Lifting Lower Bounds for Tree-Like Proofs

Alexis Maciel*
Department of Computer Science
Clarkson University

Phuong Nguyen†
School of Computer Science
McGill University

Toniann Pitassi‡
Department of Computer Science
University of Toronto

February 2, 2011

Abstract

It is known that constant-depth Frege proofs of some tautologies require exponential size. No such lower bound result is known for more general proof systems. We consider tree-like Sequent Calculus proofs in which formulas can contain modular connectives and only the cut formulas are restricted to be of constant depth. Under a plausible hardness assumption concerning small-depth Boolean circuits, we prove exponential lower bounds for such proofs. We prove these lower bounds directly from the computational hardness assumption. We start with a lower bound for cut-free proofs and "lift" it so it applies to proofs with constant-depth cuts. By using the same approach, we obtain the following additional results. We provide a much simpler proof of a known unconditional lower bound in the case where modular connectives are not used. We establish a conditional exponential separation between the power of constant-depth proofs that use different modular connectives. We show that these tree-like proofs with constant-depth cuts cannot polynomially simulate similar dag-like proofs, even when the dag-like proofs are cut-free. We present a new proof of the non-finite axiomatizability of the theory of bounded arithmetic $I\Delta_0(R)$. Finally, under a plausible hardness assumption concerning the polynomial-time hierarchy, we show that the hierarchy G_i^* of quantified propositional proof systems does not collapse.

*Email: alexis@clarkson.edu. Supported by NSF grant CCR-9877150.

†Email: pnguyen@cs.toronto.edu. Supported by the John Templeton Foundation and an NSERC Postdoctoral Research Fellowship.

‡Email: toni@cs.toronto.edu. Supported by an NSERC grant.

1 Introduction

Restricted proof systems have attracted a lot of attention, in large part due to their role in automated theorem provers. For example, Haken [10] showed that the Pigeonhole Principle, a simple, natural and ubiquitous tautology, requires exponential-size Resolution proofs. This means that any theorem prover that works by constructing a Resolution proof — and that is virtually all propositional theorem provers — will require exponential time to prove the Pigeonhole Principle, no matter how efficient it is at finding a proof.

Various extensions of Resolution have also been investigated and shown to be limited in a similar way. For example, constant-depth Frege proofs, which we call $\mathbf{AC}^0$-Frege proofs because of their relation to the circuit class $\mathbf{AC}^0$, are also unable to prove the Pigeonhole Principle in subexponential size [1, 17, 24]. And Cutting Planes have no subexponential-size proof of a certain basic principle concerning colorings of undirected graphs [12, 25].

To this day, however, no lower bound result is known for any proof system more general than $\mathbf{AC}^0$-Frege. For example, a natural extension of $\mathbf{AC}^0$-Frege is to permit the use of modulo r connectives in the proofs, for some constant r. We call this proof system $\mathbf{ACC}^0[r]$-Frege, once again because of its relation to the circuit class $\mathbf{ACC}^0[r]$. No lower bound is known for $\mathbf{ACC}^0[r]$-Frege.

The Pigeonhole Principle lower bound for $\mathbf{AC}^0$-Frege was obtained by an ingenious new model theoretic technique, together with an adaptation of the combinatorial argument used to prove that $\mathbf{AC}^0$ circuits require exponential size to compute the parity function [9, 11, 30]. It is also known that when p and q are distinct primes, $\mathbf{ACC}^0[q]$ circuits require exponential size to compute the modulo p function [27]. Therefore, it is natural to hope that the technique behind that circuit lower bound might be useful in proving a lower bound for $\mathbf{ACC}^0[q]$-Frege proofs. Unfortunately, attempts to prove the corresponding proof complexity lower bound have been unsuccessful, despite considerable effort.

On the other hand, the lower bounds for the Cutting Planes proof system were obtained by using circuit lower bounds directly, not the underlying techniques. This approach relies on the fact that the Cutting Planes proof system has the interpolation property: small Cutting Planes proofs of tautologies of a certain type yield small circuits computing a function related to the tautology. A lower bound on the size of these circuits then implies a lower bound on the size of the proofs. Unfortunately, $\mathbf{AC}^0$-Frege and all of its extensions, including $\mathbf{ACC}^0[q]$-Frege, probably do not have the interpolation property, as this would imply that Blum integers can be factored in time 2^{n^ε} for arbitrary small ε [3].

The initial goal of this research was to discover another way of obtaining proof complexity lower bounds by using circuit lower bounds directly. The hope was

that this would result in new lower bounds for classes such as $\mathbf{ACC}^0[q]$-Frege.

Our work lead us to consider a related proof system. Let $\mathbf{PK}^\star[r]$ denote tree-like Sequent Calculus proofs in which formulas contain conjunctions, disjunctions, negations and modulo r connectives of unbounded arity. Then restrict the cut formulas to be of constant depth. We call this system constant-depth $\mathbf{PK}^\star[r]$. This is a natural proof system that has at least one advantage over the usual definition of constant-depth Frege systems: it is complete for all tautologies, not just constant-depth formulas.

Note that the power of constant-depth $\mathbf{PK}^\star[r]$ is closely related to the power of $\mathbf{ACC}^0[r]$-Frege: over constant-depth tautologies, the two systems are polynomially equivalent. This means that for any constant-depth tautology, there is a polynomial relation between the size of the smallest constant-depth $\mathbf{PK}^\star[r]$ proof and the size of the smallest $\mathbf{ACC}^0[r]$-Frege proof.

The main result of this paper is a lower bound for constant-depth $\mathbf{PK}^\star[r]$. The lower bound is conditional on a plausible hardness conjecture concerning $\mathbf{ACC}^0[r]$ circuits, and uses the conjectured hardness result (directly) as a black box.

To prove the lower bound, we start with a lower bound for cut-free $\mathbf{PK}^\star[r]$ and "lift" it to get a lower bound for constant-depth $\mathbf{PK}^\star[r]$, as follows. Let S be a tautology that requires exponential-size cut-free $\mathbf{PK}^\star[r]$ proofs. Two common examples are the propositional Pigeonhole Principle [10] and the Statman tautologies [29]. Extend the tautology by replacing each of the variables in S by an AND-OR formula expressing an $\mathbf{NC}^1$ function f that is hard to approximate by $\mathbf{ACC}^0[r]$ circuits. Each of these formulas is over a separate subset of the original propositional variables. We call this tautology $S(f)$. We then essentially show that the cut formulas, which are $\mathbf{ACC}^0[r]$ formulas, are unable to help the proof figure out the value of the f formulas contained in the $S(f)$ tautology. In a sense, the proof then reduces to a cut-free proof of S, which we know requires exponential size.

Our lower bound result applies to any tautology S that satisfies certain natural conditions. We observe that these conditions guarantee an exponential lower bound for cut-free $\mathbf{PK}^\star[r]$ and then prove that these conditions imply a lower bound for constant-depth $\mathbf{PK}^\star[r]$. The Pigeonhole Principle and the Statman tautologies satisfy these conditions.

As far as we know, this is the first known lower bound result for an extension of $\mathbf{AC}^0$-Frege under a complexity assumption seemingly weaker than $\mathbf{NP}$ not closed under complementation. In addition, note that size-s constant-depth $\mathbf{PK}^\star[r]$ proofs of $\mathbf{PHP}(f)$ imply size-s $\mathbf{ACC}^0[r]$-Frege proofs of the Pigeonhole Principle. Therefore, our new lower bound is a necessary condition for a lower bound on the size of $\mathbf{ACC}^0[r]$-Frege proofs of the Pigeonhole Principle.

As mentioned above, our lower bound is conditional on a plausible hardness conjecture concerning $\mathbf{ACC}^0[r]$ circuits. This conjecture is similar to a known hardness result for $\mathbf{AC}^0$ circuits: there is a polynomial-size $\mathbf{NC}^1$ function that $\mathbf{AC}^0$ circuits of depth d and subexponential size cannot compute correctly on more than a $1/2 + 1/2^{n^{1/(d+1)}}$ fraction of the inputs [11]. In constrast, in the case of $\mathbf{ACC}^0[r]$ circuits, the strongest hardness result known is much weaker: if r is a prime power, then there is a polynomial-size $\mathbf{NC}^1$ function that $\mathbf{ACC}^0[r]$ circuits of depth d and subexponential size cannot compute correctly on more than a $1/2 + o(1)$ fraction of the inputs [27, 28]. It is natural to conjecture that a strong hardness result also holds for $\mathbf{ACC}^0[r]$ circuits, with no restriction on r: there is a polynomial-size $\mathbf{NC}^1$ function that $\mathbf{ACC}^0[r]$ circuits of depth d and subexponential size cannot compute correctly on more than a $1/2 + 1/2^{n^{1/(d+1)}}$ fraction of the inputs. Our lower bound result for constant-depth $\mathbf{PK}^\star[r]$ is conditional on this conjecture.

