Certifying polynomials for $\mathrm{AC}^0[\oplus]$ circuits, with applications

Swastik Kopparty[1] and Srikanth Srinivasan[2]

1 **Rutgers University.** swastik.kopparty@rutgers.edu
2 **DIMACS, Rutgers University.** srikanth@dimacs.rutgers.edu

Abstract

In this paper, we introduce and develop the method of certifying polynomials for proving $\mathrm{AC}^0[\oplus]$ circuit lower bounds.

We use this method to show that Approximate Majority cannot be computed by $\mathrm{AC}^0[\oplus]$ circuits of size $n^{1+o(1)}$. This implies a separation between the power of $\mathrm{AC}^0[\oplus]$ circuits of near-linear size and uniform $\mathrm{AC}^0[\oplus]$ (and even AC^0) circuits of polynomial size. This also implies a separation between randomized $\mathrm{AC}^0[\oplus]$ circuits of linear size and deterministic $\mathrm{AC}^0[\oplus]$ circuits of near-linear size.

Our proof using certifying polynomials extends the deterministic restrictions technique of Chaudhuri and Radhakrishnan, who showed that Approximate Majority cannot be computed by AC^0 circuits of size $n^{1+o(1)}$. At the technical level, we show that for every $\mathrm{AC}^0[\oplus]$ circuit C of near-linear size, there is a low degree variety V over $\mathbb{F}_2$ such that the restriction of C to V is constant.

We also prove other results exploring various aspects of the power of certifying polynomials. In the process, we show an essentially optimal lower bound of $\Omega\left(\log^{\Theta(d)} s \cdot \log \frac{1}{\epsilon}\right)$ on the degree of ϵ-approximating polynomials for $\mathrm{AC}^0[\oplus]$ circuits of size s.

1998 ACM Subject Classification F.1.1 Models of Computation, F.1.2 Modes of Computation, F.1.3 Complexity Measures and Classes

Keywords and phrases Constant-depth Boolean circuits, Polynomials over finite fields, Size hierarchies

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2012.36

1 Introduction

In this paper, we introduce and develop the method of certifying polynomials for proving circuit lower bounds. We begin by describing the motivation for the main new circuit lower bound that we show, after which we will elaborate on the the method itself, and finally we describe some other results exploring the power and limitations of this method.

1.1 The Size Hierarchy Problem for $\mathrm{AC}^0[\oplus]$

Our main result fits in the general theme of studying the relative power of constant depth circuit classes. We show a near-tight circuit lower-bound for computing Approximate Majority with AND, OR, PARITY and NOT gates. This is a first step in the direction of a uniform size-hierarchy theorem for such circuits, which is a basic open question about this well-studied class of circuits.

We first fix some notation and conventions regarding circuits for the rest of this paper. AC^0 denotes the class of bounded depth circuits with unbounded fan-in AND, OR and NOT

32nd Int'l Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2012).
Editors: D. D'Souza, J. Radhakrishnan, and K. Telikepalli; pp. 36–47

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

gates. $\mathrm{AC}^0[\oplus]$ denotes the class of bounded depth circuits with unbounded fan-in AND, OR, PARITY and NOT gates. We measure the size of a circuit by the number of gates. We use n to denote the number of input bits to a circuit.

There is a well-developed theory giving superpolynomial and even subexponential lower bounds for AC^0 and $\mathrm{AC}^0[\oplus]$ circuits [6, 1, 15, 7, 9, 12]. Our focus in this paper is on complexity theory within these classes.

An influential paper of Ragde and Wigderson [8] asked if uniform AC^0 circuits of linear size are strictly weaker as uniform AC^0 circuits of polynomial size. This was answered by Chaudhuri and Radhakrishnan [5], who showed that Approximate Majority functions do not have AC^0 circuits of near-linear size $O(n^{1+\epsilon_d})$ (where $\epsilon_d > 0$). An Approximate Majority function is any function which maps strings of Hamming weight $< n/4$ to 0 and strings of Hamming weight $> 3n/4$ to 1. Such functions were first considered in the context of AC^0 for the purpose of error-reduction for AC^0 circuits. Ajtai and Ben-Or [2] showed that Approximate Majority can be computed by polynomial-size AC^0 circuits, and later results of Ajtai [1] and Viola [14] showed that this can even be done by uniform polynomial-size AC^0 circuits of depth 3 and above (In fact these circuits can be made to have depth d and size $O(n^{1+\epsilon_d})$, where $\epsilon_d \to 0$ as $d \to \infty$ [5]). This combined with the lower-bound of [5] showed the conjectured separation of Ragde and Wigderson.

The method of proof of [5] is especially interesting to us, and we will discuss their method and our extension of it in the next subsection.

A beautiful recent result of Rossman [11] showed a size-hierarchy for AC^0: for every integer $k > 0$, uniform AC^0 circuits of size $O(n^k)$ are more powerful than non-uniform AC^0 circuits of size $O(n^{k/4})$. A striking follow-up result of Amano [3] in fact shows that depth-2 size $O(n^k)$ uniform AC^0 circuits can be more powerful than size $O(n^{k-\epsilon})$ AC^0 circuits for arbitrary $\epsilon > 0$.

In this work we study the analogous questions for uniform $\mathrm{AC}^0[\oplus]$. Our main result is that Approximate Majority cannot be computed by $\mathrm{AC}^0[\oplus]$ circuits of near-linear size. In particular this means that polynomial size uniform $\mathrm{AC}^0[\oplus]$ circuits (and even polynomial size uniform AC^0 circuits) can be more powerful than near-linear size $\mathrm{AC}^0[\oplus]$ circuits. Thus we make a first step towards a size-hierarchy theorem for $\mathrm{AC}^0[\oplus]$ circuits, analogous to the result of Chaudhuri and Radhakrishnan for AC^0. Our result also shows that randomized $\mathrm{AC}^0[\oplus]$ circuits of linear size can be more powerful than deterministic $\mathrm{AC}^0[\oplus]$ circuits of near-linear size.

