Integrality Gaps of Linear and Semi-definite Programming Relaxations for Knapsack

Anna R. Karlin* Claire Mathieu† C. Thach Nguyen*

Abstract

Recent years have seen an explosion of interest in *lift and project* methods, such as those proposed by Lovász and Schrijver [40], Sherali and Adams [49], Balas, Ceria and Cornuejols [6], Lasserre [36, 37] and others. These methods are systematic procedures for constructing a sequence of increasingly tight mathematical programming relaxations for 0-1 optimization problems.

One major line of research in this area has focused on understanding the strengths and limitations of these procedures. Of particular interest to our community is the question of how the integrality gaps for interesting combinatorial optimization problems evolve through a series of rounds of one of these procedures. On the one hand, if the integrality gap of successive relaxations drops sufficiently fast, there is the potential for an improved approximation algorithm. On the other hand, if the integrality gap for a problem persists, this can be viewed as a lower bound in a certain restricted model of computation.

In this paper, we study the integrality gap in these hierarchies for the knapsack problem. We have two main results. First, we show that an integrality gap of $2-\epsilon$ persists up to a linear number of rounds of Sherali-Adams. This is interesting, since it is well known that knapsack has a fully polynomial time approximation scheme [30, 39]. Second, we show that Lasserre's hierarchy closes the gap quickly. Specifically, after t^2 rounds of Lasserre, the integrality gap decreases to $t/(t-1)$.

Thus, we provide a second example of an integrality gap separation between Lasserre and Sherali Adams. The only other such gap we are aware of is in the recent work of Fernandez de la Vega and Mathieu [19] (respectively of Charikar, Makarychev and Makarychev [12]) showing that the integrality gap for MAXCUT remains $2-\epsilon$ even after $\omega(1)$ (respectively n^γ) rounds of Sherali-Adams. On the other hand, it is known that 2 rounds of Lasserre yields a relaxation as least as strong as the Goemans-Williamson SDP, which has an integrality gap of 0.878.

1 Introduction

Many approximation algorithms work in two phases: first, solve a linear programming (LP) or semi-definite programming (SDP) relaxation; then, round the fractional solution to obtain a feasible integer solution to the original problem. Such an algorithm is analyzed by comparing the value of the output to that of the fractional solution, and thus cannot hope to yield a better approximation ratio than the integrality gap of the relaxation.

This paradigm is amazingly powerful; in particular, under the unique game conjecture, it yields the best possible ratio for Maxcut[1], and that ratio is exactly equal to the integrality gap of the Goemans-Williamson SDP [23, 18, 32, 33, 43]. However, for any given combinatorial optimization problem, in general, there are many possible LP/SDP relaxations, and so the integrality gap seems to say more about the particular relaxation chosen than about the intrinsic difficulty of the problem. Starting from an arbitrary LP relaxation, *lift and project* methods provide a systematic procedure for constructing a sequence of increasingly tight mathematical programming relaxations for 0-1 optimization problems. A number of different procedures of this type have been proposed: by Lovász and Schrijver [40], Sherali and Adams [49], Balas, Ceria and Cornuejols [6], Lasserre [36, 37] and others. While they differ in the details, they operate in a series

*University of Washington, email: {karlin,ncthach}@cs.washington.edu. Supported by NSF Grant CCF-0635147 and a Yahoo! Faculty Research Grant.

†Brown University, email: claire@cs.brown.edu. Supported by NSF Grant CCF-0728816.

[1]It also yields the best possible ratio for a wide variety of other problems, see e.g. [45].

of rounds starting from an LP or SDP relaxation and eventually ending with the exact integer polytope. The strengthened relaxation after t rounds can typically be solved in $n^{O(t)}$ time and, roughly, satisfies the property that the values of any t variables in the original relaxation can be expressed as the projection of a convex combination of integer solutions.

A major line of research in this area has focused on understanding the strengths and limitations of these procedures. Of particular interest to our community is the question of how the integrality gaps for interesting combinatorial optimization problems evolve through a series of rounds of one of these procedures. On the one hand, if the integrality gap of successive relaxations drops sufficiently fast, there is the potential for an improved approximation algorithm (see [15, 16, 8, 9] for example). On the other hand, if the integrality gap for a problem persists, this can be viewed as a lower bound to approximability in a certain restricted model of computation. Since the various lift-and-project schemes implicitly come up in most known sophisticated approximation algorithms for NP-hard problem (such as Sparsest Cut and Maximum Satisfiability), a large integrality gap after a linear (or even logarithmic) number of rounds rules out (unconditionally) a very wide class of efficient approximation algorithms. Indeed, several very strong negative results of this type have been obtained (see [4, 2, 10, 12, 20, 44, 47, 48, 50, 46] and others).

What do these negative results really mean? They might be further evidence of the intrinsic difficulty of the problems studied, which are all well-known hard problems (Maxcut, Sparsest cut, Vertex cover, Hypergraph vertex cover, etc.). Or ... perhaps they might be evidence that the lift and project models of computation are weak! To explore this latter possibility, we focus on one problem that is well-known to be "easy" from the viewpoint of approximability: In this paper, we study the integrality gap in these hierarchies for the knapsack problem. Our main results are the following:

- We show that an integrality gap close to 2 persists up to a linear number of rounds of Sherali-Adams. (The integrality gap of the standard LP is 2.)

  This is interesting since it is well known that knapsack has a fully polynomial time approximation scheme [30, 39]. This shows that integrality gap lower bounds for hierarchies such as Sherali-Adams do not necessarily imply an inherent bound on approximability. In other words, the Sherali-Adams restricted model of computation has serious weaknesses: a lower bound in that model does not necessarily imply that it is difficult to get a good approximation algorithm.

- We show that Lasserre's hierarchy closes the gap quickly. Specifically, after t^2 rounds of Laserre, the integrality gap decreases to $t/(t-1)$.

  Thus, we provide a second example of an integrality gap separation between Lasserre and Sherali Adams. The only other such gap we are aware of is in the recent work of Fernandez de la Vega and Mathieu [19] (respectively of Charikar, Makarychev and Makarychev [12]) showing that the integrality gap for MAXCUT remains $2-\epsilon$ even after $\omega(1)$ (respectively n^{γ}) rounds of Sherali-Adams. On the other hand, it is known that 2 rounds of Lasserre yields a relaxation as least as strong as the Goemans-Williamson SDP, which has an integrality gap of 0.878.

1.1 Related Work

Many known approximation algorithms can be recognized in hindsight as starting from a naive relaxation and strengthening it using a couple of levels of lift-and-project. The original hope [3] had been to use lift and project systems as a systematic approach to designing novel algorithms with better approximation ratios, but instead, the last few years have mostly seen the emergence of a multitude of lower bounds. Indeed, lift and project systems have been studied mostly for well known difficult problems: Maxcut [12, 19, 48], Sparsest cut, [12, 13] Vertex cover [2, 3, 4, 11, 20, 21, 28, 48, 50], Hypergraph vertex cover, TSP [14], Maximum acyclic subgraph [12], CSP [47, 51], and more.

