Update Meeting on Algorithms and Complexity 10.03.08

Identity Testing of Depth Four Circuits

Lecturer: Manindra Agrawal *Scribe: Partha*

1 Arithmetic Circuits

An arithmetic circuit C over a field $\mathbb{F}$ is a circuit with addition and multiplication gates of unbounded fanin. The inputs to the circuits are either indeterminates or elements from the field. Any circuit C over $\mathbb{F}$ and having indeterminates $x_1, x_2, \cdots, x_n$ naturally computes a polynomial in $\mathbb{F}[x_1, x_2, \cdots, x_n]$. The size of an arithmetic circuit is the number of gates in it. The degree of the circuit is the degree of the polynomial computed in the output gate. In this talk we consider the degree of the circuit is bounded by the size of the circuit.

2 Polynomial Identity Testing

The polynomial Identity Testing is the following problem: Given an arithmetic circuit C computing a polynomial f in $\mathbb{F}[x_1, x_2, \cdots, x_n]$, test whether $f \equiv 0$. The polynomial identity testing problem is a classical problem in complexity theory. It is well known that this problem can be solved in randomized polynomial time using Schwartz-Zippel Lemma [MR01]. Main open question is that whether the Polynomial Identity Testing Problem can be solved in deterministic polynomial time. In 2003, Impagliazzo and Kabanets proved that any such algorithm will imply either NEXP $\not\subset P/$poly or Permanent has no polynomial size arithmetic circuit. The deterministic polynomial time algorithms for this problem are known for some restricted circuit class: for depth 2 circuit the problem is trivial. Kayal and Saxena [KS07], showed that the identity testing for depth 3 $\Sigma\Pi\Sigma$ circuit can be done in deterministic polynomial time if the fanin of the top Σ gate s constant. In 2007, Saxena [Sa07], gave a deterministic polynomial time algorithm for depth 3 diagonal circuits.

3 Black-Box derandomization

Let $\mathcal{A}_{\mathbb{F}}^n$ be the set of arithmetic circuit over $\mathbb{F}$ of size at most n. Let $f : \mathcal{N} \rightarrow (\mathbb{F}[y])$ be a function such that for all n,

$$f(n) = (p_{n,1}(y), \cdots, p_{n,n}(y))$$

where $p_{n,j}$'s are polynomials over $\mathbb{F}[y]$. Then the function f is a (n, d) pseudo-random generator against $\mathcal{A}_{\mathbb{F}}$ if the following are true:

- each $p_{n,i}$ is of degree at most d
- for any circuit $C \in \mathcal{A}_{\mathbb{F}}^n$, $C(x_1, \cdots, x_n) = 0$ if and only if $C(p_{n,1}(y), \cdots, p_{n,n}(y)) = 0$.

The function f is an optimal pseudo-random generator for $\mathcal{A}_{\mathbb{F}}^n$ if $d = n^{O(1)}$. In fact, Schwartz-Zippel Lemma implicitly tells that a random f of degree $\log n$ is a pseudo-random generator with high probability. A pseudo-random generator that can be quickly constructed is very useful.

Definition 3.1 *A function f is an efficiently computable (n, d) pseudo-random generator against $\mathcal{A}_{\mathbb{F}}^n$ if it is pseudo-random generator and can be constructed in time* $\text{poly}(n, d)$*.*

If there exists efficiently computable optimal pseudo-random generator then the following are true:

- The identity testing problem can be solved in deterministic polynomial time.
- There is a multilinear polynomial in EXP that can not be computed by a subexponential size circuit [A05].

Surprisingly, all the known deterministic polynomial time identity testing algorithm are known for only depth 3 circuits. Moreover all superpolynomial lower bounds for arithmetic circuits are also known for monotone circuits, multilinear formulas and depth 3 circuits. For example see [SW99, GR00, Raz04]. It seems that the progress in identity testing as well as in proving lower bounds seems to stop in depth 3 circuits only.

4 Main Theorem

In this section we prove that, in some sense, the identity testing of general arithmetic circuits is as hard as identity testing of depth 4 circuits. This results justifies the lack progress in either identity testing or proving lower bounds beyond depth 3 arithmetic circuits. More precisely, we prove the following theorem.

