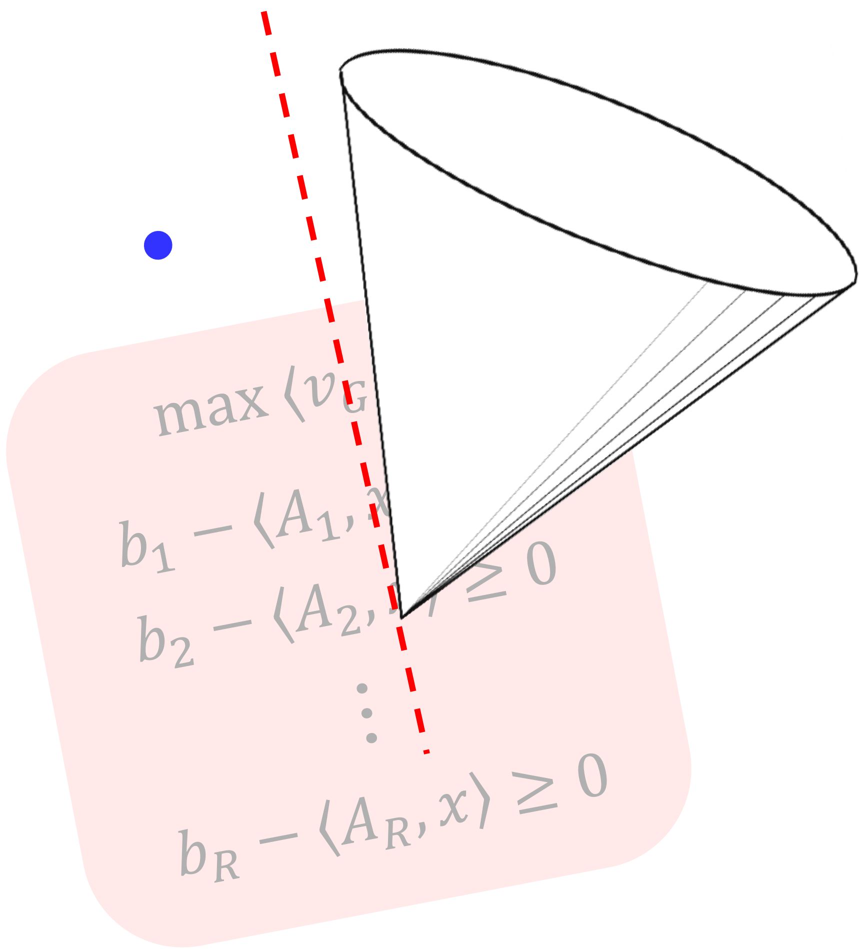
linear programming and approximate constraint satisfaction



Siu On Chan
MSR New England

James R. Lee
University of Washington

Prasad Raghavendra
UC Berkeley

David Steurer
Cornell

MAIN THEOREM:

Any polynomial-sized **linear program** for...

problem	integrality gap
MAX-CUT	$1/2$
MAX-2SAT	$3/4$
MAX-3SAT	$7/8$

holds for LPs of size

$$n^{c\frac{\log n}{\log\log n}}$$

MAIN TECHNIQUE:

For approximating MAX-CSPs, polynomial-size LPs are exactly as powerful as those arising from $O(1)$ rounds of the Sherali-Adams hierarchy.

a brief history of LP lower bounds

Specific LP hierarchies (Lovász-Schrijver, Sherali-Adams)

[Arora-Bollobás-Lovász 02]

r-round relaxation has size $n^{O(r)}$

MAX-CUT has integrality gap $1/2$ for

$\Omega(n)$ rounds of LS [Schoenbeck-Trevisan-Tulsiani 07]

$\omega(1)$ rounds of SA [Fernández de la Vega-Mathieu 07]

$n^{\Omega(1)}$ rounds of SA [Charikar-Makarychev-Makarychev 09]

Extended formulations (EF)

