Combinatorial Optimization via the Sum of Squares Hierarchy

Goutham Rajendran

University of Chicago

25th May, 2018

Contents

Contents

- Introduction:
  - Semidefinite programming
  - Sum of Squares hierarchy
- Algorithmic techniques:
  - Maximum Clique
  - Minimum Bisection
- Lower bounds:
  - Maximum K-CSP
  - Densest k-subgraph
  - Densest k-subhypergraph
  - Minimum p-Union
- Pseudoexpectations
- Pseudocalibration
- Future Work

Optimization problems

Optimization problems

- We consider discrete optimization problems.
- Examples - Maximum Clique, Densest *k*-subgraph, Maximum Cut.
- Optimum value is denoted OPT.
- For a maximization problem, an α-approximation algorithm for $\alpha \geq 1$ outputs solution with value $\geq \frac{1}{\alpha} \cdot OPT$.

Integer/Linear Programming

Integer/Linear Programming

- Input: $A \in \mathbb{R}^{m \times n}$, $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$.
- Unknown $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

$$
\begin{array}{rl}
\text{Maximize} & \boldsymbol{c}^T \boldsymbol{x} \\
\text{subject to} & A\boldsymbol{x} \leq \boldsymbol{b}
\end{array}
$$

Integer/Linear Programming

- Input: $A \in \mathbb{R}^{m \times n}$, $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$.
- Unknown $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

$$\begin{array}{ll} \text{Maximize} & \boldsymbol{c}^T \boldsymbol{x} \\ \text{subject to} & A\boldsymbol{x} \leq \boldsymbol{b} \end{array}$$

- Linear program: Optimize over $\boldsymbol{x} \in \mathbb{R}^n$
- Integer linear program: Optimize over $\boldsymbol{x} \in \mathbb{Z}^n$

Positive Semidefinite Matrices

Positive Semidefinite Matrices

- A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is *positive semidefinite* if any of these equivalent conditions is true:
  - $\boldsymbol{x}^T A \boldsymbol{x} \geq 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$.
  - All eigenvalues of A are nonnegative.
  - $A = X^T X$ for some $X \in \mathbb{R}^{d \times n}, d \leq n$.
- This is denoted $A \succeq 0$.

Semidefinite Programming

Semidefinite Programming

- Input: $C, A_1, \ldots, A_m \in \mathbb{R}^{n \times n}$, $b_i \in \mathbb{R}$.
- Unknown $Y = (y_{i,j})_{i,j \leq n} \in \mathbb{R}^{n \times n}$.
- Semidefinite program:

$$\begin{aligned} \text{Maximize} \quad & C \bullet Y = \sum_{i,j \leq n} C_{i,j} Y_{i,j} \\ \text{subject to} \quad & A_i \bullet Y \leq b_i \\ & Y \succeq 0 \\ & Y \in \mathbb{R}^{n \times n} \end{aligned}$$

Semidefinite Programming

- Input: $C, A_1, \ldots, A_m \in \mathbb{R}^{n \times n}$, $b_i \in \mathbb{R}$.
- Unknown $Y = (y_{i,j})_{i,j \leq n} \in \mathbb{R}^{n \times n}$.
- Semidefinite program:

$$\begin{aligned} \text{Maximize} \quad & C \bullet Y = \sum_{i,j \leq n} C_{i,j} Y_{i,j} \\ \text{subject to} \quad & A_i \bullet Y \leq b_i \\ & Y \succeq 0 \\ & Y \in \mathbb{R}^{n \times n} \end{aligned}$$

- Can be approximated to arbitrary precision in polynomial time, under some mild assumptions

  - Grötschel, Lovász and Schrijver[GLS88].

Maximum Cut

Maximum Cut

- Given a graph $G = (V, E)$, find a partition $(S, V - S)$ of V so that the number of edges with exactly one endpoint in S, is maximized.

Maximum Cut

- Given a graph $G = (V, E)$, find a partition $(S, V - S)$ of V so that the number of edges with exactly one endpoint in S, is maximized.
- General program:

$$
\begin{aligned}
\text{Maximize} \quad & \sum_{(u,v)\in E} \left(\frac{1}{2} - \frac{1}{2} x_u x_v\right) \\
\text{subject to} \quad & x_u^2 = 1 \\
& x_u \in \mathbb{R}
\end{aligned}
$$

Maximum Cut

- Given a graph $G = (V, E)$, find a partition $(S, V - S)$ of V so that the number of edges with exactly one endpoint in S, is maximized.
- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{(u,v)\in E} \left(\frac{1}{2} - \frac{1}{2} x_u x_v\right) \\ \text{subject to} \quad & x_u^2 = 1 \\ & x_u \in \mathbb{R} \end{aligned}$$

- Semidefinite program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{(u,v)\in E} \left(\frac{1}{2} - \frac{1}{2} \langle \boldsymbol{V}_u, \boldsymbol{V}_v \rangle\right) \\ \text{subject to} \quad & \langle \boldsymbol{V}_u, \boldsymbol{V}_u \rangle = 1 \\ & \boldsymbol{V}_u \in \mathbb{R}^d \end{aligned}$$

Goemans-Williamson algorithm

Goemans-Williamson algorithm

- Suppose $\boldsymbol{V}_u \in \mathbb{R}^d$. Sample a random unit vector $\boldsymbol{g}$ in $\mathbb{R}^d$ and set

$$x_u = \begin{cases} 1 & \text{if } \langle \boldsymbol{g}, \boldsymbol{V}_u \rangle \geq 0 \\ -1 & \text{otherwise} \end{cases}$$

- Output $S = \{u \in V \mid x_u = 1\}$.

Goemans-Williamson algorithm

- Suppose $\boldsymbol{V}_u \in \mathbb{R}^d$. Sample a random unit vector $\boldsymbol{g}$ in $\mathbb{R}^d$ and set

$$x_u = \begin{cases} 1 & \text{if } \langle \boldsymbol{g}, \boldsymbol{V}_u \rangle \geq 0 \\ -1 & \text{otherwise} \end{cases}$$

- Output $S = \{u \in V \mid x_u = 1\}$.
- Achieves ≈ 1.138 approximation.
- Above analysis is optimal for this SDP

  - Feige and Schechtman[FS02].
- Improving this approximation factor is UG-hard (UG is Unique Games)

  - Khot, Kindler, Mossel and O'Donnell[KKMO07].

Maximum Clique

Maximum Clique

- Given a graph G, find the largest clique in G.

Maximum Clique

- Given a graph G, find the largest clique in G.
- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{u \in V} x_u \\ \text{subject to} \quad & x_u x_v = 0 \qquad \forall (u,v) \notin E, u \neq v \\ & x_u \in \{0,1\} \end{aligned}$$

SoS relaxation for Maximum Clique - Intuition

SoS relaxation for Maximum Clique - Intuition

- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{u \in V} x_u \\ \text{subject to} \quad & x_u x_v = 0 \qquad \forall (u,v) \notin E, u \neq v \\ & x_u \in \{0,1\} \end{aligned}$$

SoS relaxation for Maximum Clique - Intuition

- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{u \in V} x_u \\ \text{subject to} \quad & x_u x_v = 0 \qquad \forall (u,v) \notin E, u \neq v \\ & x_u \in \{0,1\} \end{aligned}$$

- We will write a larger program to capture properties satisfied by any convex combination of optimal integer solutions.
- For all small S, introduce vectors $\boldsymbol{V}_S$ which capture the event that S is a subset of the optimal solution.
- Want $\|\boldsymbol{V}_S\|^2$ to be $\mathbb{E}[\prod_{i \in S} x_i]$ over a distribution supported on integer solutions.

