CORC REPORT 2002-01
Subset Algebra Lift Operators for 0-1 Integer Programming*
DRAFT ONLY. PLEASE DO NOT DISTRIBUTE

Daniel Bienstock and Mark Zuckerberg
Columbia University
New York, NY 10027
July, 2002

Abstract

We extend the Sherali-Adams, Lovász-Schrijver, 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 stronger relaxation of a set-covering problem than 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.

1 Introduction

Consider a $0-1$ integer programming problem

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

where

$$F = \{ x \in \{0,1\}^n : Ax \geq b \}.$$  \hfill (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 $F$ is obtained. This bound on the number of iterations is tight ([CD01], [L01], also see [GT01]). 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, indirectly, in [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 $F$. One example of a result which is derived using our operators is the following:

**Theorem 1.1** 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 $k \geq 1$ be a fixed integer, and 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 $V_k$

*This research was partially funded by NSF award 2932-5538 (CORC)
denote the set of points in $[0, 1]^n$ that satisfy all inequalities in $V_k$. Then there is a polynomial-time algorithm for solving

$$\min \, c^T x \quad \text{s.t.} \quad x \in R_k$$

where $R_k$ is a certain polyhedron satisfying

$$\{x \in \{0, 1\}^n : Ax \geq e\} \subseteq R_k \subseteq V_k.$$ 

In other words, $R_k$ is a polyhedral relaxation of the set-covering problem, all of whose points satisfy all inequalities in the set $V_k$. We note that one can produce examples of set-covering problems with exponentially many facets with coefficients 0, 1, 2. Also, Balas and Ng have completely characterized the set of facet-defining inequalities with coefficients 0, 1, 2 [BN89].

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 \textit{pitch} is the minimum $t$ such that $\sum_{j=1}^t a_j \geq a_0$. Then we have:

\textbf{Theorem 1.2} 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 $k \geq 1$ be a fixed integer, and 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 is a polytope $Q_k \subseteq R^n$ satisfying:

(a) $\{x \in \{0, 1\}^n : Ax \geq e\} \subseteq Q_k,$
(b) $a^T x \geq a_0, \ \forall x \in Q_k$ and $(a, a_0) \in P_k$
(c) $Q_k$ is the projection to $R^n$ of a polytope with polynomially many constraints and variables.

Consequently, we can optimize $c^T x$ over $Q_k$ in polynomial time.

Further, 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 $\sum_j x_j \geq 2$ has rank at least $n - 3$ for a lifting operator stronger than the Sherali-Adams and the $N_+$ procedures combined. This is a constraint of pitch 2, and in fact it has rank 1 for our algorithms.

These results, and others on set-covering problems, are presented in Sections 3.3 and 4.5. Some results on set packing problems are given in Section 4.4. The development of our algorithms starts in Section 3.

2 Background

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

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

This second membership question may be easier to answer than the original one because the vectors $z(v)$ reveal information about $F$ in a more explicit way. As pointed out in [LS91], this basic idea was implicit in earlier work on specific combinatorial problems, e.g. [BP83], [B88], [BP89], and others.

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 $\hat{v} \in \{0, 1\}^{2n}$, where

(i) the entries of $\hat{v}$ are indexed by subsets of $E_n$, and
(ii) For $S \subseteq \{1, \ldots, n\}$, $v_S = 1$ iff $v_j = 1$ for all $j \in S$.

Clearly, for $1 \leq j \leq n$, $v(j) = v_j$, so each $v \in \{0, 1\}^n$ is mapped into a distinct column of the zeta matrix $Z$ of the subset lattice $L$ of $E_n$ (see [R64]). For simplicity, we will forgo the standard lattice-theoretic notation ($\subseteq$, $\lor$, $\land$) 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 $Z$: 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 $z^p$ is defined by:

$$z^p_q = \begin{cases} 1 & \text{if } q \subseteq p \\ 0 & \text{otherwise} \end{cases}$$ (2)

Thus $Z$ is (assuming the correct ordering of columns) upper triangular with 1s along the main diagonal, and therefore invertible. Its inverse, $M$, is called the Möbius matrix of the lattice. Note that each column of $Z$ 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.

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

$$\text{conv}(\hat{F}) = \{ y \in \mathbb{R}^L : m^p y \geq 0 \ \forall p \in \hat{F}; \ m^p y = 0 \ \forall p \notin \hat{F}; \ e^T M y = 1 \}.$$ (3)

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

(a) Is $y$ a linear combination of the columns of $Z$? This is always true because $Z$ is invertible.
(b) Is $y$ an affine combination of the columns of $Z$? This requires $y_0 = 1$.
(c) Is $y$ a convex combination of the columns of $Z$? This requires $y_0 = 1$ and $\lambda = My \geq 0$.
(d) Is $y \in \text{conv}(\hat{F})$? This requires $y_0 = 1$, $My \geq 0$, and $m_r y = 0$ for all $r \notin \hat{F}$.

We stress that (c) is already a nontrivial requirement (since $Z$ is invertible and therefore $\lambda$ is unique).

Even though condition (3) completely determines $\text{conv}(\hat{F})$, 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. Here we outline the approach in [LS91].

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

$$W^y = \sum_{r \in L} \lambda_r z^r (z^r)^T.$$ (4)

We have:

$$\lambda \geq 0 \ \text{iff} \ \ W^y \succeq 0.$$ (5)

This fact is clear in one direction; for the other implication note that $\lambda_r = m^r W^y m_r$ 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 $\bar{x} \in \text{conv}(\hat{F})$ then by (4) we can lift $\bar{x}$ to an $(n+1) \times (n+1)$-matrix $M^x$ with rows and columns indexed by singletons and the empty set, such that:

(a) $M^x \succeq 0$,
(b) $M^x$ is symmetric,
(c) The zeroth row of $M^x$ is equal to its diagonal, and 
(d) The zeroth row of $M^x$ is $(1, \ x^T)$.

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

Consider one of the constraints $a_i^T x \geq b_i$ in the definition of $\mathcal{F}$. Suppose $\bar{x} \in \text{conv}(\mathcal{F})$, and let $w$ be the $k^{th}$ column of $M^x$, $0 \leq k \leq n$. Then it is easy to see that $w$ satisfies the (homogenized) constraint
\begin{equation}
(-b_i, a_i^T )w = \sum_{j=1}^{n} a_{ij}w_j - b_iw_0 \geq 0
\end{equation}
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 (4) satisfy the constraint (6).

Further, using basic properties of lattices and the Möbius matrix one can also show that the vector obtained by subtracting any column of $M^x$ from the zeroth column also satisfies each homogenized constraint. This can also be directly obtained [Z02] by expanding the formulation to include the columns $x_j = 1 - x_j$ (1 \leq j \leq n) and studying the corresponding $W$ and $M$ matrices.

We summarize these facts as
(e) Let $w^k$ indicate the $k^{th}$ column of $M^x$. Then $(-b, A)w^k \geq 0$ for $0 \leq k \leq n$, and $(-b, A)(w_0 - w^k) \geq 0$ for $1 \leq k \leq n$.

In [LS91] the lifting $\bar{x} \to M^x$ 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 
\begin{equation}
Q = \{ x \in [0,1]^n : Ax \geq b \}
\end{equation}
then we can define $M(Q)$ (resp., $M_+(Q)$) as that subset of $Q$ for which a lifting $M^x$ exists satisfying (b)-(e) (resp., (a)-(e)). Clearly $M_+(Q) \subseteq M(Q) \subseteq Q$, and both $M_+(Q)$ and $M(Q)$ are convex sets (a polytope in the second case). As shown in [LS91] after iterating $n$ times we have $M^n(Q) = \text{conv}(\mathcal{F})$. 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):
\begin{equation}
W^y_{p,q} = y_{p \cup q}, \quad \forall \ p, q \in L
\end{equation}
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 $\leq k$ we are able to make some statements about coordinates of $y$ corresponding to larger lattice elements.

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], [L01b]. Let $a_i^T x \geq b_i$ be once more one of the inequalities in $Ax \geq b$. Define $a_i \in R^L$ by $\bar{a}_i^T = (-b_i, a_i^T, 0, 0, \cdots, 0)$ where the number of appended $0$s 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
\begin{equation}
\bar{a}_i \ast y \doteq W^y \bar{a}_i
\end{equation}
satisfies
\begin{equation}
\bar{a}_i \ast y = \sum_{r \in L} \lambda_r \zeta^r (\zeta^r)^T \bar{a}_i,
\end{equation}
If, in addition, $y$ is a lifting of $\bar{x} \in \text{conv}(F)$, then the sum in (9) can be restricted to elements of $\hat{F}$, and each such term $r$ satisfies $(\zeta^T r a_i \geq 0$. Consequently, $a_i \ast y$ 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) $W^{a_i \ast y} \geq 0$ for each constraint $a_i^T x \geq b_i$ in $Ax \geq b$.

Of course, (f) involves $y$, not $x$, but notice that by definition of $a_i$, we only need the first $n+1$ columns of $W^y$ in order to compute $a_i \ast y$. Hence (through another application of (7)) condition (f) may be approximated by requiring positive semidefiniteness of appropriate minors of $W^{a_i \ast y}$.

As shown in [L01], round $t \geq 1$ of the Sherali-Adams procedure requires that for each $U \subseteq E_n$ with $|U| \leq t$, the minor of $W^y$ (resp., $W^{a_i \ast y}$) 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 that the minor of $W^y$ (resp., $W^{a_i \ast y}$) corresponding to the set of rows and columns arising from all subsets of $E_n$ of cardinality $\leq t$ be positive-semidefinite. 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:

(1) $\Upsilon(A) \geq 0$ for all $A \subseteq \mathcal{W}$;

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

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

Note that (1)-(3) imply that $\Upsilon(\emptyset) = 0$, and that $\Upsilon$ is nondecreasing.

The following result is stated in [LS91] (p. 186, “Remark”). See [Z02] for a proof.

**Theorem 2.2** Let $L$ be the subset lattice of $E_n$. Suppose $K \subseteq \{0,1\}^L$, and let $z \in R^L$. Then $z$ is a convex combination of the columns of $Z$ iff there exists

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

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

such that

$$\forall r \subseteq E_n, \quad \Upsilon(\bigcap_{j \in r} I_j) = z_r. \quad (10)$$

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 $\Upsilon$) as convenient. For example, we might choose $\mathcal{W} = R^1$.

