GCT Online Lecture series - Lower Bounds I.

Agendas:
- Some recap of models,
- Typical lb proof templates,
- Some basic lower bounds, techniques

Models of computation.

Formulas \subseteq ABPs \subseteq Circuits.

\[
\begin{align*}
\Sigma \Pi & : \sum_{i:1} \prod_{j=1} a_{ij} & & \text{ sparse polynomials.} \\
\Sigma \Pi \Sigma & : \sum_{i} \Pi \Sigma x_{i} \prod_{j=1} & & \text{ depth 3 circuits. (surprisingly powerful)} \\
\Sigma \Lambda \Sigma & : \sum_{i} \ell_{i} & & \text{ depth 3 powering circuits (or "diagonal" circuits)}
\end{align*}
\]

... and so on.

Don't panic... we'll work only with d-polynomial

(2021-11-16)
Goal: Find an explicit family \( \{ f \in \mathbb{F}[x_1, \ldots, x_n] \} \) of polynomials such that any circuit \( c \) from the class \( C \) computing \( f \) must have "large" size.

Ideally: \( C = \text{alg. circuits.} \quad \text{"large" = } \exp(n) \). \[ \text{Explicit: } \text{"you'll know it when you see it."} \]

How do lower bounds "typically" proceed?

- Identify a weakness of the class \( C \).
- Quantify this weakness using a "complexity measure".
- Upper bound this measure for any small \( C \subset C \).
- Heuristically convince yourself that the measure is "large" for a "random polynomial".
- Find a hay in the haystack.

In a lot of cases, LBS follow the following template: (sub-additive measures)

- Show that any \( C \subset C \) can be written as a sum of "building blocks" \( B_i \).
  \( t = f_1 \circ g \).
  \( C = B_1 + \cdots + B_t \) where each \( B_i \in \mathcal{B} \).
- Find a \( P: \mathbb{F}[\tilde{x}] \to \mathbb{R} \) that is subadditive
\[ \Pi(f+g) \leq \Pi(f) + \Pi(g) \]
and show that \( \Pi(B) \leq \) "small" for any \( B \in \mathcal{B} \).

- Heuristically convince yourself that \( \Pi(f) \) is "large" if \( f \) is "random".

- Find hay in the haystack.

\[ \Pi(f) = \dim \{ L(f) : L \leq L \} \]

This lecture is: - Some examples of \( \mathcal{B} \)'s that follow this template.
- Some examples that don't. (in some sense)

Starting with the basics:

- \( \Sigma \Pi \) circuits \((I \text{ know} \ldots)\)

\( \mathcal{B} = \) monomials. \( \Pi(f) = \) sparsity.
\( \Pi(B) \leq 1 \)
\( \Pi(f+g) \leq \Pi(f) + \Pi(g) \) (duh).

A random poly probably has lots of monomials in it.
\( f = (x_1+1) \ldots (x_n+1) \) is one of them hay.
\( \Pi(f) = 2^n \implies \text{we have our lower bound.} \)

- Depth 3 powering circuits: \( l_1 + \ldots + l_s \)

\( \mathcal{B} = \{ l^k : l \text{- linear, } \mathbb{E} \in \mathbb{N} \} \).
What is an exploitable weakness for such building blocks?

Obs: \( \frac{\partial^2}{\partial x_i \partial x_j} l^d = \alpha \cdot l^{d-2} \)

\[ \mathcal{E}^k(f) = \left\{ \mathcal{E}_m(f) : m = \max \deg \mathcal{E} \text{ and } k \right\} \]

**Lemma:** \( \dim \mathcal{E}^k(l^d) \leq 1 \).

Set \( \Pi_k(f) = \dim \mathcal{E}^k(f) \) (where \( k \) will be chosen later).

Let \( f \in \Sigma \times \Sigma \), then for any \( k \), \( \Pi_k(f) \leq s \).

If \( f \) is a random \( n \)-var, \( d \)-deg poly, what do we expect this to be?

There are \( \approx \binom{n+d-k}{d-k} \) vectors in \( \mathcal{E}^k(l^d) \)

each with \( \binom{n+d-k}{d-k} \) coordinates.

\[
\mathbf{M}(f) = \begin{pmatrix}
\end{pmatrix}
\]

\[
\text{derivatives of order } k.
\]

\[
\text{Wishful heuristic: } \Pi_k(f) \approx \min \left( \binom{n+k}{k} , \binom{n+d-k}{d-k} \right)
\]

for a random \( f \).
\( f = x_1 \ldots x_n, \quad P_k(f) = \binom{n}{k}. \)

If \( f \in \Sigma \Lambda \Sigma \), then \( P_k(f) \leq s \) and \( P_k(f) \geq \binom{n}{k} \)

\( \Rightarrow s \geq \binom{n}{k}. \quad k = \frac{n}{2} \Rightarrow s = 2^{\Omega(n)} \)

- LBS for depth 3 circuits.

\[ f = \sum_{i=1}^{s} \prod_{j=1}^{d_i} l_{ij} \quad \longrightarrow \text{surprisingly powerful}. \]

Hom. depth 3 circuits.

\[ f = \sum_{i=1}^{S} \prod_{j=1}^{D_i} l_{ij} \quad \text{each } l_{ij} \text{ is a linear form} \]
\noindent \text{(no constant term)}.

\[ \mathcal{B} = \left\{ \prod_{i=1}^{d} l_{ij} : l_{ij} \text{'s linear forms}\right\}. \enspace \mathcal{B}_{k+1}(l_1 \ldots l_d) \]

What is an exploitable weakness? \( x_2 \ldots x_d + k \cdot l_1 l_3 \ldots l_d \ldots \)

**Claim:** \( P_k(l_1 \ldots l_d) = \dim F^k(l_1 \ldots l_d) \leq \binom{d}{k} \)

\[ \leq \text{span} \{ \prod_{i \in S} l_i : |S| = d-k \} \]

If \( f = \sum_{i=1}^{S} l_{1i} \ldots l_{di} \) and \( f \) is "random", then

we can hope \( s \geq \min \left\{ \binom{n+k}{k}, \binom{n+d-k}{d-k} \right\} \).