[Yannakakis 88] – every symmetric EF for TSP (and matching) has exponential size

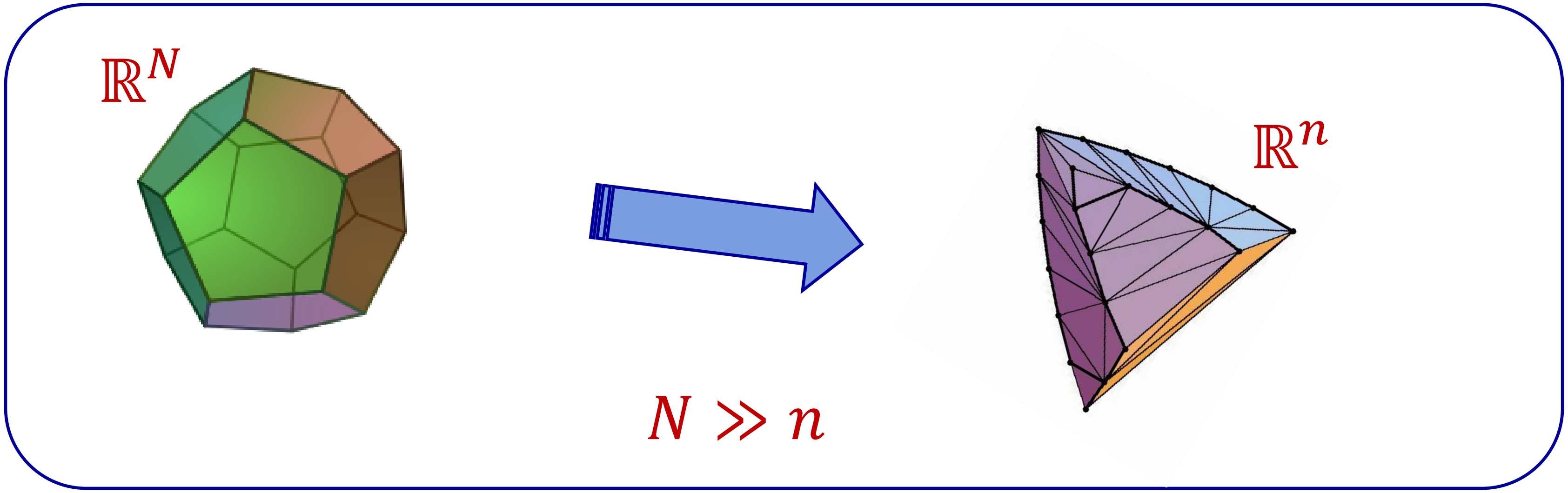
Every extended formulation for TSP has size $2^{\Omega(\sqrt{n})}$

[Fiorini-Massar-Pokutta-Tiwary-de Wolf 12]

EFs for approx. clique within $n^{\frac{1}{2}-\epsilon}$ require size $2^{n^{\epsilon}}$ [Braun-Fiorini-Pokutta-Steurer 12]

EFs for approx. clique within $n^{1-\epsilon}$ require size $2^{n^{\epsilon}}$ [Braverman-Moitra 13]

a brief history of LP lower bounds



Extended formulations (EF)

[Yannakakis 88] – every symmetric EF for TSP (and matching) has exponential size

Every extended formulation for TSP has size $2^{\Omega(\sqrt{n})}$

[Fiorini-Massar-Pokutta-Tiwary-de Wolf 12]

EFs for approx. clique within $n^{\frac{1}{2}-\epsilon}$ require size $2^{n^{\epsilon}}$ [Braun-Fiorini-Pokutta-Steurer 12]

EFs for approx. clique within $n^{1-\epsilon}$ require size $2^{n^{\epsilon}}$ [Braverman-Moitra 13]

what is a linear program for MAX-CUT?

For a graph $G = (V, E)$ and $S \subseteq V$, write $\mathrm{cut}_G(S) = \frac{|E(S,\bar{S})|}{|E|}$

so that $\mathrm{opt}(G) = \max_{S \subseteq V} \mathrm{cut}_G(S)$.

Standard relaxation:

$$\mathrm{opt}(G) = \max_{x \in \{-1,1\}^n} \sum_{i \sim j} \frac{1 - x_i x_j}{2}$$

Introduce variables $\{y_{ij}\}$ meant to represent $\left(1 - x_i x_j\right)/2$

$$\max \sum_{i \sim j} y_{ij}$$

subject to:

$$\{0 \le y_{ij} \le 1\} \qquad \{y_{ij} + y_{ik} + y_{jk} \le 2\}$$

$$\{y_{ij} \le y_{ik} + y_{jk}\} \qquad \{y_{ij} + y_{jk} + y_{k\ell} + y_{\ell h} + y_{hi} \le 4\}$$

what is a linear program for MAX-CUT?

For a graph $G = (V, E)$ and $S \subseteq V$, write $\mathrm{cut}_G(S) = \frac{|E(S, \bar{S})|}{|E|}$

so that $\mathrm{opt}(G) = \max_{S \subseteq V} \mathrm{cut}_G(S)$.

