KNAPSACK

Given: $\overbrace{c_1 \ldots c_n}^{c_i \le C}, C$ "capacities" total capacity

$v_1 \longrightarrow v_n$ "value" nonneg

Goal: $\max \sum_i v_i x_i$

s.t. $\sum_i c_i x_i \le C$

$x_i \in \{0,1\} \ \forall i$

"collect maximum value while being able to fit the items in your knapsack"

NP-Hard to solve exactly.

Algos?

Greedy: 1) WLOG assume:

$$\frac{v_1}{c_1} \geq \cdots \geq \frac{v_n}{c_n}$$

"density of value"

2) take largest j s.t $\sum_{i=1}^{j} c_i < C$

& output $\{1, \ldots, j\}$.

Lemma: Best of greedy and max value element achieves $1/2$ approx.

We will prove this by first analyzing the natural LP relaxation.

LP:

$$\max \sum_i y_i v_i$$

$$\text{s.t. } \sum_i y_i \cdot c_i \leq C$$

$$0 \leq y_i \leq 1.$$

<u>Lemma:</u> $\max\limits_{\substack{0\le y_1 \dots y_n \le 1 \\ \sum y_i c_i \le C}} \sum y_i v_i \le OPT + \max\limits_{i\le n} v_i$

<u>Proof:</u> WLOG, assume: $\frac{v_1}{c_1} \geq \dots \geq \frac{v_n}{c_n}$.

Claim: If $y_i > 0$ for some $i > 1$, then WLOG $y_1 = 1$.

<u>Proof:</u> If not, set $y_1 \leftarrow y_1 + \frac{y_i \cdot c_i}{c_1}$

& $y_i = 0$

Iterating this argument, optimal y looks like

$$y_i = \begin{cases} 1 & \text{for } i \le k-1 \\ \frac{C - \sum_{j=1}^{k-1} c_i}{c_k} & \text{for } i = k \\ 0 & \text{o/w.} \end{cases}$$

<1

<u>Thus</u>, $\sum y_i v_i < \sum\limits_{i \in [k-1]} v_i + v_k$

$\le OPT(C, V) + \max\limits_{i \ge k} v_i$

Analyzing Approx. Ratio:

$$\text{OPT} \le \text{LP-VAL}$$

$$\le \text{GREEDY-OPT} + \max_i v_i$$

$$\le 2\,\text{OPT}$$

Greedy gets $\frac{1}{2}$ approx.

LP integrality gap ≤ 2.

LP integrality gap $\geq 2-\varepsilon \ \forall \varepsilon > 0$.

<u>e.g.</u> $c_1 = c_2 = \ldots = c_n = 1$

$C = 2-\varepsilon$

$v_1 = v_2 = \ldots = v_n = 1$

LP opt: $y_1 = 1,\ y_2 = 1-\varepsilon$

value $= 2-\varepsilon$

Integral OPT $= 1.$

Better relaxations? LPs? SDPs?

Theorem: For every $\varepsilon > 0$ there is a $n^{O(1/\varepsilon)}$ time $(1+\varepsilon)$-approx. for Knapsack

Multiple proofs: ① Dynamic Programming

② I'll show you a diff method
→ introduction to "hierarchies" of convex relaxations

Strengthening Relaxations: Sum-of-Squares

Problem $\to$ basic relaxation $\to$ rounding
$\searrow$ integrality gap.

Suppose approx. ratio $\simeq$ integrality gap.

What do we do?

Prove hardness of approx. e.g. Max-Cut

What if there's a better algo?

perhaps via better relaxation?

E.g. "triangle inequality" in ARV for sparsest cut.

<u>Today</u>: "Mechanical" method to construct stronger relaxations.

Can be thought of as adding extra variables, additional constraints → often called "lift (add vars) & project".

Specifically, we will see the "sum-of-squares"

To do this, I want to show you a slightly different way to think & reason about SDPs that helps in mechanically strengthening them.

"SDP solutions as relaxations of probability distributions".

