Towards lower bounds for bounded-depth Frege proofs with modular connectives

Alexis Maciel and Toniann Pitassi

Abstract. We show that for every prime power p^k, quasipolynomial-size bounded-depth Frege proofs with mod p^k counting connectives can be simulated by quasipolynomial-size proofs of depth 3 consisting of a threshold connective at the output, mod p^k connectives on level two, and AND connectives of small fan-in on level one. We argue that this result is a plausible first step towards proving lower bounds for bounded-depth Frege proofs with modular connectives, an outstanding open problem. We also discuss possible interesting consequences for propositional theorem proving.

1. Introduction

An outstanding problem in logic and complexity theory is to prove superpolynomial lower bounds for classical propositional proof systems, known as Frege systems. This problem appears to be quite difficult, although in the last 10 years, substantial progress has been made on restricted versions of the problem. For example, it is now known that Resolution proofs as well as bounded-depth Frege proofs of the propositional pigeonhole principle require exponential size, and these results hold even if we add the mod p axiom schema for any prime p [**Hak85, Ajt88, Ajt90, BP, Rii**].

It has been noted many times that various proof systems correspond to various Boolean circuit classes, and it has been the case that the proof system lower bound has only been obtained once lower bounds for the corresponding circuit class have been firmly established. For example, bounded-depth Frege systems correspond to bounded-depth circuits, and the lower bounds for bounded-depth Frege systems use and generalize much of the technical machinery behind the AC^0 lower bounds for the parity function. In a similar manner, bounded-depth Frege systems with mod p connectives correspond to $\mathrm{ACC}^0[p]$, and Frege systems correspond to NC^1.

Since there are strong lower bounds for $\mathrm{ACC}^0[p]$ due to Razborov [**Raz87**] and Smolensky [**Smo87**], the next step in propositional proof complexity is to obtain

1991 *Mathematics Subject Classification.* Primary 03F20; Secondary 68Q15.

Part of this work was performed in the departments of Mathematics and Computer Science at the University of Pittsburgh.

The first author was supported in part by NSF Grant CCR-9457782 and CCR-9522084.

The second author was supported in part by NSF Grant CCR-9457782, US-Israel BSF Grant 95-00238, and Grant INT-9600919/ME-103 from NSF and MŠMT (Czech Republic).

exponential lower bounds for the corresponding class of bounded-depth Frege systems with mod p connectives. However, at present we cannot even show superpolynomial lower bounds for bounded-depth Frege systems with modular connectives, even under a reasonable complexity-theoretic assumption like $P \neq NP$. (Obviously $NP \neq coNP$ implies superpolynomial lower bounds for any propositional proof system, so by reasonable, we mean a weaker assumption than this.) A lot of effort has already gone towards solving this problem, and so far it has resulted in new and elegant algebraic proof systems, such as Nullstellensatz proofs ([**BIK$^+$a, BIK$^+$b**]), and Gröbner proofs [**CEI96**]. One promising approach for proving lower bounds for $\mathrm{ACC}^0[p]$ proofs is to prove lower bounds for Gröbner refutations. In fact, it can be shown that lower bounds for $\mathrm{ACC}^0[p]$ proofs imply lower bounds for Gröbner refutations over $\mathbf{Z}_p$. This important step was made very recently by Razborov [**Raza**] who showed that any Gröbner refutation of the propositional pigeonhole principle requires linear degree. (The Gröbner proof system has since been renamed the *polynomial calculus.*)

Unfortunately, lower bounds for the polynomial calculus do not appear to be enough to obtain lower bounds for $\mathrm{ACC}^0[p]$ proofs. The essential problem is that good lower bounds for Gröbner refutations seem to give good lower bounds for a very flat bounded-depth proof system having only mod p connectives at the top of each formula, rather than mod p connectives throughout the bounded-depth formula.

Surprising results from circuit complexity tell us that this may not be a problem. Allender [**All89**] (see also [**AH94**] and [**Yao90**], [**BT94**], [**GKR$^+$95**] and [**Reg93**] for related results) has shown that quasipolynomial-size $\mathrm{ACC}^0[p]$ circuits can be simulated by quasipolynomial-size, depth-3 circuits. The depth-3 circuits are of a special form: the output gate is a threshold gate, the middle layer of gates are mod p gates, and the bottom level consists of AND gates of small fan-in. Alternatively, this result can be viewed as showing that any $\mathrm{ACC}^0[p]$ circuit can be computed (with high probability for every 0/1 input) by a small-degree probabilistic polynomial over $\mathbf{Z}_p$. Thus, in the circuit world, a constant number of levels of mod p gates, intertwined with AND and OR gates, can be collapsed to a single small-degree probabilistic polynomial. The heart of the argument is that the OR function can be very well approximated by a small-degree probabilistic polynomial over $\mathbf{Z}_p$.

In this paper, we show that these collapsing results carry over to the world of proof systems. We show that any $\mathrm{ACC}^0[p]$ proof can be transformed into a quasipolynomial-size, depth-3 proof, with only one level of mod p connectives. Our result can be viewed as showing that the collapsing theorem for $\mathrm{ACC}^0[p]$ is highly constructive, since it can be formalized in the same proof system as the circuit class itself. In formal terms, the collapsing theorem is formalizable in a very restricted version of S_2^1.

The remainder of this paper is organized as follows. In section 2 we define our proof system. In Section 3 we give the general overview of the main theorem. In Section 4 we define the probabilistic polynomials used in the construction of the main theorem. In Sections 5, 6, and 7 we prove our main result, in the case of $\mathrm{ACC}^0[2]$ proofs. In Section 8 we generalize our result to $\mathrm{ACC}^0[p^k]$ proofs. In Section 9 we discuss possible extensions to ACC^0 proofs. Finally in Section 10 we conclude with a discussion of how this theorem suggests a method for proving lower bounds for $\mathrm{ACC}^0[p]$ proofs, as well as a method for obtaining a deterministic algorithm for propositional theorem proving.

2. A logic system with threshold and modular connectives

In this section, we describe a logic system with mod 2, negation, AND and OR connectives as well as threshold connectives. (The generalization of our results to mod p connectives, for p prime, will be discussed in Section 8.) Our system is an extension of the system PTK introduced by Buss and Clote [**BC96**, Section 10]. In their system, only threshold and negation connectives are allowed.

DEFINITION 2.1. Formula *depth* and *size* are defined inductively by:

1. A propositional variable x_i, $i \in \mathbf{N}$, is a formula of depth 0 and size 1.
2. If A is a formula then $\neg A$ is a formula of depth $1 + \mathrm{depth}(A)$ and size $1 + \mathrm{size}(A)$.
3. If $A_1, \ldots, A_n$ are formulas, $n \geq 0$, then $\wedge(A_1, \ldots, A_n)$ is a formula of depth $1 + \max\{\mathrm{depth}(A_i) : 1 \leq i \leq n\}$ and size $n + 1 + \sum_{1 \leq i \leq n} \mathrm{size}(A_i)$. We interpret $\wedge(A_1, \ldots, A_n)$ to be true if and only if all of the A_i's are true.
4. If $A_1, \ldots, A_n$ are formulas, $n \geq 0$, then $\vee(A_1, \ldots, A_n)$ is a formula of depth $1 + \max\{\mathrm{depth}(A_i) : 1 \leq i \leq n\}$ and size $n + 1 + \sum_{1 \leq i \leq n} \mathrm{size}(A_i)$. We interpret $\vee(A_1, \ldots, A_n)$ to be true if and only if one of the A_i's is true.
5. If $A_1, \ldots, A_n$ are formulas, $j = 0, 1$, $n \geq 0$, then $\oplus_j(A_1, \ldots, A_n)$ is a formula of depth $1 + \max\{\mathrm{depth}(A_i) : 1 \leq i \leq n\}$ and size $n + 1 + \sum_{1 \leq i \leq n} \mathrm{size}(A_i)$. We interpret $\oplus_j(A_1, \ldots, A_n)$ to be true if and only if the number of true A_i's is equal to j mod 2.
6. If $A_1, \ldots, A_n$ are formulas, $n, k \geq 0$, then $\mathrm{Th}_k(A_1, \ldots, A_n)$ is a formula of depth $1 + \max\{\mathrm{depth}(A_i) : 1 \leq i \leq n\}$ and size $(n + k) + 1 + \sum_{1 \leq i \leq k} \mathrm{size}(A_i)$. We interpret $\mathrm{Th}_k(A_1, \ldots, A_n)$ to be true if and only if the number of true A_i's is greater than or equal to k. In the following sections, we will use the more descriptive notation $\sum_{i=1}^{m} A_i \geq k$ to represent $\mathrm{Th}_k(A_1, \ldots, A_n)$.

In the above definitions, $\wedge(A_1, \ldots, A_n)$ denotes the logical AND of the multiset consisting of $A_1, \ldots A_n$, and similarly for $\vee$, $\oplus_j$ and Th_k. Thus commutativity of the connectives is implicit. A *cedent* is any sequence $A_1, \ldots, A_n$ of formulas separated by commas. Cedents will usually be designated by capital Greek letters such as Γ and Δ. A *sequent* is given by $\Gamma \rightarrow \Delta$, where Γ, Δ are arbitrary cedents. The size of a cedent $A_1, \ldots, A_n$ is $\sum_{1 \leq i \leq n} \mathrm{size}(A_i)$ and its depth is $\max_{1 \leq i \leq n}(\mathrm{depth}(A_i))$. The size of a sequent $\Gamma \rightarrow \Delta$ is $\mathrm{size}(\Gamma) + \mathrm{size}(\Delta)$ and its depth is $\max(\mathrm{depth}(\Gamma), \mathrm{depth}(\Delta))$. The intended interpretation of the sequent $\Gamma \rightarrow \Delta$ is that the conjunction of the formulas in Γ implies the disjunction of the formulas in Δ.

A proof of a sequent S in our logic system is a sequence of sequents, $S_1, ..., S_q$, such that each sequent S_i is either an initial sequent, or follows from previous sequents by one of the rules of inference, and the final sequent, S_q, is S. The size of the proof is $\sum_{1 \leq i \leq q} \mathrm{size}(S_i)$ and its depth is $\max_{1 \leq i \leq q}(\mathrm{depth}(S_i))$.

An *initial sequent* is of the following form:

1. $A \rightarrow A$ where A is any formula
2. $\rightarrow \wedge()$; $\vee() \rightarrow$
3. $\oplus_1() \rightarrow$; $\rightarrow \oplus_0()$
4. $\mathrm{Th}_k() \rightarrow$ for $k \geq 1$; $\rightarrow \mathrm{Th}_0(A_1, \ldots, A_n)$ for $n \geq 0$

The rules of inference are given by Table 1. Note that the logical rules are defined for $n \geq 1$ and $k \geq 1$.

structural rules

weak left: $$\frac{\Gamma, \Delta \rightarrow \Gamma'}{\Gamma, A, \Delta \rightarrow \Gamma'}$$ weak right: $$\frac{\Gamma \rightarrow \Gamma', \Delta'}{\Gamma \rightarrow \Gamma', A, \Delta'}$$

contract left: $$\frac{\Gamma, A, A, \Delta \rightarrow \Gamma'}{\Gamma, A, \Delta \rightarrow \Gamma'}$$ contract right: $$\frac{\Gamma \rightarrow \Gamma', A, A, \Delta'}{\Gamma \rightarrow \Gamma', A, \Delta'}$$

permute left: $$\frac{\Gamma, A, B, \Delta \rightarrow \Gamma'}{\Gamma, B, A, \Delta \rightarrow \Gamma'}$$ permute right: $$\frac{\Gamma \rightarrow \Gamma', A, B, \Delta'}{\Gamma \rightarrow \Gamma', B, A, \Delta'}$$

cut rule

$$\frac{\Gamma, A \rightarrow \Delta \qquad \Gamma' \rightarrow A, \Delta'}{\Gamma, \Gamma' \rightarrow \Delta, \Delta'}$$

logical rules

$\neg$-left: $$\frac{\Gamma \rightarrow A, \Delta}{\neg A, \Gamma \rightarrow \Delta}$$ $\neg$-right: $$\frac{A, \Gamma \rightarrow \Delta}{\Gamma \rightarrow \neg A, \Delta}$$

$\wedge$-left: $$\frac{A_1, \wedge(A_2, \ldots, A_n), \Gamma \rightarrow \Delta}{\wedge(A_1, \ldots, A_n), \Gamma \rightarrow \Delta}$$

$\wedge$-right: $$\frac{\Gamma \rightarrow A_1, \Delta \qquad \Gamma \rightarrow \wedge(A_2, \ldots, A_n), \Delta}{\Gamma \rightarrow \wedge(A_1, \ldots, A_n), \Delta}$$

$\vee$-left: $$\frac{A_1, \Gamma \rightarrow \Delta \qquad \vee(A_2, \ldots, A_n), \Gamma \rightarrow \Delta}{\vee(A_1, \ldots, A_n), \Gamma \rightarrow \Delta}$$

$\vee$-right: $$\frac{\Gamma \rightarrow A_1, \vee(A_2, \ldots, A_n), \Delta}{\Gamma \rightarrow \vee(A_1, \ldots, A_n), \Delta}$$

$\oplus$-left: $$\frac{A_1, \oplus_{i-1}(A_2, \ldots, A_n), \Gamma \rightarrow \Delta \qquad \oplus_i(A_2, \ldots, A_n), \Gamma \rightarrow A_1, \Delta}{\oplus_i(A_1, \ldots, A_n), \Gamma \rightarrow \Delta}$$

$\oplus$-right: $$\frac{A_1, \Gamma \rightarrow \oplus_{i-1}(A_2, \ldots, A_n), \Delta \qquad \Gamma \rightarrow A_1, \oplus_i(A_2, \ldots, A_n), \Delta}{\Gamma \rightarrow \oplus_i(A_1, \ldots, A_n), \Delta}$$

Th_k-left: $$\frac{\mathrm{Th}_k(A_2, \ldots, A_n), \Gamma \rightarrow \Delta \qquad A_1, \mathrm{Th}_{k-1}(A_2, \ldots, A_n), \Gamma \rightarrow \Delta}{\mathrm{Th}_k(A_1, \ldots, A_n), \Gamma \rightarrow \Delta}$$

Th_k-right: $$\frac{\Gamma \rightarrow A_1, \mathrm{Th}_k(A_2, \ldots, A_n), \Delta \qquad \Gamma \rightarrow \mathrm{Th}_{k-1}(A_2, \ldots, A_n), \Delta}{\Gamma \rightarrow \mathrm{Th}_k(A_1, \ldots, A_n), \Delta}$$

TABLE 1. Rules of inference

THEOREM 2.2. *Our proof system is sound and complete.*

PROOF. (Proof sketch.) A *truth assignment* is a mapping $\nu : \{x_i : i \in \mathbf{N}\} \rightarrow \{0, 1\}$. By induction on formula depth, it is clear how to extend ν to assign a truth value to every formula and sequent of our logic system. A sequent is *valid* if it is true for every truth assignment. By induction on the number of inferences in proofs, it is straightforward to show that every theorem of our system is valid. Therefore, our system is sound.

Our proof of completeness follows that of Buss and Clote [**BC96**]. The central idea is as follows. We will say that an inference rule of the form "S_1 and S_2 derive S_3" has the *inversion property* if for every truth assignment ν, $\nu(S_3) = 1$ implies $\nu(S_2) = 1$ and also $\nu(S_2) = 1$. In other words, soundness holds in both directions. It is not hard to check that all of the logical proof rules of our system satisfy the inversion property.

In order to prove completeness, suppose that $\Gamma \to \Delta$ is valid. We will construct a cut-free proof of $\Gamma \to \Delta$ in a top-down fashion, by applying our logical rules so as to "break up" formulas in Γ and Δ into smaller formulas. As long as there is a formula in either Γ or Δ that is not atomic, this is possible since we have logical rules which allow us to break down every connective occurring on either the left or the right side of a sequent. When we can proceed no further, we are left with leaf sequents in which every formula is atomic. The final step is to show that if $\Gamma \to \Delta$ is a leaf sequent, then it is either an initial sequent, or it can be obtained by applying weakening to an initial sequent. This final step follows directly from the inversion property: Since the original sequent was a tautology, by the inversion property, all leaf sequents must also be tautologies, and therefore can be shown to be of the right form. $\square$

DEFINITION 2.3. Let $F = \{(\Gamma_n \to \Delta_n) : n \in N\}$ be a family of sequents in which all formulas involve only the connectives $\neg$, $\wedge$, $\vee$ and $\oplus$. Then $\{R_n : n \in \mathbf{N}\}$ is a family of $\mathrm{ACC}^0[2]$ proofs for F if there exist constants c and d such that the following conditions hold: (1) Each R_n is a valid proof of $(\Gamma_n \to \Delta_n)$ in our system, and furthermore involves only the connectives $\neg$, $\wedge$, $\vee$ and $\oplus$; (2) For all i, the depth of R_n is at most d; and (3) For all n, the size of R_n is at most $(\mathrm{size}(\Gamma_n \to \Delta_n))^c$.