Showing the full size-hierarchy for uniform $\mathrm{AC}^0[\oplus]$ is still open and would be very interesting. Even the question of whether there exists a function that has uniform $\mathrm{AC}^0[\oplus]$ circuits of size $n^{\log n}$ but no polynomial-sized $\mathrm{AC}^0[\oplus]$ circuits (of possibly larger, but constant, depth) remains unanswered.

1.2 Certifying Polynomials for $\mathrm{AC}^0[\oplus]$

The main component of the [5] lower bound for Approximate Majority is a structure theorem for AC^0 circuits of near-linear size. It states that for every AC^0 circuit C of near-linear size, there is a collection of $o(n)$ variables and a fixing of them that simplifies the circuit C to a constant. Equivalently, there is a large axis-parallel subcube of $\{0,1\}^n$ on which C restricts to a constant. This structure theorem immediately implies the lower bound on Approximate Majority.

The proof of this structure theorem is by "deterministic restrictions". Going through the circuit in a bottom up fashion, one first finds a fixing of a small number of variables that simplifies the circuit into one where all the gates have small fan-in. The basic observation

is that if one considers the gates at height 1 that have large fan-in, then we can set a large number of them to constants by setting a few input variables; continuing in this way, we eventually remove all large fan-in gates of height 1 (there can't be too many of them, since C is of near-linear size), setting only a few variables in doing so. We then move on to higher levels and repeat the process, which now becomes feasible since setting gates of small fan-in to a constant reduces to setting only a few variables to constants. Once all the gates have small fan-in, the entire circuit is a function of only a few variables and hence, there is a fixing of small number of the remaining variables so that the circuit simplifies to a constant.

The main component of our lower bound is an analogous structure theorem for $\mathrm{AC}^0[\oplus]$. Clearly, the structure theorem for AC^0 is false for even a single parity gate and hence for $\mathrm{AC}^0[\oplus]$. However, here we can show that for for any $\mathrm{AC}^0[\oplus]$ circuit C of near-linear size, there is a polynomial of degree $o(n)$ such that C restricts to a constant on the zero-set of that polynomial. We call such a polynomial a *certifying polynomial* for the circuit C. The proof of this structure theorem again proceeds in a bottom up fashion, but this time finds fixings of systems of low-degree polynomials in order to simplify the circuit to one where all the AND and OR gates have small fan-in. Again, once all the AND and OR gates have small fan-in, it is easy to see that the circuit just computes a low-degree polynomial, and thus fixing this low-degree polynomial simplifies the circuit to a constant.

Given this structure theorem, it remains to see that no Approximate Majority function has this structure. This turns out to be a consequence of the general fact that a nonzero polynomial of degree d cannot vanish at every point of a Hamming ball of radius d (this follows from the fact that Hamming balls are interpolating sets for polynomials). In fact, it is even true that polynomials of degree $o(d)$ cannot vanish on all but an exponentially small fraction of a Hamming ball of radius d (this is a consequence of the p-biased version of the standard bound on the number of zeroes of a nonzero polynomial). We conclude that an Approximate Majority function cannot be constant on the zero set of a nonzero polynomial of degree $< n/4$. Combined with the structure theorem, this completes the proof of the lower bound for the $\mathrm{AC}^0[\oplus]$ complexity of Approximate Majority.

Having proved the lower bound, we then take a step back to re-examine the technique of proving lower bounds via certifying polynomials. On the face of it, it seems like this method is somewhat distinct from the Razborov-Smolensky method [9, 13] used to prove lower bounds for general $\mathrm{AC}^0[\oplus]$ circuits, which uses *polynomial approximations* to circuits. The Razborov-Smolensky method gives global, approximate structure: it shows that for any $\mathrm{AC}^0[\oplus]$ circuit C of size M, there is a polynomial of degree $\mathrm{poly}(\log(M))$ which agrees with C on most points of $\{0,1\}^n$. Our structure theorem, which only applies to circuits of near-linear size, gives local, exact structure: we get a perfect description of the values taken by an $\mathrm{AC}^0[\oplus]$ circuit on a small but structured subset of $\{0,1\}^n$.

As it turns out, however, the framework of certifying polynomials is quite robust: we demonstrate a connection between polynomial approximations and certifying polynomials for circuits. We then use this connection along with Razborov's approximating polynomials to construct certifying polynomials for general $\mathrm{AC}^0[\oplus]$ circuits. These polynomials have degree much larger than that obtained in our structure theorem, but nevertheless, their degree is small enough to be able to recover the exponential lower bound obtained by Razborov [9] for $\mathrm{AC}^0[\oplus]$ circuits computing the Majority function. We stress that most of the ideas of this lower bound proof are already present in [9, 13], and the main aim of this exercise is to show that the use of certifying polynomials is a unified framework that "explains" all previous lower bound approaches for $\mathrm{AC}^0[\oplus]$. In the course of the above proof, we also construct improved approximations to $\mathrm{AC}^0[\oplus]$ circuits in the small error regime; to the best

of our knowledge, such approximations were not known before, and may be of independent interest.

Finally, we exploit the connection between certifying polynomials and polynomial approximations in the reverse direction to prove limits on the power of polynomial approximations. We show that the low-error approximations we construct for $\mathrm{AC}^0[\oplus]$ are close to the best possible for all depths $d \geq 3$. Once again, this demonstrates the flexibility of the certifying polynomials framework.

