

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

Anna R. Karlin^{1,*}, Claire Mathieu², and C. Thach Nguyen^{1,*}

¹ University of Washington
{karlin,ncthach}@cs.washington.edu
² Brown University
claire@cs.brown.edu

Abstract. In this paper, we study the integrality gap of the Knapsack linear program in the Sherali-Adams and Lasserre hierarchies. First, we show that an integrality gap of $2 - \epsilon$ persists up to a linear number of rounds of Sherali-Adams, despite the fact that Knapsack admits a fully polynomial time approximation scheme [24, 30]. Second, we show that the Lasserre hierarchy closes the gap quickly. Specifically, after t rounds of Lasserre, the integrality gap decreases to $t/(t - 1)$. This answers the open question in [9]. Also, to the best of our knowledge, this is the first positive result that uses more than a small number of rounds in the Lasserre hierarchy. Our proof uses a decomposition theorem for the Lasserre hierarchy, which may be of independent interest.

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. This paradigm is amazingly powerful; in particular, under the unique game conjecture, it yields the best possible ratio for MaxCut and a wide variety of other problems, see e.g. [33].

However, these algorithms have a limitation. Since they are usually analyzed by comparing the value of the output to that of the fractional solution, we cannot generally hope to get a better approximation ratio than the integrality gap of the relaxation. Furthermore, for any given combinatorial optimization problem, there are many possible LP/SDP relaxations, and it is difficult to determine which relaxations have the best integrality gaps.

This has led to efforts to provide systematic procedures 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 [31], Sherali and Adams [37], Balas, Ceria and Cornuejols [5], Lasserre [28, 27] and others. While they differ in the details, they

* The first and third authors were supported by NSF Grant CCF-0635147 and a Yahoo! Faculty Research Grant.

all operate in a series of rounds starting from an LP or SDP relaxation, eventually ending with an exact integer formulation. 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 gaps of successive relaxations drop sufficiently fast, there is the potential for an improved approximation algorithm (see [15, 16, 7, 10] for example). On the other hand, a large integrality gap persisting for a large, say logarithmic, number of rounds rules out (unconditionally) a very wide class of efficient approximation algorithms, namely those whose output is analyzed by comparing it to the value of a class of LP/SDP relaxations. This implicitly contains most known sophisticated approximation algorithms for many problems including `SparsestCut` and `MaximumSatisfiability`. Indeed, several very strong negative results of this type have been obtained (see [2, 1, 11, 13, 19, 32, 35, 36, 34, 20, 21, 8] and others). These are also viewed as lower bounds of approximability in certain restricted models of computation.

How strong are these restricted models of computation? In other words, how much do lower bounds in these models tell us about the intrinsic hardness of the problems studied? To explore this question, we focus on one problem that is well-known to be “easy” from the viewpoint of approximability: `Knapsack`. We obtain the following results:

- 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 natural LP is 2.)

This is interesting since `Knapsack` has a fully polynomial time approximation scheme [24, 30]. This confirms and amplifies what has already been observed in other contexts (e.g. [13]): the Sherali-Adams restricted model of computation has serious weaknesses: a lower bound in this 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 rounds of Lasserre, the integrality gap decreases to $t/(t - 1)$.

It is known that a few rounds of Lasserre can yield better relaxations. For example, two rounds of Lasserre applied to the `MaxCut` LP yields an SDP that is at least as strong as that used by Goemans and Williamson to get the best known approximation algorithm, and the SDP in [3] which leads to the best known approximation algorithm for `SparsestCut` can be obtained by three rounds of Lasserre. However, to the best of our knowledge, this is the first positive result for more than a constant number of rounds in the Lasserre hierarchy.

Our results also answer the open question in [9], which asks for the performance of lift-and-project methods for `Knapsack`.

1.1 Related Work

Many known approximation algorithms can be recognized in hindsight as starting from a natural relaxation and strengthening it using a couple of levels of lift-and-project. The original hope [2] had been to use lift and project systems as a systematic approach to designing novel algorithms with better approximation ratios. 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` [13, 17, 36], `SparsestCut` [13, 14], `VertexCover` [1, 2, 12, 19, 20, 23, 36, 38], `MaximumAcyclicSubgraph` [13], `CSP` [35, 39], and more.

