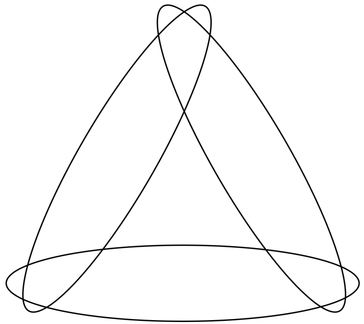
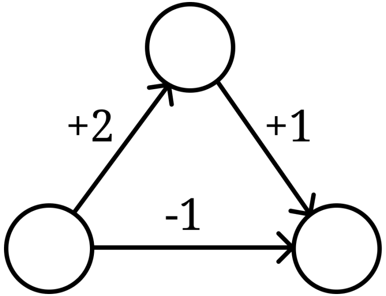
Sum-of-Squares for Unique Games

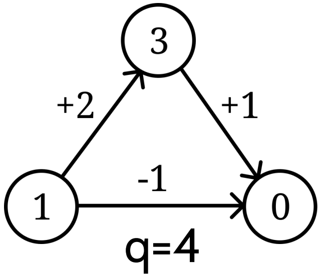
$$\widetilde{\mathbb{E}}[f^2] \geq 0$$





Chris Jones
University of Chicago
Advisor: Aaron Potechin
Committee: László Babai, Aaron Potechin, Madhur Tulsiani

Unique Games



Unique Games (UG) problem: Fix constant q. Given (G, Π) where G is a directed graph, $\Pi = (\pi_e)_{e \in E}$ is a permutation of [q] for each edge,

$$\text{maximize over } x_u \in [q] \qquad \mathbb{E}_{e=(u,v)\in E}\, 1[x_v = \pi_e(x_u)]$$

Unique Games Conjecture (UGC): For all ε, s > 0, there is q sufficiently large such that it is NP-hard to distinguish between: (G, Π) has value ≥1-ε or value ≤ s.

Lemma. WLOG constraints are *affine*, undirected, and the graph is d-regular.

"Solve UG" = when the input is (1-ε) satisfiable, find a solution with value $\Omega_\varepsilon(1)$

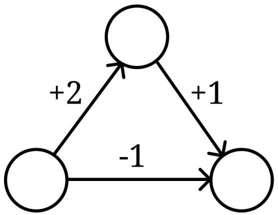
□ Drop the parameter s from here on out and assume we are given (1-ε) satisfiable (G, Π), where ε is a tiny constant

Sum-of-Squares

Our most effective algorithm for Unique Games is the Sum-of-Squares algorithm

Sum-of-Squares can be used to maximize a polynomial system

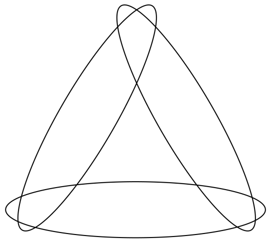
Sum-of-Squares (SoS_D) algorithm: given "degree" D, search for a pseudoexpectation $\tilde{\mathbf{E}}$ which maximizes $\tilde{\mathbf{E}}$[objective].



$\tilde{\mathbf{E}}$ looks like a real expectation over a distribution on $\mathbb{R}^{\#\text{variables}}$ with respect to:

(1) degree D/2 local reasoning

(2) $\tilde{\mathbf{E}}[p(X)^2] \geq 0$ for all degree ≤D/2 polynomials p



$\tilde{\mathbf{Pr}}$ denotes the local probability distribution, e.g. $\tilde{\mathbf{Pr}}[X_i = a]$

Sum-of-Squares for Unique Games

Given (G, Π) where G is a directed graph, $\Pi = (c_e)_{e\in E}$ is an affine shift for each edge, maximize over $x_u \in [q]$ the fraction of satisfied edges $\mathbb{E}_{e=(u,v)\in E}\, 1[x_v = x_u + c_e]$.