SoS relaxation for Maximum Clique - Local constraints

SoS relaxation for Maximum Clique - Local constraints

- Local variables - $\boldsymbol{V}_S$, for all $S \in [n]_{\leq r} = \{T \subseteq [n] \mid |T| \leq r\}$

SoS relaxation for Maximum Clique - Local constraints

- Local variables - $\boldsymbol{V}_S$, for all $S \in [n]_{\leq r} = \{T \subseteq [n] \mid |T| \leq r\}$
- Add local consistency constraints:
  - $\|\boldsymbol{V}_\phi\|^2 = 1$
  - $\langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle = \langle \boldsymbol{V}_{S_3}, \boldsymbol{V}_{S_4} \rangle$ for all $S_1, S_2, S_3, S_4 \in [n]_{\leq r}$ such that $S_1 \cup S_2 = S_3 \cup S_4$
  - $\langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \geq 0$ for all $S_1, S_2 \in [n]_{\leq r}$

SoS relaxation for Maximum Clique - Local constraints

- Local variables - $\boldsymbol{V}_S$, for all $S \in [n]_{\leq r} = \{T \subseteq [n] \mid |T| \leq r\}$
- Add local consistency constraints:
  - $\|\boldsymbol{V}_\phi\|^2 = 1$
  - $\langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle = \langle \boldsymbol{V}_{S_3}, \boldsymbol{V}_{S_4} \rangle$ for all $S_1, S_2, S_3, S_4 \in [n]_{\leq r}$ such that $S_1 \cup S_2 = S_3 \cup S_4$
  - $\langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \geq 0$ for all $S_1, S_2 \in [n]_{\leq r}$
- Replace $x_i x_j$ by $\langle \boldsymbol{V}_{\{i\}}, \boldsymbol{V}_{\{j\}} \rangle$ or $\langle \boldsymbol{V}_{\{i,j\}}, \boldsymbol{V}_\phi \rangle$.
- Replace $x_1 x_3 + x_5 \leq 10$ by $\langle \boldsymbol{V}_S, \boldsymbol{V}_{\{1,3\}} \rangle + \langle \boldsymbol{V}_S, \boldsymbol{V}_{\{5\}} \rangle \leq 10 \langle \boldsymbol{V}_S, \boldsymbol{V}_\phi \rangle$ for all $S \in [n]_{\leq r}$.

Maximum Clique - SoS relaxation

Maximum Clique - SoS relaxation

- General Program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{u \in V} x_u \\ \text{subject to} \quad & x_u x_v = 0 && \forall (u,v) \notin E, u \neq v \\ & x_u \in \{0,1\} \end{aligned}$$

- Level-r SoS relaxation:

$$\begin{aligned} \text{Maximize} \quad & \sum_{u \in V} \| \boldsymbol{V}_{\{u\}} \|^2 \\ \text{subject to} \quad & \langle \boldsymbol{V}_{\{u,v\}}, \boldsymbol{V}_S \rangle = 0 && \forall (u,v) \notin E, u \neq v, S \in [n]_{\leq r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle = \langle \boldsymbol{V}_{S_3}, \boldsymbol{V}_{S_4} \rangle && \forall S_1 \cup S_2 = S_3 \cup S_4 \text{ and } S_i \in [n]_{\leq r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \geq 0 && \forall S_1, S_2 \in [n]_{\leq r} \\ & \| \boldsymbol{V}_{\phi} \|^2 = 1 \end{aligned}$$

LP/SDP Hierarchies - Outline

LP/SDP Hierarchies - Outline

- Add more consistency constraints that an actual probability distribution over integral solutions would satisfy.
- This gives a sequence of progressively stronger relaxations of LPs/SDPs.

LP/SDP Hierarchies - Outline

- Add more consistency constraints that an actual probability distribution over integral solutions would satisfy.
- This gives a sequence of progressively stronger relaxations of LPs/SDPs.
- In particular, we add local constraints to improve the approximation factor.
- Tradeoff between approximation factor and running time.
- Need to prove that local constraints help in approximating global properties.

LP/SDP Hierarchies

LP/SDP Hierarchies

- LP hierarchies - studied by Lovász and Schrijver; and Sherali and Adams.
- SDP hierarchies - LS_+ hierarchy; Sum of Squares hierarchy (SoS) studied by Shor, Nesterov, Parrillo and Lasserre.

LP/SDP Hierarchies

- LP hierarchies - studied by Lovász and Schrijver; and Sherali and Adams.
- SDP hierarchies - LS_+ hierarchy; Sum of Squares hierarchy (SoS) studied by Shor, Nesterov, Parrillo and Lasserre.
- Can solve level-r relaxation in time $mn^{O(r)}$ where m is the number of constraints in the starting program.
- Program's optimum value is usually denoted FRAC (in this presentation).
- Integrality gap = FRAC / OPT (maximization problem) quantifies performance.

General polynomial optimization problem

General polynomial optimization problem

- General program Γ:

$$\begin{array}{lll} \text{Maximize} & p(x_1, \ldots, x_n) & \\ \text{subject to} & q_i(x_1, \ldots, x_n) \geq 0 & i = 1, 2, \ldots, m \\ & x_i \in \{0, 1\} & \end{array}$$

General polynomial optimization problem

- General program Γ:

$$\begin{array}{lll} \text{Maximize} & p(x_1,\ldots,x_n) & \\ \text{subject to} & q_i(x_1,\ldots,x_n) \geq 0 & i = 1,2,\ldots,m \\ & x_i \in \{0,1\} & \end{array}$$

- Assume p, q_i are multilinear of degree $\leq r$.
- Let $p = \sum_{T \in [n]_{\leq r}} p_T \boldsymbol{x}_T$ and $q_i = \sum_{T \in [n]_{\leq r}} (q_i)_T \boldsymbol{x}_T$ where $\boldsymbol{x}_T = \prod_{i \in T} x_i$.

General SoS relaxation

General SoS relaxation

- General program:

$$\begin{array}{lll} \text{Maximize} & p(x_1, \ldots, x_n) & \\ \text{subject to} & q_i(x_1, \ldots, x_n) \geq 0 & i = 1, 2, \ldots, m \\ & x_i \in \{0, 1\} & \end{array}$$

General SoS relaxation

- General program:

$$\begin{aligned} \text{Maximize} \quad & p(x_1, \ldots, x_n) \\ \text{subject to} \quad & q_i(x_1, \ldots, x_n) \geq 0 \qquad i = 1, 2, \ldots, m \\ & x_i \in \{0, 1\} \end{aligned}$$

- Level-r SoS relaxation:

$$\begin{aligned} \text{Maximize} \quad & \sum_{T \in [n]_{\leq r}} p_T \|\boldsymbol{V}_T\|^2 \\ \text{subject to} \quad & \sum_{T \in [n]_{\leq r}} (q_i)_T \langle \boldsymbol{V}_T, \boldsymbol{V}_S \rangle \geq 0 \quad \forall S \in [n]_{\leq r}, i = 1, \ldots, m \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle = \langle \boldsymbol{V}_{S_3}, \boldsymbol{V}_{S_4} \rangle \qquad \forall S_1 \cup S_2 = S_3 \cup S_4 \text{ and } S_i \in [n]_{\leq r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \geq 0 \qquad \forall S_1, S_2 \in [n]_{\leq r} \\ & \|\boldsymbol{V}_\phi\|^2 = 1 \end{aligned}$$

General SoS relaxation

- General program:

$$\begin{array}{lll} \text{Maximize} & p(x_1, \ldots, x_n) & \\ \text{subject to} & q_i(x_1, \ldots, x_n) \geq 0 & i = 1, 2, \ldots, m \\ & x_i \in \{0, 1\} & \end{array}$$

- Level-r SoS relaxation:

$$\begin{array}{lll} \text{Maximize} & \displaystyle\sum_{T \in [n]_{\leq r}} p_T \|\boldsymbol{V}_T\|^2 & \\ \text{subject to} & \displaystyle\sum_{T \in [n]_{\leq r}} (q_i)_T \langle \boldsymbol{V}_T, \boldsymbol{V}_S \rangle \geq 0 & \forall S \in [n]_{\leq r}, i = 1, \ldots, m \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle = \langle \boldsymbol{V}_{S_3}, \boldsymbol{V}_{S_4} \rangle & \forall S_1 \cup S_2 = S_3 \cup S_4 \text{ and } S_i \in [n]_{\leq r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \geq 0 & \forall S_1, S_2 \in [n]_{\leq r} \\ & \|\boldsymbol{V}_\phi\|^2 = 1 & \end{array}$$

- Relaxation because if general program had optimal solution $\{b_i\}_{i \leq n}$, then $\boldsymbol{V}_T = \prod_{i \in T} b_i$ gives same objective value.

