CORC REPORT 2003-09

Tree-width and the Sherali–Adams operator*

Daniel Bienstock and Nuri Ozbay
Columbia University
New York, NY 10027

November 11, 2003
version 2004-March-23

Abstract

We describe a connection between the tree-width of graphs and the Sherali–Adams reformulation procedure for 0/1 integer programs. For the case of vertex packing problems, our main result can be restated as follows: let G be a graph, let $k \geq 1$ and let $\hat{x} \in R^{V(G)}$ be a feasible vector for the formulation produced by applying the level-k Sherali–Adams algorithm to the edge formulation for $STAB(G)$. Then for any subgraph H of G, of tree-width at most k, the restriction of $\hat{x}$ to $R^{V(H)}$ is a convex combination of stable sets of H.

1 Introduction

A 0/1 *packing set* is a feasible region of the form $P^b_A = \{x \in \{0,1\}^n \, : \, Ax \leq b\}$, where A is a nonnegative, $m \times n$ matrix, and $b \in \Re^n_+$. Given such a matrix A we can define its *clique graph*, which is the graph G_A with a vertex corresponding to each column of A and an edge between two vertices j_1 and j_2 if there exists some row i with $a_{i,j_1} > 0$ and $a_{i,j_2} > 0$.

Given a vector $\alpha \in \Re^n$, denote by suppt(α) the support of α, i.e. the set $\{j \, : \, \alpha_j \neq 0\}$. We will use the notation $G_A[\alpha]$ to abbreviate $G_A[\text{suppt}(\alpha)]$, that is, the subgraph of G_A induced by suppt(α).

In this note we consider the relationship between valid inequalities $\alpha^T x \leq \beta$ that are "simple", as measured by the tree-width (defined below) of an appropriate subgraph $G_A[\alpha]$ and the strength of the relaxation provided by the Sherali–Adams operator (also defined below). Given a set of rows R of a matrix A, we denote by $A(R)$ the corresponding submatrix.

Definition 1.1 *Consider a 0/1 packing set P^b_A. The tree-width of a valid inequality $\alpha^T x \leq \beta$ is the minimum, over all subset R of rows of A such that $\alpha^T x \leq \beta$ is valid for $P^{b(R)}_{A(R)}$, of the tree-width of $G_{A(R)}[\alpha]$.*

Our main result is:

Theorem 1.2 *Consider a 0/1 packing set P^b_A. Let $k \geq 1$, and suppose that a vector $\hat{x} \in \Re^n$ satisfies the constraints imposed by the level-k Sherali–Adams operator applied to P^b_A.*

(1) $\hat{x}$ satisfies every valid inequality $\alpha^T x \leq \beta$ whose tree-width is at most $k-1$.

(2) Suppose A is 0/1 and b is integral. Then $\hat{x}$ satisfies every valid inequality $\alpha^T x \leq \beta$ whose tree-width is at most $\min\{k, \, n-1\}$.
■

In Section 3 we describe some applications of Theorem 1.2. We show how some rich families of valid inequalities for vertex packing problems (in particular, antiweb-wheel inequalities) are guaranteed to be satisfied when we apply the Sherali–Adams procedure. We also describe a class of inequalities that have fixed Sherali–Adams rank but unbounded N_0-rank.

*This research was partially funded by NSF awards CCR-ITR-0213844 and DMI-0200221.

1.1 Preliminaries

1.1.1 Tree-width

For a graph G, its vertex set is denoted by $V(G)$. Let H be a graph. A *tree-decomposition* [RS86] of H is a pair (T, X) where T is a tree and $X = \{X_t : t \in V(T)\}$ is a family of subsets of $V(H)$ such that

(i) For all $v \in V(H)$, the set $\{t \in V(T) : v \in X_t\}$ forms a subtree of T, and

(ii) For each $\{u, v\} \in E(H)$ there is a $t \in V(T)$ such that $\{u, v\} \subseteq X_t$.

The *width* of the decomposition is $\max\{|X_t| : t \in V(T)\} - 1$. The tree-width of H is the minimum width of a tree-decomposition of H. We note that tree-width has emerged as a fundamental measure of the "complexity" of a graph. A full survey is beyond the scope of this paper – see, for example [RS86], [R97]. Broadly speaking, graphs of low-tree width are "simple". This simplicity can be exploited in combinatorial optimization – combinatorial problems on graphs of small tree-width can in many cases be efficiently solved by dynamic programming. See, for example [ALS91]. Cook and Seymour [CS94], [CS03] have described sophisticated extensions and implementations of these ideas. See [H04] for a recent survey of results on tree-width, branch-width and related topics.

1.1.2 The Sherali–Adams procedure

The Sherali–Adams operator is one of several "lift-and-project" procedures that, given a formulation for a 0/1 integer program, produce a tighter formulation by adding new variables and constraints. The lifting step is that of adding the new variables, and typically (though not always) one projects the new formulation to the space of the original variables.

Let $k \geq 1$ be an integer. The Sherali–Adams level-k formulation derived from P_A^b is the following (in slightly redundant form).