Variables: X_{ua} for each $u\in V$, $a\in[q]$ — X_{ua} indicates that u is assigned a

Constraints: $X^2_{ua} = X_{ua}$ — Boolean variables, X_{ua} in {0,1}

$\Sigma_a X_{ua} = 1$ — Exactly one label per vertex

Objective: $\mathbb{E}_{e=(u,v)\in E}\, \Sigma_a X_{ua} X_{v,a+c}$

Run SoS_D to produce $\tilde{E}$ for the above system. $\tilde{E}$ is a *fake* distribution of solutions, which has pseudo-expected value at least $(1-\varepsilon)$.

Our goal is to design a rounding algorithm to "sample" from $\tilde{E}$ a *real* solution with value $\Omega_\varepsilon(1)$

How does SoS perform on UG?

Let (G, Π) be a UG instance with value at least (1-ε).

Theorem [BRS'11]. If G has threshold rank r, then rounding $SoS_{O(r^2)}$ solves UG

Theorem [BRS'11]. For general G, rounding $SoS_{n^{O(\varepsilon)}}$ solves UG

Theorem [BBKSS'21]. If G is a D-certifiable small set expander, then rounding $SoS_{O(D)}$ solves UG

Theorem [BBKSS'21]. If G is the Johnson graph, then rounding $SoS_{O(1)}$ solves UG

Open: does rounding SoS_4 solve UG?

I. Rounding low threshold rank

Theorem [BRS'11]. If G has threshold rank r, then rounding $\text{SoS}_{O(r^2)}$ solves UG

Threshold Rank

Given a d-regular graph G and a set of vertices S ($|S| \leq n/2$), the expansion of S is

$$\Phi_G(S) = E(S, V \setminus S) / d \times |S| = \Pr[\text{1-step walk leaves } S]$$

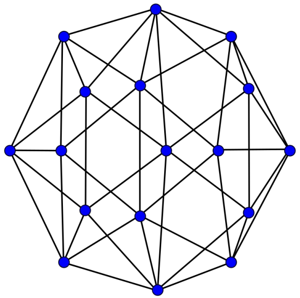
The spectrum of G are the eigenvalues of the normalized adjacency matrix A/d

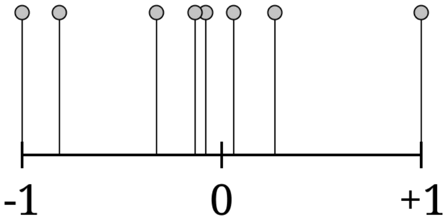
- The spectrum is a subset of [-1, +1] of size n. Recall that +1 is always an eigenvalue.

The threshold rank $\text{rank}_\tau(G)$ is the number of eigenvalues bigger than τ

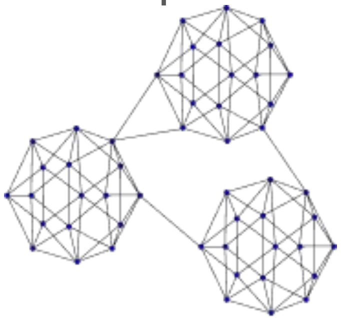
- We will always use constant τ, such as τ = 1-poly(ε)

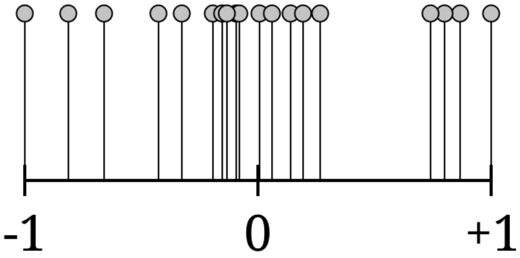
Example: expanders have $\text{rank}_{\Omega(1)}(G)=1$





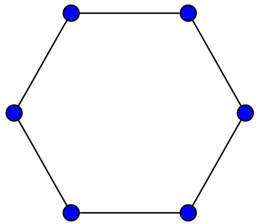
Ex: k expanders+few edges has $\text{rank}_{\Omega(1)}(G)=k+1$