We say that a formula f can be *arranged into d levels* if the connectives of f can be arranged into d groups $L_1, \ldots, L_d$ called levels such that all the inputs of every connective at some level are either propositional variables or connectives from the previous levels. Note that f can be arranged into d levels if and only if f has depth at most d. Moreover, if f has depth less than d, then some of the levels may be empty.

DEFINITION 2.4. Let $\Gamma \to \Delta$ be a sequent in which all formulas involve only the connectives $\neg$, $\wedge$, $\vee$ and $\oplus$. Then R is a size-s *flat* proof of $\Gamma \to \Delta$ if R is of size at most s and every formula in R can be arranged into three levels such that level 3 contains only $\wedge$, $\vee$ and threshold connectives, level 2 contains only $\oplus$ connectives, and level 1 contains only $\wedge$ connectives of fan-in $\log s$. If $F = \{(\Gamma_n \to \Delta_n) : n \in \mathbf{N}\}$ is a family of sequents in which all formulas involve only the connectives $\neg$, $\wedge$, $\vee$ and $\oplus$, then $\{R_n : n \in \mathbf{N}\}$ is a family of size-$s(n)$ *flat* proofs for F if, for every n, R_n is a size-$s(n)$ flat proof of $\Gamma_n \to \Delta_n$.

In the sections that follow, polynomials over $\mathbf{Z}_2$ in the propositional variables $x_1, \ldots, x_n$ will play an important role. In fact, most of the formulas in this article will be statements about such polynomials. To this end, we will adopt the following conventions. First, the symbols $\oplus$ and $\bigoplus$ will be used to denote addition of polynomials over $\mathbf{Z}_2$. They are to be distinguished from $\oplus_0$ and $\oplus_1$ which represent the parity connectives. Second, when written as part of a formula, a polynomial u will be interpreted as the natural formula expressing the fact that $u \equiv 1 \pmod 2$.

For example, $x_1 \oplus x_2x_3$ is interpreted as $\oplus_1(x_1, \wedge(x_2, x_3))$, and $1 \oplus x_1 \oplus x_2x_3$, as $\oplus_0(x_1, \wedge(x_2, x_3))$. Note that we always assume that polynomials are written in standard form, i.e., as sums of monomials.

Throughout this article, we will establish that particular sequents can be derived in our logic system by quasipolynomial-size flat proofs. This will not be done by writing detailed proofs of these sequents. Instead, we will outline how the proof should go, indicating the main steps, and leave to the reader the tedious but straightforward task of filling in the details. These missing details will usually amount to the proof of some simple fact about formulas with threshold and parity connectives. We will also leave to the reader the simple task of verifying the claimed bounds on the complexity of the proofs.

3. Overview of the simulation

Let $(\Gamma_1 \rightarrow \Delta_1), (\Gamma_2 \rightarrow \Delta_2), \ldots, (\Gamma_L \rightarrow \Delta_L)$ be a depth-d, size-s $\mathrm{ACC}^0[2]$ proof in the propositional variables $x_1, \ldots, x_n$. We assume that $s \in 2^{(\log n)^{O(d)}}$. This proof will be fixed throughout Sections 3, 4 and 5. Allender [**All89**] (see also [**AH94**]) has shown that

THEOREM 3.1 (Allender [**All89**]). *Every* $\mathrm{ACC}^0[2]$ *circuit can be simulated by a depth-three, quasipolynomial-size circuit consisting of an unweighted threshold gate at the output,* MOD_2 *gates on level two, and* AND *gates of polylog fan-in on level one.*

By viewing the propositional variables as inputs, every formula A in the $\mathrm{ACC}^0[2]$ proof can be viewed as an $\mathrm{ACC}^0[2]$ circuit. Define $\mathrm{tr}(A)$, the *translation* of A, to be a depth-three circuit that simulates A, according to the theorem. If $\Gamma = (A_1, \ldots, A_m)$, then let $\mathrm{tr}(\Gamma) = (\mathrm{tr}(A_1), \ldots, \mathrm{tr}(A_m))$. Since for every formula A, $\mathrm{tr}(A)$ and A compute the same function, it follows that $\mathrm{tr}(\Gamma_i) \rightarrow \mathrm{tr}(\Delta_i)$ is valid if and only if $\Gamma_i \rightarrow \Delta_i$ is valid.

Now replace every $\Gamma_i \rightarrow \Delta_i$ in the proof by its translation $\mathrm{tr}(\Gamma_i) \rightarrow \mathrm{tr}(\Delta_i)$. The mere fact that $\mathrm{tr}(\Gamma_i) \rightarrow \mathrm{tr}(\Delta_i)$ is valid exactly when $\Gamma_i \rightarrow \Delta_i$ is valid does not imply that $(\mathrm{tr}(\Gamma_1) \rightarrow \mathrm{tr}(\Delta_1)), \ldots, (\mathrm{tr}(\Gamma_L) \rightarrow \mathrm{tr}(\Delta_L))$ is a proof. However, we will show that this sequence constitutes the skeleton of a proof that simulates the original $\mathrm{ACC}^0[2]$ proof. We will show that whenever a rule is used to infer $\Gamma_i \rightarrow \Delta_i$ from $\Gamma_j \rightarrow \Delta_j$ and $\Gamma_k \rightarrow \Delta_k$, $i > j \geq k$, then $\mathrm{tr}(\Gamma_i) \rightarrow \mathrm{tr}(\Delta_i)$ can be derived from $\mathrm{tr}(\Gamma_j) \rightarrow \mathrm{tr}(\Delta_j)$ and $\mathrm{tr}(\Gamma_k) \rightarrow \mathrm{tr}(\Delta_k)$. Then, by putting together all these derivations, and by appending a derivation of $\Gamma_L \rightarrow \Delta_L$ from $\mathrm{tr}(\Gamma_L) \rightarrow \mathrm{tr}(\Delta_L)$, we will obtain a proof that simulates the original $\mathrm{ACC}^0[2]$ proof. As for the complexity of this new proof, we will show that all the formulas it contains are of depth three, quasipolynomial size, and of the form threshold of MOD_2's of AND's of polylog fan-in, except for those in the derivation of $\Gamma_L \rightarrow \Delta_L$ from $\mathrm{tr}(\Gamma_L) \rightarrow \mathrm{tr}(\Delta_L)$. In this last derivation, subformulas of $\Gamma_L \rightarrow \Delta_L$ will also appear. Therefore, in the case where $\Gamma_L \rightarrow \Delta_L$ is simply a DNF formula $\rightarrow C_1, \ldots, C_m$, the simulating proof is a quasipolynomial-size flat proof.

There are many proofs of Theorem 3.1 and, consequently, many possible definitions for $\mathrm{tr}(A)$. But all these definitions follow the same pattern. First, a low-degree *probabilistic polynomial* Q_* is associated with every connective $*$ occurring in our $\mathrm{ACC}^0[2]$ proof. Precise definitions will be given in the next section; for the moment, we will only say that Q_* computes the output of the connective $*$ with high probability. Then, the polynomials associated with all the connectives occurring in a formula A are composed, according to the structure of A, to yield a probabilistic polynomial P_A that computes the value of A with high probability. Finally, the formula $\mathrm{tr}(A)$ is defined as the majority, over all the possible random choices, of the probabilistic polynomial P_A. This formula has the same value as A since A is true precisely when P_A is 1 with high probability.

In the following sections, we will define the polynomials Q_* associated with the various connectives, the polynomials P_A and the formulas $\mathrm{tr}(A)$. We will then establish that several key properties of these polynomials and of $\mathrm{tr}(A)$ have depth-three, quasipolynomial-size proofs of the required form. Finally, these properties will be used to show that every inference rule used in an $\mathrm{ACC}^0[2]$ proof can be simulated by a flat proof in our logic system.

4. Probabilistic polynomials

A *probabilistic polynomial* is a polynomial in two sets of variables: ordinary variables and *probabilistic variables*. Such a polynomial P, with ordinary variables $u = u_1, \dots, u_m$ and probabilistic variables $v = v_1, \dots, v_r$, computes a function f with error ε if, for every value of u, $\mathrm{Prob}_v[P(u,v) \neq f(u)] < \varepsilon$. The probabilistic variables are assigned uniformly and independently chosen values from $\{0,1\}$.

Let $\varepsilon = 1/2^{\lceil 2+\log s\rceil}$. That is, ε is the inverse of the smallest power of 2 greater than or equal to $4s$. For every $m \in [s]$ and every connective $*$ of fan-in m occurring in our $\mathrm{ACC}^0[2]$ proof, we define a probabilistic polynomial Q_*^m over $\mathbf{Z}_2$ that computes the connective with error ε.

Given a connective of fan-in m, let $u = (u_1, \dots, u_m)$ denote its terms. The polynomial associated with the $\neg$ connective is simply $Q_\neg(u_1) = u_1 \oplus 1$. For the mod 2 connectives, we have $Q_{\oplus_1}^m(u) = \bigoplus_{i=1}^m u_i$ and $Q_{\oplus_0}^m(u) = (\bigoplus_{i=1}^m u_i) \oplus 1$. Recall that all these are polynomials over $\mathbf{Z}_2$.

Consider now an $\vee$ of fan-in m. The existence of a low-degree probabilistic polynomial $Q_\vee^m$ that computes the $\vee$ of m variables with error ε is the key step in the proof of Theorem 3.1. Many constructions are possible, each yielding a different proof of this circuit simulation. The definition that seems to be the most appropriate for our purposes, because of the complexity of the associated proofs, is in terms of universal hashing.

Let $S = \{i : u_i = 1\}$. We want $Q_\vee^m$ to determine, with high probability, if $|S| = 0$ or if $|S| \geq 1$. Now either $|S| = 0$ or $\frac{1}{8}2^k \leq |S| < \frac{1}{4}2^k$, for some $k \in \{3, \dots, \lceil \log m \rceil + 3\}$. For every such k, let $\mathcal{H}_k$ be a small universal family of hash functions with domain D containing $[m]$ and range $[2^k]$. That is, $\mathcal{H}_k \subseteq \{h : D \to [2^k]\}$, $[m] \subseteq D$, $|\mathcal{H}_k|$ is quasipolynomial in m, and, for every $i \neq j$ in D, the probability that $h(i) = h(j) = r$ is $1/2^{2k}$, when h and r are randomly chosen from $\mathcal{H}_k$ and $[2^k]$. Such families $\mathcal{H}_k$ exist: for example, see **[Wig94]** for a family of size $2^{2\lceil \log m \rceil}$. For our purposes, the particular choice of $\mathcal{H}_k$ is not important, as long as it is a small universal family of hash functions from D to $[2^k]$. For technical reasons, we will also require that $|\mathcal{H}_k|$ be a power of 2 no smaller than 4.

Such a family $\mathcal{H}_k$ can be viewed as a matrix H_k of dimensions $|\mathcal{H}_k| \times m$ with entries in $[2^k]$. Associate a function from $\mathcal{H}_k$ to each row of H_k, in some arbitrary way, and define $H_k(h,i)$ to be $h(i)$. It is then not difficult to see that the universal property of $\mathcal{H}_k$ can be equivalently stated as follows: given any $i \neq j$ in $[m]$, the number of rows in which columns i and j contain the same element is equal to $|\mathcal{H}_k|/2^k$.

The idea behind the definition of $Q^m_\vee$ is as follows. For every k, let $R_k = \bigoplus_{i:h(i)=r} u_i$, where the pair h, r is randomly and uniformly chosen from $\mathcal{H}_k \times [2^k]$. If $|S| = 0$, then $R_k = 0$. On the other hand, by using the universal property of the family of hash functions $\mathcal{H}_k$, we will show in the next section that if $\frac{1}{8}2^k \leq |S| < \frac{1}{4}2^k$, then $R_k = 1$ with probability at least $1/16$. This will imply that if $|S| \geq 1$, then, with probability at least $1/16$, $R_k = 1$ for some $k \in \{3, \ldots, \lceil \log m \rceil + 3\}$. This probability will then be amplified to $1-\varepsilon$ by taking $C = \lceil \log \frac{1}{\varepsilon} / \log \frac{16}{15} \rceil$ independent copies.

The polynomial $Q^m_\vee$ will "encode" the above in the following way. Every pair h, r with $h \in \mathcal{H}_k$ and $r \in [2^k]$ can be encoded as a binary string of length $\log |\mathcal{H}_k| + k$. The sequence of probabilistic variables w of $Q^m_\vee$ will be interpreted as a sequence of encodings for pairs $h_{l,k}, r_{l,k}$, $l = 1, \ldots, C$ and $k = 3, \ldots, \lceil \log m \rceil + 3$, where $h_{l,k} \in \mathcal{H}_k$ and $r_{l,k} \in [2^k]$. For every $i \in [m]$ and every l, k, let $F^i_{l,k}$ be a degree $\log |\mathcal{H}_k| + k$ polynomial over $\mathbf{Z}_2$ in the variables encoding $h_{l,k}, r_{l,k}$ such that $F^i_{l,k}$ takes the value 1 if $h_{l,k}(i) = r_{l,k}$, and the value 0, otherwise. Note that when the probabilistic variables are set to specific values, then $F^i_{l,k}$ is simply a constant. In that case, both as a polynomial and as a formula, $\bigoplus_{i=1}^m u_i F^i_{l,k}$ is identical to $R_{l,k} = \bigoplus_{i:h_{l,k}(i)=r_{l,k}} u_i$, the residue modulo 2 of the number of $i \in S$ such that $h_{l,k}(i) = r_{l,k}$. Recall that we always assume that polynomials are written in standard form, i.e., as sums of monomials.

DEFINITION 4.1. The polynomial $Q^m_\vee$ is defined by

$$Q^m_\vee(u, w) = 1 \oplus \prod_{l=1}^{C} \prod_{k=3}^{\lceil \log m \rceil + 3} \left(1 \oplus \bigoplus_i u_i F^i_{l,k} \right).$$

Note that $Q^m_\vee$ has polylog(n) degree and uses polylog(n) probabilistic variables.

If $|S| = 0$, then clearly $Q^m_\vee(u, w) = 0$. Otherwise, if $|S| \geq 1$, then $Q^m_\vee(u, w) = 1$ with probability $1 - \varepsilon$. Therefore, $Q^m_\vee$ does compute $\vee(u)$ with error ε; that is, for every value of u,

$$\mathrm{Prob}_w[Q^m_\vee(u, w) = 1] \begin{cases} \geq 1 - \varepsilon & \text{if } \vee(u) = 1 \\ = 0 & \text{if } \vee(u) = 0 \end{cases}$$

A key step in the proof of our main result, the depth-three simulation of $\mathrm{ACC}^0[2]$ proofs, will be to formalize the proof of this property of $Q^m_\vee$, in our logic system, using a depth-three, quasipolynomial-size proof of the form threshold of MOD_2's of AND's of polylog fan-in, i.e., a flat proof.

The polynomial $Q^m_\wedge$ associated with an $\wedge$ of fan-in m is defined from $Q^m_\vee$ as

$$Q^m_\wedge(u, w) = 1 \oplus Q^m_\vee(1 \oplus u_1, \ldots, 1 \oplus u_m, w).$$

In this case, we have that for every value of u,

$$\mathrm{Prob}_w[Q^m_\wedge(u, w) = 1] \begin{cases} = 1 & \text{if } \wedge(u) = 1 \\ \leq \varepsilon & \text{if } \wedge(u) = 0 \end{cases}$$

We now turn to the definition of the probabilistic polynomials P_A that compute, with high probability, the value of the formulas A that appear in our $\mathrm{ACC}^0[2]$ proof. Let $y_1, \ldots, y_d$ be d disjoint sequences of probabilistic variables whose length is the maximum number of probabilistic variables required by any of the polynomials Q^m_*, for $m \in [s]$. Let A be a formula of depth d appearing in the $\mathrm{ACC}^0[2]$ proof. The polynomial P_A, with ordinary variables $x = (x_1, \ldots, x_n)$ and probabilistic variables $y_1, \ldots, y_d$, is obtained by replacing each occurrence of a connective $*$ of fan-in m at level i by the corresponding probabilistic polynomial Q^m_*, using probabilistic variables y_i.