Def (Pseudo-distribution)

A pseudo-dist^n D on $\{-1,1\}^n$ is a function $D : \{-1,1\}^n \to \mathbb{R}$ s.t. $\sum_x D(x) = 1$.

(Analog of mass functions of discrete prob distributions).

For any $f : \{-1,1\}^n \to \mathbb{R}$, the pseudo-expectation of f w.r.t. a pseudo-distribution D is

$$\tilde{\mathbb{E}}_D f = \sum_x f(x) \cdot D(x)$$

Example: Let $p : \{-1,1\}^n \to \mathbb{R}_+$ s.t. $\sum_x p(x) = 1$. Then, p is a pseudo-distribution, with its pseudo expectation being simply the expectation w.r.t. p.

We now want to force some non-trivial constraints of being a prob distⁿ on a pseudo-distⁿ D.

<u>Def (Pseudo-distribution of deg d).</u>

A pseudo-distⁿ $D : \{-1,1\}^n \to \mathbb{R}$ is said to be of degree d if for <u>all polynomials f</u> of deg $\leq \frac{d}{2}$,

$$\tilde{\mathbb{E}}_D f^2 \geq 0.$$

<u>Obsⁿ</u> (Linearity): $\forall f, g$, $\tilde{\mathbb{E}}(f+g) = \tilde{\mathbb{E}}f + \tilde{\mathbb{E}}g$

Note: The mass function of a prob-dist^n is a pseudo-distribution of deg $\cancel{\infty}$. 2n

Fact: Every function $f : \{-1,1\}^n \to \mathbb{R}$ is a polynomial of deg $\leq n$.

(proof: Fourier polynomial of f).

In fact this is a characterization of prob. distributions.

Lemma (deg 2n pseudo-dist^n is a prob. dist^n).

Proof: $\forall z \in \{-1,1\}^n$, let

$\mathbb{1}_z : \{-1,1\}^n \to \{0,1\}$ be the function such that $\mathbb{1}_z(x)=1$ if & only if $z=x$. and 0 o/w.

Then, z is a polynomial of deg n.

Thus, if D is a deg $2n$ p. dist^n

then, $\tilde{\mathbb{E}}_D \mathbb{1}_z^2 \geq 0$.

Let $p_z = \tilde{\mathbb{E}}_D \mathbb{1}_z^2$.

Then $\sum_z p_z = \tilde{\mathbb{E}}_D \sum_z \mathbb{1}_z^2$

$= \tilde{\mathbb{E}}_D \mathbb{1}$

$= \sum_x D(x) = 1.$

So p_z gives you a prob dist^n on $\{-1,1\}^n$.

If $f : \{-1,1\}^n \to \mathbb{R}$ is any function,

then $f(x) = \sum_{z \in \{-1,1\}^n} f(z)\, \mathbb{1}_z(x)$.

Thus, $\tilde{\mathbb{E}} f(x) = \sum_{z \in \{-1,1\}^n} f(z)\, \tilde{\mathbb{E}}\, \mathbb{1}_z(x)$

$$= \sum_{z \in \{\pm 1\}^n} f(z)\, p(z)$$

$$= \mathbb{E}_p f \quad \checkmark$$

<u>So</u> $\tilde{\mathbb{E}}$ is really the expectation w.r.t. p. $\to$ So done.

This is an instance of "duality" between prob distributions and non-neg polynomials. In general, such results are studied in algebraic geometry: "Positivstellensatz"

<u>Lemma</u> (Cauchy-Schwarz Inequality)

Let D be a p.d. of deg $\geq 2t$ on $\{-1,1\}^n$. Then $\forall$ poly f, g of deg $\leq t$

$$\tilde{\mathbb{E}}_D f \cdot g \leq \sqrt{\tilde{\mathbb{E}}_D f^2}\sqrt{\tilde{\mathbb{E}}_D g^2}$$

<u>Proof</u>: First suppose $\tilde{\mathbb{E}} g^2 = 0$

Then $$\tilde{\mathbb{E}} \cdot f \cdot g = \frac{\tilde{\mathbb{E}}\left[(cf+g)^2\right] - \tilde{\mathbb{E}}\left[(cf-g)^2\right]}{4c}$$

$$\leq \frac{1}{4c}\left[\tilde{\mathbb{E}}\left[c^2 f^2\right] + 2c\tilde{\mathbb{E}} fg\right]$$

$$\Rightarrow \quad \tilde{\mathbb{E}} fg \leq \frac{c}{2} \tilde{\mathbb{E}} f^2$$

take limits as $c \to 0$.