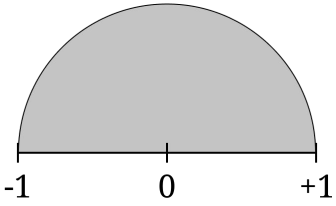




Threshold Rank

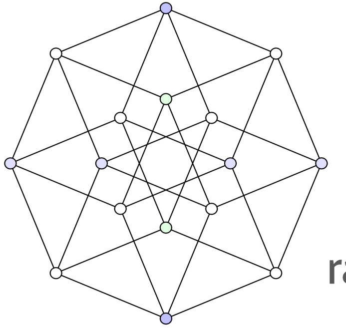
Example: cycle graph C_n

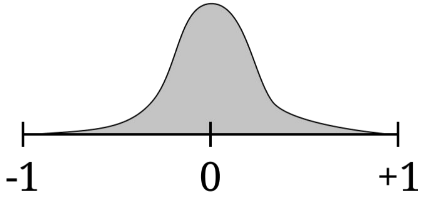




$\text{rank}_\tau(C_n) = \Theta(n)$

Example: Boolean Hypercube $\{-1,+1\}^n$





$\text{rank}_\tau(G) = C^n = N^\varepsilon$

All dense graphs have low threshold rank:

Lemma. Any d-regular graph G with $d = pn$ has $\text{rank}_\tau(G) \leq p/\tau^2 = O(1)$

Proof. $\sum_i \lambda_i^2 = \text{tr}((A/d)^2) = \sum_v \Pr[\text{2-step walk returns to } v] = n/d = O(1)$. Therefore at most $O(1)$ eigenvalues are bigger than τ. ■

Correlation Rounding on Low Threshold Rank Graphs

Theorem [BRS'11]. If G has $(1-\varepsilon^5)$-threshold rank r, then rounding $SoS_{O(r^2)}$ solves UG

Idea: we wish that one of these two rounding schemes worked:

for each v, sample v independently from its local distribution — Edges unsatisfied!

for each e = (u,v), sample (u,v) according to its local distribution — Not consistent!

Key observation: in a low threshold rank graph, after conditioning on a small number of randomly selected vertices, these become close (in total variation distance)!

We call this procedure "condition and round"

- Formally, for a random set S of size $O(r^2)$, sample an assignment X_S from the local distribution on S, then sample the assignment to u from the conditioned local distribution $\tilde{\mathbf{Pr}}[X_u \mid X_S]$.
  These distributions exist provided the SoS degree is at least |S|+1

Correlation Rounding on Low Threshold Rank Graphs

Theorem [BRS'11]. If G has $(1-\varepsilon^5)$-threshold rank r, then rounding $SoS_{O(r^2)}$ solves UG

Proof. We prove that condition+round on $O(r^2)$ random vertices works

Theorem [Raghavendra-Tan '11]. Given any boolean-valued random variables $X_1, \ldots, X_n$ there is $S \subseteq [n]$, $|S| \le O(r^2)$ such that $\mathbb{E}_{i, j \in [n]}[TV(X_i, X_j) \mid X_S] \le 1/r$

Theorem. If $\mathbb{E}_{(i,j) \in E}\, [TV(X_i, X_j)] \ge 1-2\varepsilon$, then $\mathbb{E}_{i,j \in V}\, [TV(X_i, X_j)] \ge poly(\varepsilon)/rank_{1-poly(\varepsilon)}(G)$

After conditioning, we may conclude that $\mathbb{E}_{(i,j) \in E}\, [TV(X_i, X_j)] \le 1-2\varepsilon$.