In addition to our main lower bound result, we obtain several additional results. First, our lower bound technique can be applied to other proof systems. For example, let $\mathbf{PK}^\star$ denote the restriction of $\mathbf{PK}^\star[r]$ where modular connectives are not allowed. Because constant-depth $\mathbf{PK}^\star$ and $\mathbf{AC}^0$-Frege are polynomially equivalent over constant-depth tautologies, it is known that constant-depth $\mathbf{PK}^\star$ proofs of the Pigeonhole Principle have exponential size. By our technique, we obtain a much simpler proof of the fact that constant-depth $\mathbf{PK}^\star$ proofs of $\mathbf{PHP}(\mathrm{MOD}_2)$ must have exponential size.

Second, we establish a conditional exponential separation between the power of constant-depth proofs that use different modular connectives. In particular, we show that if p is a prime that does not divide r, then, under the assumption that some function in $\mathbf{ACC}^0[p]$ is hard to approximate by $\mathbf{ACC}^0[r]$ circuits, there exists a tautology that has polynomial-size constant-depth $\mathbf{PK}^\star[p]$ proofs but requires exponential-size constant-depth $\mathbf{PK}^\star[r]$ proofs.

Third, it is known that depth-$(d+1)$ tree-like $\mathbf{ACC}^0[r]$-Frege proofs can polynomially simulate depth-d (dag-like) $\mathbf{ACC}^0[r]$-Frege proofs [14]. By applying our lower bound result to the Statman tautologies, and by using the fact that these tautologies have polynomial-size cut-free $\mathbf{PK}$ proofs, we show that such a simulation is not possible in the case of constant-depth $\mathbf{PK}[r]$ proofs: constant-depth $\mathbf{PK}^\star[r]$ proofs cannot even polynomially simulate cut-free $\mathbf{PK}$ proofs. In particular, this implies that lower bounds for constant-depth $\mathbf{PK}[r]$ do not follow automatically from lower bounds for constant-depth $\mathbf{PK}^\star[r]$.

Finally, we apply our approach to Sequent Calculus style proofs systems for quantified Boolean formulas. The system G introduced by Krajíček and Pudlák [15] and given in its present form by Cook and Morioka [7], is a proof

system for reasoning about quantified Boolean formulas. The system G_i^* is the tree-like subsystem of G obtained by restricting the cut rule to formulas with at most i alternations of quantifiers. For each i, the G_i^* proof system is essentially a nonuniform version of Buss's well-studied bounded arithmetic system S_2^i. We show that the G_i^* hierarchy does not collapse, under a hardness assumption about the polynomial-time hierarchy.

The rest of the article is organized as follows. In Section 2, we provide definitions and background, including a precise definition of the proof systems and of the Pigeonhole Principle and Statman tautologies. In Section 3, we define a class of tautologies and show that these tautologies require exponential-size cut-free tree-like proofs. In Section 4, we state our main result, the conditional lower bound for proof systems such as constant-depth $\mathbf{PK}^\star[r]$. We also provide an overview of the lower bound proof. In Section 5, we prove the lower bound. In Sections 6 and 7, we present applications of our main result. In addition to the results mentioned earlier, we prove a hierarchy theorem for constant-depth $\mathbf{PK}^\star[r]$ proofs and we give a new proof of the non-finite axiomatizability of $I\Delta_0(R)$. We conclude, in Section 8, with open problems.

This paper extends and generalizes results that appeared in earlier papers by the authors [19, 21].

2 Definitions and Background

In this section, we define several propositional proof systems based on the Sequent Calculus, as well as the Pigeonhole Principle and Statman tautologies, and establish some basic results concerning these proof systems and these tautologies. We also define related circuit classes and state known and conjectured hardness results for these classes.

2.1 The Propositional Sequent Calculus

The propositional proof systems we consider in this paper are all variants of the Sequent Calculus for AND, OR, NOT and modular connectives. (In Section 7.2 we will consider the Sequent Calculus for quantified propositional logic.) Formulas are defined as usual by using Boolean variables and the connectives $\neg, \vee, \wedge$ and $\oplus_r^b$. We allow $\vee, \wedge$ and $\oplus_r^b$ to have unbounded arity. For example, $\vee(A_1, \ldots, A_n)$ denotes the logical OR of the multiset consisting of $A_1, \ldots, A_n$. Similarly for the AND and modular connectives. Thus commutativity of the connectives is implicit. The formulas $\wedge()$ and $\vee()$ will be used as True and False values and often represented by $\top$ and $\bot$.

The fact that connectives have unbounded arity does not rule out the possibility that some occurrences of connectives may have arity two or that formulas may contain adjacent layers of identical connectives, as in $\vee(\vee(A_1, A_2), \vee(A_3, A_4))$. Binary connectives will often be written in the usual infix notation as in $A \vee B$. In addition, we will use square brackets, as in $[A_1 \vee \ldots \vee A_n]$, to emphasize the fact that $A_1 \vee \ldots \vee A_n$ denotes a formula consisting of $n-1$ binary connectives, not the formula consisting of a single connective $\vee(A_1, \ldots, A_n)$. As usual, binary OR's and AND's are left associative so that, for example, $[A_1 \wedge A_2 \wedge A_3 \wedge A_4]$ represents the formula $(((A_1 \wedge A_2) \wedge A_3) \wedge A_4)$.

The modular connective $\oplus_r^b$, for $0 \leq b < r$, is interpreted to be true if the sum of its arguments is congruent to b modulo r. In what follows, we will omit the r subscript and simply write $\oplus^b$ when there is no confusion possible.

The proof systems operate on *sequents*, which are multisets of formulas of the form $A_1, \ldots, A_r \longrightarrow B_1, \ldots, B_t$. The intended meaning of the sequent $\Gamma \longrightarrow \Delta$ is that the conjunction of the formulas in Γ implies the disjunction of the formulas in Δ. Note that the empty sequent $(\longrightarrow)$ is invalid.

Two sequents $\Gamma \longrightarrow \Delta$ and $\Gamma' \longrightarrow \Delta'$ are *equal* if $\Gamma = \Gamma'$ and $\Delta = \Delta'$ (as multisets). In other words, if each formula that appears on one side of a sequent also appears on the same side of the other sequent, and with the same frequency. In contrast, $\Gamma \longrightarrow \Delta$ and $\Gamma' \longrightarrow \Delta'$ are said to be *similar* if $\Gamma = \Gamma'$ and $\Delta = \Delta'$ *as sets*. That is, if each formula that appears on one side of one sequent also appears on the same side of the other sequent but perhaps with a different frequency. For example, $A, A \longrightarrow B$ and $A \longrightarrow B, B$ are similar but not equal.

A *proof* of a sequent S is a tree of sequents such that the root of the tree is S, the leaves of the tree are initial sequents and every non-leaf sequent in the tree follows from its children by one of the inference rules. A sequent calculus proof can also be a directed acyclic graph (dag) with similar properties.

The *initial sequents* (or axioms) are of the following form:

$$A \longrightarrow A \qquad \longrightarrow \wedge() \qquad \vee() \longrightarrow \qquad \longrightarrow \oplus_r^0() \qquad \oplus_r^b() \longrightarrow$$

where A is a formula, and $1 \leq b < r$.

The rules of inference are as follows. First we have simple structural rules such as weakening (formulas can always be added to the left or to the right of a sequent) and contraction (two copies of the same formula on the same side of a sequent can be replaced by one).