2 Results

We begin by formally defining certifying polynomials. Throughout the paper, we identify $\{0,1\}$ with $\mathbb{F}_2$.

▶ **Definition 1** (Certifying polynomial). A polynomial $P(X_1, \ldots, X_n) \in \mathbb{F}_2[X_1, \ldots, X_n]$ is a certifying polynomial for a function $f : \{0,1\}^n \to \{0,1\}$ if:

- the set $S = \{x \in \mathbb{F}_2^n \mid P(x) = 0\}$ is nonempty,
- f is constant on S.

We now define Approximate Majority.

▶ **Definition 2** (Approximate Majority). An $(a, n-a)$ Approximate Majority is a boolean function $f : \{0,1\}^n \to \{0,1\}$ such that:

- $f(x) = 0$ for every x of Hamming weight at most a.
- $f(x) = 1$ for every x of Hamming weight at least $n - a$.

If we omit the $(a, n-a)$, we assume $a = n/4$.

Ajtai and Ben-Or [2] showed that for $a \leq n/2 - n/(\log n)^{O(1)}$, there exists an $(a, n-a)$ Approximate Majority computable in AC^0. We will use a uniform and more general version of this result, due to Ajtai [1].

▶ **Theorem 3** (Ajtai [1]). *For any $n \in \mathbb{N}$, $\delta \in (0, 1/2)$ and depth $d \geq 3$, there exist $((1/2-\delta)n, (1/2+\delta)n)$ Approximate Majorities computable by uniform AC^0 circuits of size $2^{(1/\delta)^{O(1/d)}} \cdot n^{O(1)}$ and depth d.*

Our main result is:

▶ **Theorem 4.** *For every constant $d \in \mathbb{N}$, there is an $\epsilon_d > 0$ such that any depth d $\mathrm{AC}^0[\oplus]$ circuit that computes an Approximate Majority must have size $\Omega(n^{1+\epsilon_d})$.*

By Theorem 3, this implies that uniform AC^0 circuits of polynomial size are more powerful than linear-sized non-uniform $\mathrm{AC}^0[\oplus]$ circuits.

The proof of Theorem 4 yields $\epsilon_d = 1/2^{d+1}$. This is marginally better than the lower bound of $\Omega(n^{1+1/4^d})$ obtained by Chaudhuri and Radhakrishnan [5] for the case of AC^0 (the improvement is due to a slightly different deterministic restriction method: whereas [5] try to remove both high fan-in and high fan-out gates from the circuit, we handle only the high fan-in gates). Also, as already showed by [5], this lower bound cannot be substantially improved since Approximate Majorities can be computed by AC^0 circuits of depth d and size $n^{1+1/2^{\Omega(d)}}$.

The proof of Theorem 4 follows from two lemmas. The first states that every function with a near-linear $\mathrm{AC}^0[\oplus]$ circuit has a certifying polynomial of low degree, The next states that an Approximate Majority cannot have this property. We now state these lemmas formally (the proofs appear in Section 3).

▶ **Lemma 5** (Linear-size $\mathrm{AC}^0[\oplus]$ circuits have low degree certifying polynomials)**.** *For every constant $d \in \mathbb{N}$, there is an $\epsilon_d > 0$ such that for every depth-d $\mathrm{AC}^0[\oplus]$ circuit C of size $s \leq n^{1+\epsilon_d}$, C has a certifying polynomial of degree $o(n)$.*

▶ **Lemma 6** (Approximate Majority does not have any low degree certifying polynomials)**.** *For every $(a, n-a)$ Approximate Majority f, there do not exist any certifying polynomials for f of degree $\leq a$.*

Next we state our results on certifying polynomials for general $\mathrm{AC}^0[\oplus]$ circuits. This result should be contrasted with the fact that every function has a certifying polynomial of degree at most $n/2$.

▶ **Theorem 7.** *For every $s > 0$ and constant $d > 0$, every $\mathrm{AC}^0[\oplus]$ circuit C of size s and depth d has a certifying polynomial of degree at most $n/2 - n/(\log s)^{\Theta(d)}$.*

We also show that this is essentially tight.

▶ **Lemma 8.** *For every $s > n^{\Omega(1)}$, there exist $\mathrm{AC}^0[\oplus]$ circuits C on n input bits with size s, such that every certifying polynomial for C has degree at least $n/2 - n/(\log s)^{\Theta(d)}$.*

These results are proved in Section 4. The proof of Theorem 7 uses the well-studied notion of approximating polynomials.

▶ **Definition 9** (ϵ-approximating polynomial)**.** An ϵ-approximating polynomial for a function f is a polynomial P such that $\Pr_{x \in \{0,1\}^n}[f(x) = P(x)] \geq 1 - \epsilon$.

The main ingredient in the proof of Theorem 7 is the following strengthening of Razborov's original theorem on approximating polynomials.

▶ **Lemma 10.** *For any $\epsilon \in (0, 1/2)$, any $\mathrm{AC}^0[\oplus]$ circuit C of size s and depth d has an ϵ-approximating polynomial of degree at most $(c \log s)^{d-1} \cdot (\log(1/\epsilon))$.*

We also show in Section 4 how Theorem 7 gives an alternate proof of Razborov's fundamental result that Majority does not have subexponential size $\mathrm{AC}^0[\oplus]$ circuits.

Finally, we state our lower bounds for the degree of approximating polynomials for $\mathrm{AC}^0[\oplus]$ circuits, showing the near-tightness of Lemma 10.

