

## Lecture : Lovász Theta Body. Introduction to hierarchies.

Lecturer: Yuri Faenza

Scribes: Yuri Faenza

## 1 Recall: stable sets and perfect graphs

Recall that, for a graph  $G$ , we denote by  $STAB(G)$  its stable set polytope, that is, the convex hull of the characteristic vectors of its stable sets. The polytope

$$QSTAB(G) = \{x \in \mathbb{R}_n^+ : x(K) \leq 1 \text{ for all cliques } K \text{ of } G\}$$

is a relaxation of the  $STAB(G)$ , and it coincides with  $STAB(G)$  if and only if  $G$  is perfect. In a previous lecture we saw that the best known upper bound on the (linear) extension complexity of  $STAB(G)$  when  $G$  is perfect is  $O(n^{\log n})$ .

## 2 Lovász Theta Body

Let  $G(V, E)$  be a graph and  $n = |V|$ . Define the *Lovász Theta Body* to be:

$$TH(G) = \{x \in \mathbb{R}^n : \begin{array}{ll} x_i &= X_{0i} \quad \text{for } i = 1, \dots, n, \\ X_{00} &= 1 \\ X_{ii} &= X_{i0} \quad \text{for } i = 0, \dots, n, \\ X_{ij} &= 0 \quad \text{for } ij \in E, \\ X &\succeq 0 \end{array} \}.$$

**Theorem 1** *Let  $G$  be a graph. Then  $STAB(G) \subseteq TH(G) \subseteq QSTAB(G)$ . In particular, when  $G$  is perfect,  $STAB(G)$  has a semidefinite lift of size  $n + 1$ .*

**Proof** Let  $x$  be the characteristic vector of a stable set of  $G$ . Set  $U = \begin{pmatrix} 1 \\ x \end{pmatrix}$  and  $X = UU^T$ . Then  $X \succeq 0$  by construction, and one easily checks that  $X$  satisfies the linear equations. This shows  $STAB(G) \subseteq TH(G)$ .

Conversely, let  $x \in TH(G)$  and  $X$  be the semidefinite matrix certifying it. Let  $C$  be a clique of  $G$ . We show that  $\sum_{i \in C} x_i \leq 1$ . This implies  $TH(G) \subseteq QSTAB(G)$ . Wlog let  $C = \{1, \dots, k\}$ . Consider the principal submatrix  $X'$  of  $X$  indexed by rows  $0, \dots, k$ , and the vector  $v \in \mathbb{R}^{|C|+1}$  with  $v_0 = 1$  and  $v_i = -1$  for  $i \neq 0$ .  $X' \succeq 0$  by construction, hence

$$0 \leq v^T X' v = \sum_{i,j} v_i v_j X_{i,j} = 1 + \sum_{i \in C} x_i - 2 \sum_{i \in C} x_i = 1 - \sum_{i \in C} x_i,$$

since the contribution to the summation of the term with  $i = j = 0$  is 1, of the term with  $i = j \neq 0$  is  $x_i$ , of the term with  $i = 0$  and  $j \neq 0$  (resp.  $j = 0$  and  $i \neq 0$ ) is  $-x_j$  (resp.  $-x_i$ ), and all other terms give contribution 0. ■

## 3 Introduction to hierarchies

Recall the motivation behind the Chvátal closure seen at the beginning of the course: we wanted to obtain a sequence of increasingly tighter (linear) relaxations, starting from a possibly loose one. For some problems, e.g. for matching, this procedure was very successful. On the other hand, one of its drawback is that already optimizing over the first Chvátal closure is NP-Hard.

We present here some alternative techniques to obtain increasingly tighter relaxations (which are usually called *rounds*) from a starting linear program, in the special (yet very interesting) case when the original polytope is contained in the 0/1 cube. Those different techniques go under the generic name of *hierarchies* and can be presented in a unified setting. Moreover, since a fixed round of any of those hierarchy can be expressed as linear or semidefinite program of size polynomial in the size of the original polytope, the linear optimization problem over those relaxation will be polynomial-time solvable, provided the original relaxation was compact.

The general framework will therefore be the following: we are given  $Q \subseteq [0, 1]^n$  and we want to find a linear description of  $\text{conv}(Q \cap \{0, 1\}^n)$ . We assume that  $Q \cap \{x : x_i = \alpha\} \neq \emptyset$  for all  $i$  and  $\alpha = 0, 1$ , else we can reduce to a lower-dimensional problem.