The knapsack problem [41, 31] has a fully polynomial time approximation scheme [30, 39]. The natural LP relaxation (to be stated in full detail in the next section) has an integrality gap of $2-\epsilon$ [31]. Although we are not aware of previous work on using the lift and project systems for knapsack, the problem of strengthening the LP relaxation via addition of well-chosen inequalities has been much the object of much interest in the past in the mathematical programming community (stronger LP relaxations are extremely useful to speed up branch-and-bound heuristics.) The knapsack polytope was studied in detail by Weismantel [52]. Valid inequalities were studied in [5, 25, 26, 53, 7]. In particular, whenever S is a minimal set (w.r.to inclusion) that does not fit in the knapsack, then $\sum_{S\cup\{j:\forall i\in S, w_j\geq w_i\}} x_j \leq |S|-1$ is a valid inequality. Generalizations and variations were also studied in [17, 27, 54]. Thus, in spite of the existence of a dynamic program to solve the knapsack problem, the problem is fundamental enough that understanding the knapsack polytope (and its lifted tightenings) is of intrinsic interest.

Our results confirm the impression from [34, 46] for example that the Sherali-Adams lift and project is not powerful enough for the scope of our ambition. But so far, the results on integrality gaps in the Lasserre hierarchy are relatively few. The first negative results were about k-CSP [47, 51]. Our positive results leave open the possibility that the Lasserre hierarchy may have promise as a tool to capture the intrinsic difficulty of problems.

2 Preliminaries

2.1 The Knapsack problem

Our focus in this paper is on the Knapsack problem. In the Knapsack problem, we are given a set of n objects $V=[n]$ with sizes $c_1, c_2, \dots c_n$, values $v_1, v_2, \dots v_n$, and a capacity C. We assume that for every i, $c_i \leq C$. The objective is to select a subset of objects of maximum total value such that the total size of the objects selected does not exceed C.

The linear programming (LP) relaxation [31] for Knapsack is given by:

$$\begin{aligned} \max \quad & \sum_{i\in V} v_i x_i \\ \text{s.t.} \quad & \begin{cases} \sum_{i\in V} c_i x_i \leq C \\ 0 \leq x_i \leq 1 \qquad \forall i \in V \end{cases} \end{aligned} \tag{1}$$

The intended intepretation of an integral solution of this LP is obvious: $x_i = 1$ means the object i is selected, and $x_i = 0$ means it is not. The constraint can be written as $g(x) = C - \sum_i c_i x_i \geq 0$.

Let *Greedy* denote the algorithm that puts objects in the knapsack by order of decreasing ratio v_i/c_i, stopping as soon as the next object would exceed the capacity.

Lemma 1 ([?]) *Consider an instance (C,V) of* Knapsack *and its LP relaxation K given by (1). Then*

$$\textit{Value}(K) \leq \textit{Value}(\textit{Greedy}(C,V)) + \max_{i\in V} v_i.$$

2.2 The Sherali-Adams and Lasserre hierarchies

We next review the lift-and-project hierarchies that we will use in this paper. The descriptions we give here assume that the base program is linear and mostly use the notation given in the survey paper by Laurent [38]. To see that these hierarchies apply at a much greater level of generality we refer the reader to Laurent's paper [38].

Let K be a polytope defined by a set of linear constraints $g_1, g_2, \dots g_m$:

$$K = \{x \in [0,1]^n | g_\ell(x) \geq 0 \text{ for } \ell = 1, 2, \dots m\}. \tag{2}$$

We are interested in optimizing a linear objective function f over the convex hull $P = \mathsf{conv}\,(K \cap \{0,1\}^n)$ of integral points in K. Here, P is the set of convex combinations of all integral solutions of the given combinatorial problem and K is the set of solutions to its linear relaxation. For example, if K is defined by (1), then P is the set of convex combinations of valid integer solutions to Knapsack.

If all vertices of K are integral then $P = K$ and we are done. Otherwise, we would like to strengthen the relaxation K by adding additional valid constraints. The Sherali-Adams (SA) and Lasserre hierarchies are two different systematic ways to construct these additional constraints. In the SA hierarchy, all the constraints added are linear, whereas Lasserre's hierarchy is stronger and introduces a set of positive semi-definite constraints. However, for consistency, we will describe both hierarchies as requiring certain submatrices to be positive semi-definite (readers who are not familiar with the following formulation of SA are referred to Appendix A for a linear formulation of the hierarchy.)

To this end, we first state some notation. Throughout this paper we will use $\mathcal{P}(V)$ to denote the power set of V, and $\mathcal{P}_t(V)$ to denote the collection of all subsets of V whose sizes are at most t. Also, given two sets of coordinates T and S, $T \subseteq S$ and $y \in R^S$, by $y|_T$ we denote the projection of y onto T.

Next, we review the definition of the *shift operator* between two vectors $x, y \in R^{\mathcal{P}(V)}$: $x * y$ is a vector in $R^{\mathcal{P}(V)}$ such that

$$(x * y)_I = \sum_{J \subseteq V} x_J y_{I \cup J}.$$

Lemma 2 ([38]) *The shift operator is commutative: for any vectors $x, y, z \in R^{\mathcal{P}(V)}$, we have $x * (y * z) = y * (x * z)$.*

A polynomial $P(x) = \sum_{I \subseteq V} a_I \prod_{i \in I} x_i$ can also be viewed as a vector indexed by subsets of V. We define the vector $P * y$ accordingly: $(P * y)_I = \sum_{J \subseteq V} a_J y_{I \cup J}$.

Finally, let $\mathcal{T}$ be a collection of subsets of V and y be a vector in $R^{\mathcal{T}}$. We denote by $M_{\mathcal{T}}(y)$ the matrix whose rows and colums are indexed by elements of $\mathcal{T}$ such that

$$(M_{\mathcal{T}}(y))_{I,J} = y_{I \cup J}.$$

The main observation is that if $x \in K \cap \{0,1\}^n$ then $(y_I) = (\prod_{i \in I} x_i)$ satisfies $M_{\mathcal{P}(V)}(y) = yy^T \succeq 0$ and $M_{\mathcal{P}(V)}(g_\ell * y) = g_\ell(x) yy^T \succeq 0$ for all constraints g_ℓ. Thus requiring principal submatrices of these two matrices to be positive semi-definite yields a relaxation.