DEFINITION 4.2. For $i = 1, \ldots, n$, $P_{x_i} = x_i$. Next, $P_{\wedge()} = 1$, $P_{\vee()} = 0$, $P_{\oplus_1()} = 0$, $P_{\oplus_0()} = 1$. Finally, if $m \geq 1$ and $A = *(A_1, \ldots, A_m)$ is a formula of depth $d \geq 1$, then P_A is the probabilistic polynomial in $x, y_1, \ldots, y_d$ defined by $P_A = Q^m_*(P_{A_1}, \ldots, P_{A_m}, y_d)$.

For example, for

$$A = \vee(\wedge(x_1, x_2), x_3, \vee(x_1, x_4)),$$

we have

$$P_A(x_1, x_2, x_3, x_4, y_1, y_2) = Q^3_\vee(Q^2_\wedge(x_1, x_2, y_1), x_3, Q^2_\vee(x_1, x_4, y_1), y_2).$$

Note that P_A has $(\log n)^{O(d)}$ degree and uses $\mathrm{polylog}(n)$ probabilistic variables.

By using the key property of $Q^m_\vee$, it is not difficult to show that for every value of x,

$$\mathrm{Prob}_{y_1,\ldots,y_d}[P_A(x, y_1, \ldots, y_d) = 1] \begin{cases} \geq 1 - \varepsilon s_A & \text{if } A(x) = 1 \\ \leq \varepsilon s_A & \text{if } A(x) = 0 \end{cases}$$

where s_A denotes the size of A. Based on this, we define the *translation* of A, denoted $\mathrm{tr}(A)$, to be the formula that expresses the fact that A is true by saying that the probability that $P_A = 1$ is very high.

DEFINITION 4.3. If A is a propositional variable x_i, $i = 1, \ldots, n$, or one of the formulas $\wedge()$, $\vee()$, $\oplus_1()$ or $\oplus_0()$, then $\mathrm{tr}(A) = A$. If A is any other formula of depth $d \geq 1$, then $\mathrm{tr}(A)$ is the formula

$$\sum_{\sigma_1,\ldots,\sigma_d} P_A(x, \sigma_1, \ldots, \sigma_d) \geq (1 - \varepsilon s_A) 2^{|y_1| + \cdots + |y_d|}$$

where $\sigma_1, \ldots, \sigma_d$ range over all possible values of the sequences of probabilistic variables $y_1, \ldots, y_d$.

Note that $\mathrm{tr}(A)$ is indeed a depth-three formula of size $2^{(\log n)^{O(d)}}$ and of the form threshold of MOD_2's of AND's of fan-in $(\log n)^{O(d)}$. In addition, the threshold in the definition of $\mathrm{tr}(A)$ is an integer since ε is the inverse of a power of 2 and $|y_i| > \log(1/\varepsilon)$, for every i.

5. The key property of $Q^m_\vee$

In this section, we show that the key property of $Q^m_\vee$, namely, that for every value of $u = (u_1, \ldots, u_m)$,

$$\mathrm{Prob}_w[Q^m_\vee(u, w) = 1] \begin{cases} \geq 1 - \varepsilon & \text{if } \vee(u_1, \ldots, u_m) = 1 \\ = 0 & \text{if } \vee(u_1, \ldots, u_m) = 0 \end{cases}$$

has a quasipolynomial-size flat proof in our logic system. Note that in the later applications of this property, the u_i's will not be merely propositional variables. They will be polynomials of degree $(\log n)^{O(d)}$ over $\mathbf{Z}_2$ in the propositional variables $x_1, \ldots, x_n$. Therefore, special care will be necessary to ensure that the resulting proofs are flat proofs.

This property can be proved in many ways. In every case, however, the main step is to establish that if $\frac{1}{8}2^k \leq |S| < \frac{1}{4}2^k$, where $S = \{i : u_i = 1\}$, then, with probability at least $1/16$, $R_k = \bigoplus_{i:h(i)=r} u_i$ is equal to 1. (The actual numbers in the bounds on $|S|$ and in the probability may vary slightly.) Recall that R_k is equal to the residue modulo 2 of the number of $i \in S$ such that $h(i) = r$. Therefore, one way to prove such a lower bound on the probability that $R_k = 1$ is to in fact prove that, with probability at least $1/16$, there is exactly one $i \in S$ such that $h(i) = r$. This in turn can be shown either by an argument similar to the BPP $\subseteq \Sigma_2^{\mathrm{P}}$ argument (see [**Sip83**]) or by a simple application of the inclusion-exclusion principle, as suggested in [**Reg93**].

It turns out, however, that such a proof is not appropriate for our purposes. The reason is that it seems to require statements such as "the probability that $u_i = 1$ and, for all $j \neq i$, that $h(j) \neq r$ or $u_j = 0$, is at least $1/2$." Since the u_i are in fact polynomials over $\mathbf{Z}_2$ that compute the connectives from the previous level, this would lead to a depth-four proof of the form threshold of AND's of MOD$_2$'s of small AND's.

To work our way around this problem, we will give a new proof of the lower bound on the probability that $R_k = 1$. Let $C_{h,r}$ denote the number of $i \in S$ such that $h(i) = r$. The idea is to use inclusion-exclusion directly to evaluate the probability that $C_{h,r}$ is odd, instead of the probability that $C_{h,r} = 1$. Since this proof does not seem to have been previously published, we first state and prove the result without worrying about the complexity of the proof.

PROPOSITION 5.1. *If $\frac{1}{8}2^k \leq |S| < \frac{1}{4}2^k$, then*

$$\mathrm{Prob}_{h,r \in \mathcal{H}_k \times [2^k]} \left[\left(\bigoplus_{i:h(i)=r} u_i \right) = 1 \right] \geq \frac{1}{16}.$$

PROOF. Using the notation introduced earlier, we want to show that $\mathrm{Prob}_{h,r \in \mathcal{H}_k \times [2^k]}[C_{h,r} \text{ is odd}] \geq \frac{1}{16}$. For every $i \in S$, let $B_i = \{h, r : (h(i) = r) \wedge (C_{h,r} \text{ is odd})\}$. Let $B = \cup_{i \in S} B_i$ and let $V = \{h, r : C_{h,r} \text{ is odd}\}$. Then, by the principle of inclusion-exclusion, we have that

$$|V| \geq |V - B| + \sum_{i \in S} |B_i| - \sum_{i,j \in S, i<j} |B_i \cap B_j|.$$

Dividing by $|\mathcal{H}_k \times [2^k]|$, we get that

$$\begin{aligned} \mathrm{Prob}_{h,r}[C_{h,r} \text{ is odd}] \quad \geq \quad & \sum_{i \in S} \mathrm{Prob}_{h,r}[(h(i) = r) \wedge (C_{h,r} \text{ is odd})] \\ & - \sum_{i,j \in S, i<j} \mathrm{Prob}_{h,r}[(h(i) = h(j) = r) \wedge (C_{h,r} \text{ is odd})]. \end{aligned}$$

The probability in the second term is at most $\mathrm{Prob}_{h,r}[h(i) = h(j) = r]$, and this is bounded above by $1/2^{2k}$ because $\mathcal{H}_k$ is a universal family of hash functions. As for the first term, for every $i \in S$, we have that

$$\begin{aligned}
&\mathrm{Prob}_{h,r}[(h(i) = r) \wedge (C_{h,r} \text{ is odd})] \\
&= \mathrm{Prob}_{h,r}[(h(i) = r) \wedge (|\{j \in S - \{i\} : h(j) = r\}| \text{ is even})] \\
&= \mathrm{Prob}_{h,r}[h(i) = r] - \mathrm{Prob}_{h,r}[(h(i) = r) \wedge (|\{j \in S - \{i\} : h(j) = r\}| \text{ is odd})] \\
&\geq \mathrm{Prob}_{h,r}[h(i) = r] - \mathrm{Prob}_{h,r}[(\exists j \in S - \{i\})(h(i) = h(j) = r)] \\
&\geq \mathrm{Prob}_{h,r}[h(i) = r] - \sum_{j \in S - i} \mathrm{Prob}_{h,r}[h(i) = h(j) = r].
\end{aligned}$$

The first term is easily seen to be equal to $1/2^k$. Therefore, by using once again the universal property of $\mathcal{H}_k$, we get that

$$\begin{aligned}
\mathrm{Prob}_{h,r}[C_{h,r} \text{ is odd}] &\geq \sum_{i \in S} \left(\frac{1}{2^k} - \sum_{j \in S - i} \frac{1}{2^{2k}} \right) - \sum_{i,j \in S, i<j} \frac{1}{2^{2k}} \\
&\geq |S| \frac{2^k - |S|}{2^{2k}} - \frac{\frac{1}{2}|S|^2}{2^{2k}} \\
&\geq \frac{1}{8}\left(1 - \frac{1}{4}\right) - \frac{1}{32} \\
&= \frac{1}{16}.
\end{aligned}$$

□

We now proceed to show that this proof can be formalized by a flat proof. First, we prove a lemma typical of the kind of basic facts about formulas with threshold and parity connectives that will be used in the subsequent proofs.

LEMMA 5.2. *Let P and Q be arbitrary polynomials of degree $(\log n)^{O(d)}$ over $\mathbf{Z}_2$ in the propositional variables $x_1, \ldots, x_n$. Then the sequents $P, Q \rightarrow PQ$, $PQ \rightarrow P$ and $PQ \rightarrow Q$ have flat proofs of size $2^{(\log n)^{O(d)}}$.*

PROOF. Suppose that $P = \sum_{i=1}^{s} A_i + a$ and $Q = \sum_{j=1}^{t} B_j + b$, where the A_i's are distinct monomials in the variables $x_1, \ldots, x_n$, $a \in \mathbf{Z}_2$ and similarly for Q. Recall that in a sequent such as $P, Q \rightarrow PQ$, P stands for the formula $\oplus_{1-a}(A_1, \ldots, A_s)$, where A_i denotes the conjunction of all the factors in the monomial A_i. Similarly, Q stands for $\oplus_{1-b}(B_1, \ldots, B_t)$.

The polynomial $PQ = \sum_{i=1}^{s} \sum_{j=1}^{t} A_i B_j + \sum_{j=1}^{t} aB_j + \sum_{i=1}^{s} bA_i + ab$, on the other hand, does not stand for the formula

$$\oplus_{1-ab}(A_1 B_1, \ldots, A_s B_t, aB_1, \ldots, aB_t, bA_1, \ldots, bA_s),$$

where $A_i B_j$ denotes the conjunction of all the factors in A_i and B_j, aB_j denotes B_j repeated a times and similarly for bA_i. Instead, PQ must first be written in standard form, i.e., as a sum of distinct monomials. Nevertheless, it is easy to prove that

$$PQ \rightarrow \oplus_{1-ab}(A_1 B_1, \ldots, A_s B_t, aB_1, \ldots, aB_t, bA_1, \ldots, bA_s)$$

and vice-versa.

Therefore, to prove that $P, Q \rightarrow PQ$, it is sufficient to show that

$$\begin{aligned}
(5.1) \qquad &\oplus_{1-a}(A_1, \ldots, A_s), \ \oplus_{1-b}(B_1, \ldots, B_t) \\
&\rightarrow \ \oplus_{1-ab}(A_1 B_1, \ldots, A_s B_t, aB_1, \ldots, aB_t, bA_1, \ldots, bA_s)
\end{aligned}$$

and to prove $PQ \to P$, it is sufficient to show that

$$\oplus_{1-ab}(A_1B_1, \ldots, A_sB_t, aB_1, \ldots, aB_t, bA_1, \ldots, bA_s) \\ \to \oplus_{1-a}(A_1, \ldots, A_s). \tag{5.2}$$

To prove both of these sequents, we will first show that for every c and d, the sequent

$$\oplus_c(A_1, \ldots, A_s),\ \oplus_d(B_1, \ldots, B_t) \ \to\ \oplus_{cd}(A_1B_1, \ldots, A_sB_t) \tag{5.3}$$

has a quasipolynomial-size flat proof. We do this by induction on s.

In what follows, we will make frequent use of several basic facts about parity formulas, such as

$$\oplus_c(A_1, \ldots, A_s),\ \oplus_d(B_1, \ldots, B_t) \ \to\ \oplus_{c+d}(A_1, \ldots, A_s, B_1, \ldots, B_t),$$

$\to \oplus_0(A_i), \oplus_1(A_i)$ and $A_i \to \oplus_1(A_i)$.

For the base case, $s = 0$, we must prove that $\oplus_c(),\ \oplus_d(B_1, \ldots, B_t) \to \oplus_{cd}()$. This sequent follows from $\oplus_c() \to$ if $c = 1$ and from $\to \oplus_{cd}()$ if $c = 0$.

Now assume that for every c' and d' we have a proof for

$$\oplus_{c'}(A_1, \ldots, A_{s-1}),\ \oplus_{d'}(B_1, \ldots, B_t) \ \to\ \oplus_{c'd'}(A_1B_1, \ldots, A_{s-1}B_t).$$

A proof for Sequent (5.3) is outlined below.

1. $\oplus_c(A_1, \ldots, A_s), A_s \to \oplus_{c-1}(A_1, \ldots, A_{s-1})$
2. $\oplus_{c-1}(A_1, \ldots, A_{s-1}), \oplus_d(B_1, \ldots, B_t) \to \oplus_{(c-1)d}(A_1B_1, \ldots, A_{s-1}B_t)$
3. $A_s, \oplus_d(B_1, \ldots, B_t) \to \oplus_d(A_sB_1, \ldots, A_sB_t)$
4. $\oplus_{(c-1)d}(A_1B_1, \ldots, A_{s-1}B_t), \oplus_d(A_sB_1, \ldots, A_sB_t)$
   $\to \oplus_{cd}(A_1B_1, \ldots, A_sB_t)$
5. $\oplus_c(A_1, \ldots, A_s), \oplus_d(B_1, \ldots, B_t), A_s \to \oplus_{cd}(A_1B_1, \ldots, A_sB_t)$

Lines 1 and 4 are simple basic facts. Line 2 is from the inductive hypothesis. The existence of a proof for Line 3 can be established by induction on t. Line 5 follows from the previous ones by three applications of the cut rule.

6. $\oplus_c(A_1, \ldots, A_s) \to A_s, \oplus_c(A_1, \ldots, A_{s-1})$
7. $\oplus_c(A_1, \ldots, A_{s-1}), \oplus_d(B_1, \ldots, B_t) \to \oplus_{cd}(A_1B_1, \ldots, A_{s-1}B_t)$
8. $\to A_s, \oplus_0(A_sB_1, \ldots, A_sB_t)$
9. $\oplus_{cd}(A_1B_1, \ldots, A_{s-1}B_t), \oplus_0(A_sB_1, \ldots, A_sB_t) \to \oplus_{cd}(A_1B_1, \ldots, A_sB_t)$
10. $\oplus_c(A_1, \ldots, A_s), \oplus_d(B_1, \ldots, B_t) \to A_s, \oplus_{cd}(A_1B_1, \ldots, A_sB_t)$

Lines 6 and 9 are simple basic facts. Line 7 is from the inductive hypothesis. The existence of a proof for Line 8 can be established by induction on t. Line 10 follows from Lines 6 to 9 by three applications of the cut rule. Sequent (5.3) now follows from Lines 5 and 10 by the cut rule.

We now prove Sequents (5.1) and (5.2) from Sequent (5.3). Sequent (5.1) is proved as follows.

1. $\oplus_{1-a}(A_1, \ldots, A_s), \oplus_{1-b}(B_1, \ldots, B_t) \to \oplus_{(1-a)(1-b)}(A_1B_1, \ldots, A_sB_t)$
2. $\oplus_{1-a}(A_1, \ldots, A_s), \oplus_{1-b}(B_1, \ldots, B_t), \oplus_{(1-a)(1-b)}(A_1B_1, \ldots, A_sB_t)$
   $\to \oplus_{1-ab}(A_1B_1, \ldots, A_sB_t, aB_1, \ldots, aB_t, bA_1, \ldots, bA_s)$

Line 1 is from Sequent (5.3). Line 2 is a basic fact. Note that $(1-a)(1-b) + a(1-b) + b(1-a) = 1 - ab$. Sequent (5.1) follows from these two lines by the cut rule.