### 3.1 Sequential convexification

Let  $x \in P$  and suppose that  $0 < x_i < 1$ . Then  $x$  is the convex combination of two other points of  $Q$ , one with coordinate 0 and the other with coordinate 1. The set of all points of  $Q$  that satisfy this latter property is given by the polytope below.

$$C_i(Q) = \text{conv}\{x \in Q : x_i = 1\} \cup \{x \in Q : x_i = 0\}.$$

Clearly  $P \subseteq C_i(Q) \subseteq Q$ . Because of Balas' theorem on the convex hull of the union of polytopes, the size of  $C_i(P)$  is polynomial in the size of  $Q$ . We can clearly iterate this procedure over other coordinates.

**Lemma 2**  $C_n(C_{n-1}(\dots C_1(Q)) \dots) = P$ .

**Proof** We show by induction on  $i$  that all points in  $C_i(\dots)$  can be obtained as convex combination of points of  $Q$  whose coordinates  $1, \dots, i$  are 0 or 1. This is true by construction for  $i = 1$ . Suppose it true up to level  $i - 1$ , and pick  $x \in C_i(\dots)$ . Then  $x$  is the convex combinations of two points, say  $x^1$  and  $x^2$  of  $C_{i-1}(\dots)$ , having the  $i$ -th coordinate equal to 1 (resp. 0). By induction,  $x^1$  (resp.  $x^2$ ) is the convex combinations of a set of points  $X$  (resp.  $X'$ ) of  $Q$  having all coordinates  $1, \dots, i - 1$  equal to 0 or 1, and coordinate  $i$  equal to 1 (resp. 0). Hence, all points in  $X$  and  $X'$  belong to  $Q$  and have coordinates  $1, \dots, n$  integer. We can then express  $x$  as convex combination of those points, as required. ■

In order to obtain generalizations of this *sequential convexification* procedure, it will be better to formulate it in terms of cones. Recall that the *homogeneization* of a closed convex set  $R$  is the set

$$K(R) = \left\{ \begin{pmatrix} x_0 \\ x \end{pmatrix} \in \mathbb{R}^{n+1} : x_0 \geq 0 \text{ and } x \in x_0 R \right\} = \text{cone} \left\{ \begin{pmatrix} 1 \\ x \end{pmatrix} : x \in R \right\}$$

and consequently  $R = \{x \in \mathbb{R}^n : \begin{pmatrix} 1 \\ x \end{pmatrix} \in K(R)\}$ . Therefore, for all purposes, a linear description of a polytope  $R$  is as good as a linear description of  $K(R)$ .

We now provide a description of  $K(C_i(Q))$ . From now on, we will always work with cones  $K \subseteq \mathbb{R}^{n+1}$  such that  $K \cap \{x : x_0 = 1\} \subseteq \{1\} \times [0, 1]^n$ . Let

$$M_i(K) = \left\{ \hat{x} = \begin{pmatrix} x_0 \\ x \end{pmatrix}, \hat{v} = \begin{pmatrix} v_0 \\ x \end{pmatrix} \in \mathbb{R}^{n+1} : v_i = v_0 = v_i, \hat{v} \in K, \hat{x} - \hat{v} \in K \right\} \text{ and } N_i(K) = \text{proj}_{\hat{x}}(M_i(K)).$$

Clearly, both  $M_i(K)$  and  $N_i(K)$  are cones.

**Lemma 3**  $K(P) \subseteq K(C_i(Q)) = N_i(K(Q)) \subseteq K(Q)$ .

**Proof**

Since  $P \subseteq C_i(Q) \subseteq Q$ , we immediately obtain  $K(P) \subseteq K(C_i(Q)) \subseteq K(Q)$ . To show  $K(C_i(Q)) = N_i(K(Q))$ , observe:

$$\begin{aligned} C_i(Q) &= \text{conv}\{x \in \mathbb{R}^n : \exists v, w \in \mathbb{R}^n, \lambda \in [0, 1] && : v_i = \lambda, w_i = 0, \quad v \in \lambda Q, w \in (1 - \lambda)Q, x = v + w\} \\ &= \text{conv}\{x \in \mathbb{R}^n : \exists v \in \mathbb{R}^n, \lambda \in [0, 1] && : v_i = \lambda = x_i, \quad v \in \lambda Q, x - v \in (1 - \lambda)Q\} \\ &= \text{conv}\{x \in \mathbb{R}^n : \exists \hat{v} = \begin{pmatrix} v_0 \\ v \end{pmatrix} \in \mathbb{R}^{n+1}, v_0 \in [0, 1] && : v_i = v_0 = x_i, \quad \hat{v} \in K(Q), \begin{pmatrix} 1 \\ x \end{pmatrix} - \hat{v} \in K(Q)\}. \end{aligned}$$