Example 1 - Maximum Clique

Example 1 - Maximum Clique

- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{u \in V} x_u \\ \text{subject to} \quad & x_u x_v = 0 \qquad \forall (u,v) \notin E, u \neq v \\ & x_u \in \{0,1\} \end{aligned}$$

- Level-r SoS relaxation:

$$\begin{aligned} \text{Maximize} \quad & \sum_{u \in V} \| \boldsymbol{V}_{\{u\}} \|^2 \\ \text{subject to} \quad & \langle \boldsymbol{V}_{\{u,v\}}, \boldsymbol{V}_S \rangle = 0 \qquad \forall (u,v) \notin E, u \neq v, S \in [n]_{\leq r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle = \langle \boldsymbol{V}_{S_3}, \boldsymbol{V}_{S_4} \rangle \quad \forall S_1 \cup S_2 = S_3 \cup S_4 \text{ and } S_i \in [n]_{\leq r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \geq 0 \qquad \forall S_1, S_2 \in [n]_{\leq r} \\ & \| \boldsymbol{V}_{\phi} \|^2 = 1 \end{aligned}$$

Example 2 - Densest k-subgraph

Example 2 - Densest k-subgraph

- Given a graph $G = (V, E)$ and a positive integer k, find a subset W of V with exactly k vertices with maximum number of edges within.

Example 2 - Densest k-subgraph

- Given a graph $G = (V, E)$ and a positive integer k, find a subset W of V with exactly k vertices with maximum number of edges within.
- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{(u,v)\in E} x_u x_v \\ \text{subject to} \quad & \sum_{u\in V} x_u = k \\ & x_u \in \{0, 1\} \end{aligned}$$

Example 2 - Densest k-subgraph

- Given a graph $G = (V, E)$ and a positive integer k, find a subset W of V with exactly k vertices with maximum number of edges within.
- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{(u,v)\in E} x_u x_v \\ \text{subject to} \quad & \sum_{u\in V} x_u = k \\ & x_u \in \{0,1\} \end{aligned}$$

- Level-r SoS relaxation:

$$\begin{aligned} \text{Maximize} \quad & \sum_{(u,v)\in E} \|\boldsymbol{V}_{\{u,v\}}\|^2 & \\ \text{subject to} \quad & \sum_{v\in V} \langle \boldsymbol{V}_{\{v\}}, \boldsymbol{V}_S \rangle = k\|\boldsymbol{V}_S\|^2 & \forall S \in [n]_{\le r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle = \langle \boldsymbol{V}_{S_3}, \boldsymbol{V}_{S_4} \rangle & \forall S_1 \cup S_2 = S_3 \cup S_4 \text{ and } S_i \in [n]_{\le r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \ge 0 & \forall S_1, S_2 \in [n]_{\le r} \\ & \|\boldsymbol{V}_\phi\|^2 = 1 & \end{aligned}$$

Algorithmic techniques

Maximum Clique

Maximum Clique

- Hard to approximate within a factor of $n/2^{(\log n)^{3/4+\epsilon}}$ for any $\epsilon > 0$, assuming $NP \not\subseteq BPTIME(2^{(\log n)^{O(1)}})$

  - Khot and Ponnuswami[KP06]

Maximum Clique

- Hard to approximate within a factor of $n/2^{(\log n)^{3/4+\epsilon}}$ for any $\epsilon > 0$, assuming $NP \not\subseteq BPTIME(2^{(\log n)^{O(1)}})$

  - Khot and Ponnuswami[KP06]
- Interesting to study this problem for Erdös-Rényi random graphs $G \sim G(n, 1/2)$
- $G \sim G(n, 1/2)$ has no cliques of size more than $2 \log n$ with high probability

MaxClique on random graphs

MaxClique on random graphs

- Theorem: For some $c > 0$, for all $r \leq c \log n$, the level-r SoS hierarchy has $FRAC = O(\sqrt{n/2^r})$, with high probability, for $G \sim G(n, 1/2)$.

  - Feige and Krauthgamer[FK03]

MaxClique on random graphs

- Theorem: For some $c > 0$, for all $r \leq c \log n$, the level-r SoS hierarchy has $FRAC = O(\sqrt{n/2^r})$, with high probability, for $G \sim G(n, 1/2)$.

  \- Feige and Krauthgamer[FK03]

- Originally shown for the Lovász-Schrijver hierarchy but proof simplifies if we use the SoS hierarchy.
- For the Lovász-Schrijver hierarchy, they also showed $\Omega(\sqrt{n/2^r})$.

MaxClique on random graphs

- Theorem: For some $c > 0$, for all $r \leq c \log n$, the level-r SoS hierarchy has $FRAC = O(\sqrt{n/2^r})$, with high probability, for $G \sim G(n, 1/2)$.

  - Feige and Krauthgamer[FK03]
- Originally shown for the Lovász-Schrijver hierarchy but proof simplifies if we use the SoS hierarchy.
- For the Lovász-Schrijver hierarchy, they also showed $\Omega(\sqrt{n/2^r})$.
- Theorem: If $r = o(\log n)$, the level-r SoS relaxation for MaxClique will have $FRAC \geq k = n^{1/2 - O(\sqrt{r/\log n})}$ on $G \sim G(n, 1/2)$ with high probability.
  - Barak, Hopkins, Kelner, Kothari, Moitra and Potechin[BHK+16]

Minimum Bisection

Minimum Bisection

- Given a graph $G = (V, E)$ and an integer k, find a subset S of the vertices with exactly k vertices such that $\Gamma(S) = |E(S, V - S)|$, is minimized.

Minimum Bisection

- Given a graph $G = (V, E)$ and an integer k, find a subset S of the vertices with exactly k vertices such that $\Gamma(S) = |E(S, V - S)|$, is minimized.
- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{(u,v)\in E} (x_u - x_v)^2 \\ \text{subject to} \quad & \sum_{u\in V} x_u = k \\ & x_u \in \{0, 1\} \end{aligned}$$

Minimum Bisection

- Given a graph $G = (V, E)$ and an integer k, find a subset S of the vertices with exactly k vertices such that $\Gamma(S) = |E(S, V - S)|$, is minimized.
- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{(u,v)\in E} (x_u - x_v)^2 \\ \text{subject to} \quad & \sum_{u\in V} x_u = k \\ & x_u \in \{0, 1\} \end{aligned}$$

- Theorem: Consider an instance of Minimum Bisection (G, k). For any $r \in \mathbb{Z}$ and $\epsilon > 0$, we can find $R' \subseteq V$ such that
  - $|R'| \approx k$
  - $\Gamma(R') \leq \frac{1+\epsilon}{\min(1,\lambda_r(L))} \cdot OPT$

  in time $n^{O(r/\epsilon^2)}$.

- Guruswami and Sinop[GS11]

Lower bounds

Constraint satisfaction problems

Constraint satisfaction problems

- Given m constraints $C_1, \ldots, C_m$ over n variables $x_1, \ldots, x_n$ over alphabet $[q]$, find an assignment of $x_1, \ldots, x_n$ to $[q]$ such that maximum number of constraints are satisfied.

Constraint satisfaction problems

- Given m constraints $C_1, \ldots, C_m$ over n variables $x_1, \ldots, x_n$ over alphabet $[q]$, find an assignment of $x_1, \ldots, x_n$ to $[q]$ such that maximum number of constraints are satisfied.
- Each constraint C_i on subset T_i is a function from $[q]^{T_i}$ to $\{0, 1\}$.
- An assignment *satisfies* C_i if the evaluation of C_i on the assignment restricted to T_i is 1.

Constraint satisfaction problems

- Given m constraints $C_1, \ldots, C_m$ over n variables $x_1, \ldots, x_n$ over alphabet $[q]$, find an assignment of $x_1, \ldots, x_n$ to $[q]$ such that maximum number of constraints are satisfied.
- Each constraint C_i on subset T_i is a function from $[q]^{T_i}$ to $\{0, 1\}$.
- An assignment *satisfies* C_i if the evaluation of C_i on the assignment restricted to T_i is 1.
- General program:

$$
\begin{aligned}
\text{Maximize} \quad & \sum_{i=1}^{m} \sum_{\alpha \in [q]^{T_i}} C_i(\alpha) \prod_{j \in T_i} y_{(j,\alpha_j)} & \\
\text{subject to} \quad & \sum_{\alpha_j \in [q]} y_{(j,\alpha_j)} = 1 & \forall j \in [n] \\
& y_{(j,\alpha_j)} y_{(j,\alpha'_j)} = 0 & \forall \alpha_j \neq \alpha'_j \\
& y_{(j,\alpha_j)} \in \{0, 1\} &
\end{aligned}
$$

Max K-CSP - SoS hardness construction

Max K-CSP - SoS hardness construction

- In our construction, we fix a prime power q and a subset $\mathcal{C} \subseteq \mathbb{F}_q^K$. Each constraint P on the K-subset C, for some $b \in \mathbb{F}_q^K$, is of the form $P(x) = [\text{Is } x_C - b \in \mathcal{C}?]$.