Looking at the event "edge (i, j) is satisfied", we have:

$$\mathbb{E}_{\text{round vertices independently}}\, \mathbb{E}_{(i,j) \in E}\text{value} \ge \mathbb{E}_{\text{round edges independently}}\, \mathbb{E}_{(i,j) \in E}\text{value} - (1-2\varepsilon) \ge \varepsilon \quad \blacksquare$$

1. *Low threshold rank*

II. General graphs in subexponential time

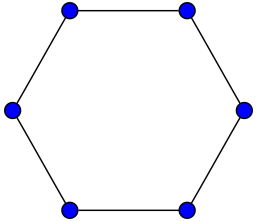
Theorem [BRS'11]. If G has threshold rank r, then rounding $SoS_{O(r^2)}$ solves UG

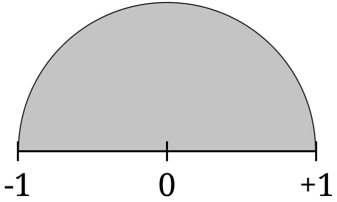
⇓

Theorem [BRS'11]. For general G, rounding $SoS_{n^{O(\varepsilon)}}$ solves UG

What about high threshold rank?

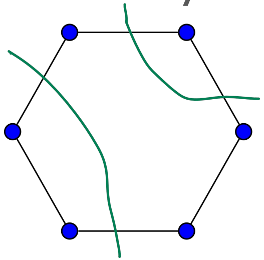
Recall: cycle graph C_n

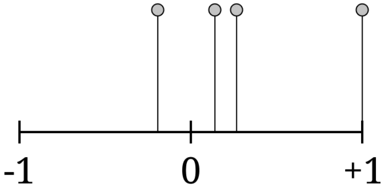




$\text{rank}_T(C_n) = \Theta(n)$

Cutting ε fraction of the edges changes the objective value by at most ε.





$\text{rank}_T(\text{piece}) = O(1)$

If you let me partition the graph by cutting $O_\varepsilon(1)$ fraction of edges, what can I do?

Lemma [ABS'10]. Any graph G can be partitioned into pieces V_i with $\text{rank}_{1-\varepsilon^5}(G[V_i]) \leq n^{100\varepsilon}$ by cutting at most $O(\varepsilon \log(1/\varepsilon))$ fraction of edges

Overall algorithm: run $\text{SoS}_{n^{100\varepsilon}}$ on the entire graph, which gives a feasible $\text{SoS}_{n^{100\varepsilon}}$ solution on each subgraph. Condition+round on each subgraph.

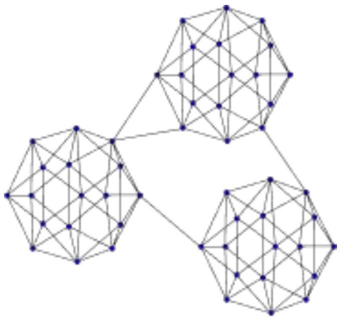
Graph Partitioning Lemma

If you let me partition the graph by cutting $O_\varepsilon(1)$ fraction of edges, what can I do?

Lemma [Arora-Barak-Steurer '10]. Any graph G can be partitioned into pieces V_i with $\mathrm{rank}_{1-\varepsilon^5}(G[V_i]) \leq n^{100\varepsilon}$ by cutting at most $O(\varepsilon \log(1/\varepsilon))$ fraction of edges

Lemma [folklore]. Any graph G can be partitioned into pieces V_i with $\Phi_G(V_i) \leq \varphi$ by cutting at most $O(\varphi \log n)$ fraction of edges

Proof idea: If G itself is a φ-expander, great!
Otherwise there is a non-expanding set S, $|S| \leq n/2$.
Partition G into S and V\S and recurse. ■



Graph Partitioning Lemma

Lemma [Arora-Barak-Steurer '10]. Any graph G can be partitioned into pieces V_i with $\text{rank}_{1-\varepsilon^5}(G[V_i]) \leq n^{100\varepsilon}$ by cutting at most $O(\varepsilon \log(1/\varepsilon))$ fraction of edges

Proof: Use the following lemma.