Now suppose $\tilde{\mathbb{E}} f^2, \tilde{\mathbb{E}} g^2 > 0$.

Let $\|f\|_2^2 = \tilde{\mathbb{E}} f^2$, $\|g\|_2^2 = \tilde{\mathbb{E}} g^2$

$$\bar{f} = f/\|f\|_2, \quad \bar{g} = g/\|g\|_2$$

then

$$\tilde{\mathbb{E}}[(\bar{f}-\bar{g})^2] \geq 0$$

$$\Downarrow$$

$$\tilde{\mathbb{E}}\bar{f}^2 + \tilde{\mathbb{E}}\bar{g}^2 \geq 2\cdot\tilde{\mathbb{E}}\bar{f}\cdot\bar{g}$$

or $\tilde{\mathbb{E}}\,\bar{f}\cdot\bar{g} \leq 1.$

or $\hat{\mathbb{E}}\,f\cdot g \leq \|f\|_2\cdot\|g\|_2$

$$\quad\quad\quad\quad\quad '' \quad\quad '' \quad\quad \sqrt{\tilde{\mathbb{E}}f^2}\sqrt{\hat{\mathbb{E}}g^2}$$

TIP: Commit to the "language"

Often can guess true results & even prove them by first thinking what's true of distn's.

$\square$

<u>Def</u> (Pseudomoments)

The degree t pseudomoments of a pseudo-dist^n D are the set of

#s $\{ \tilde{\mathbb{E}}_D \cdot x_S \mid |S| \leq t \}$

i.e. pseudo-expectations of monomials of deg $\leq t$.

Notation: $\binom{n}{\leq t} = \sum_{i \leq t} \binom{n}{i}$.

<u>Lemma</u>

$\{ Y_S \mid |S| \leq 2t \}$, $Y_\phi = 1$ are pseudomoms of a deg $2t$ pseudo-dist^n on $\{-1,1\}^n$ if and only if the matrix M_t defined by

$$M_t(S,T) = \tilde{\mathbb{E}} X_{S \Delta T}$$

satisfies $M_t \succeq 0$.

Proof :

Let $v \in \mathbb{R}^{\binom{n}{\leq t}}$ be vector. Then

$$v^T M_t v = \sum_{S,T} v_S \cdot v_T \cdot \tilde{\mathbb{E}} \cdot x_{S \Delta T}$$

$$= \sum_{S,T} v_S \cdot v_T \tilde{\mathbb{E}} \cdot x_S x_T$$

$$= \tilde{\mathbb{E}} \sum_{S,T} v_S v_T \cdot x_S x_T$$

$$= \tilde{\mathbb{E}} \Big(\sum_S v_S \cdot x_S \Big)^2$$

Thus

$$v^T M_t v \geq 0 \;\; \forall \, v \in \mathbb{R}^{\binom{n}{\leq t}}$$

$$\iff \tilde{\mathbb{E}} f(x)^2 \geq 0 \;\; \forall \, f \text{ of } \deg \leq t$$

$\square$

Thus checking if a given set of $\binom{n}{\leq 2t}$ #s form pseudomoms of a deg $\leq 2t$ pseudo-distn if & only if M_t (the moment matrix) is positive semidefinite.

Corollary [deg 2t Sum-of-Squares]

Let p be a deg $\leq t$ polynomial.

Then,
$$\max_{\substack{D:\ \text{pseudo-dist}^n \\ \text{on } \{-1,1\}^n \text{ of} \\ \deg = 2t}} \tilde{\mathbb{E}}\, p(x)$$
is reducible to semidefinite programming in time $n^{O(t)}$.