The idea behind the proof of Sequent (5.2) is to prove the sequent

$$\oplus_{-a}(A_1,\dots,A_s) \;\rightarrow\; \oplus_{-ab}(A_1B_1,\dots,A_sB_t,aB_1,\dots,aB_t,bA_1,\dots,bA_s).$$

The proof proceeds as follows.

1. $\rightarrow \oplus_{-b}(B_1,\dots,B_t), \oplus_{1-b}(B_1,\dots,B_t)$
2. $\oplus_{-a}(A_1,\dots,A_s), \oplus_{-b}(B_1,\dots,B_t) \rightarrow \oplus_{ab}(A_1B_1,\dots,A_sB_t)$
3. $\oplus_{-a}(A_1,\dots,A_s), \oplus_{-b}(B_1,\dots,B_t), \oplus_{ab}(A_1B_1,\dots,A_sB_t)$
   $\rightarrow \oplus_{-ab}(A_1B_1,\dots,A_sB_t,aB_1,\dots,aB_t,bA_1,\dots,bA_s)$
4. $\oplus_{-a}(A_1,\dots,A_s), \oplus_{-b}(B_1,\dots,B_t)$
   $\rightarrow \oplus_{-ab}(A_1B_1,\dots,A_sB_t,aB_1,\dots,aB_t,bA_1,\dots,bA_s)$

Lines 1 and 3 are simple basic facts. For Line 3, note that $ab+a(-b)+b(-a) = -ab$. Line 2 is from Sequent (5.3). Line 4 follows by one application of the cut rule.

5. $\oplus_{-a}(A_1,\dots,A_s), \oplus_{1-b}(B_1,\dots,B_t) \rightarrow \oplus_{-a(1-b)}(A_1B_1,\dots,A_sB_t)$
6. $\oplus_{-a}(A_1,\dots,A_s), \oplus_{1-b}(B_1,\dots,B_t), \oplus_{-a(1-b)}(A_1B_1,\dots,A_sB_t)$
   $\rightarrow \oplus_{-ab}(A_1B_1,\dots,A_sB_t,aB_1,\dots,aB_t,bA_1,\dots,bA_s)$
7. $\oplus_{-a}(A_1,\dots,A_s), \oplus_{1-b}(B_1,\dots,B_t)$
   $\rightarrow \oplus_{-ab}(A_1B_1,\dots,A_sB_t,aB_1,\dots,aB_t,bA_1,\dots,bA_s)$
8. $\oplus_{-a}(A_1,\dots,A_s) \rightarrow \oplus_{-ab}(A_1B_1,\dots,A_sB_t,aB_1,\dots,aB_t,bA_1,\dots,bA_s)$

Line 5 is from Sequent (5.3). Line 6 is a simple basic fact. Note that $-a(1-b) + a(1-b) + b(-a) = -ab$. Line 7 follows from Lines 5 and 6 by the cut rule. Line 8 follows from Lines 1, 4 and 7 by two applications of the cut rule.

9. $\rightarrow \oplus_{-a}(A_1,\dots,A_s), \oplus_{1-a}(A_1,\dots,A_s)$
10. $\oplus_{-ab}(A_1B_1,\dots,A_sB_t,aB_1,\dots,aB_t,bA_1,\dots,bA_s)$
    $\oplus_{1-ab}(A_1B_1,\dots,A_sB_t,aB_1,\dots,aB_t,bA_1,\dots,bA_s) \rightarrow$

Lines 9 and 10 are simple basic facts. Sequent (5.2) follows from Lines 8, 9 and 10 by two applications of the cut rule. $\square$

We now establish the inclusion-exclusion principle that is needed. Let k be an arbitrary number in $\{3,\dots,\lceil\log m\rceil+3\}$. For every pair h,r in $\mathcal{H}_k \times [2^k]$, let $\varphi_{h,r}$ denote the polynomial $\bigoplus_{i:h(i)=r} u_i$. Then, for every i, let $\varphi^i_{h,r}$ denote $\bigoplus_{j:h(j)=r} u_iu_j$ and, for every $i \neq j$, let $\varphi^{i,j}_{h,r}$ denote $\bigoplus_{l:h(l)=r} u_iu_ju_l$. Notice that $\varphi^i_{h,r}$ and $\varphi^{i,j}_{h,r}$ are equivalent, respectively, to $u_i \wedge \varphi_{h,r}$ and $u_i \wedge u_j \wedge \varphi_{h,r}$.

Lemma 5.3 (Inclusion-Exclusion). *For every number t, the sequent*

$$\sum_i \left(\sum_{h,r:h(i)=r} \varphi^i_{h,r} \right) \geq t \;\rightarrow\; \sum_{h,r} \varphi_{h,r} + \sum_{i<j} \left(\sum_{h,r:h(i)=h(j)=r} \varphi^{i,j}_{h,r} \right) \geq t$$

has a flat proof of size $2^{(\log n)^{O(d)}}$.

Proof. If $t = 0$ or if $t > m|\mathcal{H}_k|2^k$, then the proof is simple and left to the reader. So suppose that $0 < t \leq m|\mathcal{H}_k|2^k$. The sequent to be proved is the same as

$$\sum_{h,r} \left(\sum_{i:h(i)=r} \varphi^i_{h,r} \right) \geq t \;\rightarrow\; \sum_{h,r} \left(\varphi_{h,r} + \sum_{i,j:i<j,h(i)=h(j)=r} \varphi^{i,j}_{h,r} \right) \geq t.$$

First, we show, for every pair h, r and for every number t_1, that

$$\sum_{i:h(i)=r} \varphi^i_{h,r} \geq t_1 \;\rightarrow\; \varphi_{h,r} + \sum_{i,j:i<j,h(i)=h(j)=r} \varphi^{i,j}_{h,r} \geq t_1. \tag{5.4}$$

If $t_1 = 0$, then there is nothing to prove. So suppose that $t_1 \geq 1$. Then

$$\sum_{i:h(i)=r} \varphi^i_{h,r} \geq t_1 \;\rightarrow\; \varphi_{h,r},$$

since, for every i, $\varphi^i_{h,r} \rightarrow \varphi_{h,r}$. In addition, since $\varphi^i_{h,r} \rightarrow u_i$, we also have that

$$\sum_{i:h(i)=r} \varphi^i_{h,r} \geq t_1 \;\rightarrow\; \sum_{i:h(i)=r} u_i \geq t_1.$$

It is not difficult to show that

$$\sum_{i:h(i)=r} u_i \geq t_1 \;\rightarrow\; \sum_{i,j:i<j,h(i)=h(j)=r} u_i u_j \geq \binom{t_1}{2}.$$

Now, since $\varphi_{h,r}, u_i u_j \rightarrow \varphi^{i,j}_{h,r}$,

$$\varphi_{h,r}, \sum_{i,j:i<j,h(i)=h(j)=r} u_i u_j \geq \binom{t_1}{2} \;\rightarrow\; \sum_{i,j:i<j,h(i)=h(j)=r} \varphi^{i,j}_{h,r} \geq \binom{t_1}{2}.$$

Therefore,

$$\sum_{i:h(i)=r} \varphi^i_{h,r} \geq t_1 \;\rightarrow\; \sum_{i,j:i<j,h(i)=h(j)=r} \varphi^{i,j}_{h,r} \geq \binom{t_1}{2}.$$

Note that $\binom{t_1}{2} + 1 \geq t_1$ because $t_1 \geq 1$. Therefore, (5.4) follows from

$$\varphi_{h,r}, \sum_{i,j:i<j,h(i)=h(j)=r} \varphi^{i,j}_{h,r} \geq \binom{t_1}{2} \;\rightarrow\; \varphi_{h,r} + \sum_{i,j:i<j,h(i)=h(j)=r} \varphi^{i,j}_{h,r} \geq \binom{t_1}{2} + 1.$$

The proof now concludes with an induction on the number of pairs h, r. Let

$$Y_{h,r} = \sum_{i:h(i)=r} \varphi^i_{h,r}$$

and

$$Z_{h,r} = \varphi_{h,r} + \sum_{i,j:i<j,h(i)=h(j)=r} \varphi^{i,j}_{h,r}.$$

Then (5.4) says that $Y_{h,r} \geq t_1 \rightarrow Z_{h,r} \geq t_1$, and we want to show that

$$\sum_{h,r} Y_{h,r} \geq t \;\rightarrow\; \sum_{h,r} Z_{h,r} \geq t.$$

Order the pairs in $\mathcal{H}_k \times [2^k]$ in some arbitrary way. We will show, by induction, that for every pair h', r' and for every number t_2, the sequent

$$\sum_{h,r \leq h',r'} Y_{h,r} \geq t_2 \;\rightarrow\; \sum_{h,r \leq h',r'} Z_{h,r} \geq t_2 \tag{5.5}$$

has a flat proof. The inductive basis, i.e., the case when h', r' is the first pair in $\mathcal{H}_k \times [2^k]$, is immediate from (5.4).

For the inductive step, let h', r' be one of the pairs and suppose that (5.5) holds for $h, r < h', r'$. We break the proof of (5.5) into cases according to the value of $Y_{h',r'}$. We will show, for every $t_1 = 1, \ldots, t-1$, that

$$Y_{h',r'} \geq t_1, \sum_{h,r \leq h',r'} Y_{h,r} \geq t \rightarrow Y_{h',r'} \geq t_1 + 1, \sum_{h,r \leq h',r'} Z_{h,r} \geq t. \tag{5.6}$$

This is equivalent in meaning to $(Y_{h',r'} = t_1) \Rightarrow$ (5.5). Then, since

$$\sum_{h,r \leq h',r'} Y_{h,r} \geq t \rightarrow Y_{h',r'} \geq 1, \sum_{h,r < h',r'} Y_{h,r} \geq t$$

and

$$Y_{h',r'} \geq t \rightarrow Z_{h',r'} \geq t,$$

we get that

$$\sum_{h,r \leq h',r'} Y_{h,r} \geq t \rightarrow Y_{h',r'} \geq 1, \sum_{h,r \leq h',r'} Z_{h,r} \geq t$$

and

$$Y_{h',r'} \geq t \rightarrow \sum_{h,r \leq h',r'} Z_{h,r} \geq t,$$

by using the inductive hypothesis and simple facts about threshold connectives. These two sequents correspond to the cases $Y_{h',r'} \leq 0$ and $Y_{h',r'} \geq t$. Therefore, by combining with (5.6), we get a proof of (5.5).

To complete the inductive step, there only remains to prove (5.6). The idea is simple. Suppose that $Y_{h',r'} = t_1$. Then $Z_{h',r'} \geq t_1$, by (5.4). On the other hand, $\sum_{h,r<h',r'} Y_{h,r} \geq t - t_1$, which implies that $\sum_{h,r<h',r'} Z_{h,r} \geq t - t_1$, by the inductive hypothesis. Combining the two we get $\sum_{h,r \leq h',r'} Z_{h,r} \geq t$. In sequent calculus, we have that

$$Y_{h',r'} \geq t_1 \rightarrow Z_{h',r'} \geq t_1$$

and that

$$\sum_{h,r \leq h',r'} Y_{h,r} \geq t \rightarrow Y_{h',r'} \geq t_1 + 1, \sum_{h,r < h',r'} Y_{h,r} \geq t - t_1.$$

The result follows by the inductive hypothesis and since

$$Z_{h',r'} \geq t_1, \sum_{h,r<h',r'} Z_{h,r} \geq t - t_1 \rightarrow \sum_{h,r \leq h',r'} Z_{h,r} \geq t.$$

□

We now show that Proposition 5.1 has a flat proof.

Lemma 5.4. *The sequent*

$$\sum_i u_i \geq \frac{1}{8} 2^k \rightarrow \sum_i u_i \geq \frac{1}{4} 2^k, \sum_{h,r \in \mathcal{H}_k \times [2^k]} \varphi_{h,r} \geq \frac{N_k}{16},$$

where $N_k = |\mathcal{H}_k| 2^k$, *has a flat proof of size* $2^{(\log n)^{O(d)}}$.

PROOF. Note that the threshold $N_k/16$ is an integer since $k \geq 3$ and $|\mathcal{H}_k|$ is a power of 2 no smaller than 4.

First, we show that for every i,

$$u_i \;\rightarrow\; \sum_j u_j \geq \frac{1}{4}2^k,\; \sum_{h,r:h(i)=r} \varphi^i_{h,r} \geq \frac{3}{4}|\mathcal{H}_k|. \tag{5.7}$$

Since, for every h, r such that $h(i) = r$,

$$u_i,\; \left(1 \oplus \bigoplus_{j:j\neq i,h(j)=r} u_j\right) \;\rightarrow\; \varphi^i_{h,r}$$

we have that

$$u_i,\; \sum_{h,r:h(i)=r}\left(1 \oplus \bigoplus_{j:j\neq i,h(j)=r} u_j\right) \geq \frac{3}{4}|\mathcal{H}_k| \;\rightarrow\; \sum_{h,r:h(i)=r} \varphi^i_{h,r} \geq \frac{3}{4}|\mathcal{H}_k|.$$

Therefore, it is sufficient to show that

$$\rightarrow\; \sum_j u_j \geq \frac{1}{4}2^k,\; \sum_{h,r:h(i)=r}\left(1 \oplus \bigoplus_{j:j\neq i,h(j)=r} u_j\right) \geq \frac{3}{4}|\mathcal{H}_k|.$$

Let $E^i_{h,r} = \bigoplus_{j:j\neq i,h(j)=r} u_j$. For every h, r,

$$\rightarrow\; E^i_{h,r},\; 1 \oplus E^i_{h,r}.$$

Therefore,

$$\rightarrow\; \sum_{h,r:h(i)=r} E^i_{h,r} \geq \frac{1}{4}|\mathcal{H}_k|,\; \sum_{h,r:h(i)=r} (1 \oplus E^i_{h,r}) \geq \frac{3}{4}|\mathcal{H}_k|.$$

Since $E^i_{h,r} \rightarrow \sum_{j:j\neq i,h(j)=r} u_j \geq 1$, we have that

$$\sum_{h,r:h(i)=r} E^i_{h,r} \geq \frac{1}{4}|\mathcal{H}_k| \;\rightarrow\; \sum_{h,r:h(i)=r}\left(\sum_{j:j\neq i,h(j)=r} u_j\right) \geq \frac{1}{4}|\mathcal{H}_k|.$$

The formula on the right hand side of this sequent is the same as

$$\sum_{j:j\neq i}\left(\sum_{h,r:h(i)=h(j)=r} u_j\right) \geq \frac{1}{4}|\mathcal{H}_k|.$$

By the universal property of the family $\mathcal{H}_k$ of hash functions, the value of the inside sum is either 0 or $|\mathcal{H}_k|/2^k$. From this it follows that

$$\sum_{j:j\neq i}\left(\sum_{h,r:h(i)=h(j)=r} u_j\right) \geq \frac{1}{4}|\mathcal{H}_k| \;\rightarrow\; \sum_{j:j\neq i} u_j \geq \frac{1}{4}2^k.$$

Therefore,

$$\rightarrow\; \sum_j u_j \geq \frac{1}{4}2^k,\; \sum_{h,r:h(i)=r} (1 \oplus E^i_{h,r}) \geq \frac{3}{4}|\mathcal{H}_k|,$$

which completes the proof of (5.7).