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

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 $\text{conv}(F)$. In other words, if $x \in R^n$ we may think of $x_j (1 \leq j \leq n)$ as stating the probability of being in the set $I_j$. 

5
This point can be pursued further. Suppose we are trying to construct a function $\Upsilon$ so as to prove that $y$ is measure consistent. In addition to being a probability measure, the only structural condition to be satisfied by $\Upsilon$ is (10). Even though this condition appears simple, a large number of additional conditions are implied by it. In fact, there is a condition to 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 $F$ to a zeta-vector of the subset algebra of $F$, 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 $\{A_j, 1 \leq j \leq k\}$ is a collection of subsets of $S$. Let $\bar{A}_j = S - A_j$ for each $j$.

**Definition 3.1** The subset algebra $\Sigma(A_1, \cdots, A_k)$ 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, in general, $\Sigma(A_1, \cdots, A_k)$ is 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 $\Sigma$ can be written as a (finite) union of atoms. Thus, $\Sigma$ has at most $2^{2^k}$ distinct elements.

**Example 3.3** Consider $\{0, 1\}^n$. For each $1 \leq j \leq n$, let $H_j = \{x \in \{0, 1\}^n : x_j = 1\}$. 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.

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

\[\xi^p = \begin{cases} 1 & \text{if } p \subseteq q \\ 0 & \text{otherwise} \end{cases} \] (11)

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 $\Sigma$ has a member for each subset of $S$, then, as a lattice, $\Sigma$ 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:

$$F = \{x \in \{0, 1\}^n : Ax \geq b\}.$$
We will use the following notation. For $1 \leq j \leq n$ denote $H_j = \{x \in \{0,1\}^n : x_j = 1\}$, and

$$Y_j = \mathcal{F} \cap H_j, \quad N_j = Y_j^c = \mathcal{F} - Y_j.$$  \hfill (12)

In other words, $Y_j$ (resp., $N_j$) is the subset of $\mathcal{F}$ with $x_j = 1$ (resp., $x_j = 0$). Let $\Sigma$ denote the subset algebra $\Sigma(Y_1, \cdots, Y_n)$. Finally, if $z \in R^\Sigma$, its $\alpha$-entry (for $\alpha \in \Sigma$) is indicated by $z[\alpha]$, and for a matrix $B \in R^{\Sigma \times \Sigma}$ its $\alpha,\beta$-entry is indicated by $B[\alpha,\beta]$.

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

$$\big( \bigcap_{j : x_j = 1} Y_j \big) \cap \big( \bigcap_{j : x_j = 0} N_j \big).$$  \hfill (13)

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

$$x \to \xi^{\alpha(x)} (\in \{0,1\}^\Sigma).$$  \hfill (14)

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. Algorithmically, we will restrict the lifting to polynomial-sized subsets of $\Sigma$.

The following properties of the lifting are easy to verify:

(i) $\xi^\emptyset = 0$.

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

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

$$\xi^{\alpha(x)}[Y_j] = x_j, \quad \text{and} \quad \xi^{\alpha(x)}[N_j] = 1 - x_j.$$  \hfill (15)

(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 $\beta^i$ corresponds to a $\mathcal{F}^i \subseteq \mathcal{F}$ such that $\mathcal{F}^i \cap \mathcal{F}^j = \emptyset$, for all distinct $i,j \in I$. Then

$$\sum_{i \in I} \xi^{\alpha(x)}[\beta^i] \leq 1,$$  \hfill (16)

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

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

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:

**Lemma 3.4** Let $z \in \text{conv}(\hat{\mathcal{F}})$. Suppose we have a collection $\{\beta^i \in \Sigma : i \in I\}$, corresponding to pairwise disjoint subsets of $\mathcal{F}$, then

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

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

$$z \geq 0, \quad z[\emptyset] = 0, \quad \text{and} \quad z[\mathcal{F}] = 1.$$  \hfill (17)

**Lemma 3.5** Let $x \in R^n$. Then $x \in \text{conv}(\mathcal{F})$ iff there exists a vector $z \in \text{conv}(\hat{\mathcal{F}})$ such that

$$z[Y_j] = x_j, \quad z[N_j] = 1 - x_j, \quad \forall \ 1 \leq j \leq n.$$  \hfill (18)
The main question we want to address is what conditions the vector $z$ in Lemma 3.5 can be required to satisfy, in addition to (17 - 19).

To this effect, consider an arbitrary vector $f \in R^\Sigma$ of the form
\[ f = \sum_{\alpha \in \Sigma} \lambda_\alpha \xi^\alpha, \]  
(20)
where $\alpha \in R^\Sigma$ is nonnegative. Define the matrix
\[ U_f = \sum_{\alpha \in \Sigma} \lambda_\alpha \xi^\alpha (\xi^\alpha)^T. \]  
(21)

We have the following basic facts:

**Lemma 3.6** Suppose $\lambda$, $f$, and $U_f$ are as in (20 - 21). Let $\beta, \gamma \in \Sigma$. Then
\[ U_{f,\beta,\gamma} = \sum_{\alpha \subseteq \beta \cap \gamma} \lambda_\alpha. \]  
(22)

**Proof.** Let $\alpha \in \Sigma$. Then the sum in (21) contributes either zero or $\lambda_\alpha$ to $U_{f,\beta,\gamma}$, and the latter happens when both $\xi^\alpha[\beta] = 1$ and $\xi^\alpha[\gamma] = 1$. By definition of $\xi^\alpha$, this is true exactly when $\alpha \subseteq \beta$ and $\alpha \subseteq \gamma$, i.e. $\alpha \subseteq \beta \cap \gamma$, as desired.

**Corollary 3.7** For any $\beta, \gamma \in \Sigma$
\[ U_{f,\beta,\gamma} = f[\beta \cap \gamma]. \]  
(23)

This is an analogue of equation (7). As a consequence of these results,

**Lemma 3.8** Suppose $\lambda$, $f$, and $U_f$ are as in (20 - 21). Then

(i) $U_f$ is symmetric,

(ii) The main diagonal of $U_f$ and its $F$-row and -column are all equal to $f$, and

(iii) For every $v \in R^\Sigma$, $v^T U_f v \geq 0$.

**Proof.** As desired.

Note that in Lemma 3.8, (i) and (ii) are really very weak consequences of (23), 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 $\alpha$ and zero otherwise.

**Lemma 3.9** Suppose $\lambda$, $f$, and $U_f$ are as in (20 - 21). Consider a vector $\eta \in R^\Sigma$ such that
\[ \eta^T \xi^\alpha \geq 0, \ \forall \alpha \text{ with } \lambda_\alpha > 0. \]  
(24)

Then

(i) $\eta^T U_f e_{\beta} \geq 0$, for all $\beta \in \Sigma$.

(ii) $\eta^T U_f (e_F - e_{\beta}) \geq 0$, for all $\beta \in \Sigma$.

(iii) For every $v \in R^\Sigma$, $v^T U_f v \geq 0$, 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 $\xi^\alpha$, 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 \( \beta \) (i.e. an element of \( \Sigma \) whose set theoretic value is the complement of \( \beta \)'s). Then the \( F \)-column of \( U_f \) is equal to sum of the \( \beta \)-column and the \( \bar{\beta} \)-column of \( U_f \), i.e.

\[
U_f e_F = U_f (e_\beta + e_{\bar{\beta}}),
\]

(25)

because (similarly to the proof of (i))

\[
\xi^\alpha (\xi^\alpha)^T e_F = \xi^\alpha (\xi^\alpha)^T (e_\beta + e_{\bar{\beta}}), \quad \forall \alpha \in \Sigma
\]

see (16). The result follows from (25) by applying (i) to \( \bar{\beta} \).

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

\[
\kappa = \sum_{\gamma_0 > 0} \lambda_\alpha \xi^\alpha (\xi^\alpha)^T \eta
\]

(27)

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

\[
U^\kappa = \sum_{\gamma_0 > 0} \gamma_\alpha \xi^\alpha (\xi^\alpha)^T
\]

from which (iii) follows.

The properties of \( U_f \) given in Lemma 3.8 and Lemma 3.9 (i) and (ii) parallel the Lovász-Schrijver development of their operators. Property (iii) in Lemma 3.9 is similar to the Sherali-Adams and Lasserre development as shown in [L01].

We summarize what we know so far.

**Lemma 3.10** Let \( \bar{x} \in \text{conv}(F) \). Then there is a vector \( z \in R^\Sigma \) such that:

1. Equations (17) - (19) are satisfied.

2. There is a matrix \( U_z \in R^{\Sigma \times \Sigma} \) satisfying (23), conditions (i) - (iii) of Lemma (3.8), and conditions (i) - (iii) of Lemma (3.9).

In particular, consider (ii) of Lemma (3.9). One way to use this condition is to start with any inequality \( a^T x \geq a_0 \) which is valid for \( F \) and from it obtain (as in Section 1) an inequality \( \hat{a}^T z \geq 0 \) which is valid for \( \hat{F} \). (Formally, set \( \hat{a}[Y_j] = a_j \) for \( 1 \leq j \leq n \), \( \hat{a}[F] = -a_0 \), 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 \( \alpha, \beta \in \Sigma \),

\[
\begin{align*}
\min(z[\alpha \cap \beta]) & \leq \min(z[\alpha], z[\beta]), \\
\min(z[\alpha \cup \beta]) & = z[\alpha] + z[\beta] - z[\alpha \cap \beta]
\end{align*}
\]

(29)

(30)

are valid inequalities. Similar remarks can be made concerning property (iii) in Lemma 3.9.

One difference between our algorithms and the procedures in [LS91], [SA90], [BCC93] and [L01b], is that we generate variables indexed by \( \Sigma \) and not by the lattice of subsets of \( \{1, \ldots, n\} \) – which is isomorphic to a (very small) proper subset of \( \Sigma \). In addition, an iteration of our procedures generates elements of \( \Sigma \) 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 will be used by our algorithms.

**Definition 3.11** Let \( L \) be a set. A **symbol function** is a function \( S : \Sigma \to 2^L \) with the properties:

(S.1) For each \( \alpha, \beta \in \Sigma \) with \( \alpha \neq \beta \), we have \( S(\alpha) \cap S(\beta) = \emptyset \).
(S.2) \( S(\emptyset) \) and \( S(F) \) are nonempty.

For \( \alpha \in \Sigma \), the elements of \( S(\alpha) \) are called the **symbols associated with** \( \alpha \).

