Abstract

We give a formalization of $\text{AC}^0$ lower bounds based on Håstad’s switching lemma, Razborov-Smolensky’s $\text{AC}^0[p]$ lower bounds and monotone circuit lower bounds in Jeřábek’s theory of approximate counting $\text{APC}_1$. We use these formalizations to obtain short proofs and automatizability of Extended Frege system $\text{EF}$ and its extension $\text{WF}$ on various weak circuit lower bounds.

1 Introduction

Proving lower bounds on the size of Boolean circuits computing explicit Boolean functions is a fundamental problem in complexity theory. Interestingly, the known circuit lower bounds are often very constructive as captured in the notion of natural proofs by Razborov and Rudich [30]. We analyze these constructive aspects from the perspective of proof complexity.

The investigation of proof complexity of known circuit lower bounds was initiated by Razborov [27] who argued that all existing circuit lower bounds for explicit Boolean functions are derivable in the theory $\text{PV}_1$ formalizing p-time reasoning, and often below. For example, the theory $\text{U}_1$ corresponding to $\text{NC}$ reasoning proves $\text{AC}^0$ lower bounds, $\text{PV}_1$
proves $\text{AC}^0[p]$ lower bounds and $W^{1}_{1} \tau$, corresponding to reasoning with uniform p-size circuits of a suitable depth, proves monotone circuit lower bounds. Further, Krajíček [16, Theorem 15.2.3] formalized $\text{PARITY} \not\in \text{AC}^0$ with a different scaling in $\text{PV}_1 + \text{WPHP}(\text{PV})$.

The main contribution of this paper is a derivation of analogous results in the framework of theories of approximate counting developed by Jeřábek [14]. Specifically, in the theory $\text{APC}_1$ formalizing probabilistic p-time reasoning which slightly extends Krajíček’s $\text{PV}_1 + \text{WPHP}(\text{PV})$. We show that $\text{APC}_1$ proves $\text{PARITY} \not\in \text{AC}^0$ by formalizing Håstad’s switching lemma, $\text{AC}^0[p]$ lower bounds by formalizing Razborov-Smolensky’s method, and monotone circuit lower bounds by formalizing the approximation method.

A crucial difference between Razborov’s and our formalizations is in the scaling of parameters. In Razborov’s formalizations, whenever a theory $T$ proves a lower bound for a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ it is assumed that $2^n$ is the length of some number. This means that from the perspective of the theory $T$, the whole truth-table of $f$ is a feasible object. Our formalizations assume only that $n$ is the length of some number. The same scaling of parameters was used in Krajíček’s proof of $\text{PARITY} \not\in \text{AC}^0$. Consequently, the theory Razborov is working in is exponentially stronger w.r.t. his formulation of circuit lower bounds than it is w.r.t. ours. For more details see section 2.1.

We do not develope new methods for deriving circuit lower bounds, quite the opposite, we keep the original proofs as intact as possible. Some changes were, however, needed. In case of the $\text{AC}^0$ and monotone circuit lower bound, the probabilities used in the known proof are estimated by Jeřábek’s notion of approximate counting, cf. section 4. This requires a construction of surjections witnessing the approximations. More invasive changes are needed in the case of the $\text{AC}^0[p]$ lower bound. The degree lower bound in Razborov-Smolensky’s method typically requires to consider exponentially big objects (the set of all functions on $n$ inputs). In order to simulate the argument in $\text{APC}_1$ we scale it down to the functions with logarithmic input size. Secondly, we code the approximating polynomials with arithmetic circuits because the set of all coefficients representing such polynomials can be infeasible.

The presented upper bounds are to a large extent motivated by their propositional counterpart. Propositional formulas encoding circuit lower bounds like $\text{SAT} \not\in \text{P/poly}$ are considered as candidate hard tautologies for strong proof systems like Frege, cf. section 3. Razborov’s argument about the provability of known circuit lower bounds in $\text{PV}_1$ translates to p-size proofs of $2^{O(n)}$-size propositional formulas encoding known circuit lower bounds in Extended Frege system $\text{EF}$ (and often in weaker systems). Here, $n$ is the input-size of the function on which the lower bound is proven. Our formalizations yield efficient $\text{EF}$ proofs of existing circuit lower bounds expressed by propositional formulas of size $\text{poly}(n)$ under the assumption of another circuit lower bound simulating the power of $\text{APC}_1$ in $\text{EF}$.

Additionally, we show that the $\text{AC}^0$ and $\text{AC}^0[p]$ lower bounds can be naturalized within $\text{APC}_1$. Consequently, we obtain efficient algorithms generating $\text{EF}$ proofs of $\text{poly}(n)$-
size tautologies expressing $\text{AC}^0$ and $\text{AC}^0[p]$ lower bounds (assuming another circuit lower bound) for a wider class of functions. In particular, we get the so called automatizability resp. quasi-automatizability of $\text{EF}$ on many $\text{AC}^0$ and $\text{AC}^0[p]$ lower bounds. The naturalization of Razborov-Smolensky’s method gives us also $\text{WF}$ proofs of $\text{AC}^0[p]$ lower bounds without extra assumptions. Here, $\text{WF}$ is a canonical proof system corresponding to $\text{APC}_1$ on $\text{coNP}$ statements, cf. [12].

For the completeness we formalize also the natural proofs barrier itself in $\text{APC}_1$.

The paper is organized as follows. Section 2 gives the preliminaries on bounded arithmetic, propositional proof complexity, and discusses the formulation of circuit lower bounds in the language of bounded arithmetic and propositional logic. Section 3 presents some previous results concerning the provability of complexity-theoretic statements. Sections 4, 5 describe the $\text{APC}_1$ in more detail and some standard inequalities in $\text{PV}_1$. Section 6 contains the formalizations of $\text{AC}^0$, $\text{AC}^0[p]$ and monotone circuit lower bounds. Section 7 gives a naturalization of $\text{AC}^0$ and $\text{AC}^0[p]$ lower bounds in $\text{APC}_1$ which yields an automatizability of $\text{EF}$ on these circuit lower bounds. Section 7 provides also a formalization of the natural proofs barrier itself in $\text{APC}_1$. Finally, section 8 recapitulates possible improvements of our results and suggests some future research directions.

2 Bounded arithmetic and propositional logic

Theories of bounded arithmetic capture various levels of feasible reasoning and present a uniform counterpart to propositional proof systems.

The first theory of bounded arithmetic formalizing p-time reasoning was introduced by Cook [7] as an equational theory $\text{PV}$. We work with its first-order conservative extension $\text{PV}_1$ from [21]. The language of $\text{PV}_1$, denoted $\text{PV}$ as well, consists of symbols for all p-time algorithms given by Cobham’s characterization of p-time functions, cf. [6]. A $\text{PV}$-formula is a first-order formula in the language $\text{PV}$. $\Sigma^b_0$ ($=\Pi^b_0$) denotes $\text{PV}$-formulas with only sharply bounded quantifiers $\exists x, x \leq |t|$, $\forall x, x \leq |t|$, where $|t|$ is “the length of the binary representation of $t$”. Inductively, $\Sigma^b_{i+1}$ resp. $\Pi^b_{i+1}$ is the closure of $\Pi^b_i$ resp. $\Sigma^b_i$ under positive Boolean combinations, sharply bounded quantifiers, and bounded quantifiers $\exists x, x \leq t$ resp. $\forall x, x \leq t$. Predicates definable by $\Sigma^b_i$ resp. $\Pi^b_i$ formulas are in the $\Sigma^b_i$ resp. $\Pi^b_i$ level of the polynomial hierarchy, and vice versa. $\text{PV}_1$ is known to prove $\Sigma^b_0(\text{PV})$-induction,

$$A(0) \land \forall x \ (A(x) \rightarrow A(x+1)) \rightarrow \forall x A(x)$$

for $\Sigma^b_0$-formulas $A$, cf. Krajiček [16].

Buss [3] introduced the theory $\text{S}^2_2$ extending $\text{PV}_1$ with the length induction

$$A(0) \land \forall x < |a|, (A(x) \rightarrow A(x+1)) \rightarrow \forall x A(|a|)$$
for \( A \in \Sigma^b_1 \), \( S^1_2 \) proves the sharply bounded collection scheme \( BB(\Sigma^b_1) \),

\[
\forall i < |a| \exists x < a, A(i, x) \rightarrow \exists w \forall i < |a|, A(i, [w]_i)
\]

for \( A \in \Sigma^b_1 \) ([\( w \]_i is the \( i \)-th element of the sequence coded by \( w \)], which is unprovable in \( PV_1 \) under a cryptographic assumption, cf. [9]. On the other hand, \( S^1_2 \) is \( \forall \Sigma^b_1 \)-conservative over \( PV_1 \). This is a consequence of Buss’s witnessing theorem stating that \( S^1_2 \vdash \exists y, A(x, y) \) for \( A \in \Sigma^b_1 \) implies \( PV_1 \vdash A(x, f(x)) \) for some \( PV \)-function \( f \). When proving a \( \Sigma^b_2 \) formula in \( S^1_2 \) we are free to use the sharply bounded collection scheme for \( A \in \Sigma^b_2 \), denoted \( BB(\Sigma^b_2) \), because \( S^1_2 + BB(\Sigma^b_2) \) is \( \forall \Sigma^b_2 \)-conservative over \( S^1_2 \), cf. [31].

Jeřábek [14] developed a theory \( APC_1 \) capturing probabilistic p-time reasoning by means of approximate counting. The theory \( APC_1 \) is defined as \( PV_1 + dWPHP(\ PV_1) \) where \( dWPHP(\ PV_1) \) stands for the dual (surjective) pigeonhole principle for \( PV \)-functions, i.e. for the set of all formulas

\[ x > 0 \rightarrow \exists v < x(|y| + 1) \forall u < x|y|, f(u) \neq v \]

where \( f \) is a \( PV \)-function. We devote Section 4 to a more detailed description of the machinery of approximate counting in \( APC_1 \).

Any \( \Pi^b_1 \)-formula provable in \( PV_1 \) can be expressed as a sequence of tautologies \( \tau_n \) with proofs in the Extended Frege system \( EF \) which are constructible in p-time (given a string of the length \( n \)), cf. [7]. Similarly, \( \Pi^b_1 \)-formulas provable in \( APC_1 \) translate to tautologies with p-time constructible proofs in \( WF \), an extension of \( EF \) introduced by Jeřábek [12].

As it is often easier to present a proof in a theory of bounded arithmetic than in the corresponding propositional system, bounded arithmetic functions, so to speak, as a uniform language for propositional logic.

### 2.1 Formulation of circuit lower bounds

A typical formulation of a circuit lower bound for circuits of size \( s \) and a function \( f \) says that for every sufficiently big \( n \), each circuit \( C \) with \( n \) inputs and size \( s \), there exists an input \( y \) on which the circuit \( C \) fails to compute \( f(y) \).

If \( f : \{0, 1\}^n \rightarrow \{0, 1\} \) is an \( NP \) function and \( s = n^k \) for a constant \( k \), this can be written down as a \( \forall \Sigma^b_2 \) formula \( LB(f, n^k) \),

\[
\forall n, n > n_0 \forall \text{ circuit } C \text{ of size } \leq n^k \exists y, |y| = n, C(y) \neq f(y),
\]

where \( n_0 \) is a constant and \( C(y) \neq f(y) \) is a \( \Sigma^b_2 \) formula stating that a circuit \( C \) on input \( y \) outputs the opposite value of \( f(y) \).

If we want to express \( s(n) \)-size lower bounds for \( s(n) \) as big as \( 2^{O(n)} \), we add an extra assumption on \( n \) stating that \( \exists x, n = ||x|| \). The resulting formula \( LB_H(f, s(n)) \) is \( \Sigma^b_0 \) if \( f \)}
is, for instance, SAT because $n = ||x||$ implies that the quantifiers bounded by $2^{O(n)}$ are sharply bounded. Moreover, allowing $f \in \text{NE}$ lifts the complexity of $\text{LB}_{tt}(f, s(n))$ just to $\Sigma^b_1$.

To indicate sizes of objects we employ the notation $x \in \text{Log} \iff \exists y, x = |y|$ and $x \in \text{LogLog} \iff \exists y, x = ||y||$. For example, $\text{LB}(f, n^k)$ implicitly assumes $n \in \text{Log}$ while $\text{LB}_{tt}(f, n^k)$ assumes $n \in \text{LogLog}$. By chosing the scale of $n$ we are chosing the “feasible object”. In the case $n \in \text{LogLog}$, the truth-table of $f$ (and everything polynomial in it) is feasible. Assuming just $n \in \text{Log}$ means that only the objects of polynomial-size in the size of the circuit are feasible. Likewise, the theory reasoning about the circuit lower-size in the the complexity of $\text{LB}_{tt}(f, s(n))$ just to $\Sigma^b_1$.

The scaling in $\text{LB}_{tt}(f, s)$ corresponds to the choice of parameters in natural proofs and in the formalizations by Razborov [27].