▶ **Theorem 11.** *For every $s, \epsilon > 0$, and every constant $d \geq 3$, there exist $\mathrm{AC}^0[\oplus]$ circuits C of size s and depth d such that for every polynomial P which is an ϵ-approximating polynomial for C, we have*

$$\deg(P) \geq \left(\log s - O(\log\log \frac{1}{\epsilon})\right)^{\Theta(d)} \cdot \log \frac{1}{\epsilon}.$$

The proof appears in Section 5.

3 Superlinear $\mathrm{AC}^0[\oplus]$ lower bounds for computing Approximate Majority

In this section, we prove Lemma 5 and Lemma 6, thus completing the proof of Theorem 4.

3.1 Linear-size $\mathrm{AC}^0[\oplus]$ circuits have low degree certifying polynomials

We now prove Lemma 5.

It will be more convenient to work with a certifying system of polynomials as opposed to a single certifying polynomial. Given a feasible system of polynomial equations over n variables $x_1, x_2, \ldots, x_n$, say

$$\begin{aligned} p_1(x) &= 0 \\ p_2(x) &= 0 \\ &\vdots \\ p_t(x) &= 0 \end{aligned}$$

we define the degree of the system to be $\sum_{i=1}^{t} \deg(p_i)$. Clearly, the set of solutions to the above system is exactly the set of roots of $1 - \prod_{i=1}^{t}(1 - p_i)$, which is a polynomial of degree at most $\sum_{i=1}^{t} \deg(p_i)$.

Given a feasible system of polynomial equations $\mathcal{P}$, we denote by $\mathrm{Sol}(\mathcal{P})$ the non-empty set of solutions of $\mathcal{P}$; when $\mathcal{P}$ sets just a single polynomial p, we denote use $\mathrm{Sol}(p)$ instead of $\mathrm{Sol}(\mathcal{P})$. By a *restriction*, we will mean simply a feasible system of polynomial equations.

Given a restriction $\mathcal{P}$ and a boolean circuit C, we will denote by $C|_{\mathcal{P}}$ the circuit C restricted to inputs from $\mathrm{Sol}(\mathcal{P})$. We say a gate g of the circuit C is *live* under the restriction given by $\mathcal{P}$ if g takes values 0 as well as 1 under inputs from $\mathrm{Sol}(\mathcal{P})$. Note that if a gate g is not live under a restriction, we can simplify the circuit C to a smaller circuit C' which computes the same function on the restricted inputs.

We say that a circuit C is *live* under the restriction $\mathcal{P}$ if every gate of C is live under $\mathcal{P}$. The above implies that, given any circuit C and restriction $\mathcal{P}$, there exists a live circuit C' of size at most the size of C that computes the same function as C on inputs from $\mathrm{Sol}(\mathcal{P})$.

Proof of Lemma 5. The proof will proceed as follows: after restricting the given circuit C to the roots of a well-chosen low-degree polynomial restriction $\mathcal{P}$, we will obtain an equivalent circuit C' that has the property that each of the AND and OR gates of C' have very small fan-in (say n^{ϵ} for $\epsilon \ll 1/d$). At this point, the entire circuit C' computes a low-degree polynomial p and by fixing p to a feasible value, we finish the proof of the lemma.

Say we have an increasing sequence of numbers $1 < D_1 < \ldots < D_d$ (we will fix the exact values of D_i ($i \geq 1$) later). We wish to obtain a restriction $\mathcal{P}$ under which C is equivalent to a circuit C' which has the property that every AND and OR gate at height i has fan-in at most D_i. It is easy to see that this implies that the function computed by C' is a polynomial of degree at most $D_1 D_2 \cdots D_d$.

We proceed to construct a suitable restriction $\mathcal{P}$ in d steps. After the ith step, we obtain a restriction $\mathcal{P}_i$ under which there is a circuit C_i of size at most s for which the above fan-in bound holds for all heights $j \leq i$. Assuming that the $(i-1)$th step has been completed, we describe how Step i is performed for $i \geq 1$. (Note that nothing needs to be done for height 0.)

We assume that C_i is live. Otherwise, we can obtain and work with an equivalent circuit that is of at most the size of C_i and satisfies the same fan-in restrictions as C_i. Let B_i denote the "bad" gates at height i: that is, the AND and OR gates at height i that have fan-in at least D_i. We use a basic subroutine $\mathrm{Fix}(i, C_i)$ that simplifies the circuit C_i by augmenting the restriction $\mathcal{P}_i$ as follows:

$\mathrm{Fix}(i, C_i)$: Since there are at least $|B_i|D_i$ wires between gates in B_i and lower levels (which contain at most s gates), there is some gate g at height less than i that is adjacent to $|B_i|D_i/s$ gates. By the fan-in restrictions on C_i, this gate computes a polynomial p_g of degree at most $D_1 \cdots D_{i-1}$ (the empty product in the case $i = 1$ is assumed to be 1). Moreover, since the circuit C_i is live, this gate can be set to both 0 and 1. We wish to add the restriction $p_g = 0$ or $p_g - 1 = 0$ to $\mathcal{P}_i$ corresponding to setting the gate to 0 or 1 respectively. Setting the gate g to 1 sets all the OR gates that g feeds into to 1 and setting g to 0 sets all the AND gates that g feeds into to 0. Hence, there is some setting that sets at least $|B_i|D_i/2s$ many gates in B_i to constant. We set the gate g to this boolean value.