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 \( \Sigma \). As a simple example, we might generate a symbol \( \sigma_1 \) for the expression \(((Y_1 \cap N_2) \cup (N_1 \cap N_2)) \cap N_3\) and use it as one of the variable indices, e.g. impose the (valid) constraint

\[
z[\sigma_1] = z[Y_1 \cap N_2 \cap N_3] + z[N_1 \cap N_2 \cap N_3].
\]

At the same time, we might create a symbol \( \sigma_2 \) for the expression \( N_2 \cap N_3 \), and use it in constraints, e.g.

\[
z[\sigma_2] \leq \min\{z[N_2], z[N_3]\}.
\]

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

\[
z[\sigma_1] = z[\sigma_2],
\]

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

\[
(Y_1 \cap N_2) \cup (N_1 \cap N_2) = N_2,
\]

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 (31) then we end up using a relaxation, i.e. a weaker formulation than theoretically possible. In many cases this will be the case with our algorithms. Here, the set \( L \) is arbitrary (for example, in a practical implementation we might have \( L = \mathbb{Z}_+ \)). For consistency, we will use the notation \( x[i] \) to refer to the \( i \)-th entry of vector \( x \in R^{S(\Sigma)} \), for \( i \in S(\Sigma) \), and similarly we will use the notation \( M[i,j] \) to refer to the \( i,j \)-entry in a matrix \( M \in R^{S(\Sigma) \times S(\Sigma)} \).

The following Lemma formalizes the ideas we have just discussed.

**Lemma 3.12** Suppose \( S \) is a symbol function. Let \( M \subseteq \Sigma \) be such that \( \emptyset \in M \), \( F \in M \), and \( S(\alpha) \) is nonempty for each \( \alpha \in M \). Let \( U \in R^{M \times M} \). Consider the matrix

\[
\hat{U} \in R^{S(M) \times S(M)}
\]

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

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

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

\[
\sum_{\alpha \in M} \eta[\alpha] \hat{u}[j_\alpha] \geq 0.
\]

The proof of this Lemma is elementary. The import of this lemma is that, in our lifting, we may 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 (e.g. the diagonal is equal to the \( F \)-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, all of the expressions considered by our algorithms will be 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 \( F \), 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 inequalities with coefficients 0, 1, or 2.

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

$$ Ax \geq c $$

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

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 $X[F]$, and for $1 \leq j \leq n$ the variables $X[Y_j]$ and $X[N_j]$, and impose the constraint:


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

$$ \sum_{j \in A_i} X[Y_j] - X[F] \geq 0. $$

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

$$ C^{i,h} \equiv A_i \cap A_h \neq \emptyset $$

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 $X[\bigcap_{j \in C^{i,h}} N_j]$ and impose the constraints:

$$ X[N_r] - X[\bigcap_{j \in C^{i,h}} N_j] \geq 0, \quad \forall r \in C^{i,h}. $$

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

$$ X[Y_r] - X[Y_r \cap \bigcap_{j \in C^{i,h} - r} N_j] \geq 0. $$

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

$$ \sum_{j \in C^{i,h}} X[Y_j] - 2X[\tau^{i,h}] \geq 0. $$

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

$$ \bigcup_{t \geq 2} \left( \bigcup_{S \subseteq C^{i,h} : |S| = t} \left( \bigcap_{j \in S} Y_j \bigcap_{j \notin S} N_j \right) \right) \in \Sigma. $$

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 $Y$s.
(2.d) Impose the constraint:
\[ X[\bigcap_{j \in C^{i,h}} N_j] + \sum_{r \in C^{i,h}} X[Y_r \cap \bigcap_{j \in C^{i,h} - r} N_j] + X[\tau^{i,h}] - X[\mathcal{F}] = 0. \tag{40} \]

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 (34), (35), (36), (37), (38), (40), (42), (43) impose the constraints
\[ \eta^T U \geq 0. \tag{41} \]

**Step 4.** Impose:

\[ 0 \leq U[\alpha, \beta] \leq U[\mathcal{F}, \beta] \quad \forall \alpha, \beta \in \mathcal{V} \tag{42} \]
\[ X[\mathcal{F}] = 1. \tag{43} \]

End.

**Comment.** Note that the requirements in (3.1) are far weaker than what equation (23) would permit us to impose. For example, consider the expressions \( \beta_1 = Y_1 \cap N_2, \gamma_1 = Y_3 \cap Y_4 \cap N_2, \beta_2 = Y_1 \cap Y_3 \cap Y_4 \), and \( \gamma_2 = N_2 \). 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 \( Y_1 \cap N_2 \cap Y_3 \cap Y_4 \). 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.

### 3.3.1 Analysis of Algorithm C.

We will first show that Algorithm C provides a valid lifting for \( \text{conv}(\mathcal{F}) \). In what follows, let \( \mathcal{M} \) denote the set of distinct members of \( \Sigma \), 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) \( \bigcap_{j \in C^{i,h}} N_j \), all the (distinct) \( Y_r \cap \bigcap_{j \in C^{i,h} - r} N_j \), and all the (distinct) sets of the form given by the left-hand of (39). In the language of Section 3.2.1, there is an implicit symbol function \( \sigma \) such that \( \sigma(\mathcal{M}) = \mathcal{V} \), that the algorithm has constructed.

**Theorem 3.13** Suppose \( \bar{x} \in \text{conv}(\mathcal{F}) \). Then there exists a vector \( \bar{X} \) satisfying constraints (34), (35), (36), (37), (38), (40), (42), (43), and a matrix \( \bar{U} \) satisfying the conditions in Step 3, such that \( \bar{X}[Y_j] = \bar{x}_j \), for all \( 1 \leq j \leq n \).

**Proof.** Let \( z \in R^M \) be the lifting of \( \bar{x} \) that satisfies the conditions in Lemma 3.10, and \( U^z \) denote the corresponding matrix. We will show that \( z \), restricted to \( \mathcal{M} \), satisfies (34), (35), (36), (37), (38), (40), (42), (43) (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 3.4, \( z \) satisfies (43). Next, consider constraint (40) applied to a given pair \( i, h \). Now \( \{0,1\}^n \) can be partitioned into \( 2 + |C^{i,h}| \) sets:

- Those points with all coordinates in \( C^{i,h} \) equal zero. This is the subset of \( \{0,1\}^n \) corresponding to \( \bigcap_{j \in C^{i,h}} N_j \),
• Those points where one coordinate \( r \in C_i^j \) equals one, and all other coordinates equal zero. This is the subset corresponding to \( Y_r \cap \bigcap_{j \in C_i^j \setminus r} N_j \).

• The remaining points, i.e., points where at least two coordinates in \( C_i^j \) equal one. This is the subset corresponding to \( \tau^{i,j} \).

Consequently, by Lemma 3.4, \( z \in j \leq 0 \) Theorem 3.16
Consider an inequality \( \eta^T x \geq a_0 \) which is valid for \( \mathcal{F} \) and such that \( a_j \in \{0,1,2\} \), for \( 0 \leq j \leq n \). Let \((X,U)\) be a vector and matrix satisfying all the constraints imposed by Algorithm C. Then
\[
\sum_{j=1}^{n} a_j X[Y_j] \geq a_0. \tag{45}
\]

Proof. Assume without loss of generality that \( a^T x \geq a_0 \) 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 \( a^T x \geq a_0 \) is one of the constraints \( Ax \geq e \) that define \( \mathcal{F} \), and thus (45) follows by (35).

If \( a_0 = 0 \) then (45) is implied by (42).

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 \( a^T x \geq a_0 \) is of the form \( 2x(T) + x(S) \geq 2 \) for disjoint subsets \( T,S \subseteq \{1,\ldots,n\} \) where \( S \neq \emptyset \).

Recall that we indicate the \( k^{th} \) constraint of \( Ax \geq e \) by \( x(A_k) \geq 1 \). For each element \( j \in S \) it is easy to see that:

\[
\text{for some } 1 \leq k \leq m, \quad A_k \subseteq T \cup S - j. \tag{46}
\]

Consequently, since \( S \neq \emptyset \), we obtain that one of the original constraints \( x(A_k) \geq 1 \) has \( A_k \subseteq T \cup S \).

Clearly \( A_i \cap S \neq \emptyset \) – or else \( a^T x \geq 2 \) is dominated by \( x(A_i) \geq 1 \), a contradiction. Hence, if we pick any \( j \in A_i \cap S \), and apply (46) to \( j \) we find an index \( 1 \leq h \leq m \) with \( A_h \subseteq T \cup S - j \). Necessarily, \( i \neq h \).
Suppose first that $A_i \cap A_h = \emptyset$. In that case, $a^T x \geq 2$ is dominated by the sum of $x(A_i) \geq 1$ and $x(A_h) \geq 1$, and by (35) applied to $i$ and $h$, $X$ satisfies (45).

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 $\bigcap_{j \in C^{i,h}} N_j$, or by $Y_r \cap \bigcap_{j \in C^{i,h} - r} N_j$ (for some $r \in C^{i,h}$), or by $\tau^{i,h}$, then

$$\sum_{j=1}^{n} a_{j} v[Y_j] - 2v[F] \geq 0.$$  \hfill (47)

By Lemma (3.15), this will complete the proof.

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

$$v \bigl[ \bigcap_{j \in C^{i,h}} N_j \bigr] = v[F].$$  \hfill (48)

By constraint (36) imposed by the algorithm,

$$v[N_r] \geq v \bigl[ \bigcap_{j \in C^{i,h}} N_j \bigr], \quad \forall \ r \in C^{i,h}. \hfill (49)$$

Consequently, (34) and (43) applied to $v$, together with (48) and (49) imply:

$$v[Y_r] = 0, \quad \forall \ r \in C^{i,h}. \hfill (50)$$

But by Step 3.2 $v$ satisfies (35) applied to $i$. Together with (50) this implies

$$\sum_{j \in A_i - C^{i,h}} v[Y_j] - v[F] \geq 0.$$  \hfill (51)

Similarly,

$$\sum_{j \in A_h - C^{i,h}} v[Y_j] - v[F] \geq 0,$$  \hfill (52)

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

$$\sum_{j \in A_i \cup A_h} v[Y_j] - 2v[F] \geq 0,$$

which dominates (47).