It has a variable $w[Y, N]$ for every pair of disjoint subsets Y, N of $\{1, 2, \cdots, n\}$ with $|Y \cup N| \leq \min\{k+1, n\}$. In addition, it has the constraints (i)–(iv) given next:

(i)

$$w[\emptyset, \emptyset] \;=\; 1. \tag{1}$$

(ii) For all disjoint subsets Y, N of $\{1, 2, \cdots, n\}$ with $|Y \cup N| \leq \min\{k+1, n\}$,

$$0 \;\leq\; w[Y, N] \;\leq\; w[Y - j, N] \qquad \forall\, j \in Y \tag{2}$$

$$w[Y, N] \;\leq\; w[Y, N - j] \qquad \forall\, j \in N \tag{3}$$

(iii) For all disjoint subsets Y, N of $\{1, 2, \cdots, n\}$ with $|Y \cup N| \leq k$, and any $j \notin Y \cup N$,

$$w[Y \cup j, N] + w[Y, N \cup j] - w[Y, N] \;=\; 0 \tag{4}$$

(iv) Let m denote the number of rows of A. For all disjoint subsets Y, N of $\{1, 2, \cdots, n\}$ with $|Y \cup N| \leq k$, and any i, $1 \leq i \leq m$,

$$\sum_{j \notin N} a_{i,j} w[Y \cup j, N] - b_i w[Y, N] \;\leq\; 0. \tag{5}$$

Note that property (iii) implies that for any disjoint subsets Y, N of a subset U,

$$\sum_{Y \subseteq Z \subseteq U \setminus N} W[Z, U \setminus Z] = W[Y, N].$$

We say that $\hat{x} \in \Re_+^n$ satisfies the level-k Sherali–Adams constraints if there is a vector $\hat{w} \in \Re_+^M$ ($M = \sum_{j=0}^{k+1} 2^j \binom{n}{j}$) which satisfies (i)–(iv) and such that $\hat{x}_j = \hat{w}[j, \emptyset]$, for all $1 \leq j \leq n$. [Note: the above description of the operator is redundant, and $\binom{n}{k+1}$ variables suffice]. Throughout, we abbreviate $\{j\}$ as j. We will also abbreviate by SA^k the level-k operator. It is known that the SA^n operator yields the convex hull of P_A^b. Further, for fixed k, one can optimize over the formulation produced by SA^k in polynomial time.

Other examples of lift-and-project operators include the Lovász–Schrijver procedures N and N_+ [LS91], the Balas–Ceria–Cornuéjols procedure, the Lasserre procedure [L01b] and the Σ procedure in [BZ02a]. Also see [B79]. Laurent [L01b] has provided a common framework for understanding and comparing the N, N_+, SA and Lasserre operators. Also see Cook and Dash [CD01], Goemans and Tunçel [GT01]. Some recent results on the rank of cutting-plane procedures are given by Cornuéjols and Li in [CL02].

Lift-and-project operators can be complex, and the benefit of obtaining a tighter formulation can be offset by the overhead of running a larger formulation. Thus, it is of interest to describe the strength of an operator in terms of fundamental properties of a problem. The main result in this paper is of this type. Further, it is also of interest to compare the strength of the different lift-and-project operators, since they have widely different computational overhead (refer to [L01b]). In particular, it is known that the SA^k operator is at least as strong than the k-step iterated N procedure, N^k.

1.1.3 Strategy for the proof of Theorem 1.2

Consider a tree-decomposition (T, X) of a graph G. If we view T as rooted, we obtain a recipe for constructing G in a sequence of composition steps involving smaller graphs. The initial building blocks are the subgraphs of G induced by the sets X_t for each leaf t of T (other than the root of T, if it is a leaf). In each composition step, we take two graphs G_1 and G_2, each endowed with a set B_i ($i = 1, 2$) of "boundary vertices" and we identify some members of B_1 with an equal number of members of B_2. Furthermore, the set of boundary vertices of the resulting graph (to be used in future compositions) is a subset of $B_1 \cup B_2$. When we can find such a composition strategy such that the set of boundary vertices of any intermediate graph is of cardinality $\leq k+1$, then G has tree-width $\leq k$.

This insight is not new; it underlies much of the work cited above on Graph Minors, and, especially, algorithms on bounded tree-width graphs. What is the significance of having a tree-decomposition of small width, in terms of the vertex packing polyhedron of G?

Suppose again that we are considering the composition step involving G_1 and G_2 described above. Moreover, suppose that we have a fractional vector $\hat{x}$ ($\in R^{V(G)}$) such that for $i = 1, 2$, the restriction of $\hat{x}$ to $V(G_i)$ is a convex combination of incidence vectors of stable sets of G_i. Will the same be true when we restrict $\hat{x}$ to $V(G_1) \cup V(G_2)$? The difficulty here is that the two decompositions of $\hat{x}$ have, in some sense, to agree on $B_1 \cap B_2$. We cannot expect that in general this will be the case – further conditions on $\hat{x}$ are required. It turns out that the conditions imposed by the SA^k operator are enough, when $k \geq \max\{|B_1|, |B_2|\} - 1$. This is at the core of the proof provided next.

