Steurer's Conjecture and Sparsest Cut (I)

Reading Group Spring '22

Alex Tung

19 July 2022

Agenda

- Problem definitions
- Lasserre Hierarchy
- Rounding Lasserre solutions
- Correlation rounding

Agenda

- Problem definitions
- Lasserre Hierarchy
- Rounding Lasserre solutions
- Correlation rounding

Sparsest Cut

- Given graph $G = (V, E)$, define sparsest cut (aka edge expansion) as

$$\phi(G) = \min_{S \subseteq V} \frac{|E(S, S^c)|}{|S| \cdot |S^c|}$$

Approximating sparsest cut

- Spectral: $O(\sqrt{\phi})$-approximation
- LP-based: $O(\log n)$-approximation [LR]
- SDP-based: $O(\sqrt{\log n})$-approximation [ARV], [AK], [Sherman]

- Fast (almost-linear time) algorithms exist

- Can we attain a better approximation ratio?

$O(1)$-approximation?

- It is UGC-hard to approximate sparsest cut to an $O(1)$ factor
- It is unknown whether $O(1)$-approximation is possible in *subexponential* time

Steurer's conjecture

Conjecture 9.2. *For every $\varepsilon > 0$, there exists positive constants $\eta = \eta(\varepsilon)$ and $\delta = \delta(\varepsilon)$ such that the following holds: For every collection of unit vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ with $\mathbb{E}_{i,j\in[n]} \left|\langle v_i, v_j\rangle\right| \leqslant n^{-\varepsilon}$, there exists two sets $S, T \subseteq \{1, \ldots, n\}$ with $|S|, |T| \geqslant \delta n$ and $\|v_i - v_j\|^2 \geqslant \eta$ for all $i \in S$ and $j \in T$.*

In words: if a set of vectors have very low *global correlation*, there are two large cores of constant distance from each other

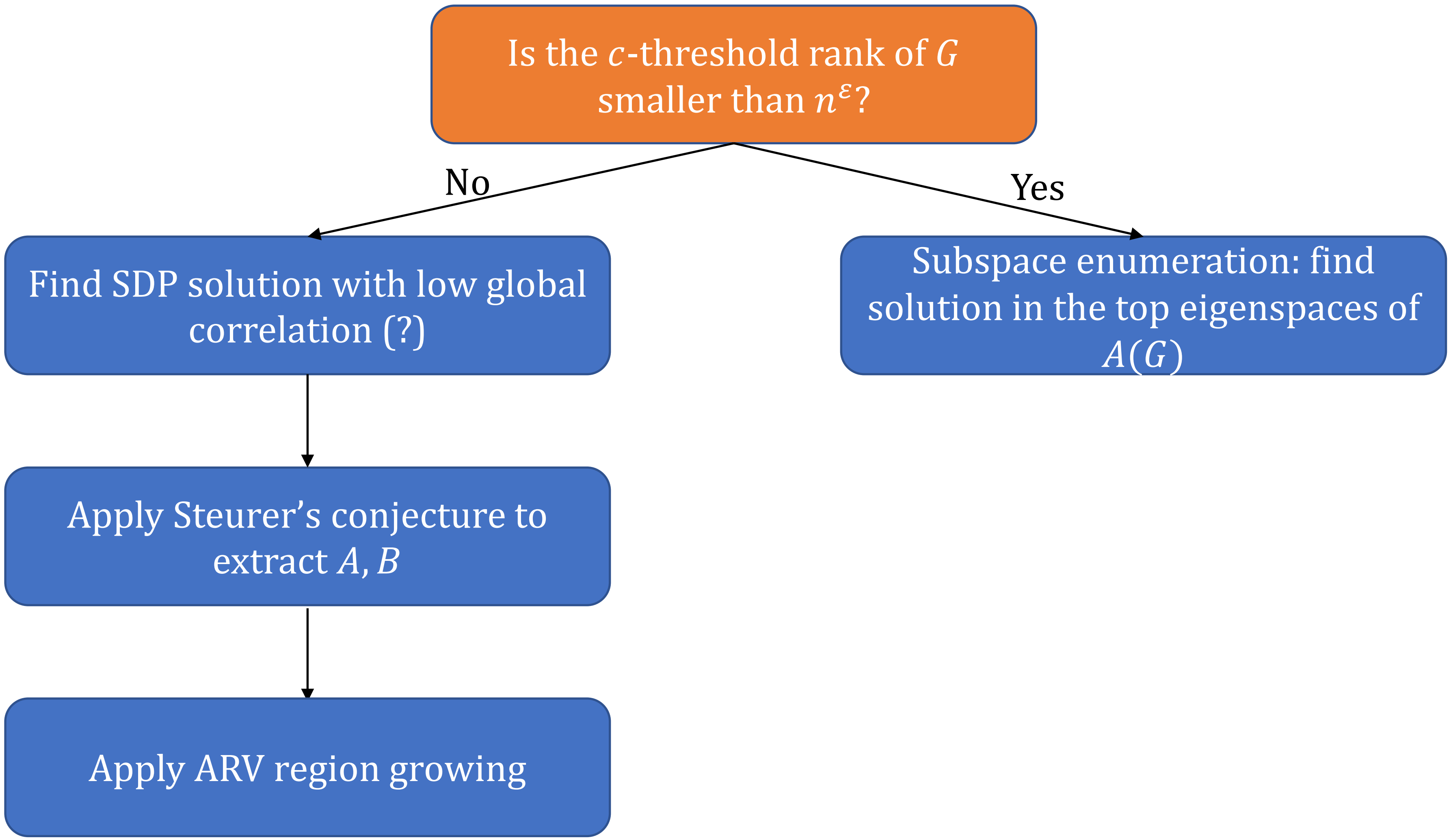
How to connect the dots?