An instance of weakening in which the formula introduced was already present in the sequent (as in $\Gamma \longrightarrow A, \Delta$ derives $\Gamma \longrightarrow A, A, \Delta$) will be called an *expansion*. We will later use the following fact: two sequents are similar if and only if they can be derived from each other using only contractions and expansions.

$$\frac{\Gamma \longrightarrow A, \Delta}{\neg A, \Gamma \longrightarrow \Delta}\ \text{NEG-left} \qquad \frac{A, \Gamma \longrightarrow \Delta}{\Gamma \longrightarrow \neg A, \Delta}\ \text{NEG-right}$$

$$\frac{A, \wedge(F), \Gamma \longrightarrow \Delta}{\wedge(A, F), \Gamma \longrightarrow \Delta}\ \text{AND-left} \qquad \frac{\Gamma \longrightarrow A, \Delta \qquad \Gamma \longrightarrow \wedge(F), \Delta}{\Gamma \longrightarrow \wedge(A, F), \Delta}\ \text{AND-right}$$

$$\frac{A, \Gamma \longrightarrow \Delta \qquad \vee(F), \Gamma \longrightarrow \Delta}{\vee(A, F), \Gamma \longrightarrow \Delta}\ \text{OR-left} \qquad \frac{\Gamma \longrightarrow A, \vee(F), \Delta}{\Gamma \longrightarrow \vee(A, F), \Delta}\ \text{OR-right}$$

$$\frac{A, \oplus_r^{b-1}(F), \Gamma \longrightarrow \Delta \qquad \oplus_r^b(F), \Gamma \longrightarrow A, \Delta}{\oplus_r^b(A, F), \Gamma \longrightarrow \Delta}\ \text{MOD-left}$$

$$\frac{A, \Gamma \longrightarrow \oplus_r^{b-1}(F), \Delta \qquad \Gamma \longrightarrow A, \oplus_r^b(F), \Delta}{\Gamma \longrightarrow \oplus_r^b(A, F), \Delta}\ \text{MOD-right}$$

$$\frac{\Gamma \longrightarrow \oplus_r^a(F), \Delta \qquad \Gamma \longrightarrow \oplus_r^b(G), \Delta}{\Gamma \longrightarrow \oplus_r^{a+b}(F, G), \Delta}\ \text{MOD-add}$$

$$\frac{\Gamma \longrightarrow \oplus_r^a(F, G), \Delta \qquad \Gamma \longrightarrow \oplus_r^b(G), \Delta}{\Gamma \longrightarrow \oplus_r^{a-b}(F), \Delta}\ \text{MOD-substract}$$

Figure 1: Logical rules

After the structural rules, we have the cut rule:

$$\frac{\Gamma, A \longrightarrow \Delta \qquad \Gamma \longrightarrow A, \Delta}{\Gamma \longrightarrow \Delta}\ \text{cut}$$

The formula A is called the *cut formula*.

The remaining rules are the logical rules, which are shown in Figure 1. These rules allow us to introduce each connective on either side of sequents. In these rules, A is an individual formula, F, G stand for a multisets of formulas and (A, F) is short for $\{A\} \cup F$. Note that even though the connectives $\wedge$, $\vee$ and $\oplus_r^b$ have unbounded arity, their introduction rules are binary rules. The rules for the modular connectives are adapted from [2]. Here we need the rules MOD-add and MOD-substract to have short derivations of the equivalences between the AND-OR formulas that compute the MOD_p function, and the formulas using the modular connectives. These equivalences are required for the proof of Theorems 6.4 and 6.5.

In this article, we will often need to perform derivations that introduce binary

$$\frac{A, B, \Gamma \longrightarrow \Delta}{(A \wedge B), \Gamma \longrightarrow \Delta}\ \text{AND-left2} \qquad \frac{\Gamma \longrightarrow A, \Delta \qquad \Gamma \longrightarrow B, \Delta}{\Gamma \longrightarrow (A \wedge B), \Delta}\ \text{AND-right2}$$

$$\frac{A, \Gamma \longrightarrow \Delta \qquad B, \Gamma \longrightarrow \Delta}{(A \vee B), \Gamma \longrightarrow \Delta}\ \text{OR-left2} \qquad \frac{\Gamma \longrightarrow A, B, \Delta}{\Gamma \longrightarrow (A \vee B), \Delta}\ \text{OR-right2}$$

Figure 2: Additional rules for binary connectives

connectives. For example,

$$\frac{A, B, \Gamma \longrightarrow \Delta}{(A \wedge B), \Gamma \longrightarrow \Delta}$$

Strictly speaking, the AND-left rule does not allows us to perform this derivation in one step. To simplify both our upper and lower bound arguments, we add to our proof systems logical rules that allow the direct introduction of binary AND's and OR's. These rules are shown in Figure 2.

Definition 2.1. *Let $F = \{(\Gamma_n \longrightarrow \Delta_n) : n \in \mathbf{N}\}$ be a family of sequents. Then $P = \{P_n : n \in \mathbf{N}\}$ is a family of $\mathbf{PK}^\star[r]$ proofs for F if, for every n, P_n is a valid (tree-like) proof of $(\Gamma_n \longrightarrow \Delta_n)$. If modular connectives are not used in P, then we say that P is a family of $\mathbf{PK}^\star$ proofs for F. If the proofs are permitted to be dag-like instead of just tree-like, then we say that P is a family of $\mathbf{PK}[r]$ or $\mathbf{PK}$ proofs, respectively.*

As usual, a formula can be represented as a tree whose leaves are the *literals* of the formula (variables and negated variables) and whose inner nodes are the connectives. The *depth* of a formula is then the maximum number of blocks of connectives of the same type along any path from the root to a leaf.

The depth of a proof is sometimes defined as the maximum depth of any formula that occurs in it. For example, an $\mathbf{ACC}^0[r]$-Frege proof is simply a $\mathbf{PK}[r]$ proof in which every formula has constant-depth. Similarly for $\mathbf{AC}^0$-Frege and $\mathbf{PK}$.

In this article, however, we are mainly interested in proofs in which only the depth of the cut formulas is limited.

Definition 2.2. *A* depth-d $\mathbf{PK}^\star[r]$ *proof is one in which all the cut formulas have depth at most d. We call these d-$\mathbf{PK}^\star[r]$ proofs. A* constant-depth $\mathbf{PK}^\star[r]$ *proof is a d-$\mathbf{PK}^\star[r]$ proof, for some constant d. Similarly, for $\mathbf{PK}^\star$, $\mathbf{PK}[r]$ and $\mathbf{PK}$.*

We will only consider tautologies consisting of AND-OR formulas. These tautologies will contain connectives of unbounded arity. Two tautologies we will consider are the Pigeonhole Principle and Statman tautologies. These will be defined

later in this section. Because the Sequent Calculus is cut-free complete, the proof systems d-$\mathbf{PK}^\star[r]$ and d-$\mathbf{PK}^\star$ are complete for all tautologies while $\mathbf{ACC}^0[r]$-Frege and $\mathbf{AC}^0$-Frege are complete only for constant-depth tautologies.

The *size* of a formula is the number of literals and connectives it contains. The *size* of a sequent is the total size of its formulas. The *size* of a proof is the total size of all the sequents it contains and that size is normally expressed in terms of the size of the conclusion. For example, if F is a family of sequents of size t_n, then a polynomial-size $\mathbf{PK}^\star[r]$ proof of F would have size $t_n^{O(1)}$.

Definition 2.3. *Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be two propositional proof systems. Then $\mathcal{P}_1$* simulates *$\mathcal{P}_2$ if whenever a tautology has a $\mathcal{P}_2$ proof of size s, then the tautology also has a $\mathcal{P}_1$ proof of size at most $s^{O(1)}$. In addition, $\mathcal{P}_1$* p-simulates *$\mathcal{P}_2$ if there is a polynomial-time function F that given a $\mathcal{P}_2$ proof outputs a $\mathcal{P}_1$ proof of the same tautology.*

The power of constant-depth $\mathbf{PK}^\star[r]$ is closely related to the power of $\mathbf{ACC}^0[r]$-Frege when we consider only tautologies of constant depth:

Theorem 2.4. *Consider the following proof systems: constant-depth $\mathbf{PK}^\star[r]$, constant-depth $\mathbf{PK}[r]$ and $\mathbf{ACC}^0[r]$-Frege. If a constant-depth tautology has a proof of size s in any of these proof systems, then it has a proof of size at most $s^{O(1)}$ in the other two. In other words, constant-depth $\mathbf{PK}^\star[r]$, $\mathbf{PK}[r]$ and $\mathbf{ACC}^0[r]$-Frege p-simulate one another with respect to constant-depth tautologies. Similarly for constant-depth $\mathbf{PK}^\star$, constant-depth $\mathbf{PK}$ and $\mathbf{AC}^0$-Frege.*

Proof. First, a constant-depth $\mathbf{PK}^\star[r]$ proof is simply a special case of a constant-depth $\mathbf{PK}[r]$ proof.

Second, all the formulas in a constant-depth $\mathbf{PK}[r]$ proof must be either subformulas of the conclusion or formulas that will be the target of a cut. Therefore, in a constant-depth $\mathbf{PK}[r]$ proof of a constant-depth tautology, all the formulas must have constant depth, which implies that such a constant-depth $\mathbf{PK}[r]$ proof is actually an $\mathbf{ACC}^0[r]$-Frege proof.

