinates.

Also note that the lifting (15) could be applied to any point $x \in \{0, 1\}^n$, not just $x \in \mathcal{F}$. However, below we will impose additional conditions on the lifting, conditions that are not guaranteed to be satisfied by points $x \in \{0, 1\}^n \setminus \mathcal{F}$, and therefore, in general, the lifting will not even exist unless $x \in \mathcal{F}$. This is fine for our purposes: we are interested in generating a relaxation that is as tight as possible. Similarly, we could have defined the target of our lifting to be the subset algebra of $\mathcal{F}$, not $\{0, 1\}^n$. Here the difference is completely semantic: if $x \in \mathcal{F}$ then, when $S \subseteq \{0, 1\}^n \setminus \mathcal{F}$, we have $\xi^{\alpha(x)}[S] = 0$, and when $S \cap \mathcal{F} \neq \emptyset$, we have $\xi^{\alpha(x)}[S] = \xi^{\alpha(x)}[S \cap \mathcal{F}]$.

The following properties of the lifting are easy to verify:

(i) $\xi^{\alpha(x)}_{\emptyset} = 0$.

(ii) $\xi^{\alpha(x)}_{\mathcal{F}} = 1$.

(iii) For $1 \leq j \leq n$,

$$\xi^{\alpha(x)}[\mathrm{Y}_j] \;=\; x_j, \quad \text{and} \quad \xi^{\alpha(x)}[\mathrm{N}_j] \;=\; 1 - x_j. \tag{16}$$

(iv) More generally, suppose we have a collection $\{\beta^i \in \Sigma \,:\, i \in I\}$, corresponding to pairwise disjoint subsets of $\mathcal{F}$, i.e., each β^i corresponds to a $\mathcal{J}^i \subseteq \mathcal{F}$ such that $\mathcal{J}^i \cap \mathcal{J}^j = \emptyset$, for all distinct $i, j \in I$. Then

$$\sum_{i \in I} \xi^{\alpha(x)}[\beta^i] \;\leq\; 1, \tag{17}$$

with equality when the $\mathcal{J}^i$ form a partition of $\mathcal{F}$.

Property (iv) follows because the point x cannot belong to more than one $\mathcal{J}^i$.

One can also prove results analogous to those outlined in Section 1. In particular, define $\hat{\mathcal{F}} \subseteq \{0, 1\}^{\Sigma}$ to be the image of $\mathcal{F}$ under our lifting. The following are straightforward results that formalize the above:

Lemma 3.5 *Let $z \in conv(\hat{\mathcal{F}})$. Suppose we have a collection $\{\beta^i \in \Sigma \, : \, i \in I\}$, corresponding to pairwise disjoint subsets of $\mathcal{F}$. Then*

$$\sum_{i \in I} z[\beta^i] \quad \leq \quad 1 \tag{18}$$

with equality when $\{\beta^i \, : \, i \in I\}$ corresponds to a partition of $\mathcal{F}$. Further,

$$z \geq 0, \quad z[\emptyset] \, = \, 0, \quad \text{and} \quad z[\mathcal{F}] \, = \, 1. \tag{19}$$

Lemma 3.6 *Let $x \in R^n$. Then $x \in conv(\mathcal{F})$ iff there exists a vector $z \in conv(\hat{\mathcal{F}})$ such that*

$$z[\mathrm{Y}_j] \, = \, x_j, \qquad z[\mathrm{N}_j] \, = \, 1 - x_j, \quad \forall \, 1 \leq j \leq n. \tag{20}$$

The main question we want to address is what conditions the vector z in Lemma 3.6 can be required to satisfy, in addition to (18 - 20).

To this effect, consider an arbitrary vector $f \in R^{\Sigma}$ of the form

$$f \quad = \quad \sum_{\alpha \in \Sigma} \lambda_\alpha \xi^\alpha, \tag{21}$$

where $\alpha \in R^{\Sigma}$ is nonnegative. Define the matrix

$$U^f \quad = \quad \sum_{\alpha \in \Sigma} \lambda_\alpha \xi^\alpha (\xi^\alpha)^T. \tag{22}$$

We have the following basic facts:

Lemma 3.7 *Suppose λ, f and U^f are as in (21 - 22). Let $\beta, \gamma \in \Sigma$. Then*

$$U^f_{\beta,\gamma} \quad = \quad \sum_{\alpha \subseteq \beta \cap \gamma} \lambda_\alpha. \tag{23}$$

Proof. Let $\alpha \in \Sigma$. Then the sum in (22) contributes either zero or λ_α to $U^f_{\beta,\gamma}$, and the latter happens when both $\xi^\alpha[\beta] = 1$ and $\xi^\alpha[\gamma] = 1$. By definition of ξ^α, this is true exactly when $\alpha \subseteq \beta$ and $\alpha \subseteq \gamma$, i.e., $\alpha \subseteq \beta \cap \gamma$, as desired. ■

Corollary 3.8 *For any $\beta, \gamma \in \Sigma$*

$$U^f_{\beta,\gamma} \quad = \quad f[\beta \cap \gamma]. \tag{24}$$

This is an analogue of equation (8). As a consequence of these results,

Lemma 3.9 *Suppose λ, f and U^f are as in (21 - 22). Then*

(i) U^f is symmetric,

(ii) The main diagonal of U^f and its $\mathcal{F}$-row and -column are all equal to f, and

(iii) $U^f \succeq 0$.

■

Note that in Lemma 3.9, (i) and (ii) are really very weak consequences of (24), which implies that many pairs of entries in U^f are equal. For $\alpha \in \Sigma$, let $e_\alpha \in R^{\Sigma}$ be the vector with a 1 in position α and zero otherwise.

Lemma 3.10 *Suppose λ, f and U^f are as in (21 - 22). Consider a vector $\eta \in R^{\Sigma}$ such that*

$$\eta^T \xi^\alpha \quad \geq \quad 0, \quad \forall \alpha \text{ with } \lambda_\alpha > 0. \tag{25}$$

Then

(i) $\eta^T U^f e_\beta \geq 0$*, for all* $\beta \in \Sigma$*.*

(ii) $\eta^T U^f (e_{\mathcal{F}} - e_\beta) \geq 0$*, for all* $\beta \in \Sigma$*.*

(iii) For every $v \in R^\Sigma$*,* $v^T U^\kappa v \geq 0$*, where* $\kappa = U^f \eta$*.*

Proof. (i) Let $\alpha \in \Sigma$ with $\lambda_\alpha > 0$. Then since $\xi^\alpha (\xi^\alpha)^T e_\beta$ is either identically zero (when $\xi^\alpha[\beta] = 0$) or else it equals ξ^α, we have $\eta^T \xi^\alpha (\xi^\alpha)^T e_\beta \geq 0$, and the result follows.

(ii) Let $\bar{\beta} \in \Sigma$ denote a negation of β (i.e., an element of Σ whose set theoretic value is that of the complement of β). Then the $\mathcal{F}$-column of U^f is equal to sum of the β-column and the $\bar{\beta}$-column of U^f, i.e.,

$$U^f e_{\mathcal{F}} = U^f (e_\beta + e_{\bar{\beta}}), \tag{26}$$

because (similarly to the proof of (i))

$$\xi^\alpha (\xi^\alpha)^T e_{\mathcal{F}} = \xi^\alpha (\xi^\alpha)^T (e_\beta + e_{\bar{\beta}}), \ \ \forall \ \alpha \in \Sigma \tag{27}$$

see (17). The result follows from (26) by applying (i) to $\bar{\beta}$.

(iii) This follows because $\kappa \in R^\Sigma$ satisfies

$$\kappa = \sum_{\lambda_\alpha > 0} \lambda_\alpha \xi^\alpha (\xi^\alpha)^T \eta \tag{28}$$

$$= \sum_{\gamma_\alpha > 0} \gamma_\alpha \xi^\alpha, \tag{29}$$

where $\gamma_\alpha = \lambda_\alpha \eta^T \xi^\alpha \geq 0$. Thus,

$$U^\kappa = \sum_{\gamma_\alpha > 0} \gamma_\alpha \xi^\alpha (\xi^\alpha)^T$$

from which (iii) follows. ■

The properties of U^f given in Lemma 3.9 and Lemma 3.10 (i) and (ii) parallel the Lovász-Schrijver development of their operators. Property (iii) in Lemma 3.10 is similar to the Sherali-Adams and Lasserre development as shown in [L01].

We summarize what we know so far.

Lemma 3.11 *Let* $\bar{x} \in conv(\mathcal{F})$*. Then there is a vector* $z \in R^\Sigma$ *such that:*

1. *Equations (18) - (20) are satisfied.*
2. *There is a matrix* $U^z \in R^{\Sigma \times \Sigma}$ *satisfying (24), conditions (i) - (iii) of Lemma (3.9), and conditions (i) - (iii) of Lemma (3.10).*

In particular, consider (ii) of Lemma (3.10). One way to use this condition is to start with any inequality $a^T x \geq a_0$ which is valid for $\mathcal{F}$ and from it obtain (as in Section 1) an inequality $\hat{a}^T z \geq 0$ which is valid for $\hat{\mathcal{F}}$. (Formally, set $\hat{a}[\mathrm{Y}_j] = a_j$ for $1 \leq j \leq n$, $\hat{a}[\mathcal{F}] = -a_0$, and $\hat{a}[\alpha] = 0$ for all other $\alpha \in \Sigma$).

In addition, there are measure-theoretic valid inequalities that one can use. For example, for $\alpha, \beta \in \Sigma$,

$$z[\alpha \cap \beta] \leq \min(z[\alpha], z[\beta]), \tag{30}$$

$$z[\alpha \cup \beta] = z[\alpha] + z[\beta] - z[\alpha \cap \beta] \tag{31}$$

are valid inequalities. Similar remarks can be made concerning property (iii) in Lemma 3.10.

One difference between our algorithms and the procedures in [LS91], [SA90], [BCC93] and [L01b], is that we generate variables indexed by Σ and not by the lattice of subsets of $\{1, \cdots, n\}$ – which is isomorphic to a (very small) proper subset of Σ. In addition, an iteration of our procedures generates elements of Σ involving possibly widely different quantities of symbols – as opposed to first generating pairs, then triples, and so on.

3.2.1 Variable replication.

We need to describe an additional feature that is required only to formally guarantee the validity of our algorithms. The need for this additional machinery is best explained with an example. In applying the Lemmas described above, our algorithms will restrict z and U^z to polynomial-sized subsets of Σ. As a simple example, we might generate a symbol σ_1 for the expression $((\mathrm{Y}_1 \cap \mathrm{N}_2) \cup (\mathrm{N}_1 \cap \mathrm{N}_2)) \cap \mathrm{N}_3$ and use it as one of the variable indices, that is, impose the (valid) constraint

$$z[\sigma_1] = z[\mathrm{Y}_1 \cap \mathrm{N}_2 \cap \mathrm{N}_3] + z[\mathrm{N}_1 \cap \mathrm{N}_2 \cap \mathrm{N}_3].$$

At the same time, we might create a symbol σ_2 for the expression $\mathrm{N}_2 \cap \mathrm{N}_3$, and use it in constraints, say

$$z[\sigma_2] \leq \min\{z[\mathrm{N}_2], z[\mathrm{N}_3]\}.$$

The algorithm certainly *would* benefit by imposing the valid requirement:

$$z[\sigma_1] = z[\sigma_2], \tag{32}$$

(or, indeed, by using a unique variable) – but the algorithm can only do this if it knows that

$$(\mathrm{Y}_1 \cap \mathrm{N}_2) \cup (\mathrm{N}_1 \cap \mathrm{N}_2) = \mathrm{N}_2,$$

and the algorithm can make this sort of deduction only by engaging in the symbolic algebra needed to determine this fact. If the algorithm does not impose (32) then we end up using a relaxation, i.e., a weaker formulation than theoretically possible. Sometimes this will be the case with our algorithms (in particular with regards to certain complex set-theoretic expressions) with the result that we end up producing duplicate symbols.

These ideas are formalized as follows.

Definition 3.12 *Let Q be a set. A* **symbol function** *is a function $\mathcal{S} : \Sigma \to 2^Q$ with the properties:*

($\mathcal{S}$.1) For each $\alpha, \beta \in \Sigma$ with $\alpha \neq \beta$, we have $\mathcal{S}(\alpha) \cap \mathcal{S}(\beta) = \emptyset$.

($\mathcal{S}$.2) $\mathcal{S}(\emptyset)$ and $\mathcal{S}(\mathcal{F})$ are nonempty.

For $\alpha \in \Sigma$, the elements of $\mathcal{S}(\alpha)$ are called the **symbols associated with** *α.*

Here, the set Q is arbitrary (for example, in a practical implementation we might have $Q = Z_+$). Let $\mathcal{S}(\Sigma)$ denote the union of all the sets $\mathcal{S}(\alpha)$, $\alpha \in \Sigma$. For consistency, we will use the notation $x[i]$ to refer to the i^{th} entry of vector $x \in R^{\mathcal{S}(\Sigma)}$, for $i \in \mathcal{S}(\Sigma)$, and similarly we will use the notation $M[i, j]$ to refer to the i, j entry in a matrix $M \in R^{\mathcal{S}(\Sigma) \times \mathcal{S}(\Sigma)}$.

Lemma 3.13 *Suppose $\mathcal{S}$ is a symbol function. Let $\mathcal{M} \subseteq \Sigma$ be such that $\emptyset \in \mathcal{M}$, $\mathcal{F} \in \mathcal{M}$, and $\mathcal{S}(\alpha)$ is nonempty for each $\alpha \in \mathcal{M}$. Let $U \in R^{\mathcal{M} \times \mathcal{M}}$. Consider the matrix*

$$\breve{U} \in R^{\mathcal{S}(\mathcal{M}) \times \mathcal{S}(\mathcal{M})}$$

defined as follows: for any $\alpha, \beta \in \mathcal{M}$, and any $i \in \mathcal{S}(\alpha)$ and $j \in \mathcal{S}(\beta)$, $\breve{U}[i, j] = U[\alpha, \beta]$. Then

(a) If U is symmetric positive-semidefinite, so is $\breve{U}$.

(b) Let $\eta \in R^{\mathcal{M}}$ be such that $\eta^T U \geq 0$. Suppose we choose, for each $\alpha \in \mathcal{M}$, a particular element $j_\alpha \in \mathcal{S}(\alpha)$. Then for each column $\breve{u}$ of $\breve{U}$ we have:

$$\sum_{\alpha \in \mathcal{M}} \eta[\alpha] \breve{u}[j_\alpha] \geq 0.$$

■

The proof of this Lemma is elementary. The significance of this lemma is that it allows our lifting to create multiple indices that correspond to the same subset of the algebra – though our algorithm may not be aware that these are duplicates of one another. At the same time, when we lift to a matrix we can impose on this matrix all the conditions discussed above (for example, the diagonal is equal to the $\mathcal{F}$-row and -column).

To summarize this section, our algorithms in general refrain from performing the manipulations needed to determine when two set-theoretic expressions are equivalent. This is done to avoid the exponential amount of work that such a certification would sometimes require. On the other hand, some equivalences are easy to check. As we will see, *any* of the expressions considered by our algorithms is explicitly of the form $B_1 \cap B_2 \cap \cdots \cap B_r$, where each B_i is of the form Y_j or N_j (for some j), or is $\mathcal{F}$, or belongs to a polynomial-size class of additional symbols. We will assume that permuting the B_i produces an equivalent expression. This is a requirement that is easy to enforce in polynomial time. Later we will discuss another simplification which is equally easy to enforce.

3.3 Example - a simple algorithm for set-covering

As a prelude to the algorithms we will describe later, here we present a simple algorithm specialized for set-covering problems that achieves provable results, which is a special case of a general algorithm to be described in Section 4. In particular, we will show that we obtain a polynomial-time algorithm for optimizing over a relaxation at least as strong as that given by the convex hull of all valid inequalities with coefficients 0, 1, or 2.

Thus, let A denote an $m \times n$ 0-1 matrix, and consider the feasible region $\mathcal{F}$ for a 0-1 set-covering problem,

$$Ax \geq e \tag{33}$$

$$x \in \{0,1\}^n \tag{34}$$

where we assume that no row of A contains another. We denote by $A_i \subseteq \{1, \cdots, n\}$ the set of indices of nonzeros in the i^{th} row of A, $1 \leq i \leq m$. The algorithm we describe next creates variables and specifies constraints that the variables must satisfy. The description we provide is a bit redundant. After presenting the algorithm we discuss its behavior.

Algorithm C

Step 0. Create the variable $X[\mathcal{F}]$, and for $1 \leq j \leq n$ the variables $X[\mathrm{Y}_j]$ and $X[\mathrm{N}_j]$, and impose the constraint:

$$X[\mathrm{Y}_j] + X[\mathrm{N}_j] - X[\mathcal{F}] = 0. \tag{35}$$

Step 1. For each $1 \leq i \leq m$ impose the constraint:

$$\sum_{j \in A_i} X[\mathrm{Y}_j] - X[\mathcal{F}] \geq 0. \tag{36}$$

Step 2. For each unordered pair of indices $i \neq h$, $1 \leq i,\, h \leq m$, where

$$C^{i,h} \doteq A_i \cap A_h,$$

if $C^{i,h} \neq \emptyset$ we do the following, provided that $C^{i,h}$ has not already been enumerated as the set $C^{\bar{i},\bar{h}}$ for some other pair $\{\bar{i}, \bar{h}\}$.

(2.a) Create the variable $X[\bigcap_{j \in C^{i,h}} \mathrm{N}_j]$ and impose the constraints:

$$X[\mathrm{N}_r] - X[\bigcap_{j \in C^{i,h}} \mathrm{N}_j] \quad \geq \quad 0, \quad \forall\, r \in C^{i,h}. \tag{37}$$

(2.b) For each $r \in C^{i,h}$, create the new variable $X[\mathrm{Y}_r \cap \bigcap_{j \in C^{i,h} - r} \mathrm{N}_j]$, and impose the constraint:

$$X[\mathrm{Y}_r] - X[\mathrm{Y}_r \cap \bigcap_{j \in C^{i,h} - r} \mathrm{N}_j] \quad \geq \quad 0. \tag{38}$$

(2.c) If $|C^{i,h}| \geq 2$, create the new variable $x[\tau^{i,h}]$, and impose the constraint:

$$\sum_{j \in C^{i,h}} X[\mathrm{Y}_j] - 2X[\tau^{i,h}] \quad \geq \quad 0. \tag{39}$$

Here, the symbol $\tau^{i,h}$ is a symbol associated with

$$\bigcup_{t \geq 2} \left(\bigcup_{S \subseteq C^{i,h}\,:\,|S|=t} \left(\bigcap_{j \in S} \mathrm{Y}_j \bigcap_{j \notin S} \mathrm{N}_j \right) \right) \quad \in \quad \Sigma. \tag{40}$$

That is to say, $\tau^{i,h}$ represents the union of all expressions involving intersections of Y and N variables indexed by all elements of $C^{i,h}$, where at least two of the variables are Ys.