Observe: Since every distⁿ is a pseudo-distⁿ,

$$\max_{\substack{D \\ \text{p.d. of deg } 2t \\ \text{on } \{-1,1\}^n}} \tilde{\mathbb{E}}_D p \geq \max_{\substack{D \\ \text{dist}^n \text{ on } \{-1,1\}^n \\ \text{of deg } 2t}} \tilde{\mathbb{E}}_D p = \max_{x \in \{\pm 1\}^n} p(x)$$

So deg 2t SoS program for maximizing $\tilde{\mathbb{E}} p$ is a relaxation of the problem of maximizing p over $\{-1,1\}^n$.

Example:

Let $p = \frac{1}{2} - \frac{1}{2} \sum_{\{i,j\} \in E} x_i x_j$

for E = edge set of graph G on $[n]$

Then, the deg 2 SoS relaxation

is
$$\max \ \tilde{\mathbb{E}}_D \, p(x)$$
$$\text{s.t.} \quad M_{2,D} \succcurlyeq 0$$

$$\iff \quad \frac{1}{2} - \frac{1}{2} \sum_{\{i,j\} \in E} M_{2,D}(i,j)$$
$$\text{s.t.} \quad M_{2,D} \succcurlyeq 0$$

Goemans Williamson SDP!

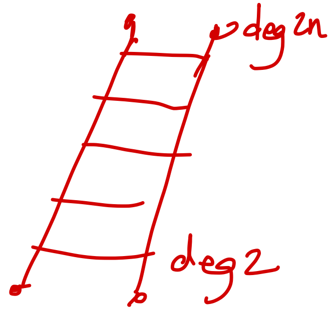
Thus GW SDP is simply the

"deg 2" SoS relaxation for Max-Cut.

OTOH, deg 2n SoS relaxation

$\equiv$ "integral" problem

"exact" problem



Increasing degree $\rightarrow$ getting tighter relaxations at the cost of larger running times.

But any O(1) deg $\equiv$ poly time.

"Mechanical" way to generate stronger relaxations

Constrained Pseudo-distribution

A deg $2t$ pseudo-dist^n D on $\{-1,1\}^n$ satisfies a constraint $q(x) \geq 0$ if

$$\tilde{\mathbb{E}} \cdot p^2 \cdot q \geq 0 \quad \forall\, p \text{ s.t.}$$

$$2\deg(p) + \deg(q) \leq 2t.$$

Corollary: For any $\underbrace{p \;\&\; q_1 \cdots q_m}_{\text{polynomials}}$, the task of finding a deg $2t$ p-d on $\{-1,1\}^n$ that max $\tilde{\mathbb{E}}\, p$ while satisfying $\{q_i \geq 0\}_{i=1}^m$ is reducible to an SDP of size $\text{poly}(m, n^{O(t)})$ in time $\text{poly}(m, n^{O(t)})$.

Low degree Transformations

Suppose D is a pseudo-dist^n on $\{-1,1\}^n$ of deg $2t$.

Define a pseudo-dist^n D' on $\{0,1\}^n$ as follows:

$$\tilde{\mathbb{E}}_{D'} X_i = \tilde{\mathbb{E}}_D \left(\frac{1+X_i}{2}\right)$$

$$\tilde{\mathbb{E}}_{D'} X_S = \tilde{\mathbb{E}}_D \prod_{i \in S} \left(\frac{1+X_i}{2}\right)$$

and extend linearly.

Then, D' is a pseudo-dist^n of deg $2t$ on $\{0,1\}^n$

More generally, holds for "low-deg poly transformations". Often referred to as the "basis invariance" of SoS relaxations.

Local Distributions

Lemma: Suppose D is a pseudo-dist^n on $\{-1,1\}^n$ of deg $\geq 2t$. Consider the restriction of D to any $\leq t$ vars S. Then, there's a prob dist^n μ on $\{-1,1\}^S$ s.t. $\forall$ $T \subseteq S$, $\tilde{\mathbb{E}}_D X_T = \mathbb{E}_\mu X_T$.