Theorem 4.1 *If a polynomial $p(x_1, x_2, \cdots, x_n)$ can be computed by an arithmetic circuit C of degree n and size $m = 2^{o(n)}$, then it can be computed by a depth 4 arithmetic circuit of degree n and size $2^{o(n)}$.*

Proof. We outline the sketch of the proof. The proof heavily depends on the depth reduction technique developed in [AJMV94]. The first step of the proof is to apply the depth reduction technique developed in [AJMV94] to construct a circuit D such that D computes the same polynomial p. Moreover, the construction of [AJMV94] guarantees that the degree of the circuit D is at most the degree of the circuit C and size is $m^{O(1)}$. We briefly describe the construction of D (details can be found in [AJMV94]).

Construction of D: Make the circuit C layered with alternating layers of $+$ and $*$ gates. Make the fanin of every multiplication gate two. Arrange the children of all the multiplication gates such that the degree of the right child is greater than or equal to the degree of the left child. Now, for every pair of gates g and h in C, introduce the gates $[g]$ and $[g, h]$ as follows: gate $[g]$ computes the same polynomial as gate g. The gate $[g, h]$ computes the zero polynomial if h does not occur in the rightmost path of a proof tree rooted at g. Gate $[g, h]$ computes the same polynomial which is the sum over all the proof trees rooted at g in which h occurs in the rightmost path of the product of leaves of the proof tree when h is replaced by 1. It can be seen that $[g] = \sum_{i=1}^{n} [g, x_i] x_i$. For a $+$ gate g with children $g_1, g_2, \cdots, g_t$, and for any other gate h, $[g, h] = \sum_{i=1}^{t} [g_i, h]$. If g is a $*$ gate with left child g_L and right child g_R and there are only $+$ gates on a path from g to h, then $[g, h] = [g_L]$. For otherwise, let g is a $*$ gate and there are multiplication gates $g_1, g_2, \cdots, g_t$ that occur in the path from g to h. Let each of the g_i has $g_{i,L}$ and $g_{i,R}$ as of their left and right child. Moreover if $\deg(g_i) \geq \frac{1}{2}(\deg(g) + \deg(h)) > deg(g_i, R)$, then $[g, h] = \sum_{i=1}^{t} [g, g_i][g_{i,L}][g_{i,R}, h]$. The gates of the circuit D are simply $[g]$ and $[g, h]$. It is easy to see that $\deg([g, h]) = \deg(g) - \deg(h)$. In the case when $\deg(g_i) \geq \frac{1}{2}(\deg(g) + \deg(h)) > deg(g_i, R)$, $\deg([g, g_i]) \leq \frac{1}{2}(\deg(g) - \deg(h))$. Also in that case, $\deg([g_{i,R}, h]) = \deg(g_{i,R}) - \deg(h) < \deg(g_i) - \deg(h) < \frac{1}{2}(\deg(g) - \deg(h))$. For the gate $[g_{i,L}]$, $\deg([g_{i,L}]) \leq \max\{\deg(g) - \deg(h), \frac{1}{2}\deg(g)\}$.

Now it is easy to see that the depth of D is $O(\log\ n)$ and the size is only $m^{O(1)}$. Now the idea is to replace the circuit D by a depth 4 circuit. The way we do this, cut the circuit D in two halves and replace each of them by a depth 2 circuit. We now explain it formally.

We cut the circuit in such a way that all the nodes in the bottom half (D_B) has degree $< e$ and all the nodes in the top half (D_T) has degree $\geq e$, where we choose e appropriately in the analysis. Let in the bottom half (D_B), $D_1, \cdots, D_\ell$ are all the circuits such that their output nodes are in the cutting plane. For each $i \in [\ell]$, write D_i as a sum of monomials. Clearly in each of the D_i the number of monomials are bounded by $\begin{pmatrix} n+e \\ e \end{pmatrix}$.

The input to the top half D_T are the outputs from the ℓ nodes. So circuit D_T computes a polynomial in the ℓ variables. Again, we will express D_T as a sum of monomials. Furthermore, an easy induction over the depth of D_T proves that the degree of D_T is at most $10\frac{d}{e}$.