Linearization: For every n, we have a natural number m and:

- For every n-vertex graph G, a vector $v_G \in \mathbb{R}^m$
- For every cut S, a vector $y_S \in \mathbb{R}^m$

  satisfying $\mathrm{cut}_G(S) = \langle v_G, y_S \rangle$

Relaxation: A polytope $P \subseteq \mathbb{R}^m$ such that $y_S \in P$ for every cut S

The **LP value** is given by $\mathcal{L}(G) = \max_{x \in P} \langle v_G, x \rangle$

Size of the relaxation = # of inequalities needed to specify P

approximation and integrality gaps

An LP relaxation $\mathcal{L}$ is a **(c, s)-approximation** for MAX-CUT if for every graph G with $\text{opt}(G) \leq s$, we have $\mathcal{L}(G) \leq c$.

Linearization: For every n, we have a natural number m and:

- For every n-vertex graph G, a vector $v_G \in \mathbb{R}^m$
- For every cut S, a vector $y_S \in \mathbb{R}^m$

  satisfying $\text{cut}_G(S) = \langle v_G, y_S \rangle$

Relaxation: A polytope $P \subseteq \mathbb{R}^m$ such that $y_S \in P$ for every cut S

The **LP value** is given by $\mathcal{L}(G) = \max_{x \in P} \langle v_G, x \rangle$

Size of the relaxation $=$ # of inequalities needed to specify P

approximation and integrality gaps

An LP relaxation $\mathcal{L}$ is a **(c, s)-approximation** for MAX-CUT if for every graph G with $\mathrm{opt}(G) \leq s$, we have $\mathcal{L}(G) \leq c$.

For the next theorem, view cut_G as a function from $\{-1,1\}^n$ to $[0,1]$.

THEOREM [Yannakakis via Farkas]: If there exists an LP relaxation of size R that is a (c, s)-approximation, then there are non-negative functions

$$q_1, q_2, \ldots, q_R : \{-1,1\}^n \to \mathbb{R}_+$$

such that for every graph G with $\mathrm{opt}(G) \leq s$, there exists $\lambda_1, \lambda_2, \ldots, \lambda_R \geq 0$ satisfying

$$c - \mathrm{cut}_G = \lambda_1 q_1 + \lambda_2 q_2 + \cdots + \lambda_R q_R$$

approximation and integrality gaps

THEOREM [Yannakakis via Farkas]: If there exists an LP relaxation of size R that is a (c, s)-approximation, then there are non-negative functions

$$q_1, q_2, \dots, q_R : \{-1,1\}^n \to \mathbb{R}_+$$

such that for every graph G with $\mathrm{opt}(G) \leq s$, there exists $\lambda_1, \lambda_2, \dots, \lambda_R \geq 0$ satisfying

$$c - \mathrm{cut}_G = \lambda_1 q_1 + \lambda_2 q_2 + \cdots + \lambda_R q_R$$

$$\max \langle v_G, x \rangle$$

$$b_1 - \langle A_1, x \rangle \geq 0$$

$$b_2 - \langle A_2, x \rangle \geq 0$$

$$\vdots$$

$$b_R - \langle A_R, x \rangle \geq 0$$

$$q_i(S) = b_i - \langle A_i, y_S \rangle$$

Farkas' Lemma says that every linear inequality valid for the polytope $\boldsymbol{P}$ can be derived from a non-negative combination of the defining inequalities.