Lemma [ABS'10]. For a graph G with $\text{rank}_{1-\varepsilon^5}(G[V_i]) > n^{100\varepsilon}$, we can find a subset S with $|S| \leq n^{1-\varepsilon}$ and $\Phi_G(S) \leq \varepsilon^2$

Recursive apply the lemma to bad pieces or until $|V_i| \leq n^{\varepsilon}$

After k subdivisions, piece has size $n^{(1-\varepsilon)^k}$. Therefore each piece is subdivided at most $k = O(\log(1/\varepsilon)/\varepsilon)$ times. Total fraction of edges cut = $\varepsilon^2 k = O(\varepsilon \log(1/\varepsilon))$ ■

III. Certified Small Set Expanders

Theorem [BBKSS'21]. If G is a D-certifiable small set expander, then rounding $SoS_{O(D)}$ solves UG

Small Set Expansion

G is a (δ, η)-small set expander (SSE) if for all $|S| \le \delta n$, $\Phi_G(S) \ge \eta$

- δ, η are fixed small constants while val(G) ≥ 1-ε where $\varepsilon \ll \delta,\eta$

Idea for rounding SoS on a small set expander:

Recall that **Ẽ** gives access to a claimed distribution of high-value solutions on G.

Suppose we sample *two independent* high-value solutions X, X'.

We claim that in a SSE these solutions will have significant overlap.

Define the (random vbl) shift partition by partitioning V on $X_v - X'_v \in [q]$.

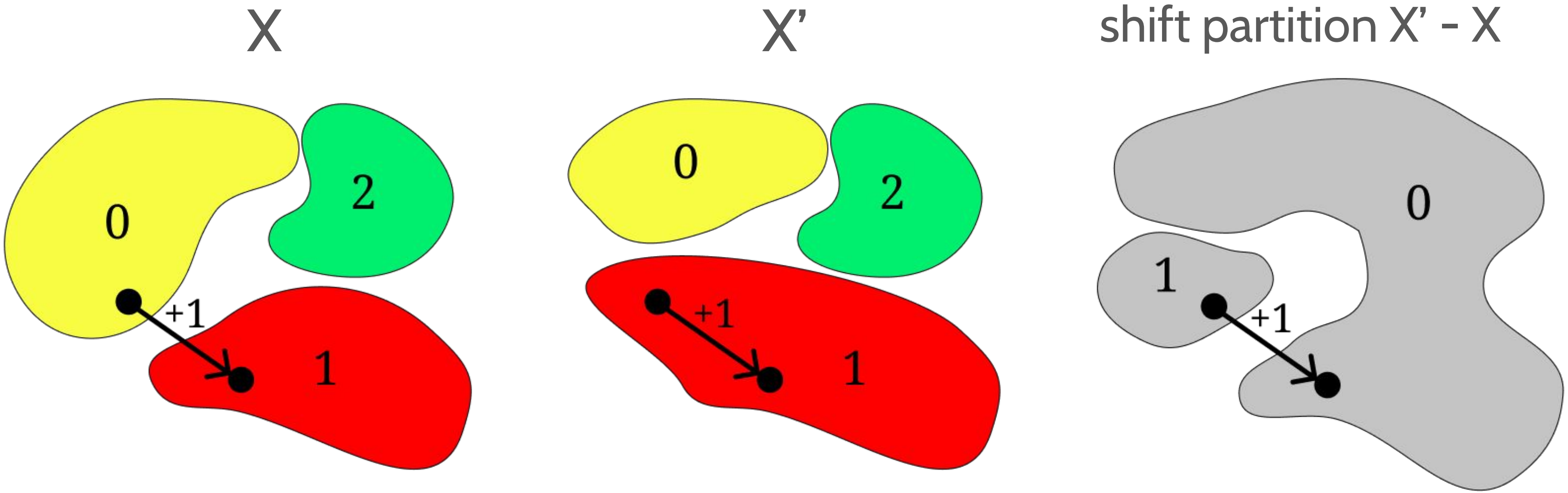
Lemma. Edges between blocks of the shift partition are violated in either X or X'

Since X, X' have value 1-ε, at most 2ε fraction of edges cross the partition.