Definition 3 *For any $1 \leq t \leq n$, the t-th* Sherali-Adams lifted polytope $\mathsf{SA}^t(K)$ *is the set of vectors $y \in [0,1]^{\mathcal{P}_t(V)}$ such that $y_\emptyset = 1$, $M_{\mathcal{P}(U)}(y) \succeq 0$ and $M_{\mathcal{P}(W)}(g_\ell * y) \succeq 0$ for all ℓ and subsets $U, W \subseteq V$ such that $|U| \leq t$ and $|W| \leq t-1$.*

We say that a point $x \in [0,1]^n$ belongs to the t-th Sherali-Adams polytope $\mathsf{sa}^t(K)$ *iff there exists a $y \in \mathsf{SA}^t(K)$ such that $y_{\{i\}} = x_i$ for all $i \in [n]$.*

Definition 4 *For any $1 \leq t \leq n$, the t-th* Lasserre lifted polytope $\mathsf{La}^t(K)$ *is the set of vectors $y \in [0,1]^{\mathcal{P}_{2t}(V)}$ such that $y_\emptyset = 1$, $M_{\mathcal{P}_t(V)}(y) \succeq 0$ and $M_{\mathcal{P}_{t-1}(V)}(g_\ell * y) \succeq 0$ for all ℓ.*

We say that a point $x \in [0,1]^n$ belongs to the t-th Lasserre polytope $\mathsf{la}^t(K)$ *if there exists a $y \in \mathsf{La}^t(K)$ such that $y_{\{i\}} = x_i$ for all $i \in V$.*

Note that $M_{\mathcal{P}(U)}(y)$ has at most 2^t rows and columns, which is constant for t constant, whereas $M_{\mathcal{P}_t(V)}(y)$ has $\binom{n+1}{t+1}$ rows and columns.

It is immediate from the definitions that $\mathsf{sa}^{t+1}(K) \subseteq \mathsf{sa}^t(K)$, and $\mathsf{la}^{t+1}(K) \subseteq \mathsf{la}^t(K)$ for all $1 \leq t \leq n-1$. In addition, $\mathsf{la}^t(K) \subseteq \mathsf{sa}^t(K)$. Sherali and Adams [49] show that $\mathsf{sa}^n(K) = P$, Thus, the sequences

$$K \supseteq \mathsf{sa}^1(K) \supseteq \mathsf{sa}^2(K) \supseteq \cdots \supseteq \mathsf{sa}^n(K) = P$$
$$K \supseteq \mathsf{la}^1(K) \supseteq \mathsf{la}^2(K) \supseteq \cdots \supseteq \mathsf{la}^n(K) = P$$

define hierarchies of polytopes that converge to P. Furthermore, the Lasserre hierarchy is "stronger" than the Sherali-Adams hierarchy. In this paper, we show that for the Knapsack problem, the Lasserre hierarchy is strictly stronger.

2.3 The integrality gap of the hierarchies

We will be studying the integrality gaps of the lifted polytopes for Knapsack. For a given Knapsack instance K, a level t and a particular hierarchy H (where H is either Sherali-Adams or Lasserre), define $H_t(K)$ to be the ratio between the optimal solution in the t-th lifted polytope (the maximum value of $\sum_i v_i y_{\{i\}}$, for y in the t-th lifted polytope on instance K) and the optimal value of the objective function for a feasible integral solution. The *integrality gap at the t-th level of the hierarchy H* is $\sup_K H_t(K)$.

3 Lower bound for the Sherali-Adams hierarchy for Knapsack

In this section, we show that the integrality gap of the t-th level of the Sherali-Adams hierarchy for $\mathsf{Knapsack}$ is 2. This lower bound even holds for the *uniform* $\mathsf{Knapsack}$ problem, in which $v_i = c_i = 1$ for all i. [2].

Theorem 5 *For every t, the integrality gap at the t-th level of the Sherali-Adams hierarchy for* $\mathsf{Knapsack}$ *is equal to 2.*

Proof. Let $t \geq 2$. Consider the instance K of $\mathsf{Knapsack}$ with n objects where $c_i = v_i = 1$ for all $i \in V$ and capacity $C = 2(1-\epsilon)$. Let $\alpha = C/(n + (t-1)(1-\epsilon))$ and consider the vector $y \in [0,1]^{\mathcal{P}_t(V)}$ defined by

$$\begin{cases} y_\emptyset = 1 \\ y_{\{i\}} = \alpha \\ y_I = 0 \ \text{ if } |I| > 1 \end{cases}$$

We claim that $y \in \mathsf{SA}^t(K)$. Consider any subset $U \subseteq V$ such that $|U| \leq t$. We have

$$M_{\mathcal{P}(U)}(y) = \begin{pmatrix} M_{\mathcal{P}_1(U)}(y) & 0 \\ 0 & 0 \end{pmatrix}, \qquad \text{with} \quad M_{\mathcal{P}_1(U)}(y) = \begin{pmatrix} 1 & \alpha & \alpha & \cdots & \alpha \\ \alpha & \alpha & 0 & \cdots & 0 \\ \alpha & 0 & \alpha & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha & 0 & 0 & \cdots & \alpha \end{pmatrix}.$$

Since $|U| \leq t < n$, $|U|\alpha \leq 1$, and it is easy to see that this implies $M_{\mathcal{P}_1(U)}(y) \succeq 0$, and so $M_{\mathcal{P}(U)}(y) \succeq 0$.

Next, let $g(x) = C - \sum_{i \in V} c_i x_i$ and consider any subset $W \subseteq V$ such that $|W| \leq t-1$. Again, we have

$$M_{\mathcal{P}(W)}(g*y) = \begin{pmatrix} M_{\mathcal{P}_1(W)}(g*y) & 0 \\ 0 & 0 \end{pmatrix}, \; M_{\mathcal{P}_1(W)}(g*y) = \begin{pmatrix} C - n\alpha & (C-1)\alpha & (C-1)\alpha & \cdots & (C-1)\alpha \\ (C-1)\alpha & (C-1)\alpha & 0 & \cdots & 0 \\ (C-1)\alpha & 0 & (C-1)\alpha & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (C-1)\alpha & 0 & 0 & \cdots & (C-1)\alpha \end{pmatrix}.$$

Since $|W| \leq t-1$, by definition of α we have $|W|(C-1)\alpha \leq C - n\alpha$, and it is easy to see that this implies $M_{\mathcal{P}_1(W)}(g*y) \succeq 0$, and so $M_{\mathcal{P}(W)}(g*y) \succeq 0$. Thus $y \in \mathsf{SA}^t(K)$.

The integer optimum has value 1, so the integrality gap is at least the value of y, which is $n\alpha = 2(1-\epsilon)/(1 + (t-1)(1-2\epsilon)/n)$. The supremum over all ϵ is $2/(1+(t-1)/n)$, and the supremum of that over all n is 2, so the integrality gap is at least 2.

On the other hand, it is well-known that the base linear program K has value at most $2OPT$ (that is an immediate consequence of Lemma 1), hence, by the nesting property, every linear program in the hierarchy has integrality gap exactly equal to 2. ■

4 A decomposition theorem for the Lasserre hierarchy

In this section, we develop the machinery we will need for our Lasserre upper bounds. It turns out that it is more convenient to work with families (z^X) of characteristic vectors rather than directly with y. We begin with some definitions and basic properties.