(We omit the proof at this point. It follows from extensions of the results presented in Section 5.2, techniques in [ARV09], and an alternative characterization of the threshold rank of a graph.)

Steurer’s idea is to combine the three:

- Subspace enumeration
- ARV region growing
- Relating *threshold rank* to global correlation of vectors

Steurer's blueprint



Our goal

- To understand how Steurer's conjecture implies subexp-time algorithm for sparsest cut, via Lasserre hierarchy

- Today: introduce Lasserre hierarchy and discuss how to round Lasserre solutions

- Next week: details and proofs, complete the picture

Agenda

- Problem definitions
- Lasserre Hierarchy
- Rounding Lasserre solutions
- Correlation rounding

SDP hierarchy: setup

- Combinatorial optimization
- Objective function $g: \{0,1\}^n \to \mathbb{R}$, extensible to $g: [0,1]^n \to \mathbb{R}$
- Subject to linear constraints: $x \in K \;\coloneqq\; \{x': Ax' \geq b\}$

- Optimizing for $x \in \{0,1\}^n \cap K$ is hard in general
- Optimizing for $x \in K$ is easy, but how to round solution?

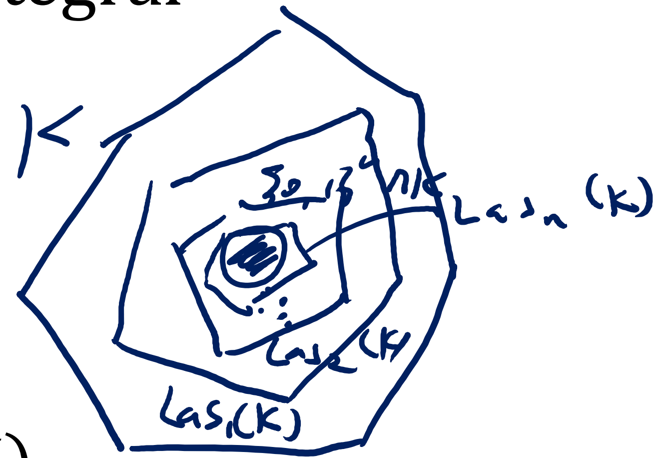
- SDP hierarchies try to balance *efficiency* and *integrality gap*

SDP hierarchy: idea

- Idea is to add additional (SDP) constraints that all integral solutions satisfy

- Finer and finer feasible region:

$$K = Las_0(K) \supseteq Las_1(K) \supseteq \cdots \supseteq Las_n(K)$$



Lasserre hierarchy

Definition 1. Let $K = \{x \in \mathbb{R}^n \mid Ax \geq b\}$. We define the *$t$-th level of the Lasserre hierarchy* $\mathrm{LAS}_t(K)$ as the set of vectors $y \in \mathbb{R}^{2^{[n]}}$ that satisfy

$$M_t(y) := (y_{I\cup J})_{|I|,|J|\leq t} \succeq 0; \quad M_t^\ell(y) := \Big(\sum_{i=1}^n A_{\ell i} y_{I\cup J\cup\{i\}} - b_\ell y_{I\cup J}\Big)_{|I|,|J|\leq t} \succeq 0 \quad \forall \ell \in [m]; \qquad y_\emptyset = 1.$$

The matrix $M_t(y)$ is called the *moment matrix of* y and the second type $M_\ell^t(y)$ is called *moment matrix of slacks*. Furthermore, let $\mathrm{LAS}_t^{\mathrm{proj}}(K) := \{(y_{\{1\}}, \ldots, y_{\{n\}}) \mid y \in \mathrm{LAS}_t(K)\}$ be the projection on the original variables.