From (5.7), we get that

$$\sum_i u_i \geq \frac{1}{8}2^k \;\rightarrow\; \sum_i u_i \geq \frac{1}{4}2^k,\; \sum_i \left(\sum_{h,r:h(i)=r} \varphi^i_{h,r} \right) \geq \frac{3}{32}N_k.$$

By the principle of inclusion-exclusion (Lemma 5.3), this implies that

$$\sum_i u_i \geq \frac{1}{8}2^k \;\rightarrow\; \sum_i u_i \geq \frac{1}{4}2^k,\; \sum_{h,r} \varphi_{h,r} + \sum_{i<j} \left(\sum_{h,r:h(i)=h(j)=r} \varphi^{i,j}_{h,r} \right) \geq \frac{3}{32}N_k.$$

Now,

$$\begin{aligned} &\sum_{h,r} \varphi_{h,r} + \sum_{i<j} \left(\sum_{h,r:h(i)=h(j)=r} \varphi^{i,j}_{h,r} \right) \geq \frac{3}{32}N_k \\ &\quad \rightarrow\; \sum_{h,r} \varphi_{h,r} \geq \frac{N_k}{16},\; \sum_{i<j} \left(\sum_{h,r:h(i)=h(j)=r} \varphi^{i,j}_{h,r} \right) \geq \frac{N_k}{32}. \end{aligned}$$

Since $\varphi^{i,j}_{h,r} \rightarrow u_i u_j$,

$$\sum_{i<j} \left(\sum_{h,r:h(i)=h(j)=r} \varphi^{i,j}_{h,r} \right) \geq \frac{N_k}{32} \;\rightarrow\; \sum_{i<j} \left(\sum_{h,r:h(i)=h(j)=r} u_i u_j \right) \geq \frac{N_k}{32}.$$

Again by the universal property of $\mathcal{H}_k$, the value of the inside sum on the right hand side is at most $|\mathcal{H}_k|/2^k$. This implies that

$$\sum_{i<j} \left(\sum_{h,r:h(i)=h(j)=r} u_i u_j \right) \geq \frac{N_k}{32} \;\rightarrow\; \sum_{i<j} u_i u_j \geq \frac{2^{2k}}{32}.$$

It is not difficult to show that for every t,

$$\sum_{i<j} u_i u_j \geq \frac{t^2}{2} \;\rightarrow\; \sum_i u_i \geq t.$$

Therefore,

$$\sum_{i<j} u_i u_j \geq \frac{2^{2k}}{32} \;\rightarrow\; \sum_i u_i \geq \frac{2^k}{4}$$

which completes the proof of the lemma. □

Finally, we arrive at the key property of $Q^m_\vee$.

Lemma 5.5. *Let $u = (u_1, \ldots, u_m)$ be a sequence of polynomials of degree $(\log n)^{O(d)}$ over $\mathbf{Z}_2$ in the propositional variables $x_1, \ldots, x_n$. The sequents*

$$\sum_\tau Q^m_\vee(u,\tau) \geq 1 \;\rightarrow\; \sum_i u_i \geq 1$$

and

$$\sum_i u_i \geq 1 \;\rightarrow\; \sum_\tau Q^m_\vee(u,\tau) \geq (1-\varepsilon)2^{|w|},$$

where τ ranges over all possible values of the probabilistic variables w of $Q^m_\vee$, have flat proofs of size $2^{(\log n)^{O(d)}}$.

Proof. First, from the sequents of Lemma 5.2, the following basic facts about polynomials over $\mathbf{Z}_2$ can be shown:

$$1 \oplus (1 \oplus P)(1 \oplus Q) \;\rightarrow\; P,\, Q \tag{5.8}$$

$$\sum_{i=1}^{M} P_i \geq K \;\rightarrow\; \sum_{i=1}^{M}\sum_{j=1}^{N}(1 \oplus (1 \oplus P_i)Q_j) \geq KN \tag{5.9}$$

$$\begin{aligned} &\textstyle\sum_{i=1}^{M}(1 \oplus P_i) \geq M - K,\; \sum_{i=1}^{N}(1 \oplus Q_i) \geq N - L \\ &\quad\rightarrow\; \textstyle\sum_{i=1}^{M}\sum_{j=1}^{N}(1 \oplus P_i Q_j) \geq MN - KL \end{aligned} \tag{5.10}$$

where P, Q, the P_i's and the Q_j's are polynomials over $\mathbf{Z}_2$ and K and L are numbers.

Consider the first sequent in the statement of the lemma. To prove this sequent, it is sufficient to show that for every τ, $Q^m_\vee(u,\tau) \rightarrow \bigvee_i u_i$. Consider an arbitrary τ. As was done earlier, let $R_{l,k}$ denote the polynomial $\bigoplus_{i:h_{l,k}(i)=r_{l,k}} u_i$. Then

$$Q^m_\vee(u,\tau) = 1 \oplus \prod_{l=1}^{C}\prod_{k=3}^{L(m)}(1 \oplus R_{l,k}),$$

where $L(m) = \lceil \log m \rceil + 3$. By repeated applications of (5.8), we get that

$$Q^m_\vee(u,\tau) \;\rightarrow\; \sum_{l,k} R_{l,k} \geq 1.$$

The first sequent follows since, for every pair l,k, $R_{l,k} \rightarrow \sum_i u_i \geq 1$.

Consider now the second sequent. By Lemma 5.4, we have that for every l,k,

$$\sum_i u_i \geq \frac{1}{8}2^k \;\rightarrow\; \sum_i u_i \geq \frac{1}{4}2^k,\; \sum_{h_{l,k},r_{l,k}} R_{l,k} \geq \frac{|\mathcal{H}_k|2^k}{16},$$

where the last sum ranges over all possible values of $h_{l,k}, r_{l,k}$. From (5.9), we get that for every l,k,

$$\begin{aligned} &\sum_{h_{l,k},r_{l,k}} R_{l,k} \geq \frac{|\mathcal{H}_k|2^k}{16} \\ &\rightarrow \sum_{(h_{l,k'},r_{l,k'})_{3\leq k'\leq L(m)}} \left(1 \oplus (1 \oplus R_{l,k})\prod_{k''\neq k}(1 \oplus R_{l,k''})\right) \geq \frac{\prod_{k'=3}^{L(m)}|\mathcal{H}_{k'}|2^{k'}}{16}. \end{aligned}$$

Therefore, for every l,

$$\sum_i u_i \geq 1 \;\rightarrow \sum_{(h_{l,k},r_{l,k})_{3\leq k\leq L(m)}} \left(1 \oplus \prod_{k'=3}^{L(m)}(1 \oplus R_{l,k'})\right) \geq \frac{\prod_{k=3}^{L(m)}|\mathcal{H}_k|2^k}{16}.$$

Then, by repeated applications of (5.10), we get that

$$\sum_i u_i \geq 1 \;\rightarrow$$

$$\sum_{(h_{l,k},r_{l,k})_{1\leq l\leq C,\ 3\leq k\leq L(m)}} \left(1 \oplus \prod_{l=1}^{C} \prod_{k'=3}^{L(m)} (1 \oplus R_{l,k'})\right) \geq (1-\varepsilon)\left(\prod_{k=3}^{L(m)} |\mathcal{H}_k| 2^k\right)^C .$$

The result follows the right hand side is precisely $\sum_\tau Q_\vee^m(u,\tau) \geq (1-\varepsilon)2^{|w|}$. □

The key property of $Q_\wedge^m$ is easily derived from the key property of $Q_\vee^m$.

LEMMA 5.6. *Let $u = (u_1, \ldots, u_m)$ be a sequence of polynomials of degree $(\log n)^{O(d)}$ over $\mathbf{Z}_2$ in the propositional variables $x_1, \ldots, x_n$. The sequents*

$$\sum_i u_i \geq m \;\rightarrow\; \sum_\tau Q_\wedge^m(u,\tau) \geq 2^{|w|}$$

and

$$\sum_\tau Q_\wedge^m(u,\tau) \geq \varepsilon 2^{|w|} + 1 \;\rightarrow\; \sum_i u_i \geq m,$$

where τ ranges over all possible values of the probabilistic variables w of $Q_\wedge^m$, have flat proofs of size $2^{(\log n)^{O(d)}}$.

PROOF. We prove the second sequent. The proof of the first one is similar. By Lemma 5.5, we have that

$$\sum_i (1 \oplus u_i) \geq 1 \;\rightarrow\; \sum_\tau Q_\vee^m(1 \oplus u,\tau) \geq (1-\varepsilon)2^{|w|},$$

where $1 \oplus u$ denotes $1 \oplus u_1, \ldots, 1 \oplus u_m$. For every i, we have that $\rightarrow u_i, 1 \oplus u_i$. This implies that

$$\rightarrow \sum_i u_i \geq m,\ \sum_i (1 \oplus u_i) \geq 1.$$

On the other hand, since, for every τ, $Q_\wedge^m(u,\tau) = 1 \oplus Q_\vee^m(1 \oplus u,\tau)$, we have that $Q_\wedge^m(u,\tau), Q_\vee^m(1 \oplus u,\tau) \rightarrow$ and that

$$\sum_\tau Q_\wedge^m(u,\tau) \geq \varepsilon 2^{|w|} + 1,\ \sum_\tau Q_\vee^m(1 \oplus u,\tau) \geq (1-\varepsilon)2^{|w|} \;\rightarrow .$$

The result follows. □

6. Properties of P_A and $\mathrm{tr}(A)$

In this section, we show that several properties of the probabilistic polynomials P_A and of the formulas $\mathrm{tr}(A)$ have flat proofs in our logic system. This will then be used in the next section to show that the rules of inference used in $\mathrm{ACC}^0[2]$ proofs can be simulated by flat proofs.

These properties all have very easy proofs if one is not concerned with the complexity of these proofs. However, some care must be taken in order to obtain flat proofs. For example, the property

$$\sum_\sigma P_A(\sigma) \geq \varepsilon s_A 2^{|y|} + 1 \;\rightarrow\; \mathrm{tr}(A),$$

can be proved from the fact that P_A computes A with error ε, a fact that is easily established by using the key property of $Q^m_\vee$ (Lemma 5.5). However, such a proof mentions explicitly the formula A and this, of course, has to be avoided. We will therefore have to prove this and the other properties in a more direct way.

In the following lemmas, by a small $\mathrm{ACC}^0[2]$ formula, we mean a formula over the connectives $\neg$, $\wedge$, $\vee$ and $\oplus$, of size at most $1/(4\varepsilon)$. Recall that in Section 3 we fixed an arbitrary $\mathrm{ACC}^0[2]$ proof of depth d and size s and that in Section 4 we defined ε to be $1/2^{\lceil 2+\log s\rceil}$. We then defined the polynomials P_A and the formulas $\mathrm{tr}(A)$ in terms of ε.

LEMMA 6.1. *For every small* $\mathrm{ACC}^0[2]$ *formula A and for every small* $\mathrm{ACC}^0[2]$ *formula of the form* $\oplus_1(A_1, \dots, A_m)$*, the sequents* $\mathrm{tr}(A), \mathrm{tr}(\neg A) \rightarrow$ *and*

$$\mathrm{tr}(\oplus_1(A_1, \dots, A_m)),\ \mathrm{tr}(\oplus_0(A_1, \dots, A_m))\ \rightarrow$$

have flat proofs of size $2^{(\log n)^{O(d)}}$.

PROOF. Suppose that A is a depth-d formula. Associated with A and $\neg A$, we have probabilistic polynomials $P_A(y_1, \dots, y_d)$ and $P_{\neg A}(y_1, \dots, y_{d+1})$. (We are omitting the sequence x of propositional variables.) Let y denote $y_1, \dots, y_d$ and let z denote y_{d+1}, so that we may write $P_A(y)$ for $P_A(y_1, \dots, y_d)$ and $P_{\neg A}(y, z)$ for $P_{\neg A}(y_1, \dots, y_{d+1})$. By definition, the sequent $\mathrm{tr}(A), \mathrm{tr}(\neg A) \rightarrow$ says that

$$\sum_\sigma P_A(\sigma) \geq (1 - \varepsilon s_A)2^{|y|},\ \sum_{\sigma,\tau} P_{\neg A}(\sigma, \tau) \geq (1 - \varepsilon s_{\neg A})2^{|y|+|z|}\ \rightarrow .$$

Recall that $P_{\neg A}(\sigma, \tau) = P_A(\sigma) \oplus 1$. For every σ, we have that $P_A(\sigma), P_A(\sigma) \oplus 1 \rightarrow$. Therefore,

$$\mathrm{tr}(A),\ \sum_\sigma (P_A(\sigma) \oplus 1) \geq \varepsilon s_A 2^{|y|} + 1\ \rightarrow .$$

Since $\sum_\tau (P_A(\sigma) \oplus 1)$ is always at most $2^{|z|}$, and since $\sum_\tau (P_A(\sigma) \oplus 1) \geq 1 \rightarrow P_A(\sigma) \oplus 1$, we have that

$$\sum_{\sigma,\tau} (P_A(\sigma) \oplus 1) \geq \varepsilon s_A 2^{|y|+|z|} + 1\ \rightarrow\ \sum_\sigma (P_A(\sigma) \oplus 1) \geq \varepsilon s_A 2^{|y|} + 1.$$

Therefore,

$$\mathrm{tr}(A),\ \sum_{\sigma,\tau} P_{\neg A}(\sigma, \tau) \geq \varepsilon s_A 2^{|y|+|z|} + 1\ \rightarrow .$$

The result follows since $1 - \varepsilon s_{\neg A} \geq \varepsilon s_A$.

The second sequent is proved similarly. □

LEMMA 6.2. *For every small* $\mathrm{ACC}^0[2]$ *formula A of the form* $\vee(A_1, \dots, A_m)$ *and for every* $i \in [m]$*, the sequents* $\mathrm{tr}(A_i) \rightarrow \mathrm{tr}(A)$ *and* $\mathrm{tr}(\bigvee_{j \neq i} A_j) \rightarrow \mathrm{tr}(A)$ *have flat proofs of size* $2^{(\log n)^{O(d)}}$.

PROOF. For the first sequent, we have to show that

$$\sum_\sigma P_{A_i}(\sigma) \geq (1 - \varepsilon s_{A_i})2^{|y|}\ \rightarrow\ \sum_\sigma \left(\sum_\tau P_A(\sigma, \tau) \right) \geq (1 - \varepsilon s_A)2^{|y|+|z|}.$$

Consider an arbitrary σ. Since $P_A(\sigma,z) = Q_\vee^m(P_{A_1}(\sigma),\dots,P_{A_m}(\sigma),z)$, from Lemma 5.5 we get that

$$P_{A_i}(\sigma) \;\rightarrow\; \sum_\tau P_A(\sigma,\tau) \geq (1-\varepsilon)2^{|z|}.$$

Therefore,

$$\sum_\sigma P_{A_i}(\sigma) \geq (1-\varepsilon s_{A_i})2^{|y|} \;\rightarrow\; \sum_\sigma \left(\sum_\tau P_A(\sigma,\tau)\right) \geq (1-\varepsilon s_{A_i})2^{|y|}(1-\varepsilon)2^{|z|},$$

which implies that $\mathrm{tr}(A_i) \rightarrow \mathrm{tr}(A)$, since $(1-\varepsilon s_{A_i})(1-\varepsilon) \geq 1-\varepsilon s_A$.

Now let $B = \bigvee_{j\neq i} A_j$. Then $P_B(y,z) = Q_\vee^{m-1}((P_{A_j}(y))_{j\neq i},z)$. We want to show that

$$\sum_\sigma \left(\sum_\tau P_B(\sigma,\tau)\right) \geq (1-\varepsilon s_B)2^{|y|+|z|} \;\rightarrow\; \sum_\sigma \left(\sum_\tau P_A(\sigma,\tau)\right) \geq (1-\varepsilon s_A)2^{|y|+|z|}.$$

Consider an arbitrary σ. By Lemma 5.5, we have that

$$\sum_\tau P_B(\sigma,\tau) \geq 1 \;\rightarrow\; \sum_{j\in S} P_{A_j}(\sigma) \geq 1$$

and that

$$\sum_j P_{A_j}(\sigma) \geq 1 \;\rightarrow\; \sum_\tau P_A(\sigma,\tau) \geq (1-\varepsilon)2^{|z|}.$$

Therefore,

$$\sum_\tau P_B(\sigma,\tau) \geq 1 \;\rightarrow\; \sum_\tau P_A(\sigma,\tau) \geq (1-\varepsilon)2^{|z|}.$$

Since the maximum value of $\sum_\tau P_B(\sigma,\tau)$ is $2^{|z|}$, it can be shown that

$$\begin{aligned}&\sum_\sigma \left(\sum_\tau P_B(\sigma,\tau)\right) \geq (1-\varepsilon s_B)2^{|y|+|z|}\\ &\quad\rightarrow\; \sum_\sigma \left(\sum_\tau P_A(\sigma,\tau)\right) \geq (1-\varepsilon s_B)2^{|y|}(1-\varepsilon)2^{|z|}.\end{aligned}$$

which implies that $\mathrm{tr}(B) \rightarrow \mathrm{tr}(A)$, since $(1-\varepsilon s_B)(1-\varepsilon) \geq 1-\varepsilon s_A$. ☐

LEMMA 6.3. *For every small* $\mathrm{ACC}^0[2]$ *formula of the form* $\oplus_i(A_1,\dots,A_m)$, *the sequent*

$$\mathrm{tr}(A_1), \mathrm{tr}(\oplus_{i-1}(A_2,\dots,A_m)) \rightarrow \mathrm{tr}(\oplus_i(A_1,\dots,A_m))$$

has a flat proof of size $2^{(\log n)^{O(d)}}$.