Consider now the case that $v$ is the column of $U$ corresponding to some $Y_r \cap \bigcap_{j \in C^{i,h} - r} N_j$. Then, by Step 3.1 $v[F] = v[Y_r \cap \bigcap_{j \in C^{i,h} - r} N_j]$, and by constraint (37) imposed by the algorithm we conclude

$$v[Y_r] \geq v[F].$$  \hfill (53)

If $a_r = 2$ we conclude that $v$ satisfies (47). If, on the other hand, $a_r = 1$, then by (46) there is a constraint $x(A_k) \geq 1$ of the original system with $A_k \subseteq T \cup S - r$. By (35), we conclude that $\sum_{j \in A_k} v[Y_j] - v[F] \geq 0$. This, together with (53) implies $v$ satisfies (47).

Finally, consider the case when $v$ is the column of $U$ corresponding to $\tau^{i,h}$. By Step 3.1 $v[\tau^{i,h}] = v[F]$, and therefore by constraint (38) $\sum_{j \in C^{i,h}} v[Y_j] - 2v[F] \geq 0$, which dominates (47).

The theorem is proved. $\blacksquare$

**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

\[
y_h + \sum_{j \in V - v} x_j \geq 1, \quad \forall v \in V, \ h \in J_v.
\]

(54)

\[
(y,x) \in \{0,1\}^W \times \{0,1\}^V.
\]

(55)

[When \( W = \emptyset \), the constraint matrix defined by these inequalities is a full-circulant matrix.] The total number of constraints of type (54) is at most \(|V| \times |W|\). We have:

**Proposition 3.17** 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 \( \text{conv}(C_{W,V}) \).

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 (39). This is necessary in the case of the symbols \( \tau^{i,h} \) because (39) has exponential length – but in a practical implementation we would also want to efficiently record some of the other indices for variables, e.g. \( Y_r \cap \bigcap_{j \in C_{i,h} - r} N_j \), 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 obstructions 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 \( N_j \); \( \bar{Y}_j \) denotes \( N_j \) and \( \bar{N}_j \) denotes \( Y_j \). The symbols \( Y_j \) and \( N_j \) will be called literals.

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

\[
\omega = M_{j_1} \cap M_{j_2} \cap \cdots \cap M_{j_h}
\]

where for each \( 1 \leq i \leq h \) we have \( 1 \leq j_i \leq n \), such that

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

(56)

Put differently, \( \omega \) is an obstructions if any \( x \in \{0,1\}^n \) satisfying

\[
x_{j_i} = \begin{cases} 
1 & \text{if } M_{j_i} = Y_{j_i} \\
0 & \text{if } M_{j_i} = N_{j_i}
\end{cases}
\]

(57)

for every \( 1 \leq i \leq h \) satisfies \( a^T x < a_0 \).

**Definition 4.2** We say that an obstructions \( \omega \) for some inequality is minimal if there is no other obstructions \( \omega' \) for the same inequality such that (as subsets of \( \{0,1\}^n \)) \( \omega \subset \omega' \).
Note: here, “minimal” refers to the set of symbols in the set-theoretic expression $\omega$, which translates into a maximal subset of $\{0,1\}^n$.

**EXAMPLES**

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

(b) Set packing. Given a constraint $\sum_{j \in S} x_j \leq 1$, with $|S| = N$, the are $\binom{N}{2}$ minimal obstructions. Each of them is of the form $Y_{i(1)} \cup Y_{i(2)}$ 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 minimal 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 which is either of the form $\sum_{j \in S} x_j \geq |S| - k$ or $\sum_{j \in S} x_j \leq |S| - k$. In the covering case, the minimal obstructions consist of $k + 1$ terms of type $N_j$, and in the packing case they consist of $k + 1$ terms of type $Y_j$. 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

$$\bigcap_{j \in A^+} N_j \cap \bigcap_{j \in A^-} Y_j$$

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, and thus the minimal ones as a byproduct.

**4.2 The algorithm**

Now we return to our general lifting procedure. The algorithm we will present is a generalization of the procedure in Section 3.3. That algorithm created variables indexed by members of $\Sigma$, 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 at least $Y_j$ terms. The particular expressions actually generated by the algorithm were not the complete set of all such expressions – rather, they reflected the structure of the problem.

The algorithm presented below generalizes this type of construction. In particular, we will create expressions which are intersections of $N_j$ and $Y_j$ terms according to some specific criteria. This necessitates some further definitions.

**Definition 4.4** Suppose we are given an expression $\gamma = \bigcap_{i=1}^{h} M_{j_i}$, 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 $\gamma$ of order $t$ is an expression of the form $\bigcap_{i=1}^{h} M_{j_i}'$, such that for exactly $t$ indices $j_i$ we have $M_{j_i}' = M_{j_i}$, and for the remaining $h - t$ indices $j_i$ we have $M_{j_i}' = M_{j_i}$.
(2) $0 \leq r \leq h$, $N(\gamma, r)$ is the set of negations of $\gamma$ of order $r$ (note: if $r = 0$ then this is just the set $\{\gamma\}$).

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

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

**Example 4.5** Consider

$$\gamma_1 = N_1 \cap Y_2 \cap Y_3$$
$$\gamma_2 = Y_4 \cap Y_5 \cap Y_6 \cap N_7$$
$$\gamma_3 = N_8 \cap Y_9.$$

Suppose

$$\gamma_1' = Y_1 \cap Y_2 \cap Y_3$$
$$\gamma_2' = N_4 \cap N_5 \cap Y_6 \cap N_7$$
$$\gamma_3' = N_8 \cap Y_9.$$

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

**Definition 4.6**

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

(b) A **tier** is an expression of the form $\gamma = \omega_1 \cap \omega_2 \cap \cdots \cap \omega_t$, and a negation $\gamma' = \omega_1' \cap \omega_2' \cap \cdots \cap \omega_t'$, and a a wall.

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

$$\gamma_1 = N_1 \cap Y_2 \cap Y_3$$
$$\gamma_2 = N_1 \cap Y_2 \cap Y_4 \cap N_5$$
$$\gamma_3 = N_1 \cap Y_4.$$

Then the wall derived from $E$ is $N_1 \cap Y_2 \cap Y_4$.

**Definition 4.8**

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

(ii) Given an expression of the form $\gamma = \omega^{>r} \cap \beta$, where $\beta = \bigcap_{i=1}^{h} M_i$, $\omega$ is a wall and $r > 0$, the simplification of $\gamma$ is the expression $\omega^{>r} \cap \beta'$, where $\beta'$ is the simplification of $\beta$.

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

(iv) Suppose we have an expression of the form $\beta = \mathcal{F} \cap \gamma$, where $\gamma$ is of the form considered in (i) or (ii). Then the simplification of $\beta$ equals the simplification of $\gamma$. 
Example 4.9 Suppose

\[ \gamma = N_1 \bigcap Y_2 \bigcap N_1 \bigcap Y_3. \]

Then the simplification of \( \gamma \) is

\[ Y_2 \bigcap Y_3 \bigcap N_1. \]

Comment. All expressions that the algorithm given below will generate will be of one of the following types

1. a literal,
2. \( F \),
3. a wall,
4. a tier,
5. a tier with some walls negated,
6. an expression of the form \( \beta \cap \gamma \) where \( \beta \) and \( \gamma \) are of types 1-5.

Any expression of this kind is of the form

\[ B_1 \bigcap B_2 \bigcap \cdots \bigcap B_r, \]

where each \( B_i \) is a literal, \( F \), or of the form \( \omega_i^+ \). Recall the discussion at the end of Section 3.2.1. We will use the assumption that any permutation of the \( B_i \)s yields an equivalent expression, and also apply the simplification operator, to enforce equivalences between the expressions generated by our algorithm.

Definition 4.10 Suppose we are given an expression

\[ \alpha = M_{j_1} \bigcap M_{j_2} \bigcap \cdots \bigcap M_{j_r}, \]

and an expression

\[ \beta = M_{i_1} \bigcap M_{i_2} \bigcap \cdots \bigcap M_{i_q}. \]

We say that \( \beta \) is a superstring of \( \alpha \) if every literal \( M_h \) appearing in the simplification of \( \alpha \) also appears in the simplification of \( \beta \).

Now we can define our basic general algorithm. 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 \( a_i x \geq b_i \).

Level-\( k \) \( \Sigma \)-Algorithm \((k \geq 2)\).

1. Generate the symbol \( F \), and for \( 1 \leq j \leq n \), generate the symbols \( Y_j \) and \( N_j \), and impose the constraint:

\[ X[Y_j] + X[N_j] - X[F] = 0. \quad (58) \]

For \( 1 \leq i \leq m \) impose the constraint:

\[ \sum_{j=1}^{n} a_{ij} X[Y_j] - b_i X[F] \geq 0. \quad (59) \]

2. For \( 1 \leq i \leq m \), enumerate all the \( k \)-small minimal obstructions to each of the constraints \( a_i x \geq b_i \). For each enumerated obstruction \( \gamma \), impose the constraint

\[ \sum_{M_j \in \gamma} X[M_j] - X[F] \geq 0. \quad (60) \]

where the notation \( "M_j \in \gamma" \) means that the sum is over all those literals \( M_j \) occurring in \( \omega_i \).

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. Every tier of the form $\theta = \bigcap_{t=1}^{\mathcal{H}} \omega_t$, where $0 \leq \mathcal{H} < k$, and $\omega_t$ is a wall derived from at most $k$ enumerated obstructions for $1 \leq t \leq \mathcal{H}$, $\theta$ is processed to generate additional symbols and constraints, as follows.