(D locally agrees with actual prob distributions)

Lemma: Suppose D is a p.d. on $\{0,1\}^n$ of deg ≥ 2. then $\tilde{\mathbb{E}} X_i \geq 0$ $\forall$ X.

Proof: $\tilde{\mathbb{E}} X_i = \tilde{\mathbb{E}} X_i^2 \geq 0$.

algo for Knapsack.

In particular, $\forall \varepsilon > 0$, can get $(1+\varepsilon)$ approx. map to knapsack in $n^{O(1/\varepsilon)}$ time

Needs 2 basic simple facts about p.d.

<u>Lemma</u>: Suppose $\forall x \in \{0,1\}^n$,

$$\sum c_i x_i \le C \Rightarrow \sum_i x_i \le k.$$

e.g. $c_1 = c_2 = 1$

$C = 2.9$, $k = 2$

$x_1 + x_2 \le 2.9 \Rightarrow x_1 + x_2 \le 2$

Then, for every pseudo-distn of deg $\ge 2k+2$ satisfying $\sum c_i x_i \le C$, it holds that $\tilde{\mathbb{E}} x_S = 0 \ \forall \ |S| \ge k+1$.

Why is this true if p.d. is a distⁿ?

Under the hyp, any distⁿ on X, restricted to S, must have support of size $\leq k$.

Thus, the prob that $X_S = 1$ $\forall\, |S| > k$ must be 0.

Let's now prove it for p.d.s.

All we will use here is that locally on all small sets ($\leq k+1$), the p.d. is in fact an actual distⁿ.

Proof: Suppose S is such that $\tilde{\mathbb{E}}_D x_S > 0$ and $|S| \geq k+1$. Consider local dist^n on S, say μ. Then, it must have $x_i = 1 \ \forall i \in S$ in its support. But for any such x, $\sum_{i \in S} c_i x_i > C$.

Now $\tilde{\mathbb{E}}_D \sum c_i x_i$

$$= \underbrace{\tilde{\mathbb{E}}_D \sum_{i \in S} c_i x_i}_{> C} + \underbrace{\tilde{\mathbb{E}}_D \sum_{i \notin S} c_i x_i}_{\geq 0}$$

$> C$ $\rightarrow$ D cannot satisfy $\sum c_i x_i \leq C$

Lemma:

Suppose D is a p.d. on $\{0,1\}^n$ of deg $\geq 2k+2$ s.t. $\forall\ |S| \geq k+1$, $\tilde{\mathbb{E}}\, x_S = 0$. Then, there's a global dist^n μ on $\{0,1\}^n$ s.t. $\tilde{\mathbb{E}}_D\, x_S = \mathbb{E}_\mu\, x_S\ \forall\ S$.

(That is D is an actual prob dist^n).

Proof: $\forall\ |S| \leq 2k+2$, set $\mathbb{E}_\mu\, x_S = \tilde{\mathbb{E}}_D\, x_S$

$\forall\ |S| > 2k+2$, set $\mathbb{E}_\mu\, x_S = 0$.

enough to prove that

$$\mathbb{E}_\mu\left[f(x)^2\right] \geq 0 \quad \forall f : \{0,1\}^n \to \mathbb{R}.$$

Write $f(x) = \Big(\underbrace{\sum_{|S| \leq 2k+2} f_S \cdot x_S}_{\text{Small}} + \underbrace{\sum_{|S| > 2k+2} f_S \cdot x_S}_{\text{large}} \Big)$

Squaring & taking $\mathbb{E}_\mu$:

$$\mathbb{E}_\mu f(x)^2 = \mathbb{E}_\mu f_{small}^2 \nearrow \geq 0 \text{ since } = \tilde{\mathbb{E}}_D f_{small}^2$$
$$+ \mathbb{E}_\mu f_{large}^2 \nearrow = 0$$
$$+ \mathbb{E}_\mu f_{large} f_{small} \;\; /\!/ \; 0$$

<u>Theorem:</u> $c_1 — c_n$ Knapsack Instance.
$v_1 — v_n$

For every p.d. D on $x_1 \dots x_n \in \{\pm 1\}$ of deg $\geq 2t$ satisfying $\sum x_i c_i \leq C$,

$$\tilde{\mathbb{E}}_D \sum_i v_i x_i \leq \left(1 + \frac{1}{t-1}\right) \text{OPT}$$

Thus, Integ gap of $n^{O(t)}$ time deg $2t$ SoS SDP is at most $(1 + 1/t-1)$.