PROOF. Let $A = \oplus_i(A_1,\dots,A_m)$ and $B = \oplus_{i-1}(A_2,\dots,A_m)$. It is easy to see that for every pair σ,τ, $P_{A_1}(\sigma)$, $P_B(\sigma,\tau) \rightarrow P_A(\sigma,\tau)$. From this, we get that

$$\begin{aligned}&\sum_{\sigma,\tau} P_{A_1}(\sigma) \geq (1-\varepsilon s_{A_1})2^{|y|+|z|},\ \sum_{\sigma,\tau} P_B(\sigma,\tau) \geq (1-\varepsilon s_B)2^{|y|+|z|}\\ &\quad\rightarrow\; \sum_{\sigma,\tau} P_A(\sigma,\tau) \geq (1-\varepsilon s_{A_1}-\varepsilon s_B)2^{|y|+|z|}.\end{aligned}$$

The result follows since $s_{A_1} + s_B \le s_A$ and

$$\sum_\sigma P_{A_1}(\sigma) \ge (1 - \varepsilon s_{A_1})2^{|y|} \;\rightarrow\; \sum_{\sigma,\tau} P_{A_1}(\sigma) \ge (1 - \varepsilon s_{A_1})2^{|y|+|z|}.$$

□

LEMMA 6.4. *For every small* $\mathrm{ACC}^0[2]$ *formula* B *of depth* d*, the sequent*

$$\sum_{\sigma_1,\dots,\sigma_d} P_B(\sigma_1,\dots,\sigma_d) \ge \varepsilon s_B 2^{|y_1|+\cdots+|y_d|} + 1 \;\rightarrow\; \mathrm{tr}(B)$$

has a flat proof of size $2^{(\log n)^{O(d)}}$.

PROOF. We use induction on the *generalized* structure of B. Call A a *generalized* subformula of B if $A = *(B_i,\dots,B_r)$ where $*(B_1,\dots,B_r)$ is a subformula of B and $1 \le i \le r$. We show that for every generalized subformula A of B, the sequent

$$\sum_{\sigma_1,\dots,\sigma_d} P_A(\sigma_1,\dots,\sigma_d) \ge \varepsilon s_A 2^{|y_1|+\cdots+|y_d|} + 1 \;\rightarrow\; \mathrm{tr}(A),$$

where d is the depth of A, has a flat proof in our logic system.

Basis. Suppose that $A = *(u_1,\dots,u_m)$, where all the u_i's are propositional variables.

Case $A = \vee(u_1,\dots,u_m)$. Then $P_A(y_1) = Q^m_\vee(u_1,\dots,u_m,y_1)$. The result follows directly from Lemma 5.5.

Case $A = \oplus_1(u_1,\dots,u_m)$. Here P_A is deterministic, i.e., P_A does not depend on y_1. Therefore, $\sum_{\sigma_1} P_A \ge 1 \rightarrow P_A$ and $P_A \rightarrow \sum_{\sigma_1} P_A \ge 2^{|y_1|}$. The result follows easily.

The other cases are similar.

Inductive step. Now assume that $A = *(A_1,\dots,A_m)$, where some of the A_i's are not merely propositional variables. Let y denote $y_1,\dots,y_{d-1}$ and let z denote y_d. Then the sequent we must prove is

$$\sum_{\sigma,\tau} P_A(\sigma,\tau) \ge \varepsilon s_A 2^{|y|+|z|} + 1 \;\rightarrow\; \mathrm{tr}(A).$$

Case $A = \neg A_1$. Recall that in this case $P_A(y,z) = P_{A_1}(y) \oplus 1$. For every σ,τ, we have that $\rightarrow P_A(\sigma,\tau), P_{A_1}(\sigma)$. Therefore,

$$\rightarrow\; \mathrm{tr}(A),\; \sum_\sigma \left(\sum_\tau P_{A_1}(\sigma) \right) \ge \varepsilon s_A 2^{|y|+|z|} + 1$$

so that

$$\rightarrow\; \mathrm{tr}(A),\; \sum_\sigma P_{A_1}(\sigma) \ge \varepsilon s_A 2^{|y|} + 1,$$

since the inside sum is at most $2^{|z|}$. Then, by the inductive hypothesis, we get that $\rightarrow \mathrm{tr}(A), \mathrm{tr}(A_1)$. On the other hand,

$$\sum_\sigma (P_{A_1}(\sigma) \oplus 1) \ge \varepsilon s_{A_1} 2^{|y|} + 1,\; \mathrm{tr}(A_1) \;\rightarrow$$

so that

$$\sum_\sigma \left(\sum_\tau P_A(\sigma,\tau) \right) \ge \varepsilon s_A 2^{|y|+|z|} + 1,\; \mathrm{tr}(A_1) \;\rightarrow,$$

since, once again, the inside sum is at most $2^{|z|}$. The result follows.

Case $A = \vee(A_1, \ldots, A_m)$. For every σ, by Lemma 5.5, we have that

$$\sum_{\tau} P_A(\sigma,\tau) \geq 1 \;\rightarrow\; \sum_i P_{A_i}(\sigma) \geq 1.$$

Therefore,

$$\sum_{\sigma}\left(\sum_{\tau} P_A(\sigma,\tau)\right) \geq \varepsilon s_A 2^{|y|+|z|}+1 \;\rightarrow\; \sum_{\sigma}\left(\sum_i P_{A_i}(\sigma)\right) \geq \varepsilon s_A 2^{|y|}+1,$$

since the value of the inside sum on the left hand side is at most $2^{|z|}$. The formula on the right hand side is the same as

$$\sum_i\left(\sum_{\sigma} P_{A_i}(\sigma)\right) \geq \varepsilon s_A 2^{|y|}+1.$$

For every i, the inductive hypothesis says that

$$\sum_{\sigma} P_{A_i}(\sigma) \geq \varepsilon s_{A_i} 2^{|y|}+1 \;\rightarrow\; \mathrm{tr}(A_i).$$

Therefore, since $\mathrm{tr}(A_i) \rightarrow \mathrm{tr}(A)$ by Lemma 6.2,

$$\sum_{\sigma} P_{A_i}(\sigma) \geq \varepsilon s_{A_i} 2^{|y|}+1 \;\rightarrow\; \mathrm{tr}(A).$$

From this, it is easy to show that

$$\sum_i\left(\sum_{\sigma} P_{A_i}(\sigma)\right) \geq \varepsilon s_A 2^{|y|}+1 \;\rightarrow\; \mathrm{tr}(A).$$

The result follows.

Case $A = \wedge(A_1, \ldots, A_m)$. Recall that $P_A(y,z) = 1 \oplus Q^m_{\vee}(1 \oplus P_{A_1}(y), \ldots, 1 \oplus P_{A_m}(y), z)$. For every σ, τ, we have both $P_A(\sigma,\tau), 1 \oplus P_A(\sigma,\tau) \;\rightarrow$ and its inverse. Therefore,

$$\sum_{\sigma,\tau} P_A(\sigma,\tau) \geq \varepsilon s_A 2^{|y|+|z|}+1,\; \sum_{\sigma,\tau}(1 \oplus P_A(\sigma,\tau)) \geq (1-\varepsilon s_A)2^{|y|+|z|} \;\rightarrow$$

and

$$\rightarrow\; \sum_{\sigma,\tau}(1 \oplus P_A(\sigma,\tau)) \geq \varepsilon s_A 2^{|y|+|z|}+1,\; \mathrm{tr}(A).$$

This implies that it is sufficient to show that

$$\sum_{\sigma,\tau}(1 \oplus P_A(\sigma,\tau)) \geq \varepsilon s_A 2^{|y|+|z|}+1 \;\rightarrow\; \sum_{\sigma,\tau}(1 \oplus P_A(\sigma,\tau)) \geq (1-\varepsilon s_A)2^{|y|+|z|}.$$

By an argument similar to the one used in the previous case ($A = \vee(A_1, \ldots, A_m)$), it can be shown that

$$\sum_{\sigma,\tau}(1 \oplus P_A(\sigma,\tau)) \geq \varepsilon s_A 2^{|y|+|z|}+1 \;\rightarrow\; \sum_i\left(\sum_{\sigma}(1 \oplus P_{A_i}(\sigma))\right) \geq \varepsilon s_A 2^{|y|}+1.$$

For every i, by the inductive hypothesis, we have that

$$\sum_{\sigma} P_{A_i}(\sigma) \geq \varepsilon s_{A_i} 2^{|y|}+1 \;\rightarrow\; \sum_{\sigma} P_{A_i}(\sigma) \geq (1-\varepsilon s_{A_i})2^{|y|}.$$

Therefore,

$$\sum_{\sigma}(1 \oplus P_{A_i}(\sigma)) \geq \varepsilon s_{A_i} 2^{|y|} + 1 \;\rightarrow\; \sum_{\sigma}(1 \oplus P_{A_i}(\sigma)) \geq (1 - \varepsilon s_{A_i}) 2^{|y|}.$$

As in the proof of the sequent $\operatorname{tr}(A_i) \rightarrow \operatorname{tr}(A)$ in Lemma 6.2, it can be shown that

$$\sum_{\sigma}(1 \oplus P_{A_i}(\sigma)) \geq (1 - \varepsilon s_{A_i}) 2^{|y|} \;\rightarrow\; \sum_{\sigma,\tau}(1 \oplus P_{A}(\sigma,\tau)) \geq (1 - \varepsilon s_{A}) 2^{|y|+|z|}.$$

Therefore,

$$\sum_{\sigma}(1 \oplus P_{A_i}(\sigma)) \geq \varepsilon s_{A_i} 2^{|y|} + 1 \;\rightarrow\; \sum_{\sigma,\tau}(1 \oplus P_{A}(\sigma,\tau)) \geq (1 - \varepsilon s_{A}) 2^{|y|+|z|}.$$

From this, we get that

$$\sum_{i}\left(\sum_{\sigma}(1 \oplus P_{A_i}(\sigma))\right) \geq \varepsilon s_{A} 2^{|y|} + 1 \;\rightarrow\; \sum_{\sigma,\tau}(1 \oplus P_{A}(\sigma,\tau)) \geq (1 - \varepsilon s_{A}) 2^{|y|+|z|}.$$

The result follows.

Case $A = \oplus_1(A_1, \ldots, A_m)$. Here, we have that $P_A(y, z) = \bigoplus_{i=1}^{m} P_{A_i}(y)$. Let $B = \oplus_1(A_2, \ldots, A_m)$ and $B' = \oplus_0(A_2, \ldots, A_m)$. For every pair σ, τ, we have that $P_A(\sigma,\tau) \rightarrow P_{A_1}(\sigma), P_B(\sigma,\tau)$ and that $P_A(\sigma,\tau) \rightarrow P_{A_1}(\sigma) \oplus 1, P_{B'}(\sigma,\tau)$.

First consider $P_A(\sigma,\tau) \rightarrow P_{A_1}(\sigma) \oplus 1, P_{B'}(\sigma,\tau)$. This implies that

$$\begin{aligned}&\sum_{\sigma,\tau} P_A(\sigma,\tau) \geq \varepsilon s_A 2^{|y|+|z|} + 1\\ &\quad\rightarrow\; \sum_{\sigma,\tau}(P_{A_1}(\sigma) \oplus 1) \geq \varepsilon s_{A_1} 2^{|y|+|z|} + 1,\; \sum_{\sigma,\tau} P_{B'}(\sigma,\tau) \geq \varepsilon s_{B'} 2^{|y|+|z|} + 1.\end{aligned}$$

Therefore, since $\operatorname{tr}(A_1), \sum_{\sigma}(P_{A_1}(\sigma) \oplus 1) \geq \varepsilon s_{A_1} 2^{|y|} + 1 \rightarrow$, and by the inductive hypothesis,

$$\sum_{\sigma,\tau} P_A(\sigma,\tau) \geq \varepsilon s_A 2^{|y|+|z|} + 1,\; \operatorname{tr}(A_1) \;\rightarrow\; \operatorname{tr}(B').$$

By Lemma 6.3, $\operatorname{tr}(A_1), \operatorname{tr}(B') \rightarrow \operatorname{tr}(A)$. Therefore,

$$\sum_{\sigma,\tau} P_A(\sigma,\tau) \geq \varepsilon s_A 2^{|y|+|z|} + 1,\; \operatorname{tr}(A_1) \;\rightarrow\; \operatorname{tr}(A).$$

Now consider $P_A(\sigma,\tau) \rightarrow P_{A_1}(\sigma), P_B(\sigma,\tau)$. This implies that

$$\begin{aligned}&\sum_{\sigma,\tau} P_A(\sigma,\tau) \geq \varepsilon s_A 2^{|y|+|z|} + 1\\ &\quad\rightarrow\; \sum_{\sigma,\tau} P_{A_1}(\sigma) \geq \varepsilon s_{A_1} 2^{|y|+|z|} + 1,\; \sum_{\sigma,\tau} P_{B}(\sigma,\tau) \geq \varepsilon s_{B} 2^{|y|+|z|} + 1.\end{aligned}$$

Therefore, by the inductive hypothesis,

$$\sum_{\sigma,\tau} P_A(\sigma,\tau) \geq \varepsilon s_A 2^{|y|+|z|} + 1 \;\rightarrow\; \operatorname{tr}(A_1),\; \operatorname{tr}(B).$$

By an argument similar to the one used in the proof of Lemma 6.3, and by using the inductive hypothesis on A_1, it can be shown that $\operatorname{tr}(B) \rightarrow \operatorname{tr}(A_1), \operatorname{tr}(A)$. Therefore,

$$\sum_{\sigma,\tau} P_A(\sigma,\tau) \geq \varepsilon s_A 2^{|y|+|z|} + 1 \;\rightarrow\; \operatorname{tr}(A_1),\; \operatorname{tr}(A).$$

The result follows.

Case $A = \oplus_0(A_1, \dots, A_m)$. Similar to the previous case. $\square$

LEMMA 6.5. *For every small* $\mathrm{ACC}^0[2]$ *formula* A *and every small* $\mathrm{ACC}^0[2]$ *formula of the form* $\oplus_1(A_1, \dots, A_m)$*, the sequents* $\rightarrow \mathrm{tr}(A), \mathrm{tr}(\neg A)$ *and*

$$\rightarrow\ \mathrm{tr}(\oplus_1(A_1, \dots, A_m)),\ \mathrm{tr}(\oplus_0(A_1, \dots, A_m))$$

have flat proofs of size $2^{(\log n)^{O(d)}}$.

PROOF. The proof of the first sequent is implicit in the proof of the previous lemma, case $A = \neg A_1$. The proof of the second sequent is similar. $\square$

LEMMA 6.6. *For every small* $\mathrm{ACC}^0[2]$ *formula* A *of the form* $\vee(A_1, \dots, A_m)$ *and for every* $i \in [m]$*, the sequent* $\mathrm{tr}(A) \rightarrow \mathrm{tr}(A_i)$, $\mathrm{tr}(\bigvee_{j \neq i} A_j)$ *has a flat proof of size* $2^{(\log n)^{O(d)}}$.