Apply to the valid inequality

$$c - \langle v_G, x \rangle \geq 0$$

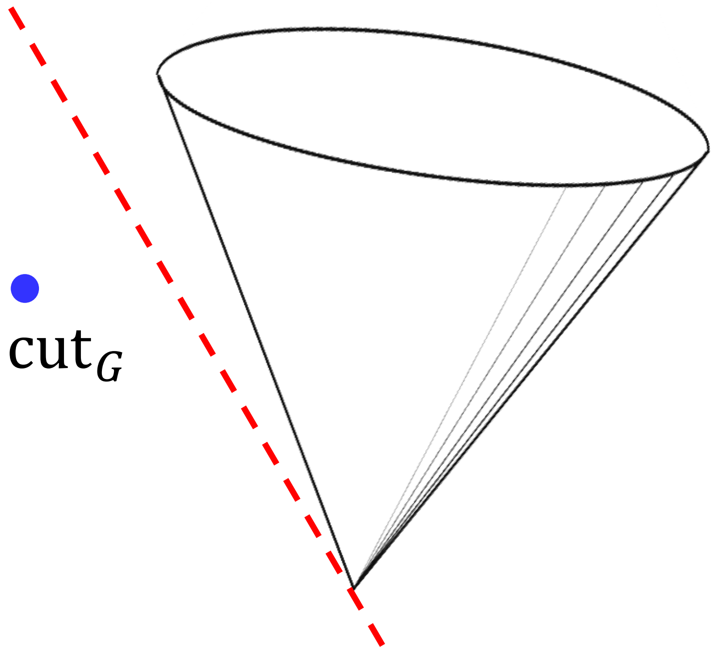
lower bounds via separating hyperplanes

THEOREM [Yannakakis via Farkas]: If there exists an LP relaxation of size R that is a (c, s)-approximation, then there are non-negative functions

$$q_1, q_2, \dots, q_R : \{-1,1\}^n \to \mathbb{R}_+$$

such that for every graph G with $\mathrm{opt}(G) \le s$, there exists $\lambda_1, \lambda_2, \dots, \lambda_R \ge 0$ satisfying

$$c - \mathrm{cut}_G = \lambda_1 q_1 + \lambda_2 q_2 + \cdots + \lambda_R q_R$$



$c - \mathrm{cut}_G$

Find a graph G and a hyperplane H such that:

$\langle H, q_i \rangle \ge 0$ for $i = 1, 2, \dots R$,

but $\langle H, c - \mathrm{cut}_G \rangle < 0$

$$H : \{-1,1\}^n \to \mathbb{R}$$

the sherali-adams hierarchy

THEOREM [Yannakakis via Farkas]: If there exists an LP relaxation of size R that is a (c, s)-approximation, then there are non-negative functions

$$q_1, q_2, \dots, q_R : \{-1,1\}^n \to \mathbb{R}_+$$

such that for every graph G with $\mathrm{opt}(G) \le s$, there exists $\lambda_1, \lambda_2, \dots, \lambda_R \ge 0$ satisfying

$$c - \mathrm{cut}_G = \lambda_1 q_1 + \lambda_2 q_2 + \cdots + \lambda_R q_R$$

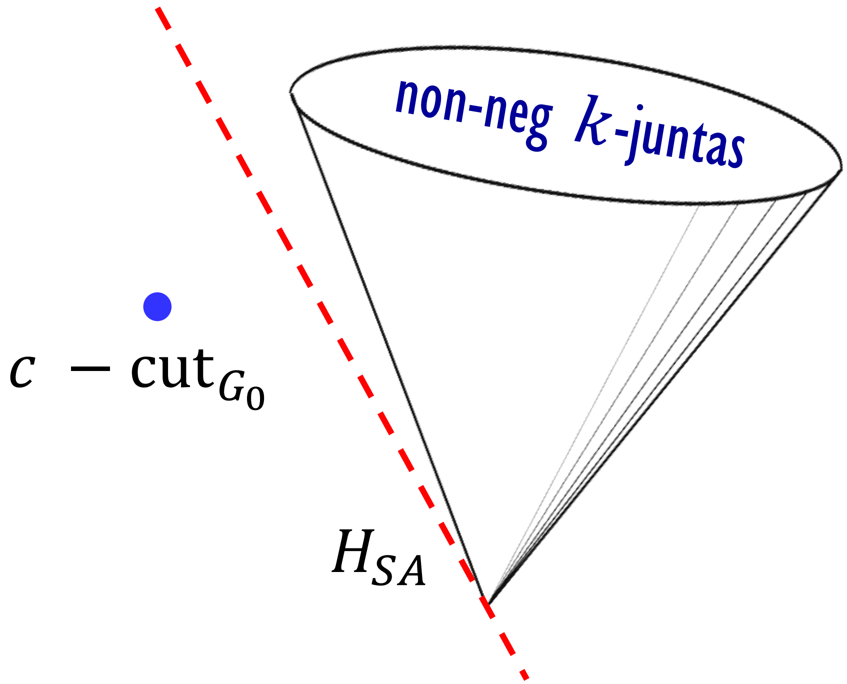
A function $q : \{-1,1\}^n \to \mathbb{R}$ is a ***k*-junta** if it only depends on k of its input coordinates.

k rounds of **Sherali-Adams** corresponds to the case when all the q_i's are k-junta's, i.e.