Pf: Let $S = \{ i \mid v_i \geq \frac{OPT}{(t-1)} \}$.

(large value items).

Then, clearly, if $\sum_{i \in S} x_i \geq t$

then $\sum_i c_i x_i > C$

(as otherwise can collect value $>$ OPT)

So, $\sum_i c_i x_i \leq C$

$$\Rightarrow \sum_{i \in S} x_i \leq t-1.$$

Thus, for any p.d. D of deg $\geq 2t$

sat $\sum_i c_i x_i \leq C$, it must hold

that $\tilde{\mathbb{E}}_D \, X_T = 0 \;\; \forall \; T \subseteq S$ of size $|T| \geq t$

Thus, D restricted to S is a global dist^n. Note that $|S|$ can be $\gg 2t$ and this still holds $\rightarrow$ so this is not the vanilla local p.d. property:

Let $D_S = \{(P_U, U)\}_{U \subseteq S}$

$P_U \hookrightarrow \Pr_{D_S}[U]$

Let $f_{S,U}(x) = \begin{cases} 1 & \text{if } x_i = 1 \ \forall i \in U \ \& \ x_i = 0 \ \forall i \in S \setminus U \\ 0 & \text{o/w} \end{cases}$

Then, $\sum_{U \subseteq S} f_{S,U}(x) = 1.$

<u>Lemma:</u> $\forall\, T \subseteq S : |T| \le 2t-1,\ \forall i \notin S$

$$\tilde{\mathbb{E}} \cdot x_T \cdot x_i = 0.$$

<u>Proof:</u> Let $T_1 \subseteq T$ be of size t.

Then $\tilde{\mathbb{E}} \cdot X_T \cdot X_i = \tilde{\mathbb{E}} \cdot X_{T_1} \cdot X_{T \setminus T_1} \cdot X_i$

$$\le \sqrt{\tilde{\mathbb{E}} \cdot X_{T_1}^2} \cdot \sqrt{\tilde{\mathbb{E}} \cdot X_{T \setminus T_1}^2 X_i}$$

$$\overset{\shortparallel}{0}$$

<u>Def (Reduced (f)) :</u> <u>In short, Red (f).</u>

For any polynomial f, $f = \sum_T f_T X_T$,

<u>define</u>: $\mathrm{Red}(f) = \sum_{T : |T \cap S| \le t} f_T \cdot X_T + \sum_{\substack{T: \\ |T \cap S| > t}} f_T \cdot X_T$

$\mathrm{Red}(f_{S,u})$ has __ —.

<u>Intuition</u>: can replace $f_{S,u}$ by $\mathrm{Red}(f_{S,u})$ - - .

<u>Observation</u>: Red is a linear operation

<u>Lemma</u>: $\sum_{u \subseteq S} \mathrm{Red}(f_{S,u}) = 1.$

<u>Proof</u>: $\sum_{u \subseteq S} \mathrm{Red}(f_{S,u}) = \mathrm{Red}\left(\sum_{u \subseteq S} f_{S,u}\right)$

$= \mathrm{Red}(1) = 1$

<u>Lemma</u>: $\sum_{u \subseteq S} \mathrm{Red}(f_{S,u}) \cdot x_i = x_i \quad \forall i$

<u>Proof</u>

$$\mathrm{Red}\left(\sum_{u \subseteq S} f_{S,u} \cdot x_i\right)$$

$$= \sum_{u \subseteq S} \mathrm{Red}(f_{S,u}) \cdot x_i$$

$$= 1 \cdot x_i = x_i$$

Define: $y_i^u = \dfrac{\tilde{\mathbb{E}} \cdot x_i \cdot \mathrm{Red}(f_{S,u})}{\tilde{\mathbb{E}} \cdot \mathrm{Red}(f_{S,u})}$.