The `Knapsack` problem has a fully polynomial time approximation scheme [24, 30]. The natural LP relaxation (to be stated in full detail in the next section) has an integrality gap of $2 - \epsilon$ [25]. 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, as stronger LP relaxations are extremely useful to speed up branch-and-bound heuristics. The knapsack polytope was studied in detail by Weismantel [40]. Valid inequalities were studied in [4, 41, 6]. 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 [18, 22, 42]. In [9], Bienstock formulated LP with arbitrary small integrality gaps for `Knapsack` using “structural disjunctions”. As mentioned earlier, this paper poses the question of analyzing lift-and-project methods for `Knapsack`. Our results answer this question.

Our results also confirm the indication from [26, 34] for example that the Sherali-Adams lift and project is not powerful enough to be an indicator of the hardness of problems. On the other hand, little is known about the Lasserre hierarchy, as the first negative results were about k -`CSP` [35, 39]. Our positive result leaves 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 standard linear programming (LP) relaxation [25] for `Knapsack` is given by:

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

The intended interpretation 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. The following lemma is folklore.

Lemma 1. *Consider an instance (C, V) of Knapsack and its LP relaxation K given by (1). Then $\text{Value}(K) \leq \text{Value}(\text{Greedy})(C, V) + \max_{i \in V} v_i$.*

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 [29]. To see that these hierarchies apply at a much greater level of generality we refer the reader to Laurent’s paper [29].

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

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

We are interested in optimizing a linear objective function f over the convex hull $P = \text{conv}(K \cap \{0, 1\}^n)$ 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.

To this end, we first state some notations. 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 ([29]). *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 columns 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 1. *For any $1 \leq t \leq n$, the t -th Sherali-Adams lifted polytope $\text{SA}^t(K)$ 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 $\text{sa}^t(K)$ iff there exists a $y \in \text{SA}^t(K)$ such that $y_{\{i\}} = x_i$ for all $i \in [n]$.

Definition 2. *For any $1 \leq t \leq n$, the t -th Lasserre lifted polytope $\text{La}^t(K)$ 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 $\text{la}^t(K)$ if there exists a $y \in \text{La}^t(K)$ 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 $\text{sa}^{t+1}(K) \subseteq \text{sa}^t(K)$, and $\text{la}^{t+1}(K) \subseteq \text{la}^t(K)$ for all $1 \leq t \leq n - 1$. Sherali and Adams [37] show that $\text{sa}^n(K) = P$, and Lasserre [28, 27] show that $\text{la}^n(K) = P$. Thus, the sequences

$$\begin{aligned} K \supseteq \text{sa}^1(K) \supseteq \text{sa}^2(K) \supseteq \dots \supseteq \text{sa}^n(K) &= P \\ K \supseteq \text{la}^1(K) \supseteq \text{la}^2(K) \supseteq \dots \supseteq \text{la}^n(K) &= P \end{aligned}$$

define hierarchies of polytopes that converge to P . Furthermore, the Lasserre hierarchy is stronger than the Sherali-Adams hierarchy: $\text{la}^t(K) \subseteq \text{sa}^t(K)$ for all $1 \leq t \leq n$. In this paper, we show that for the Knapsack problem, the Lasserre hierarchy is strictly stronger.

2.3 Proof Overview

Consider instances of Knapsack where at most $k - 1$ objects can be put into the knapsack. Examples are instances where all objects have sizes greater than C/k .

Such instances are easy: they can be solved by going through all subsets of at most $k - 1$ objects. We ask if SA and La “realize” this.

It turns out that SA does not. In fact, our lower bound instances fall into this category. For $t \geq k$, SA^t does require that $y_I = 0$ if I does not fit in the knapsack; in some senses, this means the fractional solution y has to be a combination of integral solutions. However, SA^t has very few constraints on the “interaction” of these solutions, thus fails to enforce the combination to be *convex*. In fact, the solution in Section 3 can be viewed as a combination of feasible integral solutions, but the coefficients of this combination do not sum to 1.