We will work mainly with lower bounds for restricted circuit classes like $\text{AC}^0$, constant depth circuits with a polynomial number of gates of unbounded arity, and $\text{AC}^0[p]$, $\text{AC}^0$ circuits with $\text{MOD}_p$ gates. Such circuit lower bounds can be formulated similarly without increasing the quantifying complexity of the resulting formula. For example, by $\text{LB}(f, \text{AC}^0_d, n^k)$ we denote $\text{LB}(f, n^k)$ restricted to $\text{AC}^0$ circuits of size $n^k$ and depth $d$. Analogously for $\text{LB}_{tt}$ formulation and other circuit classes.

### 2.2 Propositional version

An $s(n)$-size circuit lower bound for a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ can be expressed by a $2^{O(n)}$-size propositional formula $\text{tt}(f, s)$,

$$\bigvee_{y \in \{0, 1\}^n} f(y) \neq C(y)$$

where the formula $f(y) \neq C(y)$ says that a circuit $C$ represented by $\text{poly}(s)$ variables does not output $f(y)$ on input $y$. That is, the whole truth-table of $f$ is hard-wired in $\text{tt}(f, s)$.

A more succinct encoding follows from a result of Lipton and Young [22] who showed that whenever $f : \{0, 1\}^n \rightarrow \{0, 1\}$ is hard for circuits of size $\text{poly}(s(n))$, there is a set $S_n$ of $\text{poly}(s(n))$ $n$-bit strings such that each $s(n)$-size circuit fails to compute $f$ on some input from the “anti-checking” set $S_n$. The $s(n)$-size circuit lower bound for $f$ can be then expressed by a $\text{poly}(s(n))$-size formula $\text{lb}_A(f, s)$,

$$\bigvee_{y \in S_n} f(y) \neq C(y).$$

Even more feasible, uniform, encoding follows from translations of $\text{LB}(f, n^k)$. This requires an efficient witnessing of existential quantifiers in $\text{LB}(f, n^k)$ collapsing its complexity
to \( \forall \Sigma^b \). Such a p-time witnessing of \( \text{LB}(\text{SAT}, n^k) \) follows, for example, from the existence of one-way permutations and a function in \( E \) hard for subexponential-size circuits, cf. [23, Proposition 4.3] \(^1\). Further, by the KPT theorem [21], whenever \( PV_1 \vdash \text{LB}(f, n^k) \) we get a sequence of finitely many p-time functions \( \overline{w} = w_1, \ldots, w_c \) witnessing the existential quantifiers in \( \text{LB}(f, n^k) \). \( \text{LB}(f, n^k) \) witnessed by \( \overline{w} \) can be equivalently expressed by a sequence of \( \text{poly}(n) \)-size propositional formulas \( \text{lb}_{\overline{w}}(f, n^k) \).

Restricting \( \text{tt}(f, n^k) \) to \( \text{AC}^0 \) circuits of a depth \( d \) we obtain propositional formulas denoted \( \text{tt}(f, \text{AC}^0_d, n^k) \). Similarly for \( \text{lb}_A(f, \text{AC}^0_d, n^k) \) and \( \text{lb}_{\overline{w}}(f, \text{AC}^0_d, n^k) \).

Formulas \( \text{lb}_A(f, n^k) \) and \( \text{lb}_{\overline{w}}(f, n^k) \) seem to be harder to derive than \( \text{tt}(f, n^k) \). This intuition can be formally supported.

\[ \text{Proposition 2.1.} \quad \text{For any constant } k, \text{ if formulas } \text{tt}(f, n^k) \text{ do not have p-size constant-depth Frege proofs, then formulas } \text{lb}_A(f, n^k) \text{ do not have p-size Frege proofs. Here, } \text{tt}(f, n^k) \text{ and } \text{lb}_A(f, n^k) \text{ are assumed to be expressed as DNFs.} \]

\[ \text{Proof.} \quad \text{Suppose that Frege has p-size proofs of } \text{lb}_A(f, n^k) \text{ for some } A. \text{ A generic collapse of Frege to constant depth Frege by Filmus Pitassi and Santhanam [11] implies that there is a constant } K \text{ such that for any } d, \text{ Frege proves } \text{lb}_A(f, n^k) \text{ by proofs of size } 2^{O(dn^k/d)} \text{ and depth } d + 2. \text{Tautologies } \text{tt}(f, n^k) \text{ can be then derived by weakening which increases the size of the proofs to } 2^{O(n)}. \]

\[ \square \]

3 Prior results

3.1 Lower bounds

Assuming the existence of strong pseudorandom generators, Razborov [28, 26] showed that a theory \( S^2_2(\alpha) \) cannot prove superpolynomial circuit lower bounds for \( \text{SAT} \) in a formulation which corresponds to \( \text{LB}_{\text{tt}}(\text{SAT}, n^k) \) but with circuits coded by the oracle \( \alpha \). Unfortunately, the theory \( S^2_2(\alpha) \) is too weak w.r.t. this formulation. It is not clear how to derive results like \( \text{PARITY} \not\in \text{AC}^0 \) within \( S^2_2(\alpha) \).

For the formulation of circuit lower bounds used in this paper, Pich [23] showed that the theory \( \text{VNC}^1 \) formalizing \( \text{NC}^1 \) reasoning, cf. [8], cannot prove \( \text{LB}(\text{SAT}, n^k) \) unless functions computable by p-size circuits can be approximated by subexponential \( \text{NC}^1 \) circuits. Concerning the power of \( \text{VNC}^1 \), it seems plausible that \( \text{VNC}^1 \) could prove \( \text{NC}^1 \) lower bounds like \( \text{LB}(\text{SAT}, \text{NC}^1, n^k) \). On the other hand, we do not even know how to prove \( \text{LB}(\text{PARITY}, \text{AC}^0, n^k) \) in \( PV_1 \).

\(^1\)Proposition 4.3 in [23] shows just the existence of an S-T protocol witnessing \( \text{LB}(\text{SAT}, n^k) \) but the p-time witnessing easily follows.
For \( PV_1 \), Krajíček and Oliveira [19] obtained an unconditional unprovability result involving circuit upper bounds. Specifically, they showed that for every \( k \) there is a p-time function \( f \) such that \( PV_1 \) does not prove \( f \in \text{SIZE}(cn^k) \) for any constant \( c \). Concerning the unprovability of upper bounds, Buss’s witnessing theorem [3] implies that \( PV_1 \) cannot prove \( \text{NP} = \text{coNP} \) unless \( \text{P} = \text{NP} \). Further, any superpolynomial lower bound on the lengths of proofs in \( \text{EF} \) would imply an unprovability of \( \text{SAT} \in \text{P/poly} \) in \( PV_1 \). This implication is generic, e.g. the existing lower bounds for bounded depth Frege yield an unconditional unprovability of \( \text{SAT} \in \text{P/poly} \) in a theory \( V^0 \), cf. [18].

Razborov’s unprovability result from [28, 26] can be nicely formulated on the propositional level. Assuming strong pseudorandom generators exist, no sufficiently strong proof system admitting the so called feasible interpolation property has \( p \)-size proofs of \( \text{tt}(f, n^k) \) where \( f \) is an arbitrary function [17]. Unfortunately, stronger proof systems like Frege or even constant depth Frege do not admit feasible interpolation unless a cryptographic assumption fails [20, 2]. For weak propositional systems, lower bounds on \( \text{tt}(f, n^k) \) can be derived unconditionally. Raz [25] showed that formulas \( \text{tt}(f, n^k) \), for sufficiently big constant \( k \), have no \( p \)-size Resolution proofs and Razborov [29] obtained a \( 2^{\Omega(1)} \)-size lower bound on the lengths of proofs of \( \text{tt}(f, t) \) for \( n^2 \leq t \leq 2^n \), in en extension of Resolution operating with \( k \)-DNFs (which is not known to admit feasible interpolation). The results of Raz and Razborov work with a specific choice of encoding of \( \text{tt}(f, n^k) \) suitable for weak proof systems.

Notably, formulas \( \text{tt}(f, s) \) are special instances of proof complexity generators considered as candidate hard tautologies for strong proof systems like Frege. In fact, Razborov’s conjecture [29, Conjecture 1] implies the hardness of \( \text{tt}(f, n^k) \) for Frege assuming the existence of functions computable by \( p \)-size circuits and hard on average for \( \text{NC}^1 \).

### 3.2 Upper bounds

As already discussed in the introduction, Razborov [27] argued that the theory \( U_1 \) corresponding to \( \text{NC} \) reasoning proves \( \text{AC}^0 \) lower bounds, \( PV \) proves \( \text{AC}^0[p] \) lower bounds and \( W_1^{1,\tau} \), corresponding to reasoning with uniform \( p \)-size circuits of a suitable depth, proves monotone circuit lower bounds. Further, Krajíček [16, Theorem 15.2.3] derived \( \text{PARITY} \notin \text{AC}^0 \) with a different (the same as our) scaling in \( PV_1 + \text{WPHP}(PV) \). Since Krajíček’s \( \text{WPHP}(PV) \) is a special case of \( dWPHP(PV) \) which does not appear to imply \( dWPHP(PV) \) over \( PV_1 \), his result is stronger than ours.

The constructivity of many other parts of complexity theory has been demonstrated as well. For an illustration see Table 1.
4 Approximate counting

This section recalls a part of Jeřábek’s theory for approximate counting, cf. [14].

By a definable set we mean a collection of numbers satisfying some formula, possibly with parameters. When a number $a$ is used in a context which asks for a set it is assumed to represent the integer interval $[0, a)$, e.g. $X \subseteq a$ means that all elements of set $X$ are less than $a$. If $X \subseteq a$, $Y \subseteq b$, then $X \times Y := \{bx + y \mid x \in X, y \in Y\} \subseteq ab$ and $X \cup Y := X \cup \{y + a \mid y \in Y\} \subseteq a + b$. Rational numbers are assumed to be represented by pairs of integers in the natural way.

Let $n, m \in \log$, $C : 2^n \rightarrow 2^m$ be a circuit and $X \subseteq 2^n, Y \subseteq 2^m$ definable sets. We write $C : X \rightarrow Y$ if $\forall y \in Y \exists x \in X, C(x) = y$. Jeřábek [14] gives the following definitions in APC$_1$ but there is no need to restrict them.

**Definition 4.1.** Let $X, Y \subseteq 2^n$ be definable sets, and $\epsilon \leq 1$. The size of $X$ is approximately less than the size of $Y$ with error $\epsilon$, written as $X \preceq_\epsilon Y$, if there exists a circuit $C$, and $v \neq 0$ such that

$$C : v \times (Y \cup 2^n) \rightarrow v \times X.$$  

$X \approx_\epsilon Y$ stands for $X \preceq_\epsilon Y$ and $Y \preceq_\epsilon X$.

Since a number $s$ is identified with the interval $[0, s)$, $X \preceq_\epsilon s$ means that the size of $X$ is at most $s$ with error $\epsilon$.

**Definition 4.2.** Let $X \subseteq 2^{[t]}$ be a definable set and $0 \leq \epsilon, p \leq 1$. We define

$$\Pr_{x \leq t}[x \in X] \preceq_\epsilon p \iff X \cap t \preceq_\epsilon pt$$
and similarly for $\approx_{\epsilon}$.

The definition of $X \preceq_{\epsilon} Y$ is an unbounded $\exists \Pi^b_2$-formula even if $X, Y$ are defined by circuits so it cannot be used freely in bounded induction. Jeřábek [14] solved this problem by working in HARD$^A$, a conservative extension of APC$_1$, defined as a relativized theory PV$_1(\alpha) + dWPHP(PV(\alpha))$ extended with axioms postulating that $\alpha(x)$ is a truth-table of a function on $||x||$ variables hard on average for circuits of size $2^{|x|/4}$. In HARD$^A$ there is a PV$_1(\alpha)$ function $\text{Size}_2$ approximating the size of any set $X \subseteq 2^n$ defined by a circuit $C$ so that $X \preceq_{\epsilon} \text{Size}_2(C, 2^n, 2^{\epsilon^{-1}})$ for $\epsilon^{-1} \in \text{Log}$. If $X \subseteq 2^{[t]}$ is defined by a circuit $C$ and $\epsilon^{-1} \in \text{Log}$, we have

$$\Pr_{x < t}[x \in X] : = \frac{1}{t} \text{Size}(C, 2^{[t]}, 2^{\epsilon^{-1}}).$$

The presented definitions of approximate counting are well-behaved:

**Proposition 4.1** (Jeřábek [14]). (in PV$_1$) Let $X, X', Y, Y', Z \subseteq 2^n$ and $W, W' \subseteq 2^m$ be definable sets, and $\epsilon, \delta, \theta < 1$. Then

i) $X \subseteq Y \Rightarrow X \preceq_{0} Y$,

ii) $X \preceq_{\epsilon} Y \land Y \preceq_{\delta} Z \Rightarrow X \preceq_{\epsilon + \delta} Z$,

iii) $X \preceq_{\epsilon} X' \land W \preceq_{\delta} W' \Rightarrow X \times W \preceq_{\epsilon + \delta + \theta} X' \times W'$.

iv) $X \preceq_{\epsilon} X' \land Y \preceq_{\delta} Y'$ and $X', Y'$ are separable by a circuit, then $X \cup Y \preceq_{\epsilon + \delta} X' \cup Y'$.