Consequently,

$$\begin{aligned} K(C_i(Q)) &= \text{cone}\{\hat{x} = \begin{pmatrix} 1 \\ x \end{pmatrix} \in \mathbb{R}^{n+1} : \exists \hat{v} = \begin{pmatrix} v_0 \\ v \end{pmatrix} \in \mathbb{R}^{n+1}, v_0 \in [0, 1] : \\ &\quad v_i = v_0 = x_i, \hat{v} \in K(Q), \hat{x} - \hat{v} \in K(Q)\}. \\ &= \text{cone}\{\hat{x} = \begin{pmatrix} x_0 \\ x \end{pmatrix} \in \mathbb{R}^{n+1} : \exists \hat{v} = \begin{pmatrix} v_0 \\ v \end{pmatrix} \in \mathbb{R}^{n+1} : \\ &\quad v_i = v_0 = x_i, \hat{v} \in K(Q), \hat{x} - \hat{v} \in K(Q)\}, \end{aligned}$$

as required. ■

From Lemmas 2 and 3 we immediately deduce that repeating the convexification operation for all  $i$  produces  $K(P)$ .

**Lemma 4**  $N_n(N_{n-1}(\dots N_1(K(Q))\dots)) = K(P)$ .

### 3.2 Simultaneous convexification

Applying the convexification operation for variable  $i$  cuts off points in  $Q$  that cannot be written as the convex hull of points of the two faces of  $Q$  induced by  $x_i = 0$  and  $x_i = 1$ . Call those faces  $F_i^0$  and  $F_i^1$ . *Simultaneous convexification* excludes all the points which cannot be written as the convex combination of  $F_i^0$  and  $F_i^1$ , for *some*  $i$ . This is captured by the following cone.

$$N^0(K) = \{\hat{x} = \begin{pmatrix} x_0 \\ x \end{pmatrix} \in \mathbb{R}^{n+1} : \exists \hat{v}^1, \dots, \hat{v}^n \in \mathbb{R}^{n+1} : (\hat{x}, \hat{v}^i) \in M_i(K) \text{ for } i = 1, \dots, n\}$$

From the previous definition, we immediately obtain:

**Lemma 5**  $K(P) \subseteq N^0(K(Q)) \subseteq N_i(K(Q)) \subseteq K(Q)$  for all  $i \in [n]$ , and recursively applying  $n$  times the operator  $N^0(\cdot)$  starting from  $K(Q)$  gives  $K(P)$ .

Let us give a matrix formulation of the operator  $N^0(\cdot)$ .

**Lemma 6**  $N^0(K) = \{\hat{x} \in \mathbb{R}^{n+1} : \begin{matrix} \hat{x}_i & = & X_{0i} & \text{for } i = 0, \dots, n, \\ X \in \mathcal{X}_K & & & \end{matrix} \}$ , where for a cone  $K \subseteq \mathbb{R}^{n+1}$ , we have

$$\mathcal{X}_K = \{X \in \mathbb{R}^{(n+1) \times (n+1)} : \begin{matrix} X e_i & \in & K(Q) & \text{for } i = 1, \dots, n, \\ X(e_0 - e_i) & \in & K(Q) & \\ X_{ii} & = & X_{i0} & \text{for } i = 0, \dots, n, \end{matrix} \}.$$

**Proof** Set  $\hat{x} = X e_0$  and  $\hat{v}^i = X e_i$ , for all  $i = 1, \dots, n$ . ■

### 3.3 A different view on simultaneous convexification

We now give an equivalent procedure to deduce the operator  $N^0(\cdot)$ .