On the other hand, La^k handles these special instances. La^k also requires that $y_I = 0$ for any set I that does not fit in the knapsack. More importantly, if we extend y so that $y_I = 0$ for $|I| > k$, then the two constraints $M_{\mathcal{P}(V)}(y) \succeq 0$ and $M_{\mathcal{P}_k(V)}(y) \succeq 0$ are essentially the same, since the former matrix is the latter one padded by 0’s. The former constraint requires y to be a convex combination of integral solutions while the latter is what La^k enforces.

Now consider a general Knapsack instance K and let $y \in \text{La}^k(K)$. Let OPT be the optimal value of K , and L be the set of objects whose values are greater than OPT/k . Then no subset of more than $k - 1$ objects in L can be put into the knapsack. Thus, y is a convex combination of vectors which are integral on $\mathcal{P}_k(L)$. If these (fractional) vectors are feasible for the original Knapsack LP, then by Lemma 1, the value of each can be bounded by $\text{OPT} + \max_{i \notin L} v_i$, which is at most $\frac{k+1}{k}\text{OPT}$. Hence so is the value of y .

Proving that these vectors are feasible for the original Knapsack LP turns out to be very technical and Section 4 is dedicated to it. The section proves a stronger fact that any $y \in \text{La}^t(K)$ can be written as a convex combination of vectors which are integral on $\mathcal{P}_t(L)$ and feasible in $\text{La}^{t-k}(K)$. Section 5 then finishes the analysis of La for Knapsack.

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 Knapsack is close to 2. This lower bound even holds for the *uniform*¹ Knapsack problem, in which $v_i = c_i = 1$ for all i .

Theorem 1. *For every $\epsilon, \delta > 0$, the integrality gap at the t -th level of the Sherali-Adams hierarchy for Knapsack where $t \leq \delta n$ is at least $(2 - \epsilon)(1/(1 + \delta))$.*

Proof. (Sketch) Consider the instance K of Knapsack with n objects where $c_i = v_i = 1$ for all $i \in V$ and capacity $C = 2(1 - \epsilon)$. Then the optimal integer value is 1. On the other hand, we claim that the vector y where $y_\emptyset = 1$, $y_{\{i\}} = C/(n + (t - 1)(1 - \epsilon))$ and $y_I = 0$ for all $|I| > 1$ is in $\text{SA}^t(K)$. Thus, the integrality gap of the t th round of Sherali-Adams is at least $Cn/(n + (t - 1)(1 - \epsilon))$, which is at least $(2 - \epsilon)(1/(1 + \delta))$ when $t \leq \delta n$.

¹ This problem is also known as *Unweighted Knapsack* or *Subset Sum*.

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 3 (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.

Definition 4 (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 3 (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 $y' = \sum_{X \subseteq S} z^X$.

Lemma 4. 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 1. 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 5 (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

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

The following lemma generalizes Lemma 5 in [29]. It proves that the positive-semidefinite property carries over from y to (z^X) .

Lemma 5. 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

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

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

$$|I \cap S| \geq k \implies y_I = 0. \tag{3}$$

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:

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

To prove Theorem 2, we will need a couple more lemmas. In the first one, using assumption (3), we extend the positive semi-definite properties from y to y' , and then, using Lemma 5, from y' to z^X .

Lemma 6. Let t, y, S, k be defined as in Theorem 2, and y' be the extension of y . Let $\mathcal{T}_1 = \{A \text{ such that } |A \setminus S| \leq t - k\}$, and $\mathcal{T}_2 = \{B \text{ such that } |B \setminus S| < t - k\}$. 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 5, 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 5 again we get $M_{\mathcal{T}_2}(g_\ell * z^X) \succeq 0$.

