Sum-of-Squares for Unique Games

\[ \tilde{E}[f^2] \geq 0 \]

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 $(G, \Pi)$ where $G$ is a directed graph, $\Pi = (\pi_e)_{e \in E}$ is a permutation of $[q]$ for each edge,

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

Unique Games Conjecture (UGC): For all $\epsilon, s > 0$, there is $q$ sufficiently large such that it is NP-hard to distinguish between: $(G, \Pi)$ has value $\geq 1-\epsilon$ or value $\leq s$.

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

“Solve UG” = when the input is $(1-\epsilon)$ satisfiable, find a solution with value $\Omega_\epsilon(1)$

- Drop the parameter $s$ from here on out and assume we are given $(1-\epsilon)$ satisfiable $(G, \Pi)$, where $\epsilon$ is a tiny constant

---

0. Introduction
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 $\tilde{E}$ which maximizes $\tilde{E}[\text{objective}]$.

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

1. degree $D/2$ local reasoning
2. $\tilde{E}[p(X)^2] \geq 0$ for all degree $\leq D/2$ polynomials $p$

$\tilde{Pr}$ denotes the local probability distribution, e.g. $\tilde{Pr}[X_i = a]$
Sum-of-Squares for Unique Games

Given \((G, \Pi)\) 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 \(u \in V, a \in [q]\)

Constraints:
\[
X_{ua}^2 = X_{ua}
\]
\[
\sum_a X_{ua} = 1
\]

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

Run \(\text{SoS}_D\) to produce \(\tilde{E}\) for the above system. \(\tilde{E}\) is a \textit{fake} distribution of solutions, which has pseudo-expected value at least \((1-\epsilon)\).

Our goal is to design a \textbf{rounding algorithm} to “sample” from \(\tilde{E}\) a \textit{real} solution with value \(\Omega_\epsilon (1)\)
How does SoS perform on UG?

Let \((G, \Pi)\) be a UG instance with value at least \((1-\varepsilon)\).

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

**Theorem [BRS’11].** For general \(G\), rounding \(\text{SoS}_{nO(\varepsilon)}\) solves UG.

**Theorem [BBKSS’21].** If \(G\) is a D-certifiable small set expander, then rounding \(\text{SoS}_{O(D)}\) solves UG.

**Theorem [BBKSS’21].** If \(G\) is the Johnson graph, then rounding \(\text{SoS}_{O(1)}\) solves UG.

**Open:** does rounding \(\text{SoS}_4\) solve UG?

*0. Introduction*
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

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

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 $\tau$

- We will always use constant $\tau$, such as $\tau = 1\text{-poly}(\epsilon)$

**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 $C_n$

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

All dense graphs have low threshold rank:

**Lemma.** Any $d$-regular graph $G$ with $d = pn$ has $\tau(G) = O(1)$

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

1. Low threshold rank
Correlation Rounding on Low Threshold Rank Graphs

Theorem [BRS’11]. If $G$ has $(1-\epsilon^5)$-threshold rank $r$, then rounding $\text{SoS}_{O(r^2)}$ solves UG.

Idea: we wish that one of these two rounding schemes worked:
- for each $v$, sample $v$ independently from its local distribution
- for each $e = (u,v)$, sample $(u,v)$ according to its local distribution

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 $\tilde{\Pr}[X_u | X_S]$. These distributions exist provided the SoS degree is at least $|S|+1$.

1. Low threshold rank
Correlation Rounding on Low Threshold Rank Graphs

**Theorem [BRS’11].** If $G$ has $(1-\varepsilon^5)$-threshold rank $r$, then rounding $\text{SoS}_{O(r^2)}$ 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 $X_1, ..., X_n$ there is $S \subseteq [n], |S| \leq O(r^2)$ such that $\mathbb{E}_{i,j \in [n]}[TV(X_i, X_j) | X_S] \leq 1/r$

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

After conditioning, we may conclude that $\mathbb{E}_{(i,j) \in E} [TV(X_i, X_j)] \leq 1-2\varepsilon$. Looking at the event “edge $(i, j)$ is satisfied”, we have:

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

1. Low threshold rank
II. General graphs in subexponential time

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

**Theorem [BRS’11].** For general $G$, rounding $\text{SoS}_{nO(\epsilon)}$ solves UG.
What about high threshold rank?

Recall: cycle graph $C_n$ \\
\[ \text{rank}_T(C_n) = \Theta(n) \]

Cutting $\varepsilon$ fraction of the edges changes the objective value by at most $\varepsilon$.

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

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

2. \textit{UG in subexponential time}
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 $\text{rank}_{1-\varepsilon^5}(G[V_i]) \leq n^{100\varepsilon}$ 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 $\varphi$-expander, great! Otherwise there is a non-expanding set $S$, $|S| \leq n/2$. Partition $G$ into $S$ and $V \setminus S$ and recurse.

2. **UG in subexponential time**
Graph Partitioning Lemma

**Lemma [Arora-Barak-Steurer ’10].** Any graph $G$ can be partitioned into pieces $V_i$ with 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 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 = $\varepsilon^2k = O(\varepsilon \log(1/\varepsilon))$.