**Proposition 4.2** (Jeřábek [14]). (in APC$_1$)

1. Let $X, Y \subseteq 2^n$ be definable by circuits, $s, t, u \leq 2^n$, $\epsilon, \delta, \theta, \gamma < 1, \gamma^{-1} \in \text{Log}$. Then

i) $X \preceq_{\gamma} Y$ or $Y \preceq_{\gamma} X$,

ii) $s \preceq_{\epsilon} X \preceq_{\epsilon + \delta} t \Rightarrow s < t + (\epsilon + \delta + \gamma)2^n$,

iii) $X \preceq_{\epsilon} Y \Rightarrow 2^n \setminus Y \preceq_{\epsilon + \gamma} 2^n \setminus X$,

iv) $X \approx_{\epsilon} s \land Y \approx_{\delta} t \land X \cap Y \approx_{\theta} u \Rightarrow X \cup Y \approx_{\epsilon + \delta + \theta + \gamma} s + t - u$.

2. (Disjoint union) Let $X_i \subseteq 2^n, i < m$ be defined by a sequence of circuits and $\epsilon, \delta \leq 1, \delta^{-1} \in \text{Log}$. If $X_i \preceq_{\epsilon} X_i$ for every $i < m$, then $\bigcup_{i < m}(X_i \times \{i\}) \preceq_{\epsilon + \delta} \sum_{i < m} s_i$.

3. (Averaging) Let $X \subseteq 2^n \times 2^m$ and $Y \subseteq 2^m$ be definable by circuits, $Y \preceq_{\epsilon} t$ and $X_y \preceq_{\delta} s$ for every $y \in Y$, where $X_y := \{x \mid (x, y) \in X\}$. Then for any $\gamma^{-1} \in \text{Log}$,

$$X \cap (2^n \times Y) \preceq_{\epsilon + \delta + \epsilon \delta + \gamma} st.$$ .

It is practical to observe that for proving $\Sigma^b_1$ statements in APC$_1$ we can afford to work in $S_2^1 + dWPHP(PV) + BB(\Sigma^b_2)$ and, in fact, assuming the existence of a single hard function in PV$_1$ gives us the full power of APC$_1$.

**Lemma 4.1.** Suppose $S_2^1 + dWPHP(PV) + BB(\Sigma^b_2) \vdash \exists y A(x, y)$ for $A \in \Sigma^b_1$. Then, for every $\epsilon < 1$, there is $k$ and $g, h \in PV$ such that PV$_1$ proves

$$|f| \geq |x|^k \land \exists y, |y| = ||f||, C_h(y) \neq f(y) \to A(x, g(x, f))$$

where $C_h$ is a circuit of size $\leq 2^{|x|}$ generated by $h$ on $f, x$. Moreover, APC$_1 \vdash \exists y A(x, y).$
Proof. By [12, Corollary 4.12], $S^1_2 + dWPHP(PV) + BB(\Sigma^b_2) \vdash \exists yA(x, y)$ implies $S^1_2 + dWPHP(PV) \vdash \exists yA(x, y)$. Then, following Thapen’s proof of [32, Theorem 4.2] (cf. also [12, Proposition 1.14]), there is $\ell$ and $h \in PV$ such that $S^1_2$ proves

$$(\forall v \leq 2^{8|x|^\ell} \exists u \leq 2^{4|x|^\ell}, h(u) = v) \vee \exists yA(x, y).$$

Argue in $S^1_2$. The surjection $h : 2^m \to 2^{2m}$, where $m = 4|x|^\ell$, is computed by a circuit of size $m^\ell$ for a standard $\ell'$. Following Jeřábek’s $S^1_2$-proof of [12, Proposition 3.5], this implies that every (number) $f$ viewed as a truth-table of length $|f|$ is computed by a circuit of size $O(m|m| + m^\ell[|f|/m])$ circuit with $|f|$ inputs. For sufficiently large $k$, $|f| \geq |x|^k$ implies that this size is $\leq 2^{e||f||}$. The “moreover” part is a consequence of $APC_1 \vdash \forall n \in \text{LogLog} \exists f : 2^n \to 2, \text{LB}_{tt}(f, 2^{n/4})$, cf. [12, Corollary 3.3].

Lemma 4.1 allows us to use the $BB(\Sigma^b_2)$ collection scheme for proving $\Sigma^b_1$-statements in $APC_1$. Unfortunately, when collecting circuits witnessing $\leq_\epsilon$ predicates given by $\exists \Pi^b_2$ formulas the $BB(\Sigma^b_2)$ collection is a priori not sufficient. To overcome this complication the quantifier complexity of $\leq_\epsilon$ can be pushed down to $\Sigma^b_2$ because the circuits counting sizes of sets in $APC_1$ are invertible.

**Lemma 4.2.** (in $APC_1$) Let $X \subseteq 2^n$ be defined by a circuit and $\epsilon^{-1} \in \text{Log}$. Suppose $X \leq_\epsilon s$. Then, $X \leq_\epsilon s + 3\epsilon 2^n$ is expressible by a provable $\Sigma^b_2$ formula.

**Proof.** By [14, Theorem 2.7], there exists $t$ such that $X \approx_\epsilon t$ is witnessed by invertible circuits of size $\text{poly}(ne^{-1}S)$ where $S$ is the size of the circuit defining $X$. Applying Proposition 4.2 1.$ii$) we get $t < s + 3\epsilon 2^n$. \(\square\)

## 5 Standard inequalities in $PV_1$

For a PV-function symbol $f$ and $n \in \text{Log}$, in $PV_1$ we can define inductively $\sum_{i=0}^n f(i)$. Similarly, we can define iterated products, factorials, and binomial coefficients. It is easy to see that, by induction, $PV_1$ proves: $n \in \text{Log} \to \sum_{i=0}^n \binom{n}{i} = 2^n$. 

10
Proposition 5.1 (Stirling’s bound, cf. Jeřábek [12]). There is a $c > 1$ such that $\text{PV}_1$ proves:

$$0 < k < n \in \text{Log} \rightarrow \frac{1}{c} \binom{n}{k} < \frac{n^n}{k^k (n-k)^{n-k}} \left( \left\lceil \frac{k(n-k)}{n} \right\rceil + 1 \right)^{-1} < c \binom{n}{k}.$$ 

Proposition 5.2. For each $\epsilon > 0$ there is an $n_0$ such that $\text{PV}_1$ proves:

$$n_0 < n \in \text{Log} \rightarrow \sum_{i=0}^{[n/2]+n^{1/3}} \binom{n}{i} < (\frac{1}{2} + \epsilon) 2^n.$$ 

Proof. $\sum_{i=0}^{[n/2]-1} \binom{n}{i} = \frac{1}{2} \left( \sum_{i=0}^{[n/2]-1} \binom{n}{i} + \sum_{i=0}^{[n/2]-1} \binom{n}{n-i} \right) < 2^{n-1}$ and by Stirling’s bound, for some constant $c > 1$,

$$\sum_{i=[n/2]}^{[n/2]+n^{1/3}} \binom{n}{i} < (n^{1/3} + 1) \binom{n}{[n/2]} < 2^n 4^c \left( \binom{n^{1/3}}{[n/2]/2} + \frac{1}{[n^{1/2}/2]} \right)$$

where to verify the last inequality for odd $n$ we used also the provability of $a, b \in \text{Log}, b > 0 \rightarrow (1 + a/b) \leq 4^{a/b}$ shown in [12, Stirling’s bound, Claim 1].

Proposition 5.3. $\text{PV}_1$ proves:

$$a, b \in \text{Log}, b > a + 1 \rightarrow (b - a)^b \leq b^b/2^a.$$ 

Note that the conclusion implies $(1 - a/b) \leq 2^{-a/b}$.

Proof. Proceed as in the proof of Claim 2 in the proof of Stirling’s bound [12] but instead of Claim 1 use the inequality $b^b \leq (b + 1)^b/2$. 

6 Known circuit lower bounds

6.1 Random restrictions

In $\text{APC}_1$, for any $n \in \text{Log}$ and $0 \leq \frac{a}{b} \leq 1$ we code a restriction of $n$ variables $x_1, \ldots, x_n$ by $\rho = \sum_{i=0}^{n-1} r_{i+1} (2b)^i, r_i \in [0, 2b)$ with the following interpretation: if $r_i \in [0, 2a)$, then $\rho(x_i) = x_i$, if $r_i \in [2a, b+a)$ then $\rho(x_i) = 1$, if $r_i \in [b+a, 2b)$ then $\rho(x_i) = 0$. The notation $\rho \in R_{a/b}$ stands for $\rho < (2b)^n$. It is straightforward to construct the circuits witnessing $\Pr_{\rho \in R_{a/b}}[\rho(x_i) = x_i] \approx_0 \frac{a}{b}$ and $\Pr_{\rho \in R_{a/b}}[\rho(x_i) = 1] \approx_0 \frac{1-a/b}{2} \approx_0 \Pr_{\rho \in R_{a/b}}[\rho(x_i) = 0]$ for each $x_i \in X$. 

11
When considering predicates of the form $\Pr_{\rho_1, \rho_2} \ldots \leq_\epsilon a$ where $\rho_1 < (2b_1)^n, \rho_2 < (2b_2)^n$, the subscript $\rho_1, \rho_2$ represents $x < (2b_1)^n(2b_2)^n$ with $x$ interpreted as a pair $\rho_1, \rho_2$. Similar conventions are applied in the rest of the paper as well.

Given restrictions $\rho, \rho_1, \rho_2$ and a circuit $C$ with $n$ inputs, we denote by $C|\rho$ the circuit $C(\rho(x_1), \ldots, \rho(x_n))$ and by $C|\rho_1 \rho_2$ the circuit $C|\rho_1|\rho_2$. By the size of a circuit we mean the number of its (internal) gates. Irrational terms are assumed to be rounded down on the innermost level, e.g. $(1/n^{1/2})^c$ is $(1/\lfloor n^{1/2} \rfloor)^c$ and $2 \log n$ is $2 \lfloor \log n \rfloor$, unless specified otherwise.

**Definition 6.1.** A DNF depends on $> b$ inputs if no $\leq b$-tuple of inputs has the property that every its assignment either sets every literal in some disjunct to 1 or sets some literal in every disjunct to 0. Analogously for CNFs.

**Lemma 6.1** (Håstad’s switching lemma). For each $k$, there is $b$ and $n_0$ such that $\mathrm{APC}_1$ proves: for each $n_0 < n, \epsilon^{-1} \in \log$ and $\text{DNF or CNF } D_n(x_1, \ldots, x_n)$ of size $n^k$,

$$\Pr_{\rho_1, \rho_2} [D_n|\rho_1 \rho_2 \text{ depends on } > b \text{ inputs}] \leq_\epsilon 1/n^{2k}$$

where $\rho_1, \rho_2$ are random restrictions from $R_{1/n^{1/2}}, R_{1/n^{1/4}}$ respectively. Note that the event is defined by a circuit.

**Proof.** We follow a familiar proof of the switching lemma estimating the probabilities that formulas are reduced under random restrictions. The probabilities are approximated by Jeřábek’s notion of $\leq_\epsilon$. The extra work then boils down mainly to the construction of surjections witnessing the inequalities $\leq_\epsilon$. These constructions are postponed to the end of the proof. We prove the lemma for DNFs. The CNF case is derived analogously.