Interpretation

- The variables x_i can be regarded as $\Pr_X[X_i = 1]$
- The variables y_I ($I \subseteq [n]$) represent joint probabilities

$$y_I = \Pr[\bigwedge_{i \in I} X_i = 1]$$

- In the t-th round of Lasserre hierarchy, y_I are only constrained for $|I| \leq 2t + 1$

- They define *local distributions*

Basic properties

$\mathrm{Las}_{t-1}(K) \supseteq \mathrm{Las}_t(K)$ for $1 \le t \le n$

- Lasserre hierarchy is an increasingly tight relaxation of the original optimization problem

  $\mathrm{Las}_t(K) \supseteq K \cap \{0,1\}^n$

- The t-th level of Lasserre hierarchy has $n^{O(t)}$ variables
- For $0 \le |I| \le |J| \le t$, we have $0 \le y_J \le y_I \le 1$

  $I \subseteq J$

Proofs

$$M_t(y) := (y_{I \cup J})_{|I|,|J| \le t} \succeq 0; \quad M_t^{\ell}(y) := \Big(\sum_{i=1}^{n} A_{\ell i}\, y_{I \cup J \cup \{i\}} - b_{\ell}\, y_{I \cup J} \Big)_{|I|,|J| \le t} \succeq 0 \quad \forall \ell \in [m]; \quad y_{\emptyset} = 1.$$

Lasserre hierarchy is an increasingly tight relaxation of the original optimization problem:

$\mathrm{Las}_{t-1}(K) \supseteq \mathrm{Las}_t(K)$:

$$M_t(y) = \begin{pmatrix} \underset{|I|,|J|<t}{M_{t-1}(y)} & \\ & \end{pmatrix} \succeq 0 \Rightarrow M_{t-1}(y) \succeq 0$$

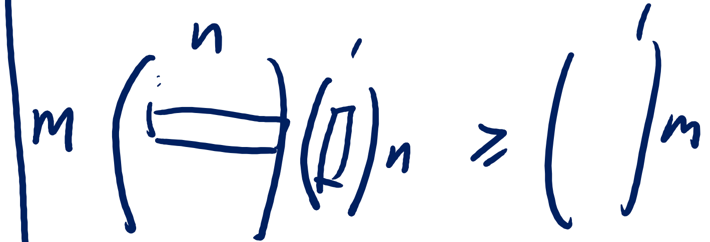
$y \in \mathrm{Las}_t(K) \Rightarrow y \in \mathrm{Las}_{t-1}(K)$.

(Almost) blank page

$$M_t(y) := (y_{I\cup J})_{|I|,|J|\le t} \succeq 0; \qquad M_t^{\ell}(y) := \left(\sum_{i=1}^{n} A_{\ell i}\, y_{I\cup J\cup\{i\}} - b_\ell\, y_{I\cup J}\right)_{|I|,|J|\le t} \succeq 0 \quad \forall \ell \in [m]; \qquad y_{\emptyset} = 1.$$

x : $x_i \in \{0,1\}$ for each i. $x \in K$.

$Ax \ge b$



$\sum A_{\ell i} x_i \ge b_\ell$

Construct $y \in Las_t(K)$

$y_i = x_i$, $y_I = ?$

$y_I := \prod_{i\in I} x_i$.

$y_{I\cup J} = \prod_{i\in I\cup J} x_i = \left(\prod_{i\in I} x_i\right)\cdot\left(\prod_{i\in J} x_i\right) = y_I y_J$

$M_t(y) = yy^T \succeq 0$.

Proofs

$$M_t(y) := (y_{I\cup J})_{|I|,|J|\le t} \succeq 0; \quad M_t^\ell(y) := \Big(\sum_{i=1}^n A_{\ell i} y_{I\cup J\cup\{i\}} - b_\ell y_{I\cup J}\Big)_{|I|,|J|\le t} \succeq 0 \quad \forall \ell \in [m]; \quad y_\emptyset = 1.$$

The t-th level of Lasserre hierarchy has $n^{O(t)}$ variables:

y_I for $|I| \le 2t+1$.