(2.d) Impose the constraint:

$$X[\bigcap_{j \in C^{i,h}} \mathrm{N}_j] + \sum_{r \in C^{i,h}} X[\mathrm{Y}_r \cap \bigcap_{j \in C^{i,h}-r} \mathrm{N}_j] + X[\tau^{i,h}] - X[\mathcal{F}] \quad = \quad 0. \tag{41}$$

where the $\tau^{i,h}$ term is only used in case $|C^{i,h}| \geq 2$.

Step 3. Let $\mathcal{V}$ denote the set of variable indices we have created so far, and let X denote the vector of variables. Create a matrix U of variables, with rows and columns indexed by $\mathcal{V}$, and impose on the variables in U the following constraints:

(3.1) U is symmetric, $U[\mathcal{F}, \mathcal{F}] = X[\mathcal{F}]$, and the main diagonal, the $\mathcal{F}$-row and the $\mathcal{F}$-column of U are all equal to X.

(3.2) For each constraint $\eta^T X \geq 0$ of the form (35), (36), (37), (38), (39), (41) impose the constraints

$$\eta^T U \quad \geq \quad 0. \tag{42}$$

Step 4. Impose:

$$0 \leq U[\alpha, \beta] \leq U[\mathcal{F}, \beta] \;\; \forall \alpha, \beta \in \mathcal{V} \tag{43}$$
$$X[\mathcal{F}] = 1. \tag{44}$$

End.

Comment 3.14 *Note that the requirements in (3.1) are far weaker than what Corollary 3.8 would permit us to impose. For example, consider the expressions* $\beta_1 = \mathrm{Y}_1 \cap \mathrm{N}_2$, $\gamma_1 = \mathrm{Y}_3 \cap \mathrm{Y}_4 \cap \mathrm{N}_2$, $\beta_2 = \mathrm{Y}_1 \cap \mathrm{Y}_3 \cap \mathrm{Y}_4$, *and* $\gamma_2 = \mathrm{N}_2$. *Then we could stipulate that* $U[\beta_1, \gamma_1] = U[\beta_2, \gamma_2]$ *since* $\beta_1 \cap \gamma_1$ *and* $\beta_2 \cap \gamma_2$ *both equal* $\mathrm{Y}_1 \cap \mathrm{N}_2 \cap \mathrm{Y}_3 \cap \mathrm{Y}_4$. *In general, many pairs of entries in* U *can be required to be equal because they are indexed by equivalent set-theoretic indices as typified by the example. However, the limited demands that we will place upon Algorithm C are such that it is not necessary to enforce such equivalences – the more general algorithm we will discuss later* **will** *make some of these requirements.*

Example 3.15 Algorithm C applied to a small problem.

Consider the set covering problem on four variables given by the following set of constraints.

$$x_1 + x_2 + x_3 \qquad \geq \quad 1, \tag{45}$$
$$x_3 + x_4 \quad \geq \quad 1, \tag{46}$$
$$x_1 + x_2 \qquad + x_4 \quad \geq \quad 1, \tag{47}$$

Then Algorithm C will produce 13 *distinct variable indices:* $\mathcal{F}$, Y_1, Y_2, Y_3, Y_4, N_1, N_2, N_3, N_4, $\mathrm{N}_1 \cap \mathrm{N}_2 (= C^{1,3})$, $\mathrm{Y}_1 \cap \mathrm{N}_2$, $\mathrm{N}_1 \cap \mathrm{Y}_2$ *and* $\tau^{1,3}$. *In Step 3, the algorithm will therefore create a* 13×13 *matrix* U *that contains all the variables; since* U *is symmetric, and its* $\mathcal{F}$*-column equals its main diagonal, and* $U[\mathcal{F}, \mathcal{F}] = X[\mathcal{F}] = 1$

(which can be eliminated), we have a total of $13 \times 6 = 78$ *distinct variables. The nontrivial constraints imposed in Steps 0, 1 and 2 are:*

$$
\begin{aligned}
X[\mathrm{Y}_j] + X[\mathrm{N}_j] &= 1, \qquad j = 1,2,3 && (48)\\
X[\mathrm{Y}_1] + X[\mathrm{Y}_2] + X[\mathrm{Y}_3] &\geq 1 && (49)\\
X[\mathrm{Y}_3] + X[\mathrm{Y}_4] &\geq 1 && (50)\\
X[\mathrm{Y}_1] + X[\mathrm{Y}_2] + X[\mathrm{Y}_4] &\geq 1 && (51)\\
\min\{X[\mathrm{N}_1], X[\mathrm{N}_2]\} - X[\mathrm{N}_1 \cap \mathrm{N}_2] &\geq 0 && (52)\\
\min\{X[\mathrm{Y}_1], X[\mathrm{N}_2]\} - X[\mathrm{Y}_1 \cap \mathrm{N}_2] &\geq 0 && (53)\\
\min\{X[\mathrm{N}_1], X[\mathrm{Y}_2]\} - X[\mathrm{N}_1 \cap \mathrm{Y}_2] &\geq 0 && (54)\\
X[\mathrm{Y}_1] + X[\mathrm{Y}_2] - 2X[\tau^{1,3}] &\geq 0 && (55)\\
X[\mathrm{N}_1 \cap \mathrm{N}_2] + X[\mathrm{Y}_1 \cap \mathrm{N}_2] + X[\mathrm{N}_1 \cap \mathrm{Y}_2] + X[\tau^{1,3}] &= 1 && (56)
\end{aligned}
$$

In addition to these constraints, Step 3 imposes other constraints on the entries of U*. For example, the (homogenized) (47) is applied to the* N_3*-column; since (by Step 3.1)* $U[\mathcal{F}, \mathrm{N}_3] = X[\mathrm{N}_3]$ *the resulting constraint reads:*

$$U[\mathrm{Y}_1, \mathrm{N}_3] + U[\mathrm{Y}_2, \mathrm{N}_3] + U[\mathrm{Y}_4, \mathrm{N}_3] - X[\mathrm{N}_3] \geq 0.$$

As mentioned above, by Corollary 3.8 we could impose additional conditions on U*: for example, we could insist that* $U[\mathrm{N}_1, \mathrm{N}_2] = X[\mathrm{N}_1 \cap \mathrm{N}_2]$*. But, again, this will not be necessary in terms of the particular result regarding Algorithm C that is proved below.*

3.3.1 Analysis of Algorithm C.

We will first show that Algorithm C provides a valid lifting for $conv(\mathcal{F})$. In what follows, let $\mathcal{M}$ denote of the set of distinct members of Σ, i.e., subsets of $\{0,1\}^n$, that arise as the set-theoretic value of variable indices $\mathcal{V}$ produced by the algorithm. Thus, $\mathcal{M}$ contains $\emptyset$, $\mathcal{F}$, all the Y_j and N_j, all the (distinct) $\bigcap_{j \in C^{i,h}} \mathrm{N}_j$, all the (distinct) $\mathrm{Y}_r \cap \bigcap_{j \in C^{i,h}-r} \mathrm{N}_j$, and all the (distinct) sets of the form given by the left-hand of (40). In the language of Section 3.2.1, there is an implicit symbol function σ such that $\sigma(\mathcal{M}) = \mathcal{V}$, that the algorithm has constructed.

Theorem 3.16 *Suppose* $\bar{x} \in conv(\mathcal{F})$*. Then there exists a vector* $\check{X}$ *satisfying constraints (35), (36), (37), (38), (39), (41), (43), (44), and a matrix* $\check{U}$ *satisfying the conditions in Step 3, such that* $\check{X}[\mathrm{Y}_j] = \bar{x}_j$*, for all* $1 \leq j \leq n$*.*

Proof. Let $z \in R^{\Sigma}$ be the lifting of $\bar{x}$ that satisfies the conditions in Lemma 3.11, and U^z denote the corresponding matrix. We will show that z, restricted to $\mathcal{M}$, satisfies (35), (36), (37), (38), (39), (41), (43), (44) (with the variable indices interpreted to obtain their true set-theoretic value) and that U^z, restricted to $\mathcal{M} \times \mathcal{M}$, satisfies the conditions in Step 3. By appealing to Lemma 3.2.1 we obtain the desired result.

First, by Lemma 3.5, z satisfies (44). Next, consider constraint (41) applied to a given pair i, h. Now $\mathcal{F} \subseteq \{0,1\}^n$ can be partitioned into $2 + |C^{i,h}|$ sets:

- The set containing those points in $\{0,1\}^n$ with all coordinates in $C^{i,h}$ equal zero. This is the subset of $\{0,1\}^n$ corresponding to $\bigcap_{j \in C^{i,h}} \mathrm{N}_j$.
- The set containing those points in $\{0,1\}^n$ where one coordinate $r \in C^{i,h}$ equals one, and all other coordinates equal zero. This is the subset corresponding to $\mathrm{Y}_r \cap \bigcap_{j \in C^{i,h}-r} \mathrm{N}_j$.
- The set containing all remaining points in $\{0,1\}^n$, i.e., points where at least two coordinates in $C^{i,h}$ equal one. This is the subset corresponding to $\tau^{i,h}$.

Consequently, by Lemma 3.5, z satisfies

$$z[\bigcap_{j \in C^{i,h}} \mathrm{N}_j] + \sum_{r \in C^{i,h}} z[\mathrm{Y}_r \cap \bigcap_{j \in C^{i,h}-r} \mathrm{N}_j] + z[\tau^{i,h}] \quad = \quad 1, \tag{57}$$

and we obtain (41), since $z[\mathcal{F}] = 1$. Similarly, it follows that z and U satisfy all other constraints. ■

Further,

Lemma 3.17 *The total number of variables and constraints created by Algorithm C is $O(m^4 n^2)$.* ■

In preparation for our main result, we need the following step.

Lemma 3.18 *Let (X, U) satisfy all the conditions set by Algorithm C. Suppose $1 \le i,\ h \le m$ are distinct rows of A with $C^{i,h} \neq \emptyset$. Let V be the submatrix of U formed by taking the column indexed by $\bigcap_{j \in C^{i,h}} \mathrm{N}_j$, and all the columns indexed by $\mathrm{Y}_r \cap \bigcap_{j \in C^{i,h}-r} \mathrm{N}_j$ (for every $r \in C^{i,h}$), and the column indexed $\tau^{i,h}$ (when $|C^{i,h}| \ge 2$).*

Suppose $\eta \in R^{\Sigma}$ is such that $\eta^T V \ge 0$. Then $\eta^T X \ge 0$.

Proof. By Step 3.2, every column of U satisfies (41). But U is symmetric, and therefore every row of U satisfies (41). Therefore, the column of U indexed by $\mathcal{F}$ is obtained by adding together the column indexed by $\bigcap_{j \in C^{i,h}} \mathrm{N}_j$, the columns indexed by $\mathrm{Y}_r \cap \bigcap_{j \in C^{i,h}-r} \mathrm{N}_j$ (for every $r \in C^{i,h}$), and the column indexed by $\tau^{i,h}$ (if $|C^{i,h}| \ge 2$).

By assumption, each column v in this sum satisfies $\eta^T v \ge 0$. We conclude that the column indexed by $\mathcal{F}$ satisfies this inequality, as well. Now we are done, by Step 3.1. ■

Finally, we now have our main theorem.

Theorem 3.19 *Consider an inequality $a^T x \ge a_0$ which is valid for $\mathcal{F}$ and such that $a_j \in \{0, 1, 2\}$, for $0 \le j \le n$. Let (X, U) be a vector and matrix satisfying all the constraints imposed by Algorithm C. Then*

$$\sum_{j=1}^{n} a_j X[\mathrm{Y}_j] \quad \ge \quad a_0. \tag{58}$$

Proof. Assume without loss of generality that $a^T x \ge a_0$ is not dominated by another valid inequality with coefficients in $\{0, 1, 2\}$.

If $a_0 = 1$ it follows that no indices j satisfy $a_j = 2$. Then $a^T x \ge a_0$ is one of the constraints $Ax \ge e$ that define $\mathcal{F}$, and thus (58) follows by (36).

If $a_0 = 0$ then (58) is implied by (43).

We are left with the case $a_0 = 2$. If all nonzero a_j equal 2, then by dividing by 2 we return to the case $a_0 = 1$. Consequently, we can assume that $a^T x \ge a_0$ is of the form $2x(T) + x(S) \ge 2$ for disjoint subsets $T, S \subseteq \{1, \cdots, n\}$ where $S \neq \emptyset$. (Here and elsewhere, for a vector v and index set J, we write $v(J) = \sum_{j \in J} v_j$).

Recall that the support of the k^{th} constraint is denoted by A_k. For each element $j \in S$ it is easy to see that:

$$\text{for some } 1 \le k \le m, \qquad A_k \subseteq T \cup S - j, \tag{59}$$

since otherwise the point $z \in \{0,1\}^n$ with $z_g = 1$ iff $g \in \{1, 2, \cdots, n\} - (T \cup S)$ or $g = j$, is in $\mathcal{F}$, but $a^T z < a_0$. Consequently, since $S \neq \emptyset$, we obtain that one of the original constraints $x(A_i) \ge 1$ has $A_i \subseteq T \cup S$.

Clearly $A_i \cap S \neq \emptyset$ – or else $a^T x \ge 2$ is dominated by $x(A_i) \ge 1$, a contradiction. Hence, if we pick any $j \in A_i \cap S$, and apply (59) to j we find an index $1 \le h \le m$ with $A_h \subseteq T \cup S - j$. Necessarily, $i \neq h$.

Suppose first that $A_i \cap A_h = \emptyset$. In that case, $a^T x \ge 2$ is dominated by the sum of $x(A_i) \ge 1$ and $x(A_h) \ge 1$, and by (36) applied to i and h, X satisfies (58).

We will therefore assume $C^{i,h} = A_i \cap A_h \neq \emptyset$. In the remainder of the proof we will show that if the vector v is the column of U indexed either by $\bigcap_{j \in C^{i,h}} \mathrm{N}_j$, or by $\mathrm{Y}_r \cap \bigcap_{j \in C^{i,h}-r} \mathrm{N}_j$ (for some $r \in C^{i,h}$), or by $\tau^{i,h}$, then

$$\sum_{j=1}^{n} a_j v[\mathrm{Y}_j] - 2v[\mathcal{F}] \quad \ge \quad 0. \tag{60}$$

By Lemma (3.18), this will complete the proof.

Consider first the case of the column v corresponding to $\bigcap_{j\in C^{i,h}} \mathrm{N}_j$. By step (3.1) of the algorithm, we have

$$v[\bigcap_{j\in C^{i,h}} \mathrm{N}_j] \quad = \quad v[\mathcal{F}]. \tag{61}$$

By constraint (37) imposed by the algorithm,

$$v[\mathrm{N}_r] \quad \geq \quad v[\bigcap_{j\in C^{i,h}} \mathrm{N}_j], \quad \forall\, r \in C^{i,h}. \tag{62}$$

Consequently, (35) and (44) applied to v, together with (61) and (62) imply:

$$v[\mathrm{Y}_r] \quad = \quad 0, \;\; \forall \;\; r \in C^{i,h}. \tag{63}$$

But by Step 3.2 v satisfies (36) applied to i. Together with (63) this implies

$$\sum_{j\in A_i - C^{i,h}} v[\mathrm{Y}_j] - v[\mathcal{F}] \quad \geq \quad 0. \tag{64}$$

Similarly,

$$\sum_{j\in A_h - C^{i,h}} v[\mathrm{Y}_j] - v[\mathcal{F}] \quad \geq \quad 0, \tag{65}$$

and since by construction of $C^{i,h}$ in Step 2 we have $(A_i - C^{i,h}) \cap (A_h - C^{i,h}) = \emptyset$, we conclude

$$\sum_{j\in A_i \cup A_h} v[\mathrm{Y}_j] - 2v[\mathcal{F}] \geq 0,$$

which dominates (60).

Consider now the case that v is the column of U corresponding to some $\mathrm{Y}_r \cap \bigcap_{j\in C^{i,h}-r} \mathrm{N}_j$. Then, by Step 3.1 $v[\mathcal{F}] = v[\mathrm{Y}_r \cap \bigcap_{j\in C^{i,h}-r} \mathrm{N}_j]$, and by constraint (38) imposed by the algorithm we conclude

$$v[\mathrm{Y}_r] \quad \geq \quad v[\mathcal{F}]. \tag{66}$$

If $a_r = 2$ we conclude that v satisfies (60). If, on the other hand, $a_r = 1$, then by (59) there is a constraint $x(A_k) \geq 1$ of the original system with $A_k \subseteq T \cup S - r$. By (36), we conclude that $\sum_{j\in A_k} v[\mathrm{Y}_j] - v[\mathcal{F}] \geq 0$. This, together with (66) implies v satisfies (60).

Finally, consider the case when v is the column of U corresponding to $\tau^{i,h}$. By Step 3.1 $v[\tau^{i,h}] = v[\mathcal{F}]$, and therefore by constraint (39) $\sum_{j\in C^{i,h}} v[\mathrm{Y}_j] - 2v[\mathcal{F}] \geq 0$, which dominates (60).

The theorem is proved. ■

Additional comments on Algorithm C

1. Note that Algorithm C does not require positive semidefiniteness of the matrix U. Nevertheless, even though the formulation generated by the algorithm is of polynomial size, it can imply exponentially many facets.

To see this, consider the following example. Suppose V and W are sets and for each $v \in V$ there is a subset $J_v \subseteq W$ such that the J_v ($v \in V$) are pairwise disjoint. Consider the system $C_{W,V}$ of constraints

$$y_h \;+\; \sum_{j\in V-v} x_j \quad \geq \quad 1, \;\; \forall\, v \in V, \;\; h \in J_v. \tag{67}$$

$$(y, x) \quad \in \quad \{0,1\}^W \times \{0,1\}^V. \tag{68}$$

[When $W = \emptyset$, the constraint matrix defined by these inequalities is a full-circulant matrix.] The total number of constraints of type (67) is at most $|V| \times |W|$. We have:

Proposition 3.20 *Suppose that for each $v \in V$ we choose an element $w(v) \in J_v$. Then*

$$\sum_{v\in V} y_{w(v)} \;+\; \sum_{v\in V} x_v \;\geq\; 2$$

defines a facet of conv$(C_{W,V})$.