Let $n$ be sufficiently big and $n, \epsilon^{-1} \in \log$. For $d = 12k$ we have,

$$\Pr_{\rho_1} [\rho_1 \text{ does not falsify all disjuncts in } D_n \text{ of size } \geq d \log n] \leq_0 n^k \left(1 - \frac{1 - 1/n^{1/2}}{2}\right)^{d \log n} \leq n^k \left(\frac{3}{4}\right)^{d \log n} \leq \frac{1}{n^{3k}}. \tag{6.1}$$

For $c = 12k + 3d + 3$,

$$\Pr_{\rho_1} [\rho_1 \text{ leaves } \geq c \text{ inputs unassigned in some disjunct in } D_n \text{ of size } \leq d \log n] \leq_0 n^k \left(\frac{1}{n^{1/2}}\right)^c 2^{d \log n} \leq \frac{1}{n^{3k}}. \tag{6.2}$$

Therefore, by Proposition 4.1 iv), the probability that $D_n|\rho_1$ after a trivial simplification is not a $c$-DNF is $\leq_0 2/n^{3k}$. Now it suffices to derive the following claim.
Claim 6.1. For any $c' \leq c$, there are $n_0, b_{c'}$ such that $\text{APC}_1$ proves: for $n_0 \leq n$, $\epsilon^{-1} \in \text{Log}$ and each $c'$-DNF $D'_n(x_1, \ldots, x_n)$,

$$\Pr[D'_n|\rho_2 \text{ depends on } > b_{c'} \text{ inputs}] \leq b_{c'} b_{c'}/n^{3k}.$$ 

To prove the claim we proceed by induction on $c'$. If $c' = 0$, the claim holds trivially. Assume that the claim holds for $(c' - 1)$-DNFs, we want to show that it holds for $c'$-DNFs. Let $S$ be a sequence of disjuncts with disjoint variables in $D'_n$ which is maximal in the sense that adding any other disjunct to $S$ would break the disjointness property. (Note that constructing the maximal set among all such sequences $S$ could be hard for $\text{APC}_1$.) If the number of disjuncts in $S$ is $\geq d' \log n$ with $d' = 4^c 4k$, we have,

$$\Pr[\text{no disjunct in } D'_n|\rho_2 \text{ equals } 1] \leq \epsilon \left(1 - \left(\frac{1 - 1/n^{1/4}}{2}\right)^{c'}\right)^{d' \log n} \leq 2 \frac{-d' \log n}{4^c} \leq \frac{1}{n^{3k}} \quad (6.3)$$

where we used the provability of $1 - x \leq 2^{-x}$ (Proposition 5.3). Otherwise, for $b_{c'} = 15k$,

$$\Pr[\rho_2 \text{ leaves } > b_{c'} \text{ variables in } S \text{ unassigned}] \leq_0 \left(\frac{1}{n^{1/4}}\right)^{b_{c'} + 1} \left(\frac{1}{b_{c'} + 1}\right) \leq \frac{1}{n^{3k}}. \quad (6.4)$$

As every disjunct outside $S$ shares a variable with some disjunct from $S$, by setting all variables in $S$ we get a $(c' - 1)$-DNF which by the induction hypothesis depends on $> b_{c' - 1}$ inputs with probability $\leq b_{c' - 1} b_{c' - 1}/n^{3k}$. Hence, by Proposition 4.1 iv), $D'_n|\rho_2$ depends on $> b_{c'} = b_{c'} + 2b_{c'} b_{c' - 1}$ inputs with probability $\leq 2^{2b_{c'} b_{c' - 1}} 2^{b_{c'} b_{c' - 1}}/n^{3k} + 1/n^{3k} \leq b_{c'}/n^{3k}$.

It remains to describe p-time algorithms witnessing the estimations (6.1)-(6.4). For example, in case of inequality (6.1), we want to map every $z \leq n^k(n^{1/2} + 1)^{d\log n}(2n^{1/2})^{n-d\log n}$ to a restriction $\rho_1 < (2n^{1/2})^n$ in such a way that any $\rho_1$ which does not falsify all disjuncts in $D_n$ of size $\geq d\log n$ is provably in the image of the mapping. Such surjections can be constructed in the following way:

(6.1) Given $z \leq n^k(n^{1/2} + 1)^{d\log n}(2n^{1/2})^{n-d\log n}$ find the triple $\langle s, p, r \rangle$ represented by $z$ with $s < n^k$, $p = \sum_{i=0}^{d\log n-1} \epsilon_i (n^{1/2} + 1)^i$, $\epsilon_i < n^{1/2} + 1$ and $r = \sum_{i=0}^{n-d\log n-1} r_i (2n^{1/2})^i$, $r_i < 2n^{1/2}$. Then output $\rho$ assigning the first $d\log n$ variables in the $s$th disjunct of $D_n$ according to $\epsilon_0, \ldots, \epsilon_{d\log n-1}$ so that the disjunct is not falsified and the rest according to $r_0, \ldots, r_{n-d\log n-1}$.

(6.2) Given $z \leq n^{k-c/2}2^{d\log n}(2n^{1/2})^n$ representing $\langle s, t, p, r \rangle \in n^k \times 2^c \times 2^{d\log n} \times (2n^{1/2})^{n-c}$ output $\rho$ assigning the first $c_0 \leq c$ variables in the $s$th disjunct of $D_n$ on the positions specified by $p$ according to $t$ (these variables remain unassigned by $\rho$), for the maximal $c_0$ possible, and the rest of variables according to $r$ together with the unused part of $t$. 

13
(6.3) First observe that for every $j$, $\Pr_{\rho_3}[j$th disjunct in $D'_n=1] \geq (1-1/n^{1/4})^{c'}$ where $\rho_3 < (2n^{1/4})$ are from $R_{1/n^{1/4}}$ and restricted only to the $s_j$ variables of the $j$th disjunct in $D'_n$. By Proposition 4.2 1.iii) (comprising dWPHP(PV)), there is a circuit $S_j$ certifying $\Pr_{\rho_3}[j$th disjunct in $D'_n=1] \leq \epsilon/(d' \log n) - (1-1/n^{1/4})^{c'}$. We can now witness $\Pr_{\rho_2}[\text{no disjunct in } D'_n=1] = \epsilon \left( 1 - \left(1-1/n^{1/4}\right)^{c'} \right)^{d' \log n}$ by mapping each $z$ coding the tuple $(z_0, \ldots, z_{d' \log n-1}, r)$, where $z_j < \left( 1 - \left(1-1/n^{1/4}\right)^{c'} \right) (2n^{1/4})^{s_j}$, $r < (2n^{1/4})^{n-S_j}$, to $\rho_2$ given by $S_j(z_j)$’s and $r$. The construction of the last witnessing circuit uses a collection scheme which can be sidestepped by realizing that there is actually only a constant number of $S_j$’s needed - one for each disjunct on $c'$ variables.

(6.4) Given $z$ coding the triple $(s,t,r) \in 2^{b'_c+1} \times (c'd' \log n) \times (2n^{1/4})^{n-b'_c-1}$, output $\rho$ assigning the first $c_0 \leq b'_c + 1$ variables in $S$ specified by the $t$-th $(b'_c+1)$-size subset of $c'd' \log n$, for the maximal $c_0$ possible, according to $s$ (these variables remain unassigned) and the rest according to $r$ together with the unused part of $s$.

$\square$

**Theorem 6.1.** For any $k,d$ there is $n_0$ such that APC$_1$ proves: for all $n_0 < n \in \log$ and each depth $d$ size $n^k$ circuit $C_n$ with $n$ inputs there is $y \in \{0,1\}^n$ such that $C_n(y) \neq \sum_{i=1}^n y_i \ (mod \ 2)$.

**Proof.** There is a PV-function transforming any $n^k$-size circuit $C_n$ of depth $d$ into an equivalent $C'_n$ circuit of size $n^{2k}$, depth $d$ and with negations appearing only at the inputs. This is proven by $\Sigma^b_0$(PV)-induction on the number of gates in $C'_n$, hence, already in PV$_1$. By Lemma 6.1, random restrictions $\rho_1, \rho_2$ simplify any DNF and CNF at the bottom level of $C'_n$ so that it depends on $>b$ inputs with probability $\leq \epsilon 1/n^{4k}$. Such DNFs, resp. CNFs, are equivalent to CNFs, resp. DNFs, of size $\leq (b+1)^{2^b}+1$. Furthermore, we have

$$\Pr_{\rho_1, \rho_2} \left[ \rho_1 \rho_2 \text{ leave } < n^{1/8} \text{ inputs unassigned} \right] \leq \epsilon n^{n^{1/8}} \left( 1 - \frac{1}{n^{3/4}} \right)^{n-n^{1/8}}$$

$$\leq n^{n^{1/8}} 2^{-(n-n^{1/8})/n^{3/4}} \leq n^{n^{1/8}} 2^{1-n^{1/4}} \leq \frac{1}{n^{2k}}$$

where we used the provability of $1-x \leq 2^{-x}$ from Proposition 5.3. The inequality $\leq_0$ is witnessed by mapping $z = (s,p,r) \in n^{1/8} \times (4n^{3/4} - 4)n^{n^{1/8}} \times (4n^{3/4})^{n^{1/8}}$ to $\rho_1, \rho_2$ which set the variables of $x_{s_0+1}, \ldots, x_{s_{n^{1/8}-1}+1}$ where $s = \sum_{i=0}^{n^{1/8}-1} s_i n^i$, $s_i < n$, according
to $r$ (in particular, $x_{s_0+1}, \ldots, x_{s_n/s-1+1}$ might be left unassigned by $\rho_1, \rho_2$) and the rest of variables according to $p$.

Therefore, applying Proposition 4.2 (Disjoint union) and Proposition 4.2 1.iii), the probability that $\rho_1, \rho_2$ simplify all CNFs and DNFs at the bottom while preserving at least $n^{1/8}$ variables free is $\geq 3e^{-1/2\sqrt{n}}$. By Proposition 4.2 1.ii) this shows that there exist restrictions $\rho_1, \rho_2$ such that $C_n^{\rho_1,\rho_2}$ is equivalent to a circuit with $\geq n^{1/8}$ inputs, depth $d - 1$ and size $\leq (b + 1)2^b + 1)n^{\log k}$. Iterating this reduction we obtain a CNF $C$ with $\geq n^{1/O(1)}$ inputs and size $\text{poly}(n)$. If $C$ computed the parity function or its negation we would get a contradiction: the parity function depends on each input, so w.l.o.g. each input appears in each conjunct of $C$, but this means that any $O(\log n)$ inputs can make $C$ evaluate to 1 independently of the rest of the inputs.

\begin{corollary}
For any $k, d$ there are $n_0, k_0$ and $w, h \in \text{PV}$ such that EF has p-size proofs of tautologies
\[
\bigvee_{y \in \{0,1\}^{k_0 \log n}} C_h(y) \neq f(y) \rightarrow \text{lb}_w(\text{PARITY}, \text{AC}_d^0, n^k)
\]
where $f$ is a Boolean function with $k_0 \log n$ inputs represented by $2^{k_0 \log n}$ variables, $w$ is a p-time witnessing function with an access to $f$, and $C_h$ is a circuit of size $2^{(k_0 \log n)/2}$ generated by $h$ on the inputs of $w$.

Tautologies \text{tt}(\text{PARITY}, \text{AC}_d^0, n^k)$ have p-size WF proofs and $2^{O(n \log n)}$-size EF proofs.
\end{corollary}

\begin{proof}
Apply Theorem 6.1, Lemma 4.1 and observe that $|f| \geq |x|^{k_0}$ translates to propositional formulas with short EF proofs.

If $n \in \text{LogLog}$, the AC$^0$ lower bound from Theorem 6.1 becomes the $\Pi_1^b$ formula $\text{LB}_{\text{tt}}(\text{PARITY}, \text{AC}_d^0, n^k)$ which for any constant $\ell$ translates to tautologies $2^n = \ell \rightarrow \text{tt}(\text{PARITY}, \text{AC}_d^0, n^k)$ with p-size WF proofs, cf. [12]. Moreover, if the assumption $2^n = \ell$ holds for constants $n$ and $\ell$, it has a trivial WF proof.

If $2^n \log n \in \text{Log}$, it is feasible to list all elements of sets $X \subseteq 2^{O(n \log n)}$ which are all the sets we need to count in Theorem 6.1.

Strengthening Corollary 6.1 to p-size WF proofs of $\text{lb}_w(\text{PARITY}, \text{AC}_d^0, n^k)$ for some witnessing functions $w$, could be achieved by a derandomization of the witnessing of the existential quantifiers in $\text{LB}(\text{PARITY}, \text{AC}_d^0, n^k)$ within APC$_1$. More precisely, the derandomized algorithm would need to be built without the need for the existence of a hard function while the proof of its properties could use APC$_1$. This might be doable at least in quasi-polynomial time by formalizing the derandomized switching lemma from [33]. In Section 7 we give a quasi-polynomial algorithm generating WF proofs of $\text{lb}_{A_n}(\text{PARITY}, \text{AC}_d^0, n^k)$ for some $A_n$ by formalizing a naturalization of Razborov-Smolensky’s lower bound.

Is it possible to formalize AC$^0$ lower bounds in PV$_1$ or perhaps even in V$^0$? For this we would need to design surjections witnessing the probabilities considered in the switching lemma and Theorem 6.1 without using $dWPHP(PV)$. In particular, we would need to
give a p-time algorithm which generates the \(i\)th restriction \(\rho\) eliminating all \(\log n\)-size disjuncts in \(D_n\) for a given DNF \(D_n\) and \(i \leq (1 - 1/n^3k)(2b)^n\).

### 6.2 Razborov-Smolensky method

Jeřábek [13, Section 4.3] gave a formalization of finite fields with their basic properties in bounded arithmetic. In particular, if \(p \in \text{Log}\) is a prime, we can construct in \(\text{PV}_1\) the finite field \(\mathbb{F}_p\) and prove that for \(a \in \mathbb{F}_p \setminus \{0\}\), \(a^{p-1} = 1 \pmod{p}\) [13, Lemma 4.3.11].

In Theorem 6.2 we want to approximate each \(\text{AC}^0[p]\) circuit by a polynomial \(p(x) \in \mathbb{F}_p[x_1, \ldots, x_n]\). Unfortunately, the sequence of coefficients coding such a polynomial \(p(x)\) can be infeasible (even if the cardinality of \(\mathbb{F}_p\) is constant). For this reason, we represent polynomials by arithmetic circuits. The degree of an arithmetic circuit is defined inductively: the degree of a constant is 0, \(\deg(x_i) = 1\), \(\deg(\sum_i C_i) = \max\{\deg(C_i)\}\), \(\deg(\prod_i C_i) = \sum_i \deg(C_i)\) where \(C_i\)'s are arithmetic circuits.