---

2. **UG in subexponential time**
III. Certified Small Set Expanders

Theorem [BBKSS'21]. If $G$ is a \textit{D-certifiable small set expander}, then rounding \textit{SoS}$_{O(D)}$ solves UG.
Small Set Expansion

G is a \((\delta, \eta)\)-small set expander (SSE) if for all \(|S| \leq \delta n\), \(\Phi_G(S) \geq \eta\)
\(-\delta, \eta\) are fixed small constants while \(\text{val}(G) \geq 1-\varepsilon\) where \(\varepsilon \ll \delta, \eta\)

Idea for rounding SoS on a small set expander:
- Recall that \(\tilde{E}\) 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 \(1-\varepsilon\), at most \(2\varepsilon\) 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, \(|\text{block}| > \delta n\).
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 $(\delta, \eta)$-SSE, there is a block of the shift partition with size $\geq \delta n$

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 $\Pr_{X'}[X_v = a \mid X_u = 0]$

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

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

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

Proof: $\mathbb{E}_Z [\text{value}(Z)] = \mathbb{E}_u \mathbb{E}_{(v,w) \in E} \Pr_{Z \mid Z_u = 0} [Z_w - Z_v = c_{vw}]$

$= \mathbb{E}_u \mathbb{E}_{(v,w) \in E} \tilde{\Pr}_{X, X'} [X_w - X_v = c_{vw} \mid X_u = X'_u = 0]$

$= \mathbb{E}_u \mathbb{E}_{(v,w) \in E} \tilde{\Pr}_{X, X'} [X_w - X_u - X'_v + X'_u = c_{vw}]$

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,

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

By a union bound,

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

$\geq 1 - (1 - \delta^2) - \varepsilon = \delta^2 - \varepsilon$
SoS-izing Rounding SSEs

Lemma. If $G$ is a $D$-certifiable $(\delta, \eta)$-SSE, then $\text{SoS}_D$ satisfies

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

[Simplified] We say that a $(\delta, \eta)$-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 $\tilde{E}_{X, X'} |\text{block } a|/n \geq \delta$ for some $a$. Write the SSE SoS proof for the block indicators $E_{va} = 1[X_v - X'_v = a]$. Apply $\tilde{E}$ to both sides of the proof. If $\tilde{E} |\text{block } a|/n < \delta$ for all $a$, then going through the argument that edges across the shift partition violate $X$ or $X'$, we conclude that the value of $\tilde{E}$ is at most $1-\eta \ll 1-\epsilon$, a contradiction.
IV. Johnson graph

Theorem [BBKSS’21]. If $G$ is the Johnson graph, then rounding $\text{SoS}_{O(1)}$ solves UG
Johnson Graph

The \((n, \ell, \alpha)\) Johnson graph has vertices \(\binom{n}{\ell}\) and edges at intersection size \((1-\alpha)\ell\)

\(\ell, \alpha\) 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 \(C = \{S : S \supseteq T\}\)

\((n, \alpha)\) Noisy Hypercube on \([-1, +1]^n\)

\[\text{r-restricted subcube} = \{x : x_1 = x_2 = \ldots = x_r = 1\}\]

Expansion \(\approx 1-(1-\alpha)^r\)

Fractional volume \(= \frac{1}{2^r}\)

\((n, \ell, \alpha)\) Johnson graph

\[\text{r-restricted subcube} = \{x : x_1 = x_2 = \ldots = x_r = 1\}\]

Expansion \(\approx 1-(1-\alpha)^r\)

Fractional volume \(\approx \frac{1}{n^r}\)
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
- $\tilde{E}$ is only changed on edges incident to the non-expanding set
- If the value of $\tilde{E}$ 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 $\delta$):

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

Find a non-expanding subcube $C$ such that “condition+round value” $\geq \delta$

Perform condition+round on $C$

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

Set remaining values arbitrarily

**Lemma.** There is a subcube with condition+round value $\geq \delta$

**Lemma.** If the value decreases by $v$, at least $\Omega(v)$ 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 $\varepsilon$ 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 $t \leq O(r^2)$ such that $\mathbb{E}_{|S| = t} \mathbb{E}_{i, j \in [n]}[TV(X_i, X_j | X_S)] \leq 1/r$

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

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

Finally, use $TV(X_i, X_j) \leq O(\sqrt{I(X_i; X_j)})$ and Jensen’s inequality.

**Theorem [Jain-Koehler-Risteski ‘18].** Cannot improve $O(r^2)$ to $o(r^2)$:
Sherrington-Kirkpatrick model
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 $\mathbb{E}_{i,j \in V} [\langle v_i, v_j \rangle^2] \geq 1/\text{rank}_{1-2\varepsilon}(G)$

**Proof sketch.** For simplicity, assume $v_i$ are scalar-valued (one-dimensional). Consider the spectral sample $\lambda_e \sim v$ by taking $\lambda_e$ 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 \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,

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

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

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

1. **Low threshold rank**
Local to Global Correlations

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

Proof sketch. Let $v_{\{ia\}} = w_{\{ia\}} + c_{\{ia\}}1$ be the SDP vectors. We have $TV(X_i, X_j) = \sum_{a,b} |<w_{\{ia\}}, w_{\{jb\}}>|$. Construct $v_i$ such that $<v_i, v_j> = poly(TV(X_i, X_j))$ and apply the Lemma on $v_i$.

Specifically, let $v_i = \sum_a w_{\{ia\}}^{\otimes 2} / \|w_{ia}\|$