Circuit Complexity, Session 11

Lecturer: Sam Buss
Scribe: Marco Carmosino

December 4, 2013

Today we prove another lower bound on parity, this time against circuits that can count modulo some prime $p \neq 2$. This result was originally proved independently by Razborov [RE86] and Smolensky [Smo87]. In this lecture, we followed Smolensky [Smo87]. The argument uses low-degree polynomials over finite fields to closely approximate low-depth circuits.

Definition 1 (AC^0)**.** Constant-depth, unbounded-fanin circuits over the $\wedge, \vee, \neg$ basis of polynomial size.

Definition 2 ($\mathrm{Mod}_{i,m}$)**.**

$$\mathrm{Mod}_{i,m}(x) = \begin{cases} 1 & \text{if } x \equiv i \mod m \\ 0 & \text{otherwise} \end{cases}$$

Definition 3 ($\mathsf{AC}^0[m]$)**.** For $m \in \mathbb{Z}$, the class of AC^0 circuits with gates for $\mathrm{Mod}_{i,m}$

Note that in general, we can use $\mathrm{Mod}_{i,m}$ to compute $\mathrm{Mod}_{j,m}$ for $j \neq i$ by padding the input with extra 1 entries, so the $\mathrm{Mod}_{0,m}$ suffice to count to any quantity modulo m.

Now we consider how to represent Boolean-valued functions using functions that take values in a finite field.

Definition 4 (Representation)**.** Let $\mathbb{F}_p$ be a finite field of characteristic p and let $f_{\mathcal{B}}$ be a Boolean function. We say that $f : \{0,1\}^n \to \mathbb{F}_p$ *represents* $f_{\mathcal{B}}$ if:

$$f_{\mathcal{B}}(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & \text{if } f(x_1, x_2, \ldots, x_n) \neq 0 \\ 0 & \text{if } f(x_1, x_2, \ldots, x_n) = 0 \end{cases}$$

Now we will represent a few functions using polynomials with coefficients in $\mathbb{F}_p$. First, consider the *projection function*, $X_i(x_1, x_2, \ldots, x_n) = x_i$. This is easily represented as the polynomial $f(x_1, x_2, \ldots, x_n) := x_i$. The negation of a single variable is similar; $X_i^{\mathrm{neg}}(x_1, x_2, \ldots, x_n) = \neg x_i$ is represented by the polynomial $f_\neg(x_1, x_2, \ldots, x_n) := 1 - x_1$. The x_i all take values in $\{0,1\}$ because they are binary inputs.

Now, we make the assumption that polynomials $g_1, g_2, \ldots, g_k : \{0,1\}^n \to \mathbb{F}_p$ represent Boolean-valued functions, with the additional property that each g_i only takes values $\{0,1\} \in \mathbb{F}_p$. To represent $\text{AND}(g_1, g_2, \ldots, g_k)$, we take $f_\wedge := \prod_{i=1}^k g_i$. Notice that this representation of AND preserves the property of $\{0,1\} \in \mathbb{F}_p$ output, if the g_i have this property.

To represent OR, we use our representation of X_i^{neg} and De Morgan's law:

$$f_\vee := 1 - \prod_{i=1}^k (1 - g_i)$$

In both polynomials $f_\vee$ and $f_\wedge$ we assumed that their inputs were $\{0,1\} \in \mathbb{F}_p$. We could remove this assumption by taking $(g_i)^{p-1}$ each time g_i appears in the above polynomials, because for prime p and $j \in \mathbb{F}_p$:

$$j^{p-1} \mod p = \begin{cases} 0 & \text{if } j \equiv 0 \mod p \\ 1 & \text{otherwise} \end{cases}$$

Our complexity measure on these representations is the degree of the polynomials involved. Let $\mathcal{D} := \max_i(\deg(g_i))$. Then $\deg(f_\wedge) \leq k\mathcal{D}$ and $\deg(f_\vee) \leq k\mathcal{D}$. If we use the $(p-1)$ powering trick, then $\deg(f_{\vee,\wedge}) \leq k\mathcal{D}(p-1)$.

To represent $\text{Mod}_{i,p}$ is particularly straightforward:

$$f_\oplus = (g_1 + g_2 + \cdots + g_k - i)^{p-1}$$

And note that $\deg(f_\oplus) \leq (p-1)\max_i[\deg(g_i)]$, so it is particularly cheap to compute in this representation. Intuitively, that should be unsurprising – we are using "native" addition over $\mathbb{F}_p$ which does not increase the degree of the polynomials involved.