Total # of relevant variables

$\le \sum_{i \le 2t+1} \binom{n}{i} \le n^{O(t)}$.

Proofs

$$M_t(y) := (y_{I\cup J})_{|I|,|J|\le t} \succeq 0; \quad M_t^\ell(y) := \Big(\sum_{i=1}^n A_{\ell i} y_{I\cup J\cup\{i\}} - b_\ell y_{I\cup J}\Big)_{|I|,|J|\le t} \succeq 0 \quad \forall \ell \in [m]; \quad y_\emptyset = 1.$$

For $0 \le |I| \le |J| \le t$, we have $0 \le y_J \le y_I \le 1$:

$y_I \in [0,1]$: $\begin{matrix} & \phi & I \\ \phi & y_\phi & y_I \\ I & y_I & y_I \end{matrix} \succeq 0$

$\Rightarrow y_I - y_I^2 \ge 0 \Rightarrow y_I \in [0,1]$

$y_J \le y_I$ if $I \subseteq J$: $\begin{matrix} & I & J \\ I & y_I & y_J \\ J & y_J & y_J \end{matrix} \succeq 0$

$\Rightarrow y_I y_J - y_J^2 \ge 0 \Rightarrow y_I \ge y_J$

($\because y_J \ge 0$ and $y_I \ge 0$)

Geometric interpretation of $M_t(y)$

- We have $M_t(y) \coloneqq \left(y_{I \cup J}\right)_{|I|,|J| \leq t} \succcurlyeq 0$
- Meaning: there exists vectors $\{v_I\}_{|I| \leq t} \subseteq \mathbb{R}^{2^n}$ such that $\langle v_I, v_J \rangle = y_{I \cup J}$
- For example we immediately have $\|v_I\|^2 = y_I$

Agenda

- Problem definitions
- Lasserre Hierarchy
- Rounding Lasserre solutions
- Correlation rounding

Basic rounding mechanism

- Start with Lasserre solution $y \in Las_t(K)$
- $x_i = y_{\{i\}}$ are between 0 and 1
- Goal: round to an integral solution

- We don't quite know how to do that without suffering loss in objective…

“Lift and Project”

- But we can have t integral entries!
- Idea: project solution $y \in Las_t(K)$ to solution $y' \in Las_{t-1}(K)$, such that $x_i' = y_{\{i\}}' \in \{0, 1\}$

- In the end, t entries will be in $\{0, 1\}$
- Round the rest according to the marginal distribution

Projection lemma

Lemma 2. *For $t \geq 1$, let $y \in \text{LAS}_t(K)$ and $i \in [n]$ be a variable with $0 < y_i < 1$. If we define*

$$z_I^{(1)} := \frac{y_{I \cup \{i\}}}{y_i} \quad \text{and} \quad z_I^{(0)} := \frac{y_I - y_{I \cup \{i\}}}{1 - y_i}$$

$$\Pr[X_I = 1 \mid X_i = 1] = \Pr[X_I = 1 \text{ and } X_i = 1] / \Pr[X_i = 1]$$

then we have $y = y_i \cdot z^{(1)} + (1 - y_i) \cdot z^{(0)}$ with $z^{(0)}, z^{(1)} \in \text{LAS}_{t-1}(K)$ and $z_i^{(0)} = 0, z_i^{(1)} = 1$.

$\in \mathbb{R}^{2^n}$

This should remind us of *conditional probability*

Proof

Lemma 2. *For $t \geq 1$, let $y \in \text{LAS}_t(K)$ and $i \in [n]$ be a variable with $0 < y_i < 1$. If we define*

$$z_I^{(1)} := \frac{y_{I\cup\{i\}}}{y_i} \quad \text{and} \quad z_I^{(0)} := \frac{y_I - y_{I\cup\{i\}}}{1-y_i}$$

$$z_{\{i\}}^{(0)} = \frac{y_{\{i\}} - y_{\{i\}}}{1 - y_{\{i\}}} = 0$$

then we have $y = y_i \cdot z^{(1)} + (1-y_i)\cdot z^{(0)}$ with $z^{(0)}, z^{(1)} \in \text{LAS}_{t-1}(K)$ and $z_i^{(0)} = 0, z_i^{(1)} = 1$.

$$M_t(y) := (y_{I\cup J})_{|I|,|J|\leq t} \succeq 0; \quad M_t^\ell(y) := \Big(\sum_{i=1}^n A_{\ell i} y_{I\cup J\cup\{i\}} - b_\ell y_{I\cup J}\Big)_{|I|,|J|\leq t} \succeq 0 \quad \forall \ell \in [m]; \quad y_\emptyset = 1.$$

$y \in \text{Las}_t(K)$, $M_t(y) \succeq 0 \Rightarrow \exists (v_I \in \mathbb{R}^{2^n})_{|I|\leq t}$ s.t. $y_{I\cup J} = \langle v_I, v_J \rangle$.