$$c - \mathrm{cut}_G \in \mathrm{cone}(\text{non-negative } k\text{-juntas})$$

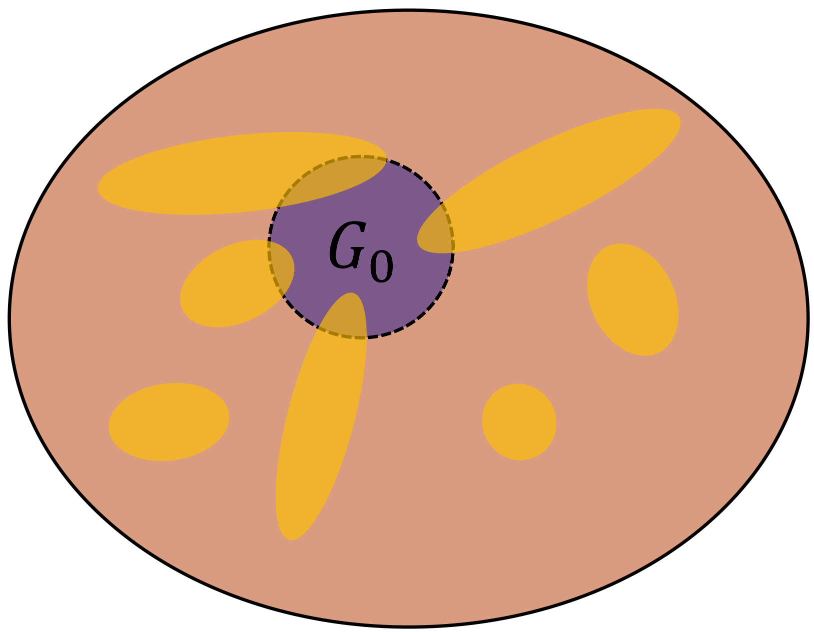
junta reduction

Let G_0 be a (c, s) gap instance for k rounds of Sherali-Adams.



$|G_0| = m$

$q_1, q_2, \ldots, q_R: \{-1,1\}^n \to \mathbb{R}_+$
are $n^{0.2}$-juntas



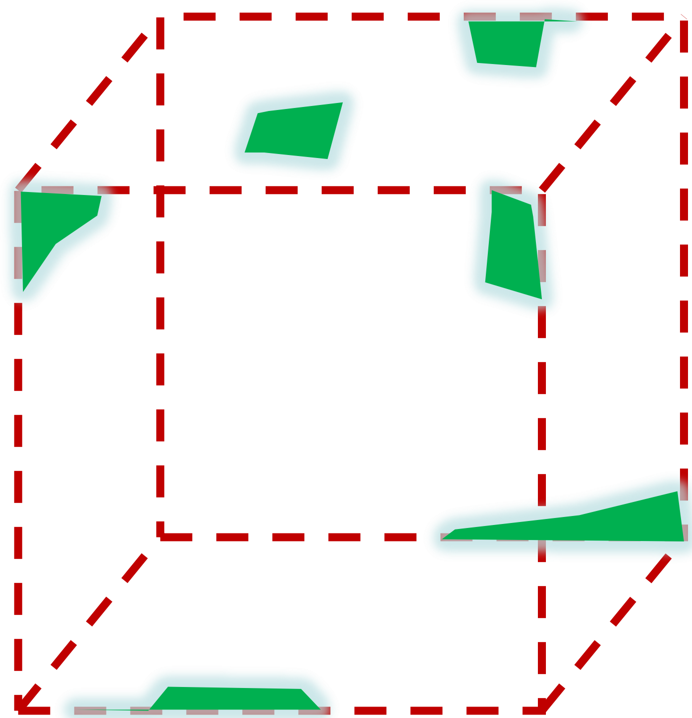
$|G| = n \gg m \sim \sqrt{n}$

Works for $R \sim n^{0.3k}$

smoothing the $q_i's$

Normalize $q_1, q_2, \dots, q_R : \{-1,1\}^n \to \mathbb{R}_+$ so that $\mathbb{E}(q_i) = 1$

Consider all the points at which $q_i(x) > R^2$ for some i



By Markov's inequality, total measure of such points is $< \frac{1}{R}$

Zero out the separating functional H on these points.

Uses: $||H_{SA}||_\infty$ small

structure lemma

LEMMA:

Suppose $q : \{-1,1\}^n \to \mathbb{R}_+$ satisfies $\mathbb{E}(q) = 1$ and $\|q\|_\infty < R^2$.