Therefore, at least one block of the shift partition must be non-expanding.

In a SSE, this block must be large, |block| > δn.

Shift Partitions



Since edges across the shift partition are violated, but X,X' have high value, at least one block of the shift partition is non-expanding

Rounding SSEs

Takeaway: in a (δ, η)-SSE, there is a block of the shift partition with size ≥ δn

This implies the following rounding algorithm succeeds: condition on one random vertex, then round the remaining vertices independently

- Use Z_u to denote the output assignment, while X_u and X'_u denote the variables of the SoS program
- Set $Z_u = 0$, then sample Z_v independently from $\tilde{\mathbf{Pr}}_X[X_v = a \mid X_u = 0]$

Lemma. $\mathbb{E}_{\text{rounding } Z}[\text{value}(Z)] \geq \delta^2 - \varepsilon = \Omega(1)$

Lemma. For *symmetrized* $\tilde{\mathbf{E}}$, the conditional dist $X_v \mid X_u = 0$ is the same as $X_v - X_u$

Rounding SSEs

Lemma. $\mathbb{E}_{\text{rounding } Z}[\text{value}(Z)] \geq \delta^2 - \varepsilon$

Proof:

$$\mathbb{E}_Z[\text{value}(Z)] = \mathbb{E}_u \, \mathbb{E}_{(v,w)\in E} \Pr_{Z \mid Zu = 0}[Z_w - Z_v = c_{vw}]$$
$$= \mathbb{E}_u \, \mathbb{E}_{(v,w)\in E} \tilde{\mathbf{Pr}}_{X, X'}[X_w - X'_v = c_{vw} \mid X_u = X'_u = 0]$$
$$= \mathbb{E}_u \, \mathbb{E}_{(v,w)\in E} \tilde{\mathbf{Pr}}_{X, X'}[X_w - X_u - X'_v + X'_u = c_{vw}]$$

If u, v are in the same block of the shift partition of X and X', and v,w is a satisfied edge of X, then this event will occur. Hence,

$$\geq \mathbb{E}_u \, \mathbb{E}_{(v,w)\in E} \tilde{\mathbf{Pr}}_{X, X'}[\text{u,v in same block of shift partition, (v,w) is satisfied in X}]$$

By a union bound,

$$\geq 1 - \mathbb{E}_{u,v} \tilde{\mathbf{Pr}}_{X,X'}[\text{u,v not in same block of shift partition}] - \mathbb{E}_{v,w} \tilde{\mathbf{Pr}}_X[\text{(v,w) unsat in X}]$$
$$\geq 1 - (1 - \delta^2) - \varepsilon = \delta^2 - \varepsilon$$

■

SoS-izing Rounding SSEs

Lemma. If G is a D-certifiable (δ, η)-SSE, then SoS_D satisfies

$$\mathbb{E}_{u,v} \tilde{\mathbf{Pr}}_{X, X'}[u,v \text{ in same block of shift partition}] \geq \delta^2$$

[Simplified] We say that a (δ, η)-SSE G is D-certifiable if there is a degree-D SoS proof of

$$X_v^2 = X_v \quad \Rightarrow \quad \mathbb{E}_{(v,w)\in E} X_v(1-X_w) \geq \eta\, \mathbb{E}_v X_v + 0.1(\mathbb{E}_v X_v)(\delta - \mathbb{E}_v X_v)$$

$$\Pr[\text{edge crosses } S] \geq \eta \Pr[\text{edge starts in } S] + c\,(|S|/n)\,(\delta - |S|/n)$$