Let's prove $z_\emptyset^{(1)} \in \text{Las}_{t-1}(K)$. . .

$\vdash$: $M_{t-1}(z^{(1)}) \succeq 0$. Want to construct $v_I^{(1)}$ s.t. $z_{I\cup J}^{(1)} = \langle v_I^{(1)}, v_J^{(1)} \rangle$.

$v_I^{(1)} := \frac{1}{\sqrt{y_i}} v_{I\cup\{i\}}$. Check: $\langle v_I^{(1)}, v_J^{(1)} \rangle = \langle \frac{1}{\sqrt{y_i}} v_{I\cup\{i\}}, \frac{1}{\sqrt{y_i}} v_{J\cup\{i\}} \rangle$

$$= \frac{1}{y_i} \cdot \langle v_{I\cup\{i\}}, v_{J\cup\{i\}} \rangle$$

$$= \frac{1}{y_i} \cdot y_{I\cup J\cup\{i\}} = z_{I\cup J}^{(1)}.$$

Remark

- The order of conditioning does not matter!

Condition on S_1 to be 1

S_0 to be 0

Agenda

- Problem definitions
- Lasserre Hierarchy
- Rounding Lasserre solutions
- Correlation rounding

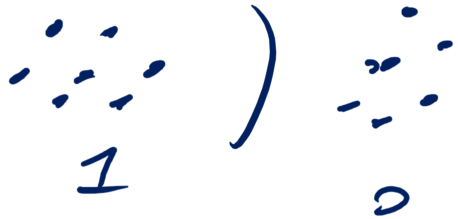
Correlation rounding

- From $y \in Las_t(K)$, we can obtain solutions with at most t integral entries
- For non-integral entries, we just use marginals to round randomly

- When does this yield a good integral solution?

Good rounding when?

- Ideal case: entries are independent
- Close to ideal case: entries are almost independent



- For graph CSP's, only (X_i, X_j) over $(i, j) \in E$ matter

- We expect rounding to be good if *local correlation* is small

Definitions

- From $y \in Las_{t+2}(K)$, we can obtain solutions with t integral entries

- Define local correlation as

$$\left|\Pr[X_i = X_j = 1] - \Pr[X_i = 1] \cdot \Pr[X_j = 1]\right|$$

$$\mathbb{E}_{(i,j)\in E}[|y_{ij} - y_i y_j|]$$

Uniform measure

- Define global correlation as

$$\mathbb{E}_{i,j\in V}[|y_{ij} - y_i y_j|]$$

Global correlation

- Use variance $\sum_i x_i(1 - x_i)$ as potential function
- If global correlation is high, can project one level down while decreasing variance by a lot

- Variance is bounded in the beginning and is always ≥ 0
- Therefore, at some point global correlation must be low

Local-to-global correlation

- We expect rounding to be good if *local correlation* is small
- At some point *global correlation* must be low

- If low global correlation -> low local correlation, then we are good!

  high local correlation → high global correlation.
- For example, expanders has this property
- In fact low threshold-rank graphs have this property too!

  λ_k is large

Back to sparsest cut…

- Suppose we have an appropriate Lasserre hierarchy for sparsest cut

- At some point *global correlation* must be low

- If we can extract a vector configuration from the respective Lasserre solution, we can apply Steurer's conjecture!

Summary

We have seen:

- What Lasserre hierarchy is
- How to round a Lasserre solution
- An overview of correlation rounding

Next time…

- Proving main results about correlation rounding
- Application to sparsest cut
- How Steurer’s conjecture -> subexp-time sparsest cut

References

- David Steurer, "On the Complexity of Unique Games and Graph Expansion"
- Thomas Rothvoss, "The Lasserre hierarchy in Approximation algorithms"
- Lap Chi Lau, "Lasserre Hierarchy", Lecture 13 of CSCI5060