Then there is an $O(k(\log R)\, n^{0.2})$-junta q' such that every degree-k Fourier coefficient of $q - q'$ is at most $n^{-0.1}$

Tells us nothing about the high-degree Fourier coefficients of $q - q'$.

That's OK. The Sherali-Adams(k) functional $H_{SA} : \{-1,1\}^n \to \mathbb{R}$ is degree-k (as a multi-linear polynomial).

LEMMA:

Suppose $q : \{-1,1\}^n \to \mathbb{R}_+$ satisfies $\mathbb{E}(q) = 1$ and $\|q\|_\infty < R^2$.

Then there is an $O(k(\log R)\, n^{0.2})$-junta q' such that every degree-k Fourier coefficient of $q - q'$ is at

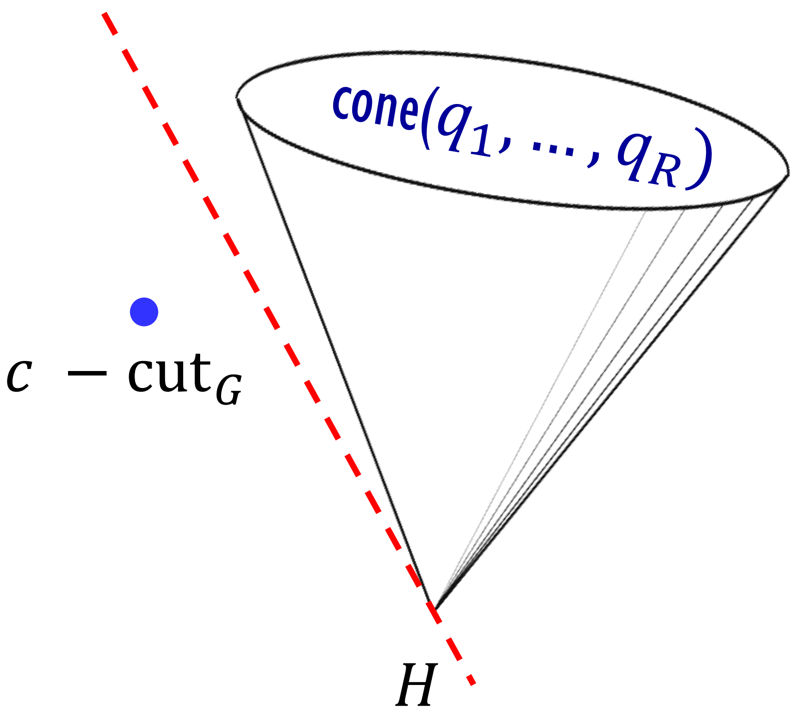
Tells us nothing about the high-degr

That's OK. The Sherali-Adams(k) fur degree-k (as a multi-linear polynom

$q - q'$

recap

$q_1, q_2, \dots, q_R : \{-1,1\}^n \to \mathbb{R}_+$

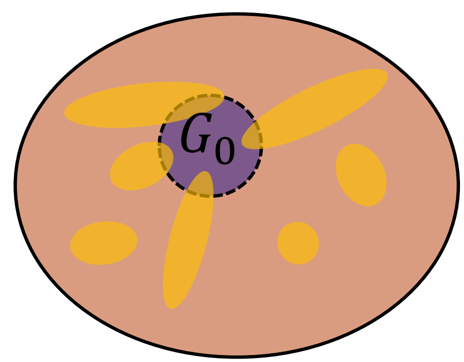


$H : \{-1,1\}^n \to \mathbb{R}$

(i) By zeroing H on a small set, can assume that $\mathbb{E}(q_i) = 1$ and $\|q_i\|_\infty < R^2$

(ii) Every such q_i can be approximated by an $n^{0.2}$-junta q_i' so that $q_i - q_i'$ has small degree-k Fourier coefficients.

(iii) When randomly planting G_0, each q_i' becomes a k-junta on the support of G_0



(iv) The Sherali-Adams functional H_{SA} is degree-k. Cannot see the high-degree discrepancy between q_i and q_i'.

future directions

- For CSPs, does the connection between Sherali-Adams(k) and general LPs hold for $k \sim n^{\epsilon}$?
- Can our method be extended beyond CSPs? (TSP, Vertex Cover, ...)
- Can it be used to resolve the long-standing open problem: Do there exist polynomial-size extended formulations of the perfect matching polytope?
- Is there a similar connection between SDPs and the Lasserre hierarchy?

thanks.