Definition 6 (extension) *Let $\mathcal{T}$ be a collection of subsets of V and let y be a vector indexed by sets of $\mathcal{T}$. We define the* extension *of y to be the vector y', indexed by all subsets of V, such that y'_I equals y_I if $I \in \mathcal{T}$ and equals 0 otherwise.*

[2] Some people call this problem *Unweighted Knapsack* or *Subset Sum*.

Definition 7 (characteristic polynomial) *Let S be a subset of V and X a subset of S. We define the* characteristic polynomial P^X *of X with respect to S as*

$$P^X(x) = \prod_{i \in X} x_i \prod_{j \in S \setminus X} (1 - x_j) = \sum_{J : X \subseteq J \subseteq S} (-1)^{|J \setminus X|} \prod_{i \in J} x_i.$$

Lemma 8 (inversion formula) *Let y' be a vector indexed by all subsets of V. Let S be a subset of V and, for each X subset of S, let $z^X = P^X * y'$:*

$$z_I^X = \sum_{J : X \subseteq J \subseteq S} (-1)^{|J \setminus X|} y'_{I \cup J}.$$

Then $y' = \sum_{X \subseteq S} z^X$.

Proof. Fix a subset I of V. Substituting the definition of z_I^X in $\sum_{X \subseteq S} z_I^X$, and changing the index of summation, we get

$$\sum_{X \subseteq S} z_I^X = \sum_{A \subseteq S} \sum_{J \subseteq A} (-1)^{|J|} \; y'_{I \cup A}.$$

For $A \neq \emptyset$ the inner sum is 0, so only the term for $A = \emptyset$, which equals y'_I, remains. ■

Lemma 9 *Let y' be a vector indexed by all subsets of V, S be a subset of V and X be a subset of S. Then*

$$\begin{cases} z_I^X = z_{I \setminus X}^X & \text{for all } I \\ z_I^X = z_\emptyset^X & \text{if } I \subseteq X \\ z_I^X = 0 & \text{if } I \cap (S \setminus X) \neq \emptyset \end{cases}$$

Proof. Let $I' = I \setminus X$ and $I'' = I \cap X$. Using the definition of z_I^X and noticing that $X \cup I'' = X$ yields $z_I^X = z_{I'}^X$. This immediately implies that for $I \subseteq X$, $z_I^X = z_\emptyset^X$.

Finally, consider a set I that intersects $S \setminus X$ and let $i \in I \cap (S \setminus X)$. In the definition of z_I^X, we group the terms of the sum into pairs consisting of J such that $i \notin J$ and of $J \cup \{i\}$. Since $I = I \cup \{i\}$, we obtain:

$$\sum_{J : X \subseteq J \subseteq S} (-1)^{|J \setminus X|} y'_{I \cup J} = \sum_{J : X \subseteq J \subseteq S \setminus \{i\}} \left((-1)^{|J \setminus X|} + (-1)^{|J \setminus X| + 1} \right) y'_{I \cup J} = 0.$$

■

Corollary 10 *Let y' be a vector indexed by all subsets of V, S be a subset of V and X be a subset of S. Let w^X be defined as $z^X / z_\emptyset^X$ if $z_\emptyset^X \neq 0$ and defined as 0 otherwise. Then, if $z_\emptyset^X \neq 0$, then $w_{\{i\}}^X$ equals 1 for elements of X and 0 for elements of $S \setminus X$.*

Definition 11 (closed under shifting) *Let S be an arbitrary subset of V and $\mathcal{T}$ be a collection of subsets of V. We say that $\mathcal{T}$ is* closed under shifting by S *if*

$$Y \in \mathcal{T} \implies \forall X \subseteq S, \quad X \cup Y \in \mathcal{T}.$$

The following lemma generalizes Lemma 5 in [38]. It proves that the positive-semidefinite property carries over from y to (z^X).

Lemma 12 *Let S be an arbitrary subset of V and $\mathcal{T}$ be a collection of subsets of V that is closed under shifting by S. Let y be a vector indexed by sets of $\mathcal{T}$. Then*

$$M_{\mathcal{T}}(y) \succeq 0 \implies \forall X \subseteq S, \quad M_{\mathcal{T}}(z^X) \succeq 0.$$

Proof. Since $M_{\mathcal{T}}(y)\succeq 0$, there exist vectors v_I, $I \in \mathcal{T}$, such that $\langle v_I, v_J\rangle = y_{I\cup J}$. Fix a subset X of S. For each $I \in \mathcal{T}$, let

$$w_I = \sum_{H\subseteq S\setminus X} (-1)^{|H|} v_{I\cup X\cup H},$$

which is well-defined since $\mathcal{T}$ is closed under shifting by S.

Let $I, J \in \mathcal{T}$. It is easy to check that $\langle w_I, w_J\rangle = (z^X)_{I\cup J}$. Indeed,

$$\langle w_I, w_J\rangle = \sum_{H\subseteq S\setminus X}\sum_{L\subseteq S\setminus X} (-1)^{|H|+|L|}\langle v_{I\cup X\cup H}, v_{J\cup X\cup L}\rangle \tag{3}$$

$$= \sum_{H\subseteq S\setminus X}\sum_{L\subseteq S\setminus X} (-1)^{|H|+|L|} y_{I\cup J\cup X\cup H\cup L} \tag{4}$$

by definition of v_I, v_J and since $\mathcal{T}$ is closed under shifting by S (so that this is well-defined). Consider a non-empty subset H of $S \setminus X$ and let $i \in H$. We group the terms of the inner sum into pairs consisting of L such that $i \notin L$ and of $L \cup \{i\}$. Since $H = H \cup \{i\}$, we obtain:

$$\sum_{L\subseteq S\setminus X} (-1)^{|H|+|L|} y_{I\cup J\cup X\cup H\cup L} = \sum_{L\subseteq (S\setminus X\setminus\{i\})} \left((-1)^{|H|+|L|} + (-1)^{|H|+|L|+1}\right) y_{I\cup J\cup X\cup H\cup L} = 0.$$

Thus, the expression in (4) becomes

$$\langle w_I, w_J\rangle = \sum_{L\subseteq S\setminus X} (-1)^{|L|} y_{I\cup J\cup X\cup L} = (z^X)_{I\cup J}.$$

This implies that $M_{\mathcal{T}}(z^X)\succeq 0$. ∎

In the rest of the section, we prove a decomposition theorem for the Lasserre hierarchy, which allows us to "divide" the action of the hierarchy and think of it as using the first few rounds on some subset of variables, and the other rounds on the rest. We will use this theorem to prove that the Lasserre hierarchy closes the gap for the Knapsack problem in the next section.