Let the new depth 4 circuit that is obtained by writing D_T and D_B as sum of monomials (and thus replacing by depth 2 circuits) is E. So the size of E is bounded by $\begin{pmatrix} n+e \\ e \end{pmatrix} + \begin{pmatrix} \ell + 10\frac{d}{e} \\ 10\frac{d}{e} \end{pmatrix}$. Since $\ell = m^{O(1)}$, we can bound the size of E by $m^{c(\frac{cn}{e}e)} + m^{\frac{cn}{e}}$ for an appropriate constant c.

Now to bound the size of E further we use the fact that $m = 2^{o(n)}$. Let $m = 2^{\frac{n}{g(n)}}$ and choose $e = n/\sqrt{g(n)}$. Now it is easy to see that the size of E is bounded by $2^{o(n)}$. This completes the proof of the theorem. ■

5 Identity Testing for the Depth four circuits

Suppose there exists an optimal efficiently computable pseudo random generator against depth four circuits over $\mathbb{F}$. Let f be such a generator with $f(n) = (p_{n,1}(y), \cdots, p_{n,n}(y))$. Moreover assume that the degree of $p_{n,i}$'s are bounded by $d = n^{O(1)}$. Fix $\ell = \log\ n$. Define the polynomial $r_{2\ell}$ as follows:

$$r(x_1, x_2, \cdots, x_{2\ell}) = \sum_{S \subseteq [1, 2\ell]} c_S \prod_{i \in S} x_i$$

such that the coefficients c_S satisfy the following relation: $\sum_{S \subseteq [1,2\ell]} c_S \prod_{i \in S} p_{n,i} = 0$. Simply counting the number of c_S's and the number of homogeneous constraints, one can easily argue that a non zero polynomial $r_{2\ell}$ always exists whose coefficients satisfy the above relation. Furthermore, $r_{2\ell}$ can easily be computed by solving a system of $2^{O(\ell)}$ linear equations and so the computation can be performed in EXP. Can the polynomial $r_{2\ell}$ be computed by a depth four circuit of size of size $2^{\epsilon\ell}$ for some $\epsilon > 0$. We will use the pseudo random generator $f(n)$ for depth four circuits to argue that $r_{2\ell}$ can not be computed by a depth four circuit of size $2^{\epsilon\ell}$. or otherwise let C be such a circuit. By the definition of $r_{2\ell}$, we know that $C(p_{n,1}, \cdots, p_{n,n}) = 0$. But $r_{2\ell} = C(x_1, \cdots, x_n)$ is a non zero polynomial. That contradicts the assumption that f is a pseudo random generator against the depth four circuits. Now it follows directly from the Theorem 4.1 that the polynomial $r_{2\ell}$ requires a circuit of size $2^{\Omega(\ell)}$. Recall that $\ell = \log\ n$. Now we will use this hard polynomial $r_{2\ell}$ to construct a $(n, n^{\log\ n})$ pseudo random generator against the entire class of arithmetic circuits over $\mathbb{F}$.

5.1 Construction of the pseudo random generator

Fix n. Let b be an appropriately chosen large constant. Define the standard Nisan-Wigderson design [NW94] $S_1, S_2, \cdots, S_n$ such that:

- $S_i \subseteq [1, b\log\ n]$.
- $S_i = a\log\ n$ for constant $a < b$.
- $|S_i \cup S_j| \leq \log\ n$.

Define the polynomial $q_i(z_1, z_2, \cdots, z_{b\log\ n}) = r_{a\log\ n}(z|S_i)$. Let $p_{n,i}(y) = q_i(y, y^{n+1}, y^{(n+1)^2}, \cdots, y^{(n+1)^{b\log\ n-1}})$ and define $f(n) = (p_{n,1}(y), \cdots, p_{n,n}(y))$. We claim that $f(n)$ is a required pseudo random generator.