Thus, we have a family of $\Pi_{v\in V}|J_v|$ facets, all with coefficients in $\{0,1,2\}$, and therefore implied by Algorithm C.

Suppose we have a pure full-circulant example, i.e., $W = \emptyset$. In the Appendix we present a procedure that is stronger than both the N_+ and the Sherali-Adams operators, and yet needs at least $|V|-3$ rounds to prove the valid inequality $\sum_{j\in|V|} x_j \geq 2$.

2. In Step (2.c) we use the symbol $\tau^{i,h}$ to represent the set-theoretic expression in (40). This is necessary in the case of the symbols $\tau^{i,h}$ because (40) has exponential length – but in a practical implementation we would also want to efficiently record some of the other indices for variables, for example, $\mathrm{Y}_r \cap \bigcap_{j\in C^{i,h}-r} \mathrm{N}_j$, which only requires four symbols to store, in addition to those used to store $C^{i,h}$.

4 Main algorithm

4.1 Obstructions

The algorithm we describe here generalizes the approach presented in the last section. The notion of an *obstruction* plays a central role. First we need some notation.

Notation. Let $1 \leq j \leq n$. In what follows, the notation M_j will be used to denote a symbol that is either Y_j or N_j; $\bar{\mathrm{Y}}_j$ denotes N_j and $\bar{\mathrm{N}}_j$ denotes Y_j. The symbols Y_j and N_j will be called **literals**.

Definition 4.1 *An* **obstruction for** $a^T x \geq a_0$ *is an element* $\omega \in \Sigma$ *of the form*

$$\omega = \mathrm{M}_{j_1} \cap \mathrm{M}_{j_2} \cap \cdots \cap \mathrm{M}_{j_h}$$

where for each $1 \leq i \leq h$ *we have* $1 \leq j_i \leq n$ *and* $a_{j_i} \neq 0$, *and such that*

$$\xi^{\alpha(x)}[\omega] \quad = \quad 0, \quad \forall \ x \in \{0,1\}^n \ \textit{with} \ a^T x \geq a_0. \tag{69}$$

Put differently, ω is an obstruction for $a^T x \geq a_0$ if any $x \in \{0,1\}^n$ satisfying

$$x_{j_i} = \begin{cases} 1 & \text{if } \mathrm{M}_{j_i} = \mathrm{Y}_{j_i} \\ 0 & \text{if } \mathrm{M}_{j_i} = \mathrm{N}_{j_i} \end{cases} \tag{70}$$

for every $1 \leq i \leq h$ satisfies $a^T x < a_0$. Obstructions of this type are closely related to *covers* of knapsacks, see [NW88].

Definition 4.2 *For* $1 \leq j \leq n$, $\mathrm{Y}_j \cap \mathrm{N}_j$ *is called a* **trivial** *obstruction.*

Note: using Definition 4.1, $\mathrm{Y}_j \cap \mathrm{N}_j$ is an obstruction to $x_j + (1 - x_j) = 1$.

EXAMPLES

(a) Set-covering. Given a constraint $\sum_{j\in S} x_j \geq 1$, its unique obstruction is $\bigcap_{j\in S} \mathrm{N}_j$.

(b) Set-packing. Given a constraint $\sum_{j\in S} x_j \leq 1$, with $|S| = N$, the are $\binom{N}{2}$ minimal obstructions (here "minimal" refers to the set of literals in the obstruction). Each of them is of the form $\mathrm{Y}_{i(1)} \cap \mathrm{Y}_{i(2)}$ for some pair $i(1)$, $i(2)$ of distinct indices from S.

(c) Set-partitioning. Given a constraint $\sum_{j\in S} x_j = 1$, its set of obstructions are those obtained by viewing the constraint as a set-covering and as a set-packing constraint.

(d) Multi-covering or -packing. Consider a constraint which is either of covering type, $\sum_{j\in S} x_j \geq |S| - k$, or of packing type, $\sum_{j\in S} x_j \leq |S| - k$. In the covering case, the minimal obstructions consist of intersections of $k+1$ literals N_j, and in the packing case they consist of intersections of $k+1$ literals Y_j. For k fixed, we can therefore enumerate all minimal obstructions in polynomial time.

(e) Mixed covering and packing. Consider a constraint of the form

$$\sum_{j\in I} \alpha_j x_j + \sum_{j\in J} \beta_j (1 - x_j) \geq b,$$

where I and J are disjoint, $\alpha_j > 0$ and $\beta_j > 0$ for each j, and $b \geq 0$. This generalizes the examples in (a) - (d). Clearly, no minimal obstruction contains symbols of the form Y_j, $j \in I$ or N_j, $j \in J$.

Definition 4.3 *Let $k \geq 0$. Consider an inequality of the form*

$$\sum_{j \in J^+} a_j x_j - \sum_{j \in J^-} a_j x_j \geq b$$

*where $a_j > 0 \ \forall j \in J^+ \cup J^-$. An obstruction to this inequality is called k-**small** if it is of the form*

$$\bigcap_{j \in A^+} \mathrm{N}_j \cap \bigcap_{j \in A^-} \mathrm{Y}_j$$

where $A^+ \subseteq J^+$ and either $|A^+| \geq |J^+| - k$ or $|A^+| \leq k$; and $A^- \subseteq J^-$ and either $|A^-| \geq |J^-| - k$ or $|A^-| \leq k$.

Note: for any for fixed k we can enumerate all k-small obstructions to any inequality in polynomial time. Also note that all minimal obstructions for set-covering, set-packing and set-partitioning constraints are k-small for $k \leq 2$.

4.2 The algorithm

Now we return to our general lifting procedure. The algorithm we will present is a generalization of Algorithm C given in Section 3.3. That algorithm created variables indexed by members of Σ, all of which were either intersections of N_j terms, or intersections of N_j terms and one Y_j term, or more complicated expressions: unions of intersections of N_j terms and some number of Y_j terms. The particular expressions actually generated by the algorithm were not the complete set of all possible such expressions – rather, they reflected the structure of the problem.

The algorithm presented below generalizes this approach. Broadly speaking, the algorithm will perform three steps:

(a) Generate a (polynomial-size) family of set-theoretic indices, $\mathcal{V}$. This family will include all Y_j and Y_j as well as more complex expressions defined below that we call "walls" and "tiers", both of which are derived from obstructions to the constraints.

(b) Write constraints that a lifting of any point $x \in \mathcal{F}$ to $\{0,1\}^{\Sigma}$ must satisfy. These constraints include all (homogenized) inequalities present in the continuous relaxation of $\mathcal{F}$, and also set-theoretic constraints with support in $\mathcal{V}$. (The support of a vector v, denoted *suppt*(v), is the set of indices j such that $v_j \neq 0$). In particular, there will be constraints that generalize constraint (41) in Algorithm C.

(c) Create a matrix of variables $U \in R^{\mathcal{V} \times \mathcal{V}}$. This matrix is designed to capture the implications of Lemma 3.11. So, for example, we can impose on its columns all constraints generated in (b).

The expressions created in (a) are clearly critical for the success of such an approach. We introduce them next.

Definition 4.4 *Suppose we are given an expression $\gamma = \bigcap_{i=1}^{h} \mathrm{M}_{j_i}$, where for $1 \leq i \leq h$, we have $1 \leq j_i \leq n$.*

(0) We write $|\gamma| = h$.

*(1) For $0 \leq t \leq h$, a **negation of γ of order** t is an expression of the form $\bigcap_{i=1}^{h} \mathrm{M}'_{j_i}$, such that for exactly t indices j_i we have $\mathrm{M}'_{j_i} = \bar{\mathrm{M}}_{j_i}$, and for the remaining $h - t$ indices j_i we have $\mathrm{M}'_{j_i} = \mathrm{M}_{j_i}$.*

(2) $0 \leq r \leq h$, $\mathcal{N}(\gamma, r)$ is the set of negations of γ of order r (note: if $r = 0$ then this is just the set $\{\gamma\}$).

*(3) For $0 \leq t < h$, the **negation of γ of order greater than** t is the expression $\gamma^{>t}$ defined by*

$$\gamma^{>t} \doteq \bigcup_{r=t+1}^{h} \left(\bigcup_{\beta \in \mathcal{N}(\gamma, r)} \beta \right).$$

Example 4.5 *Consider*

$$\begin{aligned}\gamma_1 &= \mathrm{N}_1 \bigcap \mathrm{Y}_2 \bigcap \mathrm{Y}_3\\ \gamma_2 &= \mathrm{Y}_4 \bigcap \mathrm{Y}_5 \bigcap \mathrm{Y}_6 \bigcap \mathrm{N}_7\\ \gamma_3 &= \mathrm{N}_8 \bigcap \mathrm{Y}_9.\end{aligned}$$

Suppose

$$\begin{aligned}\gamma'_1 &= \mathrm{Y}_1 \bigcap \mathrm{Y}_2 \bigcap \mathrm{Y}_3\\ \gamma'_2 &= \mathrm{N}_4 \bigcap \mathrm{N}_5 \bigcap \mathrm{Y}_6 \bigcap \mathrm{N}_7\\ \gamma'_3 &= \mathrm{N}_8 \bigcap \mathrm{Y}_9.\end{aligned}$$

Then $|\gamma_1| = 3$, $|\gamma_2| = 4$, $|\gamma_3| = 2$, *and* γ'_i *is a negation of* γ_i *of order* o_i, *where* $o_1 = 1$, $o_2 = 2$ *and* $o_3 = 0$.

The concept of negation will be of central importance to our algorithms given later. Consider the case of γ_2 above, and let $t = 1$. Then we can **partition** $\mathcal{F}$ as

$$\begin{aligned}\mathcal{F} \;=\; & \gamma_2 \bigcup \left(\mathrm{N}_4 \bigcap \mathrm{Y}_5 \bigcap \mathrm{Y}_6 \bigcap \mathrm{N}_7\right) \bigcup \left(\mathrm{Y}_4 \bigcap \mathrm{N}_5 \bigcap \mathrm{Y}_6 \bigcap \mathrm{N}_7\right)\\ & \left(\mathrm{Y}_4 \bigcap \mathrm{Y}_5 \bigcap \mathrm{N}_6 \bigcap \mathrm{N}_7\right) \bigcup \left(\mathrm{Y}_4 \bigcap \mathrm{Y}_5 \bigcap \mathrm{Y}_6 \bigcap \mathrm{Y}_7\right) \bigcup \gamma_2^{>1}. \qquad (71)\end{aligned}$$

Since this is a partition, if $x \in \mathcal{F}$ is lifted to $z \in R^{\Sigma}$ we will have:

$$\begin{aligned}1 \;=\; & z[\gamma_2] + z[\mathrm{N}_4 \bigcap \mathrm{Y}_5 \bigcap \mathrm{Y}_6 \bigcap \mathrm{N}_7] + z[\mathrm{Y}_4 \bigcap \mathrm{N}_5 \bigcap \mathrm{Y}_6 \bigcap \mathrm{N}_7]\\ & z[\mathrm{Y}_4 \bigcap \mathrm{Y}_5 \bigcap \mathrm{N}_6 \bigcap \mathrm{N}_7] + z[\mathrm{Y}_4 \bigcap \mathrm{Y}_5 \bigcap \mathrm{Y}_6 \bigcap \mathrm{Y}_7] + z[\gamma_2^{>1}]. \qquad (72)\end{aligned}$$

This valid constraint satisfied by z can be viewed as a multi-way disjuntion (see [B75]). Our algorithm will systematically generate a (polynomial) number of such disjunctions, which are driven by the set of obstructions to the constraints in the problem. The following definitions lay the ground for this.

Definition 4.6

(a) Consider a set of expressions $E = \{\gamma_1, \gamma_2, \cdots, \gamma_t\}$ *where each* γ_i *is of the form* $\gamma_i = \bigcap_{r=1}^{h_i} \mathrm{M}_{j(r,i)}$ *and* $1 \le j(r,i) \le n$ *for* $1 \le r \le h_i$. *Then the* **wall derived from** E *is the expression* $\Omega(E) = \bigcap \mathrm{M}_j$ *containing every literal* M_j *occurring in* **more than one** γ_i. *If no such symbol exists, we will say that the wall is* **empty**.

(b) A **tier** *is an expression of the form* $\gamma = \omega_1 \cap \omega_2 \cap \cdots \cap \omega_t$, *where each* ω_i *is a negation of a wall (of some order* t, *or greater than* t, *for some* $t \ge 0$*).*

Example 4.7 *Suppose* $E = \{\gamma_1, \gamma_2, \gamma_3\}$, *where*

$$\begin{aligned}\gamma_1 &= \mathrm{N}_1 \bigcap \mathrm{Y}_2 \bigcap \mathrm{Y}_3\\ \gamma_2 &= \mathrm{N}_1 \bigcap \mathrm{Y}_2 \bigcap \mathrm{Y}_4 \bigcap \mathrm{N}_5\\ \gamma_3 &= \mathrm{N}_1 \bigcap \mathrm{Y}_4.\end{aligned}$$

Then the wall derived from E *is* $\omega_1 = \mathrm{N}_1 \bigcap \mathrm{Y}_2 \bigcap \mathrm{Y}_4$.
Suppose in addition that $H = \{\gamma_1, \gamma_4\}$, *where*

$$\gamma_4 = \mathrm{Y}_2 \bigcap \mathrm{Y}_3 \bigcap \mathrm{N}_4.$$

Then the wall derived from H *is* $\mathrm{Y}_2 \bigcap \mathrm{Y}_3$, *and* $\omega_2 = \mathrm{Y}_2 \bigcap \mathrm{N}_3$ *is a negation of it. Thus,* $\omega_1 \bigcap \omega_2 = \mathrm{N}_1 \bigcap \mathrm{Y}_2 \bigcap \mathrm{Y}_4 \bigcap \mathrm{Y}_2 \bigcap \mathrm{N}_3$ *is a tier.*

Note that the last expression in this example is redundant. To handle this and similar situations we introduce the following formal procedure.

Definition 4.8

(i) Given an expression $\gamma = \bigcap_{i=1}^{h} \mathrm{M}_{j_i}$*. The* **simplification** *of* γ *is the expression of the form* $\bigcap \mathrm{M}_j$ *obtained by taking exactly one copy of each* distinct *literal* M_j *appearing in* γ*.*

(ii) Given an expression of the form $\gamma = \beta \cap \omega^{>r}$*, where* $\beta = \bigcap_{i=1}^{h} \mathrm{M}_{j_i}$*,* ω *is a wall and* $r > 0$*, the simplification of* γ *is the expression* $\beta' \cap \omega^{>r}$*, where* β' *is the simplification of* β*.*

(iii) Suppose we have an expression of the form $\beta = \gamma \cap \omega_1^{>r_1} \cap \omega_2^{>r_2}$*, where* ω_1 *and* ω_2 *are walls and* r_1, r_2 *are positive, and* γ *is as in (i). Let* γ' *be the simplification of* γ*. Then the simplification of* β *is the expression* $\gamma' \cap \omega_1^{>r_1} \cap \omega_2^{>r_2}$*, if* $\omega_1 \neq \omega_2$*, and it equals* $\gamma' \cap \omega_1^{>\max\{r_1, r_2\}}$ *otherwise.*

(iv) Suppose we have an expression of the form $\beta = \mathcal{F} \cap \gamma$*, where* γ *is of the form considered in (i) or (ii). Then the simplification of* β *equals the simplification of* γ*.*

Example 4.9 *Suppose*

$$\gamma = \mathrm{N}_1 \bigcap \mathrm{Y}_2 \bigcap \mathrm{N}_1 \bigcap \mathrm{Y}_3.$$

Then the simplification of γ *is*

$$\mathrm{Y}_2 \bigcap \mathrm{Y}_3 \bigcap \mathrm{N}_1.$$

Comment 4.10 *All expressions that the algorithm given below will generate will be of one of the following types:*

Σ*.1. a literal,*

Σ*.2.* $\mathcal{F}$*,*

Σ*.3. a wall,*

Σ*.4. a tier,*

Σ*.5. an expression of the form* $\beta \cap \gamma$ *where* β *and* γ *are of types* Σ*.1-*Σ*.4.*

Any expressions of type Σ*.1-*Σ*.5 is of the form* $B_1 \cap B_2 \cap \cdots \cap B_r$*, where each* B_i *is a literal,* $\mathcal{F}$*, or of the form* $\omega^{>r}$*. Recall the discussion at the end of Section 3.2.1. We will use the assumption that any permutation of the* B_i *yields an equivalent expression, and also apply the simplification operator, to enforce equivalences between the symbols generated by our algorithms.*

Definition 4.11 *Suppose we are given an expression*

$$\alpha = \mathrm{M}_{j_1} \bigcap \mathrm{M}_{j_2} \bigcap \cdots ... \bigcap \mathrm{M}_{j_p}$$

and an expression

$$\beta = \mathrm{M}_{i_1} \bigcap \mathrm{M}_{i_2} \bigcap \cdots ... \bigcap \mathrm{M}_{i_q}.$$

We say that β *is a* **superstring** *of* α *if every literal* M_h *appearing in the simplification of* α *also appears in the simplification of* β*.*

Now we can define our basic general algorithm. This algorithm embodies one particular formalization of the idea of lifting to variables indexed by the subset algebra – several other such formalizations are possible. The primary reason we use the algorithm below is that we can show that it has provably good properties. However, in a practical implementation most likely one would use only some of the algorithmic ideas we present. The algorithm is presented in somewhat redundant form.

Let $k \geq 2$ be a fixed integer, and as before we have a feasible set $\mathcal{F} = \{\, x \in \{0,1\}^n \,:\, Ax \geq b \,\}$, where A is $m \times n$. The i^{th} constraint is denoted $a_i^T x \geq b_i$.

Σ^k-Algorithm ($k \geq 2$).