Theorem 13 *Let $t > 1$ and $y \in \mathsf{La}^t(K)$. Let $k < t$ and S be a subset of V and such that*

$$|I\cap S| \geq k \implies y_I = 0. \tag{5}$$

Consider the projection $y|_{\mathcal{P}_{2t-2k}(V)}$ of y to the coordinates corresponding to subsets of size at most $2t-2k$ of V. Then there exist subsets $X_1, X_2, \dots, X_m$ of S such that $y|_{\mathcal{P}_{2t-2k}(V)}$ is a convex combination of vectors w^{X_i} with the following properties:

- $w^{X_i}_{\{j\}} = \begin{cases} 1 & \text{if } j \in X_i \\ 0 & \text{if } j \in S \setminus X_i; \end{cases}$
- *$w^{X_i} \in \mathsf{La}^{t-k}(K)$; and*
- *if K_i is obtained from K by setting $x_j = w^{X_i}_{\{j\}}$ for $j \in S$, then $w^{X_i}|_{\mathcal{P}_{2t-2k}(V\setminus S)} \in \mathsf{La}^{t-k}(K_i)$.*

To prove Theorem 13, we will need a couple more lemmas. In the first one, using assumption (5), we extend the positive semi-definite properties from y to y', and then, using Lemma 12, from y' to z^X.

Lemma 14 *Let t, y, S, k be defined as in Theorem 13, and y' be the extension of y. Let $\mathcal{T}_1 = \{A$ such that $|A\setminus S| \leq t-k\}$, and $\mathcal{T}_2 = \{B$ such that $|B\setminus S| < t-k\}$. Then for all $X \subseteq S$, $M_{\mathcal{T}_1}(z^X)\succeq 0$ and, for all ℓ, $M_{\mathcal{T}_2}(g_\ell * z^X)\succeq 0$.*

Proof. We will first prove that $M_{\mathcal{T}_1}(y')\succeq 0$ and, for all ℓ, $M_{\mathcal{T}_2}(g_\ell * y')\succeq 0$. Order the columns and rows of $M_{\mathcal{T}_1}(y')$ by subsets of non-decreasing size. By definition of $\mathcal{T}_1$, any $I \in \mathcal{T}_1$ of size at least t must have $|I\cap S| \geq k$, and so $y'_I = 0$. Thus

$$M_{\mathcal{T}_1}(y') = \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix},$$

where M is a principal submatrix of $M_{\mathcal{P}_t(V)}(y)$.Thus $M \succeq 0$, and so $M_{\mathcal{T}_1}(y') \succeq 0$.

Similarly, any $J \in \mathcal{T}_2$ of size at least $t-1$ must have $|J \cup \{i\} \cap S| \geq k$ for every i as well as $|J \cap S| \geq k$, and so, by definition of $g_\ell * y'$ we must have $(g_\ell * y')_J = 0$. Thus

$$M_{\mathcal{T}_2}(g_\ell * y') = \begin{pmatrix} N & 0 \\ 0 & 0 \end{pmatrix},$$

where N is a principal submatrix of $M_{\mathcal{P}_{t-1}(V)}(g_\ell * y)$. Thus $N \succeq 0$, and so $M_{\mathcal{T}_2}(g_\ell * y') \succeq 0$.

Observe that $\mathcal{T}_1$ is closed under shifting by S. By definition of z^X and Lemma 12, we thus get $M_{\mathcal{T}_1}(z^X) \succeq 0$.

Similarly, observe that $\mathcal{T}_2$ is also closed under shifting by S. By Lemma 2, we have $g_\ell * (P^X * y') = P^X * (g_\ell * y')$, and so by Lemma 12 again we get $M_{\mathcal{T}_2}(g_\ell * z^X) \succeq 0$. ∎

Lemma 15 *Let t, y, S, k be defined as in Theorem 13, and y' be the extension of y. Then for any $X \subseteq S$:*

1. *$z^X_\emptyset \geq 0$.*
2. *If $z^X_\emptyset = 0$ then $z^X_I = 0$ for all $|I| \leq 2t - 2k$.*

Proof. Let $\mathcal{T}_1$ be defined as in Lemma 14. By Lemma 14 $M_{\mathcal{T}_1}(z^X) \succeq 0$ and $z^X_\emptyset$ is a diagonal element of this matrix, hence $z^X_\emptyset \geq 0$.

For the second part, start by considering $J \subseteq V$ of size at most $t-k$. Then $J \in \mathcal{T}_1$, and so the matrix $M_{\{\emptyset, J\}}(z^X)$ is a principal submatrix of $M_{\mathcal{T}_1}(z^X)$, hence is also positive semidefinite. Since $z^X_\emptyset = 0$,

$$M_{\{\emptyset, J\}}(z^X) = \begin{pmatrix} 0 & z^X_J \\ z^X_J & z^X_J \end{pmatrix} \succeq 0,$$

hence $z^X_J = 0$.

Now consider any $I \subseteq V$ such that $|I| \leq 2t - 2k$, and write $I = I_1 \cup I_2$ where $|I_1| \leq t-k$ and $|I_2| \leq t-k$. $M_{\{I_1, I_2\}}(z^X)$ is a principal submatrix of $M_{\mathcal{T}_1}(z^X)$, hence is also positive semidefinite. Since $z^X_{I_1} = z^X_{I_2} = 0$, Since

$$M_{\{I_1, I_2\}}(z^X) = \begin{pmatrix} 0 & z^X_I \\ z^X_I & 0 \end{pmatrix} \succeq 0,$$

hence $z^X_I = 0$. ∎

We now have what we need to prove Theorem 13.

Proof of Theorem 13. By definition, Lemma 8 and the second part of Lemma 15, we have

$$y|_{\mathcal{P}_{2t-2k}(V)} = y'|_{\mathcal{P}_{2t-2k}(V)} = \sum_{X \subseteq S} z^X|_{\mathcal{P}_{2t-2k}(V)} = \sum_{X \subseteq S} z^X_\emptyset w^X|_{\mathcal{P}_{2t-2k}(V)}.$$

By Lemma 8 and by definition of y, we have $\sum_{X \subseteq S} z^X_\emptyset = y_\emptyset = 1$, and the terms are non-negative by the first part of Lemma 15, so $y|_{\mathcal{P}_{2t-2k}(V)}$ is a convex combination of w^X's, as desired.

Consider $X \subseteq S$ such that $z^X_\emptyset \neq 0$. By Lemma 14, $M_{\mathcal{T}_1}(z^X) \succeq 0$ and $M_{\mathcal{T}_2}(g_\ell * z^X) \succeq 0$ for all ℓ, and so this also holds for their principal submatrices $M_{\mathcal{P}_{t-k}(V)}(z^X)$ and $M_{\mathcal{P}_{t-k-1}(V)}(g_\ell * z^X)$. Scaling by the positive quantity $z^X_\emptyset$, by definition of w^X this also holds for $M_{\mathcal{P}_{t-k}(V)}(w^X)$ and $M_{\mathcal{P}_{t-k-1}(V)}(g_\ell * w^X)$. In other words, $w^X|_{\mathcal{P}_{2t-2k}(V)} \in \mathsf{La}^{t-k}(K)$.