Max K-CSP - SoS hardness construction

- In our construction, we fix a prime power q and a subset $\mathcal{C} \subseteq \mathbb{F}_q^K$. Each constraint P on the K-subset C, for some $b \in \mathbb{F}_q^K$, is of the form $P(x) = [\text{Is } x_C - b \in \mathcal{C}?]$.
- τ, η, ζ are parameters.
  - $\mathcal{C}$ is $(\tau - 1)$-wise uniform
  - ηn is roughly the number of levels of SoS
  - ζ is slack, think $1/\log n$

Max K-CSP - SoS hardness construction

- In our construction, we fix a prime power q and a subset $\mathcal{C} \subseteq \mathbb{F}_q^K$. Each constraint P on the K-subset C, for some $b \in \mathbb{F}_q^K$, is of the form $P(x) = [\text{Is } x_C - b \in \mathcal{C}?]$.
- τ, η, ζ are parameters.
  - $\mathcal{C}$ is $(\tau - 1)$-wise uniform
  - ηn is roughly the number of levels of SoS
  - ζ is slack, think $1/\log n$
- *Random instance*: For a fixed $\mathcal{C}$, choose the m constraints independently as follows - Choose the K-subset u.a.r. and choose $b \in \mathbb{F}_q^K$ u.a.r.

Max K-CSP - associated graphs

Max K-CSP - associated graphs

- *Factor Graph* G_I: Bipartite graph with
  - $L = \{C_i \mid i \in [m]\}$
  - $R = \{x_j \mid j \in [n]\}$
  - (C_i, x) is an edge $\iff x \in C_i$.

Max K-CSP - associated graphs

- *Factor Graph* G_I: Bipartite graph with
  - $L = \{C_i \mid i \in [m]\}$
  - $R = \{x_j \mid j \in [n]\}$
  - (C_i, x) is an edge $\iff x \in C_i$.
- *The Label Extended Factor graph* $H_{I,\beta}$: Bipartite graph with
  - $L = \{(C_i, \alpha) \mid i \in [m], \alpha \in [q]^K, C_i(\alpha) = 1\}$
  - $R = \{(x_i, \alpha_{x_i}, j) \mid i \in [n], \alpha_{x_i} \in [q], j \in [\beta]\}$
  - $((C_i, \alpha), (x, \alpha_x, j))$ is an edge $\iff x \in C_i$ and α assigns x to α_x.

Max K-CSP - Plausibility Assumption

Max K-CSP - Plausibility Assumption

- *Density* of instance $\Delta = m/n$.
- τ-*subgraph*: A subgraph of G_I with no isolated vertices, such that each constraint vertex has degree at least τ.

Max K-CSP - Plausibility Assumption

- *Density* of instance $\Delta = m/n$.
- τ-*subgraph*: A subgraph of G_I with no isolated vertices, such that each constraint vertex has degree at least τ.
- A τ-subgraph H with c constraint vertices, v variable vertices and e edges is *plausible* if $v \geq e - \frac{\tau - \zeta}{2} c$
- *Plausibility assumption*: All τ-subgraphs H of G_I with at most $2\eta n$ constraint variables are plausible.

Max K-CSP - Plausibility Assumption

- *Density* of instance $\Delta = m/n$.
- τ*-subgraph*: A subgraph of G_I with no isolated vertices, such that each constraint vertex has degree at least τ.
- A τ-subgraph H with c constraint vertices, v variable vertices and e edges is *plausible* if $v \geq e - \frac{\tau-\zeta}{2}c$
- *Plausibility assumption*: All τ-subgraphs H of G_I with at most $2\eta n$ constraint variables are plausible.
- Theorem: With high probability, G_I for a random Max K-CSP instance will satisfy the Plausibility assumption with $\eta = \frac{1}{K}\left(\frac{1}{2^{K/(\tau-2)}}\right)^{O(1)} \cdot \frac{1}{\Delta^{2/(\tau-2-\zeta)}}$

  - Kothari, Mori, O'Donnell, Witmer[KMOW17]

Max K-CSP - SoS Hardness

Max K-CSP - SoS Hardness

- Theorem[KMOW17]: If the Plausibility assumption holds, then, for a degree $O(\zeta\eta n)$ SoS relaxation, $FRAC = m$.
- $OPT \approx m|\mathcal{C}|/q^K$ with high probability.

Max K-CSP - SoS Hardness

- Theorem[KMOW17]: If the Plausibility assumption holds, then, for a degree $O(\zeta\eta n)$ SoS relaxation, $FRAC = m$.
- $OPT \approx m|\mathcal{C}|/q^K$ with high probability.
- Corollary: For a random Max K-CSP instance, the level $O\left(\frac{1}{K}\left(\frac{1}{2^{K/(\tau-2)}}\right)^{O(1)} \cdot \frac{n}{\Delta^{2/(\tau-2-\zeta)}}\right)$ SoS relaxation will have $FRAC = m$, with high probability.

Max K-CSP - SoS Hardness

- Theorem[KMOW17]: If the Plausibility assumption holds, then, for a degree $O(\zeta\eta n)$ SoS relaxation, $FRAC = m$.
- $OPT \approx m|\mathcal{C}|/q^K$ with high probability.
- Corollary: For a random Max K-CSP instance, the level $O\left(\frac{1}{K}\left(\frac{1}{2^{K/(\tau-2)}}\right)^{O(1)} \cdot \frac{n}{\Delta^{2/(\tau-2-\zeta)}}\right)$ SoS relaxation will have $FRAC = m$, with high probability.
- Theorem: For a random Max K-CSP instances over boolean predicates, the level $\tilde{O}(n/\Delta^{2/(\tau-2)})$ SoS relaxation will have $FRAC < m$, with high probability.

  - Allen, O'Donnell and Witmer[AOW15]; Raghavendra, Rao and Schramm[RRS17]

Max K-CSP for superconstant K

Max K-CSP for superconstant K

- We had $\eta n = O\left(\frac{1}{K}\left(\frac{1}{2^{K/(\tau-2)}}\right)^{O(1)} \cdot \frac{n}{\Delta^{2/(\tau-2-\zeta)}}\right)$.
- Exponential dependence on K, not suitable for some applications like Densest k-subgraph.

Max K-CSP for superconstant K

- We had $\eta n = O\left(\frac{1}{K}\left(\frac{1}{2^{K/(\tau-2)}}\right)^{O(1)} \cdot \frac{n}{\Delta^{2/(\tau-2-\zeta)}}\right)$.
- Exponential dependence on K, not suitable for some applications like Densest k-subgraph.
- Theorem: If $\mathcal{C}$ supports a pairwise independent distribution, and if
  - $10 \leq K \leq \sqrt{n}$.
  - $n^{\nu-1} \leq O(1/((\Delta K^{D+0.75})^{2/(D-2)})$ for some $\nu > 0$.

  Then, with high probability, for a random Max K-CSP instance, the level $O\left(\frac{n}{(\Delta K^{D})^{2/(D-2)}}\right)$ SoS relaxation will have $FRAC = m$.

  - Bhaskara, Charikar, Guruswami, Vijayaraghavan, Zhou[BCG$^+$12];

Max K-CSP for superconstant K - Our results

Max K-CSP for superconstant K - Our results

- Theorem: If
  - $\tau \geq 4$.
  - $0 < \zeta < 0.99(\tau - 2)$.
  - $10 \leq K \leq \sqrt{n}$.
  - $n^{\nu-1} \leq 1/(10^8(\Delta K^{\tau-\zeta+0.75})^{2/(\tau-\zeta-2)})$ for some $\nu > 0$.

  Then, with high probability, for a random Max K-CSP instance, the level $O\left(\frac{n}{(\Delta K^{\tau-\zeta})^{2/(\tau-\zeta-2)}}\right)$ SoS relaxation will have $FRAC = m$.

Max K-CSP for superconstant K - Our results

- Theorem: If
  - $\tau \geq 4$.
  - $0 < \zeta < 0.99(\tau - 2)$.
  - $10 \leq K \leq \sqrt{n}$.
  - $n^{\nu-1} \leq 1/(10^8(\Delta K^{\tau-\zeta+0.75})^{2/(\tau-\zeta-2)})$ for some $\nu > 0$.

