Lecture 8: SOS Lower Bound for 3-XOR

Lecture Outline

- Part I: SOS Lower Bounds from Pseudo-expectation Values
- Part II: Random 3-XOR Equations and Pseudo-expectation Values
- Part III: Proving PSDness
- Part IV: Analyzing Parameter Regimes
- Part V: Gaussian Elimination and SOS
- Part VI: Further Work

Part I: SOS Lower Bounds from Pseudo-expectation Values

Positivstellensatz Proofs Review

- Recall: a degree d Positivstellensatz proof that constraints $s_1(x_1, \ldots, x_n) = 0, s_1(x_1, \ldots, x_n) = 0,$ etc. are infeasible is an expression of the form $-1 = \sum_i f_i s_i + \sum_j g_j^2$ where:
  1. $\forall i, \deg(f_i) + \deg(s_i) \leq d$
  2. $\forall j, \deg(g_j) \leq \frac{d}{2}$
- How do we show that there is no degree d Positivstellensatz proof of infeasibility?

Positivstellensatz Proofs Review

- Recall: a degree d Positivstellensatz proof that $h(x_1, \ldots, x_n) \geq c$ given constraints $s_1(x_1, \ldots, x_n) = 0, s_1(x_1, \ldots, x_n) = 0$, etc. is an expression of the form $h = c + \sum_i f_i s_i + \sum_j g_j^2$ where:
  1. $\forall i, \deg(f_i) + \deg(s_i) \leq d$
  2. $\forall j, \deg(g_j) \leq \frac{d}{2}$
- How do we show that there is no degree d Positivstellensatz proof that $h(x_1, \ldots, x_n) \geq c$?

Pseudo-expectation Values Review

- Recall: Given constraints $s_1(x_1, \dots, x_n) = 0, s_1(x_1, \dots, x_n) = 0$, etc., degree d Pseudo-expectation values consist of a linear map $\tilde{E}$ from polynomials of degree $\leq d$ to $\mathbb{R}$ such that:
  1. $\tilde{E}[1] = 1$
  2. $\forall f, i, \tilde{E}[f s_i] = 0$ whenever $\deg(f_i) + \deg(s_i) \leq d$
  3. $\forall g, \tilde{E}\left[g^2\right] \geq 0$ whenever $\deg(g) \leq \frac{d}{2}$
- The third condition is equivalent to $M \succcurlyeq 0$ where M is the moment matrix with entries $M_{pq} = \tilde{E}[pq]$

SOS Lower Bound Strategy

- Recall: degree d pseudo-expectation values imply there is no degree d Positivstellensatz proof of infeasibility
- Analogously, degree d pseudo-expectation values with $\tilde{E}[h] < c$ imply there is no degree d Positivstellensatz proof that $h \geq c$.
- Proof: can assume both exist and get the following contradiction:

$$c > \tilde{\mathrm{E}}[h] = \tilde{E}[c] + \sum_i \tilde{E}[f_i s_i] + \sum_j \tilde{E}\left[g_j^2\right] \geq c$$

SOS Lower Bound Strategy

- To prove an SOS lower bound, we generally do the following:
  1. Come up with pseudo-expectation values $\tilde{E}$ which obey the required linear equations
  2. Show that the moment matrix M is PSD
- In the examples we'll see, part 1 is relatively easy and the technical part is part 2.
- That said, for several very important problems, we're stuck on part 1!

Part II: Random 3-XOR Equations and Pseudo-expectation Values

Equations for Random 3-XOR

- Want each $x_i \in \{-1,1\}$
- 3-XOR constraint: $x_i x_j x_k = 1$ or $x_i x_j x_k = -1$
- We will take m 3-XOR constraints at random
- Problem equations:
  1. $\forall i, x_i^2 = 1$
  2. $\forall a \in [1,m], x_{i_a} x_{j_a} x_{k_a} = c_a$ where $\forall a \in [1,m]$, $i_a, j_a, k_a \in [1,n]$ and $c_a \in \{-1,1\}$

SOS Lower Bound for Random 3-XOR

- Problem equations:
  1. $\forall i, x_i^2 = 1$
  2. $\forall a \in [1, m], x_{i_a} x_{j_a} x_{k_a} = c_a$ where $\forall a \in [1, m]$, $i_a, j_a, k_a \in [1, n]$ and $c_a \in \{-1,1\}$
- Theorem [Gri02], rediscovered by [Sch08]: If $m \leq \frac{n^{\frac{3}{2}-\epsilon}}{\sqrt{d}}$ then w.h.p., degree d SOS does not refute these equations.