**Theorem 6.2** (Approximation by low-degree polynomials). For any \(d, S \geq dW PHP(\text{PV}) + BB(\Sigma^b_2)\) proves: for each \(0 < \ell \in \text{LogLog}\), prime \(p \in \text{Log}\), \(\epsilon^{-1} \in \text{Log}\), each depth \(d\) size \(s \in \text{Log}\) circuit \(C\) with \(n\) inputs and \(\text{MOD}_p\) gates, there is an arithmetic circuit of degree \((p - 1)\ell^d\) representing a polynomial \(p(x) \in \mathbb{F}_p[x_1, \ldots, x_n]\) such that

\[
\Pr_{x \in \{0,1\}^n} [p(x) \neq C(x)] \leq \epsilon (1/2^{\ell-1} + \epsilon).
\]

**Proof.** As \(p \in \text{Log}\), applying Fermat's little theorem, \(\text{MOD}_p\) with \(m \leq s\) inputs can be computed by an arithmetic circuit of degree \(p - 1\),

\[
\text{MOD}_p(x_1, \ldots, x_m) = 1 - \left(\sum_{i=1}^{m} x_i\right)^{p-1} \pmod{p}.
\]

We want to define polynomials approximating also Boolean connectives. First, observe that for any nonzero \(x \in \{0,1\}^m\), \(\Pr_{S \subseteq [m]} [\sum_{i \in S} x_i = 0 \mod p] \leq 0.5\). Consequently,

\[
\Pr_{x \in \{0,1\}^m, S_1, \ldots, S_\ell \subseteq [m]} \left[ \text{OR}(x_1, \ldots, x_m) \neq 1 - \prod_{j=1}^{\ell} \left(1 - \left(\sum_{i \in S_j} x_i\right)^{p-1}\right) \right] \leq 0 \frac{1}{2^\ell}.
\]

By an averaging argument (Proposition 4.2), we can fix some sets \(S_1, \ldots, S_\ell \subseteq [m]\) preserving the probability \(\leq \epsilon\) \(1/2^{\ell-1}\) (trading a factor \(1/2\) for the error in approximate counting).

Analogously, there are sets \(S_1, \ldots, S_\ell\) such that

\[
\Pr_{x \in \{0,1\}^m} \left[ \text{AND}(x_1, \ldots, x_m) \neq \prod_{j=1}^{\ell} \left(1 - \left(\sum_{i \in S_j} (1 - x_i)\right)^{p-1}\right) \right] \leq \epsilon \frac{1}{2^{\ell-1}}.
\]
Given a circuit $C(x)$ of depth $d$ and size $s$ we now construct an arithmetic circuit of degree $((p - 1)\ell)^d$ representing polynomial $p(x)$ by replacing NOT$(x)$ gates by $1 - x$ and the other gates by their approximating polynomials of degree $(p - 1)\ell$ described above. This is possible because by Lemma 4.2 the probabilities of errors on the respective gates of $C$ can be expressed by $\Sigma^b_2$ formulas with extra error $3\epsilon$ so we can collect sets $S_1, \ldots, S_l$ for all gates with $BB(\Sigma^b_2)$ and use the resulting sequence in the inductive construction of $p(x)$. The fact that $p(x)$ errs in computing $C(x)$ with probability $\leq \epsilon\frac{s}{2^\ell - 1} + 3\epsilon$ is witnessed by mapping $z = z_0\left(\frac{1}{2^\ell - 1} + 3\epsilon\right)2^n + r < s\left(\frac{1}{2^\ell - 1} + 3\epsilon\right)2^n$, $r < \left(\frac{1}{2^\ell - 1} + 3\epsilon\right)2^n$ to $B(r)$ where $B$ is the circuit witnessing the probability of error on the $(z_0 + 1)$th gate. The collection of circuits $B$ applied in the last step is also $BB(\Sigma^b_2)$.

To derive an $AC^0[p]$ lower bound, one usually proceeds further by showing that any polynomial approximating $MOD_q$ with high probability must have degree $\Omega(n^{1/2})$. The simplest proof of this theorem is obtained by comparing the number of all functions on $n$ variables to the number of low-degree polynomials. As this argument is infeasible, we reproduce it on functions with only $\log O(1)$ inputs. This results in a weaker degree lower bound which, however, still suffices for an $AC^0[p]$ lower bound.

**Theorem 6.3** (Degree lower bound). For any $d$ and primes $p \neq q$, there is an $n_0$ such that $APC_1$ proves: if $n_0 < 2^{\log^3 n}$, $\epsilon^{-1} \in \text{Log}$, every arithmetic circuit representing a polynomial $p(x) \in \mathbb{F}_p[x_1, \ldots, x_n]$ such that

$$\Pr_{x \in \mathbb{F}_p^n}[p(x) \neq MOD_q(x_1, \ldots, x_n)] \leq \epsilon\frac{1}{5q2^q}$$

must have degree $\geq \log^d n$.

**Proof.** If $p \neq q$ are primes, then $p^{q-1} = 1 \mod q$ and the field $\mathbb{F}_{p^{q-1}}$ contains (a multiplicative subgroup of order $p^{q-1} - 1$ and) the $q$-th root of unity $\omega \neq 1$, i.e. $\omega^q = 1$. This is trivially $PV_1$ provable because $p, q$ are constant.

Assume that an arithmetic circuit of degree $\log^d n$ fails to compute $MOD_q$ with probability $\leq \epsilon\frac{1}{5q2^q}$. Using the substitution $y = \frac{x-1}{\omega-1}$ (which maps $\omega \mapsto 1$ and $1 \mapsto 0$) we can construct arithmetic circuits $p_i(x_1, \ldots, x_{n-q})$ of degree $\log^d n$ such that for $x \in \{\omega, 1\}^n$,

$$p_i(x) = 1 \text{ if } \prod_{j=1}^{n-q} x_j = \omega^i \text{ and } p_i(x) = 0 \text{ otherwise, with probability } \geq \epsilon - 1/4q.$$  

Then the polynomial $p'(x_1, \ldots, x_{n-q}) = \sum_{i=0}^{q-1} p_i \omega^i$ of degree $\log^d n$ satisfies $p'(x) = \prod_{i=1}^{n-q} x_i$ for $x \in \{\omega, 1\}^n$ with probability $\geq \frac{3}{4}$. Let $m = \log^3 n$. By an averaging argument fix $a \in \{\omega, 1\}^{n-q-m}$ and $S \subseteq \{\omega, 1\}^m$, $|S| \geq 2^m$ such that $p'(x, a) = \prod_{i=1}^m x_i \prod_{i=1}^{n-q-m} a_i$ for $x \in S$. Define $p''(x) := p'(x, a)(\prod_{i=1}^{n-q-m} a_i)^{-1}$.

Now, consider an arbitrary function $f : \{\omega, 1\}^m \to \mathbb{F}_{p^{q-1}}$. We can express $f$ as

$$f(x) = \sum_{y \in \{\omega, 1\}^m} f(y) \prod_{i=1}^m \frac{2x_iy_i - (1 + \omega)(x_i + y_i) + 1 + \omega^2}{(1 - \omega)^2} = \sum_{y \in \{\omega, 1\}^m} f(y) \prod_{i=1}^m \frac{x_i y_i + t_{i,1} + t_{i,2}}{(1 - \omega)^2}$$

equals 1 if $x = y$ and 0 otherwise.
where $t_{i,1} = 2y_i - (1 + \omega)$ and $t_{i,2} = -(1 + \omega) y_i + 1 + \omega^2$. Since, for $x \in S$,

$$
\prod_{i=1}^{m} (x_i t_{i,1} + t_{i,2}) = \sum_{T \subseteq [m], |T| \leq m} \prod_{i \in [m] \setminus T} x_i t_{i,1} \prod_{i \in [m] \setminus T} t_{i,2} + p''(x) \sum_{T \subseteq [m], |T| > m} \prod_{i \in [m] \setminus T} t_{i,1} \prod_{i \in [m] \setminus T} t_{i,2} x_i^{q-1}
$$

and $x_i^{q-1} = \sum_{z \in \{0,1\}} z^{q-1} x_i^z (1 + \omega)^{(x_i + z) + 1} + \omega^2 (1 - \omega)^2$, we conclude that $f$ can be defined by a polynomial of degree $\lceil \frac{m}{2} \rceil + m^{1/3} + 1$. Note that the arithmetic circuit representing polynomial $p''(x)$ can be expanded to the sum of $\leq 2^m \in \text{Log}$ monomials so the polynomial representing $f$ can be coded by the sequence of its coefficients. By Proposition 5.2, the number of such polynomials is $\leq 0$

$$
|\mathbb{F}_{p^{d-1}}| \sum_{i=0}^{\lfloor m/2 \rfloor + m^{1/3} + 1} \binom{m}{i} < |\mathbb{F}_{p^{d-1}}|(5/9)2^m
$$

while the number of all functions $f : S \to \mathbb{F}_{p^{d-1}}$ is $\geq 0 |\mathbb{F}_{p^{d-1}}|(2/3)2^m$.

**Corollary 6.2.** For any $d$ and primes $p \neq q$, there is an $n_0$ such that APC$_1$ proves: if $2^{d \log^{3d} n} \in \text{Log}$ and $n > n_0$, no depth $d$ circuit with MOD$_p$ gates and size $n^{\log n}$ computes MOD$_q(x, \ldots, x_n)$.

**Proof.** As the statement we want to prove in APC$_1$ is $\forall \Sigma^b_1$, by Lemma 4.1, we are free to work in $S^b_1 + dWPHP(PV) + B^b \Sigma_2$. Let $C$ be a circuit with MOD$_p$ gates, depth $d$, and size $s(n) \in \text{Log}$ computing MOD$_q(x_1, \ldots, x_n)$. By Theorem 6.2 with $\ell = 12q + \log s(n)$, there is an arithmetic circuit representing a polynomial $p(x) \in \mathbb{F}_p[x_1, \ldots, x_n]$ of degree $((12q + \log s(n))(p - 1))^d$ such that for $\epsilon \leq 1 / (10q2^q s)$,

$$
\Pr_{x} [p(x) \neq C(x)] \leq 1 / 5q2^q.
$$

By Theorem 6.3, $((12q + \log s(n))(p - 1))^d \geq \log^{3d} n$.

**Corollary 6.3.** For any $d$ and primes $p \neq q$, there are $n_0, k_0$ and $w, h \in PV$ such that EF has $p$-size proofs of $n^{O(\log^{3d-1} n)}$-size tautologies

$$
\bigvee_{y \in \{0,1\}^{k_0 \log^{9d} n}} C_h(y) \neq f(y) \to \text{lb}_w(\text{MOD}_q, \text{AC}_d^0[p], n^{\log n})
$$

where $f$ is a Boolean function with $k_0 \log^{9d} n$ inputs represented by $2^{k_0 \log^{9d} n}$ variables, $w$ is a $p$-time witnessing function with an access to $f$, and $C_h$ is a circuit of size $2^{(k_0 \log^{9d} n)/2}$ generated by $h$ on the inputs of $w$.

WF has $p$-size proofs of tautologies $\text{tt}(\text{MOD}_q, \text{AC}_d^0[p], n^{\log n})$.  

18
Proof. Proceed as in Corollary 6.1.

A weakness of Theorem 6.3 and Corollary 6.2 is in the assumption $2^{\log^{O(1)} n} \in \text{Log}$. This results in $n^{O(\log n)}$-size EF proofs in Corollary 6.3. If $n^k$-size $\text{AC}_0^d[p]$ lower bounds were obtained assuming just $n \in \text{Log}$, we would get $n^{O(1)}$-size EF proofs $\forall y \in \{0,1\}^k_{\log n} C_h(y) \neq f(y) \rightarrow \text{lb}_w(\text{MOD}_q, \text{AC}_d^0[p], n^k)$ for some constant $k_0$. In Section 7 we give a quasi-polynomial algorithm generating WF proofs of $\text{lb}_{A_n}(\text{MOD}_q, \text{AC}_d^0[p], n^{\log n})$ for some $A_n$.

6.3 Monotone circuits

Theorem 6.4. There is an $n_0$ such that $\text{APC}_1$ proves: for any $n > n_0$ and $k \leq n^{1/4}$ such that $n^k \in \text{Log}$, no monotone circuit of size $2^{\sqrt{k}}$ with $\binom{n}{2}$ inputs accepts exactly (the adjacency matrices of) the $n$-vertex graphs containing a clique of size $k$.

Proof. We follow the presentation from [1]. The only difference is that we need to observe that all surjections witnessing the estimated probabilities can be constructed in $\text{APC}_1$.