Say \( k = d/2 \)

Numerator \( \approx \binom{n}{d/2} \)

Denom: \( \approx 2^d \)
Fact: If $f \in \ESYM_{n,d}$, then $\dim \min^k(f) \approx \text{wishful bound}$.

Any hom. $\Sigma \Pi \Sigma$ chf computing $\ESYM_{n,d}$ requires size $(\frac{2n}{d})^{d/2}/2^d$.

On the other hand, there are non-hom. $\Sigma \Pi \Sigma$ chfs for $\ESYM_{n,d}$ of size just $O(n^2)$ [Ben-Or]

- Multilinear $\Sigma \Pi \Sigma$ circuits.

$$f = \sum_{i=1}^{g} l_1 \ldots l_d$$

where $l_1, \ldots, l_d$ are variable disjoint.

$\mathbb{B} = \{ l_1, \ldots, l_d : \text{variable disjoint linear poly} \}$.

Alice \hspace{1cm} Bob

$X = Y \cup Z$

On: Is $\text{rank}(M(f))$ a good measure?

$\text{rel.-rank}(f) = \frac{\text{rank}(M_{Y,2}(f))}{\sqrt{\#\text{rows} \cdot \#\text{cols}}} = \frac{\text{rank}(M_{Y,2}(f))}{\sqrt{2(Y_1+Z_1)^2}}.$
Fact: \( f, g \) var disjoint \( \Rightarrow \) rel\( \text{Rank}(fg) \leq \text{rel\( \text{Rank}(f) \cdot \text{rel\( \text{Rank}(g) \))} \)

\[ M(fg) = M(f) \otimes M(g) \]

\( n = 2d \)

What is \( \mathbb{E} \left[ \text{rel\( \text{Rank}(l_1, \ldots, l_k) \)} \right] \)

\[ \leq \frac{d^{1/2}}{\binom{n}{k}} \mathbb{E} \left[ \text{rel\( \text{Rank}(l_i) \)} \right] \]

Suppose \( \text{Var}(l_i) > 3 \). \( \Rightarrow \) rel\( \text{Rank}(l) \leq \frac{2}{2^{3/2}} = \frac{1}{\sqrt{2}} \)

\( \text{Var}(l_i) = 2. \) \( \mathbb{E} \left[ \text{rel\( \text{Rank} \right) \leq \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot 1 = \frac{3}{4} \)

\[ \mathbb{E} \left[ \text{rel\( \text{Rank}(l_1, \ldots, l_k) \)} \right] \leq \left( \frac{3}{4} \right)^{d/2} \]

For a random poly, we expect \( M(f) \) to be full rank.

\( \Rightarrow \) We have a \( \Omega(d) \) lower bound.

\([\text{Raz}] \) \( n^{\Omega(\log n)} \) lb for ML formulas computing Det\Perm.

\([\text{CELS}] \) Depth hierarchy thinner for ML formulas.

Some lbs that don't follow this template:

- Baur-Strassen: \( \Omega(n \log d) \) lb for general circuits computing \( x_1 + \cdots + x^d \).
Can we fix an output $\overline{a} \in \mathbb{F}^n$ and ask how many input result in that output?

There is one quadratic eqn for each gate.

# solutions to $n$-quadratic eqns.

[Bezout] If this is finite, then $\leq 2^n$.

If $f = x_1^{d_1} + \ldots + x_n^{d_n}$

$\overline{a} = (d_1, \ldots, d_n)$ How many inputs lead to this?

$\Rightarrow$ each $x_i$ is a $(d-1)$th root of unity.

$\Rightarrow$ We have $(d-1)^n$ inputs.

$(d-1)^n \leq 2^n \Rightarrow s \geq \Omega(n \log d)$
Determinental complexity LB. [Mignon-Ressayre]

\[ f = \text{Det} \begin{bmatrix} l_1 & \cdots & l_m \\ l_{m1} & \cdots & l_{mn} \end{bmatrix} \]

\[ l_i \in \mathbb{F}[x_1, \ldots, x_n] \]

linear.

Want an LB for \( m \).

\[ f(x) = \text{det}(A(x)) \]

let \( x_0 \) : \( f(x_0) = 0 \).

\[ f(x + x_0) = \text{det}(A(x + x_0)) \]

\[ = \text{det}(L(x) + B) \]

\[ \exists v: \quad Bv = 0 \]

\[ \implies \text{det}(B) = 0 \]

Suppose \( x' \) s.t. \( L(x') v = 0 \)

then \( L(x') + B \) is also singular!

\[ \implies f(x' + x_0) = 0 \]

\[ V = \{ x : L(x') v = 0 \} \implies \text{vector space of dim } \geq N - m. \]

Claim: If \( f(x_0) = 0 \) and \( V \) is a vector space of dim \( n-r \)

\[ f(x_0 + x) = 0 \forall x \in V, \text{ then } \]

\[ \text{rank} (\text{Hess}(f)(x_0)) \leq 2r. \]

Claim: There is an \( n \times n \) matrix \( x_0 \) s.t. \( \text{Perm}(x_0) = 0 \)

and \( \text{rank} (\text{Hess}(\text{Perm})(x_0)) = n^2. \)
Cor: \( \text{Det}_m (A(x)) = \text{Perm}_n \Rightarrow n \geq \frac{n^2}{2} \).