Since $M_{\mathcal{P}_{t-k}(V)}(w^{X_i}) \succeq 0$, by taking a principal submatrix, we infer $M_{\mathcal{P}_{t-k}(V \setminus S)}(w^{X_i}) \succeq 0$. Similarly, $M_{\mathcal{P}_{t-k}(V)}(g_\ell * w^{X_i}) \succeq 0$ and so $M_{\mathcal{P}_{t-k}(V \setminus S)}(g_\ell * w^{X_i}) \succeq 0$. Let g'_ℓ be the constraint of K_i obtained from g_ℓ by setting $x_j = w^{X_i}_{\{j\}}$ for all $j \in S$. We claim that for any $I \subseteq V \setminus S$, $(g'_\ell * z^{X_i})_I = (g_\ell * z^{X_i})_I$; scaling implies that $M_{\mathcal{P}_{t-k}(V \setminus S)}(g'_\ell * w^{X_i}) = M_{\mathcal{P}_{t-k}(V \setminus S)}(g_\ell * w^{X_i})$ and we are done.

To prove the claim, let $g_\ell(x) = \sum_{j \in V} a_j x_j + b$. Then, by Corollary 10, $g'_\ell = \sum_{j \in V \setminus S} a_j x_j + (b + \sum_{j \in X_i} a_j)$. Let $I \subseteq V \setminus S$. We see that

$$(g_\ell * w^{X_i})_I - (g'_\ell * w^{X_i})_I = \sum_{j \in X_i} a_j w^{X_i}_{I \cup \{j\}} + \sum_{j \in S \setminus X_i} a_j w^{X_i}_{I \cup \{J\}} - \sum_{j \in X_i} a_j w^{X_i}_I.$$

By Lemma 9, $w^{X_i}_{I \cup \{j\}} = w^{X_i}_I$ for $j \in X_i$ and $w^{X_i}_{I \cup \{j\}} = 0$ for $j \in S \setminus X_i$. The claim follows. ∎

5 Upper bound for the Lasserre hierarchy for Knapsack

In this section, we use Theorem 13 to prove that for the Knapsack problem the gap of $\mathsf{La}^t(K)$ approaches 1 quickly as t grows, where K is the LP relaxation of (1). First, we show that there is a set S such that every feasible solution in $\mathsf{La}^t(K)$ satisfies the condition of the Theorem.

Given an instance (C, V) of Knapsack, Let $OPT(C, V)$ denote the value of the optimal integral solution.

Lemma 16 *Consider an instance (C, V) of Knapsack and its linear programming relaxation K given by (1). Let $t > 1$ and $y \in \mathsf{La}^t(K)$. Let $k < t$ and $S = \{i \in V | v_i > OPT(C, V)/k\}$. Then:*

$$\sum_{i \in I \cap S} c_i > C \implies y_I = 0.$$

Proof. There are three cases depending on the size of I:

1. $|I| \leq t-1$. Recall the capacity constraint $g(x) = C - \sum_{i \in V} c_i x_i \geq 0$. On the one hand, since $M_{\mathcal{P}_{t-1}(V)}(g * y) \succeq 0$, the diagonal entry $(g*y)_I$ must be non-negative. On the other hand, writing out the definition of $(g * y)_I$ and noting that the coefficients c_i are all non-negative, we infer $(g * y)_I \leq Cy_I - \left(\sum_{i \in I} c_i\right) y_I$. But by assumption, $\sum_{i \in I} c_i > C$. Thus we must have $y_I = 0$.

2. $t \leq |I| \leq 2t-2$. Write $I = I_1 \cup I_2 = I$ with $|I_1|, |I_2| \leq t-1$ and $|I_1 \cap S| \geq k$. Then $y_{I_1} = 0$. Since $M_{\mathcal{P}_t(y)} \succeq 0$, its 2-by-2 principal submatrix $M_{\{I_1, I_2\}}(y)$ must also be positive semi-definite.

$$M_{\{I_1, I_2\}}(y) = \begin{pmatrix} 0 & y_I \\ y_I & y_{i_1} \end{pmatrix},$$

   and it is easy to check that we must then have $y_I = 0$.

3. $2t-1 \leq |I| \leq 2t$. Write $I = I_1 \cup I_2 = I$ with $|I_1|, |I_2| \leq t$ and $|I_1 \cap S| \geq k$. Then $y_{I_1} = 0$ since $t \leq 2t-2$ for all $t \geq 2$. By the same argument as in the previous case, we must then have $y_I = 0$.

■

The following theorem shows that the integrality gap of the t^{th} level of the Lasserre hierarchy for Knapsack reduces quickly when t increases.

Theorem 17 *Consider an instance (C, V) of Knapsack and its LP relaxation K given by (1). Let $t \geq 2$. Then*

$$Value(\mathsf{La}^t(K)) \leq (1 + \frac{1}{t-1}) OPT,$$

and so the integrality gap at the t-th level of the Lasserre hierarchy is at most $1 + 1/(t-1)$.

Proof. Let $S = \{i \in V | v_i > OPT(C, V)/(t-1)\}$. Let $y \in \mathsf{La}^t(K)$. If $|I \cap S| \geq t-1$, then the elements of $I \cap S$ have total value greater than $OPT(C, V)$, so they must not be able to fit in the knapsack: their total capacity exceeds C, and so by Lemma 16 we have $y_I = 0$. Thus the condition of Theorem 13 holds for $k = t-1$.

Therefore, $y|_{\mathcal{P}_2(V)}$ is a convex combination of w^{X_i} with $X_i \subseteq S$, thus $\text{Value}(y) \leq \max_i \text{Value}(w^{X_i})$. By the first and third properties of the Theorem, we have:

$$\text{Value}(w^{X_i}) \leq \sum_{j \in X_i} v_j + \text{Value}(\mathsf{La}^1(K_i)).$$

By the nesting property of the Lasserre hierarchy, Lemma 1, and definition of S,

$$\text{Value}(\mathsf{La}^1(K_i)) \leq \text{Value}(K_i) \leq OPT(C - \text{Cost}(X_i), V \setminus S)) + OPT(C, V)/(t-1).$$

By the second property of the Theorem, w^{X_i} is in $\mathsf{La}^{t-k}(K) \subseteq K$, so it must satisfy the capacity constraint, so $\sum_{i \in X_i} c_i \leq \sum_{i \in I} c_i \leq C$, so X_i is feasible. Thus:

$$\text{Value}(y) \quad \leq \quad \max_{\text{feasible } X \subseteq S} \left(\sum_{j \in X} v_j + OPT(C - \text{Cost}(X), V \setminus S) \right) + OPT(C, V)/(t-1)$$

The first expression in the right hand side is equal to $OPT(C, V)$, hence the Theorem. ■

Remark 18 *We note that the analysis above suggests a new PTAS (albeit an inefficient one). For the straightforward details, see Appendix B.*

6 Conclusion

We have shown that for the Knapsack problem, an integrality gap of $2 - \epsilon$ persists up to a linear number of rounds of Sherali-Adams, whereas Lasserre's hierarchy closes the gap quickly. (In fact, we have observed that large gaps persist under Sherali-Adams for a number of other "easy" problems as well, including minimum and maximum spanning trees and a scheduling problem, among others.)