Finally, any $\mathbf{ACC}^0[r]$-Frege proof of size s and depth d can be transformed into a tree-like $\mathbf{ACC}^0[r]$-Frege proof of size $s^{O(1)}$ and depth $d+1$ [13]. Such a proof is a special case of a constant-depth $\mathbf{PK}^\star[r]$ proof. □

In this article, we are mainly interested in the d-$\mathbf{PK}^\star[r]$ and d-$\mathbf{PK}^\star$ proof systems, but our main theorem will be more general: it will apply to any version of $\mathbf{PK}^\star[r]$ or $\mathbf{PK}^\star$ in which the cuts are limited to a set $\mathcal{C}$. We denote these proof systems by $\mathbf{PK}^\star[r](\mathcal{C})$ and $\mathbf{PK}^\star(\mathcal{C})$. For example, d-$\mathbf{PK}^\star = \mathbf{PK}^\star(\mathcal{C})$ when $\mathcal{C}$ is the set of depth-d AND-OR formulas.

One final note: the addition of the rules for binary connectives (Figure 2) does not significantly alter the power of the proof systems we consider in this paper. The reason is simple: each of these rules can be easily simulated in three steps by using the original rules. In addition, lower bounds for proof systems that include these extra rules obviously imply lower bounds for systems that include only the original rules.

2.2 Closure under Restrictions

Throughout this paper, we will apply partial truth assignments (also called restrictions) to sequents and proofs. In this section, we show that if S is a sequent that has a small proof and we apply a partial truth assignment to S and then simplify S, then the resulting sequent also has a small proof. In fact, we will show how that proof can be obtained by adapting the original proof of S.

First, we define precisely what we mean by applying a partial truth assignment to a sequent and then simplifying it.

Definition 2.5. (Restriction of a formula) *Let f be a formula and ρ a partial truth assignment to the variables of f. Then $f|_\rho$,* the restriction of f by ρ, *is defined inductively as follows.*

1. *If f is a variable, then $f|_\rho$ is either the value assigned to that variable or the variable itself, in case the variable is given no value by ρ.*

2. *If $f = \neg A$, then consider $\neg(A|_\rho)$, the result of replacing A by $A|_\rho$ in f. If $A|_\rho = \top$, then $f|_\rho = \bot$. If $A|_\rho = \bot$, then $f|_\rho = \top$. Otherwise, $f|_\rho = \neg(A|_\rho)$.*

3. *If $f = \vee(F)$, where F is a multiset of formulas, then consider $\vee(F')$, the result of replacing each argument B in F by its restriction $B|_\rho$. If F' contains $\top$, then $f|_\rho = \top$. Remove every $\bot$ from F'. If F' is empty, then $f|_\rho = \bot$. If exactly one $B|_\rho$ is left in F', then $f|_\rho = B|_\rho$. Otherwise, $f|_\rho = \vee(F')$.*

4. *If $f = \wedge(F)$, then $f|_\rho$ is defined in a similar way but with $\bot$ and $\top$ interchanged.*

5. *If $f = \oplus^b_r(F)$, then consider $\oplus^b_r(F')$ with F' defined as before. Remove every $\bot$ from F'. If any $B|_\rho = \top$, remove it from F' and subtract 1 from b (modulo r). If F' is empty and $b = 0$, then $f|_\rho = \top$. If F' is empty and $b \neq 0$, then $f|_\rho = \bot$. Otherwise, $f|_\rho = \oplus^b_r(F')$.*

We then extend this definition to sequents as follows.

Definition 2.6. (Restriction of a sequent) *Let $S = \Gamma \longrightarrow \Delta$ be a sequent and ρ a partial truth assignment. Then $S|_\rho$,* the restriction of S by ρ*, is defined as follows. Consider $\Gamma' \longrightarrow \Delta'$, the result of replacing every formula A in S by its restriction $A|_\rho$. If Γ' contains $\bot$ or Δ' contains $\top$, then $S|_\rho$ is the axiom $\longrightarrow \top$. Otherwise, remove every $\top$ from Γ' and every $\bot$ from Δ'. Then $S|_\rho$ is $\Gamma' \longrightarrow \Delta'$.*

We now show that if a sequent has a small proof, then all of its restrictions also have small proofs. The proof of this result uses the fact that we do not remove duplicate formulas in defining the restriction of a sequent. (Later we will show that the same result holds for the quantified proof system $\mathbf{G}$.)

Definition 2.7. (Closure under restrictions) *A proof system $\mathcal{P}$ is* closed under restrictions *if for any tautology S and any partial truth assignment ρ, if S has a $\mathcal{P}$ proof of size t, then $S|_\rho$ has a $\mathcal{P}$ proof of size at most t.*

Lemma 2.8. *All of the proof systems defined above are closed under restrictions.*

Proof. Suppose that $\mathcal{P}$ is one of these proof systems. To prove the lemma, we will show that if ρ is any restriction, then any $\mathcal{P}$ proof P can be transformed into a $\mathcal{P}$ proof P' whose sequents are the restrictions of the sequents of the original proof.

Let P' be the result of replacing every sequent S in P by its restriction $S|_\rho$. We must show that P' is a valid proof.

If S is an initial sequent, then it is easy to verify that $S|_\rho$ is also an initial sequent. For example, suppose that S is $x \longrightarrow x$ and that ρ sets x to $\bot$. Then $S|_\rho$ is $\longrightarrow \top$.

Now suppose that S is the result of an inference in P. The argument splits into cases depending on the rule used to infer S.

Suppose that S is inferred by an application of the OR-left rule from S_1 and S_2:

$$\frac{S_1 \qquad S_2}{S} = \frac{A, \Gamma \longrightarrow \Delta \qquad \vee(F), \Gamma \longrightarrow \Delta}{\vee(A, F), \Gamma \longrightarrow \Delta}$$

where A is a formula and F is a multiset of formulas. We will show that $S|_\rho$ can be inferred from $S_1|_\rho$ and $S_2|_\rho$, which we know have replaced S_1 and S_2 in P'.

Consider how the restriction acts on these sequents. If $A|_\rho = \top$, then $A|_\rho$ is removed from $S_1|_\rho$ and $\vee(A, F)|_\rho$ is removed from $S|_\rho$. In that case, we have that $S|_\rho = S_1|_\rho$ and those two sequents can be collapsed in P'. If $A|_\rho = \bot$, then $A|_\rho$ is removed from $\vee(A, F)|_\rho$, which implies that $S|_\rho = S_2|_\rho$ and those two sequents can be collapsed in P'.

Now, let F' be the result of replacing each B in F by its restriction $B|_\rho$. If F' contains $\top$, then $\vee(F)|_\rho$ is removed from $S_2|_\rho$ and $\vee(A, F)|_\rho$ is removed from $S|_\rho$, which implies that $S|_\rho = S_2|_\rho$ and those two sequents can be collapsed in

P'. Remove every $\bot$ from F'. If F' is empty, then $\vee(A,F)|_\rho = A|_\rho$, which implies that $S_1|_\rho = S|_\rho$. If F' contains a single $B|_\rho$, then $\vee(F)|_\rho = B|_\rho$ and $\vee(A,F)|_\rho = \vee(A|_\rho, B|_\rho)$, which implies that $S|_\rho$ can be inferred from $S_1|_\rho$ and $S_2|_\rho$ by using the OR-left2 rule. If $|F'| \geq 2$, then $S|_\rho$ can be inferred from $S_1|_\rho$ and $S_2|_\rho$ by an application of the OR-left rule.

The cases where S is inferred in P by using other rules can be handled in a similar way. The details are left to the reader.

Note that P' does not contain any connectives that were not already present in P, and thus the size of P' is no greater than the size of P. In addition, if P is tree-like, then so is P'. The depth of any formula in P', including the cut formulas, is no greater than the depth of the corresponding formula in P. Therefore, since $\mathcal{P}$ is one of the proofs systems defined earlier, the fact that P is a $\mathcal{P}$ proof implies that P' is also a $\mathcal{P}$ proof. This proves the lemma. □

2.3 Hard Propositional Formulas

As mentioned in the introduction, the main result of this paper is a lower bound that applies to every tautology that satisfies certain conditions. The Pigeonhole Principle and Statman tautologies, which we define in this subsection, are good examples of such tautologies.