1. Generate the symbol $\mathcal{F}$, and for $1 \leq j \leq n$, generate the symbols Y_j and N_j, and impose the constraint:

$$X[\mathrm{Y}_j] + X[\mathrm{N}_j] - X[\mathcal{F}] = 0. \tag{73}$$

For $1 \leq i \leq m$ impose the constraint:

$$\sum_{j=1}^{n} a_{ij} X[\mathrm{Y}_j] - b_i X[\mathcal{F}] \geq 0. \tag{74}$$

2. For $1 \leq i \leq m$, enumerate all the k-small obstructions to each of the constraints $a_i^T x \geq b_i$, as well as all trivial obstructions. For each enumerated obstruction γ, impose the constraint

$$\sum_{\mathrm{M}_j \in \gamma}^{n} X[\bar{\mathrm{M}}_j] - X[\mathcal{F}] \geq 0. \tag{75}$$

where the notation "$\mathrm{M}_j \in \gamma$" means that the sum is over all those literals M_j occurring in γ.

3. For every set E of distinct enumerated obstructions with $1 < |E| \leq k$, compute $\Omega(E)$, the wall derived from E. If $\Omega(E) \neq \emptyset$ and $\Omega(E)$ does not contain both terms Y_j and N_j for some j $(1 \leq j \leq n)$ then generate a symbol for $\Omega(E)$.

4. For every $1 \leq \mathcal{H} < k$, every tier of the form $\theta = \bigcap_{t=1}^{\mathcal{H}} \omega_t$, where ω_t is an enumerated wall for $1 \leq t \leq \mathcal{H}$, is processed to generate additional symbols and constraints, as follows.

Enumerate every $\mathcal{H}$-vector of integers $o = (o_1, o_2, \cdots, o_{\mathcal{H}})$ such that $0 \leq o_t \leq \min\{|\omega_t|, k\}$ for $1 \leq t \leq \mathcal{H}$ and $\sum_{t=1}^{\mathcal{H}} o_t < k$. Given o, enumerate every $\mathcal{H}$-tuple $\omega' = (\omega'_1, \omega'_2, \cdots, \omega'_{\mathcal{H}})$, where ω'_t is a negation of ω_t of order o_t for $1 \leq t \leq \mathcal{H}$. For each such enumerated pair (o, ω'),

(i) Create a symbol for

$$\theta_o^{\#} = \bigcap_{t=1}^{\mathcal{H}} \omega'_t$$

and, for each $1 \leq j \leq n$, and each literal M_j such that M_j is one of the literals in $\theta_o^{\#}$, impose the constraint

$$X[\mathrm{M}_j] - X[\theta_o^{\#}] \geq 0. \tag{76}$$

Further, if the simplification σ of $\theta_o^{\#}$ is a superstring of some enumerated obstruction, or if it contains both symbols Y_j and N_j for some $1 \leq j \leq n$, impose

$$X[\theta_o^{\#}] = 0. \tag{77}$$

Finally, impose

$$X[\theta_o^{\#}] \; - \sum_{\mathrm{M}_j \in \sigma} X[\mathrm{M}_j] + (|\sigma| - 1) X[\mathcal{F}] \geq 0. \tag{78}$$

(ii) If $\sum_{t=1}^{\mathcal{H}} o_t = k - 1$ and $o_{\mathcal{H}} < |\omega_{\mathcal{H}}|$ create a symbol for

$$\theta_o^{>} = \left(\bigcap_{t=1}^{\mathcal{H}-1} \omega'_t \right) \cap \omega_{\mathcal{H}}^{> o_{\mathcal{H}}},$$

where as before $\omega_i^{> o_i}$ is the negation of ω_i of order greater than o_i. Further, for each $1 \leq j \leq n$, and each literal M_j such that M_j is one of the symbols in any of the ω'_t, for $1 \leq t \leq \mathcal{H} - 1$, impose the constraint

$$X[\mathrm{M}_j] - X[\theta_o^{>}] \geq 0. \tag{79}$$

Moreover, impose the constraint

$$\sum_{\mathrm{M}_j \in \omega_{\mathcal{H}}} X[\bar{\mathrm{M}}_j] - (1 + o_{\mathcal{H}}) X[\theta_o^{>}] \quad \geq \quad 0. \tag{80}$$

Further, if the simplification of $\bigcap_{t=1}^{\mathcal{H}-1} \omega'_t$ is a superstring of some enumerated obstruction, or if it contains both symbols Y_j and N_j for some $1 \leq j \leq n$, impose

$$X[\theta_o^>] = 0. \tag{81}$$

Note: strictly speaking, the notation $\theta_o^\#$ and $\theta_o^>$ is not complete, since we need to specify *which* literals are negated, but the correct interpretation will be clear from the context.

5. For each expression $\theta_o^\# = \bigcap_{t=1}^{\mathcal{H}} \omega'_t$ enumerated in Step 4, such that $\mathcal{H} < k-1$ and ω'_t is a negation of order o_t of ω_t, for $1 \leq t \leq \mathcal{H}$; and for each enumerated wall ω, impose the following constraints:

(a) If $|\omega| + \sum_{t=1}^{\mathcal{H}} o_t < k$ then impose the constraint

$$X[\theta_o^\#] - \sum_{r=0}^{|\omega|} \left(\sum_{\beta \in \mathcal{N}(\omega,r)} X[\theta_o^\# \cap \beta] \right) \quad = \quad 0. \tag{82}$$

(b) If $|\omega| + \sum_{t=1}^{\mathcal{H}} o_t \geq k$ then setting $R = k - 1 - \sum_{t=1}^{\mathcal{H}} o_t$, impose the constraint

$$X[\theta_o^\#] \; - \; X[\theta_o^\# \cap \omega^{>R}] - \sum_{r=0}^{R} \left(\sum_{\beta \in \mathcal{N}(\omega,r)} X[\theta_o^\# \cap \beta] \right) \quad = \quad 0, \tag{83}$$

where, as previously, $\mathcal{N}(\omega, r)$ denotes the set of negations of ω of order r.

Similarly, for each enumerated wall ω, if $|\omega| < k$ then impose

$$X[\mathcal{F}] - \sum_{r=0}^{|\omega|} \left(\sum_{\beta \in \mathcal{N}(\omega,r)} X[\beta] \right) \quad = \quad 0, \tag{84}$$

and if $|\omega| \geq k$ then impose

$$X[\mathcal{F}] \; - \; X[\omega^{>k-1}] - \sum_{r=0}^{k-1} \left(\sum_{\beta \in \mathcal{N}(\omega,r)} X[\beta] \right) \quad = \quad 0. \tag{85}$$

6. Let $\mathcal{V}$ denote the set of variable indices created so far, and let X denote the vector of variables. Create a matrix U of variables, with rows and columns indexed by $\mathcal{V}$, and impose on the variables in U the following constraints:

(6.1) U is symmetric, $U[\mathcal{F}, \mathcal{F}] = X[\mathcal{F}]$, and the main diagonal, the $\mathcal{F}$-row and the $\mathcal{F}$-column of U are all equal to X.

(6.2) For each constraint $\eta^T X \geq 0$ of the form (74-85) (when they apply) impose the constraints

$$\eta^T U \quad \geq \quad 0. \tag{86}$$

(6.3) Suppose $\beta_1, \gamma_1, \beta_2, \gamma_2$ are elements of $\mathcal{V}$ such that $\beta_1 \cap \gamma_1$ and $\beta_2 \cap \gamma_2$ have the same simplification. Then require that $U[\beta_1, \gamma_1] = U[\beta_2, \gamma_2]$. Suppose β, γ are elements of $\mathcal{V}$ such that $\beta \cap \gamma$ is a superstring of some enumerated obstruction, or contains both Y_j and N_j for some j. Then impose $U[\beta, \gamma] = 0$.

(6.4) Impose

$$0 \;\leq\; U[\theta, \beta] \;\leq\; U[\mathcal{F}, \beta] \quad \forall\, \theta, \beta \in \mathcal{V} \tag{87}$$

$$X[\mathcal{F}] = 1. \tag{88}$$

7. (Optional) Impose $U \succeq 0$.

8. (Optional) For each constraint $\eta^T X \geq 0$ of the form (74), (75), (78) or (80), let $f^\eta \doteq U\eta$. Form the matrix W_η with rows and columns indexed by $\mathcal{V}$, such that for each pair β, γ of symbols in $\mathcal{V}$, the β, γ-entry of W_η equals $f^\eta[\beta \cap \gamma]$. Impose $W_\eta \succeq 0$.

9. End.

Note: in what follows, a vector $o = (o_1, o_2, \cdots, o_{\mathcal{H}})$ as in Step 4 will be called a *negation vector*.

Example 4.12 *Application with $k = 4$. Consider the following sytem of constraints:*

$$\begin{array}{lrrrrrrrrl} C_1: & -x_1 & +x_2 & +\frac{1}{2}x_3 & & & & & \geq & 0 \\ C_2: & -x_1 & +x_2 & & -x_4 & -x_5 & & & \geq & -2 \\ C_3: & -x_1 & & & +x_4 & & +x_6 & +\frac{4}{5}x_7 & \geq & 0 \\ C_4: & x_1 & & -x_3 & & +x_5 & +x_6 & +x_7 & \geq & 0 \\ C_5: & & x_2 & -x_3 & & +x_5 & +x_6 & +x_7 & \geq & 0 \\ C_6: & -x_1 & & +x_3 & & & & +x_7 & \geq & 0 \\ C_7: & -x_1 & & +x_3 & & & & -x_7 & \geq & -1 \end{array} \tag{89}$$

The algorithm enumerates, among others, the (2-small) obstruction γ_i to C_i $(1 \leq i \leq 7)$ given next:

$$\begin{aligned} \gamma_1 &= \mathrm{Y}_1 \cap \mathrm{N}_2 \cap \mathrm{N}_3 \\ \gamma_2 &= \mathrm{Y}_1 \cap \mathrm{N}_2 \cap \mathrm{Y}_4 \cap \mathrm{Y}_5 \\ \gamma_3 &= \mathrm{Y}_1 \cap \mathrm{N}_4 \cap \mathrm{N}_6 \\ \gamma_4 &= \mathrm{N}_1 \cap \mathrm{Y}_3 \cap \mathrm{N}_5 \cap \mathrm{N}_6 \cap \mathrm{N}_7 \\ \gamma_5 &= \mathrm{N}_2 \cap \mathrm{Y}_3 \cap \mathrm{N}_5 \cap \mathrm{N}_6 \cap \mathrm{N}_7 \\ \gamma_6 &= \mathrm{Y}_1 \cap \mathrm{N}_3 \cap \mathrm{N}_7 \\ \gamma_7 &= \mathrm{Y}_1 \cap \mathrm{N}_3 \cap \mathrm{Y}_7 \end{aligned}$$

Thus, constraint (75) applied to γ_1 yields

$$X[\mathrm{N}_1] + X[\mathrm{Y}_2] + X[\mathrm{Y}_3] - X[\mathcal{F}] \geq 0.$$

In Step 3, the algorithm will generate, among others, the wall derived from $\{\gamma_1, \gamma_2, \gamma_3\}$, which is

$$\omega_1 = \mathrm{Y}_1 \cap \mathrm{N}_2,$$

the wall derived from $\{\gamma_4, \gamma_5\}$, which is

$$\omega_2 = \mathrm{Y}_3 \cap \mathrm{N}_5 \cap \mathrm{N}_6 \cap \mathrm{N}_7,$$

and the wall derived from $\{\gamma_6, \gamma_7\}$

$$\omega_3 = \mathrm{Y}_1 \cap \mathrm{N}_3.$$

In Step 4, the algorithm processes the tier

$$\theta \;=\; \omega_1 \bigcap \omega_2 \;=\; (\mathrm{Y}_1 \cap \mathrm{N}_2) \bigcap (\mathrm{Y}_3 \cap \mathrm{N}_5 \cap \mathrm{N}_6 \cap \mathrm{N}_7),$$

and in Step 4 (i) it creates (among several others) the symbol

$$\theta_o^{\#} \;=\; (\mathrm{Y}_1 \cap \mathrm{Y}_2) \bigcap (\mathrm{Y}_3 \cap \mathrm{N}_5 \cap \mathrm{N}_6 \cap \mathrm{Y}_7),$$

obtained from θ by negating the N_2 appearing in ω_1 and the N_7 appearing in ω_2. This corresponds to the negation vector $o = (1, 1)$. In this case, the collection of all constraints (76) can be abbreviated as:

$$\min\{X[\mathrm{Y}_1], X[\mathrm{Y}_2], X[\mathrm{Y}_3], X[\mathrm{N}_5], X[\mathrm{N}_6], X[\mathrm{Y}_7]\} - X[\theta_o^{\#}] \geq 0,$$

and since $\theta_o^{\#}$ is simplified, constraint (78) is:

$$X[\theta_o^{\#}] - X[\mathrm{Y}_1] - X[\mathrm{Y}_2] - X[\mathrm{Y}_3] - X[\mathrm{N}_5] - X[\mathrm{N}_6] - X[\mathrm{Y}_7] + 5X[\mathcal{F}] \geq 0.$$

Further, in Step 4 (ii) the algorithm, using the negation vector $p = (2, 1)$ and $\omega_1' = \mathrm{N}_1 \cap \mathrm{Y}_2$ (i.e., both literals in ω_1 are negated) will also generate the expression

$$\theta_p^{>} = (\mathrm{N}_1 \cap \mathrm{Y}_2) \bigcap \omega_2^{>1}$$

in which case (80) becomes:

$$X[\mathrm{N}_3] + X[\mathrm{Y}_5] + X[\mathrm{Y}_6] + X[\mathrm{Y}_7] - 2X[\theta_p^{>}] \geq 0.$$

In Step 4 the algorithm will also generate the tier

$$\omega_1 \bigcap \omega_2 \bigcap \omega_3 = (\mathrm{Y}_1 \cap \mathrm{N}_2) \bigcap (\mathrm{Y}_3 \cap \mathrm{N}_5 \cap \mathrm{N}_6 \cap \mathrm{N}_7) \bigcap (\mathrm{Y}_1 \cap \mathrm{N}_3) = \theta \bigcap \omega_3,$$

and, in Step 4 (ii) the algorithm will generate, using the negation vector $(1,1,1)$ *the expression*

$$(\mathrm{Y}_1 \cap \mathrm{Y}_2) \bigcap (\mathrm{Y}_3 \cap \mathrm{N}_5 \cap \mathrm{N}_6 \cap \mathrm{Y}_7) \bigcap \omega_3^{\geq 1} =$$

$$\theta_o^{\#} \bigcap \omega_3^{\geq 1}.$$

In this case (83) applies, and it imposes

$$X[\theta_o^{\#}] - X[\theta_o^{\#} \bigcap \omega_3^{\geq 1}] - X[\theta_o^{\#} \bigcap (\mathrm{N}_1 \cap \mathrm{N}_3)] - X[\theta_o^{\#} \bigcap (\mathrm{Y}_1 \cap \mathrm{Y}_3)] = 0.$$

Also, a tier obtained from the wall derived from $\{\gamma_6, \gamma_7\}$*, and the wall derived from* $\{\gamma_1, \gamma_2\}$*, for the negation vector* $(1,0)$ *(negating* N_3 *in* $\mathrm{Y}_1 \cap \mathrm{N}_3$*) is* $(\mathrm{Y}_1 \cap \mathrm{Y}_3) \bigcap (\mathrm{Y}_1 \cap \mathrm{N}_2)$*, whose simplification is* $\sigma = \mathrm{Y}_1 \cap \mathrm{Y}_3 \cap \mathrm{N}_2$*. In this case constraint (78) becomes:*

$$X[(\mathrm{Y}_1 \cap \mathrm{Y}_3) \bigcap (\mathrm{Y}_1 \cap \mathrm{N}_2)] - X[\mathrm{Y}_1] - X[\mathrm{N}_2] - X[\mathrm{Y}_3] + 2X[\mathcal{F}] \geq 0.$$

Finally, when constructing the matrix U*, note that by Step (6.3)* $U[\omega_2, \omega_3] = 0$*, as* $\omega_2 \cap \omega_3$ *contains both* Y_3 *and* N_3*.*

Similarly, one of the symbols we will generate will be $\omega_2' = \mathrm{N}_3 \cap \mathrm{N}_5 \cap \mathrm{N}_6 \cap \mathrm{N}_7$ *(obtained from* ω_3 *by negating* Y_3*), and so Step 6 (ii) will force* $U[\omega_2', \omega_3] = 0$ *since* $\omega_2' \cap \omega_3$ *is a superstring of* $\mathrm{Y}_1 \cap \mathrm{N}_3 \cap \mathrm{N}_7 = \omega_6$*.*

Comments on the algorithm.

1. In step (6.2) some of the constraints imposed on U are of the form "=", rather than "$\geq$" as stated in the algorithm – then in (86) we should use "=".
2. Note that the simplification step in (6.3) implies that, for example, if β_1 and β_2 are elements of $\mathcal{V}$ with the same simplification, then the columns (and rows) of U corresponding to β_1 and β_2 are identical. In such a case one could simply keep a single row and column corresponding to the equivalence class of symbols that have the same simplification. To keep the analysis simple, in what follows we will assume that we are using the entire matrix.
3. In steps 5(a), (b), note that the expressions $\theta_o^{\#} \cap \beta$ are obtained by negating walls in a tier enumerated in step 4. Hence the variables $X[\theta_o^{\#} \cap \beta]$ will be created by the algorithm, and similarly with all other variables in (82-85).
4. The algorithm can be made far more efficient. For example, it is not strictly necessary to use both Y_j and N_j variables. More to the point, it is possible to use far fewer variables.
5. Step 7 is M_+-like, using of course a very different set of variables. Similarly, Step 8 is Lasserre-like (refer to Lemma 3.10 (iii)).

Notation. In what follows we will use the notation $col[\theta]$ to mean the column of U corresponding to the symbol $\theta \in \mathcal{V}$. The matrix U will always be clear from the context. Note that $col[\theta] = Ue_\theta$, where $e_\theta \in R^{\mathcal{V}}$ has a 1 in position θ and zeros elsewhere, but we prefer the "col" notation.

The following result, which is straightforward, will be used in many of the proofs below, and we record it for reference.

Proposition 4.13 *Let* $\mathcal{V}$ *and* U *be the set of symbols and the matrix produced by a run of the* Σ^k *algorithm. Suppose there is an equation* $\sum_{\theta \in \mathcal{V}} \lambda[\theta] u[\theta] = 0$ *which is satisfied by every column* u *of* U*, where* $\lambda \in R^{\mathcal{V}}$*. Then* $\sum_{\theta \in \mathcal{V}} \lambda[\theta] col[\theta] = \vec{0}$*, where* $\vec{0} \in R^{\mathcal{V}}$ *is the 0-vector.*

Proof. Since U is symmetric, it follows that $\sum_{\theta \in \mathcal{V}} \lambda[\theta] r[\theta] = 0$ is satisfied by every row r of U, which implies the result. ■

4.3 Some basic properties of the Σ^k-algorithm

In this section we prove that the algorithm generates a valid formulation, and that it runs in polynomial-time for fixed k.