This means that we can replace any constant-depth circuit with a polynomial. Suppose we are given $\mathcal{C}$, a circuit over $\bigvee, \bigwedge, x_i, \bar{x_i}$ of size s and depth d. Then the fanin of $\mathcal{C}$ is trivially $\leq s$, and so d-many times we can repeat the simple constructions above, resulting in a degree $\leq s^d$ polynomial. This is unsatisfactory, as any polynomial is equivalent to a polynomial of degree $\leq n$.

What we would really like is a polynomial of degree $\log(s)^{O(1)}$ *approximating* $\mathcal{C}$. We will obtain this polynomial by altering the construction for OR.

Let $g_1, g_2, \ldots, g_\ell$ be Boolean functions, and suppose that $f_1, f_2, \ldots, f_\ell$ are $\mathbb{F}_p$ polynomials that approximately-represent $g_1, g_2, \ldots, g_\ell$. For $k \geq 1$ denote by OR_k the simple $f_\vee$ that computes the logical or of k inputs exactly, discussed above. Then we approximate the OR of $g_1, g_2, \ldots, g_\ell$ by the following polynomial:

$$f := \text{OR}_k[(c_{i1}f_1 + c_{i2}f_2 + \cdots + c_{i\ell}f_\ell)^{(p-1)}]$$

Where the c_{ij} are chosen uniformly at random from $\mathbb{F}_p$, with i ranging from 1 to k. Consider when the above polynomial works properly: if each $f_i = 0$, then we will have $f = 0$ and there will be no error.

If, on the other hand, f should be equal to 1, then $\exists f_j = 1$. We will compute this correctly, then, if at least one value i has:

$$\sum_{s=1}^{\ell} c_{is} f_s \neq 0$$

The above will be equal to zero if and only if we have that $c_{ij} = -f_j^{-1}(\sum_{s \neq j} c_{is} f_s)$. There is only one such c_{ij}, and so only one value that can erroneously zero out this argument to OR_k. Since we chose c_{ij} at random from $\mathbb{F}_p$, the probability of this error is $\frac{1}{p}$. Fixing a particular input $\bar{x}$, then, we will have $f(\bar{x})$ correct with probability $\geq 1 - \frac{1}{p^k}$, since we only need one of the k simplistic OR arguments to hit a nonzero value.

By an averaging argument, there exist choices of the c_{ij} such that the number of inputs $\bar{x}$ where f makes an error is $\leq 2^n(1 - \frac{1}{p^k})$. Now, all that remains is to set a value for k, and obtain degree bounds for our construction when applied to an entire circuit. Each linear combination contributes at most $\deg(f) \leq k(p-1)\max_i[\deg(f_i)]$.

Lemma 5. *Let $\mathcal{C}$ be a circuit over $\bigvee, \bigwedge, \oplus_p, x_i, \bar{x}_i$ of size s and depth d, and let $k = 2\log(s)$. Then $\mathcal{C}$ is approximated by an $\mathbb{F}_p$-polynomial f of degree $\leq (k(p-1))^d$ such that $f(\bar{x}) = \mathcal{C}(\bar{x})$ on all but a $\frac{1}{s}$-fraction of the possible inputs.*

Proof. By upper-bounding the number the number of inputs where the approximation makes an error:

$$\#\mathrm{err} \leq 2^n \left(\frac{1}{p^k}\right) s \tag{1}$$

$$\leq 2^n \frac{1}{s^2} s \tag{2}$$

$$\leq \frac{1}{s} 2^n \tag{3}$$

□

We have been using the "standard representation" of Boolean values over a field, where we represent True by 1 and False by 0. Now, let $p > 2$, and we will switch to the "Fourier basis", representing True by -1 and False by +1. We will use $x_i \in \{0, 1\}$ to denote variables in the standard representation, and $y_i \in \{-1, 1\}$ to denote variables in the Fourier representation. Observe that there is a simple linear transformation between these representations:

$$y_i = 1 - 2x_i \tag{4}$$

$$x_i = 2^{-1}(1 - y_i) \tag{5}$$

This transformation does not change the degree of a polynomial over $\mathbb{F}_p$. Note also that, over the Fourier basis, we have that $\text{Parity}_n(y_1, y_2, \ldots, y_n) = \prod_i y_i$

Corollary 6. *We restate the lemma above in the case of the Fourier basis. Suppose* Parity_n *has a circuit over* $\bigvee, \bigwedge, \oplus_p, x_i, \bar{x}_i$ *of constant depth* d *and size* $2^{n^{o\left(\frac{1}{2d}\right)}}$. *Then there exists an* $\mathbb{F}_p$ *polynomial* $f(\bar{y})$ *of degree* $o(\sqrt{n})$ *such that* $f(y_1, y_2, \ldots, y_n) = \prod_i y_i$ *for all but fraction* $\frac{1}{s}$ *of the inputs.*