The (injective) Pigeonhole Principle with m pigeons and n holes, for $m > n$, intuitively states that if m pigeons are placed into n holes, then (at least) one hole must receive more than one pigeon. This tautology can be expressed as the following sequent, which we denote by $\mathbf{PHP}_n^m$:

$$\longrightarrow \bigwedge_{j=1}^{n} \neg p_{1j},\ \ldots,\ \bigwedge_{j=1}^{n} \neg p_{mj},\ p_{11} \wedge p_{21},\ p_{11} \wedge p_{31},\ \ldots,\ p_{(m-1)n} \wedge p_{mn}$$

When m is much larger than n, typically when $m \geq 2n$, we refer to this tautology as the Weak Pigeonhole Principle. The case $m = n+1$ is usually what is meant simply by the Pigeonhole Principle. We will use $\mathbf{PHP}_n$ to denote the corresponding tautology $\mathbf{PHP}_n^{n+1}$.

Exponential lower bounds have been proved on the size of $\mathbf{AC}^0$-Frege proofs of the Pigeonhole Principle (for $m = n+1$) [1, 17, 24]. By Theorem 2.4, this also implies an exponential lower bound on the size of constant-depth $\mathbf{PK}$ and constant-depth $\mathbf{PK}^\star$ proofs of the Pigeonhole Principle.

Statman's tautologies express a form of strong induction. The tautology for strong induction up to n has variables p_i, q_i, $i \leq n$, and is given by the following sequent, which we denote $\mathbf{STATMAN}_n$:

$$\longrightarrow (\neg p_1 \wedge \neg q_1),\ [\gamma_1 \wedge \neg p_2 \wedge \neg q_2],\ \ldots,\ [\gamma_{n-1} \wedge \neg p_n \wedge \neg q_n],\ \gamma_n \qquad (1)$$

where

$$\gamma_i = [(p_1 \vee q_1) \wedge \ldots \wedge (p_i \vee q_i)]$$

For example, $\mathbf{STATMAN}_2$ is the sequent

$$\longrightarrow (\neg p_1 \wedge \neg q_1),\ [(p_1 \vee q_1) \wedge \neg p_2 \wedge \neg q_2],\ [(p_1 \vee q_1) \wedge (p_2 \vee q_2)]$$

It is easy to see that the Statman sequents are tautologies. Let $A_i = p_i \vee q_i$. Then $\mathbf{STATMAN}_n$ essentially states that if it is not the case that A_i is true for all i, then there is j such that A_k is true for all $k < j$ but A_j is false. This is clearly a tautology: simply let j be the smallest i for which A_i is false.

The Statman sequents are known to require exponential-size cut-free $\mathbf{PK}^\star$ proofs [29, 5, 6]. This lower bound will be proved in Section 3. It is the basic lower bound that we will "lift" in order to obtain the main result of this paper.

In contrast, it is also known that the Statman sequents have polynomial-size cut-free $\mathbf{PK}$ proofs. We prove this result here for completeness.

Theorem 2.9. ([29, 5, 6]) *The sequent* $\mathbf{STATMAN}_n$ *has a cut-free* $\mathbf{PK}$ *proof of size polynomial in* n.

Proof. We construct a cut-free $\mathbf{PK}$ proof inductively. It will be clear from our construction that the proof has size polynomial in n. For $n = 1$ we have the following proof of $\mathbf{STATMAN}_1$:

$$\dfrac{\dfrac{\dfrac{\dfrac{p_1 \longrightarrow p_1}{p_1 \longrightarrow p_1, q_1}\text{ weakening}}{\longrightarrow \neg p_1, p_1, q_1}\text{ NEG-right} \qquad \dfrac{\dfrac{\dfrac{q_1 \longrightarrow q_1}{q_1 \longrightarrow p_1, q_1}\text{ weakening}}{\longrightarrow \neg q_1, p_1, q_1}\text{ NEG-right}}{\longrightarrow (\neg p_1 \wedge \neg q_1), p_1, q_1}\text{ AND-right}}{\longrightarrow (\neg p_1 \wedge \neg q_1), (p_1 \vee q_1)}\text{ OR-right}$$

For the inductive step, suppose that we have a proof of $\mathbf{STATMAN}_{n-1}$:

$$\longrightarrow (\neg p_1 \wedge \neg q_1),\ \ldots,\ [\gamma_{n-2} \wedge \neg p_{n-1} \wedge \neg q_{n-1}],\ \gamma_{n-1}$$

The following is a proof of $\mathbf{STATMAN}_n$:

1. Apply NEG-right to $p_n \longrightarrow p_n$:

$$\longrightarrow \neg p_n,\ p_n$$

2. Apply weakening to (1) and AND-right2 with $\mathbf{STATMAN}_{n-1}$:

$$\longrightarrow (\neg p_1 \wedge \neg q_1),\ \ldots,\ [\gamma_{n-2} \wedge \neg p_{n-1} \wedge \neg q_{n-1}],\ [\gamma_{n-1} \wedge \neg p_n],\ p_n$$

3. Apply NEG-right to $q_n \longrightarrow q_n$:

$$\longrightarrow \neg q_n,\ q_n$$

4. Apply weakening to (3) and AND-right2 with (2):

$$\longrightarrow (\neg p_1 \wedge \neg q_1),\ \ldots,\ [\gamma_{n-2} \wedge \neg p_{n-1} \wedge \neg q_{n-1}],\ [\gamma_{n-1} \wedge \neg p_n \wedge \neg q_n],\ p_n,\ q_n$$

5. Apply OR-right to (4):

$$\longrightarrow (\neg p_1 \wedge \neg q_1),\ \ldots,\ [\gamma_{n-2} \wedge \neg p_{n-1} \wedge \neg q_{n-1}],\ [\gamma_{n-1} \wedge \neg p_n \wedge \neg q_n],\ (p_n \vee q_n)$$

6. Apply weakening to $\mathbf{STATMAN}_{n-1}$ and AND-right2 with (5):

$$\longrightarrow (\neg p_1 \wedge \neg q_1),\ \ldots,\ [\gamma_{n-1} \wedge \neg p_n \wedge \neg q_n],\ [\gamma_{n-1} \wedge (p_n \vee q_n)]$$

This last sequent is $\mathbf{STATMAN}_n$ as desired. □

Even though the Statman sequents require exponential-size cut-free $\mathbf{PK}^\star$ proofs, it can be shown that they have small $\mathbf{PK}^\star$ proofs if cut-formulas of depth 1 are allowed.

Theorem 2.10. *The sequent* $\mathbf{STATMAN}_n$ *has a polynomial-size* 1-$\mathbf{PK}^\star$ *proof.*

Proof. Start by deriving $\Gamma_i \longrightarrow \gamma_i$, for $1 \leq i \leq n$, where

$$\Gamma_i = \{(p_1 \vee q_1),\ \ldots,\ (p_i \vee q_i)\}$$

Then, consider the following sequents:

$$\begin{array}{rcll} & \longrightarrow & (\neg p_1 \wedge \neg q_1),\ (p_1 \vee q_1) & \\ \Gamma_{i-1} & \longrightarrow & [\gamma_{i-1} \wedge \neg p_i \wedge \neg p_i],\ (p_i \vee q_i) & (2 \leq i \leq n) \\ \Gamma_n & \longrightarrow & \gamma_n & \end{array}$$

The first two groups can be derived in a way similar to some of the sequents in the proof of Theorem 2.9. The sequent $\mathbf{STATMAN}_n$ can then be obtained from these sequents by repeated cuts on the formulas $(p_i \vee q_i)$, starting with $i = n$. □

2.4 Constant-Depth Boolean Circuits

In this article, we will consider the standard Boolean circuit classes $\mathbf{AC}^0$, $\mathbf{ACC}^0[r]$, for constant r, and $\mathbf{NC}^1$. $\mathbf{AC}^0$ and $\mathbf{ACC}^0[r]$ circuits are of constant depth and consist of gates of unbounded fan-in. $\mathbf{AC}^0$ circuits allow only AND, OR and NOT gates. $\mathbf{ACC}^0[r]$ also permit MOD_r gates. These gates output 1 when the sum of their inputs is divisible by r. $\mathbf{NC}^1$ circuits are of logarithmic depth but allow only NOT and binary AND and OR gates.