5.1.1 Correctness of the generator

We use the standard *hybrid argument* to prove that $f(n)$ is indeed a pseudo random generator for the entire class of arithmetic circuits C over $\mathbb{F}$. Let C be any arithmetic circuit of size and degree at most n such that:

$$C(p_{n,1}(y), \cdots, p_{n,n}(y)) = 0.$$

So,

$$C(q_1(y, y^{n+1}, y^{(n+1)^2}, \cdots, y^{(n+1)^{b\log\ n-1}}), \cdots, q_n(y, y^{n+1}, y^{(n+1)^2}, \cdots, y^{(n+1)^{b\log\ n-1}})) = 0.$$

Now the univariate transformation that is developed in [AB03], indeed implies that

$$C(q_1(z_1, \cdots, z_{b\log\ n}), \cdots, q_n(z_1, \cdots, z_{b\log\ n})) = 0.$$

Let j be the largest number such that $C(q_1, \cdots, q_{j-1}, x_j, \cdots, x_n) \neq 0$ but $C(q_1, \cdots, q_{j-1}, q_j, x_{j+1}, \cdots, x_n) = 0$. Fix values to the variables $x_{j+1} = a_{j+1}, \cdots, x_n = a_n$ such that $C(q_1, \cdots, q_{j-1}, x_j)|_{x_{j+1}=a_{j+1}, \cdots, x_n=a_n} \neq 0$ (By the Schwartz-Zippel Lemma we know such fixing for x_k's, $j+1 \leq j \leq n$ exist). Similarly fix values for all z_i's except those in S_j such that $C(q_1, \cdots, q_{j-1}, x_j) \neq 0$. Let the resulting circuit be $\hat{C}(z|_{S_j}, x_j)$. Since the size of the circuits for q_i's are bounded by n, the size of the circuit $\hat{C}$ is bounded by n^2. Moreover, since $\hat{C}(z|_{S_j}, q_j) = 0$, it implies that $(x_j - q_j)$ is a factor of the polynomial computed by the circuit $\hat{C}(z|_{S_j}, x_j)$. By the result of [Kal89], $x_j - q_j$ can be computed by a circuit of size n^c (for some $c > 0$), that uses the circuit $\hat{C}$. This circuit computes $r_{a \log\ n}$ which is a contradiction to the hardness of $r_{a \log\ n}$. This completes the correctness proof of the pseudo random generator.

6 Open Problem

Can we improve the construction of the pseudo random generator described in the section 5.1 to a optimal easily computable pseudo random generator ? The generator that we have described is a $(n, n^{O(\log\ n)})$ generator. Another very interesting problem is to design an optimal generator for the depth four circuits.

References

[A05] MANINDRA AGRAWAL. Proving Lower Bounds Via Pseudo-Random Generators. *FSTTCS 2005.*, 92-105

[AB03] M. AGRAWAL AND S. BISWAS. Primality and identity testing via Chinese remaindering. *J. ACM.*, 50(4):429-443, 2003.

[AJMV94] E. ALLENDER, J. JIAO, M. MAHAJAN AND V. VINAY Non-commutative arithmetic circuits: depth reduction and size lower bounds. *Theoretical Computer Science.*, 209: 47-86, 1998.

[GR00] D. GRIGORIEV AND A. RAZBOROV. Exponential Lower Bounds for Depth 3 Arithmetic Circuits in Algebras of Functions over Finite Fields. *Appl. Algebra Eng. Commun. Comput.*,10(6):465-487, 2000

[MR01] R. MOTWANI AND P. RAGHAVAN. Randomized Algorithm. Cambridge, 2001.

[NW94] N. NISAN AND A. WIGDERSON Hardness vs. randomness. *J. Comput. Syst. Sci.*, 49(2):149-167, 1994.

[Kal89] E. KALTOFEN.Factorization of polynomials given by straight-line programs. *In S. Micali, editor, Randomness in Computation.*, vol 5, pages 375-412, 1989.

[KS07] N. KAYAL AND N. SAXENA. Polynomial Identity Testing for Depth 3 Circuits. *Computational Complexity.*, 16(2):115-138, 2007.

[Raz04] R. RAZ. Multilinear formulas for permanent and determinant are of super-polynomial size. *In Proc. of the thirty-sixth annual ACM symposium on Theory of computing.*, pages 633-641, 2004.

[Sa07] N. SAXENA. Diagonal Circuit Identity Testing and Lower Bounds. Manusript 2007.

[SW99] A. SHPILKA AND A. WIGDERSON Depth-3 Arithmetic Formulae over Fields of Characteristic Zero. *IEEE Conference on Computational Complexity* 1999.