Let  $Q = \{x \in \mathbb{R}^n : Ax \leq b\}$ . Write the  $\ell$ -th inequality from  $Ax \leq b$  as  $g_\ell^T (1 \ x^T) = a_0 + a^T x \geq 0$ . Then we can write

$$K(Q) = \{\hat{x} = \begin{pmatrix} x_0 \\ x \end{pmatrix} \in \mathbb{R}^{n+1} : g_\ell^T \hat{x} \geq 0\}.$$

Note that inequalities  $x_i(g_\ell^T \begin{pmatrix} 1 \\ x \end{pmatrix}) \geq 0$  and  $(1-x_i)(g_\ell^T \begin{pmatrix} 1 \\ x \end{pmatrix}) \geq 0$  are valid for  $P$ , where as usual we write  $P = \text{conv}(Q \cap \{0, 1\}^n)$ , for all  $i \in [n]$ . The set of all such inequalities gives a quadratic system:

$$\begin{aligned} S = \{x \in \mathbb{R}^n : & x_i(g_\ell^T \begin{pmatrix} 1 \\ x \end{pmatrix}) &= g_\ell^T \begin{pmatrix} 1 \\ x \end{pmatrix} (1 \ x^T)e_i &\geq 0 & \text{for all } i \in [n] \text{ and } \ell \\ & (1-x_i)(g_\ell^T \begin{pmatrix} 1 \\ x \end{pmatrix}) &= g_\ell^T \begin{pmatrix} 1 \\ x \end{pmatrix} (1 \ x^T)(e_0 - e_i) &\geq 0 & \text{for all } i \in [n] \text{ and } \ell\}. \end{aligned}$$

The previous system can be linearized setting  $X = \begin{pmatrix} 1 \\ x \end{pmatrix} (1 \ x^T)$ . We then obtain exactly  $N^0(K(Q))$ .

### 3.4 Adding semidefinite constraints - Lovász-Schrijver's $N^+$ operator

The matrix formulation from Lemma 6 allows us to impose "for free" some additional constraints to the system.

**Lemma 7** *Let  $Q \subseteq [0, 1]^n$  be a polytope and  $P = \text{conv}(Q \cap \{0, 1\}^n)$ . Then  $K(P) \subseteq N^+(K(Q)) \subseteq N^0(K(Q))$ , where*

$$N^+(K) = \{\hat{x} \in \mathbb{R}^{n+1} : \begin{array}{l} \hat{x}_i = X_{0i} \quad \text{for } i = 0, \dots, n, \\ X \in \mathcal{X}_K^+ \end{array} \},$$

and  $\mathcal{X}_K^+ = \{X \in \mathcal{X}_K \text{ and } X \succeq 0\}$ .

**Proof** The only statement that requires a proof is  $K(P) \subseteq N^+(K(Q))$ . Fix a point  $x \in Q \cap \{0, 1\}^n$ , let  $\hat{x} = \begin{pmatrix} 1 \\ x \end{pmatrix}$  and  $X = \hat{x}\hat{x}^T$ . Then  $X \succeq 0$ . Using the fact that  $x \in \{0, 1\}^n$ , we deduce that  $Xe_i = \hat{x}$ ,  $X(e_0 - e_i)$  are either equal to  $\hat{x}$  or to 0, and  $X_{ii} = X_{i0}$ . ■

Note that  $N^+(K(Q))$  is not a polyhedron anymore. However, the conic setting in which we phrase it allows us to iteratively apply the  $N^+(\cdot)$  operator.

Let us see an example: recall the *fractional relaxation* of the stable set polytope of a graph  $G(V, E)$ :

$$Q = \{x \in \mathbb{R}_+^n : x_i + x_j \leq 1 \text{ for all } ij \in E\}.$$

We have  $K(Q) = \{\begin{pmatrix} x_0 \\ x \end{pmatrix} \in \mathbb{R}_+^{n+1} : x_i + x_j \leq x_0\}$ . We show that  $N^+(K(Q)) \subseteq K(\text{TH}(G))$ . Note that  $K(\text{TH}(G)) \subseteq \mathbb{R}_+^{n+1}$  is a cone in the variables  $\begin{pmatrix} x_0 \\ x \end{pmatrix}$  whose constraints are exactly those describing  $\text{TH}(G)$ , with the exception of  $X_{00} = 1$  which is replaced by  $X_{00} = x_0$ . Hence, in order to conclude