It is known that $\mathbf{AC}^0$ and $\mathbf{ACC}^0[q^k]$ circuits of subexponential size cannot compute the MOD_p function if p and q are distinct primes [11, 27]. In addition, these circuits cannot approximate MOD_p very well:

Theorem 2.11. (Håstad [11]) *If C is a depth-d $\mathbf{AC}^0$ circuit of size $2^{n^{1/(d+1)}}$, then, for sufficiently large n, C cannot compute MOD_p correctly on more than a $(p-1)/p + 1/2^{n^{1/(d+1)}}$ fraction of the inputs.*

Theorem 2.12. (Smolensky [27, 28]) *Suppose that p and q are distinct primes. If C is a depth-d $\mathbf{ACC}^0[q^k]$ circuit of size $2^{o(n^{1/2d})}$, then, for sufficiently large n, C cannot compute MOD_p correctly on more than a $(p-1)/p + o(1)$ fraction of the inputs.*

Note how the $\mathbf{AC}^0$ hardness result is stronger than the one for $\mathbf{ACC}^0[q^k]$. It is natural to conjecture that a stronger hardness result also holds for $\mathbf{ACC}^0[q^k]$, when q is prime, and even for $\mathbf{ACC}^0[r]$, with no restriction on r.

More precisely, some of the results in this paper are conditional on the following two conjectures. We say that a Boolean function is *balanced* if evaluates to 0 and 1 on the same number of inputs.

Conjecture 2.13. *Let p be a prime number that does not divide r. There exists a balanced polynomial-size $\mathbf{ACC}^0[p]$ function f such that if C is a depth-d $\mathbf{ACC}[r]$ circuit of size $2^{n^{1/(d+1)}}$, then, for sufficiently large n, C cannot compute f correctly on more than a $1/2 + 1/2^{n^{1/(d+1)}}$ fraction of the inputs.*

Conjecture 2.14. *There exists a balanced polynomial-size $\mathbf{NC}^1$ function f such that if C is a depth-d $\mathbf{ACC}^0[r]$ circuit of size $2^{n^{1/(d+1)}}$, then, for sufficiently large n, C cannot compute f correctly on more than an $1/2 + 1/2^{n^{1/(d+1)}}$ fraction of the inputs.*

The first conjecture implies the separation of the $\mathbf{ACC}^0[r]$ circuit classes for various r. When $p = 2$, MOD_2 is a reasonable candidate for a hard function.

The second conjecture is weaker since the hard function f is only required to be in $\mathbf{NC}^1$. A balanced version of the majority function is a reasonable candidate for a hard function.

It is well known that a function has a polynomial-size $\mathbf{NC}^1$ circuit if and only if it has a polynomial-size AND-OR formula. Therefore, the second conjecture states that there is a balanced polynomial-size AND-OR formula that is hard to approximate by $\mathbf{ACC}^0[r]$ circuits.

3 Basic Lower Bound for Statman's Sequents

In the previous section, we saw that the Statman sequents have polynomial-size cut-free $\mathbf{PK}$ proofs. Those proofs were clearly not tree-like because in the induction step, the sequent $\mathbf{STATMAN}_{n-1}$ was used more than once.

In this section, we will show that this is necessary: any cut-free $\mathbf{PK}^\star$ proof of the Statman sequents must be of exponential size [29, 5, 6]. We then define a class of tautologies and point out that this lower bound applies to all of these tautologies. The proof of this lower bound will provide the backbone for the proof of the main result of this paper.

Theorem 3.1. (Statman lower bound) *Any cut-free $\mathbf{PK}^\star$ proof of the sequent $\mathbf{STATMAN}_n$ has size at least 2^n.*

We will use the following lemma in the proof of the theorem.

Lemma 3.2. *Consider the sequent $\mathbf{STATMAN}_n$, which is of the form*

$$\longrightarrow (A_1 \wedge B_1),\ (A_2 \wedge B_2),\ \ldots,\ (A_t \wedge B_t)$$

Suppose that T is similar to $\mathbf{STATMAN}_n$ and that T' is the result of modifying T by replacing one of the formulas $A \wedge B$ by either A, $\wedge(A)$, B or $\wedge(B)$. Then there exists a partial truth assignment ρ such that $T'|\rho$ is similar to $\mathbf{STATMAN}_{n-1}$, modulo a possible renaming of the variables.

Proof of Theorem 3.1. It will be easier to prove the lower bound for all sequents that are similar to $\mathbf{STATMAN}_n$. We will prove a lower bound on the number of sequents in the proof, which, of course, is a lower bound on the size of the proof. The proof is by induction on n.

The base case, for $n = 1$, is obvious since any sequent similar to $\mathbf{STATMAN}_1$ cannot be an axiom.

For the induction step, suppose that the lower bound holds for all sequents similar to $\mathbf{STATMAN}_{n-1}$. Consider a cut-free $\mathbf{PK}^\star$ proof of a sequent S similar to $\mathbf{STATMAN}_n$. Once again, S cannot be an axiom. So S must be derived by either a contraction, weakening or an AND-right rule. In addition, moving up the proof from the root S, we must eventually reach a sequent T derived by either

an AND-right rule or by an instance of weakening that is not just an expansion. (Recall that an expansion is an instance of weakening that introduces another copy of a formula that is already present in the sequent.)

Consider the weakening case and suppose that T is derived from T'. The sequent T must be similar to $\mathbf{STATMAN}_n$. This implies that T' is similar to $\mathbf{STATMAN}_n$ except for the fact that one of the $A \wedge B$ formulas is missing. It is not hard to see that such a sequent cannot be a tautology, which implies that the weakening case cannot occur.

Therefore, T is derived from two sequents T' and T'' by an AND-right rule. But then, by the lemma, there are partial truth assignments ρ' and ρ'' such that $T'|_{\rho'}$ and $T''|_{\rho''}$ are both similar to $\mathbf{STATMAN}_{n-1}$. By induction, these restrictions require proofs of size at least 2^{n-1}. Therefore, by Lemma 2.8, T' and T'' each require a proof of that size, which implies that the total size of the proof of S is at least 2^n, as desired. □

We now prove the lemma.

Proof of Lemma 3.2. Recall that $\mathbf{STATMAN}_n$ is the sequent

$$\longrightarrow (\neg p_1 \wedge \neg q_1),\ [\gamma_1 \wedge \neg p_2 \wedge \neg q_2],\ \ldots,\ [\gamma_{n-1} \wedge \neg p_n \wedge \neg q_n],\ \gamma_n$$

where

$$\gamma_n = [(p_1 \vee q_1) \wedge \ldots \wedge (p_n \vee q_n)]$$

Suppose that T is similar to $\mathbf{STATMAN}_n$. There are two cases to consider depending on which formula is broken up.

The first case is when an occurrence of $[\gamma_{i-1} \wedge \neg p_i \wedge \neg q_i]$ is broken up, for some $i \leq n$. This means that T' is an expansion of $\mathbf{STATMAN}_n$ with one occurrence of $[\gamma_{i-1} \wedge \neg p_i \wedge \neg q_i]$ replaced by one of the following: $\neg q_i$, $\wedge(\neg q_i)$, $(\gamma_{i-1} \wedge \neg p_i)$ or $\wedge(\gamma_{i-1} \wedge \neg p_i)$. In all cases, let ρ set both p_i and q_i to $\top$.

In $T'|_\rho$, if $j < i$, then every occurrence of $[\gamma_{j-1} \wedge \neg p_j \wedge \neg q_j]$ is unchanged. If $j > i$, then $(p_i \vee q_i)$ is deleted from every occurrence of $[\gamma_{j-1} \wedge \neg p_j \wedge \neg q_j]$. The formula $(p_i \vee q_i)$ is also deleted from every occurrence of γ_n. In addition, any remaining occurrence of $[\gamma_{i-1} \wedge \neg p_i \wedge \neg q_i]$ is deleted from $T'|_\rho$. For $j > i$, rename every p_j and q_j as p_{j-1} and q_{j-1}, respectively. The sequent $T'|_\rho$ is now similar to $\mathbf{STATMAN}_{n-1}$.

The second case is when an occurrence of γ_n is broken up. Then T' is similar to $\mathbf{STATMAN}_n$ but with one occurrence of γ_n replaced by one of the following: $(p_n \vee q_n)$, $\wedge(p_n \vee q_n)$, γ_{n-1} or $\wedge(\gamma_{n-1})$. In the first two subcases, let ρ set both p_n and q_n to $\bot$. Then $T'|_\rho$ is similar to $\mathbf{STATMAN}_{n-1}$.