The obvious conclusion is: Sherali-Adams is weak (hence good algorithms based on Sherali-Adams are surprising, negative results are unsurprising). Lasserre might be strong, hence negative results in the Lasserre hierarchy might have meaningful implications on the intrinsic difficulty of the problem under study.

References

[1] N.E. Aguilera, S.M. Bianchi and G.L. Nasini. Lift and project relaxations for the matching and related polytopes. *Discrete Applied Mathematics* **134** (2004), pp. 193–212.

[2] M. Alekhnovich, S. Arora and I. Tourlakis. Towards strong non-approximability results in the Lovász-Schrijver hierarchy. *Proceedings of the 37th ACM Symposium on Theory of Computing*, 2005, pp. 294–303.

[3] S. Arora, B. Bollobás and L. Lovász. Proving integrality gaps without knowing the linear program. *Proceedings of the 43rd ACM Symposium on Theory of Computing*, 2002, pp. 313–322.

[4] S. Arora, B. Bollobás, L. Lovász and I. Tourlakis. Proving integrality gaps without knowing the linear program. *Theory of Computing* **2** (2006), pp. 19–51.

[5] E. Balas Facets of the Knapsack Polytope. (1998)

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

[7] E. Balas and E. Zemel, "Facets of the Knapsack Polytope from Minimal Covers," SIAM Journal on Applied Mathematics, 34 (1978), 119-148.

[8] MohammadHossein Bateni, Moses Charikar and Venkatesan Guruswami. MaxMin allocation via degree lower-bounded arborescences. ACM STOC, 2009.

[9] Daniel Bienstock and Nuri Ozbay Tree-width and the SheraliAdams operator. Discrete Optimization Volume 1, Issue 1, 15 June 2004, Pages 13-21.

[10] J. Buresh-Oppenheim, N. Galesi, S. Hoory, A. Magen and T. Pitassi. Rank bounds and integrality gaps for cutting plane procedures. *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science*, 2003, pp. 318–327.

[11] M. Charikar. "On semidefinite programming relaxations for graph coloring and vertex cover." *Proceedings of the 13th ACM-SIAM Symposium on Discrete Algorithms*, 2002, pp. 616–620.

[12] M. Charikar, K. Makarychev and Y. Makarychev. "Integrality gaps for Sherali-Adams relaxations." ACM STOC, 2009.

[13] Jeff Cheeger, Bruce Kleiner and Assaf Naor. A $(\log n)^{\Omega(1)}$ integrality gap for the Sparsest Cut SDP. IEEE FOCS, 2009.

[14] Kevin K. H. Cheung. On Lovsz-Schrijver Lift-and-Project Procedures on the Dantzig-Fulkerson-Johnson Relaxation of the TSP. SIAM Journal on Optimization, Volume 16 , Issue 2 (2005), pp. 380 - 399.

[15] E. Chlamtac. Approximation algorithms using hierarchies of semidefinite programming relaxations. *Proceedings of the 48th IEEE Symposium on Foundations of Computer Science*, 2007, pp. 691–701.

[16] E. Chlamtac and G. Singh. Improved Approximation Guarantees through Higher Levels of SDP Hierarchies. Proceedings of the 11th International Workshop, APPROX 2008, pp. 4962.

[17] Laureano F. Escudero, Araceli Garn, Gloria Prez An O(n log n) procedure for identifying facets of the knapsack polytope. Oper. Res. Lett. 31(3): 211-218 (2003)

[18] Uriel Feige and Gideon Schechtman On the optimality of the random hyperplane rounding technique for MAX CUT (2000)

[19] W. Fernandez de la Vega and C. Kenyon-Mathieu. Linear programming relaxations of Maxcut. *Proceedings of the 18th ACM-SIAM Symposium on Discrete Algorithms*, 2007.

[20] K. Georgiou, A. Magen, T. Pitassi and I. Tourlakis. Integrality gaps of $2 - o(1)$ for vertex cover SDPs in the Lovász-Schrijver hierarchy. *Proceedings of the 48th IEEE Symposium on Foundations of Computer Science*, 2007.

[21] K. Georgiou, A. Magen and I. Tourlakis. Vertex Cover Resists SDPs Tightened by Local Hypermetric Inequalities. 13th Conference on Integer Programming and Combinatorial Optimization (IPCO 2008).

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

[23] M.X. Goemans and D.P. Williamson. "Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming," J. ACM, 42, 1115–1145, 1995.

[24] D. Grigoriev, E. Hirsch and D. Pasechnik. "Complexity of semialgebraic proofs." *Moscow Mathematical Journal* **2** (2002), pp. 647–679.

[25] P. L. Hammer, E. L. Johnson and Uri N. Peled, Facets of regular 0-1 polytopes, Math. Programming, 8, 1975, 179–206.

[26] P. L. Hammer and Uri N. Peled, Computing low capacity 0-1 knapsack polytopes, Z. Oper. Res. Ser. A-B , 26, 1982, 243–249.

[27] D. Hartvigsen and E. Zemel, "On the Computational Complexity of Facets and Valid Inequalities for the Knapsack Problem," Discrete Applied Math., 39 (1992), 113-123.

[28] H. Hatami, A. Magen, E. Markakis. Integrality gaps of semidefinite programs for Vertex Cover and relations to ℓ_1 embeddability of Negative Type metrics. APPROX, pp. 164-179, 2007.

[29] S.-P. Hong and L. Tunçel. Unification of lower-bound analyses of the lift-and-project rank of combinatorial optimization polyhedra. *Discrete Applied Mathematics* **156** (2008), pp. 25–41.

[30] O H Ibarra, C E Kim. Fast approximation algorithms for the knapsack and sum of subset problems (1975). Journal of the ACM.

[31] Hans Kellerer, Ulrich Pferschy and David Pisinger. "Knapsack problems." Springer, 2003.

[32] S. Khot. On the power of unique 2-prover 1-round games. In Proceedings of the 34th ACM Symposium on Theory of Computing, pages 767–775, 2002.

[33] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other two-variable CSPs? SIAM J. of Computing 37(1), pp. 319-357, 2007.

[34] Subhash Khot and Rishi Saket. SDP Integrality Gaps with Local ℓ_1-Embeddability. IEEE FOCS, 2009.