  Then, with high probability, for a random Max K-CSP instance, the level $O\left(\frac{n}{(\Delta K^{\tau-\zeta})^{2/(\tau-\zeta-2)}}\right)$ SoS relaxation will have $FRAC = m$.
- Proof idea:
  - Use a lemma implicitly shown in [BCG$^+$12], on the expansion properties of G_I.
  - Prove that these expansion properties imply the Plausibility assumption.

Densest k-subgraph

Densest k-subgraph

- Theorem: The level $O(1/\epsilon)$ SoS relaxation gives a $n^{1/4+\epsilon}$ approximation for any $\epsilon > 0$.
  - Bhaskara, Charikar, Chlamtáč, Feige, Vijayaraghavan[BCC$^+$12]

Densest k-subgraph

- Theorem: The level $O(1/\epsilon)$ SoS relaxation gives a $n^{1/4+\epsilon}$ approximation for any $\epsilon > 0$.
  - Bhaskara, Charikar, Chlamtáč, Feige, Vijayaraghavan[BCC+12]
- Theorem: The integrality gap of the level $n^{\Omega(\epsilon)}$ SoS relaxation is at least $\Omega(n^{1/14-\epsilon})$ for any $\epsilon > 0$.
  - Bhaskara, Charikar, Guruswami, Vijayaraghavan, Zhou[BCG+12];

  Manurangsi[Man15]

Densest k-subgraph - SoS Hardness

Densest k-subgraph - SoS Hardness

- Idea: Reduction from Max K-CSP.
- Integrality gap construction: For a random instance I of Max K-CSP, consider an instance Γ of Densest k-subgraph with the graph being $G = H_{I,\Delta}$ and $k = 2m$.

Densest k-subgraph - SoS Hardness

- Idea: Reduction from Max K-CSP.
- Integrality gap construction: For a random instance I of Max K-CSP, consider an instance Γ of Densest k-subgraph with the graph being $G = H_{I,\Delta}$ and $k = 2m$.
- Completeness lemma[BCG$^+$12]: If level-r SoS relaxation for I has $FRAC = m$, then the level r/K SoS relaxation for Γ has $FRAC' \geq \Delta m K$.
- Soudness lemma[Man15]: For suitable choice of parameters, Γ has $OPT' \leq O(\Delta m K \ln q/q)$ with high probability.

Densest k-subhypergraph

Densest k-subhypergraph

- Given a hypergraph G and a positive integer k, find a subset W of vertices with exactly k vertices that maximizes the number of edges $e \in E$ with $e \subseteq W$.

Densest k-subhypergraph

- Given a hypergraph G and a positive integer k, find a subset W of vertices with exactly k vertices that maximizes the number of edges $e \in E$ with $e \subseteq W$.
- For 3-uniform hypergraphs, there is a $O(n^{4(4-\sqrt{3})/13+\epsilon})$ approximation.
    - Chlamtáč, Dinitz, Konrad, Kortsarz and Rabanca[CDK$^+$16]

Densest k-subhypergraph

- Given a hypergraph G and a positive integer k, find a subset W of vertices with exactly k vertices that maximizes the number of edges $e \in E$ with $e \subseteq W$.
- For 3-uniform hypergraphs, there is a $O(n^{4(4-\sqrt{3})/13+\epsilon})$ approximation.
  - Chlamtáč, Dinitz, Konrad, Kortsarz and Rabanca[CDK+16]
- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{F \in E} \prod_{u \in F} x_u \\ \text{subject to} \quad & \sum_{u \in V} x_u = k \\ & x_u \in \{0, 1\} \end{aligned}$$

Densest k-subhypergraph - SoS Hardness - Our results

Densest k-subhypergraph - SoS Hardness - Our results

- Theorem: Integrality gap of level-r SoS relaxation for Densest k-subgraph $= \alpha(n) \implies$ Integrality gap of level-r SoS relaxation for Densest k-subhypergraph of arity 2^t is $\geq (\alpha(n)/2^{t+2})^{2^{t-1}}$

Densest k-subhypergraph - SoS Hardness - Our results

- Theorem: Integrality gap of level-r SoS relaxation for Densest k-subgraph $= \alpha(n) \implies$ Integrality gap of level-r SoS relaxation for Densest k-subhypergraph of arity 2^t is $\geq (\alpha(n)/2^{t+2})^{2^{t-1}}$
- Idea: Reduction from Densest k-subgraph
- Construction:
  - Take instance $I = ((V, E), k)$ of Densest k-subgraph.
  - Construct hypergraph $G' = (V, E')$ where each element of E' is obtained by taking union of 2^{t-1} edges in E.
  - We consider the instance $J = (G', k)$ on n vertices.

Densest k-subhypergraph - SoS Hardness proof

Densest k-subhypergraph - SoS Hardness proof

- Completeness lemma: $FRAC' \geq \frac{FRAC^{2^{t-1}}}{(2^t)^{2^t}}$
- Main claim: For an integer $p \geq 0$, let $T = E^{2^p}$ be the set of ordered tuples of 2^p edges. Then,

$$\sum_{(f_1,\ldots,f_{2^p})\in T} \|\boldsymbol{V}_{f_1\cup\ldots\cup f_{2^p}}\|^2 \geq FRAC^{2^p}$$

- Soundness lemma: $OPT' \leq OPT^{\rho}$

Densest k-subhypergraph - SoS Hardness proof

- Completeness lemma: $FRAC' \geq \frac{FRAC^{2^{t-1}}}{(2^t)^{2^t}}$
- Main claim: For an integer $p \geq 0$, let $T = E^{2^p}$ be the set of ordered tuples of 2^p edges. Then,

$$\sum_{(f_1,\ldots,f_{2^p}) \in T} \|\boldsymbol{V}_{f_1 \cup \ldots \cup f_{2^p}}\|^2 \geq FRAC^{2^p}$$

- Soundness lemma: $OPT' \leq OPT^{\rho}$
- Corollary: For any integer $\rho \geq 2$, $n^{\Omega(\epsilon)}$ levels of the SoS hierarchy has an integrality gap of at least $\Omega(n^{(2^{\lfloor \log \rho \rfloor}/28)}) \geq \Omega(n^{\rho/56})$ for Densest k-subhypergraph on n vertices of arity ρ

Minimum p-Union

Minimum p-Union

- Given integer p and m subsets $S_1, \ldots, S_m$ of $[n]$, choose exactly p of these sets such that the size of their union is minimized.

Minimum p-Union

- Given integer p and m subsets $S_1, \ldots, S_m$ of $[n]$, choose exactly p of these sets such that the size of their union is minimized.
- SSBVE formulation: Given integer l and a bipartite graph $G = (L, R, E)$, choose exactly l vertices from L such that the size of the neighborhood of these l vertices is minimized.

Minimum p-Union

- Given integer p and m subsets $S_1, \ldots, S_m$ of $[n]$, choose exactly p of these sets such that the size of their union is minimized.
- SSBVE formulation: Given integer l and a bipartite graph $G = (L, R, E)$, choose exactly l vertices from L such that the size of the neighborhood of these l vertices is minimized.
- $O(m^{1/4})$ approximation by Chlamtáč, Dinitz and Makarychev[CDM17]
- General program:

$$
\begin{aligned}
\text{Minimize} \quad & \sum_{v \in R} x_v \\
\text{subject to} \quad & \sum_{u \in L} x_u = l \\
& x_u \leq x_v \qquad \forall (u, v) \in E, u \in L, v \in R \\
& x_u, x_v \in \{0, 1\}
\end{aligned}
$$

Minimum p-Union - SoS Hardness - Our results

Minimum p-Union - SoS Hardness - Our results

- Theorem: The integrality gap of the level $m^{\Omega(\epsilon)}$ SoS relaxation is at least $\Omega(m^{1/18-\epsilon})$ for any $\epsilon > 0$.
- Idea: Reduction from Max K-CSP.

Minimum p-Union - SoS Hardness - Our results

- Theorem: The integrality gap of the level $m^{\Omega(\epsilon)}$ SoS relaxation is at least $\Omega(m^{1/18-\epsilon})$ for any $\epsilon > 0$.
- Idea: Reduction from Max K-CSP.
- Construction:
  - Take a random instance I of Max K-CSP and consider the label extended factor graph $H_{I,\Delta}$.
  - Subdivide the edges to obtain H.
  - The new instance of SSBVE is $J = (H, l)$ where $l = \Delta mK$.