PROOF. Let $B = \bigvee_{j \neq i} A_j$. Consider an arbitrary σ. By Lemma 5.5,

$$\sum_\tau P_A(\sigma, \tau) \geq 1 \ \rightarrow\ \sum_j P_{A_j}(\sigma) \geq 1$$

which implies that

$$\sum_\tau P_A(\sigma, \tau) \geq 1 \ \rightarrow\ P_{A_i}(\sigma),\ \sum_{j \neq i} P_{A_j}(\sigma) \geq 1.$$

In addition, again by Lemma 5.5,

$$\sum_{j \neq i} P_{A_j}(\sigma) \geq 1 \ \rightarrow\ \sum_\tau P_B(\sigma, \tau) \geq (1 - \varepsilon) 2^{|z|}.$$

Therefore, for every σ,

$$\sum_\tau P_A(\sigma, \tau) \geq 1 \ \rightarrow\ P_{A_i}(\sigma),\ \sum_\tau P_B(\sigma, \tau) \geq (1 - \varepsilon) 2^{|z|}.$$

From this, it can be shown that

$$\begin{aligned} \sum_\sigma \left(\sum_\tau P_A(\sigma, \tau) \right) &\geq (1 - \varepsilon s_A) 2^{|y|+|z|} \\ \rightarrow\ \sum_\sigma P_{A_i}(\sigma) &\geq \frac{1}{2}(1 - \varepsilon s_A) 2^{|y|}, \\ \sum_\sigma \left(\sum_\tau P_B(\sigma, \tau) \right) &\geq \frac{1}{2}(1 - \varepsilon s_A)(1 - \varepsilon) 2^{|y|+|z|}, \end{aligned}$$

since the inside sum on the left hand side is at most $2^{|z|}$. Note that $\frac{1}{2}(1 - \varepsilon s_A)(1 - \varepsilon) \geq 1/4 \geq \max\{\varepsilon s_{A_i}, \varepsilon s_B\}$. Therefore, by Lemma 6.4, $\mathrm{tr}(A) \rightarrow \mathrm{tr}(A_i), \mathrm{tr}(B)$. $\square$

LEMMA 6.7. *For every small* $\mathrm{ACC}^0[2]$ *formula* A *of the form* $\wedge(A_1, \dots, A_m)$ *and for every* $i \in [m]$*, the sequents* $\mathrm{tr}(A) \rightarrow \mathrm{tr}(A_i)$, $\mathrm{tr}(A) \rightarrow \mathrm{tr}(\bigwedge_{j \neq i} A_j)$ *and* $\mathrm{tr}(A_i)$, $\mathrm{tr}(\bigwedge_{j \neq i} A_j) \rightarrow \mathrm{tr}(A)$ *have flat proofs of size* $2^{(\log n)^{O(d)}}$.

PROOF. All three sequents are proved by using the proofs of the corresponding sequents for the $\vee$ connective (see Lemmas 6.2 and 6.6). We show how this is done for the first sequent; the arguments are similar for the other two.

We have to show that

$$\sum_{\sigma, \tau} P_A(\sigma, \tau) \geq (1 - \varepsilon s_A) 2^{|y|+|z|} \ \rightarrow\ \sum_\sigma P_{A_i}(\sigma) \geq (1 - \varepsilon s_{A_i}) 2^{|y|}.$$

Since, for every σ,τ, $P_A(\sigma,\tau) = 1 \oplus Q^m_\vee(1\oplus P_{A_1}\sigma(),\ldots,1\oplus P_{A_m}(\sigma),\tau)$, we have that

$$\sum_{\sigma,\tau} P_A(\sigma,\tau) \geq (1-\varepsilon s_A)2^{|y|+|z|},\ \sum_{\sigma,\tau}(1\oplus P_A(\sigma,\tau)) \geq \varepsilon s_A 2^{|y|+|z|}+1 \ \rightarrow$$

which implies that

$$\sum_{\sigma,\tau} P_A(\sigma,\tau) \geq (1-\varepsilon s_A)2^{|y|+|z|},\ \sum_{\sigma,\tau}(1\oplus P_A(\sigma,\tau)) \geq (1-\varepsilon s_A)2^{|y|+|z|} \ \rightarrow .$$

In addition,

$$\sum_{\sigma} P_{A_i}(\sigma) \geq \varepsilon s_{A_i}2^{|y|}+1 \ \rightarrow \ \sum_{\sigma} P_{A_i}(\sigma) \geq (1-\varepsilon s_{A_i})2^{|y|},$$

by Lemma 6.4, and

$$\rightarrow \ \sum_{\sigma} P_{A_i}(\sigma) \geq \varepsilon s_{A_i}2^{|y|}+1,\ \sum_{\sigma}(1\oplus P_{A_i}(\sigma)) \geq (1-\varepsilon s_{A_i})2^{|y|}.$$

Therefore, it is sufficient to show that

$$\sum_{\sigma}(1\oplus P_{A_i}(\sigma)) \geq (1-\varepsilon s_{A_i})2^{|y|} \ \rightarrow \ \sum_{\sigma,\tau}(1\oplus P_A(\sigma,\tau)) \geq (1-\varepsilon s_A)2^{|y|+|z|}.$$

This in turn can be proved by the same argument that was used in Lemma 6.2 to prove that $\mathrm{tr}(A_i) \rightarrow \mathrm{tr}(A)$ when $A = \vee(A_1,\ldots,A_m)$. ☐

LEMMA 6.8. *For every small* $\mathrm{ACC}^0[2]$ *formula of the form* $\oplus_i(A_1,\ldots,A_m)$, *the sequents*

$$\mathrm{tr}(\oplus_i(A_2,\ldots,A_m)) \ \rightarrow \ \mathrm{tr}(A_1),\ \mathrm{tr}(\oplus_i(A_1,\ldots,A_m))$$

$$\mathrm{tr}(A_1),\ \mathrm{tr}(\oplus_i(A_1,\ldots,A_m)) \ \rightarrow \ \mathrm{tr}(\oplus_{i-1}(A_2,\ldots,A_m))$$

and

$$\mathrm{tr}(\oplus_i(A_1,\ldots,A_m)) \ \rightarrow \ \mathrm{tr}(A_1),\ \mathrm{tr}(\oplus_i(A_2,\ldots,A_m))$$

have flat proofs of size $2^{(\log n)^{O(d)}}$.

PROOF. By Lemma 6.3, we have that

$$\mathrm{tr}(A_1),\ \mathrm{tr}(\oplus_{i-1}(A_2,\ldots,A_m)) \rightarrow \mathrm{tr}(\oplus_i(A_1,\ldots,A_m)).$$

By an argument similar to the one used in the proof of that lemma, and by using Lemma 6.4 for A_1, it can be shown that

$$\mathrm{tr}(\oplus_i(A_2,\ldots,A_m)) \rightarrow \mathrm{tr}(A_1),\ \mathrm{tr}(\oplus_i(A_1,\ldots,A_m)).$$

The other two sequents follow from these two by Lemmas 6.1 and 6.5. ☐

7. Main result: the simulation of $\mathrm{ACC}^0[2]$ proofs

THEOREM 7.1. *If the family of sequents $F = \{(\Gamma_n \to \Delta_n) : n \in \mathbf{N}\}$ has a depth-d $\mathrm{ACC}^0[2]$ proof, then $\mathrm{tr}(F) = \{(\mathrm{tr}(\Gamma_n) \to \mathrm{tr}(\Delta_n)) : n \in \mathbf{N}\}$ has a flat proof of size $2^{(\log n)^{O(d)}}$.*

PROOF. It is sufficient to show that every rule of inference that can occur in an $\mathrm{ACC}^0[2]$ proof can be simulated by a flat proof in the sense that whenever $\Gamma \to \Delta$ can be inferred from $\Gamma' \to \Delta'$ and $\Gamma'' \to \Delta''$, then $\mathrm{tr}(\Gamma) \to \mathrm{tr}(\Delta)$ can be derived from $\mathrm{tr}(\Gamma') \to \mathrm{tr}(\Delta')$ and $\mathrm{tr}(\Gamma'') \to \mathrm{tr}(\Delta'')$ by using a flat proof. It is easy to verify that this can be done by using the sequents from the previous section. □

COROLLARY 7.2. *If the family of DNF formulas $F = \{(\to C_{n1}, \dots, C_{nm}) : n \in \mathbf{N}\}$ has an $\mathrm{ACC}^0[2]$ proof, then F has a flat proof of size $2^{(\log n)^{O(d)}}$.*

PROOF. Given the previous theorem, there only remains to show that $\to C_{n1}, \dots, C_{nm}$ can be derived from $\to \mathrm{tr}(C_{n1}), \dots, \mathrm{tr}(C_{nm})$ by using a flat proof. Clearly, it suffices to show that for every formula A of the form $\wedge(u_1, \dots, u_m)$, where all the u_i's are propositional variables, the sequent $\mathrm{tr}(A) \to A$ has a flat proof.

By Lemma 6.7, we have that for every i, $\mathrm{tr}(A) \to \mathrm{tr}(u_i)$. Since $\mathrm{tr}(u_i)$ is simply u_i, this is the same as $\mathrm{tr}(A) \to u_i$. The sequent $\mathrm{tr}(A) \to A$ can now be derived by repeated applications of the $\wedge$-right rule. □

8. Generalization to $\mathrm{ACC}^0[p^k]$ proofs

We now explain how the simulation of $\mathrm{ACC}^0[2]$ proofs can be generalized to $\mathrm{ACC}^0[p]$ proofs.

First, redefine our logic system by using the symbol $\oplus_j$ to denote a mod p connective instead of a mod 2 connective. For $j = 0, \dots, p-1$, the formula $\oplus_j(A_1, \dots, A_m)$ is interpreted to be true if and only if the number of true A_i's is equal to j mod p. To the initial sequents $\oplus_1() \to$ and $\to \oplus_0()$, add $\oplus_j() \to$ for $j = 2, \dots, p-1$. The rules of inference remain the same. $\mathrm{ACC}^0[p]$ proofs can now be defined precisely as in Definition 2.3 and flat proofs now have mod p connectives on level two (see Definition 2.4).

The simulation of $\mathrm{ACC}^0[2]$ proofs made essential use of polynomials over $\mathbf{Z}_2$. Naturally, for $\mathrm{ACC}^0[p]$ proofs, polynomials over $\mathbf{Z}_p$ will be used instead. The symbols $\oplus$ and $\bigoplus$ will now denote addition of polynomials over $\mathbf{Z}_p$. Statements about polynomials over $\mathbf{Z}_p$ will be of the form $u \equiv a \pmod p$, where u is a polynomial over $\mathbf{Z}_p$ in the propositional variables and $a \in \mathbf{Z}_p$. Suppose that $u = a_0 + a_1 M_1 + \cdots + a_N M_N$, where the M_i's are nonempty products of variables and the a_i's are coefficients in $\mathbf{Z}_p$. Then a statement such as $u \equiv a \pmod p$ will be realized by the formula $\oplus_{a-a_0}(M_1, \dots, M_1, M_2, \dots, M_2, \dots, M_N, \dots, M_N)$, where each M_i is repeated a_i times. Whenever a polynomial u appears on its own in a formula, then $u \equiv 1 \pmod p$ is assumed. In particular, $u \oplus 1$ is equivalent to $u \equiv 0 \pmod p$.

Probabilistic polynomials over $\mathbf{Z}_p$ are associated with the various connectives as follows. Let $\varepsilon = 1/2^{\lceil 1+\log p+\log s\rceil}$. For the $\neg$ connective, let $Q_\neg(u_1) = 1 \oplus (p-1)u_1$. For the mod p connectives, let

$$Q^m_{\oplus_j}(u) = 1 \oplus (p-1)\left((p-j) \oplus \bigoplus_{i=1}^{m} u_i\right)^{p-1}.$$

The fact that $Q^m_{\oplus_j}(u)$ computes $\oplus_j(u)$ follows from Fermat's Little Theorem: $a^{p-1} \equiv 1 \pmod p$ if and only if $a \not\equiv 0 \pmod p$.

Consider now an $\vee$ connective. Recall that the main idea behind the definition of $Q^m_\vee$ in the mod 2 case was the fact that if $\frac{1}{8}2^k \leq |S| < \frac{1}{4}2^k$, then $R_k = \bigoplus_{i:h(i)=r} u_i$ is equal to 1 with probability at least 1/16. That is, the number of $i \in S$ such that $h(i) = r$ is odd with probability at least 1/16. In the mod p case, this is replaced by "the number of $i \in S$ such that $h(i) = r$ is not divisible by p." Accordingly, R_k is now $(\bigoplus_{i:h(i)=r} u_i)^{p-1}$ and $Q^m_\vee$ is defined by

$$Q^m_\vee(u, w) = 1 \oplus (p-1) \prod_{l=1}^{C} \prod_{k=3}^{\lceil \log m \rceil + 3} \left(1 \oplus (p-1) \left(\bigoplus_i u_i F^i_{l,k} \right)^{p-1} \right).$$

Naturally, for the $\wedge$ connective we now have

$$Q^m_\wedge(u, w) = 1 \oplus (p-1) Q^m_\vee(1 \oplus (p-1)u_1, \ldots, 1 \oplus (p-1)u_m, w).$$

The lemmas of Sections 5 and 6, and their proofs, have to be adapted to the new definitions and the new context. Only minor changes are required and they usually involve basic properties of polynomials over $\mathbf{Z}_p$.

In Section 5, the key property of $Q^m_\vee$ now requires an additional hypothesis, namely, the fact that the inputs to the $\vee$ connective, the polynomials u_i, are such that $u_i \equiv 1 \pmod p$ or $u_i \equiv 0 \pmod p$. Consequently, in Lemmas 5.3 to 5.6, instead of "the sequent ... has a flat proof of size $2^{(\log n)^{O(d)}}$", we now have "the sequent ... can be derived from the sequents $\rightarrow u_i, u_i \oplus 1$, for $i = 1, \ldots, m$, by a flat proof of size $2^{(\log n)^{O(d)}}$". Recall that $u_i \oplus 1$ stands for $u_i \equiv 0 \pmod p$.

Some of the other changes are as follows. In the paragraph preceding Lemma 5.3, let $\varphi_{h,r} = (\bigoplus_{i:h(i)=r} u_i)^{p-1}$, $\varphi^i_{h,r} = u_i \varphi_{h,r}$ and $\varphi^{i,j}_{h,r} = u_i u_j \varphi_{h,r}$. In the proof of Lemma 5.4, $E^i_{h,r}$ should be defined as $1 \oplus (1 \oplus \bigoplus_{j:j \neq i, h(j)=r} u_j)^{p-1}$. Then, for every h, r,

$$\rightarrow E^i_{h,r}, \left(1 \oplus \bigoplus_{j:j \neq i, h(j)=r} u_j \right)^{p-1}.$$

The proof then proceeds as in the mod 2 case, with $1 \oplus E^i_{h,r}$ replaced by $(1 \oplus \bigoplus_{j:j \neq i, h(j)=r} u_j)^{p-1}$.

In Section 6, we first need an additional lemma.

LEMMA 8.1. *For every small* $\mathrm{ACC}^0[p]$ *formula* A*, the sequent*

$$\rightarrow P_A(\sigma_1, \ldots, \sigma_d), 1 \oplus P_A(\sigma_1, \ldots, \sigma_d)$$

has a flat proof of size $2^{(\log n)^{O(d)}}$.

This is easily proved by induction on the structure of A.

Lemma 6.1 should be for $\mathrm{tr}(\oplus_i(A_1, \ldots, A_m))$, $\mathrm{tr}(\oplus_j(A_1, \ldots, A_m)) \rightarrow$, $i, j \in \{0, \ldots, p-1\}$, $i \neq j$, instead of just $\mathrm{tr}(\oplus_1(A_1, \ldots, A_m))$, $\mathrm{tr}(\oplus_0(A_1, \ldots, A_m)) \rightarrow$. Similarly, Lemmas 6.3 and 6.8 should be for $i \in \{0, \ldots, p-1\}$. As for Lemma 6.5, the sequent $\rightarrow \mathrm{tr}(\oplus_1(A_1, \ldots, A_m))$, $\mathrm{tr}(\oplus_0(A_1, \ldots, A_m))$ should be replaced by

$$\rightarrow \mathrm{tr}(\oplus_0(A_1, \ldots, A_m)),\ \mathrm{tr}(\oplus_1(A_1, \ldots, A_m)), \ldots, \mathrm{tr}(\oplus_{p-1}(A_1, \ldots, A_m)).$$

The proof of this sequent uses the fact that $\varepsilon \leq 1/(ps)$. In Lemma 6.4, cases have to be added for $\oplus_i(A_1, \ldots, A_m)$, $i \in \{2, \ldots, p-1\}$, but these are handled as in the mod 2 case.

Finally, we arrive at the main result, the simulation of $\mathrm{ACC}^0[p]$ proofs by quasipolynomial-size flat proofs.

THEOREM 8.2. *If the family of sequents* $F = \{(\Gamma_n \to \Delta_n) : n \in \mathbf{N}\}$ *has a depth-*d $\mathrm{ACC}^0[p]$ *proof, then* $\mathrm{tr}(F) = \{(\mathrm{tr}(\Gamma_n) \to \mathrm{tr}(\Delta_n)) : n \in \mathbf{N}\}$ *has a flat proof of size* $2^{(\log n)^{O(d)}}$.

This is proved exactly as Theorem 7.1.