Note that $\mathrm{Fix}(i, C_i)$ reduces the number of live bad gates to at most $|B_i|(1 - D_i/2s)$. We are now ready to describe Step i. Until the set of bad nodes B_i is empty, we repeatedly call the subroutine, $\mathrm{Fix}(i, C'_i)$ where C'_i represents the circuit we currently have. After an application of the subroutine $\mathrm{Fix}(i, C'_i)$ adds another equation to our current restriction $\mathcal{P}'_i$, we fix the non-live nodes and simplify the circuit until it becomes live again (this process, of course, does not increase the fan-in of any node). Note that since we are only fixing live nodes, the system of polynomial equations $\mathcal{P}'_i$ we maintain is feasible. Moreover, since the size of B_i is falling by a factor of at most $(1 - D_i/2s)$ after each application of $\mathrm{Fix}(i, C'_i)$ and $|B_i| \leq s \leq n^{O(1)}$, we need to apply $\mathrm{Fix}(i, C'_i)$ at most $\frac{2s \log |B_i|}{D_i} = O(s \log n/D_i)$ times to reduce B_i to the empty set.

Let us analyze the total degree of the equations added to the restriction during the ith step. Each equation added is a polynomial of degree at most $D_1 D_2 \cdots D_{i-1}$. Hence, the total degree of the added equations is $O(s \log n D_1 D_2 \cdots D_{i-1}/D_i)$.

At the end of Step d, we have a circuit C_d computing a polynomial of degree at most $D' = D_1 D_2 \cdots D_d$ that agrees with the original circuit C on a restriction of degree at most

$$D'' = O(s \log n) \left(\frac{1}{D_1} + \frac{D_1}{D_2} + \frac{D_1 D_2}{D_3} + \cdots + \frac{D_1 D_2 \cdots D_{d-1}}{D_d} \right)$$

We would like to set $D_1, \cdots, D_d$ such that both D' and D'' to be $o(n)$. We will choose K and the D_is such that $D_1 D_2 \cdots D_{i-1}/D_i = K$ for each i. This implies that $D_i = K^{2^{i-1}}$. Furthermore, we have $D' \leq K^{2^d}$ and $D'' \leq O(s \log n/K)$.

Setting $K = n^{1/(2^d+1)}$ and $\epsilon_d = \frac{1}{2^{d+1}}$, we get D' as well as D'' are $o(n)$ as long as $s \leq n^{1+\epsilon_d}$. Thus, by setting the polynomial p computed by the circuit C_d to some feasible value, we obtain a restriction of degree $D' + D'' = o(n)$ under which the circuit C becomes constant.

As mentioned above, this implies that there is a certifying polynomial for C of degree $o(n)$. ◀

3.2 Approximate Majority does not have any low degree certifying polynomials

We now prove Lemma 6.

Proof of Lemma 6. Let p be any polynomial of degree $d \leq a$ that takes the value 0 at some point of $\mathbb{F}_2^n$. We will show that it cannot be that f is constant on $\mathrm{Sol}(p)$.

Our intermediate claim is that $\mathrm{Sol}(p)$ intersects every Hamming ball of radius a. By translating p if necessary, we may assume that the Hamming ball is centered at the origin, and thus we seek to prove that there is a point of Hamming weight at most a where p vanishes.

Given the intermediate claim, it follows that there exist $x_0, x_1 \in \mathbb{F}_2^n$ with $p(x_0) = p(x_1) = 0$ such that the Hamming weight of x_0 is at most a, and the Hamming weight of x_1 is at least $n - a$. Thus f cannot be constant on $\mathrm{Sol}(p)$.

Now we prove the claim. Let $\tilde{p}$ denote the unique multilinear polynomial which agrees with p on $\mathbb{F}_2^n$. Since $\deg(p) \leq a$, we have $\deg(\tilde{p}) \leq a$. Now let q be the polynomial $1 - \tilde{p}$. Notice that q is multilinear and has degree at most a. Since $\mathrm{Sol}(p)$ is nonempty, we see that q is non-zero. Consider the monomials of q. Since $q \neq 0$, there must be a minimal $S \subseteq [n]$ (possibly empty) such that the monomial $\prod_{i \in S} X_i$ appears in q (i.e., has a non-zero coefficient) but no monomial $\prod_{i \in T} X_i$ for $T \subsetneq S$ appears in q. Let $x \in \mathbb{F}_2^n$ be the input that takes value 1 at exactly the indices in S. It is easy to see that $q(x) = 1$ and hence $\tilde{p}(x) = p(x) = 0$. Moreover, the Hamming weight of x is equal to the size of S which is at most $\deg(q) \leq a$. Hence, we see that $\mathrm{Sol}(p)$ does intersect the Hamming ball of radius a. This completes the proof of the claim, and hence the proof of Lemma 6.

◀

4 Certifying polynomials for general $\mathrm{AC}^0[\oplus]$ circuits

Given the results of the previous section, it makes sense to ask what are the lowest degree certifying polynomials we can obtain for general (i.e. significantly larger than linear-sized) $\mathrm{AC}^0[\oplus]$ circuits. Using an easy linear algebraic argument, it can be shown that *every* function, irrespective of its complexity, has a certifying polynomial of degree at most $n/2$ (and in this generality, it cannot be improved). In this section, we use Razborov's approximations for $\mathrm{AC}^0[\oplus]$ circuits by probabilistic polynomials to derive somewhat better certifying polynomials for functions with small $\mathrm{AC}^0[\oplus]$ circuits. In particular, we show that polynomial-sized $\mathrm{AC}^0[\oplus]$ circuits have certifying polynomials of degree $n/2 - n/(\log n)^{O(1)}$.