In the remaining two subcases, let ρ set both p_n and q_n to $\top$. Then the every occurrence of $[\gamma_{n-1} \wedge \neg p_n \wedge \neg q_n]$ is deleted from $T'|_\rho$ and that sequent is similar to $\mathbf{STATMAN}_{n-1}$. □

The lower bound of Theorem 3.1 can be generalized to apply to all tautologies that satisfy the following set of conditions:

Definition 3.3. (Statman property) *We say that a sequent S has the* Statman property of order n *if it satisfies the following conditions:*

1. *S is of the form $\longrightarrow \Gamma$ where Γ is not empty and consists of nonempty conjunctions.*

2. *Removing from S every occurrence of any of these conjunctions results in an invalid sequent.*

3. *If $n \geq 2$, then for all sequents T similar to S the following condition holds. For any formula $\wedge(F)$ of T (for a multiset of formulas F) and for any $A \in F$, let T' be obtained from T by replacing simultaneously all occurrences of $\wedge(F)$ by either A or $\wedge(F')$, where F' is F with one occurrence of A removed. Then there is a partial truth assignment ρ such that $T'|\rho$ has the Statman property of order $n-1$, modulo a possible renaming of the variables.*

We say that a family of sequents $\{S_n\}$ has the Statman property *if, for every n, S_n has the Statman property of order n.*

This definition is for sequents that have all their formulas on the right. All of our tautologies will have that form but this is only done for convenience. The Statman property, as well as all the results presented later in this paper, can be generalized to sequents that have formulas on both sides. In that case, the $\vee$ connective would play on the left the role that the $\wedge$ connective plays on the right.

The following lemma will be useful and follows directly from the definition.

Lemma 3.4. *If S has the Statman property of order n, then every sequent similar to S also has the Statman property of order n.*

It is easy to verify that the Statman sequents have the Statman property. This can be proved by induction on n with the third condition following from Lemmas 3.2 and 3.4.

It is also easy to see that our proof of Theorem 3.1 applies not just to the Statman sequents but to all sequents that have the Statman property.

Theorem 3.5. *If S has the Statman property of order n, then any cut-free $\mathbf{PK}^\star$ proof of S requires size 2^n.*

As mentioned earlier, the proof of this lower bound will provide the backbone for the proof of the main result of this paper. In a sense, the lower bound for cut-free proof systems will be lifted to apply to proof systems with limited cuts. The lower bound will apply to certain extensions of any tautology that has the Statman property. We therefore end this section by noting that the Pigeonhole Principle also has the Statman property. (Note that the lower bound of Theorem 3.5 was already known to apply to the Pigeonhole Principle.)

Lemma 3.6. *The Pigeonhole Principle has the Statman property.*

Proof. Recall that $\mathbf{PHP}_n$ is the sequent

$$\longrightarrow \bigwedge_{k=1}^{n} \neg p_{1k},\ \ldots,\ \bigwedge_{k=1}^{n} \neg p_{(n+1)k},\ p_{11} \wedge p_{21},\ p_{11} \wedge p_{31},\ \ldots,\ p_{nn} \wedge p_{(n+1)n}$$

The proof of the lemma is by induction on n. For every $n \geq 1$, it is clear that $\mathbf{PHP}_n$ is of the form specified in Definition 3.3. It is also easy to see that if any of the conjunctions of $\mathbf{PHP}_n$ is removed, then we can find an assignment that falsifies all of the remaining conjunctions. In particular, this establishes that $\mathbf{PHP}_1$ has the Statman property of order 1.

Now suppose that $n \geq 2$. All that remains to show is that Part 3 of the definition holds for $\mathbf{PHP}_n$. Suppose that T is similar to $\mathbf{PHP}_n$. We will consider two cases, depending on which formula is broken up.

First, suppose that this formula is a conjunction associated with a pigeon i, saying that pigeon i is not mapped to any hole. In this case, in T', that formula will be replaced by either $\neg p_{ir}$ or $\bigwedge_{k \neq r} \neg p_{ik}$, for some hole r. Suppose it is $\neg p_{ir}$. Let ρ set p_{ir} to true, all other p_{ik} to false, and all other p_{jr} to false. In other words, ρ maps pigeon i to hole r, and nowhere else, and no other pigeon goes to hole r. Then $T'|_\rho$ becomes similar to the pigeonhole principle with one less pigeon (pigeon i) and one less hole (hole r). The inductive hypothesis and Lemma 3.4 imply that $T'|_\rho$ has the Statman property of order $n-1$. The partial truth assignment that works for the case of $\bigwedge_{k \neq r} \neg p_{ik}$ is similar: it sends pigeon i to some hole other than r.

Second, suppose that the formula that is broken up is of the form $(p_{ir} \wedge p_{jr})$, saying that two pigeons i and j are both mapped to the same hole r. In this case, in T', $(p_{ir} \wedge p_{jr})$ will be replaced by either p_{ir}, $\wedge(p_{ir})$, p_{jr} or $\wedge(p_{jr})$. Suppose it is p_{ir}. Let ρ be the restriction that maps pigeon j to r and nowhere else, and no other pigeon goes to hole r. Then $T'|_\rho$ once again becomes similar to the pigeonhole principle with one less pigeon and one less hole. The other cases are handled similarly. □

4 Main Lower Bound Result and Overview of Proof

The Statman lower bound of the previous section is for cut-free $\mathbf{PK}^\star$. We now want to "lift" this lower bound so that it holds for stronger proof systems. Let $\mathcal{C}$ be a set of formulas. Our main theorem is a lower bound for systems of the form $\mathbf{PK}^\star[r](\mathcal{C})$ and $\mathbf{PK}^\star(\mathcal{C})$. Recall that these are versions of $\mathbf{PK}^\star[r]$ and $\mathbf{PK}^\star$ in which the cuts are limited to $\mathcal{C}$. For example, d-$\mathbf{PK}^\star = \mathbf{PK}^\star(\mathcal{C})$ where $\mathcal{C}$ is the set of depth-d AND-OR formulas.

Theorem 3.5 essentially says that a sequent S with the Statman property is hard for cut-free proofs. We will obtain sequents that are hard for proofs with $\mathcal{C}$ cuts by replacing each variable in S by a formula that is hard for $\mathcal{C}$. For technical reasons, we restrict ourselves to functions that are balanced.

Definition 4.1. *Let $f(x_1, \ldots, x_m)$ be a balanced Boolean function on m variables. Let $\mathcal{C}$ be a set of circuits. Then f is (σ, ϵ)-hard with respect to $\mathcal{C}$ if the following holds. Suppose that $B(x_1, \ldots, x_m, y_1, \ldots, y_k)$ is any conjunction of circuits that are either in $\mathcal{C}$ or are negations of circuits in $\mathcal{C}$, with $k \geq 0$. If the total size of B is at most $2^{\sigma(m)}$, then when f is viewed as a function of $x_1, \ldots, x_m, y_1, \ldots, y_k$, neither B nor $\neg B$ compute f correctly on more than a $1/2 + \epsilon(m)$ fraction of the inputs.*

For example, let $\mathcal{C}$ be the set of depth-d AND-OR formulas. By Theorem 2.11, the parity function is (σ, ϵ)-hard with respect to $\mathcal{C}$ where $\sigma(m) = m^{1/(d+1)}$ and $\epsilon(m) = 1/2^{m^{1/(d+1)}}$.

As another example, let q be prime and let $\mathcal{C}$ be the set of depth-d formulas with AND, OR, NOT and MOD_{q^k} connectives. By Theorem 2.12, the parity function is (σ, ϵ)-hard for $\mathcal{C}$ where $\sigma = m^{1/(d+1)}$ and $\epsilon = o(1)$. In addition, Conjecture 2.14 asserts there is a balanced polynomial-size $\mathbf{NC}^1$ function that is (σ, ϵ)-hard for $\mathcal{C}$ where now $\sigma = m^{1/(d+1)}$ and $\epsilon = 1/2^{m^{1/(d+1)}}$.

Definition 4.2. *Let S be a sequent with variables $p_1, \ldots, p_n$ and f a formula on m variables. Then $S(f)$ denotes the sequent obtained from S by replacing each variable p_i by $f(x_1^i, x_2^i, \ldots, x_m^i)$ for a new set of variables $x_1^i, x_2^i, \ldots, x_m^i$.*

We are now ready to state our main theorem. Let $\mathcal{P}$ be either the $\mathbf{PK}^\star[r]$ or $\mathbf{PK}^\star$ proof systems and let $\mathcal{C}$ be a set of formulas. Suppose that S has the Statman property of order n and that f is (σ, ϵ)-hard for $\mathcal{C}$. We will prove a lower bound of 2^n on the size of any $\mathcal{P}(\mathcal{C})$ proof of $S(f)$.

In applications, this lower bound is useful only if 2^n is at least superpolynomial in the size of $S(f)$. Therefore, informally, the functions σ and ϵ must satisfy the following requirement: for sufficiently large n, there exists m such that

1. m is not too large, so that 2^n is superpolynomial in the size of $S(f)$, and

2. m is not too small, so that Condition (2) below holds.