For appropriate choice of parameters, we have

  - $FRAC' \geq 2m$
  - $OPT' \geq O(m\sqrt{q}/\sqrt{\ln q})$

Pseudoexpectations - Alternate view of SoS

Pseudoexpectations - Alternate view of SoS

- $P^{\leq r}[x_1, \ldots, x_n]$ - Set of polynomials of degree at most r in $\mathbb{R}[x_1, \ldots, x_n]$

Pseudoexpectations - Alternate view of SoS

- $P^{\leq r}[x_1, \ldots, x_n]$ - Set of polynomials of degree at most r in $\mathbb{R}[x_1, \ldots, x_n]$
- $\tilde{E} : P^{\leq 2r}[x_1, \ldots, x_n] \longrightarrow \mathbb{R}$ is a degree $2r$ *pseudoexpectation operator* if
  - Normalization: $\tilde{E}[1] = 1$
  - Linearity: $\tilde{E}$ is linear.
  - Positivity: $\tilde{E}[p^2] \geq 0$ for every $p \in P^{\leq r}[x_1, \ldots, x_n]$

SoS relaxation

SoS relaxation

- General program Γ:

$$\begin{array}{lll} \text{Maximize} & p(x_1, \ldots, x_n) & \\ \text{subject to} & q_i(x_1, \ldots, x_n) = 0 & i = 1, 2, \ldots, m \\ & x_i \in \{0, 1\} & \end{array}$$

- Level-r SoS relaxation $\mathcal{P}_r$:

$$\begin{array}{lll} \text{Maximize} & \sum_{T \in [n]_{\leq r}} p_T \|\boldsymbol{V}_T\|^2 & \\ \text{subject to} & \sum_{T \in [n]_{\leq r}} (q_i)_T \langle \boldsymbol{V}_T, \boldsymbol{V}_S \rangle = 0 & \forall S \in [n]_{\leq r}, i = 1, \ldots, m \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle = \langle \boldsymbol{V}_{S_3}, \boldsymbol{V}_{S_4} \rangle & \forall S_1 \cup S_2 = S_3 \cup S_4 \text{ and } S_i \in [n]_{\leq r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \geq 0 & \forall S_1, S_2 \in [n]_{\leq r} \\ & \|\boldsymbol{V}_\phi\|^2 = 1 & \end{array}$$

Pseudoexpectation operator program

Pseudoexpectation operator program

- General program Γ:

$$\begin{array}{lll} \text{Maximize} & p(x_1, \ldots, x_n) & \\ \text{subject to} & q_i(x_1, \ldots, x_n) = 0 & i = 1, 2, \ldots, m \\ & x_i \in \{0, 1\} & \end{array}$$

- Degree $2r$ pseudoexpectation operator program $\mathcal{Q}_{2r}$:

$$\begin{array}{lll} \text{Maximize} & \tilde{E}[p] & \\ \text{subject to} & \tilde{E}[q_i h] = 0 & \forall h \text{ such that } q_i h \in P^{\leq 2r}[x_1, \ldots, x_n], i \in [m] \\ & \tilde{E}[(x_i^2 - x_i)h] = 0 & \forall h \in P^{\leq 2r-2}[x_1, \ldots, x_n], i \in [n] \\ & \tilde{E} \text{ is a degree } 2r \text{ pseudoexpectation operator} & \end{array}$$

Equivalence between SoS and Pseudoexpectations

Equivalence between SoS and Pseudoexpectations

- SoS to Pseudoexpectation programs:

  $\mathcal{P}_{2r}$ has a feasible solution of value $FRAC$

  $\implies \mathcal{Q}_{2r}$ has a feasible solution of value $FRAC$

- Pseudoexpectation programs to SoS:

  $\mathcal{Q}_{4r}$ has a feasible solution of value $FRAC$

  $\implies \mathcal{P}_{r}$ has a feasible solution of value $FRAC$

Equivalence between SoS and Pseudoexpectations

- SoS to Pseudoexpectation programs:

$$\mathcal{P}_{2r} \text{ has a feasible solution of value } FRAC \implies \mathcal{Q}_{2r} \text{ has a feasible solution of value } FRAC$$

- Pseudoexpectation programs to SoS:

$$\mathcal{Q}_{4r} \text{ has a feasible solution of value } FRAC \implies \mathcal{P}_{r} \text{ has a feasible solution of value } FRAC$$

- Means we can work with either program interchangeably upto a constant loss in the level

SoS hardness for MaxClique

SoS hardness for MaxClique

- Theorem[BHK$^+$16]: If $r = o(\log n)$, the level-r SoS relaxation for MaxClique will have $FRAC \geq k = n^{1/2 - O(\sqrt{r/\log n})}$ on $G \sim G(n, 1/2)$ with high probability.

SoS hardness for MaxClique

- Theorem[BHK+16]: If $r = o(\log n)$, the level-r SoS relaxation for MaxClique will have $FRAC \geq k = n^{1/2 - O(\sqrt{r/\log n})}$ on $G \sim G(n, 1/2)$ with high probability.
- Idea: Exhibit a degree $2r$ pseudoexpectation operator $\tilde{E}$, that satisfies the following w.h.p. when $G \sim G(n, 1/2)$
  - $\tilde{E}$ is linear and $\tilde{E}[1] = 1$
  - $\tilde{E}[(x_u^2 - x_u)h] = 0$ for all $h \in P^{\leq 2r-2}[x_1, \ldots, x_n], u \in [n]$
  - $\tilde{E}[x_u x_v h] = 0$ for all $(u, v) \notin E, u \neq v, h \in P^{\leq 2r-2}[x_1, \ldots, x_n]$
  - $\sum_{u=1}^{n} \tilde{E}[x_u] = k$
  - $\tilde{E}[h^2] \geq 0$ for all $h \in P^{\leq r}[x_1, \ldots, x_n]$

Pseudocalibration for MaxClique - Planted distribution

Pseudocalibration for MaxClique - Planted distribution

- Think of $\tilde{E}$ as a computationally bounded solver
- $\tilde{E}$ "thinks" that $G(n, 1/2)$ has a clique of size k for $k \gg 2\log n$

Pseudocalibration for MaxClique - Planted distribution

- Think of $\tilde{E}$ as a computationally bounded solver
- $\tilde{E}$ "thinks" that $G(n, 1/2)$ has a clique of size k for $k \gg 2\log n$
- Assume $\tilde{E}$ cannot distinguish the following distributions:
  - Random distribution $G(n, 1/2)$ - G sampled from the Erdös-Rényi random graph distribution
  - Planted distribution $G(n, 1/2, k)$ - Sample $G \sim G(n, 1/2)$ and plant a clique on a random subset of k vertices.

Pseudocalibration for MaxClique - Heuristic 1

Pseudocalibration for MaxClique - Heuristic 1

- $\tilde{E}$ is unable to distinguish $G(n, 1/2)$ from $G(n, 1/2, k)$

Pseudocalibration for MaxClique - Heuristic 1

- $\tilde{E}$ is unable to distinguish $G(n, 1/2)$ from $G(n, 1/2, k)$
  - Expectations of $\tilde{E}[f]$ are the same for both distributions for any $f \in P^{\leq 2r}[x_1, \ldots, x_n]$.

$$\mathbb{E}_{G \sim G(n,1/2)} \tilde{E}_G[f] = \mathbb{E}_{G \sim G(n,1/2,k)} \tilde{E}_G[f]$$

Pseudocalibration for MaxClique - Heuristic 1

- $\tilde{E}$ is unable to distinguish $G(n, 1/2)$ from $G(n, 1/2, k)$
  - Expectations of $\tilde{E}[f]$ are the same for both distributions for any $f \in P^{\leq 2r}[x_1, \dots, x_n]$.

$$\mathbb{E}_{G \sim G(n,1/2)} \tilde{E}_G[f] = \mathbb{E}_{G \sim G(n,1/2,k)} \tilde{E}_G[f]$$

  - Correlations of $\tilde{E}[f]$ with low degree $g : \{\pm 1\}^{n(n-1)/2} \longrightarrow \mathbb{R}$ are the same for both distributions for any $f \in P^{\leq 2r}[x_1, \dots, x_n]$

$$\mathbb{E}_{G \sim G(n,1/2)}[\tilde{E}_G[f] g(G)] = \mathbb{E}_{G \sim G(n,1/2,k)}[\tilde{E}_G[f] g(G)]$$