Choosing Pseudo-expectation Values

- How do we choose the pseudo-expectation values?
- Many choices are fixed.
- Example: If $x_1x_2x_3 = 1$ and $x_1x_4x_5 = -1$ then $x_1^2x_2x_3x_4x_5 = x_2x_3x_4x_5 = -1$
- However, we only want to make these deductions at low degrees…

Choosing Pseudo-expectation Values

- Def: Define $x_I = \prod_{i \in I} x_i$
- Proposition: $\forall I, J, x_I x_J = x_{I \Delta J}$ where $I \Delta J = (I \cup J) \setminus (I \cap J)$ is the disjoint union of I and J.
- To decide which x_I have fixed values:
  1. Keep track of a collection of equations $\{x_I = c_I\}$ starting with the problem constraints.
  2. If we have equations $x_I = c_I$ and $x_J = c_J$ where I, J, and $I \Delta J$ all have size at most d, then we add the equation $x_{I \Delta J} = c_I c_J$ (if we don't have it already)

Choosing Pseudo-expectation Values

- Set $\tilde{E}[x_I] = c_I$ if our collection has $x_I = c_I$
- What if we don't have an equation for x_I?
- If we have no equation for x_I, set $\tilde{E}[x_I] = 0$
- Set $\tilde{E}\left[x_i^2 f\right] = \tilde{E}[f]$ for all f of degree $\leq d - 2$
- These pseudo-expectation values are well-defined as long as we never have both the equations $x_I = 1$ and $x_I = -1$.

Part III: Proving PSDness

To-Do List

- Here we assume that $\tilde{E}$ is well defined. We will analyze when this holds w.h.p. in the next section.
- Need to check linear equations. This follows from the definitions:
  - Whenever we have a constraint $x_I = c_I$, for all J of size $\leq d-3$, either $\tilde{E}\left[x_I x_J\right] = c_I c_J = c_I \tilde{E}[x_J]$ or $\tilde{E}\left[x_I x_J\right] = c_I \tilde{E}\left[x_J\right] = 0$
  - $\forall i, f: \deg(f) \leq d-2, \tilde{E}\left[x_i^2 f\right] = \tilde{E}[f]$
- Need to check moment matrix is PSD.

Restriction to Multilinear Indices

- Observation: Whenever we have constraints $x_i^2 = x_i$ or $x_i^2 = 1$, it is sufficient to consider the entries of M indexed by multilinear monomials.
- Reason: Given any g of degree $\leq \frac{d}{2}$, $\exists$ multilinear g' such that $\tilde{E}\left[g'^2\right] = \tilde{E}[g^2]$.
- Proof idea: Any non-multilinear term $x_i^2 f$ in g can be replaced by f.
- Corollary: $\tilde{E}\left[g^2\right] \geq 0$ for all g of degree $\leq d/2$ $\Leftrightarrow$ $\tilde{E}[g^2]$ for all multilinear g of degree $\leq d/2$.

Key Idea: Equivalence Classes

- Definition: For sets I, J of size $\leq \frac{d}{2}$, we say $x_I \sim x_J$ if $x_I x_J = x_{I \Delta J}$ is determined
- Proposition: If $x_I \sim x_J$ and $x_J \sim x_K$ then $x_I \sim x_K$.
- Proof: If $x_I \sim x_J$ and $x_J \sim x_K$ then $x_{I\Delta J}$ and $x_{J\Delta K}$ are determined. Now $x_{I\Delta J} x_{J\Delta K} = x_I x_J^2 x_K = x_{I\Delta K}$ is determined. Thus, $x_I \sim x_K$
- Remark: We carefully chose which deductions to make so that this would work.