Proof. By Lemma 5 we have that we can take the Parity_n circuit and approximate it with $f(\bar{x})$ over the standard representation, where $f(\bar{x})$ has degree $\leq (k(p-1))^d$. If we then change the representation of f by applying the linear transformation (4) above to each x_i in $f(\bar{x})$ to obtain $f(\bar{y})$ over the Fourier basis, we incur no extra cost in $\deg(f(\bar{y}))$ and obtain $f(y_1, y_2, \ldots, y_n) = \prod_i y_i$ for all but fraction $\frac{1}{s}$ of the inputs, because over the Fourier basis $\text{Parity}_n = \prod_i y_i$. Below we compute $\deg(f(\bar{y}))$:

$$\begin{aligned}
\deg(f(\bar{y})) &\leq (k(p-1))^d && (6)\\
&\leq (2\log(s)(p-1))^d && (7)\\
&\leq \left(2(p-1)n^{o\left(\frac{1}{2d}\right)}\right)^d && (8)\\
&\leq n^{o\left(\frac{1}{2}\right)} && (9)\\
&\leq o(\sqrt{n}) && (10)
\end{aligned}$$

□

Corollary 7. *Suppose the conditions of the previous Corollary hold. Then any* $h : \{-1, 1\}^n \to \mathbb{F}_p$ *has an approximating* $\hat{h}$, *a* $\mathbb{F}_p$*-polynomial of degree* $\leq \frac{n}{2} + o(\sqrt{n})$ *such that* $h(\bar{y}) = \hat{h}(\bar{y})$ *for all but fraction* $\frac{1}{s}$ *of the inputs.*

Proof. Assume h is a monomial. Let $h = y_{i1}y_{i2}\ldots y_{i\ell}$. Then with f as the degree $o(\sqrt{n})$ approximation of Parity from above, we take:

$$\hat{h} = y_{i1}y_{i2}\ldots y_{i\ell}f(y_1, \ldots, y_\ell)(y_1y_1\ldots y_\ell)$$

If $\deg(h) > \frac{n}{2}$, then $\deg(\hat{h}) \leq \frac{n}{2} + \deg(f)$ because more than half of the y's cancel out.

If $\deg(h) < \frac{n}{2}$, we just take $\hat{h} = h$. □

The proof above also shows that the only places where $\hat{h} \neq h$ is where f makes an error. Therefore, $\exists G \subset \{-1, 1\}^n$ with $|G| = 2^n(1 - o(1))$ such that $\forall y \in G h(\bar{y}) = \hat{h}(\bar{y})$.

Now consider the set of functions from G to $\mathbb{F}_p$. Formally let $V = \{f : G \to \mathbb{F}_p\}$. V is a vector space over $\mathbb{F}_p$, with $\dim(V) = |G| = 2^n(1 - o(1)))$.

But we have that a basis for V is a sum of monomials of degree $\frac{n}{2} + o(\sqrt{n})$, by the previous lemma, so the dimension of V should be $\leq$ the number of such monomials. We calculate this:

$$\#\text{monomials} = \sum_{i=0}^{\frac{n}{2}} \binom{n}{i} \tag{11}$$

$$= \frac{1}{2}2^n + \sum_{i=\frac{n}{2}}^{\frac{n}{2}+o(\sqrt{n})} \binom{n}{i} \tag{12}$$

$$\leq 2^{n-1} + o(\sqrt{n})\binom{n}{\frac{n}{2}} \tag{13}$$

$$\tag{14}$$

Which by approximating $n!$ as $\left(\frac{n}{e}\right)^n \sqrt{2\pi n}$ works out to $2^{n-1}(1 + o(1)) << |G| = 2^n(1 + o(1)$.

Corollary 8. *Let $p > 2$ be prime.* Parity_n *does not have* $\bigvee, \bigwedge, \oplus_p, x_i, \bar{x}_i$ *circuits* $\mathcal{C}$ *of size* $2^{n^{o\left(\frac{1}{2d}\right)}}$ *and depth* d.

References

[RE86] A. A. Razborov and Mark Alan Epstein. Lower bounds for the size of circuits of bounded depth in basis and, parity, 1986.

[Smo87] Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In Alfred V. Aho, editor, *STOC*, pages 77–82. ACM, 1987.