Lemma 7. *Let t, y, S, k be defined as in Theorem 2, and y' be the extension of y . Then for any $X \subseteq S$:*

1. $z_\emptyset^X \geq 0$.
2. If $z_\emptyset^X = 0$ then $z_I^X = 0$ for all $|I| \leq 2t - 2k$.

Proof. Let \mathcal{T}_1 be defined as in Lemma 6. By Lemma 6 $M_{\mathcal{T}_1}(z^X) \succeq 0$ and z_\emptyset^X is a diagonal element of this matrix, hence $z_\emptyset^X \geq 0$.

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 $z_\emptyset^X = 0$,

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

hence $z_J^X = 0$.

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 $z_{I_1}^X = z_{I_2}^X = 0$,

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

hence $z_I^X = 0$.

We now have what we need to prove Theorem 2.

Proof (of Theorem 2). By definition, Lemma 3 and the second part of Lemma 7, we have

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

By Lemma 3 and by definition of y , we have $\sum_{X \subseteq S} z_\emptyset^X = y_\emptyset = 1$, and the terms are non-negative by the first part of Lemma 7, so $y|_{\mathcal{P}_{2t-2k}(V)}$ is a convex combination of w^X 's, as desired.

Consider $X \subseteq S$ such that $z_\emptyset^X \neq 0$. By Lemma 6, $M_{\mathcal{T}_1}(z^X) \succeq 0$ 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_\emptyset^X , 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, $w^X|_{\mathcal{P}_{2t-2k}(V)} \in \text{La}^{t-k}(K)$.

Since $M_{\mathcal{P}_{t-k}(V)}(w^{X_i}) \succeq 0$, by taking a principal submatrix, we infer that $M_{\mathcal{P}_{t-k}(V \setminus S)}(w^{X_i}) \succeq 0$. 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 $x_j = w_{\{j\}}^{X_i}$ 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 1, $g'_\ell = \sum_{j \in V \setminus S} a_j x_j + (b + \sum_{j \in X_i} a_j)$. Let $I \subseteq V \setminus S$. We see that

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

By Lemma 4, $w_{I \cup \{j\}}^{X_i} = w_I^{X_i}$ for $j \in X_i$ and $w_{I \cup \{j\}}^{X_i} = 0$ 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 2 to prove that for the Knapsack problem the gap of $\text{La}^t(K)$ 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 $\text{La}^t(K)$ satisfies the condition of the Theorem.

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

Lemma 8. *Consider an instance (C, V) of Knapsack and its linear programming relaxation K given by (1). Let $t > 1$ and $y \in \text{La}^t(K)$. Let $k < t$ and $S = \{i \in V \mid v_i > \text{OPT}(C, V)/k\}$. 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 $(g * y)_I \leq C y_I - (\sum_{i \in I} c_i) y_I$. 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.

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

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

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

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

Proof. Let $S = \{i \in V \mid v_i > \text{OPT}(C, V)/(t - 1)\}$. Let $y \in \text{La}^t(K)$. If $|I \cap S| \geq t - 1$, then the elements of $I \cap S$ have total value greater than $\text{OPT}(C, V)$, so they must not be able to fit in the knapsack: their total capacity exceeds C , and so by Lemma 8 we have $y_I = 0$. Thus the condition of Theorem 2 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:

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

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

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

By the second property of the Theorem, w^{X_i} is in $\text{La}^{t-k}(K) \subseteq K$, 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:

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

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

6 Conclusion

We have shown that for **Knapsack**, an integrality gap of $2 - \epsilon$ persists up to a linear number of rounds in the Sherali-Adams hierarchy. This broadens the class of problems for which Sherali-Adams is not strong enough to capture the intrinsic difficulty of problems.

On the other hand, the positive result for Lasserre opens the possibility that lower bounds in the Lasserre hierarchy are good indicators of the intrinsic difficulty of the problems at hand. Further investigations into the performance of the Lasserre Hierarchy on other “easy” problems such as **SpanningTree**, **BinPacking**, etc. to either confirm or reject this possibility would be of interest.