PSD Decomposition

- Proposition: $\tilde{E}\left[x_I x_J\right] \neq 0$ if and only $I \sim J$.
- Choose a representative I_E from every equivalence class E.
- Take $v_E(x_I) = \tilde{E}\left[x_I x_{I_E}\right]$
- $v_E(x_I) = c_{I\Delta I_E}$ if $x_I \in E$. Otherwise, $v_E(x_I) = 0$
- $v_E(x_I)v_E\left(x_J\right) = c_{I\Delta I_E}c_{\mathrm{J}\Delta I_E} = c_{I\Delta J}$ if $I, J \in E$. Otherwise, $v_E(x_I)v_E\left(x_J\right) = 0$

PSD Decomposition

- $v_E(x_I)v_E(x_J) = c_{I\Delta I_E}c_{\mathrm{J}\Delta I_E} = c_{I\Delta J}$ if $I, J \in E$. Otherwise, $v_E(x_I)v_E(x_J) = 0$
- Corollary: $\forall I, J, \sum_E v_E(x_I)v_E(x_J) = \tilde{E}[x_I x_J]$
- Corollary: $M = \sum_E v_E v_E^T \succcurlyeq 0$

Part IV: Analyzing Parameter Regimes

Parameter Regimes

- How large does m have to be before the random 3-XOR constraints are unsatisifable w.h.p.?
- For which m will the pseudo-expectation values be well-defined w.h.p., giving us the SOS lower bound?

Unsatisfiability of 3-XOR Constraints

- For any given possible solution $(x_1, \ldots, x_n)$, the probability it is valid if there are m random 3-XOR constraints is 2^{-m}.
- Using a union bound, $P[\exists solution] \leq 2^{n-m}$
- Equations are unsatisfiable w.h.p. if $m \gg n$
- In fact, not hard to show that

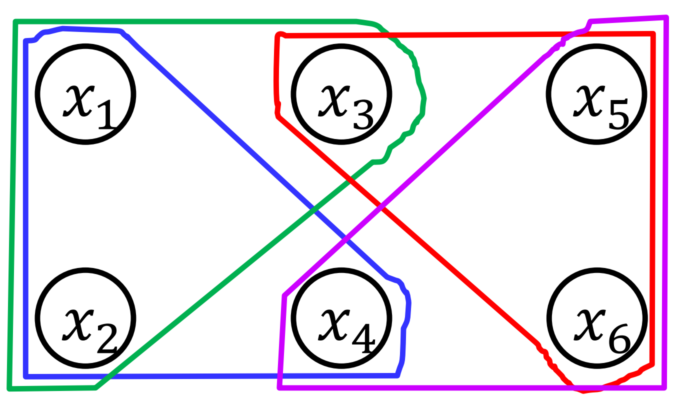
$\forall \epsilon > 0, \exists C, n_0 > 0$: if $m \geq Cn, n \geq n_0$ then w.h.p. there is no solution satisfying $\frac{1}{2} + \epsilon$ of the constraints

Local Consistency

- If $\tilde{E}$ is not well-defined then we must be able to derive the contradiction $-1 = 1$ without going to degree higher than $2d$.
- Multiplying all of the constraints involved in such a contradiction, every variable appears an even number of times.

Local Contradiction Picture

- Draw a triangle $(x_{i_a}, x_{j_a}, x_{k_a})$ for each constraint $x_{i_a}x_{j_a}x_{k_a} = c_a$ involved in the contradiction.
- Every vertex is covered an even number of times
- Example: If we have the constraints $x_1x_2x_3 = 1$, $x_4x_5x_6 = 1$, $x_1x_2x_4 = 1$, $x_3x_5x_6 = 1$, we get the following picture:



Probabilistic Analysis

- What is the probability that there is some contradiction involving D vertices where each variable appears twice?
- There are $\binom{n}{D} \leq \left(\frac{en}{D}\right)^D$ ways to choose the D vertices.
- Now choose the triangles one by one, starting at any vertex which has not yet been covered twice and choosing the other two vertices. This gives $\leq D^2$ choices for each of the $\frac{2D}{3}$ triangles.