Intuition: "y_i^u is the conditional expectation of x_i conditioned on $f_{S,u}=1$."

Lemma: ① $0 \le y_i^u \le 1 \quad \forall i$

② $\sum y_i^u \cdot c_i \le C - \sum_{i \in U} c_i^u$

for every $U \subseteq S$

We will use the following lemma.

Lemma: $\forall U \subseteq S$, $\tilde{\mathbb{E}} \cdot \mathrm{Red}(f_{S,u})^2 \cdot X_i = \tilde{\mathbb{E}} \cdot \mathrm{Red}(f_{S,u})^2$.

Proof:

Observe: $\quad f_{S,u}^2 x_i = f_{S,u} x_i \quad \forall i \quad \forall u \subseteq S$

Thus $\quad \mathrm{Red}(f_{S,u}^2 \cdot x_i) = \mathrm{Red}(f_{S,u} \cdot x_i)$

$$\|$$

$$\Rightarrow \quad \mathrm{Red}(f_{S,u}^2) \cdot x_i = \mathrm{Red}(f_{S,u})\, x_i$$

To show the lemma, we need:

$$\tilde{\mathbb{E}} \cdot \mathrm{Red}(f_{S,u}^2) \cdot x_i = \tilde{\mathbb{E}}\, \mathrm{Red}(f_{S,u})^2 \cdot x_i$$

<u>Note:</u> $\quad \mathrm{Red}(f_{S,u})^2 = \mathrm{Red}(f_{S,u}^2)$ + High-degree

degree $\geq t$

From lemma before $\quad \tilde{\mathbb{E}}\, x_T \cdot x_j = 0 \quad \forall\ T \subseteq S$, $|T| \geq t$.

Thus,

$$\tilde{\mathbb{E}} \cdot \text{High-degree} \cdot x_i = 0. \qquad \square$$

Proof of lemma analyzing y_i^u:

Let's prove that $y_i^u \le 1$.

If $\tilde{\mathbb{E}} \cdot Red(f_{S,u}) = 0 \Rightarrow \tilde{\mathbb{E}}\, Red(f_{S,u}) \cdot x_i$

$$\le \sqrt{\tilde{\mathbb{E}} \cdot Red(f_{S,u})^2} \cdot \sqrt{\tilde{\mathbb{E}} x_i^2}$$

$$= \sqrt{\tilde{\mathbb{E}} \cdot Red(f_{S,u})} \sqrt{\tilde{\mathbb{E}} x_i} = 0.$$

o/w:

$$y_i = \frac{\tilde{\mathbb{E}}\, Red(f_{S,u}) \cdot x_i}{\tilde{\mathbb{E}}\, Red(f_{S,u})} = \frac{\tilde{\mathbb{E}} \cdot Red(f_{S,u})^2 x_i}{\tilde{\mathbb{E}}\, Red(f_{S,u})}.$$

Using the lemma above; we have:

$$\tilde{\mathbb{E}} \cdot Red(f_{S,u})^2 \cdot (1 - x_i) = \tilde{\mathbb{E}}\, Red(f_{S,u})^2 (1 - 2x_i + x_i^2)$$

$$= \tilde{\mathbb{E}} \cdot \mathrm{Red}(f_{S,u})^2 \cdot (1-x_i)^2 \geq 0.$$

Thus $\tilde{\mathbb{E}} \cdot \mathrm{Red}(f_{S,u})^2 \geq \tilde{\mathbb{E}}\, \mathrm{Red}(f_{S,u})^2 \cdot x_i$

or $y_i^u \leq 1.$

2) Next let's prove that $y_i^u \geq 0$

$$\tilde{\mathbb{E}} \cdot \mathrm{Red}(f_{S,u}) = \tilde{\mathbb{E}}\, \mathrm{Red}(f_{S,u})^2 \geq 0$$

$$\tilde{\mathbb{E}}\, \mathrm{Red}(f_{S,u}) x_i = \tilde{\mathbb{E}}\, \mathrm{Red}(f_{S,u})^2 \cdot x_i^2$$