$N^+(K(Q)) \subseteq K(TH(G))$ , it is enough to show that  $X_{ij} = 0$  for  $ij \in E$  is implied by constraints of  $N^+(K(Q))$ . Fix  $ij \in E$ . We have

$$Xe_i \in N^+(K(Q)) \Rightarrow X_{ii} + X_{ij} \leq X_{i0} \text{ and } X_{ij} \geq 0$$

and using  $X_{ii} = X_{i0}$  we conclude  $X_{ij} = 0$ , as required. So, in particular,  $N^+(K(Q))$  implies all clique inequalities. Those inequalities are known to be hard to obtain for Chvátal closure (in particular, no constant round of Chvátal suffice, in general, to deduce all clique inequalities).

### 3.5 The Sherali-Adams and Lasserre hierarchies

The proof of Lemma 7 suggests us to define matrices associated to vectors in  $\mathbb{R}^n$  with the property that, when  $x$  is a 0/1 point in  $Q$ , the matrix is somehow "well-behaved". Imposing this "good behaviour" on the matrix will therefore cut off some points of  $Q$  that are not points of  $P$ .

Consider the vector  $y \in \mathbb{R}^{2^n}$  indexed by subsets of  $[n]$ . Define the matrix  $Y(y) \in \mathbb{R}^{2^n \times 2^n}$  as follows:

$$Y(y)_{I,J} = y_{I \cup J} \text{ for all } I, J \subseteq [n].$$

**Figure 1:** Matrix  $Y(y) \in \mathbb{R}^{2^3 \times 2^3}$ . For  $r = 2$ , the submatrix induced by  $U = \{1, 2\}$  (rows and columns in in green) is a principal submatrix of the matrix induced by  $\mathcal{P}_2(\{3\})$  (rows and columns in green and red).

$$\begin{array}{c}
 \emptyset \\
 1 \\
 2 \\
 3 \\
 12 \\
 13 \\
 23 \\
 123
 \end{array}
 \begin{pmatrix}
 \emptyset & 1 & 2 & 3 & 12 & 13 & 23 & 123 \\
 y_\emptyset & y_1 & y_2 & y_3 & y_{12} & y_{13} & y_{23} & y_{123} \\
 y_1 & y_1 & y_{12} & y_{13} & y_{12} & y_{13} & y_{123} & y_{123} \\
 y_2 & y_{12} & y_2 & y_{23} & y_{12} & y_{123} & y_{23} & y_{123} \\
 y_3 & y_{13} & y_{23} & y_3 & y_{123} & y_{13} & y_{23} & y_{123} \\
 y_\emptyset & y_{12} & y_{12} & y_{123} & y_{12} & y_{123} & y_{123} & y_{123} \\
 y_{13} & y_{13} & y_{123} & y_{13} & y_{123} & y_{13} & y_{123} & y_{123} \\
 y_{23} & y_{123} & y_{23} & y_{23} & y_{123} & y_{123} & y_{23} & y_{123} \\
 y_{123} & y_{123} & y_{123} & y_{123} & y_{123} & y_{123} & y_{123} & y_{123}
 \end{pmatrix}$$

For an inequality  $g(x) = a_0 + a^T x \geq 0$  and a vector  $y \in \mathbb{R}^{2^n}$ , consider the *shift operator*

$$* : \mathbb{R}^{n+1} \times \mathbb{R}^{2^n} \rightarrow \mathbb{R}^{2^n} \text{ defined as } (g * y)_I = \sum_{i \in [n]_0} a_i y_{I \cup \{i\}}, \text{ where } [n]_0 = [n] \cup \{0\}$$

Now, for a point  $x \in Q \cap \{0, 1\}$ , set  $y_I = \prod_{i \in I} x_i$  and note that  $Y(y) = yy^T$  is positive semidefinite, and satisfies  $Y(y)_{\emptyset, \emptyset} = 1$ .