2 Proof of Theorem 1.2

Lemma 2.1 *Let H be a graph of tree-width k. Then there is a tree-decomposition of (T, X) of H of width k such that every vertex of T has degree ≤ 3, and further satisfying:*

(1) If t is a vertex of T with neighbors u, v, w then $X_u = X_v = X_w = X_t$, and u, v, w all have degree 2.

(2) If u and v are neighbors in T, both of degree ≤ 2, then either $X_u \subseteq X_v$, or $X_v \subseteq X_u$, or $X_u \cap X_v = \emptyset$.

Proof. That T can be assumed to have vertices of degree at most 3 is straightforward. To obtain (1), whenever t is a degree-3 vertex of T then subdivide the three edges incident with t to introduce new vertices t_i with $X_{t_i} = X_t,\ 1 \leq i \leq 3$. (2) follows similarly. ■

The following result shows how the assumption $A \geq 0$ plays a role in our proof.

Lemma 2.2 *Let $k \geq 1$ and A a nonnegative matrix. Suppose $\hat{x} \in \Re^n_+$ satisfies the requirements of the SA^k procedure applied to P^b_A. Let C be a subset of the columns of A and R a subset of the rows of A. Then the restriction of $\hat{x}$ to C satisfies the constraints generated by the level-$\min\{k, |C|\}$ Sherali–Adams procedure when applied to $\left\{x \in \{0,1\}^C \ : \ \sum_{j \in C} a_{ij} x_j \leq b_i \quad \forall\ i \in R\right\}$.*

Proof. Let $\hat{w}$ satisfy (1)–(5), and $\hat{x}_j = \hat{w}[j, \emptyset]$ for $1 \leq j \leq n$. Then for each pair of disjoint subsets Y, N with $Y \cup N \subseteq C$ and $|Y \cup N| \leq k$, and for any $i \in R$,

$$\sum_{j \in C - N} a_{ij} \hat{w}[Y \cup j, N] \quad \leq \quad \sum_{j \notin N} a_{ij} \hat{w}[Y \cup j, N] \quad \leq b_i \hat{w}[Y, N], \tag{6}$$

where the first inequality follows since $A \geq 0$. ■

In the remainder of the proof of Theorem 1.2, we will assume that R is the set of all rows – this assumption is warranted by Lemma 2.2. We will also consider:

A.1. A fixed $m \times n$ nonnegative matrix A, a fixed $\hat{x} \in \Re^n$, a fixed $k \geq 1$, and a fixed $\hat{w} \in \Re^M$ ($M = \sum_{j=0}^{k+1} 2^j \binom{n}{j}$) which satisfies conditions (i)–(iv) of the level-k Sherali–Adams procedure applied to P_A^b, and such that $\hat{x}_j = \hat{w}[j, \emptyset]$, for all $1 \leq j \leq n$,

Our ultimate goal is to show $\alpha^T \hat{x} \leq \beta$, for any inequality $\alpha^T x \leq \beta$ valid for P_A^b, of tree-width $\leq k-1$ in case (1) of Theorem 1.2, and of tree-width $\leq \min\{k,\, n-1\}$ in case (2). To simplify nomenclature, we call such an inequality *appropriate*. There is one way that α can be further constrained.

Lemma 2.3 *If $\hat{x}$ satisfies every appropriate inequality $\alpha^T x \leq \beta$ such that $G[\alpha]$ is connected, then $\hat{x}$ satisfies every appropriate inequality.*

Proof. Consider an appropriate inequality $\alpha^T x \leq \beta$ such that $G[\alpha]$ is not connected. Then we can write $G = G^1 \cup G^2$ where the G^i have disjoint vertex sets. For $i = 1, 2$ let V^i be the vertex set of G^i, let $\beta^i = \max_{x \in P_A^b} \sum_{j \in V^i} \alpha_j x_j$, and let $\bar{x}^i \in P_A^b$ attain this maximum. Since $A \geq 0$, without loss of generality we have $\bar{x}^i_j = 0$ for $j \notin V^i$. As a result, $\beta \geq \beta^1 + \beta^2$ – this holds because the vector z defined by

$$z_j = \begin{cases} \bar{x}^i_j, & \text{if } j \in V^i \\ 0, & \text{otherwise,} \end{cases} \tag{7}$$

is in P_A^b, since the V^i are disjoint. Thus the two valid inequalities $\sum_{j \in V^i} \alpha_j x_j \leq \beta_i$ together dominate $\alpha^T x \leq \beta$. Since the tree-width of each G^i cannot exceed that of G, it suffices to show that $\sum_{j \in V^i} \alpha_j \hat{x}_j \leq \beta^i$ in order to obtain $\alpha^T \hat{x} \leq \beta$. ■

In the remainder of the analysis, we will also consider

A.2. A fixed appropriate inequality $\alpha^T x \leq \beta$ such that $G[\alpha]$ is connected.