$$\geq 0. \qquad \square.$$

3) $\tilde{\mathbb{E}} \sum_{i \in S} x_i \cdot c_i \cdot \mathrm{Red}(f_{S,u})$

$$= \mathbb{E}_{\mu} \sum_{i \in S} c_i \cdot x_i \cdot f_{S,u} \rightarrow *$$

$$= \sum_{i \in u} c_i \cdot \tilde{\mathbb{E}}\, \mathrm{Red}(f_{S,u}).$$

OTOH,

$$\tilde{\mathbb{E}} \sum_{i\in[n]} x_i \cdot c_i \cdot \mathrm{Red}(f_{S,u})$$

$$= \tilde{\mathbb{E}} \sum_{i\in[n]} x_i\, c_i \cdot \mathrm{Red}(f_{S,u})^2$$

Now note that $\mathrm{Red}(f_{S,u})^2$ is a square of a degree t poly.

Thus, $\tilde{\mathbb{E}}_D \mathrm{Red}(f_{S,u})^2 \cdot \left(C - \sum_{i\in[n]} c_i x_i\right)$

≥ 0 pseudo-dist^n satisfying constraint!

So $C\, \tilde{\mathbb{E}}\, \mathrm{Red}(f_{S,u})^2 \geq \sum_{i\in[n]} c_i x_i \mathrm{Red}(f_{S,u})$

$\tilde{\mathbb{E}}\, \mathrm{Red}(f_{S,u})$ —— $**$

Subtracting * from **,

$$\tilde{\mathbb{E}} \sum_{i \notin S} \mathrm{Red}(f_{S,u}) \cdot x_i \cdot c_i$$

$$\leq C \cdot \tilde{\mathbb{E}}\, \mathrm{Red}(f_{S,u}) - \left(\sum_{i \in u} c_i\right) \tilde{\mathbb{E}}\, \mathrm{Red}(f_{S,u})$$

<u>So:</u> $\quad \sum y_i \cdot c_i \leq \left(C - \sum_{i \in u} c_i\right)$

Completing Knapsack Analysis

$$\tilde{\mathbb{E}} \cdot \sum_i v_i \cdot x_i$$

$$= \tilde{\mathbb{E}} \sum_{u \subseteq S} \mathrm{Red}(f_{S,u}) \cdot \sum_i v_i \cdot x_i$$

$$= \tilde{\mathbb{E}} \sum_{u \subseteq S} \Big[\sum_{i \in S} \mathrm{Red}(f_{S,u}(x)) \cdot v_i \cdot x_i + \sum_{i \notin S} \mathrm{Red}(f_{S,u}(x)) \cdot v_i \, x_i \Big]$$

Consider $y_i = \dfrac{\tilde{\mathbb{E}}\, \mathrm{Red}(f_{S,u}(x)) \cdot x_i}{\tilde{\mathbb{E}}\, \mathrm{Red}\, f_{S,u}}$

1. these are ≤ 1

② Satisfy $\sum_i c_i \cdot y_i \le (C + \sum_{i\in u} c_i^i)$

So: $\sum_{i \notin S} y_i \cdot v_i \le OPT(C - \sum_{i\in u} c_i, V|S) + \frac{OPT}{(t-1)}$.

So

$$\sum_{u \subseteq S} \left(Red(f_{S,u}) \cdot \sum_{i\in S} x_i v_i + \sum_{i\notin S} Red(f_{S,u}(x)) \cdot x_i v_i \right)$$

$$\le \sum_{u \subseteq S} \tilde{\mathbb{E}}_{Red(f_{S,u})} \left[\max_{u \subseteq S} \left(\sum_{i\in u} v_i + OPT(C - \sum_{i \in u} x_i, V|S) \right) + \frac{OPT}{t-1} \right]$$

$$\leq \left(1+\frac{1}{t-1}\right) OPT \left(\sum_{u \leq s} \tilde{\mathbb{E}}\, redof_{s,u}\right)$$

$$= \left(1+\frac{1}{t-1}\right) OPT$$