In what follows, let $\mathcal{M}$ denote the set of distinct members of Σ, i.e., subsets of $\{0,1\}^n$, that arise as the set-theoretic value of variable indices $\mathcal{V}$ produced by the algorithm.

Lemma 4.14 *Let $\hat{x} \in \mathcal{F}$. Then $\hat{x}$ can be lifted to a vector $\breve{X} \in \{0,1\}^{\mathcal{V}}$ and matrix $\breve{U} \in \{0,1\}^{\mathcal{V}\times\mathcal{V}}$ that satisfy the constraints imposed by the algorithm.*

Proof. We will show that $\hat{x}$ can be lifted to a vector $\hat{x} \in R^{\mathcal{M}}$ and a matrix $\mathcal{M}$ which satisfy the desired constraints. By appealing to Lemma 3.2.1 we will be done.

Let $\alpha(\hat{x}) \in \Sigma$ denote the atom corresponding to $\hat{x}$, i.e.,

$$\alpha(\hat{x}) \;=\; (\bigcap_{j\,:\,\hat{x}_j=1} \mathrm{Y}_j) \cap (\bigcap_{j\,:\,\hat{x}_j=0} \mathrm{N}_j).$$

Define $\hat{X} \in \{0,1\}^{\mathcal{M}}$ to be the restriction of $\xi^{\alpha(\hat{x})} \in \{0,1\}^{\Sigma}$ to the coordinates in $\mathcal{M}$. Recall here that $\xi^{\alpha(\hat{x})}$ is a 0-1 vector such that for each $\beta \in \Sigma$, $\xi^{\alpha(\hat{x})}[\beta]$ equals one iff the subset of $\mathcal{F}$ defined by β contains the point $\hat{x}$. Further, let $\hat{U}$ be defined by $\hat{U}[\beta,\gamma] = \hat{X}[\beta \cap \gamma]$.

Let $\breve{X}$ denote the restriction of $\hat{X}$ to $\{0,1\}^{\mathcal{V}}$ and $\breve{U} = \hat{U}$. We claim that $\breve{X}$ and $\breve{U}$ are as desired.

Clearly (73), (74) and (75) are satisfied. Consider now expressions θ, $\theta_o^{\#}$ as in Step 4 of the algorithm. Since $\theta_o^{\#}$ is of the form $\bigcap_{j\in J} \mathrm{M}_j$ for some set J, clearly (76) is satisfied (that is, the subset of $\mathcal{F}$ defined by $\theta_o^{\#}$ is contained in the subset defined by M_j for $j \in J$). Similarly, (79) is satisfied. Constraint (78) holds because if $|\sigma|$ terms $\breve{X}[\mathrm{M}_j]$ equal 1, then so does $\breve{X}[\theta_o^{\#}]$, and hence in any case the left-hand of (78) is nonnegative. (77) and (79) are clear. Finally, in constructing $\theta_o^{>}$ at least $1 + o_{\mathcal{H}}$ symbols are negated in $\omega_{\mathcal{H}}$, and consequently (80) holds as well.

Consider now (82 - 85). Suppose that case 5(a) applies. If $\hat{X}[\mathrm{M}_j] = 0$, for at least one literal M_j which appears in at least one of the expressions ω'_t, for $1 \le t \le \mathcal{H}$, then by construction both terms in (82) are equal to zero and (82) holds. Suppose now that $\hat{X}[\theta_o^{\#}] = 1$. Let $0 \le K \le |\omega|$ be the number of indices j, such that either Y_j is a symbol in ω and $\hat{x}_j = 0$, or N_j is a symbol in ω and $\hat{x}_j = 1$. By construction in step 3 of the algorithm, K is well defined; exactly one term $X[\theta_o^{\#} \cap \beta]$ with $\beta \in \mathcal{N}(\omega, K)$ will equal 1, and all remaining terms $X[\theta_o^{\#} \cap \beta]$ will be zero, and (82) follows. The proof that (83 - 85) hold, when they apply, is similar.

Finally, by construction, the conditions in (6.1), (6.2) and (6.4) apply; and the fact that the conditions in the optional steps 7 and 8 hold follow from Lemma 3.10. ■

As a corollary of this Lemma, we now have:

Theorem 4.15 *Suppose $\tilde{x} \in conv(\mathcal{F})$. Then there exists a vector $\tilde{X} \in R^{\mathcal{V}}$ and a matrix $\tilde{U} \in R^{\mathcal{V}\times\mathcal{V}}$ satisfying the conditions imposed by the Σ^k-algorithm, and such that $\tilde{X}[\mathrm{Y}_j] = \tilde{x}_j$, for all $1 \le j \le n$.*

Proof. This follows from Lemmas 3.11 and 4.14. ■

Lemma 4.16 *For k fixed, the Σ^k-algorithm generates polynomially many variables and constraints.*

Proof. Suppose that we start with a formulation with n variables. Then there are at most $O(mn^k)$ k-small obstructions. Thus at most $O\left(m^k n^{k^2}\right)$ walls, and at most $O\left(m^{k^2-k} n^{k^3-k}\right)$ tiers θ are enumerated. As a result, at most $O\left(2^{3k} k!\, m^{k^2-k} n^{k^3}\right)$ symbols $\theta_o^{\#}$ and $\theta_o^{>}$ will be created. ■

The following result is not needed in what follows, and we only state it for completeness, although its proof is not difficult. Let $\mathcal{R}_k^n$ denote the set of those vectors $x \in [0,1]^n$ such that x can be lifted to a pair (X, U) which satisfies the Σ^k-algorithm constraints.

Lemma 4.17 *(a) Let $k \geq 3$. Then $\mathcal{R}_k^n \subseteq \mathcal{R}_{k-1}^n$. (b) Suppose that each variable appears in at least one constraint. Then $\mathcal{R}_{n+1}^n = conv(\mathcal{F})$.* ■

Comment 4.18 *For a proof, see [Z03]. In particular, in (b), the U matrix generated by the Σ^k algorithm ($k \geq 3$) has the property that every column indexed by a literal satisfies the Σ^{k-1} constraints. Further, One can prove general conditions where fewer than n rounds are possible.*

Observation. Some of the expressions produced by the Σ^k algorithm are of the form $\cap_{j=1}^h M_{i_j}$ where the M_{i_j} are literals. These expressions amount to simple conjunctions of variables and their negations, and as such will "eventually" be enumerated by all the other lifting procedures we have discussed, for example, the Sherali-Adams operator. The "eventually" here is important, because the quantity of literals in the expression, h, may be very large – in general, our algorithm will enumerate expressions with a number of literals that is unbounded as a function of k. To put it differently, in polynomial time we enumerate expressions that, using the other operators, are not enumerated until after an exponential amount of work. In addition, the expressions that we do enumerate are obtained from the obstructions to the constraints, i.e., are driven by the structure of the problem. Furthermore, expressions of the form $\gamma^{>t}$ are **never** enumerated by the other algorithms.

The nature of the expressions we enumerate, and of the constraints placed upon them, are concrete algorithmic details that distinguish our procedure from the other lifting procedures. In the following sections we will present specific examples of inequalities that have unbounded rank for (say) the N_+ operator but which the Σ^k-algorithm satisfies for (small) fixed k.

4.4 Applications to set covering

In this section we consider a set-covering problem with feasible region $\mathcal{F} = \{ x \in \{0,1\}^n \ : \ Ax \geq e \}$, where the 0-1 matrix A is $m \times n$. We would like to prove, as generalization of Theorem 3.19 that the Σ^k algorithm implies all valid inequalities with coefficients $0, 1, 2, \cdots, k$. However, a direct inductive proof (on k) does not work, and we need a stronger inductive assumption.

This is provided by the concept of *pitch*: we will prove that the Σ^k algorithm implies all inequalities of pitch $\leq k$ which are valid for $\mathcal{F}$. For completeness, we redefine pitch here.

Definition 4.19 *Given an inequality $a^T x \geq a_0$ with indices ordered so that $0 < a_1 \leq a_2 \leq \cdots \leq a_J$ and $a_j = 0$ for $j > J$, its pitch is the minimum integer $\pi = \pi(a, a_0)$ such that $\sum_{j=1}^{\pi} a_j \geq a_0$.*

In fact, not only is pitch $\leq k$ a the right inductive assumption, but it is a more appropriate concept, since it serves to parameterize all valid inequalities, and any valid inequality has pitch $\leq n$. We note the following result, which follows directly from the definition of pitch.

Proposition 4.20 *Consider a valid inequality $a^T x \geq a_0$. Let $I \subseteq suppt(a)$ be such that $\sum_{i \in I} a_i < a_0$. Then the pitch of $a^T x \geq a_0$ is at least $|I| + 1$.* ■

We denote by A_i the support of the i^{th} row of A. We will assume, without loss of generality, that no A_i contains another.

Observation. If $\sum_{j \in V} x_j \geq 1$ is valid for $\mathcal{F}$ for some set V, then there is a row h of A with $A_h \subseteq V$.

Proposition 4.21 *Suppose $\alpha^T x \geq \alpha_0$ is a valid inequality for $\mathcal{F}$ with nonnegative coefficients. Consider the set-covering problem with feasible region $\mathcal{G} = \{x \in \{0,1\}^{suppt(\alpha)} \ : \ \hat{A}x \geq e\}$, where $\hat{A}$ is the submatrix of A obtained by*

(i) Using the columns in $suppt(\alpha)$, and

(ii) Using those rows i such that $A_i \subseteq suppt(\alpha)$.

Let $\bar{\alpha}$ be the restriction of α to $suppt(\alpha)$. Then $\bar{\alpha}^T x \geq \alpha_0$ is valid for $\mathcal{G}$.

Proof. Assume by contradiction that we can find a point $\hat{x} \in \mathcal{G}$ with $\alpha^T\hat{x} < \alpha_0$. Define $\breve{x} \in \{0,1\}^n$ by setting $\breve{x}_j = \hat{x}_j$ if $j \in \mathit{suppt}(\alpha)$, and $\breve{x}_j = 1$ otherwise. Then by construction $\breve{x} \in \mathcal{F}$, but $\alpha^T\breve{x} < \alpha_0$, a contradiction. ■

The following result describes a key structural property of valid inequalities for set covering problems, and will be at the core of our proof of Theorem 1.2.

Proposition 4.22 *Suppose $\alpha^T x \geq \alpha_0$ is a valid inequality for $\mathcal{F}$ and with $\alpha \geq 0$ and $\alpha_0 > 0$. Then there is a subset $\mathcal{C} = \mathcal{C}(\alpha, \alpha_0)$ of the rows of A with $|\mathcal{C}| \leq \pi(\alpha, \alpha_0)$, such that*

$$(i) \qquad A_i \subseteq \mathit{suppt}(\alpha) \quad \forall \ i \in \mathcal{C} \tag{90}$$

$$(ii) \qquad \Delta_i \doteq A_i - \bigcup_{r \in \mathcal{C}-i} A_r \neq \emptyset, \quad \forall \ i \in \mathcal{C}, \ \ \mathit{and} \tag{91}$$

$$(iii) \qquad \sum_{i \in \mathcal{C}} \min\{\alpha_j \ : \ j \in \Delta_i\} \ \geq \ \alpha_0. \tag{92}$$

Proof. The proof of the Lemma will be by induction on $\pi = \pi(\alpha, \alpha_0)$. Since $\alpha_0 > 0$ we must have that $\sum_{j \in suppt(\alpha)} x_j \geq 1$ is valid for $\mathcal{F}$, and thus the case $\pi = 1$ is proved.

In the remainder of the proof we assume that $A_i \subseteq \mathit{suppt}(\alpha)$ for every row i; without loss of generality we can make this assumption by Proposition 4.21.

Assume now that $\pi > 1$. Suppose first that some row i of A satisfies $A_i = \mathit{suppt}(\alpha)$. Since by assumption no row of A contains another, we have a set-covering problem with one constraint, namely $\sum_{j \in suppt(\alpha)} x_j \geq 1$. Thus, either $\pi = 1$, or $\alpha^T x \geq \alpha_0$ is not valid for $\mathcal{F}$, a contradiction in either case.

Hence, choose any row $i(1)$; $A_{i(1)}$ will then be properly included in $\mathit{suppt}(\alpha)$. Let $j(1)$ be an index in $A_{i(1)}$ with minimum coefficient $\alpha_{j(1)}$. Assume $\alpha_{j(1)} < \alpha_0$ (or else we are done by setting $\mathcal{C} = \{i(1)\}$).

Let $K = \mathit{suppt}(\alpha) - A_{i(1)}$. Consider the set-covering problem with feasible region $\mathcal{H} = \{x \in \{0,1\}^K \ : \ \bar{A}x \geq e\}$, where $\bar{A}$ is the submatrix of A where

1. $\bar{A}$ has column set K.
2. for any row h of A, with $j(1) \notin A_h$, $\bar{A}$ will have a row with $\bar{A}_h = A_h \cap K$.

We claim that the inequality

$$\sum_{j \in K} \alpha_j x_j \quad \geq \quad \alpha_0 - \alpha_{j(1)} \tag{93}$$

is valid for $\mathcal{H}$. For otherwise, we can find $\bar{x} \in \mathcal{H}$ with $\sum_{j \in K} \alpha_j \bar{x}_j < \alpha_0 - \alpha_{j(1)}$. In that case, define $\breve{x} \in \{0,1\}^n$ by setting $\breve{x}_j = \bar{x}_j$ if $j \in K$, $\breve{x}_{j(1)} = 1$ and $\breve{x}_j = 0$ for all $j \in A_{i(1)} - j(1)$ (recall that we are assuming $A_i \subseteq \mathit{suppt}(\alpha)$ for all i). Clearly $\alpha^T\breve{x} < a_0$. But $\breve{x} \in \mathcal{F}$, because no row of A is contained in another (and, in particular, not contained in $A_{i(1)}$). This contradiction shows that (93) is indeed valid for $\mathcal{H}$.

Since $\alpha_{j(1)} > 0$ it follows that the pitch of (93) is less than π. Since $\alpha_{j(1)} < \alpha_0$, the result now follows by induction as $j(1)$ is not contained in any set $\bar{A}_i$. ■

Corollary 4.23 *Suppose the inequality $a^T x \geq a_0$ is valid for $\mathcal{F}$ and has pitch ≤ 1. Then $a^T x \geq a_0$ is dominated by $a_i x \geq 1$ for some row i of A, or is dominated by the nonnegativity constraints.* ■

In the rest of this section we will consider a fixed vector $\tilde{X} \in R^{\mathcal{V}}$ and matrix $\tilde{U} \in R^{\mathcal{V} \times \mathcal{V}}$ that satisfy the constraints imposed by the Σ^k-algorithm. Our goal is to show that Theorem 1.2 holds, i.e., that for every inequality $a^T x \geq a_0$ with $\pi(a, a_0) \leq k$ which is valid for $\mathcal{F}$, $\tilde{X}$ satisfies the (homogenized) inequality

$$\sum_j a_j X[\mathrm{Y}_j] - a_0 X[\mathcal{F}] \quad \geq \quad 0, \tag{94}$$

which implies Theorem 1.2 since by construction $\tilde{X}[\mathcal{F}] = 1$.

4.4.1 Brief outline of the proof of Theorem 1.2

We will next outline our strategy towards this goal, which relies on using induction and on the fact that the algorithm enforces constraints (82-85).

Suppose that we could express $\tilde{X}$ as a sum of other columns of $\tilde{U}$, each of which satisfies (94). Since (94) is homogeneous, it follows that $\tilde{X}$ satisfies (94) as well. This "decomposition" approach mirrors the strategy we followed in the algorithm given in Section 3.3.

To fix ideas, consider an instance of constraint (84) arising from a particular choice of ω. Since this constraint is satisfied by every column of $\tilde{U}$, by Proposition 4.13 we have

$$\tilde{X} \quad = \quad \sum_{r=0}^{|\omega|} \left(\sum_{\beta \in \mathcal{N}(\omega,r)} col[\beta] \right). \tag{95}$$

Now (95) (or a similar decomposition if (85) applies instead of (84)) holds for every wall ω enumerated in step 5. Consequently, by the discussion in the previous paragraph, our task would be complete if we could select ω so that each column in the sum satisfies (94).

In our proof, the particular ω that will give the desired decomposition will be supplied by Proposition 4.22. In order to show that each resulting β-column of $\tilde{U}$ satisfies (94) we will use a special proof for the case $r = 0$, and for the column corresponding to $\omega^{>k-1}$ if (85) applies. For the cases where $r > 0$ we will use induction.

The induction is applied as follows. Consider the term β corresponding to a particular $r > 0$. Let $\tilde{u} = col[\beta]$. We want to show that

$$\sum_j a_j \tilde{u}[\mathrm{Y}_j] - a_0 \tilde{u}[\mathcal{F}] \quad \geq \quad 0. \tag{96}$$

Since we are dealing with a set-covering problem and $r > 0$, each of the negations of order r yields a Y_j literal appearing in β. As we will argue, for each such j we have $\tilde{u}[\mathrm{Y}_j] = \tilde{u}[\mathcal{F}]$. Thus, instead of having to show (96) we will have to show

$$\sum_{j : \mathrm{Y}_j \notin \beta} a_j \tilde{u}[\mathrm{Y}_j] - (a_0 - \sum_{j : \mathrm{Y}_j \in \beta} a_j) \tilde{u}[\mathcal{F}] \geq 0$$

It turns out that this is a weaker condition, because, as we will show, the inequality

$$\sum_{j : \mathrm{Y}_j \notin \beta} a_j x_j \geq (a_0 - \sum_{j : \mathrm{Y}_j \in \beta} a_j)$$

is valid for $\mathcal{F}$ and has pitch strictly less than the pitch of $a^T x \geq a_0$. Hence, we can apply induction, after another use of Proposition 4.22. Subsequent inductive steps will use (82) and (83) instead of (84) and (85). Finally, for the basis of the induction we have to handle valid inequalities of pitch ≤ 1 – but all such inequalities are satisfied by *all* columns of $\tilde{U}$, as implied by Corollary 4.23.