A.3. A fixed tree-decomposition (X, T) of $G_A[\alpha]$ that satisfies the conditions of Lemma 2.1, of width $\leq k-1$ in case (1) of Theorem 1.2, and of width $\leq \min\{k,\, n-1\}$ in case (2). In addition, we view T as *rooted* at an arbitrary, but fixed, **leaf** $r \in V(T)$.

In order to describe our strategy for the proof, we need one further definition.

Definition 2.4 *Let W and U be such that $W \subseteq U \subseteq$ suppt(α). We say that U is W-decomposable if there exist 0/1 vectors $v^i \in P_A^b$ and weights $0 \leq \lambda^i \leq 1$, for $1 \leq i \leq I$ (I some positive integer) such that*

(a) suppt$(v^i) \subseteq U$ for all $1 \leq i \leq I$,

(b) $\sum_{i=1}^{I} \lambda^i = 1$,

(c) $\hat{x}_j = \sum_i \lambda^i v^i_j$ for all $j \in U$, and

(d) $\hat{w}[Y, N] = \sum\{\lambda^i \,:\, v^i_j = 1 \;\forall j \in Y, \text{ and } v^i_j = 0 \;\forall j \in N\}$, for all disjoint subsets Y, N of W such that $|Y \cup N| \leq k$.

Informally, U is W-decomposable if the restriction of $\hat{x}$ to U can be decomposed as a convex combination of 0/1 feasible vectors, and furthermore the decomposition is more precise when restricted to W – it carries over to the level-k Sherali–Adams iterates $\hat{w}$.

Our strategy to prove Theorem 1.2 will be as follows. For each vertex $t \in V(T)$, let $D_t = \bigcup\{X_s : s \text{ a descendant of } t\}$ (note: t is a descendant of itself). We will show that D_r is X_r-decomposable. To see that this implies Theorem 1.2, let v^i, λ^i ($1 \leq i \leq I$) satisfy the conditions of Definition 2.4. Then $\alpha^T v^i \leq \beta$ for $1 \leq i \leq I$. But $D_r = \text{suppt}(\alpha)$, and so condition (c) of Definition 2.4 completes the proof.

In order to prove that D_r is X_r-decomposable we will prove that D_t is X_t-decomposable for each $t \in V(T)$, by induction, starting at the leaves of T and moving up toward r. To handle the start of the induction we have the following.

Lemma 2.5 *For each $t \in V(T)$, X_t is X_t-decomposable.*

Proof. Let $t \in V(T)$. Note that in both cases of Theorem 1.2 the SA^k operator produces variables $w[Y, N]$ for each pair Y, N such that $Y \cup N = X_t$. For each such pair with $\hat{w}[Y, N] > 0$, define a vector $v^{Y,N} \in \{0, 1\}^n$ where $v_j^{Y,N} = 1$ if and only if $j \in Y$. For each such pair also define $\lambda^{Y,N} = \hat{w}[Y, N]$.

Requirement (a) of Definition 2.4 holds by construction. By the properties of the SA^k algorithm, in particular property (iii), requirements (b)–(d) are satisfied. For example, (b) follows since $\sum_Y \hat{w}[Y, N] = \hat{w}[\emptyset, \emptyset] = 1$. What remains to be shown is that each $v^{Y,N} \in P_A^b$. First consider case (1) of Theorem 1.2, and let i be any row of A. We know that $\hat{w}$ satisfies the constraint obtained when we apply condition (iv) of the SA^k procedure to row i, and pair Y, N, which is:

$$\sum_{j \in Y} a_{ij} \hat{w}[Y, N] + \sum_{j \notin Y \cup N} a_{ij} \hat{w}[Y \cup j, N] \le b_i \, \hat{w}[Y, N],$$

and therefore

$$\sum_{j \in Y} a_{ij} \le b_i.$$

Thus $v^{Y,N} \in P_A^b$, as desired.

For case (2) of Theorem 1.2, let i again denote a row of A. If $a_{ij} = 1$ for at most b_i elements $j \in Y$, then clearly $v^{Y,N}$ satisfies constraint i and we are done. If $a_{ij} = 1$ for more than b_i elements $j \in Y$, take any fixed element $u \in Y$ with $a_{iu} = 1$. We have that $\hat{w}$ satisfies the constraint obtained when we apply condition (iv) of the SA^k procedure to the pair $[(Y - u), N]$ and row i, which is:

$$\hat{w}[Y, N] + \sum_{j \in Y - u} a_{ij} \hat{w}[Y - u, N] + \sum_{j \notin Y \cup N} a_{ij} \hat{w}[(Y - u) \cup j, N] \le b_i \, \hat{w}[Y - u, N].$$

By assumption the second term in the left-hand side has at least b_i terms. It follows that $\hat{w}[Y, N] = 0$. This concludes the proof. ■

To handle the general inductive step of Theorem 1.2, let $t \in V(T)$ be a non-leaf vertex such that the inductive hypothesis applies to every child of t. Recall that (T, X) is assumed to satisfy the conditions of Lemma 2.1.

Clearly, for any child s of t we must have $X_s \cap X_t \neq \emptyset$, or else $G[\alpha]$ is not connected.