[35] A. Kojevnikov and D. Itsykson. "Lower bounds for static Lovász-Schrijver calculus proofs for Tseitin tautologies." *Proceedings of ICALP* 2006, pp. 323–334.

[36] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization* **11** (2001), pp. 796–817.

[37] J.B. Lasserre. An explicit exact SDP relaxation for nonlinear 0-1 programs. *Proceedings of the 8th International Conference on Integer Programming and Combinatorial Optimization*, 2001, Springer Lecture Notes in Computer **2081**, pp. 293–303.

[38] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0-1 programming. *Mathematics of Operations Research* **28** (2003), pp. 470–496.

[39] E L Lawler. Fast Approximation Algorithms for Knapsack Problems (1977). Proceedings of 18th Annual Symposium on Foundations of Computer Science

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

[41] Silvano Martello and P. Toth Knapsack Problems: Algorithms and Computer Implementations, John Wiley & Sons, ChichesterNew York, 1990

[42] Claire Mathieu, Alistair Sinclair. Sherali-adams relaxations of the matching polytope. ACM STOC 2009.

[43] Elchanan Mossel, Ryan O'Donnell, Krzysztof Oleszkiewicz. Noise stability of functions with low infuences invariance and optimality. FOCS 2005: 21-30

[44] T. Pitassi and N. Segerlind. Exponential lower bounds and integrality gaps for tree-like Lovász-Schrijver procedures. ACM-SIAM SODA, 2009.

[45] Prasad Raghavendra. Optimal Algorithms and Inapproximability Results for every CSP?. STOC 2008.

[46] Prasad Raghavendra and David Steurer. Integrality gaps for Strong SDP Relaxations of Unique Games. IEEE FOCS, 2009.

[47] G. Schoenebeck. Linear level Lasserre lower bounds for certain k-CSPs. *Proceedings of the 49th IEEE Symposium on Foundations of Computer Science*, 2008.

[48] G. Schoenebeck, L. Trevisan and M. Tulsiani. Tight integrality gaps for Lovász-Schrijver SDP relaxations of vertex cover and max cut. *Proceedings of the 39th ACM Symposium on Theory of Computing*, 2007, pp. 302–310.

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

[50] I. Tourlakis. New lower bounds for vertex cover in the Lovász-Schrijver hierarchy. *Proceedings of the 21st IEEE Conference on Computational Complexity*, 2006.

[51] M. Tulsiani CSP Gaps and Reductions in the Lasserre Hierarchy. ACM STOC, 2009.

[52] Robert Weismantel On the 0/1 knapsack polytope. Math. Program. 77: 49-68 (1997)

[53] Laurence A. Wolsey. Valid inequalities for 0-1 knapsacks and mips with generalised upper bound constraints. Discrete Applied Mathematics 29(2-3): 251-261 (1990)

[54] E. Zemel, "Easily Computable Facets of the Knapsack Problem," Mathematics of Operations Research, 14 (1989), 760-774.

Appendix

A A linear formulation of the Sherali-Adams hierarchy

For any constraint $g_\ell(x) \geq 0$ in the definition of the base polytope K and any subsets $I, J \subseteq V$, the following constraint is a consequence of the fact that $x \in [0,1]^n$:

$$g_\ell(x) \prod_{i \in I} x_i \prod_{j \in J} (1 - x_j) \geq 0. \tag{6}$$

If x is indeed integral, then $x_i^k = x_i$ for any $k \geq 1$. Thus, the constraint obtained by expanding (6) and replacing x_i^k by x_i holds in P and can be added to strengthen the relaxation. However, this constraint is not linear. To preserve the linearity of the system, each product $\prod_{i \in I} x_i$ is replaced by a variable y_I.

In addition, to keep the number of variables from growing exponentially, we restrict ourselves to only variables y_I such that $|I| \leq t$. By this, we "lift" the polytope K to a polytope $\mathsf{SA}^t(K) \subseteq [0,1]^{\mathcal{P}_t(V)}$.

Definition 19 *Let K be a polytope defined as in equation 2. For any $1 \leq t \leq n$, the t-th* Sherali-Adams lifted polytope $\mathsf{SA}^t(K)$ *is defined by*

$$\mathsf{SA}^t(K) = \left\{ y \in \mathcal{P}_t([n]) \middle| y_\emptyset = 1, \text{ and } g'_{\ell,I,J}(y) \geq 0 \text{ for any } \ell \text{ and } I, J \subseteq V \text{ s.t. } |I \cup J| \leq t-1 \right\}$$

where $g'_{\ell,I,J}(y)$ is obtained by:

1. *multiplying $g_\ell(x)$ by $\prod_{i \in I} x_i \prod_{j \in J}(1 - x_j)$;*
2. *expanding the result and replacing each x_i^k by x_i; and*
3. *replacing each $\prod_{i \in S} x_i$ by y_S.*

We say that a point $x \in [0,1]^V$ belongs to the t-th Sherali-Adams polytope $\mathsf{sa}^t(K)$ *iff there exists a $y \in \mathsf{SA}^t(K)$ such that $y_{\{i\}} = x_i$ for all $i \in V$.*

In particular, in the case of Knapsack, $\mathsf{SA}^t(K)$ is the set of all points in $[0,1]^{\mathcal{P}_t(V)}$ that satisfy the following constraints for any $I, J \subseteq V$ such that $I \cap J = \emptyset$ and $|I| + |J| \leq t - 1$:

$$\sum_{i=1}^{n} c_i \sum_{L \subseteq J} (-1)^{|L|} y_{I \cup L \cup \{i\}} \leq C \sum_{L \subseteq J} (-1)^{|L|} y_{I \cup L}, \tag{7}$$

and

$$0 \leq \sum_{L \subseteq J} (-1)^{|L|} y_{I \cup L \cup \{i\}} \leq \sum_{L \subseteq J} (-1)^{|L|} y_{I \cup L}, \quad \forall i \in V.$$

For a proof that this definition is equivalent to Definition 3, we refer the reader to Laurent's paper [38].

B A New PTAS for Knapsack

As a side product of the analysis of the Lasserre hierarchy, we obtain a new PTAS, albeit a rather inefficient one, for Knapsack. First, solve the SDP of the t-th level of the Lasserre hierarchy for Knapsack to obtain the vector $y \in \{0,1\}^{\mathcal{P}_{2t}(V)}$. Let S be the set of objects of value at least $\frac{OPT}{t-1}$ (Assume we know OPT by binary search?). By the analysis, we know that $y|_{\mathcal{P}(S)}$ is a convex combination of the integral vectors in $\{0,1\}^{\mathcal{P}(OPT)}$. Then, for each integral vector w that appears in the convex combination with positive coefficient, we compute the value of X_i and construct the greedy solution on the reduced instance $C - \text{cost}(X_i), V \setminus S$. We output the best of all the solutions thus constructed.