For each $\mathcal{H}$-tuple of integers $o_1, o_2, \ldots, o_{\mathcal{H}}$ such that $0 \leq o_t \leq \min\{|\omega_t|, k\}$ for $1 \leq t \leq \mathcal{H}$, $\sum_{t=1}^{\mathcal{H}} o_t < k$ and $o_t > 0$ for $1 \leq t < \mathcal{H}$, and each $\mathcal{H}$-tuple $(\omega'_1, \omega'_2, \ldots, \omega'_\mathcal{H})$, where $\omega'_t$ is a negation of $\omega_t$ of order $o_t$ for $1 \leq t \leq \mathcal{H}$,

(i) We create a symbol for

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

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

$$X[M_j] - X[\theta^\#] \geq 0. \quad (61)$$

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

$$X[\theta^\#] = 0. \quad (62)$$

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

$$\theta^> = \omega_{\mathcal{H}}^{o_{\mathcal{H}} \cap} \bigcap_{t=1}^{\mathcal{H}-1} \omega'_t,$$

where as before $\omega_{\mathcal{H}}^{o_{\mathcal{H}}}$ is the negation of $\omega_t$ of order greater than $o_t$. 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 $\omega'_t$, for $1 \leq t \leq \mathcal{H} - 1$, impose the constraint

$$X[M_j] - X[\theta^>] \geq 0. \quad (63)$$

Moreover, we impose the constraint

$$\sum_{M_j \in \omega_{\mathcal{H}}} X[M_j] - (1 + o_{\mathcal{H}}) X[\theta^>] \geq 0. \quad (64)$$

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

$$X[\theta^>] = 0. \quad (65)$$

5. For each wall $\omega$ derived from at most $k$ enumerated obstructions, and for each expression $\theta^\# = \bigcap_{t=1}^{\mathcal{H}} \omega'_t$ enumerated in Step 4, such $\omega'_t$ is a negation of order $0 < o_t$ of some wall $\omega_t$, for $1 \leq t \leq \mathcal{H}$, we impose the following constraints:

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

$$X[\theta^\#] - \left( \sum_{r=0}^{|\omega|} \left( \sum_{\beta \in \mathcal{N}(|\omega|, r)} X[\beta \cap \theta^\#] \right) \right) = 0. \quad (66)$$

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

$$X[\theta^\#] - X[\omega^{>R} \cap \theta^\#] - \left( \sum_{r=0}^{R} \left( \sum_{\beta \in \mathcal{N}(|\omega|, r)} X[\beta \cap \theta^\#] \right) \right) = 0, \quad (67)$$

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

19
Similarly, for each wall $\omega$ derived from at most $k$ enumerated obstructions, if $|\omega| < k$ then we impose

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

(68)

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

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

(69)

6. 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:

(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 (59-69) (when they apply) we impose the constraints

$$\eta^T U \geq 0.$$

(70)

(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 we require that $U[\beta_1, \gamma_1] = U[\beta_2, \gamma_2]$.

(6.4) Impose

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

(71)

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

(72)

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

8. (Optional) For each constraint $\eta^T X \geq 0$ of the form (59), (60), or (64), let $f^\eta \doteq U \eta$. Form the matrix $W_{\eta}$ with rows and columns indexed by $\mathcal{V}$, such that for each pair $\beta, \gamma$ of symbols in $\mathcal{V}$, the $\beta, \gamma$-entry of $W_{\eta}$ equals $f^\eta[\beta \cap \gamma]$. Impose $W_{\eta} \succeq 0$.

9. End.

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 (70) we should use “$=$”.

2. Note that the simplification step in (6.3) implies that e.g. if $\beta_1$ and $\beta_2$ are elements of $\mathcal{V}$ with the same simplification, then the columns (and rows) of $U$ corresponding to $\beta_1$ and $\beta_2$ 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 $\beta \cap \theta^#$ are obtained by negating walls in a tier enumerated in step 4. Hence the variables $X[\beta \cap \theta^#]$ will be created by the algorithm, and similarly with all other variables in (66-69).

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 $\theta^#$ variables (among others). Also, when there is more than one minimal obstruction, (67) can be made stronger and (60) is not needed.

5. Step 7 is $M_\omega$-like, using of course a very different set of variables. Similarly, Step 8 is Lasserre-like (refer to Lemma 3.9 (iii)).
4.3 Properties of the level-$k$ Σ-algorithm

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

Lemma 4.11 For $k$ fixed, the level-$k$ Σ-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 minimal obstructions. Thus at most $O(m^kn^k)$ walls, and at most $O(m^k - kn^k)$ tiers $\theta$ are enumerated. As a result, at most $O(2^k k! m^k - kn^k)$ symbols $\theta^#$ and $\theta^>$ will be created.

Comment. By using a more streamlined version of the algorithm, the number of symbols becomes $O((mn)^{O(k)})$.

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

Lemma 4.12 Let $\hat{x} \in \mathcal{F}$. Then $\hat{x}$ can be lifted to a vector $\hat{X} \in \{0,1\}^V$ and matrix $\hat{U} \in \{0,1\}^{V \times V}$ that satisfy the constraints imposed by the algorithm.

Proof. We will show that $\hat{x}$ can be lifted to a vector $\hat{x} \in \mathbb{R}^M$ 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.

$$\alpha(\hat{x}) = \bigcap_{j : \hat{x}_j = 1} Y_j \cap \bigcap_{j : \hat{x}_j = 0} N_j.$$  

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 $\beta$ contains the point $\hat{x}$. Further, let $\hat{U}$ be defined by $\hat{U}[\beta, \gamma] = \hat{X}[\beta \cap \gamma]$. We claim that $\hat{X}$ and $\hat{U}$ are as desired.

Clearly (58), (59) and (60) are satisfied. Consider now expressions $\theta$, $\theta^#$ as in Step 4 of the algorithm. Since $\theta^#$ is of the form $\bigcap_{j \in J} M_j$ for some set $J$, clearly (61) is satisfied (e.g. the subset of $\mathcal{F}$ defined by $\theta^#$ is contained in the subset defined by $M_j$ for $j \in J$). Similarly, (63) is satisfied. Finally, in constructing $\theta^>$ at least $1 + \mathcal{O}_\mathcal{H}$ symbols are negated in $\omega_\mathcal{H}$, and consequently (64) holds as well.

Consider now (66 - 69). Suppose that (66) applies. If $\hat{x}[M_j] = 0$, for at least one literal $M_j$ which appears in at least one of the expressions $\omega^t$, for $1 \leq t \leq \mathcal{H}$, then by construction both terms in (66) are equal to zero and (66) holds. Suppose now that $\hat{X}[\theta^#] = 1$. Let $0 \leq K$ be the number of indices $j$, such that either $Y_j$ is a symbol in $\omega$ and $\hat{x}_j = 0$, or $N_j$ is a symbol in $\omega$ and $\hat{x}_j = 1$. By construction in step 3 of the algorithm, $K$ is well defined and (66) follows. The proof that (67 - 69) 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 3.9.

As a corollary of this Lemma, we now have:

Theorem 4.13 Suppose $\hat{x} \in \text{conv}(\mathcal{F})$. Then there exists a vector $\hat{X} \in \mathbb{R}^V$ and a matrix $\hat{U} \in \mathbb{R}^{V \times V}$ satisfying the conditions imposed by the level-$k$ Σ-algorithm, and such that $\hat{X}[Y_j] = \hat{x}_j$, for all $1 \leq j \leq n$.

Proof. This follows from Lemmas 3.10 and 4.12.

4.4 Packing problems

Stable set problems are prominently studied in [LS91], where it is shown that the $N_+$-rank of clique inequalities (among others) is 1, whereas their $N$-rank is much higher. Thus, positive-semidefiniteness can help in a concrete way. In this section we present some results concerning the application of the level-$k$ Σ-algorithm, with small $k$, to set packing problems.
Consider the following packing system. We are given a family of disjoint sets $J_i$, $1 \leq i \leq r$, and an integer $k > 1$. We consider the system of constraints

$$ \sum_{i \in S} \sum_{j \in A_i} x_j \leq \sum_{i \in S} |J_i| - 1, \quad \forall S \subseteq \{1, 2, \cdots, r\} \text{ with } |S| = k. \tag{73} $$

Thus, if $k = 2$ and $|J_i| = 1$ for all $i$, we obtain a stable set system for an $r$-clique. Clearly,

$$ \sum_{i=1}^{r} \sum_{j \in J_i} x_j \leq \sum_{i \in S} |J_i| - r + k - 1 \tag{74} $$

is a valid inequality which generalizes the standard clique inequality.

**Lemma 4.14** Suppose we have a packing problem containing a subsystem of the form (73). The level-$k$ $\Sigma$-algorithm using Step 7 implies (74). On the other hand, suppose $k = 2$ and $|J_1| = |J_2| = \cdots = |J_r| = m$, say. Then the $N_+$-rank of (74) is at least $\min \{m, r\} - 1$. ■

Comment: the second part of the Lemma can be made considerably stronger. Instead of $N_+$, the same result holds using the much stronger operator described in the appendix.

(Further material will be added to this section).

### 4.5 Set covering

In this section we consider a set-covering problem with feasible region $F = \{ x \in \{0,1\}^n : Ax \geq e \}$, where the $0-1$ matrix $A$ is $m \times n$. We will prove that the $k$-level algorithm implies all inequalities of pitch $\leq k$ which are valid for $F$. For completeness, we redefine pitch here.

**Definition 4.15** 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 $\pi = \pi(a_0)$ such that $\sum_{j=1}^\pi a_j \geq a_0$.

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.

For a vector $v$, we denote its support by $\text{supp}(v)$. The following results are needed for later proofs.

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

**Proposition 4.16** Suppose $a^T x \geq a_0$ is a valid inequality for $F$ with nonnegative coefficients. Consider the set-covering problem with feasible region $G = \{ x \in \{0,1\}^{\text{supp}(\alpha)} : \hat{A}x \geq e \}$, where $\hat{A}$ is the submatrix of $A$ obtained by

(i) Using the columns in $\text{supp}(\alpha)$, and

(ii) Using those rows $i$ such that $A_i \subseteq \text{supp}(\alpha)$.

Let $\hat{\alpha}$ be the restriction of $\alpha$ to $\text{supp}(\alpha)$. Then $\hat{\alpha}^T x \geq \alpha_0$ is valid for $G$.

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

**Lemma 4.17** Suppose $\alpha^T x \geq \alpha_0$ is a valid inequality for $F$ and $\alpha_0 > 0$. Then there is a subset $C = C(\alpha, \alpha_0)$ of the rows of $A$ with $|C| \leq \pi(\alpha, \alpha_0)$, such that

(i) $A_i \subseteq \text{supp}(\alpha)$ $\forall$ $i \in C$ \hspace{1cm} (75)

(ii) $\Delta_i \equiv A_i - \bigcup_{r \in C^{-i}} A_r \neq \emptyset$, $\forall$ $i \in C$, and \hspace{1cm} (76)

(iii) $\sum_{i \in C} \min \{ \alpha_j : j \in \Delta_i \} \geq \alpha_0$. \hspace{1cm} (77)
Proof. The proof of the Lemma will be by induction on $\pi = \pi(\alpha, \alpha_0)$. Since $\alpha_0 > 0$ we must have $\sum_{j \in \text{suppt}(\alpha)} x_j \geq 1$ is valid for $\mathcal{F}$, and thus the case $\pi = 1$ is proved.