Pseudocalibration for MaxClique - Heuristic 1

- $\tilde{E}$ is unable to distinguish $G(n, 1/2)$ from $G(n, 1/2, k)$
  - Expectations of $\tilde{E}[f]$ are the same for both distributions for any $f \in P^{\leq 2r}[x_1, \ldots, x_n]$.

$$\mathbb{E}_{G \sim G(n,1/2)} \tilde{E}_G[f] = \mathbb{E}_{G \sim G(n,1/2,k)} \tilde{E}_G[f]$$

  - Correlations of $\tilde{E}[f]$ with low degree $g : \{\pm 1\}^{n(n-1)/2} \longrightarrow \mathbb{R}$ are the same for both distributions for any $f \in P^{\leq 2r}[x_1, \ldots, x_n]$

$$\mathbb{E}_{G \sim G(n,1/2)}[\tilde{E}_G[f] g(G)] = \mathbb{E}_{G \sim G(n,1/2,k)}[\tilde{E}_G[f] g(G)]$$

- In the second condition, $\tilde{E}[f]$ is treated as a function on graphs, from $\{\pm 1\}^{n(n-1)/2}$ to $\mathbb{R}$.

Pseudocalibration for MaxClique - Heuristic 2

Pseudocalibration for MaxClique - Heuristic 2

- $\tilde{E}$ is the correct expectation on $G \sim G(n, 1/2, k)$ with a unique support being the indicator vector $\boldsymbol{x} \in \mathbb{R}^n$ of the planted clique

$$\mathbb{E}_{G\sim G(n,1/2,k)}[\tilde{E}_G[f]g(G)] = \mathbb{E}_{(G,\boldsymbol{x})\sim G(n,1/2,k)}[f(\boldsymbol{x})g(G)]$$

Pseudocalibration for MaxClique - Heuristic 2

- $\tilde{E}$ is the correct expectation on $G \sim G(n, 1/2, k)$ with a unique support being the indicator vector $\boldsymbol{x} \in \mathbb{R}^n$ of the planted clique

$$\mathbb{E}_{G \sim G(n,1/2,k)}[\tilde{E}_G[f]g(G)] = \mathbb{E}_{(G,\boldsymbol{x}) \sim G(n,1/2,k)}[f(\boldsymbol{x})g(G)]$$

- For all $f \in P^{\leq 2r}[x_1, \ldots, x_n]$ and low degree $g : \{\pm 1\}^{n(n-1)/2} \longrightarrow \mathbb{R}$,

$$\mathbb{E}_{G \sim G(n,1/2)}[\tilde{E}_G[f]g(G)] = \mathbb{E}_{(G,\boldsymbol{x}) \sim G(n,1/2,k)}[f(\boldsymbol{x})g(G)]$$

Pseudocalibration for MaxClique - Combining the heuristics

Pseudocalibration for MaxClique - Combining the heuristics

- Enough to define $\tilde{E}[\boldsymbol{x}_S]$ for all $S \in [n]_{\leq 2r}$ where $\boldsymbol{x}_S(\boldsymbol{x}) = \prod_{i \in S} x_i$.

Pseudocalibration for MaxClique - Combining the heuristics

- Enough to define $\tilde{E}[\boldsymbol{x}_S]$ for all $S \in [n]_{\leq 2r}$ where $\boldsymbol{x}_S(\boldsymbol{x}) = \prod_{i \in S} x_i$.
- For edge $e \in [n(n-1)/2]$, let

$$G_e = \begin{cases} 1 & \text{if } e \in E \\ -1 & \text{if } e \notin E \end{cases}$$

- Consider Fourier basis $\chi_T(G)$ for $T \subseteq [n(n-1)/2]$ where $\chi_T(G) = \prod_{e \in T} G_e$.

Pseudocalibration for MaxClique - Combining the heuristics

- Enough to define $\tilde{E}[\boldsymbol{x}_S]$ for all $S \in [n]_{\leq 2r}$ where $\boldsymbol{x}_S(\boldsymbol{x}) = \prod_{i \in S} x_i$.
- For edge $e \in [n(n-1)/2]$, let

$$G_e = \begin{cases} 1 & \text{if } e \in E \\ -1 & \text{if } e \notin E \end{cases}$$

- Consider Fourier basis $\chi_T(G)$ for $T \subseteq [n(n-1)/2]$ where $\chi_T(G) = \prod_{e \in T} G_e$.
- Suffices to ensure, for all $S \in [n]_{\leq 2r}$ and all $T \subseteq [n(n-1)/2]$,

$$\mathbb{E}_{G \sim G(n,1/2)}[\tilde{E}_G[\boldsymbol{x}_S]\chi_T(G)] = \mathbb{E}_{(G,\boldsymbol{x}) \sim G(n,1/2,k)}[\boldsymbol{x}_S(\boldsymbol{x})\chi_T(G)]$$

Pseudocalibration for MaxClique - Fourier coefficients

Pseudocalibration for MaxClique - Fourier coefficients

- For a fixed S,

$$\tilde{E}_G[\mathbf{x}_S] = \sum_{T \subseteq [n(n-1)/2]} \widehat{\tilde{E}[\mathbf{x}_S](T)} \chi_T(G)$$

Pseudocalibration for MaxClique - Fourier coefficients

- For a fixed S,

$$\tilde{E}_G[\boldsymbol{x}_S] = \sum_{T \subseteq [n(n-1)/2]} \widehat{\tilde{E}[\boldsymbol{x}_S](T)} \chi_T(G)$$

$$\begin{aligned} \widehat{\tilde{E}[\boldsymbol{x}_S](T)} &= \mathbb{E}_{G \sim G(n,1/2)}[\tilde{E}_G[\boldsymbol{x}_S] \chi_T(G)] \\ &= \mathbb{E}_{(G,\boldsymbol{x}) \sim G(n,1/2,k)}[\boldsymbol{x}_S(\boldsymbol{x}) \chi_T(G)] \end{aligned}$$

Pseudocalibration for MaxClique - Fourier coefficients

- For a fixed S,

$$\tilde{E}_G[\mathbf{x}_S] = \sum_{T \subseteq [n(n-1)/2]} \widehat{\tilde{E}[\mathbf{x}_S](T)} \chi_T(G)$$

$$\begin{aligned}
\widehat{\tilde{E}[\mathbf{x}_S](T)} &= \mathbb{E}_{G \sim G(n,1/2)}[\tilde{E}_G[\mathbf{x}_S]\chi_T(G)] \\
&= \mathbb{E}_{(G,\mathbf{x}) \sim G(n,1/2,k)}[\mathbf{x}_S(\mathbf{x})\chi_T(G)] \\
&= \Pr[\text{Planted Clique contains } S \cup V(T)] \\
&= \frac{\binom{n - |S \cup V(T)|}{k - |S \cup V(T)|}}{\binom{n}{k}} \\
&\approx \left(\frac{k}{n}\right)^{|S \cup V(T)|}
\end{aligned}$$

Pseudocalibration for MaxClique - Final pseudoexpectation

Pseudocalibration for MaxClique - Final pseudoexpectation

- One more heuristic: Set $\widehat{\tilde{E}[x_S](T)} = 0$ for all subsets T such that $|S \cup V(T)| > \tau$
- Threshold τ restricts "power" of $\tilde{E}$
- [BHK$^+$16] set $\tau \approx r/\epsilon$ where $k \approx n^{1/2-\epsilon}$

Pseudocalibration for MaxClique - Final pseudoexpectation

- One more heuristic: Set $\widehat{\tilde{E}[\boldsymbol{x}_S](T)} = 0$ for all subsets T such that $|S \cup V(T)| > \tau$
- Threshold τ restricts "power" of $\tilde{E}$
- [BHK+16] set $\tau \approx r/\epsilon$ where $k \approx n^{1/2-\epsilon}$
- Final pseudoexpectation: If $f(\boldsymbol{x}) = \sum_{S \in [n]_{\leq 2r}} c_S \boldsymbol{x}_S$, then

$$\tilde{E}[f] = \sum_{S \in [n]_{\leq 2r}} c_S \sum_{|S \cup V(T)| \leq \tau, T \subseteq [n(n-1)/2]} \left(\frac{k}{n}\right)^{|S \cup V(T)|} \chi_T(G)$$

for the graph G.