Moreover, for any linear inequality  $g(x) = a_0 + a^T x \geq 0$  valid for  $Q$ , we have:

$$(g * y)_I = \sum_{i \in [n]} a_i y_{I \cup \{i\}} + a_0 y_I = \sum_{i \in [n]} a_i y_i y_I + a_0 y_I = g(x) y_I \text{ and}$$

$$Y(g * y) = g(x) Y(y) \succeq 0.$$

Hence, we can safely impose constraints

$$Y_{\emptyset, \emptyset} = 1, Y(y) \succeq 0 \text{ and } Y(g * y) \succeq 0 \text{ for all inequalities } g(x) \geq 0 \text{ describing } Q. \quad (1)$$

But those constraint are in an exponential number, and we want a compact relaxation. The Sherali-Adams and Lasserre hierarchies impose some weaker conditions on the matrix  $Y$ : this will trade exactness for compactness. For a set  $U$ , let  $\mathcal{P}(U)$  be the power set of  $U$ ,  $P_r(U)$  the collection of all subsets of  $U$  of size at most  $r$ , and  $g^\ell(x) \geq 0$ ,  $\ell = 1, \dots, m$ , be the inequalities defining  $Q$ . We write  $Y$  for  $Y(y)$  and, for a matrix  $Y$  and a family  $\mathcal{F} \subseteq 2^{[n]}$ , we write  $Y_{\mathcal{F}}$  for its principal submatrix whose rows and columns range over  $\mathcal{F}$ . Let  $r \in [n]_0$ .

$$\text{Let } SA_r(Q) = \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^{2^n} : \begin{array}{ll} x_i & = Y_{\emptyset, \{i\}} & \text{for } i = 1, \dots, n, \\ Y_{\emptyset, \emptyset} & = 1 \\ Y_{\mathcal{P}(U)} & \succeq 0 & \text{for } [n], |U| \leq r \\ Y_{\mathcal{P}(W)}(g^\ell * y) & \succeq 0 & \text{for } |W| \leq r \text{ and all } \ell \end{array} \} \text{ and}$$

$$Las_r(Q) = \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^{2^n} : \begin{array}{ll} x_i & = Y_{\emptyset, \{i\}} & \text{for } i = 1, \dots, n, \\ Y_{\emptyset, \emptyset} & = 1 \\ Y_{\mathcal{P}_r(V)} & \succeq 0 \\ Y_{\mathcal{P}_r(V)}(g^\ell * y) & \succeq 0 & \text{for all } \ell \end{array} \}.$$

Note that, in the previous definitions, even if we are asking for a vector  $y \in \mathbb{R}^{2^n}$ , we are constructing the submatrices of  $Y$  from a subvector of  $y$  of soze polynomial in  $n$  for fixed  $r$ . Hence, the associated semidefinite programs are of polynomial size.

**Lemma 8** *For each  $r \in \mathbb{N}$ ,  $P \subseteq Las_r(Q) \subseteq SA_r(Q) \subseteq N^{0,r}(K(Q)) \subseteq Q$ , where we denoted by  $N^{0,r}(K(Q))$  the convex set obtained by applying  $r$  times the operator  $N^0$ , and then projecting to the original  $x$  variables. Moreover,  $P = Las_n(Q) = SA_n(Q)$ .*

**Proof** Note that all the matrices that are imposed to be semidefinite in  $Las_r(Q)$  are principal submatrices of the matrices  $Y(y)$  and  $Y(g * y)$ . Hence, the corresponding inequalities are relaxation of inequalities (1), showing  $P \subseteq Las_r(Q)$ . A similar argument shows  $Las_r(Q) \subseteq SA_r(Q)$ .

We defer the proof of  $SA_r(Q) \subseteq K^{-1}(N^{0,r}(K(Q)))$  to the exercise session. Then using Lemma 5,  $P = Las_n(Q) = SA_n(Q)$  follows. ■

## 4 Application of hierarchies

### 4.1 Set cover

By definition, when we convexify  $Q$  on variable  $i$ , we have that any point in  $C_i(Q)$  can be written as convex combination of points  $x^1, x^2 \in Q$  such that  $x^1_i = 1$  and  $x^2_i = 0$ .

In simultaneous convexification, we have that, for any point  $x \in proj_x(N^0(Q))$ , for any  $i \in [n]$ ,  $x$  can be written as the convex combination of points in  $Q$  with coordinate  $i$  being 0 and 1. Since in the proof of Lemma 8 we argued that Sherali-Adams operator is stronger than simultaneous convexification, let us state this fact in terms of  $SA_t$ .

**Lemma 9** *Let  $x \in SA_t(Q)$ ,  $i \in [n]$  with  $0 < x_i < 1$ . Then  $x \in conv\{z \in SA_{t-1}(Q) : z_i \in \{0, 1\}\}$ .*

As done in the proof of Lemma 2, we can iterate the previous lemma and conclude the following.