One obstacle along this line is the fact that the second positive semidefinite constraint of the hierarchy ($M_{\mathcal{P}(t)V}(g_\ell * y) \succeq 0$) is notoriously hard to deal with, especially when g_ℓ contains many variables (in the lower bounds for k -CSPs [35, 39], the authors are able to get around this by constructing vectors for only valid assignments, an approach that is possible only when all the constraints are “small”.) Clearly, both lower bounds and upper bounds for the Lasserre hierarchy for problems with large constraints remain interesting to pursue.

Acknowledgement

Clare Mathieu would like to thank Eden Chlamtac for stimulating discussions.

References

1. Alekhovich, M., Arora, S., Toulakis, I.: Towards strong non-approximability results in the Lovász-Schrijver hierarchy. In: ACM STOC (2005)
2. Arora, S., Bollobás, B., Lovász, L., Toulakis, I.: Proving integrality gaps without knowing the linear program. *Theory of Computing* 2, 19–51 (2006)
3. Arora, S., Rao, S., Vazirani, U.V.: Expander flows, geometric embeddings and graph partitioning. *J. ACM* 56(2) (2009)
4. Balas, E.: *Facets of the Knapsack Polytope* (1998)
5. Balas, E., Ceria, S., Cornuéjols, G.: A lift-and-project cutting plane algorithm for mixed 0-1 programs. *Mathematical Programming* 58, 295–324 (1993)
6. Balas, E., Zemel, E.: Facets of the knapsack polytope from minimal covers. *SIAM Journal on Applied Mathematics* 34, 119–148 (1978)
7. Bateni, M.H., Charikar, M., Guruswami, V.: MaxMin allocation via degree lower-bounded arborescences. In: ACM STOC (2009)
8. Benabbas, S., Magen, A.: Extending SDP integrality gaps to sherali-adams with applications to QUADRATIC PROGRAMMING and MAXCUTGAIN. In: Eisenbrand, F., Shepherd, F.B. (eds.) IPCO 2010. LNCS, vol. 6080, pp. 299–312. Springer, Heidelberg (2010)
9. Bienstock, D.: Approximate formulations for 0-1 knapsack sets. *Operational Research Letters* 36(3), 317–320 (2008)
10. Bienstock, D., Ozbay, N.: Tree-width and the Sherali-Adams operator. *Discrete Optimization* 1(1), 13–21 (2004)
11. Buresh-Oppenheimer, J., Galesi, N., Hoory, S., Magen, A., Pitassi, T.: Rank bounds and integrality gaps for cutting plane procedures. In: IEEE FOCS (2003)