Denote by $C_S$ a function on $\binom{n}{2}$ inputs which outputs 1 on a graph $G$ if and only if $S$ is a clique in $G$. Let $P$ be a set of all graphs containing a clique on some $K \subseteq [n]$ of size $k$ and no other edges, and let $N$ be the multiset of all graphs $G_c$ given by functions $c : [n] \rightarrow [k-1]$ so that $G_c$ has an edge between the vertex $i$ and $j$ if and only if $c(i) \neq c(j)$. Further, for a $p$-time predicate $A$, let $\Pr_{G \in P}[A(G)] \geq_p 0$ denote $\{G; G \in P \cap A\} \geq_p p(\binom{n}{k})$ and let $\Pr_{G \in N}[A(G)] \geq_p 0$ denote $\{c : [n] \rightarrow [k-1]; G_c \in A\} \geq_p p(k-1)^n$.

Claim 6.2. There is an $n_0$ such that $\text{PV}_1$ proves: if $n_0 < n \in \text{Log}, k \leq n^{1/4}, n^k \in \text{Log}$ and $S \subseteq [n]$, then $\Pr_{G \in N}[C_S(G) = 1] \geq_p 0.9$ or $\Pr_{G \in P}[C_S(G) = 1] \geq_p 0.9$.

Claim 6.2 is derived by considering two cases. If $|S| \leq l := \sqrt{k-1}/10$ then the probability that a random $f : S \rightarrow [k-1]$ contains a collision is $\leq \left(\binom{|S|}{2}\right)^{1/[k-1]} < 0.1$. Since $k^{l|S|} \in \text{Log}$, it is feasible to list all functions $f : S \rightarrow [k-1]$, what allows us to construct also the surjection witnessing $\Pr_{G \in N}[C_S(G) = 1] \geq_p 0.9$. If $|S| > l$, the probability that $S \subseteq K$ for a random set $K$ of size $k$ is $\leq_p (n-1)/\binom{n}{k} < (k/n)^l < n^{-1/20}$ for sufficiently big $n$. Again, as $n^k \in \text{Log}$, we can count the probability precisely.

Claim 6.3. There is an $n_0$ such that $S_2^1 + dWPHP(\text{PV}_1)$ proves: if $n^k, \epsilon^{-1} \in \text{Log}$, then for any monotone circuit $C$ of size $s \leq 2^{\sqrt{k}}$ where $k \leq n^{1/4}$, there exist $m < n^{\sqrt{k}/20}$ sets $S_i$ of size $\leq l$ such that

$$\Pr_{G \in P}[\bigvee_i C_{S_i}(G) \geq C(G)] \geq_p 0.9$$

$$\Pr_{G \in N}[\bigvee_i C_{S_i}(G) \leq C(G)] \geq_p 0.9$$

where empty $\bigvee_i C_{S_i}(G)$ with $m = 0$ is defined as the constant 0.
Claims 6.2 and 6.3 imply Theorem 6.4: The statement we want to prove in APC_1 is \( \forall \Sigma^b \) so by Lemma 4.1 we are free to work in \( S^2 + dWPHP(\Sigma^1) \). If \( m = 0 \), then \( \Pr_{G \in P}[\bigvee_i C_{S_i}(G) \geq C(G)] \geq 0 \) forces \( C \) to err on some \( G \in P \) by Proposition 4.2 1.\( \text{ii} \). Otherwise, using Proposition 4.2 1.\( \text{iv} \) and 1.\( \text{ii} \), \( \Pr_{G \in N}[\bigvee_i C_{S_i}(G) = 1] \geq 0 \) and \( \Pr_{G \in N}[\bigvee_i C_{S_i}(G) \leq C(G)] \geq c \geq 0 \) imply that \( C \) errs on some \( G \in N \).

In the rest of the proof we derive Claim 6.3.

Let \( l = \sqrt{k - 1}/10, p = 10\sqrt{k \log n} \) and \( m = (p - 1)^{l}! \in \text{Log} \). The gates of the circuit \( C \) compute functions \( f_1, \ldots, f_s \) from \( \{0,1\}^{|S|} \) to \( \{0,1\} \). We will approximate \( f_1, \ldots, f_s \) by functions \( \tilde{f}_1, \ldots, \tilde{f}_s \) such that each \( \tilde{f}_k \) is an \((l,m)\)-function: i.e. a disjunction of at most \( m \) functions \( C_{S_i} \) with \( |S_i| \leq l \).

The functions \( \tilde{f}_1, \ldots, \tilde{f}_s \) are constructed by induction. For \( 1 \leq k \leq s \), if \( f_k \) is an input, then \( \tilde{f}_k = f_k \). If \( f_k = f_{k'} \lor f_{k''} \), then \( \tilde{f}_k = \tilde{f}_{k'} \lor \tilde{f}_{k''} \), and if \( f_k = f_{k'} \land f_{k''} \), then \( \tilde{f}_k = \tilde{f}_{k'} \land \tilde{f}_{k''} \) where the operations \( \sqcup, \sqcap \) are defined as follows.

- **f \sqcup g:** for \((m, l)\)-functions \( f = \bigvee_{i=1}^{m} C_{S_i}, g = \bigvee_{i=1}^{m} C_{T_i}, \) let \( h = \bigvee_{i=1}^{2m} C_{Z_i} \), where \( Z = S \) and \( Z_{m+i} = T_j \) for \( 1 \leq i,j \leq m \). Next we make \( h \) into an \((m, l)\)-function: as long as there are more than \( m \) distinct sets, find \( p \) subsets \( Z_{i_1}, \ldots, Z_{i_p} \) that form a sunflower, i.e. there exists a set \( Z \) such that for \( j \neq j', Z_{i_j} \cap Z_{i_j'} = \emptyset \). Replace \( C_{Z_{i_1}}, \ldots, C_{Z_{i_p}} \) in \( h \) by \( C_Z \). Once we obtain an \((m, l)\)-function \( h' \), we define \( f \lor g \) to be \( h' \). By the Sunflower lemma (below) we will not get stuck.

- **f \sqcap g:** for \((m, l)\)-functions \( f = \bigvee_{i=1}^{m} C_{S_i}, g = \bigvee_{i=1}^{m} C_{T_i}, \) let \( h = \bigvee_{1 \leq i,j \leq m} C_{S_i \cup T_j} \). Discard from \( h \) every \( C_Z \) with \( |Z| > l \) and reduce the number of disjuncts to \( m \) by applying the Sunflower lemma as above.

**Lemma 6.2 (Sunflower lemma).** \( \Sigma^1 \) proves: let \( Z \) be a collection of distinct sets \( |Z| = l + 1 \) at most \( l \) with \( |Z| \in \text{Log} \). If \( |Z| > (p - 1)^l \), then there exist \( p \) sets \( Z_1, \ldots, Z_p \in Z \) and a set \( Z_0 \) such that \( Z_i \cap Z_j = Z_0 \) for \( 1 \leq i \neq j \leq p \).

Lemma 6.2 is proven by induction on \( l \). The case \( l = 1 \) is trivial since distinct sets of size \( 1 \) form a sunflower with an empty center. For \( l > 1 \), let \( M \) be a set of disjoint sets from \( Z \) such that \( \bigcup_{N \in M} N \cap Z_i \neq \emptyset \) for every \( Z_i \in Z \). We can assume that \( |M| < p \) since otherwise \( M \) is a sufficiently large sunflower. As \( |\bigcup_{N \in M} N| \leq (p - 1)^l \), there is an \( x \) that appears in at least \( 1/(p - 1) \) of all sets in \( Z \). Let \( Z_1, \ldots, Z_t \) be the sets containing \( x \). Note that \( t > (p - 1)^{l - 1}(l - 1)! \). Thus, by the induction hypothesis, there are \( p \) sets among \( Z_i \) forming a sunflower. Adding back \( x \) we get the desired sunflower among the original sets. This completes the proof of Lemma 6.2.

Now we show that the operations \( \sqcup \) and \( \sqcap \) approximate \( \lor \) and \( \land \), respectively:

- \( \Pr_{G \in P}[f \lor g < f \lor g] \leq 0 \)

If \( Z \subseteq Z_i \), then for any \( G \), \( C_Z(G) = 0 \) implies \( C_{Z_i}(G) = 0 \), and therefore, \( \sqcup \) cannot introduce any “false 0”.

20
• \( \Pr_{G \in P}[f \cap g < f \wedge g] \leq \epsilon 1/(10s) \)

A graph \( G \in P \) is a clique over some set \( K \). Thus, \( C_{S_i}(G) \wedge C_{T_j}(G) = 1 \iff S_i, T_j \subseteq K \iff C_{S_i \cup T_j}(G) = 1 \). This means that \( f \wedge g = \bigvee_{1 \leq i,j \leq m} C_{S_i \cup T_j} \). Discarding \( C_Z \) with \(|Z| > l \) might introduce “false 0s”. However, by Claim 6.2, for any \( Z \) with \(|Z| > l \), \( \Pr_{G \in P}[C_Z(G) = 1] \leq n^{-\sqrt{k}/20} < 1/(10sm^2) \) for big enough \( n \). As we discard at most \( m^2 \) such sets and applying the Sunflower lemma cannot introduce any “false 0”, the inequality follows. The last step collects \( \leq m^2 \) circuits. This is just \( BB(\Sigma^b_1) \) collection because all the respective probabilities can be counted precisely and the circuits witnessing them are efficiently invertible.

\( BB(\Sigma^b_1) \) can be used again to compose the circuits witnessing the probability of error on the respective gates of \( C \) and conclude that \( \Pr_{G \in P}[\bigvee_i C_{S_i}(G) < C(G)] \leq 0.1 \) for some \( \leq m \) sets \( S_i \) of size \( \leq l \). As \( n^k \in \text{Log} \), the circuits count the probability precisely and can be turned into witnessing of \( \Pr_{G \in P}[\bigvee_i C_{S_i}(G) \geq C(G)] \geq 0.9 \).

It remains to show that a similar approximation holds for graphs in \( N \):

• \( \Pr_{G \in N}[f \cup g > f \wedge g] \leq \epsilon 1/(10s) \)

Replacing \( C_{Z_1}, \ldots, C_{Z_p} \) with \( C_Z \) can introduce a “false 1” if \( C_Z(G) = 1 \) while \( C_{Z_i}(G) = 0 \) for every \( i \). Each \( G \in N \) is specified by a function \( c : [n] \to [k-1] \). Thus, we get a “false 1” only if \( c \) is one-to-one on \( Z \) but not one-to-one on \( Z_i \)’s. Denote this event by \( A \). For every \( i \), since \(|Z_i| \leq l \), \( \Pr_c[c \text{ is not one-to-one on } Z_i \setminus Z] \leq 1/2 \). As \( Z_i \setminus Z \)’s are disjoint sets, \( \Pr_{G \in N}[A] \leq 2^{-p} < 1/(10sm) \) for big enough \( n \). We apply the reduction step at most \( m \) times so the inequality follows by Proposition 4.2 (Disjoint union).

• \( \Pr_{G \in N}[f \cap g > f \wedge g] \leq \epsilon 1/(10s) \)

Since \( C_{S \cup T}(G) = 1 \) implies \( C_S(G) = 1 \) and \( C_T(G) = 1 \), a “false 1” can be introduced only when we apply the Sunflower lemma. We bound the probability of such error in the same way as in the previous case.

Applying Proposition 4.2 (Disjoint union), the estimated probabilities can be used to conclude \( \Pr_{G \in N}[\bigvee_i C_{S_i}(G) > C(G)] \geq 2\epsilon 0.1 \) for some \( \leq m \) sets \( S_i \). Hence, by Proposition 4.2 1.iii), \( \Pr_{G \in N}[\bigvee_i C_{S_i}(G) \leq C(G)] \leq 3\epsilon 0.9 \).

It is not hard to see that Theorem 6.4 scales down so that poly-size lower bounds are provable assuming only \( n \in \text{Log} \). More precisely, for every \( k \) there is an \( n_0 \) such that \( PV_1 \) proves that for any \( n_0 < n \in \text{Log} \), no monotone circuit of size \( n^k \) with \( \binom{n}{2} \) inputs accepts exactly the \( n \)-vertex graphs containing a clique of size \( 20k^3 \). Denote by \( \text{lb}_w(\text{Clique}(n, 20k^3), \text{monotone}, n^k) \) the propositional translation of this \( \Sigma^b_1 \) formula witnessed by a p-time function \( w \). Similarly as in Corollary 6.1 we get
Corollary 6.4. For any $k$ there are $n_0, k_0$ and $w, h \in \mathsf{PV}$ such that $\mathsf{EF}$ has p-size proofs of tautologies

$$\bigvee_{y \in \{0,1\}^{k_0 \log n}} C_h(y) \neq f(y) \rightarrow \text{lb}_w(\text{Clique}(n, 20k^3), \text{monotone}, n^k)$$

where $f$ is a Boolean function with $k_0 \log n$ inputs represented by $2^{k_0 \log n}$ variables, $w$ is a p-time witnessing function with an access to $f$, and $C_h$ is a circuit of size $2^{(k_0 \log n)/2}$ generated by $h$ on the inputs of $w$.

Tautologies $\mathsf{tt}(\text{Clique}(n, 20k^3), \text{monotone}, n^k)$ have p-size EF proofs.