4.4.2 Formal proof of Theorem 1.2

Now we continue with the formal proof. Note that since we are dealing with a set covering problem all obstructions will be of the form $\bigcap_{j \in A_h} \mathrm{N}_j$ for some h.

The following observation will be of use later.

Proposition 4.24 *Let $\theta = \bigcap_{t=1}^{\mathcal{H}} \omega_t$ be a tier enumerated in step 4 of the algorithm. Consider one of the expressions $\theta_o^{\#}$ obtained from θ, and let $\tilde{u} = col[\theta_o^{\#}]$. Suppose $1 \leq j \leq n$ is such that Y_j appears in $\theta_o^{\#}$. Then*

$$\tilde{u}[\mathrm{Y}_j] \quad = \quad \tilde{u}[\mathcal{F}]. \tag{97}$$

Similarly, if $\theta_o^{>} = (\bigcap_{t=1}^{\mathcal{H}-1} \omega_t') \cap \omega_{\mathcal{H}}^{>o_{\mathcal{H}}}$ is obtained from θ, and Y_j appears in $\bigcap_{t=1}^{\mathcal{H}-1} \omega_t'$, then (97) holds, as well.

Proof. This follows from (76), (79), step (6.1) and (87). ■

First, we will prove the base for the induction.

Lemma 4.25 *Suppose $\sum_{j=1}^{n} \alpha_j x_j \geq \alpha_0$ is an inequality valid for $\mathcal{F}$ of pitch ≤ 1. Let $\tilde{u}$ be any column of $\tilde{U}$. Then $\sum_{j=1}^{n} \alpha_j \tilde{u}[\mathrm{Y}_j] - \alpha_0 \tilde{u}[\mathcal{F}] \geq 0$.*

Proof. This follows from Corollary 4.23 and the fact that due to Step 6 of the algorithm, every column of $\tilde{U}$ satisfies each of the set covering constraints $Ax \geq e$, homogenized. ■

In order to develop the inductive proof outlined in the previous section, we need to parameterize those expressions generated by the algorithm. This will be done through a value that essentially counts (in reverse) the number of negations in an expression. In what follows, we use the following notation. Consider expressions $\theta = \bigcap_{t=1}^{\mathcal{H}} \omega_t$ and $\theta_o^{\#} = \bigcap_{t=1}^{\mathcal{H}} \omega_t'$ generated in step 4 of the algorithm, where each ω_t' is a negation of order o_t of ω_t.

Definition 4.26 *We will say that $\theta_o^{\#}$ is* **positive** *if $0 < o_t$ for all t, $1 \leq t \leq \mathcal{H}$.*

We let

$$\mathrm{Y}(\theta_o^{\#}) = \left\{1 \leq j \leq n \,:\, \mathrm{Y}_j \text{ appears in } \theta_o^{\#}\right\}.$$

Note that since we are dealing with a set-covering problem, every $j \in \mathrm{Y}(\theta_o^{\#})$ appears as N_j in some ω_t and is thus negated in constructing $\theta_o^{\#}$, i.e., it appears as Y_j in ω_t'. Finally, we write

$$\rho(\theta_o^{\#}) \quad = \quad k - \sum_{t=1}^{\mathcal{H}} o_t. \tag{98}$$

Note: strictly speaking, we should write $\rho(\theta_o^{\#}, \theta)$, but as θ will always be clear from the context we will omit it. Also note that $\rho(\theta_o^{\#}) < k$ when $\theta_o^{\#}$ is positive. Furthermore, it is important to notice that the definition of ρ uses $\theta_o^{\#}$ itself, and not its simplification. We will also write

$$\mathrm{Y}(\mathcal{F}) = \emptyset,$$

and

$$\rho(\mathcal{F}) = k.$$

The ρ parameter will be used to drive the induction. Next we have the main result, which implies Theorem 1.2.

Theorem 4.27 *Consider an expression γ so that either $\gamma = \mathcal{F}$, or γ is some positive $\theta_o^{\#}$ enumerated in Step 4. Let*

$$\sum_{j=1}^{n} \alpha_j x_j \quad \geq \quad \alpha_0 \tag{99}$$

be an inequality valid for $\mathcal{F}$ with $\pi(\alpha, \alpha_0) \leq \rho(\gamma)$. Then the column $\tilde{u}$ of $\tilde{U}$ corresponding to γ satisfies

$$\sum_{j=1}^{n} \alpha_j \tilde{u}[\mathrm{Y}_j] - \alpha_0 \tilde{u}[\mathcal{F}] \quad \geq \quad 0. \tag{100}$$

Proof. The proof will be by induction on $\rho(\gamma)$, which, as just discussed, is positive. Lemma 4.25 therefore handles the base of this induction.

By Proposition 4.24, if $j \in \mathrm{Y}(\gamma)$ then $\tilde{u}[\mathrm{Y}_j] = \tilde{u}[\mathcal{F}]$, and as a result

$$\sum_{j \in \mathrm{Y}(\gamma)} \alpha_j \tilde{u}[\mathrm{Y}_j] = \left(\sum_{j \in \mathrm{Y}(\gamma)} \alpha_j \right) \tilde{u}[\mathcal{F}].$$

Consequently, if we could also prove that the inequality

$$\sum_{j \notin \mathrm{Y}(\gamma)} \alpha_j X[\mathrm{Y}_j] \; - \; \left(\alpha_0 - \sum_{j \in \mathrm{Y}(\gamma)} \alpha_j \right) X[\mathcal{F}] \quad \geq \quad 0 \tag{101}$$

is satisfied by setting $X = \tilde{u}$, we would complete the proof of the theorem. This is what we will do next.

Rewrite (101) as

$$\sum_{j=1}^{n} \bar{\alpha}_j X[\mathrm{Y}_j] \; - \; \bar{\alpha}_0 X[\mathcal{F}] \quad \geq \quad 0, \tag{102}$$

where, for $1 \leq j \leq n$

$$\bar{\alpha}_j = \begin{cases} \alpha_j, & \text{if } j \notin \mathrm{Y}(\gamma) \\ 0, & \text{otherwise.} \end{cases} \tag{103}$$

and $\bar{\alpha}_0 = \alpha_0 - \sum_{j \in \mathrm{Y}(\gamma)} \alpha_j$. The inequality

$$\bar{\alpha}^T x \quad \geq \quad \bar{\alpha}_0 \tag{104}$$

is clearly valid for $\mathcal{F}$ (since (99) is); without loss of generality $\bar{\alpha}_0 \geq 0$ (else (102) follows trivially) and by definition of pitch, $\pi \doteq \pi(\bar{\alpha}, \bar{\alpha}_0) \leq \pi(\alpha, \alpha_0)$.

First we claim that, in the case that $\gamma = \theta^{\#}_{o}$, without loss of generality γ is a conjunction of strictly fewer than $k-1$ negated walls. Say $\gamma = \bigcap_{t=1}^{\mathcal{H}} \omega'_t$, where ω'_t is a negation of ω_t of order o_t for $1 \leq t \leq \mathcal{H}$. We want to argue that, without loss of generality, $\mathcal{H} < k-1$. But γ is positive, by assumption in this theorem. Thus $o_t > 0$ for $1 \leq t \leq \mathcal{H}$. Then by definition of ρ, we have that $\rho(\gamma) \leq k - \mathcal{H}$. Hence, if $\mathcal{H} \geq k-1$ the Theorem is proved by Lemma 4.25. In what follows we assume that $\mathcal{H} < k-1$.

Suppose we apply Proposition 4.22 to (104). Let $\mathcal{C}$ denote the resulting subset of rows of A. Thus we obtain a set of obstructions, one from each row of $\mathcal{C}$; since $|\mathcal{C}| \leq \pi \leq \rho(\gamma) \leq k$, the wall ω derived from this set of obstructions will be enumerated in Step 3.

Hence, whether $\gamma = \mathcal{F}$ or $\gamma = \theta^{\#}_{o}$, in Step 5 the algorithm will impose either

$$X[\gamma] \quad = \quad \sum_{r=0}^{|\omega|} \left(\sum_{\beta \in \mathcal{N}(\omega, r)} X[\gamma \cap \beta] \right) \tag{105}$$

if $|\omega| < \rho(\gamma)$ (c.f. (98), and see the rule that determines for example that 5(a) applies), or

$$X[\gamma] \quad = \quad X[\gamma \cap \omega^{>R}] + \sum_{r=0}^{R} \left(\sum_{\beta \in \mathcal{N}(\omega, r)} X[\gamma \cap \beta] \right) \tag{106}$$

otherwise, where $R = \rho(\gamma) - 1$.

Now one of (105) or (106) applies; the algorithm enforces that constraint on all columns of $\tilde{U}$, hence by Proposition 4.13

$$col[\gamma] \quad = \quad \sum_{r=0}^{|\omega|} \left(\sum_{\beta \in \mathcal{N}(\omega, r)} col[\gamma \cap \beta] \right) \tag{107}$$

or

$$col[\gamma] \quad = \quad col[\gamma \cap \omega^{>R}] + \sum_{r=0}^{R} \left(\sum_{\beta \in \mathcal{N}(\omega, r)} col[\gamma \cap \beta] \right) \tag{108}$$

will hold.

We will next show, using induction, that all columns of the form $col[\gamma \cap \beta]$ arising from values $r > 0$ satisfy (102). Using a special proof, we will show the same fact for the case $r = 0$. Finally, assuming (108) applies, we will show the same for $col[\gamma \cap \omega^{>R}]$. This will complete the proof of the theorem.

Case $r > 0$**.** Consider a fixed value $r > 0$, and consider a particular $\beta \in \mathcal{N}(\omega, r)$. Thus, β is of the form

$$\bigcap_{j \in S} \mathrm{Y}_j \cap \bigcap_{j \in T} \mathrm{N}_j$$

where $|S| = r$, $\omega = \bigcap_{j \in S \cup T} \mathrm{N}_j$, and S and T are disjoint. Let $\hat{u} = col[\gamma \cap \beta]$. We want to show that $\hat{u}$ satisfies (102).

By Proposition 4.24, $\hat{u}[\mathrm{Y}_j] = \hat{u}[\mathcal{F}]$ for each $j \in S$. Thus, it suffices to prove that $\hat{u}$ satisfies the inequality

$$\sum_{j \notin S} \bar{\alpha}_j X[\mathrm{Y}_j] - \left(\bar{\alpha}_0 - \sum_{j \in S} \bar{\alpha}_j \right) X[\mathcal{F}] \quad \geq \quad 0 \tag{109}$$

in order to prove that $\hat{u}$ satisfies (102).

But notice that by construction in Proposition 4.22, all rows in the set $\mathcal{C}$ have support contained in $suppt(\bar{\alpha})$. Hence $\bar{\alpha}_j > 0$ $\forall j \in S$, and as a result the pitch of (109) is at most the pitch of (104) minus r, i.e., $\leq \rho(\gamma) - r < \rho(\gamma)$. Consequently, if we can show that

$$\rho(\gamma \cap \beta) \geq \rho(\gamma) - r$$

then by induction we will have that $\hat{u}$ satisfies (109).

Consider the case $\gamma = \theta^{\#}_o$. We have that $\gamma \cap \beta$ is obtained from $\theta \cap \omega$ by (i) negating literals in θ to obtain $\theta^{\#}$, and (ii) since $\gamma \in \mathcal{N}(\omega, r)$, by negating exactly r literals in ω. Thus $\rho(\gamma \cap \beta) = \rho(\gamma) - r$, as desired. In the case $\gamma = \mathcal{F}$ this is also clear.

Case $r = 0$**.** Let $\hat{u} = col[\gamma \cap \beta] = col[\gamma \cap \omega]$. For convenience, we restate here the properties satisfied by the set of rows $\mathcal{C}$ which we are using in this proof, produced by applying Proposition 4.22 to $\bar{\alpha}^T x \geq \bar{\alpha}_0$:

$$\begin{array}{ll} (i) & A_i \subseteq suppt(\bar{\alpha}) \ \ \forall \ i \in \mathcal{C} \\ (ii) & \Delta_i \doteq A_i - \bigcup_{r \in \mathcal{C} - i} A_r \neq \emptyset, \quad \forall \ i \in \mathcal{C}, \ \text{and} \\ (iii) & \sum_{i \in \mathcal{C}} \min \{ \bar{\alpha}_j \ : \ j \in \Delta_i \} \ \geq \ \bar{\alpha}_0. \end{array}$$

Further each $i \in \mathcal{C}$ gives rise to one obstruction and ω is the wall derived from these obstructions. Thus, $\omega = \bigcap_{j \in J} \mathrm{N}_j$, where $J \subseteq \{1, 2, \cdots, n\}$ is the set of all j appearing in at least two A_i, $i \in \mathcal{C}$. Note that, by (76) and Step 6 of the algorithm,

$$\hat{u}[\mathrm{N}_j] \ \geq \ \hat{u}[\gamma \cap \omega] \ = \ \hat{u}[\mathcal{F}] \quad \forall \ j \ \in \ J,$$

and consequently

$$\hat{u}[\mathrm{Y}_j] \quad = \quad 0, \quad \forall \ j \ \in \ J, \tag{110}$$

Further, consider any row $i \in \mathcal{C}$. By step (6.2) of the algorithm,

$$\sum_{j=1}^{n} a_{ij} \hat{u}[\mathrm{Y}_j] - \hat{u}[\mathcal{F}] \geq 0.$$

Combining this with (110) we obtain

$$\sum_{j \in \Delta_i} a_{ij} \hat{u}[\mathrm{Y}_j] - \hat{u}[\mathcal{F}] \geq 0$$

and consequently

$$\sum_{j \in \Delta_i} \bar{\alpha}_j \hat{u}[\mathrm{Y}_j] \quad \geq \quad (\min \{ \bar{\alpha}_j \ : \ j \in \Delta_i \}) \, \hat{u}[\mathcal{F}]. \tag{111}$$

By construction, the sets Δ_i are pairwise disjoint. So if we sum (111) over all $i \in \mathcal{C}$ we obtain:

$$\begin{aligned} \sum_{j=1}^{n} \bar{\alpha}_j \hat{u}[\mathrm{Y}_j] \quad & \geq \quad \left(\sum_{i \in \mathcal{C}} \min \{ \bar{\alpha}_j \ : \ j \in \Delta_i \} \right) \hat{u}[\mathcal{F}] \\ & \geq \quad \bar{\alpha}_0 \hat{u}[\mathcal{F}], \end{aligned} \tag{112}$$

where the last inequality follows by property (iii) of the set $\mathcal{C}$. This concludes the proof in this case.

Case $\omega^{>R}$. Let $\hat{u} = col[\gamma \cap \omega^{>R}]$ and write $\omega = \bigcap_{j \in J} \mathrm{N}_j$. Recall that $R = \rho(\gamma) - 1$. Now,

$$\begin{aligned} \sum_j \bar{\alpha}_j \hat{u}[\mathrm{Y}_j] &= \sum_{j \in suppt(\bar{\alpha})} \bar{\alpha}_j \hat{u}[\mathrm{Y}_j] \\ &\geq \sum_{j \in J} \bar{\alpha}_j \hat{u}[\mathrm{Y}_j]. \end{aligned} \tag{113}$$

Further, in step (6.2), the algorithm imposes the constraint (80), applied to $\gamma \cap \omega^{>R}$, i.e., on the column $\hat{u}$. This amounts to imposing:

$$\sum_{j \in J} \hat{u}[\mathrm{Y}_j] \; - \; (1+R)\hat{u}[\gamma \cap \omega^{>R}] \;\; \geq \;\; 0. \tag{114}$$

Now $1 + R = \rho(\gamma)$, by definition of R. Also, by step (6.1), $\hat{u}[\gamma \cap \omega^{>R}] = \hat{u}[\mathcal{F}]$. So (114) is simply:

$$\sum_{j \in J} \hat{u}[\mathrm{Y}_j] \; - \; \rho(\gamma)\hat{u}[\mathcal{F}] \;\; \geq \;\; 0. \tag{115}$$

Thus, combining (113), (115) we obtain that

$$\sum_j \bar{\alpha}_j \hat{u}[\mathrm{Y}_j]$$

is at least $\hat{u}[\mathcal{F}]$, times the sum of the $\rho(\gamma)$ smallest positive coefficients $\bar{\alpha}_j$. But we know that $\pi(\bar{\alpha}, \bar{\alpha}_0) \leq \pi(\alpha, \alpha_0)$ which by assumption is at most $\rho(\gamma)$. Thus

$$\sum_j \bar{\alpha}_j \hat{u}[\mathrm{Y}_j] \geq \bar{\alpha}_0 \hat{u}[\mathcal{F}],$$

as desired. This completes the proof. ■

Comment 4.28 *Note that positive-semidefiniteness (Steps 7 and 8 of the algorithm) is not used in this theorem.*

4.5 Positive-semidefiniteness and set packing problems

As discussed previously, positive-semidefiniteness arises as a natural feature in all of the lifting algorithms we described. An important question concerns how much positive-semidefiniteness provably adds to the strength of the algorithms. This question was taken up in [LS91], with special attention paid to vertex-packing problems. Among many other related results, it was shown in [LS91] that the N_+-rank of clique inequalities is 1, whereas their N-rank is in general much higher, thus showing that positive-semidefiniteness can indeed help in a concrete way.

Before describing our results, we note that one can always rewrite a linear inequality in the form

$$\sum_{j \in J^+} a_j x_j + \sum_{j \in J^-} a_j (1 - x_j) \geq b$$

where all the a_j are nonnegative and $J^+ \cap J^- = \emptyset$. In the context of the Σ-algorithms, this is a natural step since we introduce both variables Y_j and N_j for every j. By reformulating the problem in this manner one obtains more general results than those we will describe here. See [Z03]. In particular, one can show how some high-pitch valid inequalities for set-covering problems are enforced by the Σ^k-algorithm for small k (we will touch on this again at the end of this section). However, for the sake of simplicity, in this section we will only explicitly deal with set-packing problems.

Consider the following packing system. Let I be an index set of cardinality at least three, and suppose that for each $i \in I$ we have a set $S_i \subseteq \{1, 2, \cdots, n\}$ such that the S_i are pairwise disjoint.

Consider the system of constraints

$$x(S_i) + x(S_j) \;\; \leq \;\; |S_i| + |S_j| \; - 1, \;\; \forall \, i \neq j. \tag{116}$$

where, as we use the notation $v(H) = \sum_{j \in H} v_j$. Thus, if $|S_i| = 1$ for all i, we obtain a vertex-packing system for an t-clique. Write $S = \cup_i S_i$, $N = |S| = \sum_i |S_i|$ and $t = |I|$.