Suppose first that there is a child s of t, of degree three in T. Then the result follows by Lemma 2.1 (1), since $X_s = X_t$ and thus $D_s = D_t$. Hence we may assume that every child of t has degree ≤ 2.
We organize the analysis into three cases:

(a) t has degree ≤ 2, and for some child s of t, $X_t \subseteq X_s$.

(b) t has degree ≤ 2, and for every child s of t, $X_s \subseteq X_t$.

(c) t has degree 3.

In case (a) we have $D_t = D_s$. Pick any $j \in X_s \backslash X_t$. Then condition (iii) of the Sherali–Adams operator implies that that D_s is $(X_s - j)$-decomposable, since D_s is X_s-decomposable by the inductive assumption. Repeating this step, we will obtain that D_s $(= D_t)$ is X_t-decomposable.

The remaining two cases, (b) and (c), are handled by the following result, as we will see.

Lemma 2.6 *For $h = 1, 2$, suppose U^h is W^h-decomposable, where $W^h \subseteq U^h \subseteq suppt(\alpha)$. Assume that*

(1) $W^1 \subseteq W^2$, and $|W^2| \le k$,

(2) $U^1 \cap U^2 = W^1$, and

(3) there are no edges of $G_A[\alpha]$ with one end in $U^1 \backslash W^1$ and the other in $U^2 \backslash W^1$.

Then $U^1 \cup U^2$ is W^2-decomposable.

Postponing the proof of Lemma 2.6 for the moment, let us see how to apply the Lemma to cases (b) and (c) above. For case (b), note that t has only one child, since the root of T was assumed to be a leaf. Denoting this child by s, we set $W^1 = X_s$, $U^1 = D_s$ and $W^2 = D^2 = X_t$. Then clearly (1) and (2) in Lemma 2.6 hold. Further, (3) holds by the definition of tree decomposition, since for any edge $\{u, v\}$ with $u \in D_s$ and $v \in X_t$, we must have that either $u \in X_s$ or $v \in X_s$, and hence no edge as in (3) exists. Similarly, for case (c), let p and q be the two children of t. Then we set $W^1 = W^2 = X_t$ $(= X_p = X_q)$ and $U^1 = D_p$ and $U^2 = D_q$.

Proof of Lemma 2.6. Let p^i be 0/1 vectors and λ^i be reals ($1 \le i \le I$) that satisfy the conditions of Definition 2.4 with respect to W^1 and U^1. Similarly, let q^j be 0/1 vectors and μ^j be reals ($1 \le j \le J$) that satisfy the conditions of Definition 2.4 with respect to W^2 and U^2.

For each pair i, j with $1 \le i \le I$ and $1 \le j \le J$ such that

$$\text{suppt}(p^i) \cap W^1 \quad = \quad \text{suppt}(q^j) \cap W^1$$

we do the following. Define the vector $v^{i,j} \in \{0,1\}^n$ so that $\text{suppt}(v^{i,j}) = \text{suppt}(p^i) \cup \text{suppt}(q^j)$. Thus, $\text{suppt}(v^{i,j}) \subseteq U^1 \cup U^2 \subseteq \text{suppt}(\alpha)$. Since $p^i \in P_A^b$ and $q^j \in P_A^b$, assumption (3) of this Lemma implies that $v^{i,j} \in P_A^b$ as well. In addition, if we write

$$Z^{i,j} \quad \doteq \quad \text{suppt}(v^{i,j}) \cap W^1$$

then

$$Z^{i,j} \quad = \quad \text{suppt}(p^i) \cap W^1 \ = \ \text{suppt}(q^j) \cap W^1$$

by construction. Define the real $\gamma^{i,j}$ by the following rule

$$\gamma^{i,j} \quad = \quad 0, \ \text{ if } \hat{w}[Z^{i,j}, W^1 - Z^{i,j}] = 0, \tag{8}$$

and

$$\gamma^{i,j} \quad = \quad \frac{\lambda^i \mu^j}{\hat{w}[Z^{i,j}, W^1 - Z^{i,j}]} \tag{9}$$

otherwise.

We claim that the family of all vectors $v^{i,j}$ (at most $|I||J|$ vectors) together with the reals $\gamma^{i,j}$ show that $U^1 \cup U^2$ is W^2-decomposable. To see this, note that

$$\begin{aligned}
\sum_{i,j} \gamma^{i,j} \quad &= \quad \sum_{Z \subseteq W^1} \sum \left\{ \frac{\lambda^i \mu^j}{\hat{w}[Z, W^1 - Z]} \ : \ Z = \text{suppt}(p^i) \cap W^1 = \text{suppt}(q^j) \cap W^1 \right\} &(10)\\
&= \quad \sum_{Z \subseteq W^1} \left(\sum_{j \,:\, Z = \text{suppt}(q^j) \cap W^1} \mu^j \right) \left(\sum_{i \,:\, Z = \text{suppt}(p^i) \cap W^1} \frac{\lambda^i}{\hat{w}[Z, W^1 - Z]} \right) &(11)\\
&= \quad \sum_{Z \subseteq W^1} \left(\sum_{j \,:\, Z = \text{suppt}(q^j) \cap W^1} \mu^j \right) &(12)\\
&= \quad \sum_j \mu^j \ = \ 1. &(13)
\end{aligned}$$