A complication in improving Theorem 6.4 to $\mathsf{PV}_1$ is that it is unclear how to efficiently generate some $G$ satisfying $\bigvee_i C_{S_i}(G) \leq C(G)$.

7 Natural proofs

7.1 Naturalization of $\mathsf{AC}^0$ and $\mathsf{AC}^0[p]$ lower bound

(Automatizability of EF on $\mathsf{AC}^0$ and $\mathsf{AC}^0[p]$ lower bounds)

Razborov and Rudich [30] showed that the known circuit lower bounds on explicit Boolean functions actually work for a random function with high probability. Moreover, there are p-size circuits recognizing truth-tables of the functions for which the lower bounds work.

We are interested in a more constructive version of circuit lower bounds, so we formalize their naturalization on functions $f$ given by sequences of input/output tuples $(x, f(x))$, not necessarily by the whole truth-table of $f$. That is, instead of proving formulas $\mathsf{tt}(f, n^k)$ we want to prove $\mathsf{lb}_{A_n}(f, n^k)$. We present the formalization already on propositional level. As a consequence, in case of Razborov-Smolensky's method we obtain short WF proofs of formulas $\mathsf{lb}_{A_n}(\text{MOD}_q, \mathsf{AC}^0_d[p], n^k)$ for some small sets $A_n$ and $p \neq q$, thus getting rid of the implicational form of Corollary 6.3.

To further motivate the quest for automatizing the provability of formulas $\mathsf{lb}_{A_n}(f, s)$ consider a basic learning task. Given bits $f(x_1), \ldots, f(x_k)$ for $k$ $n$-bit strings $x_1, \ldots, x_k$ we want to predict the value of $f$ on a new input $x_{k+1} \in \{0,1\}^n$. Predicting $f(x_{k+1})$ makes sense only if the minimal circuit $C$ coinciding with $f$ on $x_1, \ldots, x_k$ determines the value $f(x_{k+1})$. Say that the size of the minimal circuit $C$ is $s$. Then the task to predict the value $f(x_{k+1})$ can be formulated as the task to prove an $s$-size circuit lower bound of the form $\bigvee_{i=1,\ldots,k} C(x_i) \neq f(x_i) \lor C(x_{k+1}) \neq \epsilon$ for $\epsilon \in \{0,1\}$. A more sophisticated connection between circuit lower bounds and learning algorithms was recently demonstrated in [5].

Our naturalization of $\mathsf{AC}^0$ lower bounds contains an extra assumption stating that a function $g_1$ with $m$ inputs is hard on average for circuits of size $2^{m/4}$, i.e. no circuit of size $2^{m/4}$ computes $g_1$ on $\geq 2^m/2 + 2^{(1-1/4)m}$ inputs. The assumption might be reducible to
the worst-case hardness of $g_1$ but we omit a deeper analysis of the approximate counting and hardness amplification in $\text{PV}_1$. In fact, the proof of Theorem 7.1 already asks for a slightly deeper knowledge of approximate counting so we give just a sketch. Further, it is also unclear for how many functions the lower bound actually works. These issues do not arise in the naturalization of $\text{AC}^0[p]$ lower bounds in Theorem 7.2.

**Theorem 7.1.** For any $k, d$, there are constants $k_0, k_1, b$ such that

1. There is a probabilistic $p$-time algorithm which for any string of the length $n$ with probability $\geq 3/4$ generates (i.e. lists all elements of) a set $S_n$ of restrictions of $n$ variables leaving at least $n^{1/b}$ variables unassigned.

2. There is a $p$-time algorithm which given tuples $(x, f(x))$, where $x \in A_n \subseteq \{0, 1\}^n$, $f(x) \in \{0, 1\}$, $n$ sufficiently big, such that for any $\rho \in S_n$ there are $x_1, x_2 \in A_n$ extending $\rho$ and satisfying $f(x_1) \neq f(x_2)$, outputs an $\text{EF}$ proof of

$$C_{h_1} \not\Rightarrow g_1 \land \bigvee_{y \in \{0, 1\}^{k_0 \log n}} C_{h_0}(y) \neq g_0(y) \rightarrow \text{lb}_{A_n}(f, \text{AC}^0_d, n^k)$$

where $g_0$ is a Boolean function with $k_0 \log n$ inputs represented by $2^{k_0 \log n}$ variables, $g_1$ is a Boolean function with $k_1 \log(n \log n)$ inputs represented by $2^{k_1 \log(n \log n)}$ constants which is hard on average for circuits of size $2^{(k_1 \log(n \log n))/4}$, $C_{h_0}$ is a circuit of size $2^{(k_0 \log n)/2}$ generated by a $p$-time algorithm $h_0$ on $g_0, g_1$ together with the variables of $\text{lb}_{A_n}$, $C_{h_1}$ is a circuit of size $2^{(k_1 \log(n \log n))/4}$ generated by a $p$-time algorithm $h_1$ on $g_0, g_1$ together with the variables of $\text{lb}_{A_n}$, and $C_{h_1} \not\Rightarrow g_1$ is a propositional formula stating that $C_{h_1}$ does not compute $g_1$ on $\geq 2^{(k_1 \log(n \log n))/2 + 2(1-1/4)(k_1 \log(n \log n))}$ inputs.

**Proof (Sketch).** The proof of Theorem 6.1 shows that for every $k, d$, $\text{APC}_1$ proves: if $n$ is sufficiently big, then for any $n^k$-size circuit $C_n$ of depth $d$ there is an equivalent $n^{2k}$-size circuit $C'_n$ such that for some constant $b$, a random sequence of restrictions $\rho_1, \ldots, \rho_{2d}$, where $\rho_{2i+1} \in R_{1/n^{1/2}}, \rho_{2i} \in R_{1/n^{1/4}}$, leaves $< n^{1/b}$ variables unassigned or makes $C'_n|\rho_1 \ldots \rho_{2d}$ depend on $> b$ inputs with probability $\leq 2^{2d} 2d \frac{d}{n^{2d}}$. Applying one more restriction $\rho_0 \in R_{1/n^{1/2}},$

$$\Pr_{\rho_0, \ldots, \rho_{2d}}[C'_n|\rho_1 \ldots \rho_{2d}\rho_0 \text{ depends on } 0 \text{ inputs and } \geq n^{1/8b} \text{ inputs remain unassigned}] \leq (2d+1)e$$

$$1 - \left( \frac{2d}{n^{2k}} + \frac{b}{n^{1/2}} + \frac{1}{n^{2k/b}} \right).$$

In $\text{APC}_1$ the probability is approximated by generating random restrictions $\rho = \rho_1 \ldots \rho_{2d}\rho_0$ using a Nisan-Wigderson generator with a seed of the length $O(\log(n \log n))$, cf. [14, Theorem 2.7]. The Nisan-Wigderson generator is based on a function $g_1$ with $k_1 \log(n \log n)$ inputs which is hard on average for circuits of size $2^{k_1 \log(n \log n)/4}$. Therefore, by Lemma

23
4.1, we can generate in p-time EF proofs of tautologies stating that if $C_{h_1} \not\sim g_1 \land \bigvee_{y \in \{0,1\}^{k_0 \log n}} C_{h_0}(y) \neq g_0(y)$ where both $g_0, g_1$ are given by free variables, then a p-time algorithm with access to $g_1$ generates a restriction $\rho$ collapsing $C'_n$ to a constant while leaving $\geq n^{1/8b}$ inputs unassigned. The restrictions $\rho$ are generated by an algorithm which does not depend on $C'_n$. As a random boolean function on $k_1 \log(n \log n)$ inputs is hard on average for circuits of size $2^{(k_1 \log(n \log n))/2}$ with probability $\geq 3/4$, this defines the set $S_n$ and yields a p-time algorithm generating EF proofs of tautologies stating that if $C_{h_1} \not\sim g_1 \land \bigvee_{y \in \{0,1\}^{k_0 \log n}} C_{h_0}(y) \neq g_0(y)$ where now $g_1$ are fixed constants, then any $n^{2k}$-size circuit $C'_n$ is collapsed by some restriction $\rho \in S_n$. Hence, any function $f$ which is not collapsed by any restriction $\rho \in S_n$ when considering inputs from $A_n$ extending $\rho$ satisfies $\text{lb}_{A_n}(f, \text{AC}^0_d, n^k)$.

If $A_n = \{0,1\}^n$, i.e. the whole truth-table of $f$ is given as input, we get a p-time algorithm generating WF proofs of tautologies $\text{tt}(f, \text{AC}^0_d, n^k)$ for $2^{2^{\omega} - O(n)}$ functions $f$.

**Corollary 7.1.** For any $k, d$, there is $b$ and a p-time algorithm which given the truth-table of a function $f : \{0,1\}^n \to \{0,1\}$, $n$ sufficiently big, such that for any restriction $\rho$ leaving at least $n^{1/b}$ variables unassigned there are $x_1, x_2 \in \{0,1\}^n$ extending $\rho$ with $f(x_1) \neq f(x_2)$, outputs a WF proof of $\text{tt}(f, \text{AC}^0_d, n^k)$. Analogously, EF proofs can be generated in $2^{O(n \log n)}$-time.

**Proof.** Proceed as in the proof of Theorem 7.1 with $n \in \text{LogLog}$ resp. $2^{O(n \log n)} \in \text{Log}$ and the set $R_n$ being the set of all restrictions leaving $\geq n^{1/8b}$ variables unassigned. □

**Theorem 7.2.** For any $d$ and primes $p \neq q$, there is a constant $k$ and an $n^{O(m)}$-time algorithm, $m = \log^{10d} n$ which

- given tuples $\left\langle x, f(x) \right\rangle$, where $x \in A_n \subseteq \{0,1\}^n$, $f(x) \in \{0,1\}$, such that for some restriction $\rho$ leaving $m + q$ variables unassigned, $A_n$ contains all $x \in \{0,1\}^n$ extending $\rho$, and for the multilinear polynomial $p(x)$ satisfying $p(x) = f'(x)$ where $x \in \{\omega,1\}^{m+q}$, $\omega \neq 1$ is the $q$-th root of unity in $\mathbb{F}_{p^q-1}$ and $f'$ is $f|\rho$ under the inputwise substitution $y = \frac{x-1}{\omega-1}$, the $2^{m+q} \times 2^{m+q}$ matrix $P := \{P(x)\}_{x,p}$ where $x \in \{\omega,1\}^{m+q}$ and $P$ is a term from

  \[
  \left\{ \prod_{i \in T} x_i \right\}_{T \subseteq [m+q], |T| \leq \frac{m+q}{2}} \cup \left\{ p(x) \prod_{i \in [m+q] \setminus T} x_i \right\}_{T \subseteq [m+q], |T| > \frac{m+q}{2}}
  \]

  has rank $\geq \frac{3\omega m}{4}$,

- outputs an EF proof of

  \[
  \bigvee_{y \in \{0,1\}^{km}} C_h(y) \neq g(y) \to \text{lb}_{A_n}(f, \text{AC}^0_d[p], n^{\log n})
  \]

24
where \( g \) is a Boolean function with \( km \log n \) inputs represented by \( 2^{km \log n} \) variables, and \( C_h \) is a circuit of size \( 2^{km \log n/2} \) generated by an \( n^{O(m)} \)-time algorithm \( h \) on \( g, f \) and the variables of \( \text{lb}_{A_n} \). Moreover, the algorithm outputs a \( \text{WF} \) proof of

\[
\text{lb}_{A_n} (f, \text{AC}_d^0[p], n^{\log n}).
\]

Note that for \( f \) being the \( \text{MOD}_q \) function, it is easy to construct a suitable set \( A_n \) so that Theorem 7.2 gives a quasi-polynomial algorithm generating \( \text{WF} \) proofs of \( \text{lb}_{A_n}(\text{MOD}_q, \text{AC}_d^0[p], n^k) \) for \( p \neq q \).

**Proof.** We reason in \( S^2_2 + dWPHP(\text{PV}) \). Let a sequence of tuples \( \langle x, f(x) \rangle \) satisfy the assumptions of Theorem 7.2, so \( \rho \) can be found in time \( n^{O(m)} \) and \( f' \) can be expressed by a multilinear polynomial \( p(x) \). If \( f \) can be computed on \( A_n \) by a circuit \( C \) with \( \text{MOD}_p \) gates, depth \( d \) and size \( n^{\log n} \in \text{Log} \), then as in Corollary 6.2 we obtain a polynomial \( p'(x) \) of degree \( ((5 + q + \log^2 n)(p - 1))^d \) such that,

\[
\Pr_{x \in \mathbb{F}_{2^{m+q}}} [p'(x) \neq f(x)] \leq 1/2^{q+4}.
\]

The probability can be counted exactly assuming \( 2^m \in \text{Log} \) so \( BB(\Sigma^b_2) \) is not needed. Consequently, there is a polynomial \( p''(x) \) of degree \( ((5 + q + \log^2 n)(p - 1))^d \) and a set \( S' \subseteq \{\omega, 1\}^{m+q} \) of size \( (1 - 1/2^{q+4})2^{m+q} \) such that \( p''(x) = p(x) \) for \( x \in S' \).