The inequality

$$\sum_{j \in S} x_j \quad \leq \quad N - t + 1 \tag{117}$$

is valid: for $x \in \{0,1\}^S$, if $\sum_{j \in S} x_j = N - t + 2$ then there are at least 2 distinct sets S_i such that $x_j = 1$ for all $j \in S_i$, violating (116). We will refer to (117) as the *set-clique* constraint.

Lemma 4.29 *Consider a system of inequalities that contains, possibly among others, all constraints (116) for some family $\{S_i : i \in I\}$ of pairwise disjoint sets. Then the vector X produced by the Σ^2-algorithm, using the positive-semidefiniteness condition 7, satisfies (117).*

Proof. Pick a pair i, k of distinct elements of I. Clearly,

$$\bigcap_{j \in S_i} \mathrm{Y}_j \cap \bigcap_{j \in S_k} \mathrm{Y}_j$$

is an obstruction to that constraint (116) corresponding to the pair i, k.

Consequently, for any $i \in I$, $\bigcap_{j \in S_i} \mathrm{Y}_j$ is a wall enumerated by the algorithm, and hence a tier, and, by condition (78) imposed by the algorithm in Step 4, we have

$$X[\bigcap_{j \in S_i} \mathrm{Y}_j] \quad \geq \quad \sum_{j \in S_i} X[\mathrm{Y}_j] \; - \; |S_i| \; + \; 1, \tag{118}$$

and as a result

$$\sum_{i \in I} X[\bigcap_{j \in S_i} \mathrm{Y}_j] \quad \geq \quad \sum_{j \in \cup_i S_i} X[\mathrm{Y}_j] \; - \; N \; + \; t. \tag{119}$$

Suppose that in addition we could show that

$$1 \quad \geq \quad \sum_{i \in I} X[\bigcap_{j \in S_i} \mathrm{Y}_j] \tag{120}$$

holds. Then, using (119) we obtain the desired result. The rest of the proof shows that (120) holds.

Let U be the matrix produced by the algorithm. For $i \in I$,

$$U[\mathcal{F}, \bigcap_{j \in S_i} \mathrm{Y}_j] \; = \; U[\bigcap_{j \in S_i} \mathrm{Y}_j, \mathcal{F}] \; = \; U[\bigcap_{j \in S_i} \mathrm{Y}_j, \bigcap_{j \in S_i} \mathrm{Y}_j] \; = \; X[\bigcap_{j \in S_i} \mathrm{Y}_j],$$

by condition (6.1) imposed by the algorithm in Step 6. Further, for a pair i, k of distinct indices in I,

$$U[\bigcap_{j \in S_i} \mathrm{Y}_j, \bigcap_{j \in S_k} \mathrm{Y}_j] \; = \; U[\bigcap_{j \in S_k} \mathrm{Y}_j, \bigcap_{j \in S_i} \mathrm{Y}_j] \; = \; 0,$$

by condition (6.5) imposed by the algorithm. Finally, $U[\mathcal{F}, \mathcal{F}] = 1$.

Let $\mu \in R^{\mathcal{V}}$ be defined by

$$\mu_\beta = \begin{cases} -1, & \text{if } \beta = \bigcap_{j \in S_i} \mathrm{Y}_j \text{ for some } i \in I, \\ 1, & \text{if } \beta = \mathcal{F}, \\ 0, & \text{otherwise.} \end{cases} \tag{121}$$

Then $\mu^T U \mu = 1 - \sum_{i \in I} X[\bigcap_{j \in S_i} \mathrm{Y}_j]$. Since $U \succeq 0$ this concludes the proof. ■

In contrast to this result, we also have:

Lemma 4.30 *Suppose $|S_i| = t$ for every $i \in I$. For $t \geq 3$ the N_+-rank and the Sherali-Adams rank of the set-clique constraint are both at least $t - 2$.*

Proof sketch. This proof is similar to those in the Appendix. Note that the x vector has $n = t^2$ coordinates if the S_i comprise all the variables. Define

$$T = (t-2)t^{t-1} + 1, \;\; w_1 = \frac{t-2}{t} T^{-1} \quad \text{and} \quad w_2 = \frac{2}{(t-1)t^{t-1}} T^{-1}. \tag{122}$$

Consider the function λ that assigns, to each $v \in \{0,1\}^n$ the value $\lambda(v)$ defined using the following rules:

(a) Suppose $\sum_{j \in S_i} v_j = t-1$, for $t-1$ indices $i \in I$; and $\sum_{j \in S_i} v_j = t$ for the remaining index i. Then $\lambda(v) = w_1$.

(b) Suppose $\sum_{j \in S_i} v_j = t-1$, for $t-2$ indices $i \in I$; and $\sum_{j \in S_i} v_j = t$ for each of the two remaining indices i. Then $\lambda(v) = w_2$.

(c) In all other cases we set $\lambda(v) = 0$.

Note that points of type (a) are feasible for system (116), while those of type (b) are not. Also, there are t^t points of type (a), and $\frac{t(t-1)}{2} t^{t-2} = \frac{(t-1)t^{t-1}}{2}$ points of type (b), and thus

$$\sum_{v \in \{0,1\}^n} \lambda(v) \;\; = \;\; 1. \tag{123}$$

As a result, λ can be extended to a *measure* on $\{0,1\}^n$, as follows: for a subset $A \subseteq \{0,1\}^n$, define

$$\lambda(A) = \sum_{v \in A} \lambda(v).$$

Further, if we define the vector $\bar{x} \in R^n$ by

$$\bar{x}_j = \lambda(\{v \in \{0,1\}^n \, : \, v_j = 1\}) \; = \; \sum_{v_j = 1} \lambda(v)$$

for $1 \le j \le n$, then a quick calculation shows that $\bar{x}$ satisfies all inequalities (116) but violates (117).

Further, $\bar{x}$ can be lifted to a vector $\bar{y}$ indexed by the subset lattice of $\{1, 2, \cdots, n\}$. For $p \subseteq \{1, 2, \cdots, n\}$ set

$$\bar{y}_p = \sum \{\lambda(v) \, : \, v_j = 1 \quad \forall \, j \in p\},$$

and a calculation shows that $\bar{y}$ (and the corresponding matrix $U^{\bar{y}}$) satisfy the requirements for both the rank $(t-3)$ N_+ operator and the level $(t-3)$ Sherali-Adams operator. ■.

This result can be strengthened, see [Z03]. In particular, when $|S_1| = |S_2| = \lfloor \frac{n}{3} \rfloor$, and $|S_i| = 1$ for all other i, then the N_+ and Sherali-Adams rank of (117) is $\lfloor \frac{n}{3} \rfloor$.

The results above can be generalized in a different way. Let G be an undirected graph. Suppose that to each vertex i we assign a nonempty set of 0-1 variables $x_k, k \in S_i$, where the index sets S_i are pairwise disjoint, and we impose the constraints:

$$x(S_i) + x(S_j) \le |S_i| + |S_j| - 1, \;\; \forall \{i,j\} \in E(G).$$

This construction generalizes the standard vertex-packing polyhedron. A question of interest concerns the rank of the corresponding generalizations of classical valid inequalities. As we have shown above, the set-clique inequality has unbounded N_+-rank, but is guaranteed by the Σ^2- algorithm. In a similar way one obtains, using the obvious notation (see [Z03] for details), and using the appropriate set-generalization of Lemma 1.5 of [LS91]:

Lemma 4.31 *The set-odd-hole, set-odd-antihole, and set-odd-wheel inequalities are all guaranteed by the* Σ^2*-algorithm.* ■

4.6 Further remarks

Consider the feasible system $\mathcal{F}$ introduced in [CCH89]:

$$\sum_{j \in S} x_j + \sum_{j \notin S} (1 - x_j) \;\; \geq \;\; \frac{1}{2}, \quad \forall\, S \subseteq \{1, \cdots, n\} \tag{124}$$

$$x \in \{0,1\}^n \tag{125}$$

This system was was analyzed in [CCH89], [CD01], [GT01], [L01] to show that either n or $n-1$ iterations of various procedures (Sherali-Adams, the N_+ operator) are required to prove that $\mathcal{F}$ is empty.

Here we will show that running the Σ^3-algorithm proves the same result. We will denote by $C(S)$ the inequality (124) corresponding to the set S. Note that corresponding to $C(S)$ there is the (unique) 0-small obstruction

$$\bigcap_{j \in S} \mathrm{N}_j \cap \bigcap_{j \notin S} \mathrm{Y}_j$$

and conversely, any expression of this form is the 0-small obstruction to some constraint. In fact, each expression of this form corresponds to an atom of the subset algebra of $\{0,1\}^n$ and every atom gives rise to such an expression; thus in a sense our result is not surprising.

Lemma 4.32 *Suppose we run the Σ^3-algorithm. Consider an expression of the form*

$$\gamma = \bigcap_{j \in A} \mathrm{N}_j \cap \bigcap_{j \in B} \mathrm{Y}_j$$

where $A \cap B = \emptyset$ and $0 < |A \cup B|$.

(a) If $|A \cup B| < n$, γ is negation of order one of a wall enumerated by the algorithm.

(b) γ is a tier enumerated by the algorithm.

Proof. Assume without loss of generality that $A \neq \emptyset$ and let $k \in A$. (a) The wall derived from the obstruction to $C(\{1, 2, \cdots, n\} - B - \{k\})$ and the obstruction to $C(A - \{k\})$ is

$$\bigcap_{j \in A - \{k\}} \mathrm{N}_j \cap \bigcap_{j \in B \cup \{k\}} \mathrm{Y}_j,$$

and γ is a negation of order one of this wall. (b) In this case, $\gamma = \left(\bigcap_{j \in A-k} \mathrm{N}_j \cap \bigcap_{j \in B} \mathrm{Y}_j \right) \cap \mathrm{N}_k$ and by (a), applied to the expression within parentheses, we are done. ■

Theorem 4.33 *The linear system produced by the Σ^3-algorithm is infeasible.*

Proof. Let $\bar{X}$ be the vector produced by the Σ^3 and $\bar{U}$ the corresponding matrix. We will prove that any expression of the form

$$\gamma = \bigcap_{j \in A} \mathrm{N}_j \cap \bigcap_{j \in B} \mathrm{Y}_j$$

where $A \cap B = \emptyset$ is such that

$$\bar{X}[\gamma] \;\; = \;\; 0. \tag{126}$$

Pending the proof of this claim, this completes the proof of the theorem, because by picking any j with $1 \leq j \leq n$, and using step 1 of the algorithm, we will get

$$\bar{X}[\mathcal{F}] \; = \; \bar{X}[\mathrm{Y}_j] + \bar{X}[\mathrm{N}_j] \; = \; 0,$$

and since the algorithm also enforces $\bar{X}[\mathcal{F}] = 1$ (step 6) we have an infeasible system, as desired.

We will prove equation (126) by induction on $n - |A \cup B|$. Suppose first that $n = |A \cup B|$. Since γ is an obstruction, the algorithm enforces equation (77) and we conclude $\bar{X}[\gamma] = 0$.

For the general inductive step, suppose we have proved the result for expressions with more symbols than γ. Pick any index $h \notin A \cup B$. Consider the expression

$$\mathrm{Y}_h \bigcap \gamma.$$

By Lemma 4.32 this expression is a wall, and thus a tier; in step 4(iii)(a) the algorithm will impose:

$$\bar{X}[\gamma] = \bar{X}[\mathrm{Y}_h \bigcap \gamma] + \bar{X}[\mathrm{N}_h \bigcap \gamma]$$

and by the inductive hypothesis we are done. ■

Note: Laurent [L01] has conjectured that the Lasserre rank of (124 - 125) is $n - 1$. Also see [L02].

One can also show that the inequality shown in [CD01] to have N_+-rank n is implied by the Σ^2 algorithm:

Lemma 4.34 *Consider the region* $\{x \in \{0,1\}^n : \sum_j x_j \geq \frac{1}{2}\}$. *The* Σ^2*-algorithm proves* $\sum_j x_j \geq 1$.

Proof. This follows from constraint (75) imposed in Step 2 of the algorithm. ■

5 Future work

One critical area that we plan to address concerns how to make our algorithms practical. A simple idea would be to project our formulations to the space of the original variables. However, we prefer the use of additional variables indexed by the subset algebra as they reveal useful structure of the problem. We note that Haus, Köppe and Weismantel [HKW01] have introduced algorithms for solving general integer programs which rely on explicitly adding new variables, though in a rather different form than our algorithms. Also, we point out the result in Section 3.3 – a polynomial enlargement of a formulation can imply an exponential number of facets, even without requiring positive-semidefiniteness (already known in a different context).

A better idea would be to apply column generation so as to implement the Σ^k-algorithm. More precisely: the real difficulty in implementing the algorithm, even for small values of k, is that the number of variables and constraints (though polynomially bounded) may be too large to actually ask a Linear Programming solver to handle (let alone a semidefinite programming solver). However, the number may be small enough that we can generate the formulation, at least in some implicit format. Column generation would then be used to select those variables and constraints that we actually want to use. Roughly speaking, then, we obtain algorithms that enlarge formulations by adding variables and constraints related to elements of the algebra, but without necessarily following the formal hierarchies described in this paper.

Another topic that relates to this work concerns problems that are not, strictly speaking, integer programming formulations in the standard sense, but are instead feasibility problems that are given by lists of forbidden configurations, i.e., obstructions. The Σ-algorithms can be directly adapted to handle such problems, and, more generally, to satisfiability problems. See [Z03]. There is a clear and well-known connection between set theory and logic, and, further, there is a rich literature on the connection between logic and optimization, and on incorporating techniques of logic to integer programming. See for example [H00], also see [BH02]. This topic appears to be underlaid by Balas' work on disjunctive programming, but we have not found in the literature prior work relating logic to, in particular, the combination of the Lovász-Schrijver technique of matrix lifting and the use of variables indexed by the subset algebra.

Finally, a question of interest concerns the relationship between our cuts and Gomory cuts. This question has some substance since it is known that Gomory mixed-integer cuts are equivalent to general disjunctive cuts. See [CL01b].

Acknowledgment We thank Monique Laurent and Bruce Shepherd for many useful comments. This paper was partially written at Bell Laboratories, whose support to us is gratefully acknowledged.

Appendix – Circulant matrices and set covering

In this section we consider a set-covering problem defined by a full-circulant matrix, i.e., a feasible region of the form

$$\sum_{j \neq k} x_j \geq 1, \qquad \text{for each } \ k, \ 1 \leq k \leq n \tag{127}$$
$$x \in \{0,1\}^n$$

for $n > 1$, and show that the valid inequality

$$\sum_j x_j \quad \geq \quad 2 \tag{128}$$

has rank at least $n-3$ for a procedure stronger than the Sherali-Adams and the N_+ procedures combined.

First, we review the Sherali-Adams procedure. Suppose we have a feasible region of the form

$$\begin{aligned} Ax &\geq b \\ x &\in \{0,1\}^n. \end{aligned} \tag{129}$$

For $t \geq 0$, the level t Sherali-Adams operator is obtained as follows. For each polynomial of the form

$$f(Q,P) \doteq \prod_{j \in Q} x_j \cdot \prod_{j \in P} (1 - x_j)$$

where Q and P are disjoint subsets of $\{1, 2, \cdots, n\}$ and $|Q| + |P| = t$, we multiply each constraint r of (129) by $f(Q,P)$, obtaining a (valid) polynomial inequality of the form

$$(a_r x - b_r) f(Q,P) \quad \geq \quad 0. \tag{130}$$

(where a_r denotes the r^{th} row of A). In addition, we write the (valid) inequality

$$f(Q,P) \quad \geq \quad 0. \tag{131}$$

Next, we linearize the constraints (130) and (131): we replace x_j^2 with x_j for each j, and then each product of the form $\prod_{j \in H} x_j$ ($H \subseteq \{1, 2, \cdots, n\}$) is replaced by a new variable $y(H)$ (including the cases $H = \{j\}$ for some j).

Clearly, projecting the feasible region to the space of the x variables yields a valid relaxation. It is shown in [SA90] that increasing values of t yield stronger relaxations. In [L01] it is also shown that each level of the Sherali-Adams operator is at least as strong as the corresponding iteration of the Lovász-Schrijver operator N (though not N_+), and several known results imply that with $t = n$ we obtain the convex hull of the feasible region.

The following equivalent restatement of the Sherali-Adams operator will be useful below. This follows from the fact that the linearization step imposes $x_j^2 = x_j$, i.e., $x_j(1 - x_j) = 0$ for all j. This restatement includes some redundant steps, included to make the presentation easier, and is essentially given in [SA90], in any case.

Step S1. For each j, $1 \leq j \leq n$ create a new variable $\bar{x}_j$, and add the new constraint

$$x_j + \bar{x}_j \quad = \quad 1. \tag{132}$$

Step S2. For each polynomial of the form

$$g(Q,P) \doteq \prod_{j \in Q} x_j \cdot \prod_{j \in P} \bar{x}_j$$

where Q and P are disjoint subsets of $\{1, 2, \cdots, n\}$ and $|Q| + |P| = t - 1$, we multiply each constraint of type (129) or (132) by $g(Q,P)$, obtaining a (valid) polynomial inequality of the form

$$(\alpha^T x - \alpha_0) g(Q,P) \quad \geq \quad 0. \tag{133}$$

(note: in the case of (132) this is an equation). In addition, we write the (valid) inequality

$$g(Q,P) \quad \geq \quad 0. \tag{134}$$

Step S3. Linearize the constraints (133) and (134) by setting $x_j^2 = x_j$, $\bar{x}_j^2 = \bar{x}_j$ and $x_j \bar{x}_j = 0$ for all j, and then replacing each product of the form $\prod_{j \in K} x_j \prod_{j \in L} \bar{x}_j$ (K and L disjoint subsets of $\{1, 2, \cdots, n\}$ with possibly $|K \cup L| = 1$) with a new variable variable $w(K, L)$. [Note: In the case of (134) we simply get $w(Q, P) \geq 0$.]

We leave it to the reader to verify that indeed this is an equivalent restatement of the Sherali-Adams level t operator. Note: the variables $y(H)$ in the standard Sherali-Adams operators correspond to the variables $w(H, \emptyset)$ in this new formulation.