Proof sketch. It suffices to show that $\tilde{\mathbf{E}}_{X, X'}$ |block a|/n ≥ δ for some a.
Write the SSE SoS proof for the block indicators $E_{va} = 1[X_v - X'_v = a]$.
Apply $\tilde{\mathbf{E}}$ to both sides of the proof. If $\tilde{\mathbf{E}}$ |block a|/n < δ for all a, then going through the argument that edges across the shift partition violate X or X', we conclude that the value of $\tilde{\mathbf{E}}$ is at most 1-η ≪ 1-ε, a contradiction. ■

IV. Johnson graph

Theorem [BBKSS'21]. If G is the Johnson graph, then rounding $SoS_{O(1)}$ solves UG

Johnson Graph

The (n, ℓ, α) Johnson graph has vertices ([n] choose ℓ) and edges at intersection size $(1-\alpha)\ell$

- ℓ,α are constants and $\alpha \in [0,1]$ is the "noise parameter"

Slice of the hypercube $\{-1, +1\}^n$

The Johnson graph is not SSE. There are n+1 eigenspaces. Non-exp'ing sets are subcubes

- For $T \subseteq [n]$, the subcube for T (also known as link) is $C = \{S : S \supseteq T\}$

(n, α) Noisy Hypercube on $\{-1, +1\}^n$

r-restricted subcube = $\{x : x_1 = x_2 = \ldots = x_r = 1\}$

Expansion ≈ $1-(1-\alpha)^r$

Fractional volume = $1/2^r$

(n,ℓ,α) Johnson graph

r-restricted subcube = $\{x : x_1 = x_2 = \ldots = x_r = 1\}$

Expansion ≈ $1-(1-\alpha)^r$

Fractional volume ≈ $1/n^r$

Rounding the Johnson Graph

If we sample two high-value solutions, the shift partition must have a non-expanding set, but it's not necessarily large anymore

Idea: apply condition+round on *just this set*, fix those vertices, and repeat

- $\tilde{\mathbf{E}}$ is only changed on edges incident to the non-expanding set
- If the value of $\tilde{\mathbf{E}}$ changes, *should* be able to satisfy some incident edges
- Since the set is non-expanding, C&R satisfies nontrivial fraction of incident edges

Several pieces of the analysis are specific to the Johnson graph

- Proof that shift partition is correlated with a subcube requires degree O(1) SoS proof
- How to find the non-expanding subcube? Brute force search over all subcubes in poly(n) time
- Need that non-expanding sets chosen in the future have small overlap with previous ones

Rounding the Johnson Graph

Formal rounding algorithm (for a carefully chosen parameter δ):

while $\tilde{E}$ has value at least 1 - 2ε:

Find a non-expanding subcube C such that "condition+round value" ≥ δ

Perform condition+round on C

Rerandomize $\tilde{E}$ on C: $\tilde{E}[X] \leftarrow 1/q^{|C|} \sum_{\sigma \in [q]^C} \tilde{E}[\prod_{v \in C} X_{v,\sigma(v)} X]$

Set remaining values arbitrarily

Lemma. There is a subcube with condition+round value ≥ δ

Lemma. If the value decreases by v, at least Ω(v) fraction of edges become sat

Conclusion and Open Problems

UG is easy on: low threshold rank, certified SSEs, graphs with small number of distinct large eigenvalues/simple non-expanding sets

UG is unknown on: graphs with less structured spectra

Solve UG on the hypercube graph

Construction of a non-SoS-certifiable SSE

Other graph decompositions cutting ε fraction of edges?

Smarter ways to round SoS?