12. Charikar, M.: On semidefinite programming relaxations for graph coloring and vertex cover. In: ACM-SIAM SODA (2002)
13. Charikar, M., Makarychev, K., Makarychev, Y.: Integrality gaps for Sherali-Adams relaxations. In: ACM STOC (2009)
14. Cheeger, J., Kleiner, B., Naor, A.: A $(\log n)^{\Omega(1)}$ integrality gap for the Sparsest Cut SDP. In: IEEE FOCS (2009)
15. Chlamtac, E.: Approximation algorithms using hierarchies of semidefinite programming relaxations. In: IEEE FOCS (2007)
16. Chlamtac, E., Singh, G.: Improved approximation guarantees through higher levels of SDP hierarchies. In: Goel, A., Jansen, K., Rolim, J.D.P., Rubinfeld, R. (eds.) APPROX and RANDOM 2008. LNCS, vol. 5171, pp. 49–62. Springer, Heidelberg (2008)
17. Fernandez de la Vega, W., Kenyon-Mathieu, C.: Linear programming relaxations of MaxCut. In: ACM-SIAM SODA (2007)
18. Escudero, L.F., Garn, A.: An $o(n \log n)$ procedure for identifying facets of the knapsack polytope. *Operational Research Letters* 31(3), 211–218 (2003)
19. Georgiou, K., Magen, A., Pitassi, T., Toulakis, I.: Integrality gaps of $2 - o(1)$ for vertex cover SDPs in the Lovász-Schrijver hierarchy. In: IEEE FOCS (2007)
20. Georgiou, K., Magen, A., Toulakis, I.: Vertex cover resists sDPs tightened by local hypermetric inequalities. In: Lodi, A., Panconesi, A., Rinaldi, G. (eds.) IPCO 2008. LNCS, vol. IPCO, pp. 140–153. Springer, Heidelberg (2008)
21. Georgiou, K., Magen, A., Tulsiani, M.: Optimal sherali-adams gaps from pairwise independence. In: Dinur, I., Jansen, K., Naor, J., Rolim, J. (eds.) APPROX 2009. LNCS, vol. 5687, pp. 125–139. Springer, Heidelberg (2009)
22. Hartvigsen, D., Zemel, E.: On the computational complexity of facets and valid inequalities for the knapsack problem. *Discrete Applied Math.* 39, 113–123 (1992)
23. Hatami, H., Magen, A., Markakis, E.: Integrality gaps of semidefinite programs for vertex cover and relations to ℓ_1 embeddability of negative type metrics. In: Charikar, M., Jansen, K., Reingold, O., Rolim, J.D.P. (eds.) RANDOM 2007 and APPROX 2007. LNCS, vol. 4627, pp. 164–179. Springer, Heidelberg (2007)
24. Ibarra, O.H., Kim, C.E.: Fast approximation algorithms for the knapsack and sum of subset problems. *Journal of the ACM* (1975)
25. Kellerer, H., Pferschy, U., Pisinger, D.: *Knapsack problems*. Springer, Heidelberg (2003)
26. Khot, S., Saket, R.: SDP integrality gaps with local ℓ_1 -embeddability. In: IEEE FOCS (2009)
27. Lasserre, J.B.: An explicit exact SDP relaxation for nonlinear 0-1 programs. In: Aardal, K., Gerards, B. (eds.) IPCO 2001. LNCS, vol. 2081, p. 293. Springer, Heidelberg (2001)
28. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization* 11, 796–817 (2001)
29. Laurent, M.: A comparison of the Sherali-Adams, Lovász-schrijver and Lasserre relaxations for 0-1 programming. *Mathematics of Operations Research* 28, 470–496 (2003)
30. Lawler, E.L.: Fast approximation algorithms for knapsack problems. In: IEEE FOCS (1997)
31. Lovász, L., Schrijver, A.: Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization* 1, 166–190 (1991)
32. Pitassi, T., Segerlind, N.: Exponential lower bounds and integrality gaps for tree-like Lovász-Schrijver procedures. In: ACM-SIAM SODA (2009)

33. Raghavendra, P.: Optimal algorithms and inapproximability results for every CSP? In: ACM STOC (2008)
34. Raghavendra, P., Steurer, D.: Integrality gaps for strong SDP relaxations of unique games. In: IEEE FOCS (2009)
35. Schoenebeck, G.: Linear level Lasserre lower bounds for certain k-csps. In: IEEE FOCS (2008)
36. Schoenebeck, G., Trevisan, L., Tulsiani, M.: Tight integrality gaps for Lovász-Schrijver SDP relaxations of vertex cover and max cut. In: ACM STOC, pp. 302–310 (2007)
37. Sherali, H.D., Adams, W.P.: A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics* 3, 411–430 (1990)
38. Toulakis, I.: New lower bounds for vertex cover in the Lovász-Schrijver hierarchy. In: IEEE CCC (2006)
39. Tulsiani, M.: CSP gaps and reductions in the Lasserre hierarchy. In: ACM STOC (2009)
40. Weismantel, R.: On the 0/1 knapsack polytope. *Mathematical Programming* 77, 49–68 (1997)
41. Wolsey, L.A.: Valid inequalities for 0-1 knapsacks and mip with generalised upper bound constraints. *Discrete Applied Mathematics* 29(2-3), 251–261 (1990)
42. Zemel, E.: Easily computable facets of the knapsack problem. *Mathematics of Operations Research* 14, 760–774 (1989)