**Lemma 10** *Let  $x \in SA_t(Q)$ ,  $i \in [n]$ ,  $S \subseteq [n]$  with  $|S| \leq t$ . Then*

$$x \in conv\{z \in SA_{t-|S|}(Q) : z_i \in \{0, 1\} \text{ for } i \in S\}.$$

We now use the previous lemma to show an algorithm that runs in subexponential time and achieves an approximation of  $(1 - \varepsilon) \log n + O(1)$  for set cover. Note that it is widely believed that there is no polynomial time algorithm that achieves this approximation.

Recall that in set cover, we are given  $n$  objects and  $m$  sets  $\mathcal{S} = S_1, \dots, S_m$  with costs  $c_1, \dots, c_m$ , each containing some of those objects. The goal is to find a collection of sets whose union contain all objects at minimum total cost.

The algorithm is based on solving a linear optimization problem over  $SA_{n^\varepsilon}(Q)$ , where  $Q$  is the following LP relaxation.

$$\sum_{i: x_j \in S_i} x_i \geq 1 \quad \text{for all } j \quad (2)$$

$$\begin{aligned} \sum_{i \in [m]} c_i x_i &\leq OPT \\ x &\in [0, 1] \end{aligned} \quad (3)$$

and rounding its optimum solution. Note that we are assuming that we know the optimum value  $OPT$ . This is not cheating: standard techniques in approximation algorithms imply that we can round costs and guess a "small" number of values of  $OPT$ , and repeat the algorithm for all those values, picking the best solution.

*First part:* Let  $x$  be the optimum solution of  $\min c^T x : x \in SA_{n^\varepsilon}(Q)$ . Among  $i$  with  $x_i > 0$ , pick the one whose associated set covers the biggest set of objects. Then we know that  $x$  can be written as convex combination of points  $x^1, x^2$  in  $SA_{n^\varepsilon-1}$ , with  $x_i^1 = 1$ . Replace  $x$  with  $x^1$  and iterate  $n^\varepsilon$  times, each time choosing the set that covers the biggest number of uncovered items, among the sets whose corresponding variable is fractional. Note that, if at some point no variable is fractional, we have found a feasible integral solution of value at most  $OPT$ , hence we are done. Suppose this is not the case, and let  $\mathcal{S}'$  be the set picked with this procedure,  $\bar{x}$  the final point obtained. By Lemma 10,  $x \in Q$ , hence  $\sum_{i \in [m]} c_i x_i \leq OPT$ .

*Second part:* Now consider the residual set cover problem, once all sets in  $\mathcal{S}$  have been picked. Note that each of the remaining sets covers at most  $n^{(1-\varepsilon)}$  elements, since otherwise, together with all the  $n^\varepsilon$  sets picked before, which cover at least as many elements as it does, we would have covered more than  $n$  elements. We can then exploit the following well-known fact.

**Theorem 11** *Let  $\bar{x}$  satisfy the linear relaxation of a set cover instance given by (2) and (3). Then an integral solution satisfying (2) and (3) of cost at most  $(1 + \log k) \sum_{i=1}^m c_i \bar{x}_i$  can be found in polynomial time, where  $k$  is the size of the largest set in the set cover instance.*

Consider the solution to the original set cover instance that picks the sets  $\mathcal{S}''$  given by the previous theorem, plus those in  $\mathcal{S}'$  selected in the first part. Clearly it is feasible. Its cost is

$$c(\mathcal{S}') + c(\mathcal{S}'') \leq c^T \bar{x} + (1 + \log(n^{1-\varepsilon}))OPT \leq ((1 - \varepsilon) \log n + O(1))OPT.$$

## 4.2 The decomposition theorem for the Lasserre hierarchy with an application to max Knapsack

Note that so far we did not use any property specific of Lasserre in deriving previous results. Next theorem, whose proof we omit, holds for Lasserre hierarchy only.