Counting Unique Games vs #BIS

Correlation Rounding on Low Threshold Rank Graphs

Define TV(X_u, X_v) = $\frac{1}{2} \sum_{a,b \in [q]} |\Pr[X_u=a, X_v=b] - \Pr[X_u=a]\Pr[X_v=b]|$

Conditioning reduces the average pairwise correlation of the variables:

Theorem [Raghavendra-Tan '11]. For all r and all boolean-valued random variables X_1, ..., X_n there is $t \le O(r^2)$ such that $\mathbb{E}_{|S|=t}\mathbb{E}_{i,j\in[n]}[TV(X_i, X_j) \mid X_S] \le 1/r$

Proof. Claim: there is $t \le r$ such that $\mathbb{E}_{|S|=t}\mathbb{E}_{i,j\in[n]}[\, I(X_i; X_j \mid X_S)] \le 1/r$

$I(X;Y) = H(X) - H(X|Y)$

$$\sum_{t=0}^{r-1} \mathbb{E}_{|S|=t}\mathbb{E}_{i,j\in[n]}[\, I(X_i; X_j \mid X_S)] = \mathbb{E}_{i\in[n]}[H(X_i)] - \mathbb{E}_{|R|=r}\mathbb{E}_{i\in[n]}[H(X_i \mid X_R)] \le 1$$

Finally, use TV(Xi, Xj) ≤ O(√I(Xi; Xj)) and Jensen's inequality. ■

Theorem [Jain-Koehler-Risteski '18]. Cannot improve $O(r^2)$ to $o(r^2)$:

Sherrington-Kirkpatrick model

1. Low threshold rank

Local to Global Correlations

In an expander or low threshold rank graph, local correlation implies global correlation

Theorem. If $\mathbb{E}_{i\in V}\|v_i\|^2 = 1$ and $\mathbb{E}_{(i,j)\in E}[\langle v_i, v_j\rangle] \geq 1-\varepsilon$, then $\mathbb{E}_{i,j\in V}[\langle v_i, v_j\rangle^2] \geq 1/\mathrm{rank}_{1-2\varepsilon}(G)$

Proof sketch. For simplicity, assume v_i are scalar-valued (one-dimensional). Consider the spectral sample $\lambda_e \sim v$ by taking λ_e with probability $\langle v, b_e\rangle^2$.

Local correlation: $\mathbb{E}_{(i,j)\in E}\, v_i v_j = v^T(A/d)v\, /\, n = \mathbb{E}_{\lambda e\sim v}[\lambda_e]$

Global correlation: $\mathbb{E}_{i,j\in V}\,(v_i v_j)^2 \geq \|p(\lambda_e)\|_2^{\,2}$

Using Cauchy-Schwarz,

$$\Pr_{\lambda e\sim v}[\lambda_e \geq 1-2\varepsilon] \leq \mathrm{rank}_{1-2\varepsilon}(G)\, \|p(\lambda_e)\|_2^{\,2}$$

Compare this with $\mathbb{E}[\lambda_e]$ using the inequality below, then rearrange,

$$\mathbb{E}_{\lambda e\sim v}[\lambda_e] \leq \Pr_{\lambda e\sim v}[\lambda_e \geq 1-2\varepsilon] + (1-2\varepsilon)(1-\Pr_{\lambda e\sim v}[\lambda_e \geq 1-2\varepsilon])$$ ∎

Local to Global Correlations

Passing to TV...

Theorem. If $\mathbb{E}_{(i,j)\in E}\,[\mathrm{TV}(X_i, X_j)] \geq \varepsilon$, then $\mathbb{E}_{i,j\in V}\,[\mathrm{TV}(X_i, X_j)] \geq \mathrm{poly}(\varepsilon)/\mathrm{rank}_{1\text{-}\mathrm{poly}(\varepsilon)}(G)$

Proof sketch. Let v_{ia} = w_{ia}+c_{ia}1 be the SDP vectors.

We have TV(Xi, Xj) = sum_{a,b} |<w_{ia}, w_{jb}>|.

Construct vi such that <vi, vj> = poly(TV(Xi, Xj)) and apply the Lemma on v_i.

Specifically, let vi = sum_a w_{ia}^{otimes 2} / ||w_ia||