Assume now that $\pi > 1$. By Proposition 4.16 we can assume that $A_i \subseteq \text{suppt}(\alpha)$ for every row $i$. Suppose first that some row $i$ of $A$ satisfies $A_i = \text{suppt}(\alpha)$. Since by assumption no row of $A$ contains another, we have a set-covering problem with one constraint, namely $\sum_{j \in \text{suppt}(\alpha)} x_j \geq 1$. 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 $\text{suppt}(\alpha)$. 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 $C = \{i(1)\}$).

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

1. $A$ has column set $K$.
2. for any row $h$ of $A$, with $j(1) \notin A_h$, $\tilde{A}$ will have a row with $A_h = A_h \cap K$ if this set is nonempty (and otherwise $A$ does not have a row corresponding to $h$).

We claim that the inequality

$$\sum_{j \in K} \alpha_{j} x_j \geq \alpha_0 - \alpha_{j(1)} \tag{78}$$

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 $\bar{x} \in \{0,1\}^n$ by setting $\bar{x}_j = \bar{x}_j$ if $j \in K$, and $\bar{x}_j = 1$ if either $j \notin \text{suppt}(\alpha)$ or $j = j(1)$. Clearly $\alpha^T \bar{x} < \alpha_0$, but $\bar{x} \in \mathcal{F}$, a contradiction. The latter fact follows because no row of $A$ is contained in another (and, in particular, not contained in $A_i(1)$).

Since $\alpha_{j(1)} > 0$ it follows that the pitch of (78) is less than $\pi$. Since $\alpha_{j(1)} < \alpha_0$, the result now follows by induction as $j(1)$ is not contained in any set $A_i$.

Corollary 4.18 Suppose the inequality $a^T x \geq \alpha_0$ is valid for $\mathcal{F}$ and has pitch $\leq 1$. Then there is a row $i$ of $A$ such that $a_i x \geq 1$ dominates $a^T x \geq \alpha_0$. ■

In the rest of this section we will consider a fixed a fixed vector $\tilde{X} \in R^V$ and matrix $\tilde{U} \in R^{V \times V}$ that satisfy the constraints imposed by the level-$k \Sigma$-algorithm. Our goal is to show that for every inequality $a^T x \geq \alpha_0$ with $\pi(a, \alpha_0) \leq k$ which is valid for $\mathcal{F}$, $\tilde{X}$ satisfies the (homogenized) inequality

$$\sum_{j} a_j X[Y_j] - \alpha_0 X[\mathcal{F}] \geq 0, \tag{79}$$

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

4.5.1 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 (66-69).

Suppose that we could express $\tilde{X}$ as a sum of other columns of $\tilde{U}$, each of which satisfies (79). Since (79) is homogeneous, it follows that $\tilde{X}$ satisfies (79) 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 inequality (68) arising from a particular choice of $\omega$. By step 6, this inequality is satisfied by every column of $\tilde{U}$. Since $\tilde{U}$ is symmetric, this instance of (68) is therefore satisfied by every row of $\tilde{U}$, and as a result $\tilde{X}$ is a sum of columns of $\tilde{U}$, each of which corresponds to some negation $\beta$ of $\omega$. Formally,

$$\tilde{X} = \sum_{r=0}^{\left|\omega\right|} \left( \sum_{\beta \in N(\omega, r)} \text{col}[\beta] \right), \tag{80}$$

23
where by col[β] we denote the column of ˜U indexed by β. Now (80) (or a similar decomposition if (69) applies instead of (68)) 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 (79).

In our proof, the particular ω that will give the desired decomposition will be supplied by Lemma 4.17. In order to show that each resulting β-column of ˜U satisfies (79) we will use a special proof for the case r = 0 (and for the column corresponding to ω^k−1 when (69) applies). For the cases r > 0 we will use induction.

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

$$\sum_j a_j ˜u[Y_j] - a_0 ˜u[F] \geq 0.$$  \hspace{1cm} (81)

Since we are dealing with a set-covering problem, each of the r negations yields a Y_j factor appearing in β. As we will argue, for each such j we have ˜u[Y_j] = ˜u[F]. Thus, instead of having to show (81) we will have to show

$$\sum_{j: Y_j \notin \beta} a_j ˜u[Y_j] - (a_0 - \sum_{j: Y_j \in \beta} a_j) ˜u[F] \geq 0$$

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

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

is valid for F and has pitch strictly less than the pitch of a^T x ≥ a_0. Hence, we can apply induction, after another use of Lemma 4.17. Subsequent inductive steps will use (66) and (67) instead of (68) and (69). 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 ˜U, as implied by Corollary 4.18.

4.5.2 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 minimal obstructions will be of the form ∩_{j∈A_h} N_j for some h.

The following observation will be of use later.

Proposition 4.19 Let θ = ∩_{i=1}^n, ω_i be a tier enumerated in step 4 of the algorithm. Consider one of the expressions θ# obtained from θ, and et ˜u be the column of ˜U indexed by some expression θ#. Suppose 1 ≤ j ≤ n is such that M_j appears in θ#. Then

$$\tilde{u}[Y_j] = \tilde{u}[F].$$  \hspace{1cm} (82)

Similarly, if θ' = ∩_{i=1}^{n-1} ω_i and ϕ is obtained from θ, and Y_j appears in ∩_{i=1}^{n-1} ω_i, then (82) holds, as well.

Proof. This follows from (61), (63), step (6.1) and (71).

First, we will prove the base for the induction.

Lemma 4.20 Suppose \sum_{j=1}^n a_j x_j ≥ a_0 is an inequality valid for F of pitch ≤ 1. Let ˜u be any column of ˜U. Then \sum_{j=1}^n a_j ˜u[Y_j] - a_0 ˜u[F] ≥ 0.

Proof. This follows from Corollary 4.18 and the fact that due to Step 6 of the algorithm, every column of ˜U satisfies each of the set covering constraints Ax ≥ c, homogenized.

In what follows, we use the following notation. Consider expressions θ = ∩_{i=1}^n ω_i and θ# = ∩_{i=1}^n ω_i generated in step 4 of the algorithm, where each ω_i is a negation of order o_i of ω_i. We will say that θ# is simple if 0 < o_i for all i. Since by construction 0 < o_i for t ≤ H−1, this simply states that 0 < o_H. We let

$$Y(\theta#) = \{1 ≤ j ≤ n : Y_j appears in \theta# \}.$$
Every \( j \in Y(\theta^\#) \) appears as \( N_j \) in some \( \omega_t \) and is negated in constructing \( \theta^\# \), i.e. it will appear as \( Y_j \) in \( \omega'_t \). Finally, we write

\[
\rho(\theta^\#) = k - \sum_{t=1}^N \alpha_t.
\]  

(83)

Note: strictly speaking, we should write \( \rho(\theta^\#, \theta) \), but as \( \theta \) will always be clear from the context we will omit it. Also note that \( 0 < \rho(\theta^\#) \) when \( \theta^\# \) is simple. We will also write

\[
Y(F) = \emptyset,
\]  

and

\[
\rho(F) = k.
\]

Now we have the main result, which implies Theorem 1.2.

**Theorem 4.21** Consider an expression \( \gamma \) so that either \( \gamma = F \) or \( \gamma \) is some \( \theta^\# \) enumerated in Step 4 and is simple. Let

\[
\sum_{j=1}^n \alpha_j x_j \geq \alpha_0
\]  

(84)

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

\[
\sum_{j=1}^n \alpha_j \tilde{u}[Y_j] - \alpha_0 \tilde{u}[F] \geq 0.
\]  

(85)

**Proof.** The proof will be by induction on \( \rho(\gamma) \), which, as just discussed, is positive. Lemma 4.20 therefore handles the base of this induction.

By Proposition 4.19, if \( j \in Y(\gamma) \) then \( \tilde{u}[Y_j] = \tilde{u}[F] \), and as a result

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

Consequently, if we could also prove that the inequality

\[
\sum_{j \notin Y(\gamma)} \alpha_j X[Y_j] - \left( \alpha_0 - \sum_{j \in Y(\gamma)} \alpha_j \right) X[F] \geq 0
\]  

(86)

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

Rewrite (86) as

\[
\sum_{j=1}^n \tilde{\alpha}_j X[Y_j] - \tilde{\alpha}_0 X[F] \geq 0,
\]  

(87)

where, for \( 1 \leq j \leq n \)

\[
\tilde{\alpha}_j = \begin{cases} 
\alpha_j, & \text{if } j \notin Y(\gamma) \\
0, & \text{otherwise}
\end{cases}
\]  

(88)

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

\[
\tilde{\alpha}^T X \geq \tilde{\alpha}_0
\]  

(89)

is clearly valid for \( F \) (since (84) is), and by definition of pitch, \( \pi \triangleq \pi(\tilde{\alpha}, \tilde{\alpha}_0) \leq \pi(\alpha, \alpha_0) \).

Suppose we apply Lemma 4.17 to (89). Let \( C \) denote the resulting subset of rows of \( A \). Thus we obtain a set of obstructions, one from each row of \( C \); since \( |C| \leq \pi \leq \rho(\gamma) \leq k \), the wall \( \omega \) derived from this set of obstructions will be enumerated in Step 5.
Hence, according to (66 - 69) the algorithm will impose either

\[ X[\gamma] = \sum_{r=0}^{\lfloor \omega \rfloor} \left( \sum_{\beta \in \mathcal{N}(\omega,r)} X[\beta \cap \gamma] \right) \]

if \(|\omega| < \rho(\gamma)\) (c.f. (83)), or

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

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

Since \(\bar{U}\) is required to be symmetric, and since whichever of (90) or (91) applies is enforced on all columns of \(\bar{U}\), we have that all rows of \(\bar{U}\) satisfy this constraint. Hence, denoting by \(\text{col}[\sigma]\) the column of \(\bar{U}\) corresponding to a symbol \(\sigma\), one of

\[ \text{col}[\gamma] = \sum_{r=0}^{\lfloor \omega \rfloor} \left( \sum_{\beta \in \mathcal{N}(\omega,r)} \text{col}[\beta \cap \gamma] \right) \]

or

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

will apply.

We will next show, using induction, that all columns of the form \(\text{col}[\beta \cap \gamma]\) arising from values \(r > 0\) satisfy (87). Using a special proof, we will show the same fact for the case \(r = 0\). Finally, assuming (93) applies, we will show the same for \(\text{col}[\omega^R \cap \gamma]\). 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, \(\beta\) is of the form

\[ \bigcap_{j \in S} Y_j \cap \bigcap_{j \in T} N_j \]

where \(|S| = r\), \(\omega = \bigcap_{j \in S \cup T} N_j\), and \(S\) and \(T\) are disjoint. Let \(\hat{\alpha} = \text{col}[\beta \cap \gamma]\).