Similarly, assume Y, N are disjoint subsets of W^2. Then

$$\begin{aligned}
\sum \left\{ \gamma^{i,j} \ : \ v_h^{i,j} = 1 \ \forall h \in Y, \text{ and } v_h^{i,j} = 0 \ \forall h \in N \right\} \quad &= &(14)\\
\sum_{Z \,:\, Y \subseteq Z \subseteq W^2 - N} \left(\sum_{j \,:\, \text{suppt}(q^j) \cap W^2 = Z} \left\{ \sum_{i \,:\, \text{suppt}(p^i) \cap W^1 = Z \cap W^1} \frac{\lambda^i \mu^j}{\hat{w}[Z \cap W^1, W^1 - Z \cap W^1]} \right\} \right) \quad &= &(15)\\
\sum_{Z \,:\, Y \subseteq Z \subseteq W^2 - N} \left(\sum_{j \,:\, \text{suppt}(q^j) \cap W^2 = Z} \mu^j \right) \quad &= &(16)\\
\hat{w}[Y, N], & &(17)
\end{aligned}$$

as desired, where the last equation follows by property (d) of Definition 2.4. One similarly shows that for any $h \in U^1 \cup U^2$,

$$\hat{x}_h \quad = \quad \sum \left\{ \gamma^{i,j} \ : \ v_h^{i,j} = 1 \right\}. \tag{18}$$

This completes the proof. ■

At the end of the following section we will examine in what sense Theorem 1.2 is tight. In the rest of this section we show how Theorem 1.2 can extend to valid inequalities $\alpha^T x \le \beta$ of high tree-width. This result relies on a generalization to packing problems of a technique introduced in [CC97] in the context of the vertex-packing problem, which is itself related to the classical "vertex multiplication" technique.

Definition 2.7 *Let $A' \geq 0$ be $m \times n$. Let p and q be columns of A' and let A be the $m \times (n-1)$ matrix obtained from A' by replacing columns p and q with a single column, denoted by (pq), which is equal to the sum of p and q. We say that A is obtained from A' by identifying p and q into (pq).*

The following is clear:

Lemma 2.8 *Let the matrix A be obtained from A' by identifying columns p and q into the column (pq). Suppose $\sum_{j=1}^{n} \alpha_j x_j \leq \beta$ is valid for $P^b_{A'}$. Then*

$$(\alpha_p + \alpha_q) x_{(pq)} + \sum_{j \neq p,\, q} \alpha_j x_j \leq \beta$$

is valid for P^b_A. ■

Lemma 2.8 can be applied to a sequence of column identifications. Starting with an inequality $\alpha^T x \leq \beta$ valid for P^b_A, of small tree-width, the tree-width of the final inequality can be arbitrarily high. We have the following result.

Lemma 2.9 *Let A be the matrix obtained from A' by identifying columns p and q into the column (pq). Let $1 \leq k \leq n-1$, and suppose that $\hat{x} \in \Re^{n-1}_{+}$ satisfies the conditions imposed by the SA^k procedure applied to $Ax \leq b$. Then the vector $\hat{x}' \in \Re^n$ defined by $\hat{x}'_p = \hat{x}'_q = \hat{x}_{(pq)}$ and $\hat{x}'_j = \hat{x}_j$ for all $j \neq (pq)$, satisfies the conditions imposed by the SA^k procedure applied to $A'x \leq b$.*

Proof. We will show that $\hat{x}'$ can be lifted to a vector $\hat{w}' \in \Re^L$, ($L == \sum_{j=0}^{k+1} 2^j \binom{n}{j}$) that satisfies conditions (i)–(iv) of the SA^k procedure applied to $A'x \leq b$.

This is done as follows. Let $\hat{w} \in \Re^M_+$ ($M = \sum_{j=0}^{k+1} 2^j \binom{n-1}{j}$) satisfy conditions (i)–(iv) of the SA^k procedure applied to $Ax \leq b$, with $\hat{w}[j, \emptyset] = \hat{x}_j$ for all j. In order to construct $\hat{w}'$, suppose Y', N' are disjoint subsets of the columns of A' with $|Y' \cup N'| \leq k$.

(i) If $|Y' \cap \{p, q\}| = 1$ and $|N' \cap \{p, q\}| = 1$ then we set $\hat{w}'[Y', N'] = 0$.

(ii) If $(Y' \cup N') \cap \{p, q\} = \emptyset$, we set $\hat{w}'[Y', N'] = \hat{w}[Y', N']$.

(iii) Finally, suppose that $Y' \cap \{p, q\} \neq \emptyset$ and $N' \cap \{p, q\} = \emptyset$. Then we set

$$\hat{w}'[Y', N'] = \hat{w}[\,(Y' - p - q) \cup (pq)\,,\, N'].$$

(and correspondingly when $N' \cap \{p, q\} \neq \emptyset$ and $Y' \cap \{p, q\} = \emptyset$).