Future work

Future work

- The Level $O(1/\epsilon)$ Lovász-Schrijver hierarchy gives a $n^{1/4+\epsilon}$ approximation for the Densest k-subgraph problem. Open to analyze performance of SoS for the Densest k-subhypergraph problem.

Future work

- The Level $O(1/\epsilon)$ Lovász-Schrijver hierarchy gives a $n^{1/4+\epsilon}$ approximation for the Densest k-subgraph problem. Open to analyze performance of SoS for the Densest k-subhypergraph problem.
- Known lower bounds on integrality gap for the polynomial level SoS relaxation:
  - $n^{1/14-\epsilon}$ for Densest k-subgraph.
  - $m^{1/18-\epsilon}$ for Minimum p-Union.
- Both are not tight.

Future work

- The Level $O(1/\epsilon)$ Lovász-Schrijver hierarchy gives a $n^{1/4+\epsilon}$ approximation for the Densest k-subgraph problem. Open to analyze performance of SoS for the Densest k-subhypergraph problem.
- Known lower bounds on integrality gap for the polynomial level SoS relaxation:
  - $n^{1/14-\epsilon}$ for Densest k-subgraph.
  - $m^{1/18-\epsilon}$ for Minimum p-Union.
- Both are not tight.
- Known lower bounds on integrality gap for the $\Omega(\log n/\log\log n)$ Sherali-Adams relaxation:
  - $n^{1/4}$ for Densest k-subgraph.
  - $m^{1/4}$ for Minimum p-Union.

Future work

- The Level $O(1/\epsilon)$ Lovász-Schrijver hierarchy gives a $n^{1/4+\epsilon}$ approximation for the Densest k-subgraph problem. Open to analyze performance of SoS for the Densest k-subhypergraph problem.
- Known lower bounds on integrality gap for the polynomial level SoS relaxation:
  - $n^{1/14-\epsilon}$ for Densest k-subgraph.
  - $m^{1/18-\epsilon}$ for Minimum p-Union.
- Both are not tight.
- Known lower bounds on integrality gap for the $\Omega(\log n/\log\log n)$ Sherali-Adams relaxation:
  - $n^{1/4}$ for Densest k-subgraph.
  - $m^{1/4}$ for Minimum p-Union.
- Pseudocalibration could be applied but it is open to analyze the operators so obtained.

Thank You

Minimum Bisection - SoS relaxation

- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{(u,v)\in E} (x_u - x_v)^2 \\ \text{subject to} \quad & \sum_{u\in V} x_u = k \\ & x_u \in \{0,1\} \end{aligned}$$

- Level-r SoS relaxation:

$$\begin{aligned} \text{Minimize} \quad & \sum_{(u,v)\in E} \|\boldsymbol{V}_{\{u\}} - \boldsymbol{V}_{\{v\}}\|^2 \\ \text{subject to} \quad & \sum_{v\in V} \langle \boldsymbol{V}_{\{v\}}, \boldsymbol{V}_S \rangle = k\|\boldsymbol{V}_S\|^2 && \forall S \in [n]_{\leq r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle = \langle \boldsymbol{V}_{S_3}, \boldsymbol{V}_{S_4} \rangle && \forall S_1 \cup S_2 = S_3 \cup S_4 \in [n]_{\leq r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \geq 0 && \forall S_1, S_2 \in [n]_{\leq r} \\ & \|\boldsymbol{V}_\phi\|^2 = 1 \end{aligned}$$

Max K-CSP - SoS relaxation

- Level-r SoS relaxation:

$$
\begin{aligned}
\text{Maximize} \quad & \sum_{i=1}^{m} \sum_{\alpha \in [q]^{T_i}} C_i(\alpha) \| \boldsymbol{V}_{(T_i,\alpha)} \|^2 & \\
\text{subject to} \quad & \langle \boldsymbol{V}_{(S_1,\alpha_1)}, \boldsymbol{V}_{(S_2,\alpha_2)} \rangle = 0 & \forall \alpha_1(S_1 \cap S_2) \neq \alpha_2(S_1 \cap S_2), S_1, S_2 \in [n]_{\leq r} \\
& \langle \boldsymbol{V}_{(S_1,\alpha_1)}, \boldsymbol{V}_{(S_2,\alpha_2)} \rangle = \langle \boldsymbol{V}_{(S_3,\alpha_3)}, \boldsymbol{V}_{(S_4,\alpha_4)} \rangle & \forall S_1 \cup S_2 = S_3 \cup S_4, \alpha_1 \circ \alpha_2 = \alpha_3 \circ \alpha_4, S_i \in [n]_{\leq r} \\
& \sum_{\alpha \in [q]} \langle \boldsymbol{V}_{\{j\},[j \to \alpha]}, \boldsymbol{V}_S \rangle = \| \boldsymbol{V}_S \|^2 & \forall S \in [n]_{\leq r}, j \in [n] \\
& \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \geq 0 & \forall S_1, S_2 \in [n]_{\leq r} \\
& \| \boldsymbol{V}_{\phi} \|^2 = 1 &
\end{aligned}
$$

Densest k-subhypergraph - SoS relaxation

- General program:

$$\begin{aligned} \text{Maximize} \quad & \sum_{F \in E} \prod_{u \in F} x_u \\ \text{subject to} \quad & \sum_{u \in V} x_u = k \\ & x_u \in \{0, 1\} \end{aligned}$$

- Level-r SoS relaxation:

$$\begin{aligned} \text{Maximize} \quad & \sum_{F \in E} \| \boldsymbol{V}_F \|^2 & \\ \text{subject to} \quad & \sum_{v \in V} \langle \boldsymbol{V}_{\{v\}}, \boldsymbol{V}_S \rangle = k \| \boldsymbol{V}_S \|^2 & \forall S \in [n]_{\leq r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle = \langle \boldsymbol{V}_{S_3}, \boldsymbol{V}_{S_4} \rangle & \forall S_1 \cup S_2 = S_3 \cup S_4 \text{ and } S_i \in [n]_{\leq r} \\ & \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \geq 0 & \forall S_1, S_2 \in [n]_{\leq r} \\ & \| \boldsymbol{V}_\phi \|^2 = 1 & \end{aligned}$$

Minimum p-Union - SoS relaxation

- General program:

$$\begin{aligned}
\text{Minimize} \quad & \sum_{v \in R} x_v \\
\text{subject to} \quad & \sum_{u \in L} x_u = l \\
& x_u \leq x_v \qquad \forall (u,v) \in E, u \in L, v \in R \\
& x_u, x_v \in \{0,1\}
\end{aligned}$$

- Level-r SoS relaxation:

$$\begin{aligned}
\text{Minimize} \quad & \sum_{v \in R} \|\boldsymbol{V}_{\{v\}}\|^2 \\
\text{subject to} \quad & \sum_{u \in L} \langle \boldsymbol{V}_{\{u\}}, \boldsymbol{V}_S \rangle = l \|\boldsymbol{V}_S\|^2 \quad \forall S \in [n]_{\leq r} \\
& \langle \boldsymbol{V}_{\{u\}}, \boldsymbol{V}_S \rangle \leq \langle \boldsymbol{V}_{\{v\}}, \boldsymbol{V}_S \rangle \quad \forall (u,v) \in E, u \in L, v \in R, S \in [n]_{\leq r} \\
& \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle = \langle \boldsymbol{V}_{S_3}, \boldsymbol{V}_{S_4} \rangle \quad \forall S_1 \cup S_2 = S_3 \cup S_4 \text{ and } S_i \in [n]_{\leq r} \\
& \langle \boldsymbol{V}_{S_1}, \boldsymbol{V}_{S_2} \rangle \geq 0 \quad \forall S_1, S_2 \in [n]_{\leq r} \\
& \|\boldsymbol{V}_\phi\|^2 = 1
\end{aligned}$$

Low threshold-rank graphs

- Graph G is *low threshold-rank* if the normalized adjacency matrix A has very few eigenvalues more than a positive constant.
- Example: Only one eigenvalue more than 0.5 means graph is an expander.
- Low threshold rank graphs roughly look like a union of expanders.
  - Gharan and Trevisan[GT14]
- Good approximation for many graph theoretic problems on such graphs due to Guruswami and Sinop[GS11]; and Barak, Raghavendra and Steurer[BRS11]