By Proposition 4.19, \(\hat{\alpha}[Y_j] = \hat{\alpha}[\mathcal{F}]\) for each \(j \in S\). Thus, it suffices to prove that \(\hat{\alpha}\) satisfies the inequality

\[ \sum_{j \notin S} \alpha_j X[Y_j] - \left( \sum_{j \in S} \alpha_j - \sum_{j \in S} \bar{\alpha}_j \right) X[\mathcal{F}] \geq 0 \]

in order to prove that \(\hat{\alpha}\) satisfies (87).

But notice that by construction in Lemma 4.17, all rows in the set \(\mathcal{C}\) have support contained in \(\text{supp}(\bar{\alpha})\). Hence \(\bar{\alpha}_j > 0\) \(\forall j \in S\), and as a result the pitch of (94) is at most the pitch of (89) minus \(r\), i.e. \(\leq \rho(\gamma) - r\).

Further, in the case \(\gamma = \theta^R\), we have that \(\beta \cap \gamma\) is obtained from \(\omega \cap \theta\) by negating exactly \(r\) additional literals than were negated in deriving \(\theta^R\) from \(\theta\), i.e. \(\rho(\beta \cap \gamma) = \rho(\gamma) - r\). In the case \(\gamma = \mathcal{F}\) this is also clear. Thus, in either case, \(\rho(\beta \cap \gamma)\) is at least as large as the pitch of (94), and by induction we have that \(\hat{\alpha}\) satisfies (94), as desired.

**Case** \(r = 0\). Let \(\hat{\alpha} = \text{col}[\beta \cap \gamma]\). 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 Lemma 4.17 to \(\bar{\alpha}^T x \geq \bar{\alpha}_0\):

\[ (i) \quad A_i \subseteq \text{supp}(\bar{\alpha}) \quad \forall \ i \in \mathcal{C} \]

\[ (ii) \quad \Delta_i \triangleq A_i - \bigcup_{r \in \mathcal{C} - \mathcal{i}} A_r \neq \emptyset, \quad \forall \ i \in \mathcal{C}, \quad \text{and} \]

\[ (iii) \quad \sum_{i \in \mathcal{C}} \min \{ \bar{\alpha}_j : j \in \Delta_i \} \geq \bar{\alpha}_0. \]
Further each \( i \in C \) gives rise to one obstruction and \( \omega \) is the wall derived from these obstructions. Thus, \( \omega = \bigcap_{j \in J} N_j \), where \( J \subseteq \{1, 2, \ldots, n\} \) is the set of all \( j \) appearing in at least two \( A_i, i \in C \).

Consider any row \( i \in C \). By step (6.2) of the algorithm,

\[
\sum_{j=1}^{n} a_{ij} \hat{u}[Y_j] - \hat{u}[F] \geq 0
\]

But since \( r = 0 \), by Proposition 4.19

\[
\hat{u}[Y_j] = 0, \quad \forall \ j \in J
\]

Combining these two inequalities, we obtain

\[
\sum_{j \in \Delta_i} a_{ij} \hat{u}[Y_j] - \hat{u}[F] \geq 0
\]

and consequently

\[
\sum_{j \in \Delta_i} \bar{\alpha}_j \hat{u}[Y_j] \geq \left( \min_{j \in \Delta_i} \{ \bar{\alpha}_j : j \in \Delta_i \} \right) \hat{u}[F]. \tag{95}
\]

By construction, the sets \( \Delta_i \) are pairwise disjoint. So if we sum (95) over all \( i \in C \) we obtain:

\[
\sum_{j=1}^{n} \bar{\alpha}_j \hat{u}[Y_j] \geq \left( \min_{i \in C} \{ \bar{\alpha}_j : j \in \Delta_i \} \right) \hat{u}[F] \geq \bar{\alpha}_0 \hat{u}[F], \tag{96}
\]

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

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

\[
\sum_{j} \bar{\alpha}_j \hat{u}[Y_j] = \sum_{j \in \text{supp}(\bar{\alpha})} \bar{\alpha}_j \hat{u}[Y_j] \geq \sum_{j \in J} \bar{\alpha}_j \hat{u}[Y_j]. \tag{97}
\]

Further, in step (6.2), the algorithm imposes the constraint (64), applied to \( \omega^{>R} \cap \gamma \), on the column \( \hat{u} \). This amounts to imposing:

\[
\sum_{j \in J} \hat{u}[Y_j] - (1 + R)\hat{u}[\omega^{>R} \cap \gamma] \geq 0. \tag{98}
\]

Now \( 1 + R = \rho(\gamma) \), by definition of \( R \). Also, by step (6.1), \( \hat{u}[\omega^{>R} \cap \gamma] = \hat{u}[F] \). So (98) is simply:

\[
\sum_{j \in J} \hat{u}[Y_j] - \rho(\gamma)\hat{u}[F] \geq 0. \tag{99}
\]

Finally, by inequality (72) imposed by the algorithm in step (6.4)

\[
\hat{u}[F] \geq \hat{u}[Y_j], \quad \forall j \in J. \tag{100}
\]

Thus, combining (97), (99) and (100) we obtain that

\[
\sum_j \bar{\alpha}_j \hat{u}[Y_j]
\]

is at least \( \hat{u}[F] \), 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

\[
\sum_j \bar{\alpha}_j \hat{u}[Y_j] \geq \bar{\alpha}_0 \hat{u}[F],
\]

as desired. This completes the proof. \( \blacksquare \)

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

(Further material will be added to this section).

4.7 Logic problems

There is a clear and well-known connection between set theory and logic. Further, here is a rich literature on the connection between logic and optimization, and on incorporating techniques of logic to integer programming. See e.g., [H00], also see [BH01] (more references to be added). This topic appears to be underlaid by Balas’ work on disjunctive programming.

Our work clearly touches on this connection. At the same time, 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.

The Σ-algorithm can naturally be restated so that it deals with problems directly presented by a list of obstructions.

(Further material will be added to this section).

4.8 Further remarks

Consider the feasible system \( F \) introduced in [CCH89]:

\[
\sum_{j \in S} x_j + \sum_{j \notin S} (1 - x_j) \geq \frac{1}{2}, \text{ } \forall S \subseteq \{0,1,\ldots,n\}
\]

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

Here we will show that running the level-3 Σ-algorithm proves the same result. We will denote by \( C(S) \) the inequality (101) corresponding to the set \( S \). Note that corresponding to \( C(S) \) there is a unique minimal obstruction, namely

\[
\bigcap_{j \in S} N_j \cap \bigcap_{j \notin S} Y_j
\]

and conversely, any expression of this form is the minimal 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. Further, notice that each minimal obstruction is 0-small.

Lemma 4.22 Suppose we run the level-3 Σ-algorithm. Consider an expression of the form

\[
\gamma = \bigcap_{j \in A} N_j \cap \bigcap_{j \in B} Y_j
\]

where \( A \cap B = \emptyset \) and \( |A \cup B| > 0 \). Then there is a tier \( \theta \) generated by the algorithm, such that \( \gamma \) is one of the enumerated \( \theta^\# \), with one negation.

Proof. Suppose first that \( |A \cup B| = n \). Either \( A \neq \emptyset \) or \( B \neq \emptyset \); assuming e.g. the first case pick any \( j_0 \in A \), and set

\[
\theta = \bigcap_{j \in A \setminus j_0} N_j \cap \bigcap_{j \in B \cup j_0} Y_j.
\]

Then \( \theta \) is an obstruction, and thus a wall enumerated by the algorithm, and \( \gamma \) will be enumerated as one of the expressions \( \theta^\# \) (\( \gamma \) is a negation of order 1 of \( \theta \)).

Suppose next that \( |A \cup B| < n \). Assume again without loss of generality that \( A \neq \emptyset \) and pick any \( j_0 \in A \). Write \( \tilde{G} = \{1,2,\ldots,n\} - (A \cup B) \). Let \( \theta \) be the wall derived from two obstructions: the obstruction to constraint \( C((A \setminus j_0) \cup \tilde{G}) \) and the obstruction to \( C(A \setminus j_0) \). Then \( \theta \) is as desired (we negate \( j_0 \) to obtain \( \gamma \) from \( \theta \)). ■
Theorem 4.23 The linear system produced by the level-3 $\Sigma$-algorithm is infeasible.

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

$$\gamma = \bigcap_{j \in A} N_j \cap \bigcap_{j \in B} Y_j$$

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

$$\bar{X}[\gamma] = 0.$$  \hfill (102)

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

$$\bar{X}[\mathcal{F}] = \bar{X}[Y_j] + \bar{X}[N_j] = 0,$$

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

We will prove equation (102) by induction on $n - |A \cup B|$. Suppose first that $n = |A \cup B|$. Since $\gamma$ is an obstruction, the algorithm enforces equation (62) and we conclude $\bar{X}[\gamma] = 0$.

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

$$Y_h \bigcap \gamma.$$

By Lemma 4.22 this expression is of the form $\theta^\#$ for some $\theta$, with one negation involved; in step 4(iii)(a) the algorithm will impose:

$$\bar{X}[\gamma] = \bar{X}[Y_h \bigcap \gamma] + \bar{X}[N_h \bigcap \gamma]$$

and by the inductive hypothesis we are done. \blacksquare

One can also show (with simpler proof) that the inequality shown in [CD01] to have $N_+\text{-rank } n$ is implied by the level-2 $\Sigma$ algorithm.

Lemma 4.24 Consider the region \( \{x \in \{0,1\}^n : \sum_j x_j \geq \frac{1}{4} \} \). The level-2 $\Sigma$-algorithm proves $\sum_j x_j \geq 1$.

\blacksquare

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 level-$k$ $\Sigma$-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. This is an approach that we plan to explore.