By definition of identification, $\hat{w}'$ satisfies condition (iv) imposed by the SA^k procedure. The other conditions trivially hold. ■

3 Applications

In this section we examine Theorem 1.2 in the context of vertex packing problems, and show how it can be used to produce polynomial-size formulations whose feasible solutions are guaranteed to satisfy all inequalities of several types studied by other authors.

Given a graph G, we denote by $STAB(G)$ the stable set polytope for G, and by $SA^k(G)$ the formulation produced by the SA^k operator. Note that in this formulation, each variable $w[Y, N]$ is such that the sets Y and N correspond to disjoint subsets of the vertices. We denote by $\bar{SA}^k(G)$ the projection of $SA^k(G)$ to the space of the variables $w[j, \emptyset]$.

The *odd-hole* and *odd-wheel* inequality are among the first known inequalities for the stable set polytope. Note that it was shown in [LS91] that the points in the Lovász–Schrijver polytope $N(G)$, which equals $\bar{SA}^1(G)$, satisfy the odd-hole inequalities, and similarly it can be shown that the points in $\bar{SA}^2(G)$ satisfy the odd-wheel inequalities.

Cheng and Cunningham [CC97] consider several generalizations of odd-hole and odd-wheel constraints. All these inequalities are of the generic form

$$\sum_{j \in V(H)} \alpha_j x_j \quad \leq \quad \beta, \tag{19}$$

where H is a certain subgraph of G. The simplest case they consider is that where H is an appropriate subdivision of a wheel. A more complex case is that of a (again, possibly subdivided) p-wheel, which is similar to a wheel except that the hub is replaced with a p-clique, all of whose members are adjacent to all of the vertices on the rim of the wheel.

Cheng and de Vries [CdV02] first consider a more complex class of graphs, the t-antiwebs, defined as follows. Let n and t be integers such that $t \geq 2$, $n \geq 2t-1$ and $n \neq 0 \,(\text{mod } t)$. An (n,t)-antiweb is a graph with vertices $\{1, 2, \cdots, n\}$ and such that $\{i, j\}$ is an edge if and only if $\min\{i-j, n+j-i\} \leq t-1$. The antiwebs were introduced, in slightly different form, by Trotter [T75], and were also studied by other authors ([EJR87], [MS96], [S96]). [CdV02] also consider more complex graphs, the antiweb-wheels (similar to antiwebs, but with an additional vertex, the hub, which is adjacent to all other vertices). Finally, they also introduce another class of graphs, the t-antiweb-s-wheels. For integers n, t and s, an (n,t)-antiweb-s-wheel is the graph obtained by starting with an (n,t)-antiweb, and adding to it an s-clique using new vertices, all of which are adjacent to each vertex in the antiweb.

Two remarks about the above graphs. First, [CC97] and [CdV02] consider *subdivisions* of these graphs. For carefully constructed subdivisions, they obtain inequalities of the form (19) that can be shown to be facet-defining. These inequalities can also be separated in polynomial-time, under appropriate conditions: in the case of the t-antiwebs, t must be bounded, and in the case of the t-antiweb-s-wheels, both t and s must be bounded. The term "subdivision" is also dropped from their notation, e.g. when they refer to a t-antiweb inequality it is implicit that this refers to an appropriate subdivision. We will use this convention in what follows.

Second, in the language of [CdV02], the inequalities we described are *simple*. We will define this term below, where we will also extend our analysis to non-simple inequalities.

Of the above graphs, the antiweb-wheels are the most complex in the sense that each of the other graphs is a subgraph of an antiweb-wheel.

Lemma 3.1 *The tree-width of an t-antiweb-s-wheel is at most $2t + s - 2$.* ■

The proof of this result is routine – note that subdividing a graph does not change its tree-width.

Corollary 3.2 *Let $k \geq 2$. Then any point in $\bar{SA}^k(G)$ satisfies all simple odd-hole inequalities, wheel inequalities, t-antiweb inequalities (for $k \geq 2t-2$) and t-antiweb-s-wheel inequalities (for $k \geq 2t+s-2$).* ■

Next we consider the non-simple inequalities. These were considered in [CC97], [CdV02], where the following is proved:

Proposition 3.3 *Let G' be a graph, and suppose $\sum_{i=1}^{n} a_i x_i \leq b$ is valid for $STAB(G')$. Suppose vertices v_1 and v_2 are nonadjacent, and let G be the graph obtained from G' by identifying v_1 and v_2 so as to obtain a vertex $v_{1,2}$. Then $(a_1 + a_2)x_{1,2} + \sum_{i=3}^{n} a_i x_i \leq b$ is valid for $STAB(G)$.*

Note that this is a special case of Lemma 2.8. Applying this result repeatedly, one obtains inequalities of the form (19) valid for $STAB(G)$ where H is obtained from e.g. a t-antiweb-s-wheel by identifying nonadjacent vertices. Even if s and t are fixed values, the tree-width of H can be arbitrarily large. In the terminology of [CdV02], the resulting inequalities are called *nonsimple* while the standard wheel, etc., inequalities are termed *simple*.
Using Lemma 2.9 we have the following result:

Corollary 3.4 *Let $k \geq 2$. Then any point in $\bar{SA}^k(G)$ satisfies all simple and non-simple odd-hole inequalities, wheel inequalities, t-antiweb inequalities (for $k \geq 2t-2$) and t-antiweb-s-wheel inequalities (for $k \geq 2t+s-2$).* ■

Another set of interesting graphs is considered by Lipták and Tunçel [LL03]. They consider the strength of the N_0, N and N_+ operators of [LS91] when applied to the stable set problem, and they conjecture that the N_0- and N-rank of any graph are the same. For integer k odd, they consider the graph G_k whose vertex set is $\{1, 2, \cdots, 3k\}$, and which is the union of the cycle

$$(1,\, 4,\, 7, \cdots, 3k-2,\, 3k-1,\, 3k,\, 3k-3,\, 3k-6, \cdots, 3)$$

as well as a star with edges

$$\{2+3i,\, 1+3i\}, \{2+3i,\, 3+3i\}, \{2+3i,\, 4+3i\}$$

for each $0 \leq i \leq k-2$. In [LL03] it is shown that the N_0-rank of this graph equals $\lfloor \log_2 \frac{k+1}{3} + 2 \rfloor$. We have:

Proposition 3.5 *The tree-width of G_k is 3. Consequently, the SA-rank of G_k is at most 3.* ■

Thus, Proposition 3.5 provides an example where the N_0-rank of an inequality is arbitrarily larger than the SA-rank – we are not aware of an older example where the gap is unbounded. It is an interesting open question whether Lipták and Tunçel's conjecture is true.

Is Theorem 1.2 best possible? Consider the following (easy) result:

Proposition 3.6 *Consider a set P_A^b where A is 0/1 and b integral, and let $\alpha^T x \leq \beta$ be a valid inequality. Suppose the set $X \subseteq V(G_A[\alpha])$ is such that $G_A[\alpha] - X$ is bipartite. Then the SA rank of $\alpha^T x \leq \beta$ is at most $|X|$.* ■

In the case of a stable set problem, Proposition 3.6 can give a tighter bound than Theorem 1.2 – an example is that of an odd-hole inequality, which has SA rank 1 (and not 2, which is what Theorem 1.2 gives). On the other hand, the case of (n, t)-antiweb for $t > 2$ and n large is an example where Theorem 1.2 is much stronger than Proposition 3.6. It seems conceivable that in the case of A 0/1 and b integral, the bound provided by Theorem 1.2 could be improved by 1 unit.
Acknowledgment We thank anonymous referees whose suggestions helped us improve the presentation.

References

[ALS91] S. Arnborg, J. Lagergren and D. Seese, Easy problems for tree-decomposable graphs, *J. Algorithms* **2** (1991), 308-340.

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

[BZ02a] D. Bienstock and M. Zuckerberg, Subset algebra lift operators for 0-1 integer programming, to appear, *SIAM J. Optimization.*

[CC97] E. Cheng and W. H. Cunningham, Wheel inequalities for stable set polytopes, *Math. Programming* **77** (1997), 389–421.

[CdV02] E. Cheng and S. de Vries, Antiweb-wheel inequalities and their separation problems over the stable set polytopes, *Math. Programming* **92** (2002), 153–175.

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

[CS94] W.J. Cook and P.D. Seymour, An algorithm for the ring-router problem, Technical Report, Bellcore (1994).

[CS03] W.J. Cook and P.D. Seymour, Tour merging via branch decomposition, *INFORMS J. Computing* **21** (2003).

[CL02] G. Cornuéjols and Y. Li, On the rank of mixed 0-1 polyhedra, *Math. Programming* **92** (2002), 391-397.

[EJR87] R. Euler, M. Jünger and G. Reinelt, Generalizations of cliques, odd cycles and anticyles and their relation to independence system polyhedra, *Math. Oper. Res.* **12** (1987), 451–462.

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

[H04] I.V. Hicks, Branch Decompositions and Minor Containment, *Networks* **43** (2004) 1-9.

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

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

[LL03] L. Lipták and L. Tunçel, Stable set problems and graph ranks, *Math. Programming* **98** (2003), 319-353.

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

[MS96] R. Müller and A.S. Schulz, Transitive packing, In: Cunningham, McCormick, Queyranne, eds., Proc. 5th IPCO, *LNCS* **1084** (1996), 430-444.

[R97] B. Reed, Tree width and tangles: a new connectivity measure and some applications, in *London Math. Soc. Lecture Note* Ser. 241, 87-162, Cambridge Univ. Press, Cambridge, 1997.

[RS86] N. Robertson and P.D. Seymour, Graph Minors II: Algorithmic aspects of tree-width, *J. Algorithms* **7** (1986), 309-322.

[S96] A.S. Schulz, Polytopes and scheduling, PhD thesis, TU Berlin.

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

[T75] L.E. Trotter, A class of facet producing graphs for vertex packing polyhedra, *Disc. Math.* **12** (1975), 373–388.