Jeřábek [13, Theorem 4.3.19] showed that a \( \text{PV} \) function \( \text{PV}_{11} \)-provably computes a solution to a system of linear equations over a finite field if one exists, and a basis for the space of solutions of a homogeneous linear system over a finite field. Hence, we can conclude in \( S^2_2 + dWPHP(\text{PV}) \) that if the rank of \( \mathcal{P} \) is \( \geq 3/4 2^m \), all functions \( h : S \to \mathbb{F}_{p^{q-1}} \), where \( S \subseteq \{\omega, 1\}^m \) is a set of size \( \geq 3/2 2^m \), are expressible by a polynomial of degree \( [m/2] + m^{1/3} + 1 \). This is a contradiction (and the only place where it is crucial to apply \( dWPHP(\text{PV}) \)).

Therefore, by Lemma 4.1, we can generate in \( p \)-time \( \text{EF} \) proofs of tautologies stating that if \( \bigvee_{y \in \{0,1\}^m} C_h(y) \neq g(y) \), then a function \( f \) given by tuples \( \langle x, f(x) \rangle \) either satisfies \( \text{lb}_{A_n}(f, \text{AC}_d^0[p], n^{\log n}) \), or the rank of \( \mathcal{P} \) is \( < 3/4 2^m \). Since we assume that \( \mathcal{P} \) has rank \( \geq 3/4 2^m \) we obtain \( \text{EF} \) proofs of \( \bigvee C_h(y) \neq g(y) \rightarrow \text{lb}_{A_n}(f, \text{AC}_d^0[p], n^{\log n}) \).

The \( \text{WF} \) proof is obtained by realizing that the antecedent \( \bigvee C_h(y) \neq g(y) \) has the form of a special axiom in \( \text{WF} \), cf. [12, Definition 2.6], and its variables do not occur in \( \text{lb}_{A_n}(f, \text{AC}_d^0[p], n^{\log n}) \). \( \square \)

Unlike in the case of Theorem 7.1, we can observe that the proofs of \( \text{lb}_{A_n}(f, \text{AC}_d^0[p], n^{\log n}) \) with \( p \neq 2 \) can be generated for many functions \( f \); for at least half of all functions \( f \) defined on \( A_n \) with the property that \( A_n \) contains all \( x \in \{0,1\}^n \) extending some restriction \( \rho \) which leaves \( m + q \) variables unassigned. Here we fix \( q = 2 \).

Specifically, we claim that for any function \( f \) and the multilinear polynomial \( p(x) \) defined as in Theorem 7.2, either the rank of \( \mathcal{P}_0 \) defined as \( \mathcal{P} \) in Theorem 7.2 but with
\[ p(x) \text{ substituted by } r(x) := (\omega - 1)p(x) + 1, \text{ or the rank of } P_1 \text{ defined as } P \text{ with } p(x) \text{ substituted by } r^{q-1}\prod_{i \in [m+q]} x_i, \text{ is } \geq \frac{3}{4}2^{m+q}. \] 

As \( q = 2, \omega = -1, \) polynomials \( r^{q-1}\prod x_i \) represent Boolean functions, and hence, at least half of all functions \( f \) on \( A_n \) are hard.

To see this, identify a set of polynomials \( U \) with the vector space generated by the column vectors of the \( 2^{m+q} \times |U| \) matrix \( \{u(x)\}_{x \in \{0,1\}^{m+q}, u \in U}. \) For a polynomial \( p, \) denote by \( pU \) the set \( \{pu, u \in U\} \) and put \( L := \{(\prod_{i \in T} x_i)_{T \subseteq [m+q], |T| \leq \frac{m+q}{2}}. \) If the dimension \( \dim(L \cup rL) \) of the vector space \( L \cup rL, \) which is equal to the rank of \( P_0, \) is \( < \frac{3}{4}2^{m+q}, \) then

\[
\dim\left(\frac{r^{q-1}\prod x_iL \cup L}{L}\right) \geq \dim\left(\frac{\prod x_iL \cup L}{rL}\right) \\
\geq \dim\left(\frac{\prod x_iL \cup rL \cup L}{(rL \cup L)}\right) \geq \frac{2^{m+q}}{4}
\]

where the first inequality follows because we multiply every row vector in the matrix \( rL \) resp. \( \prod x_iL \cup rL \) by a nonzero constant \((r(x))^{q-1}\) which does not change the dimension of the vector space generated by the row vectors and hence neither the dimension of the column vectors. Therefore, the rank of \( P_1 \) is \( \dim(L \cup r^{q-1}\prod x_iL) \geq \frac{3}{4}2^{m+q}. \)

If \( A_n = \{0,1\}^n, \) we get in particular a \( p \)-time algorithm generating WF proofs of tautologies \( \text{tt}(f, \text{AC}_d[p], n\log n) \) for \( 2^n - O(n) \) functions \( f. \)

### 7.2 Natural proofs barrier

**Theorem 7.3.** For any \( c', d' \geq 1; c, \delta > 0 \) there is an \( m_0 \) such that the theory \( \text{HARD}^A \) proves: given any \( \epsilon^{-1}, 2^{k^\delta} \in \text{Log}, \, \epsilon \leq 1/(18(2^{d'}m)), \) \( m = \lceil k^{\delta/2} \rceil \geq m_0, \) if a circuit \( C \) defines a \( \text{P/poly-natural property useful against circuits of size } (c + 4)m^{(1+2c/\delta)}, \) meaning

1. (Constructivity) \( C \) has \( 2^m \) inputs and size \( 2^{c'm}, \)

2. (Largeness) \( \Pr_x[C_{2^m}(x) = 1] \geq \epsilon 1/2^{d'm}, \)

3. (Usefulness) for \( C_{2^m}(x) = 1, \) \( x \) is a truth-table of a function on \( m \) inputs which is not computable by a circuit of size \( (c + 4)m^{1+2c/\delta}, \)

then no \( ck^c \)-size circuit \( G_k \) defines a strong pseudorandom generator safe against circuits of size \( 2^{k^\delta}, \) meaning that no \( ck^c \)-size circuit \( G_k : \{0,1\}^k \mapsto \{0,1\}^{2k} \) satisfies that for all circuits \( C \) of size \( 2^{k^\delta}, \)

\[
\left| \Pr_x[C(G_k(x)) = 1] - \Pr_y[C(y) = 1] \right| \leq \frac{1}{2^{k^\delta}}.
\]
Proof. Let \( c', d' \geq 1; c, \delta > 0; \epsilon^{-1}, 2^k \in \text{Log}, \epsilon \leq 1/(18(2^d m)) \) and \( m = \lceil k^{3/2} \rceil \). Suppose \( C_{2m} \) defines a \( P/\text{poly} \)-natural property against circuits of size \((c + 4)m^{1+c/\delta}\) and \( G_k : \{0, 1\}^k \mapsto \{0, 1\}^{2k} \) is a \( ck^c \)-size circuit. We will show that there is a circuit \( C \) of size \( 2^{(d'+d_0)m} \) recognizing \( G_k \) with advantage \( > 1/2^{(d'+d_0)m} \) for an absolute constant \( d_0 \).

Let \( G^0, G^1 : \{0, 1\}^k \mapsto \{0, 1\}^k \) be the first and the last \( k \) bits of \( G_k \), respectively. For any \( y \in \{0, 1\}^m \) define \( G^y : \{0, 1\}^k \mapsto \{0, 1\}^k \) by \( G^{y_m} \circ G^{y_{m-1}} \cdots \circ G^{y_1} \) and for \( x \in \{0, 1\}^k \) let \( f(x)(y) \) be the first bit of \( G^y(x) \).

For any fixed \( x \in \{0, 1\}^k \), \( f(x)(y) \) is computable by circuits of size \((c + 4)m^{(1+2c/\delta)}\), more precisely, by \( m \) copies of \( G_k \) with \( m \) circuits of size \( 4k \) choosing between the first resp. last \( k \) bits of \( G_k \). Hence, \( Pr_x[C_{2m}(f(x)) = 1] \approx \epsilon \). As the circuit \( C_{2m} \) of size \( 2^{c'm} \) defining a natural property satisfies \( Pr_z[C_{2m}(z) = 1] \geq 2^m 1/2^{d'm} \), it distinguishes \( f(x) \) from random functions:

\[
Pr_z[C_{2m}(z) = 1] - Pr_x[C_{2m}(f(x)) = 1] \geq 1/2^{d'm} - 3\epsilon. \tag{7.1}
\]

Consider now the binary tree \( T \) of height \( m \). Its internal nodes \( v_1, \ldots, v_{2^m-1} \) are arranged so that if \( v_i \) is a son of \( v_j \), then \( i < j \). The last level of \( T \) contains \( 2^m \) leaves corresponding to the elements of \( \{0, 1\}^m \). Let \( T_i \) be the union of subtrees of \( T \) whose nodes are \( \{v_1, \ldots, v_i\} \) along with all the leaves. For a leaf \( y \), let \( v_i(y) \) be the root of the subtree in \( T_i \) containing \( y \). Denote by \( h(i, y) \) the distance between \( y \) and \( v_i(y) \).

For \( x_{v_i(y)} \in \{0, 1\}^k \), define \( f_{i,m}(y) \) to be the first bit of \( G^{y_m} \circ \cdots \circ G^{y_{m-h(i,y)+1}}(x_{v_i(y)}) \). Given a random assignment \( x_{v_0(y)} \in \{0, 1\}^k \), \( f_{0,m} \) is a random function.

Since \( f_{2^m-1,m} \) is \( f(x) \), by (7.1), for some \( i \),

\[
Pr_{\{x_{v_i(y)}\}}[C_{2m}(f_{i,m}) = 1] - Pr_{\{x_{v_{i+1}(y)}\}}[C_{2m}(f_{i+1,m}) = 1] \geq 1/2^{(d'+1)m} - 3\epsilon/2^m
\]

where \( \{x_{v_i(y)}\} \) is the set of all assignments \( x_{v_j(y)} \in \{0, 1\}^k \) with \( v_j(y) \) a root in \( T_i \).

Fix all \( x_{v_i(y)} \) other than those with \( v_i(y) \in \{v_{i+1}, v', v''\} \) where \( (v', v'') \) are the two sons of \( v_{i+1} \), so that the bias \( 1/2^{(d'+1)m} - 3\epsilon/2^m \) is preserved. The existence of such a fixation \( \text{Fix} \) follows from an application of Proposition 4.2 (averaging), which implies \( Pr_{\{x_{v_i(y)}\} \subseteq \text{Fix}}[C_{2m}(f_{i,m}) = 1] \geq Pr_{\{x_{v_i(y)}\}}[C_{2m}(f_{i,m}) = 1] - 3\epsilon \) and a similar approximation of \( Pr_{\{x_{v_{i+1}(y)}\} \subseteq \text{Fix}}[C_{2m}(f_{i+1,m}) = 1] \geq Pr_{\{x_{v_{i+1}(y)}\}}[C_{2m}(f_{i+1,m}) = 1] - 3\epsilon \) for a circuit of size \( 2^{(d'+d_0)m} \) with a sufficiently big \( d_0 \), distinguishing between \( G_k(x_{v_{i+1}}) \) and \( (x_{v'}, x_{v''}) \).

\[
8 \quad \text{Conclusion}
\]

We showed that \( \text{AC}^0, \text{AC}^0[p] \) and monotone circuit lower bounds are provable in \( \text{APC}_1 \). By Lemma 4.1 our formalizations imply randomized p-time (resp. quasipolynomial-time in case of \( \text{AC}^0[p] \)) algorithms witnessing errors of \( \text{AC}^0, \text{AC}^0[p], \) monotone circuits of small size attempting to compute the corresponding hard function.
If it was possible to derandomize these witnessing algorithms provably in APC₁ we could express AC⁰, AC⁰[p] and monotone circuit lower bounds by Σ₀ᵇ formulas and derive short WF proofs of their propositional translations, thus getting rid of the extra assumption on the hardness of some function in Corollaries 6.1, 6.3, 6.4 at the cost of moving from EF just to WF. In Theorem 7.2 we managed to generate such WF proofs of AC⁰[p] lower bounds in quasi-polynomial time. It seems that quasipolynomial-size WF proofs of tautologies lbₜ(PARITY, AC₀ᵈ[nᵏ]) could be obtained also by formalizing the derandomized switching lemma from [33].

A more challenging problem is a derandomization of AC⁰, AC⁰[p] and monotone circuit lower bounds, that is proving them in the theory PV₁. Eventually, we would like to know if it is possible to derive e.g. AC⁰ circuit lower bounds within AC⁰ reasoning, i.e. in the theory V₀, cf. [8].

Another natural question is the improvement of the quasipolynomial-size proofs of AC⁰[p] lower bounds from Corollary 6.3 to polynomial-size proofs, resp. proving Corollary 6.2 assuming just n ∈ Log.

Acknowledgement

We thank Igor Carboni Oliveira for helpful discussions. Both authors are supported by the Austrian Science Fund (FWF) under project number P 28699. Some initial parts of this work have been undertaken while the second author was supported by the NSERC.

References