Though the improvement over the trivial $n/2$ bound above might seem small, the existence of such certifying polynomials is quite powerful: we demonstrate this by showing how this fact, along with Lemma 6, can be used to give a (slightly) conceptually different proof of Razborov's result that Majority does not have subexponential size $\mathrm{AC}^0[\oplus]$ circuits. We note that the proof is essentially unchanged at a technical level from the proofs of [9, 13], but the higher-order concepts involved seem curiously different. More specifically, this seems to provide a different 'constructive' (in the sense of Razborov and Rudich [10]) lower bound criterion for lower bounds against $\mathrm{AC}^0[\oplus]$ which is reminiscent of the work of Aspnes et al. [4].

The main theorem of this section is the following.

▶ Theorem 7 (Restated from Section 2). For every $s > 0$ and constant $d > 0$, every $\mathrm{AC}^0[\oplus]$ circuit C of size s and depth d has a certifying polynomial of degree at most $n/2 - n/(\log s)^{\Theta(d)}$.

The above theorem shows that functions computed by small subexponential size $\mathrm{AC}^0[\oplus]$ circuits have nontrivial certifying polynomials.

We will need to use probabilistic polynomials in the proof.

▶ **Definition 12** (Probabilistic polynomials)**.** An ϵ-error probabilistic polynomial of degree D for a function $f : \{0,1\}^n \to \{0,1\}$ is a random polynomial $\mathbf{P}$ of degree at most D (chosen according to some distribution over polynomials of degree at most D) such that for any $x \in \{0,1\}^n$, we have $\Pr_{\mathbf{P}}[f(x) = \mathbf{P}(x)] \geq 1 - \epsilon$.

Clearly, if a function f has an ϵ-error probabilistic polynomial $\mathbf{P}$ of degree D, then by averaging, it has an ϵ-approximating polynomial P of degree D as well.

We need the following well-known theorem, due to Razborov, on the existence of ϵ-error probabilistic polynomials for $\mathrm{AC}^0[\oplus]$. A node of a circuit C is said to be *internal* if it is not a leaf.

▶ **Theorem 13** (Razborov [9]). *For any $\epsilon \in (0, 1/2)$, any $\mathrm{AC}^0[\oplus]$ circuit C with at most s internal nodes and depth $d \geq 1$ has an ϵ-error probabilistic polynomial of degree at most $(\log(s/\epsilon))^d$. In particular, C has an ϵ-approximating polynomial of degree at most $(\log(s/\epsilon))^d$.*

Using Theorem 13 directly in our arguments would only give us a version of Theorem 7 with weaker parameters. To obtain the parameters mentioned above, we need a strengthening of Theorem 13 that does better for small ϵ. The proof follows quite simply from Razborov's theorem above, though to the best of our knowledge, this has not been observed in the literature.

▶ Lemma 10 (Restated from Section 2 in a stronger form). For any $\epsilon \in (0, 1/2)$, any $\mathrm{AC}^0[\oplus]$ circuit C of size s and depth d has an ϵ-error probabilistic polynomial of degree at most $(c \log s)^{d-1} \cdot (\log(1/\epsilon))$ for some absolute constant $c > 0$. In particular, C has an ϵ-approximating polynomial of degree at most $(c \log s)^{d-1} \cdot (\log(1/\epsilon))$.

Proof. Let C be an $\mathrm{AC}^0[\oplus]$ circuit of size s and depth d. Let g be the output gate of the circuit and let $C_1, \ldots, C_k$ ($k \leq s$) be the depth $d-1$ subcircuits of C feeding into g. By Theorem 13, we know that each C_i ($i \in [k]$) has a $(1/10s)$-approximating polynomial $\mathbf{P}_i$ of degree at most $(O(\log s))^{d-1}$. Also by Theorem 13, we know that the function computed by g has an $\mathrm{AC}^0[\oplus]$ circuit with just one *internal* node and hence has a $(1/10)$-approximating polynomial $\mathbf{P}$ of degree $O(1)$. The probabilistic polynomial $\mathbf{P'} := \mathbf{P}(\mathbf{P}_1, \ldots, \mathbf{P}_k)$ is a $1/5$-error probabilistic polynomial for C, since for any $x \in \{0,1\}^n$,

$$\begin{aligned}
\Pr_{\mathbf{P'}}[C(x) \neq \mathbf{P'}(x)] \leq & \Pr_{\mathbf{P}_1,\ldots,\mathbf{P}_k}[\exists i \in [k] :\ C_i(x) \neq \mathbf{P}_i(x)] + \\
& \Pr_{\mathbf{P}}[g(C_1(x), \ldots, C_k(x)) \neq \mathbf{P}(C_1(x), \ldots, C_k(x))] \\
\leq & \sum_{i \in [k]} \Pr_{\mathbf{P}_i}[C_i(x) \neq \mathbf{P}_i(x)] + \\
& \Pr_{\mathbf{P}}[g(C_1(x), \ldots, C_k(x)) \neq \mathbf{P}(C_1(x), \ldots, C_k(x))] \\
\leq & k/10s + 1/10 \leq 1/10 + 1/10 = 1/5
\end{aligned}$$

Note that $\mathbf{P'}$ has degree at most $(O(\log s))^{d-1}$. Let $\ell = c' \log(1/\epsilon)$ for a constant c' that we will choose later in the proof. Let $\mathbf{P}'_1, \ldots, \mathbf{P}'_\ell$ be ℓ independent copies of the probabilistic polynomial $\mathbf{P'}$. Let $\mathbf{Q}$ denote the probabilistic polynomial $\mathrm{Maj}(\mathbf{P}'_1, \ldots, \mathbf{P}'_\ell)$, where Maj is just the polynomial of degree at most ℓ that computes the majority of ℓ bits. Clearly, $\mathbf{Q}$ is of degree at most $(O(\log s))^{d-1} \cdot \ell = (O(\log s))^{d-1} \cdot \log(1/\epsilon)$. We claim that $\mathbf{Q}$ is an ϵ-error probabilistic polynomial for C, which will finish the proof of the corollary.