Probabilistic Analysis Continued

- We have $\leq \left(D^2\right)^{\frac{2D}{3}} \left(\frac{en}{D}\right)^D$ choices for the structure of the constraints. For a given structure, the probability it appears is $\left(\frac{m}{n^3}\right)^{\frac{2D}{3}}$. Thus, the probability of such a contradiction is at most $\left(\frac{mD^2}{n^3}\right)^{\frac{2D}{3}} \left(\frac{en}{D}\right)^D = \frac{m^{\frac{2D}{3}} D^{\frac{D}{3}} e^D}{n^D} = e\sqrt[3]{m^2 D / n^3}$
- This is much less than 1 if $m \ll \frac{n^{\frac{3}{2}}}{\sqrt{D}}$

Analysis Subtleties

- Note: Can have $D > d$ variables involved in a contradiction without going to degree more than d (by ignoring vertices which have already been covered twice)
- However, must have a constraint graph on $\geq \frac{D}{3}$ vertices where at most d vertices appear an odd number of times.
- Can take $D = O(d)$ and show w.h.p. this does not happen.

Analysis Subtleties

- Note: Also have to consider the cases where variables appear more than twice in the clauses.
- These cases can be analyzed in a similar way.

Part V: Gaussian Elimination and SOS

Disproving Perfect Completeness

- As stated, the 3-XOR problem is actually easy, it's a system of linear of linear equations mod 2
- Map $\{-1,1\}$ to $\{1,0\}$ and multiplication to addition mod 2. Example: $x_i x_j x_k = -1$ becomes $x_i + x_j + x_k = 1 \, mod \, 2$
- Can use Gaussian elimination!

Noise Gives NP-hardness

- While disproving perfect completeness is easy, it is NP-hard to distinguish between the case when $(1-\epsilon)$ of the constraints can be satisfied and the case when at most $\left(\frac{1}{2}+\epsilon\right)$ of the constraints can be satisfied.
- Problem reformulation: Given constraints $\{x_{i_a}x_{j_a}x_{k_a} = c_a : a \in [1,m]\}$, problem becomes: Maximize $\sum_{a=1}^{m} c_a x_{i_a}x_{j_a}x_{k_a}$ subject to
  1. $\forall i, x_i^2 = 1$

SOS Robustness

- Why doesn't SOS capture Gaussian elimination?
- One explanation: SOS is inherently robust to noise, so it cannot capture techniques which are not robust, like Gaussian elimination.
- This explanation has merit, though the fact remains that Gaussian elimination is an algorithm not captured by SOS.

Part VI: Further Work

k-wise Independent Distributions

- Definition: A distribution of solutions for a clause is balanced k-wise independent if for all indices $i_1, \ldots, i_k$ and all $b_1, \ldots, b_k \in [0,1]$,
$$P\left[\forall j \in [1,n], x_{i_j} = b_j\right] = 2^{-k}$$
- Example: For a 3-XOR clause $x_i + x_j + x_k = b$ mod 2, the uniform distribution of solutions is balanced 2-wise independent.

Further Work

- These ideas have been vastly generalized to show tight SOS upper and lower bounds on CSPs with balanced k-wise independent distributions [BCK15], [KMDW17].
- Note: Balanced pairwise independence implies UGC-hardness [AM08], NP-hardness is only known if there is a balanced pairwise independent subgroup [Cha13].

References

- [AM08] P. Austrin, E. Mossel. Approximation Resistant Predicates From Pairwise Independence. https://arxiv.org/abs/0802.2300 . 2008
- [BCK15] B. Barak, S. O. Chan, and P. Kothari. Sum of squares lower bounds from pairwise independence. STOC 2015.
- [Cha13] S. O. Chan. Hardness of Maximum Constraint Satisfaction. Ph.D. thesis at Berkeley.
- [KMDW17] P. Kothari, R. Mori, R. O'Donnell, D. Witmer. Sum of squares lower bounds for refuting any CSP. STOC 2017.