It is possible to further generalize the simulation to $\mathrm{ACC}^0[p^k]$ proofs. These are defined exactly as $\mathrm{ACC}^0[p]$ proofs except that we now have connectives $\oplus_{j \bmod p^e}$, for $j = 0, \ldots, p^e - 1$ and $e = 1, \ldots, k$: the formula $\oplus_{j \bmod p^e}(A_1, \ldots, A_m)$ is true if and only if the number of true A_i's is equal to j mod p^e. The simulation then proceeds exactly as for $\mathrm{ACC}^0[p]$ proofs. In particular, polynomials over $\mathbf{Z}_p$ are still used and the simulating flat proofs still have only mod p connectives on level two (no mod p^e connectives). The only difference is in the handling of the mod p^e connectives.

The idea behind the translation of mod p^e connectives into low-degree polynomials over $\mathbf{Z}_p$ comes from a simple fact that was used, for example, by Beigel and Tarui [**BT94**] in their simulation of ACC^0 circuits (see Section 9 for more on their result): $u_1 + \cdots + u_m \equiv 0 \pmod{p^e}$ if and only if $\binom{u_1 + \cdots + u_m}{p^l} \equiv 0 \pmod{p}$ for all $l \in \{0, \ldots, e-1\}$. Accordingly, let

$$Q^m_{\oplus_{j \bmod p^e}}(u) = \prod_{l=0}^{e-1} \left(1 \oplus (p-1) \left(\sum_{a=0}^{p^l} \left(\binom{p^e - j}{a} \sum_{S \subseteq [m],\, |S| = p^l - a} \prod_{i \in S} u_i \right) \right)^{p-1} \right).$$

It is not hard to see that the summation over a in this polynomial is congruent to $0 \pmod{p}$ if and only if $\binom{u_1 + \cdots + u_m + p^e - i}{p^l} \equiv 0 \pmod{p}$.

The lemmas of Section 6 can be easily adapted to these new polynomials. One fact that must be used is that $\binom{p^e}{a} \equiv 0 \pmod{p}$ for every $a \in \{1, \ldots, p^e - 1\}$.

9. Possible extensions to ACC^0 proofs

Our depth-three simulation of $\mathrm{ACC}^0[p]$ proofs relies in an essential way on the corresponding simulation of $\mathrm{ACC}^0[p]$ circuits. As we have seen, most of the simulation consists in the formalization, in our logic system, using quasipolynomial-size flat proofs, of the key step in the circuit simulation, i.e., the computation of OR gates by low-degree probabilistic polynomials over $\mathbf{Z}_p$, and of several related basic facts. Fixed-depth simulations are also known for general ACC^0 circuits. For example, let SYM^+ denote the class of depth-two circuits with symmetric gates at the output and AND gates of polylog(n) fan-in on level one. Beigel and Tarui [**BT94**] have shown that ACC^0 circuits can be simulated by quasipolynomial-size SYM^+ circuits, which implies a simulation by depth-three threshold circuits. (See also [**GKR$^+$95**].) Therefore, it is natural to ask if our simulation of $\mathrm{ACC}^0[p]$ proofs can be extended to ACC^0 proofs, perhaps by using the simulation results for ACC^0 circuits.

An important characteristic of the simulation of $\mathrm{ACC}^0[p]$ circuits, one that was not mentioned explicitly earlier, is that this simulation is "bottom-up". This means that the simulation of a circuit $A = *(A_1, \dots, A_m)$ is done inductively by first simulating each subcircuit A_i and then combining with the $*$ gate to get the simulation of A. This characteristic plays an important role in our simulation because the rules of inference of the logic system involve the top-most connectives of the formulas. To illustrate this point, consider the $\vee$-right rule

$$\frac{\Gamma \to A_1, \vee(A_2, \dots, A_m), \Delta}{\Gamma \to \vee(A_1, \dots, A_m), \Delta}.$$

It is easy to see that this rule can be simulated by a flat proof if and only if the two sequents

$$\mathrm{tr}(A_1) \;\to\; \mathrm{tr}(\vee(A_1, \dots, A_m))$$

and

$$\mathrm{tr}(\vee(A_2, \dots, A_m)) \;\to\; \mathrm{tr}(\vee(A_1, \dots, A_m))$$

have flat proofs themselves. Now the proof of these two sequents, which was done in Lemma 6.2, relies crucially on the fact that the polynomial $P_{\vee(A_1,\dots,A_m)}$ computing the formula $\vee(A_1, \dots, A_m)$ is defined in terms of the polynomials P_{A_i} computing the A_i's.

All of the known simulations of ACC^0 circuits, however, are "top-down". That is, the simulation of a circuit A of the form $B(G_1, \dots, G_m)$, where the G_i's are the gates at level one, is done by first simulating B and then combining with the G_i's to get a simulation of A. As a consequence, it is not clear how a sequent such as $\mathrm{tr}(A_1) \to \mathrm{tr}(\vee(A_1, \dots, A_m))$ could be proved without having to "translate back" to A_1, derive $\vee(A_1, \dots, A_m)$ and then "translate" again to $\mathrm{tr}(\vee(A_1, \dots, A_m))$. The reason is that if $A = *(A_1, \dots, A_m) = B(G_1, \dots, G_m)$, then, in a top-down simulation, $\mathrm{tr}(A)$ would be defined in terms of $\mathrm{tr}(B)$ and not in terms of the $\mathrm{tr}(A_i)$'s.

We are therefore led to the following question: is it possible to do a "bottom-up", fixed-depth simulation of ACC^0 circuits? This is an interesting question on its own, but a positive answer might also lead to a fixed-depth simulation of ACC^0 proofs. One way of obtaining such a "bottom-up" simulation would be to show that a circuit of the form MOD_r of SYM^+ can be simulated by a quasipolynomial-size SYM^+ circuit. Notice that this would imply that ACC^0 circuits with one level of arbitrary symmetric gates inserted anywhere could be simulated by quasipolynomial-size SYM^+ circuits. Such a result would be surprising; however, since no superpolynomial lower bound is known for SYM^+ circuits, this possibility cannot be ruled out. In addition, a bottom-up simulation of ACC^0 circuits—at least one that would proceed as indicated at the beginning of this section—would imply a simulation of ACC^0 circuits with an extra level of arbitrary symmetric gates at the input. The idea is to simply start with a stronger inductive basis. Therefore, the simulation of MOD_r of SYM^+ circuits by quasipolynomial-size SYM^+ circuits might not only be sufficient but also necessary for a bottom-up simulation of ACC^0 circuits. Whether this should be taken as an approach for obtaining such a simulation or as evidence in favor of its impossibility is open for debate.

10. Towards $\mathrm{ACC}^0[p]$-proof lower bounds

In this section, we outline an approach for obtaining exponential lower bounds for $\mathrm{ACC}^0[p]$ proofs.

In this article, we have shown that any quasipolynomial-size $\mathrm{ACC}^0[p]$ proof can be quasipolynomially simulated by a flat Frege proof. In fact, each formula can be viewed as a small-degree probabilistic polynomial. Thus, if we could succeed in obtaining lower bounds for our flat system, then $\mathrm{ACC}^0[p]$-proof lower bounds would follow.

A very successful recent method for obtaining lower bounds for very small depth propositional proof systems is the *interpolation method.* The general idea is as follows. Let $A(x,z) \wedge B(y,z)$ be a 3CNF formula, where $A(x,z)$ is a 3CNF formula involving the variables x and z, and $B(y,z)$ is a 3CNF formula involving the variables y and z. (Here, x, y, and z each denote a vector of propositional variables.) Let S be a particular propositional proof system. Then S has a *feasible interpolation theorem* if for every 3CNF formula F of the above split form, there exists a circuit, $C(z)$, such that: $C(a) = 1$ only if $A(x,a)$ is unsatisfiable, and $C(a) = 0$ only if $B(y,a)$ is unsatisfiable. Furthermore, the size of the circuit C must be polynomial in the size of the shortest S-refutation of F. Note that for any assignment a to z, at least one of $A(x,a)$ or $B(y,a)$ must be unsatisfiable because F is unsatisfiable. So the function computed by $C(z)$ exists, and we are interested in the case where the function is computable by polynomial-size circuits whenever F has a short refutation in S. There is also a monotone version of the above definition. Define $A(x,z) \wedge B(y,z)$ to be monotone if $A(x,z)$ involves only positive occurrences of z variables. In this case, S has a *feasible monotone interpolation theorem* if for every 3CNF formula of the monotone split form, there exists a monotone, polynomial-size circuit $C(z)$ computing the above (monotone) function.

Now suppose that we can show that S has a feasible interpolation theorem. Then, using a result due to Razborov [**Razb**] (and later generalized by Krajíček [**Kraa**]) this implies lower bounds for S, assuming a standard cryptographic conjecture. Furthermore, if we can show that S has a feasible monotone interpolation theorem, then using the method of [**BPR95**] (generalized by [**Kraa**]), together with exponential lower bounds on the size of monotone circuits, we get unconditional lower bounds for our proof system S. Thus, feasible interpolation theorems are quite important because they give rise to lower bounds.

This method has been quite successful in obtaining new lower bounds. For example, it has been used to obtain exponential lower bounds for Resolution and Cutting Planes [**BPR95, Krab, Pud, CH**], as well as for Nullstellensatz proofs [**PS**]. It has also been applied to get conditional lower bounds for bounded arithmetic [**Razb**], for Gröbner refutations [**PS**], and for generalizations of Cutting Planes [**IP, Krab**].

When does a proof system have an effective interpolation theorem, and how hard is it to show this? This is a very interesting question that does not have a satisfactory answer at present. For example, we do not know if $\mathrm{ACC}^0[p]$ proofs have effective interpolation theorems. An interesting observation, due to Impagliazzo, can be used to obtain interpolation theorems. Suppose that system S can be *deterministically simulated efficiently.* This means that there is an efficient procedure testing whether or not there is a polynomial-size S proof of a given formula F. The existence of an efficient deterministic simulation not only suggests a good deterministic propositional theorem prover, but it also implies an effective interpolation theorem for S. This approach was used to obtain interpolation theorems for the Nullstellensatz system [**CEI96, PS**], as well as for the polynomial calculus [**PS**].

In light of the interpolation technology discussed above, it would be very interesting to obtain a subexponential-time deterministic simulation of flat proofs. First, this would give rise to a powerful and possibly efficient deterministic theorem prover. Second, it would imply superpolynomial lower bounds for $\mathrm{ACC}^0[p^k]$ proofs (assuming a standard cryptographic assumption).

As mentioned above, [**CEI96**] shows how to deterministically simulate polynomial calculus refutations using a variation of the Gröbner basis algorithm; our flat proofs can be viewed as a generalization of small-degree polynomial calculus refutations in the following way. A small-degree polynomial calculus refutation is a refutation in our system where each line is a small-degree polynomial over the main variables x. One derives new polynomials as small-degree linear combinations of previously generated polynomials. This linearity of the rules is exactly what enables the Gröbner basis algorithm to be used to efficiently find a small-degree polynomial calculus refutation, if one exists. In contrast, a flat refutation can be viewed as a *parametric* polynomial calculus refutation: each line is a small-degree polynomial but now over two types of variables, the main variables x and the symbolic parameters y. The intended interpretation is that $f(x, y)$ is true if for almost all values to the y variables, $f(x, y)$ evaluates to 1, and $f(x, y)$ is false if for almost all values of the y variables, $f(x, y)$ evaluates to 0. It is open whether or not there is a suitable generalization of the Gröbner basis algorithm which would give rise to a subexponential-time simulation of flat proofs.

References

[AH94] E. Allender and U. Hertrampf, *Depth reduction for circuits of unbounded fan-in*, Inform. and Comput. **108** (1994), 217–238.

[Ajt88] M. Ajtai, *The complexity of the pigeonhole principle*, Proceedings of the 29th IEEE Symposium on Foundations of Computer Science, 1988, pp. 346–355.

[Ajt90] M. Ajtai, *Parity and the pigeonhole principle*, Feasible Mathematics (S.R. Buss and P.J. Scott, eds.), Birkhauser, 1990, pp. 1–24.

[All89] E. Allender, *A note on the power of threshold circuits*, Proceedings of the 30th IEEE Symposium on Foundations of Computer Science, 1989, pp. 580–584.

[BC96] S. Buss and P. Clote, *Cutting planes, connectivity, and threshold logic*, Arch. Math. Logic **35** (1996), no. 1, 33–62.

[BIK+a] P. Beame, R. Impagliazzo, Krajíček, T. Pitassi, and Pudlák P., *Lower bounds on Hilbert's nullstellensatz and propositional proofs*, To appear.

[BIK+b] S. Buss, R. Impagliazzo, J. Kraíček, P. Pudlák, A.A. Razborov, and J. Sgall, *Proof complexity in algebraic systems and bounded-depth Frege systems with modular counting*, Manuscript, 1996.

[BP] P. Beame and T. Pitassi, *An exponential separation between the matching principle and the pigeonhole principle*, To appear in *Annals of Pure and Applied Logic.*

[BPR95] M. Bonet, T. Pitassi, and R. Raz, *Lower bounds for cutting planes proofs with small coefficients*, Proceedings of the 27th ACM Symposium on Theory of Computing, 1995, Full version to appear in *Journal of Symbolic Logic.*

[BT94] R. Beigel and J. Tarui, *On ACC*, Comput. Complexity **4** (1994), 350–366.

[CEI96] M. Clegg, J. Edmonds, and R. Impagliazzo, *Using the Gröbner basis algorithm to find proofs of unsatisfiability*, Proceedings of the 28th ACM Symposium on Theory of Computing, 1996.

[CH] S.A. Cook and A. Haken, *An exponential lower bound for the size of monotone real circuits*, Manuscript, 1995.

[GKR+95] F. Green, J. Köbler, K. Regan, T. Schwentik, and J. Torán, *The power of the middle bit of a #P function*, J. Comput. System Sci. **50** (1995), 456–467.

[Hak85] A. Haken, *The intractability of resolution*, Theoret. Comput. Sci. **39** (1985), 297–308.

[IP] R. Impagliazzo and T. Pitassi, *Interpolation theorems for generalized cutting planes*, Manuscript, 1996.

[Kraa] J. Krajíček, *Bounded arithmetic, propositional logic and complexity theory*, Cambridge University Press.

[Krab] J. Krajíček, *Interpolation theorems, lower bounds for proof systems and independence results for bounded arithmetic*, To appear in the *Journal of Symbolic Logic*.

[PS] P. Pudlák and J. Sgall, *Algebraic models of computation and interpolation for algebraic proof systems*, Manuscript, 1996.

[Pud] P. Pudlák, *Lower bounds for resolution and cutting planes proofs and monotone computation*, To appear in *Journal of Symbolic Logic*.

[Raza] A.A. Razborov, *Lower bounds for the polynomial calculus*, Manuscript, 1996.

[Razb] A.A. Razborov, *Unprovability of lower bounds on the circuit size in certain fragments of bounded arithmetic*, Izvestiya of the R.A.N. **59**, no. 1, 201–224.

[Raz87] A.A. Razborov, *Lower bounds for the size of circuits of bounded depth with basis* $\{\wedge, \oplus\}$, Mathematical Notes of the Academy of Sciences of the USSR **41** (1987), no. 4, 333–338.

[Reg93] K. Regan, *Efficient reductions from NP to parity using error-correcting codes*, Tech. Report 93–24, Department of Computer Science, State University of New York at Buffalo, Buffalo, NY, U.S.A., 1993.

[Rii] S. Riis, *Count(q) does not imply Count(p)*, Manuscript, 1995.

[Sip83] M. Sipser, *A complexity theoretic approach to randomness*, Proceedings of the 15th ACM Symposium on Theory of Computing, 1983, pp. 330–335.

[Smo87] R. Smolensky, *Algebraic methods in the theory of lower bounds for boolean circuit complexity*, Proceedings of the 19th ACM Symposium on Theory of Computing, 1987, pp. 77–82.

[Wig94] A. Wigderson, *Lectures on the fusion method and derandomization*, Tech. Report SOCS-95.2, School of Computer Science, McGill University, Montréal, Québec, Canada, 1994.

[Yao90] A.C.-C. Yao, *On ACC and threshold circuits*, Proceedings of the 31st IEEE Symposium on Foundations of Computer Science, 1990, pp. 619–627.

Department of Computer Science and UMIACS, University of Maryland, College Park, MD 20742, U.S.A.

Current address: Department of Mathematics and Computer Science, Clarkson University, Potsdam, NY 13699-5815, U.S.A.

E-mail address: `alexis@cs.umd.edu`

Department of Computer Science, University of Arizona, Tucson, AZ 85721, U.S.A.

E-mail address: `toni@cs.arizona.edu`