For any input $x \in \{0,1\}^n$, each $\mathbf{P}'_j(x)$ predicts the value of $C(x)$ correctly with probability $4/5$. Now, for $\mathbf{Q}(x)$ to predict $C(x)$ incorrectly, a *majority* of the $\mathbf{P}'_j$ ($j \in [\ell]$) must predict the value of $C(x)$ incorrectly and by a Chernoff bound, the probability of this is bounded by $\exp\{-\Omega(\ell)\}$, which is at most ϵ for a large enough constant $c' > 0$. ◀

The next lemma shows that functions with low-degree ϵ-approximating polynomials also have low-degree certifying polynomials.

▶ **Lemma 14.** *Suppose $f : \{0,1\}^n \to \{0,1\}$ has a degree D ϵ-approximating polynomial. Then f has a certifying polynomial of degree at most $\frac{n}{2} - c_1\sqrt{n \log \frac{1}{\epsilon}} + D$, where c_1 is an absolute constant.*

Proof. Let P be the given ϵ-approximating polynomial. Let S be the set of points where P differs from f. We have $|S| \leq \epsilon \cdot 2^n$.

Let D_0 be the smallest integer such that

$$\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{D_0} > |S|.$$

By linear algebra, there is a non-zero polynomial Q of degree at most D_0 that vanishes on S. Note that one of $Q \cdot P$ and $Q \cdot (1-P)$ is a non-zero polynomial. Moreover, for any input x s.t. $Q(x) = 1$, $P(x) = f(x)$.

Thus, it follows that one of $1 - Q \cdot P$ or $1 - Q \cdot (1-P)$ is a certifying polynomial for f with degree at most $D_0 + D$ (provided $D_0 + D < n$; if not then the result is vacuously true). To finish the proof, we note that $D_0 \leq \frac{n}{2} - c_1\sqrt{n \log \frac{1}{\epsilon}}$. ◀

Proof of Theorem 7. Combining Lemma 10 and Lemma 14, we conclude that for every $\epsilon > 0$. C has a certifying polynomial of degree at most

$$\frac{n}{2} - c_1\sqrt{n \cdot \log \frac{1}{\epsilon}} + (c_2 \log s)^{d-1} \cdot \log(1/\epsilon),$$

where $c_1, c_2 > 0$ are absolute constants. In particular, setting $\epsilon = \exp\{-n/(\log s)^{\Theta(d)}\}$ above, we get that C has a certifying polynomial of degree at most $n/2 - n/(\log s)^{\Theta(d)}$. ◀

Combining Theorem 7 with Lemma 6 (and using the fact that Majority is an $(n/2 - 1, n/2+1)$ Approximate Majority), we get an alternate proof of the fact that Majority cannot be computed by $\mathrm{AC}^0[\oplus]$ circuits of size smaller than $\exp(n^{\Omega(1/d)})$.

Finally, we show that the bound of Theorem 7 is essentially tight.

▶ **Lemma 15** (Restated from Section 2). *For every $s > n^{\Omega(1)}$, there exist $\mathrm{AC}^0[\oplus]$ circuits C on n input bits with size s, such that every certifying polynomial for C has degree at least $n/2 - n/(\log s)^{\Theta(d)}$.*

Proof. Let δ be a parameter to be specified later. Let C be the $\mathrm{AC}^0[\oplus]$ circuit for $((1/2 - \delta)n, (1/2+\delta)n)$ Approximate Majority given by Theorem 3. Then we have $|C| = 2^{(1/\delta)^{O(\frac{1}{d})}} \cdot n^{O(1)}$. We choose δ so that $|C| = s$; this gives $\delta = \frac{1}{(\log s - c \log n)^{\Omega(d)}}$.

By Lemma 6, any certifying polynomial for C has degree at least $n \cdot (\frac{1}{2} - \delta) = \frac{n}{2} - \frac{n}{(\log s)^{\Theta(d)}}$. ◀

5 Lower bounds for approximating polynomials

We now use the tools of the previous two sections to show near-optimal lower bounds on the degree of approximating polynomials for $\mathrm{AC}^0[\oplus]$ circuits. It is a folklore fact that ϵ-approximations for $\mathrm{AC}^0[\oplus]$ circuits of size s and depth d are required to have degree at least $\max\{(\log s)^{\Omega(d)}, \log(1/\epsilon)\}$. In this section, we show a stronger lower bound of $\Theta((\log s)^{\Omega(d)} \cdot \log(1/\epsilon))$, which essentially matches the upper bound obtained in Lemma 10. Our lower bound example is just a suitable Approximate Majority and thus holds even for AC^0 circuits.

We prove the lower bound by exploiting Lemma 14 in the contrapositive. Since there are Approximate Majorities that are efficiently computable in AC^0, by Lemma 6, we know that AC^0 circuits can compute functions that do not have efficient certifying polynomials. We can then use Lemma 14 to infer a lower bound on the degree of ϵ-approximations to AC^0 circuits.