Acknowledgment This paper was partially written at Bell Laboratories. We thank Bruce Shepherd for many useful comments.
Appendix – Circulant matrices and set covering

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

\[
\sum_{j \neq k} x_j \geq 1, \quad \text{for each } k, \ 1 \leq k \leq n
\]

(103)

\[
x \in \{0,1\}^n
\]

for \(n > 1\), and show that the valid inequality

\[
\sum_j x_j \geq 2
\]

(104)

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

\[
Ax \geq b
\]

(105)

\[
x \in \{0,1\}^n
\]

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

\[
f(Q,P) = \prod_{j \in Q} x_j \cdot \prod_{j \in P} (1 - x_j)
\]

where \(Q\) and \(P\) are disjoint subsets of \(\{1, 2, \cdots, n\}\) and \(|Q| + |P| = t\), we multiply each constraint of (105) by \(f(Q, P)\), obtaining a (valid) polynomial inequality of the form

\[
(a_r x - b)f(Q, P) \geq 0.
\]

(106)

In addition, we write the (valid) inequality

\[
f(Q, P) \geq 0.
\]

(107)

Next, we linearize the constraints (106) and (107): we replace \(x_j^2\) with \(x_j\) for each \(j\), and then replace 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 we 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

\[
x_j + \bar{x}_j = 1.
\]

(108)

**Step S2.** For each polynomial of the form

\[
g(Q, P) = \prod_{j \in Q} x_j \cdot \prod_{j \in P} \bar{x}_j
\]

where \(Q\) and \(P\) are disjoint subsets of \(\{1, 2, \cdots, n\}\) and \(|Q| + |P| = t\), we multiply each constraint of type (105) or (108) by \(g(Q, P)\), obtaining a (valid) polynomial inequality of the form

\[
(a_r x - b)g(Q, P) \geq 0.
\]

(109)
In addition, we write the (valid) inequality

\[ g(Q, P) \geq 0. \]  \hfill (110)

**Step S3.** Linearize the constraints (109) and (110) by setting \( x_j^1 = x_j, \ x_j^2 = \bar{x}_j \) and \( x_j \bar{x}_j = 0 \) for all \( j \), and then replacing each product of the form \( \prod_{i \in K} x_i \prod_{j \in L} \bar{x}_j \) (\( K \) and \( L \) disjoint subsets of \( \{1, 2, \ldots, n\} \)) with possibly \( |K \cup L| = 1 \) with a new variable variable \( w(K, L) \). [Note: In the case of (110) 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-1 operator. Note: the variables \( y \) 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 (103). Denote \( E_n = \{1, 2, \ldots, n\} \). We will show that (104) 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 \( |Q \cup P| \leq n - 2 \) which satisfy the constraints imposed in Steps S2 and S3, and which however violate (104). 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 \( j \notin Q \cup P \),

\[ w(Q \cup j, P) + w(Q, P \cup j) - w(Q, P) = 0, \]  \hfill (111)

(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 \( k \).

\[
\begin{align*}
\sum_{j \in Q - k} w(Q, P) + \sum_{j \in E_n - (Q \cup P)} w(Q \cup j, P) - w(Q, P) \geq 0, & \text{ if } k \in Q \hfill (112) \\
\sum_{j \in Q} w(Q, P) + \sum_{j \in E_n - (Q \cup P)} w(Q \cup j, P) - w(Q, P) \geq 0, & \text{ if } k \in P \hfill (113) \\
\sum_{j \in Q} w(Q, P) + \sum_{j \in E_n - (Q \cup P \cup k)} w(Q \cup j, P) - w(Q, P) \geq 0, & \text{ if } k \in E_n - (Q \cup P). \hfill (114)
\end{align*}
\]

Note: there is no constraint (112) when \( Q = \emptyset \). In addition, we require

\[ w(\emptyset, \emptyset) = 1. \]  \hfill (115)

Finally, besides satisfying all these inequalities, we also want:

\[ \sum_j w(j, \emptyset) < 2, \]  \hfill (116)

i.e. (104) is violated.

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:

\[
\begin{align*}
(z + 1, 1) + (z, p + 1) - z(q, p) = 0 & \text{ if } q + p \leq n - 1 \hfill (117) \\
(q - 2)z(q, p) + (n - q - p)z(q + 1, p) \geq 0 & \text{ if } q + p \leq n - 3 \text{ and } q > 0 \hfill (118) \\
(n - p - 1)z(1, p) - z(0, p) \geq 0 & \text{ if } p \leq n - 3 \hfill (119) \\
z(0, 0) = 1 & \hfill (120) \\
z(1, 0) < 2 & \hfill (121)
\end{align*}
\]

Then there is a set of nonnegative values \( g(Q, P) \) that satisfy (111), (112), (113), (114), (115) and (116). In addition, (111) 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 $|Q \cup P| \leq n - 2$, set

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

Then (111) follows from (117), (115) from (120) and (116) from (121). It remains to show that (112) - (114) 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

$$(q - 1)z(q,p) + (n - q - p)z(q + 1, p) - z(q,p) \geq 0 \tag{122}$$

$$qz(q,p) + (n - q - p)z(q + 1, p) - z(q,p) \geq 0 \tag{123}$$

$$qz(q,p) + (n - q - p - 1)z(q + 1, p) - z(q,p) \geq 0 \tag{124}$$

corresponding, respectively to (112), (113) and (114). Clearly, (122) is more binding than (123), and (122) is more binding than (124) because, by (111), $z(q,p) \geq z(q + 1, p)$. Hence, (122), (123) and (124) hold if and only if (122) does; but this is equivalent to (118). The case $q > 0$ is completed.

Suppose next that $q = 0$. Then (112) is not a constraint, and corresponding to (113) and (114) we have to satisfy

$$(n - p)z(1, p) - z(0, p) \geq 0, \quad \text{and} \tag{125}$$

$$(n - p - 1)z(1, p) - z(0, p) \geq 0. \tag{126}$$

and this holds because of (119). 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 do something stronger: we will define $z(q,p)$ for all for all $q, p$ with $q + p \leq n$, and our choice of values will satisfy (117) for all $q, p$ with $q + p \leq n - 1$, and not just $q + p \leq n - 3$. This is done as follows.

Define

$$\kappa = \frac{2}{n^2 - n + 2} \tag{127}$$

and set:

$$z(q,p) = 0, \quad \forall \ p \geq 3 \ \text{and} \ q \leq n - p \tag{128}$$

$$z(2,p) = \kappa, \quad \forall \ p \leq n - 2 \tag{129}$$

$$z(1,p) = (n - 1 - p)\kappa, \quad \forall \ p \leq n - 1 \tag{130}$$

$$z(0,p) = \left(\frac{(n - p)(n - 1 - p)}{2} + 1\right)\kappa, \quad \forall \ p \leq n \tag{131}$$

Now we have:

**Lemma 5.2** The choice of values $z$ as in (127 - 131) satisfies the conditions of Lemma 5.1.

Proof. We show first that (117) and (118) hold. These are homogeneous inequalities, hence satisfied for $q \geq 3$ by (128).

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 (117) holds, and (118) 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 (117) holds. Also, (118) is equivalent to $-(n - p - 1)\kappa + (n - p - 1)\kappa \geq 0$, which is true.
Finally, consider the case $q = 0$. Then (117) is equivalent to:

\[(n - p - 1) + \frac{(n - p - 1)(n - 2 - p)}{2} + 1 = \frac{(n - p)(n - 1 - p)}{2} + 1, \quad (132)\]

which is true by inspection.

In order to show that (119) holds, notice that it is equivalent to

\[(n - p - 1)^2 \geq \frac{(n - p)(n - 1 - p)}{2} + 1 \quad (133)\]

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

Next, note that $z(0, 0) = \left(\frac{n(n-1)}{2} + 1\right)\kappa = 1$, so (120) holds.

Finally,

\[nz(1, 0) = n(n - 1)\kappa < 2 \quad (135)\]

by definition of $\kappa$, hence (121) 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 (104).

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

**Theorem 5.3** The Sherali-Adams rank of constraint (104) 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 (104) is at least $n - 3$. But how about its $N_+\text{-rank}$? In order to answer this question, consider the matrix $W^w$. This is the matrix whose rows and columns are indexed by subsets of $E_n$, and whose $I,J$ entry is defined by:

\[W_{I,J}^w = w(I \cup J, \emptyset). \quad (136)\]

where the $w$ are the values we construct in Lemma 5.1 using the $v$ values we described above. We will prove below that $W \succeq 0$. 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) 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 [Z02]. Moreover, one can also similarly show that our Lemmas above imply that at least $n - 3$ rounds of $\hat{S}$ are needed to satisfy (104). We conjecture that the Lasserre-rank of (104) also grows as a function of $n$.

Now we return to the proof of $W \succeq 0$. 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** $W \succeq 0$. 33
Proof. We will use Theorem 2.2. Consider abstract events $I_j$, $1 \leq j \leq n$, and a probability measure $\Upsilon$, defined as follows: on an atom

\[ \bar{\alpha} = \left( \bigcap_{j \in Q} I_j \right) \cap \left( \bigcap_{j \in E_n - Q} \bar{I}_j \right) \]

we set

\[ \Upsilon(\alpha) = w(Q, E_n - Q), \quad (137) \]

and then we extend $\Upsilon$ 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 $\alpha$ contained in $A$).

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

\[ \Upsilon\left( \bigcap_{j \in Q} I_j \right) = w(Q, \emptyset), \quad (138) \]

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

\[ \Upsilon\left( \bigcap_{j \in Q} I_j \cap \bigcap_{j \in P} \bar{I}_j \right) = w(Q, P). \quad (139) \]

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 (139) follows by definition (137).

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

\[ \Upsilon\left( \bigcap_{j \in Q} I_j \cap \bigcap_{j \in P} \bar{I}_j \right) = \Upsilon\left( \bigcap_{j \in Q \cup k} I_j \cap \bigcap_{j \in P \cup k} \bar{I}_j \right) = w(Q \cup k, P) + w(Q, P \cup k) \quad (by \ induction) \]

\[ = w(Q, P) \quad (by \ (111)) \]

as desired. The theorem is proved.

\section{Theorem 5.5} $W \succeq 0$.

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 (104), 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. (4) and (5)) 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 $\alpha \geq 0$ and $\sum_{r \in L} \alpha_r = 1$, and such that $Wy = W$.

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 $\alpha$:

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

Note that

\[ y_H = \begin{cases} 0 & \text{if } |H| > 2 \\ \kappa & \text{if } |H| = 2 \\ (n - 1)\kappa & \text{if } |H| = 1 \\ 1 & \text{if } H = \emptyset \end{cases} \quad (143) \]

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

\[ \blacksquare \]
References


[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.


[GT01] M.X. Goemans and L. Tunçel, When does the positive semidefinite-ness constraint help in lifting procedures, manuscript (2000).