Now we return to the full-circulant set-covering example (127). Denote $E_n = \{1, 2, \cdots, n\}$. We will show that (128) has Sherali-Adams rank (at least) $n-3$. In order to do this, we have to produce nonnegative values $w(Q, P)$ for each pair of disjoint subsets Q, P of E_n with $|Q \cup P| \leq n-3$ which satisfy the constraints imposed in Steps S2 and S3, and which however violate (128). In detail, this is done as follows.

Consider a fixed pair of disjoint subsets Q, P of E_n with $|Q \cup P| \leq n-3$. Then when we multiply $x_j + \bar{x}_j - 1 = 0$ times $g(Q, P) \geq 0$ and linearize we get, when $j \notin Q \cup P$,

$$w(Q \cup j, P) + w(Q, P \cup j) - w(Q, P) \quad = \quad 0, \tag{135}$$

(where we abbreviate $\{j\}$ as j) and if either $j \in Q$ or $j \in P$ we get the identity $w(Q, P) = w(Q, P)$. When we multiply $\sum_{j \neq k} x_j - 1 \geq 0$ times $g(Q, P)$ and linearize, we get three different cases depending on k.

$$\sum_{j \in Q-k} w(Q,P) \quad + \sum_{j \in E_n - (Q \cup P)} w(Q \cup j, P) \quad - w(Q,P) \quad \geq \quad 0, \quad \text{if } k \in Q \tag{136}$$

$$\sum_{j \in Q} w(Q,P) \quad + \sum_{j \in E_n - (Q \cup P)} w(Q \cup j, P) \quad - w(Q,P) \quad \geq \quad 0, \quad \text{if } k \in P \tag{137}$$

$$\sum_{j \in Q} w(Q,P) \quad + \sum_{j \in E_n - (Q \cup P \cup k)} w(Q \cup j, P) \quad - w(Q,P) \quad \geq \quad 0, \quad \text{if } k \in E_n - (Q \cup P). \tag{138}$$

Note: there is no constraint (136) when $Q = \emptyset$. In addition, we require $w \geq 0$, and

$$w(\emptyset, \emptyset) \quad = \quad 1. \tag{139}$$

These constraints guarantee that the vector w satisfies the level t Sherali-Adams constraints. In order to show that the vector violates (128), we also require

$$\sum_j w(j, \emptyset) \quad < \quad 2, \tag{140}$$

The next series of Lemmas shows how to construct such values in a symmetric way, i.e., $w(Q, P) = w(Q', P')$ if $|Q| = |Q'|$ and $|P| = |P'|$.

Lemma 5.1 *Suppose there exist nonnegative values $z(q, p)$, for every pair q, p of nonnegative integers so that:*

$$\begin{aligned}
z(q+1,p) + z(q,p+1) - z(q,p) &\;=\; 0 \quad \text{if } q+p \leq n-1 && (141)\\
(q-2)z(q,p) + (n-q-p)z(q+1,p) &\;\geq\; 0 \quad \text{if } q+p \leq n-3 \text{ and } q > 0 && (142)\\
(n-p-1)z(1,p) - z(0,p) &\;\geq\; 0 \quad \text{if } p \leq n-3 && (143)\\
z(0,0) &\;=\; 1 && (144)\\
nz(1,0) &\;<\; 2. && (145)
\end{aligned}$$

Then there is a set of nonnegative values $g(Q, P)$ that satisfy (135), (136), (137), (138), (139) and (140). In addition, (135) holds for all disjoint Q, P with $|Q| + |P| \leq n-1$.

Proof. For each pair of disjoint subsets Q, P of E_n with $|Q \cup P| \leq n-3$, set

$$w(Q,P) = z(|Q|,|P|).$$

Then (135) follows from (141), (139) from (144) and (140) from (145). It remains to show that (136) - (138) hold.

In order to show this, consider a pair Q,P, and write $q = |Q|$ and $p = |P|$. Assume first that $q > 0$. We have to show that

$$(q-1)z(q,p) \quad +(n-q-p)z(q+1,p) \quad -z(q,p) \quad \geq \quad 0 \tag{146}$$

$$qz(q,p) \quad +(n-q-p)z(q+1,p) \quad -z(q,p) \quad \geq \quad 0 \tag{147}$$

$$qz(q,p) \quad +(n-q-p-1)z(q+1,p) \quad -z(q,p) \quad \geq \quad 0 \tag{148}$$

corresponding, respectively to (136), (137) and (138). Clearly, (146) is more binding than (147), and (146) is more binding than (148) because, by (135), $z(q,p) \geq z(q+1,p)$. Hence, (146), (147) and (148) hold if and only if (146) does; but this is equivalent to (142). The case $q > 0$ is completed.

Suppose next that $q = 0$. Then (136) is not a constraint, and corresponding to (137) and (138) we have to satisfy

$$(n-p)z(1,p) \quad -z(0,p) \quad \geq \quad 0, \quad \text{and} \tag{149}$$

$$(n-p-1)z(1,p) \quad -z(0,p) \quad \geq \quad 0. \tag{150}$$

and this holds because of (143). The proof is completed. ■

Now we will set about constructing values $z(q,p)$ that satisfy the hypotheses of Lemma 5.1. In fact, we will define $z(q,p)$ for all for all q,p with $q+p \leq n$. This is done as follows.

Define

$$\kappa \quad = \quad \frac{2}{n^2-n+2} \tag{151}$$

and set:

$$z(q,p) \quad = \quad 0, \quad \forall \ q \geq 3 \ \text{and} \ \ p \leq n-q \tag{152}$$

$$z(2,p) \quad = \quad \kappa, \quad \forall \ p \leq n-2 \tag{153}$$

$$z(1,p) \quad = \quad (n-1-p)\kappa, \quad \forall \ p \leq n-1 \tag{154}$$

$$z(0,p) \quad = \quad \left(\frac{(n-p)(n-1-p)}{2} + 1\right)\kappa, \quad \forall \ p \leq n \tag{155}$$

Now we have:

Lemma 5.2 *The choice of values z as in (151 - 155) satisfies the conditions of Lemma 5.1.*

Proof. We show first that (141) and (142) hold. These are homogeneous inequalities, hence satisfied for $q \geq 3$ by (152).

Suppose next that $q = 2$. If $p \leq n-3$ then $z(3,p) = 0$ and $z(2,p+1) = \kappa = z(2,p)$, hence (141) holds, and (142) is trivial.

Next, assume $q = 1$. Then $z(1,p) = (n-1-p)\kappa = (n-2-p)\kappa + \kappa = z(1,p+1) + z(2,p)$, so (141) holds. Also, (142) is equivalent to $-(n-p-1)\kappa + (n-p-1)\kappa \geq 0$, which is true.

Finally, consider the case $q = 0$. Then (141) is equivalent to:

$$(n-p-1) + \frac{(n-p-1)(n-2-p)}{2} + 1 \quad = \quad \frac{(n-p)(n-1-p)}{2} + 1, \tag{156}$$

which is true by inspection.

In order to show that (143) holds, notice that it is equivalent to

$$(n-p-1)^2 \quad \geq \quad \frac{(n-p)(n-1-p)}{2} + 1 \tag{157}$$

$$= \quad \frac{(n-p-1)^2 + (n-p+1)}{2} \tag{158}$$

which holds since $p \leq n-3$.

Next, note that $z(0,0) = \left(\frac{n(n-1)}{2} + 1\right)\kappa = 1$, so (144) holds.

Finally,

$$nz(1,0) \quad = \quad n(n-1)\kappa \; < \; 2 \tag{159}$$

by definition of κ, hence (145) holds, as well. ■

Note: if we set $\bar{x}_j = z(1,0) = (n-1)\kappa$ we obtain the fractional vector that fails to satisfy (128).

As a corollary of the above Lemmas we can now state:

Theorem 5.3 *The Sherali-Adams rank of constraint (128) is at least* $n-3$. ■

In fact, there is more than can be said about this inequality. As shown in [LS91], Theorem 5.3 implies that the N-rank of (128) is at least $n-3$. But how about its N_+-rank? In order to answer this question, consider the $2^n \times 2^n$-matrix W^w, whose rows and columns are indexed by subsets of E_n, and whose I, J entry is defined by:

$$W^w_{I,J} \quad = \quad w(I \cup J, \emptyset). \tag{160}$$

where the w are the values we construct in Lemma 5.1 using the z values we described above. We will prove below that $W^w \succeq 0$. Pending the proof of this fact, we can make some observations.

Consider a lifting procedure, denoted by $\hat{S}$, which is essentially like the Sherali-Adams procedure, with an additional positive semidefiniteness requirement. Unlike the standard Sherali-Adams procedure, which at level t will create variables $y(H)$ for each subset $H \subseteq E_n$ of cardinality $\min\{t+1, n\}$, $\hat{S}$ creates variables y_H for *every* subset $H \subseteq E_n$. Specifically, at level t,

(i) The Sherali-Adams level t constraints are imposed on y (thus, this only concerns subsets of cardinality t and $\min\{t+1, n\}$.

(ii) Denoting by W^y the $2^n \times 2^n$-matrix whose I, J entry is $y_{I \cup J}$, we require $W^y \succeq 0$.

Clearly, $\hat{S}$ is at least as strong as the Sherali-Adams procedure. In fact, with a little work one can show that $\hat{S}$ is at least as strong as $n-3$ rounds of the Lovász-Schrijver operator N_+. This follows because $W^y \succeq 0$ implies that several submatrices of W^y are also positive semi-definite (see, for example, [L01], Lemma 5).

Thus, $\hat{S}$ is at least as strong as Sherali-Adams and N_+ combined. In fact, it is far stronger: it solves vertex-packing problems at level 1 [Z03]. Moreover, one can also similarly show that our Lemmas above imply that at least $n-3$ rounds of $\hat{S}$ are needed to satisfy (128). We conjecture that the Lasserre-rank of (128) also grows as a function of n. Also, note that Letchford [Le01] describes a disjunctive procedure (on the original space of variables) that proves (127) in the full-circulant example.

Now we return to the proof of $W^w \succeq 0$. We will provide two proofs of this fact. The first one is actually longer, but we also feel it is more revealing.

Theorem 5.4 $W^w \succeq 0$.

Proof. We will use Theorem 2.2. Consider abstract events I_j, $1 \le j \le n$, and a probability measure Υ, defined as follows: on an atom

$$\alpha = (\bigcap_{j \in Q} I_j) \cap (\bigcap_{j \in E_n - Q} \bar{I}_j)$$

we set

$$\Upsilon(\alpha) \quad = \quad w(Q, E_n - Q), \tag{161}$$

and then we extend Υ to the subset-algebra generated by the I_j (i.e., for any event A, $\Upsilon(A)$ is the sum of $\Upsilon(\alpha)$, over all atoms α contained in A).

We claim that for each $Q \subseteq E_n$,

$$\Upsilon(\bigcap_{j \in Q} I_j) \quad = \quad w(Q, \emptyset), \tag{162}$$

from which the theorem follows, by Theorem 2.2. In order to prove this, we will prove the stronger statement that for disjoint subsets Q, P of E_n,

$$\Upsilon(\bigcap_{j \in Q} I_j \cap \bigcap_{j \in P} \bar{I}_j) \quad = \quad w(Q, P). \tag{163}$$

This claim will be proved by induction on $t = n - |Q| - |P|$. If $t = 0$ then $|P| = E_n - Q$ and $\bigcap_{j \in Q} I_j \cap \bigcap_{j \in P} \bar{I}_j$ is an atom; consequently (163) follows by definition (161).

Suppose now that $t > 0$, and let $k \in E_n - Q - P$. Then

$$\begin{aligned} \Upsilon(\bigcap_{j \in Q} I_j \cap \bigcap_{j \in P} \bar{I}_j) &= \Upsilon(\bigcap_{j \in Q \cup k} I_j \cap \bigcap_{j \in P} \bar{I}_j) + \Upsilon(\bigcap_{j \in Q} I_j \cap \bigcap_{j \in P \cup k} \bar{I}_j) && (164) \\ &= w(Q \cup k, P) + w(Q, P \cup k) \quad \text{(by induction)} && (165) \\ &= w(Q, P) \quad \text{(by (135))} && (166) \end{aligned}$$

as desired. The theorem is proved. ■

Theorem 5.5 $W^w \succeq 0$.

Proof. Consider the vector $\bar{x} \in R^n$ where $\bar{x}_j = z(1,0) = (n-1)\kappa$ for all j, which fails to satisfy (128), as shown in the proof of Lemma 5.2. Let L denote the subset lattice of E_n. By the results described in Section 2 (in particular, consider eqs. (5) and (6)) the theorem will follow if we can show that $\bar{x}$ can be lifted to a vector $y \in R^L$, such that y is a convex combination of zeta vectors for L, i.e.,

$$y = \sum_{r \in L} \alpha_r \zeta^r,$$

where $\alpha \ge 0$ and $\sum_{r \in L} \alpha_r = 1$, and such that $W^y = W^w$.

In order to achieve this end, we choose y in the obvious way: for $H \subseteq E_n$, set $y_H = w(H, \emptyset)$. Consider the following vector α:

1. $\alpha_\emptyset = \kappa$.
2. For every pair $p = \{i, j\}$ with $i \ne j$ we set $\alpha_p = \kappa$.

Note that

$$y_H = \begin{cases} 0 & \text{if } |H| > 2 \\ \kappa & \text{if } |H| = 2 \\ (n-1)\kappa & \text{if } |H| = 1 \\ 1 & \text{if } H = \emptyset \end{cases} \tag{167}$$

and the same holds true for $\sum_{r \in L} \alpha_r \zeta^r$: the first two statements trivially because of our choice of α, the third because each singleton is contained in exactly $n-1$ pairs, and lastly $\sum_{r \in L} \alpha_r = (1 + n(n-1)/2)\kappa = 1$. ■

A interesting question concerns the Lasserre rank of (128). Part of the difficulty here is that, despite the characterization in [L01], it is still nontrivial to provide a simple description of what the Lasserre algorithm actually does. Our numerical experiments suggest that the Lasserre rank of (128) is at least $n/4$, and we have produced examples where it is larger than 2.

References

[B75] E. Balas, Disjunctive Programs: Cutting Planes from Logical Conditions, in O.L. Mangasarian et al., eds., *Nonlinear Programming* **2** (1975) Academic Press, 279 – 312.

[B79] E. Balas, Disjunctive Programming. *Annals of Discrete Math.* **5** (1979) 3–51.

[BCC93] E. Balas, S. Ceria and G. Cornuéjols, A lift-and-project cutting plane algorithm for mixed 0-1 programs, *Mathematical Programming* **58** (1993), 295 – 324.

[BCC96] E. Balas, S. Ceria and G. Cornuéjols, Mixed 0-1 Programming by Lift-and-Project in a Branch-and-cut Framework, *Management Science* **42** (1996), 1229 – 1246.

[BN89] E. Balas and S.M. Ng, On the set covering polytope: I. All the facets with coefficients in $\{0, 1, 2\}$, *Mathematical Programming* **45** (1989), 1 – 20.

[CD01] W. Cook and S. Dash, On the matrix-cut rank of polyhedra, *Mathematics of Operations Research* **26** (2001), 19 – 30.

[BP83] E. Balas and W. Pulleyblank, The perfect matcheable subgraph polytope of a bipartite graph, *Networks* **13** (1983), 495 – 516.

[BP89] E. Balas and W. Pulleyblank, The perfect matcheable subgraph polytope of an arbitrary graph, *Combinatorica* **9** (1989), 321 – 327.

[B93] F. Barahona, Reducing matching to polynomial size linear programming, *SIAM Journal on Optimization* **3** (1993), 688 – 695.

[BH02] E. Boros and P.L. Hammer, Pseudo-Boolean Optimization, *Discrete Appl. Math.* **123** (2002), 155 – 225.

[CCH89] V. Chvátal, W. Cook and M. Hartmann, On cutting-plane proofs in combinatorial optimization, *Linear Algebra and its Applications* **114** (1989), 455 – 499.

[C74] L. Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions*, Reidel (1974).

[CL01] G. Cornuéjols and Y. Li, On the rank of mixed 0-1 polyhedra. In K. Aardal and A.M.H. Gerards (eds.), IPCO 2001, *Lecture Notes in Computer Science* **2081** (2001), 71 – 77.

[CL01b] G. Cornuéjols and Y. Li, Elementary closures for integer programs. *Oper. Res. Letters* **28** (2001), 1 – 8.

[DL97] M. Deza and M. Laurent, *Geometry of cuts and metrics*, Springer (1997).

[F66] W. Feller, *An introduction to probability theory and its applications*, Wiley, New York (1966).

[GT01] M.X. Goemans and L. Tunçel, When does the positive semidefiniteness constraint help in lifting procedures?, *Mathematics of Operations Research* **26** (2001), 796 – 815.

[H00] J. Hooker, Logic, Optimization and Constraint Programming, *Informs J. on Computing* (2002).

[HKW01] U.-U. Haus, M. Köppe and R. Weismantel, The Integral Basis Method for integer programming. *Mathematical Methods of Operations Research* **3** (2001), 353–361.

[L01b] J. B. Lasserre, An explicit exact SDP relaxation for nonlinear 0-1 programs, in *Lecture Notes in Computer Science* (K. Aardal and A.M.H. Gerards, eds.) (2001), 293-303.

[LS91] L. Lovász and A. Schrijver, Cones of matrices and set-functions and 0-1 optimization, *SIAM J. on Optimization* **1** (1991), 166-190.

[L01] M. Laurent, A Comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre Relaxations for 0-1 Programming, *Mathematics of Operations Research* **28** (2003), 470–496.

[L02] M. Laurent, Lower bound for the number of iterations in semidefinite relaxations for the cut polytope, *Mathematics of Operations Research* **28** (2003), 871–883.

[Le01] A.N. Letchford, On disjunctive cuts for combinatorial optimization, *J. of Comb. Opt.* **5** (2001) 299-315.

[NW88] G.L. Nemhauser and L.A. Wolsey, *Integer and Combinatorial Optimization*, Wiley, New York (1988).

[R64] G.-C. Rota, On the foundations of combinatorial theory I. Theory of Möbius functions, *Z. Wahrsch. Verw. Gebiete* **2** (1964), 340-368.

[SA90] S. Sherali and W. Adams, A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems, *SIAM J. on Discrete Mathematics* **3** (1990), 411-430.

[Z03] M. Zuckerberg, Ph.D. Thesis, Columbia University (2003).