▶ Theorem 11 (Restated from Section 2). For every $s, \epsilon > 0$, and every constant $d \geq 3$, there exist $\mathrm{AC}^0[\oplus]$ circuits C of size s and depth d such that for every polynomial P which is an ϵ-approximating polynomial for C, we have

$$\deg(P) \geq \left(\log s - O(\log\log \frac{1}{\epsilon})\right)^{\Theta(d)} \cdot \log \frac{1}{\epsilon}.$$

Proof. Let δ and m be parameters (to be specified later). Let C be an $\mathrm{AC}^0[\oplus]$ circuit on m inputs which computes a $((\frac{1}{2}-\delta)m, (\frac{1}{2}+\delta)m)$-approximate majority. By Theorem 3, such an AC^0 circuit can be taken to have depth d and size at most $2^{(1/\delta)^{O(\frac{1}{d})}} \cdot m^{O(1)}$. We will choose m and δ so that this size equals s.

Suppose P is an ϵ-approximating polynomial for C with degree D. By Lemma 14, there is a degree $\frac{m}{2} - c_1\sqrt{m \log \frac{1}{\epsilon}} + D$ polynomial Q which is a certifying polynomial for C.

But since C is a $((\frac{1}{2}-\delta)m, (\frac{1}{2}+\delta)m)$ Approximate Majority, Lemma 6 tells us that that $\deg(Q) \geq (\frac{1}{2}-\delta)\cdot m$.

Putting this together, we get that $D \geq c_1\sqrt{m \log \frac{1}{\epsilon}} - \delta \cdot m$.

We now choose m, δ so that $c_1\sqrt{m \log \frac{1}{\epsilon}} = 2\delta \cdot m$ and $s = 2^{(1/\delta)^{O(\frac{1}{d})}} \cdot m^{O(1)}$. Thus:

$$m = \left(\log s - O(\log\log \frac{1}{\epsilon})\right)^{\Theta(d)} \cdot \log \frac{1}{\epsilon}.$$

We therefore get

$$D \geq \left(\log s - O(\log\log \frac{1}{\epsilon})\right)^{\Theta(d)} \cdot \log \frac{1}{\epsilon},$$

as desired. ◀

6 Discussion and Open Questions

We have seen that certifying polynomials are a natural and useful notion in the context of lower bounds for $\mathrm{AC}^0[\oplus]$ circuits. We also saw that they have a rather interesting interaction with the well-studied notion of approximating polynomials for $\mathrm{AC}^0[\oplus]$ circuits.

The fundamental question we would like to answer is whether we can prove a size-hierarchy theorem for $\mathrm{AC}^0[\oplus]$ analogous to the results of Rossman [11] and Amano [3] for AC^0. It would even be interesting to obtain the weaker separation of uniform $\mathrm{AC}^0[\oplus]$ circuits of size $n^{\log n}$ from polynomial-sized $\mathrm{AC}^0[\oplus]$ circuits? Good candidates for proving these separations seem to be the parity of the number of k-cliques in a graph for the former, and the elementary symmetric polynomial of degree $\log n$ for the latter. We have taken the first step in this direction by demonstrating a function that has polynomial-sized uniform AC^0 circuits but not near-linear sized $\mathrm{AC}^0[\oplus]$ circuits.

Another question that we leave open is to prove lower bounds on the degree for ϵ-approximating polynomials for depth 2 $\mathrm{AC}^0[\oplus]$ circuits. Our lower bound utilized small $\mathrm{AC}^0[\oplus]$ circuits for Approximate Majority, which only exist for depth 3 and higher.

It would be interesting to see whether certifying objects (analogous to the certifying polynomials studied here) exist for other, more powerful, circuit classes, and if they can be used to prove new circuit lower bounds.

Acknowledgements

We would like to thank Albert Atserias for asking us about the tradeoff between degree and error for approximating polynomials for $AC^0[\oplus]$ circuits. We would also like to thank the reviewers of FSTTCS 2012 for pointing out some errors and improving the quality of the exposition.

References

1 Miklós Ajtai. *Approximate counting with uniform constant-depth circuits.*, pages 1–20. Providence, RI: American Mathematical Society, 1993.
2 Miklós Ajtai and Michael Ben-Or. A theorem on probabilistic constant depth computations. In *STOC*, pages 471–474, 1984.
3 Kazuyuki Amano. k-subgraph isomorphism on AC^0 circuits. *Computational Complexity*, 19(2):183–210, 2010.
4 James Aspnes, Richard Beigel, Merrick L. Furst, and Steven Rudich. The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994.
5 Shiva Chaudhuri and Jaikumar Radhakrishnan. Deterministic restrictions in circuit complexity. In *STOC*, pages 30–36, 1996.
6 Merrick L. Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17(1):13–27, 1984.
7 Johan Håstad. *Computational limitations of small-depth circuits.* The MIT Press, Cambridge(MA)-London, 1987.
8 Prabhakar Ragde and Avi Wigderson. Linear-size constant-depth polylog-treshold circuits. *Inf. Process. Lett.*, 39(3):143–146, 1991.
9 Alexander A. Razborov. Lower bounds on the size of constant-depth networks over a complete basis with logical addition. *Mathematicheskie Zametki*, 41(4):598–607, 1987.
10 Alexander A. Razborov and Steven Rudich. Natural proofs. *Journal of Computer and System Sciences*, 55(1):24–35, 1997.
11 Benjamin Rossman. On the constant-depth complexity of k-clique. In *STOC*, pages 721–730, 2008.
12 Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing*, pages 77–82, 1987.
13 Roman Smolensky. On representations by low-degree polynomials. In *FOCS*, pages 130–138, 1993.
14 Emanuele Viola. On approximate majority and probabilistic time. *Computational Complexity*, 18(3):337–375, 2009.
15 Andrew Chi-Chih Yao. Separating the polynomial-time hierarchy by oracles (preliminary version). In *Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science*, pages 1–10, 1985.