**Theorem 12** *[Decomposition Theorem for Lasserre] Let  $x \in Las_t(Q)$ ,  $S \subseteq [n]$  with  $|S \cap \text{supp}(x)| \leq k$  for all  $x \in Q \cap \{0, 1\}^n$ . Then  $x \in \text{conv}\{z \in Las_{t-k}(Q) : z_i \in \{0, 1\} \text{ for } i \in S\}$ .*

The previous theorem allows us to show the integrality gap of the Lasserre hierarchy quickly converges to zero for the max knapsack problem. This separates Lasserre from Sherali-Adams hierarchy, for which this number stays arbitrarily close to 2 after linearly many rounds.

Recall that the *max knapsack* problem is defined as

$$\begin{aligned} \max \quad & c^T x \\ \text{st} \quad & w^T x \leq \beta \\ & x \geq 0 \\ & x \in \{0, 1\}^n, \end{aligned}$$

and the *integrality gap*  $IG_t$  of the  $t$ -th round of Lasserre is  $\frac{\max_{x \in Las_t(Q)} c^T x}{OPT}$ , where  $Q$  is the classical linear relaxation of the Knapsack polytope. It is known well-known that one has  $\max_{x \in Q} c^T x \leq OPT + \max_i c_i$ .

**Theorem 13** *For any knapsack instance and any  $t \geq 2$ ,  $IG_t \leq 1 + o(\frac{1}{t})$ .*

**Proof** Let  $\bar{x}$  be optimum solution of  $\max_{x \in Las_t(Q)} c^T x$ . Let  $S = \{i \in [n] : w_i > \frac{OPT}{t+1}\}$ . Note that, for any integral solution to the knapsack problem, one has  $|I \cap S| \leq t$ , else the profit of  $I$  is strictly bigger than the profit of an optimum solution. Hence we can apply Theorem 12 with  $k = t$ . That is,  $\bar{x}$  is the convex hull of feasible points of  $Q$  that have integer values in the coordinates corresponding to  $S$ . Since the objective function is linear, one of those, say  $x'$ , has profit greater or equal than the profit of  $\bar{x}$ . Let  $S'$  be the coordinates from  $S$  that are set to 1 in  $x'$ . We deduce

$$w^T \bar{x} \leq w^T x' \leq \sum_{i \in S'} w_i + LP',$$

where  $LP'$  is the solution of max profit of the original knapsack relaxation, among those that take elements in  $S'$  and not elements in  $S \setminus S'$ . Using the bound on the integrality gap at the first round and that all objects not in  $S$  have bounded profit, we obtain

$$LP' \leq OPT' + \frac{OPT'}{t+1},$$

where  $OPT'$  is the optimum solution of the integer version of  $LP'$ .

We conclude

$$w^T \bar{x} \leq \sum_{i \in S'} w_i + OPT' + \frac{OPT'}{t+1} \leq OPT + \frac{OPT}{t+1}.$$

■

## 5 Notes

Lovász' Theta Body has been studied in many papers. Results presented in here were proved in [3, 6, 7] and we follow the presentation given in [2]. We refer the reader to [5] for an introduction on the different hierarchies, to [1] for the application to set cover, and to [4] for results from Section 4.2. Our presentation of hierarchies partly follows [8, 9].



## References

- [1] E. Chlamtac, Z. Friggstad, and K. Georgiou. Lift-and-Project Methods for Set Cover and Knapsack. WADS 2013, pp. 256-267.
- [2] M. Conforti, G. Cornuejols, and G. Zambelli. Integer Programming. Springer, 2014.
- [3] M. Grötschel, L. Lovász and A. Schrijver . Relaxations of vertex packing. Journal of Combinatorial Theory (B) 40, pp. 330–343, 1986.
- [4] A. Karlin, C. Mathieu, and C. Nguyen. Integrality gaps of linear and semidefinite programming relaxations for knapsack. In IPCO, pp. 301-314, 2011.
- [5] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming. Mathematics of Operations Research, 28(3), pp. 470-496, 2003.
- [6] L. Lovász. On the Shannon capacity of a graph. IEEE Transactions on Information Theory 25, pp. 1–7, 1979.
- [7] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. SIAM Journal on Optimization, 1(2), pp. 166-190, 1991.
- [8] T. Rothvoss. The lasserre hierarchy in approximation algorithms. Lecture Notes for the MAPSP 2013 - Tutorial, 2013.
- [9] . Zuckerberg. A Set Theoretic Approach to Lifting Procedures for 0,1 Integer Programming. Ph. D. Dissertation, Columbia University, 2009.