Theorem 4.3. (Main theorem) *Let S be a sequent with the Statman property of order n and let k denote the number of variables in S. Let $\mathcal{P}$ be either the $\mathbf{PK}^\star[r]$ or $\mathbf{PK}^\star$ proof systems. Let f be a Boolean formula in m variables and suppose that, as a Boolean function, f is balanced and (σ, ϵ)-hard for some set $\mathcal{C}$ of formulas that is closed with respect to subformulas and restrictions. Suppose that m is such that*

$$m > 3k + n^2, \qquad \sigma(m) \geq n, \qquad \epsilon(m) < \frac{1}{2^{n^2} 4^k} \tag{2}$$

Then, the sequent $S(f)$ requires $\mathcal{P}(\mathcal{C})$ proofs of size 2^n.

The hypothesis of the main theorem is satisfied, for example, when S is $\mathbf{STATMAN}_n$, $\mathcal{P}$ is $\mathbf{PK}^\star$, $\mathcal{C}$ is the set of depth-d AND-OR formulas (so that $\mathcal{P}(\mathcal{C})$ is d-$\mathbf{PK}^\star$) and f is a polynomial-size AND-OR parity formula. In this case, $k = 2n$, $\sigma(m) = m^{1/(d+1)}$, $\epsilon(m) = 1/2^{m^{1/(d+1)}}$, so Condition (2) is satisfied for any sufficiently large n by letting $m = 2n^{2(d+1)}$. Then the size N of $S(f)$ is $n^{O(d)}$ and $2^n \geq 2^{N^{1/O(d)}}$, which is not only superpolynomial but exponential in N.

We end this section with an overview of the proof of the main theorem. The complete proof will be given in the next section.

Suppose that S has the Statman property of order n and let $\mathcal{P}$, $\mathcal{C}$, f and m satisfy the conditions of the theorem. In particular, f is hard with respect to $\mathcal{C}$. Recall that S must have the form $\longrightarrow \Gamma$ where each formula in Γ is a nonempty conjunction. To keep things simple, suppose that all the formulas of $S(f)$ are distinct and that the contraction rule is not used. Now suppose, by contradiction, that π is a small $\mathcal{P}(\mathcal{C})$ proof of $S(f)$.

First, note that $S(f)$ is not an axiom. So $S(f)$ must be derived by either weakening, an AND-right rule or a cut on a $\mathcal{C}$ formula. The first two cases can be handled in essentially the same way as in the Statman lower bound (Theorem 3.1). So we will focus on the third case in this overview.

Suppose that $S(f)$ is derived from $g \longrightarrow \Gamma(f)$ and $\longrightarrow g, \Gamma(f)$ by a cut on $g \in \mathcal{C}$. In this context, we call g a *side formula*. It could be that one of those two sequents is easy to prove. A trivial example is when $g = \vee()$. In that case, $g \longrightarrow \Gamma(f)$ can be derived from the axiom $\vee() \longrightarrow$ by weakening. But then the validity of $\longrightarrow g, \Gamma(f)$ would essentially depend on $\Gamma(f)$ since $g = \vee()$ is false for every possible truth assignment. So $\longrightarrow g, \Gamma(f)$ should be just as hard to prove as the original sequent $\longrightarrow \Gamma(f)$.

In general, with respect to $\Gamma(f)$, we say that an assignment is *critical* for $g \longrightarrow \Gamma(f)$ if it satisfies g and critical for $\longrightarrow g, \Gamma(f)$ if it falsifies g. Clearly, at least

half the assignments will be critical for one of those two sequents. Suppose it is $\longrightarrow g, \Gamma(f)$. Then the fact that g does not approximate f very well will allow us to show that every truth assignment to the variables of Γ can be achieved by a large number of critical truth assignments to the variables of $\longrightarrow g, \Gamma(f)$.

For example, consider any variable p of Γ. In $\Gamma(f)$, p is replaced by f. Since f is hard for g, at least 1/4 of the assignments that falsify g satisfy f and at least 1/4 of the assignments that falsify g falsify f. We will later see how to extend this to all the variables of Γ.

Intuitively, what all this seems to indicate is that cuts on $\mathcal{C}$ formulas don't help in a proof of $S(f)$. This intuition will be formalized as follows. From the root of π, follow all paths until one of the following is reached: an axiom, a sequent where the first occurrence of one of the formulas of $S(f)$ is introduced by weakening, or a sequent where one of the formulas of $S(f)$ is introduced by an AND-right rule (but not necessarily the first occurrence). This defines a subtree π' of π in which all sequents are of the form $\Lambda \longrightarrow \Delta, \Gamma(f)$ with all the formulas in Λ and Δ belonging to $\mathcal{C}$.

Generalizing the earlier definitions, we say that the formulas in Λ and Δ are *side formulas* (with respect to $\Gamma(f)$) and that an assignment is *critical* for a sequent of this form if is satisfies all side formulas on the left and falsifies all side formulas on the right.

All assignments are critical for the root sequent $S(f)$. In addition, critical assignments are preserved as we go up π' from the root: if T is derived from T' and T'', then every assignment critical for T is also critical for at least one of T' and T''. This is essentially because of the soundness of the inference rules.

Now, if π' has at least 2^n leaves, then we are done: we have shown that π is large. Otherwise, a $1/2^n$ fraction of all assignments is critical for some leaf L of π'. Note that this is a large number of assignments since, by Condition (2), the total number of assignments is at least 2^{kn^2}.

We can now use on L essentially the same argument that was used in the proof of the Statman lower bound. For example, suppose that L is derived from L' and L'' by an application of one of the AND-right rules that introduces a formula of $\Gamma(f)$. The fact that L is of the form $\Lambda \longrightarrow \Delta, \Gamma(f)$ implies that L' must be of the form $\Lambda \longrightarrow \Delta, \Gamma'(f)$ where Γ' contains all the formulas of Γ but with some $\wedge(F)$ replaced by either A or $\wedge(F')$, and similarly for L'':

$$\frac{L' \qquad L''}{L} \qquad = \qquad \frac{\Lambda \longrightarrow \Delta, \Gamma'(f) \qquad \Lambda \longrightarrow \Delta, \Gamma''(f)}{\Lambda \longrightarrow \Delta, \Gamma(f)}$$

In addition, all the partial truth assignments that are critical for L (with respect to $\Gamma(f)$) are also critical for L' and L'' (with respect to $\Gamma'(f)$ and $\Gamma''(f)$, respectively).

Because $\longrightarrow \Gamma$ has the Statman property of order n, there is a partial truth assignment ρ' to the variables of $\longrightarrow \Gamma$ such that $(\longrightarrow \Gamma')|_{\rho'} = (\longrightarrow \Psi')$ has the Statman property of order $n-1$. As explained earlier, the fact that f is hard with respect to the side formulas allows us to achieve ρ' with a large number of critical truth assignments to the variables of L'. In particular, as we will show later, there is a partial truth assignment τ' to the variables of L' that is consistent with ρ' and such that $L'|_{\tau'} = (\Lambda|_{\tau'} \longrightarrow \Delta|_{\tau'}, \Psi'(f))$ still has a large number of critical assignments (with respect to $\Psi'(f)$).

The same holds for L'': there is a partial truth assignment τ'' to the variables of L'' such that $L''|_{\tau''} = (\Lambda|_{\tau''} \longrightarrow \Delta|_{\tau''}, \Psi''(f))$, where $\longrightarrow \Psi''$ has the Statman property of order $n-1$, and such that $L''|_{\tau''}$ has a large number of critical assignments (with respect to $\Psi''(f)$). The large number of critical assignments of both $L'|_{\tau'}$ and $L''|_{\tau''}$ allows us to repeat the argument on these sequents and inductively show that each of these sequents requires a proof of size 2^{n-1}. Therefore, as in the Statman lower bound, π must be of size 2^n.

As we said earlier, in the next section, we will turn this overview into a complete proof of the main theorem. This will require careful calculations of numbers of critical assignments. We will also address the possibility that contractions may be used in the proof.

5 Proof of Main Theorem

First, we precisely define the concepts of side formula and critical assignment.

Definition 5.1. *Let L be a sequent of the form $\Lambda \longrightarrow \Delta, \Gamma$. With respect to Γ, the formulas of Λ and Δ are called* side formulas *and we say that a truth assignment is* critical *for L (still with respect to Γ) if it satisfies all the side formulas in Λ and falsifies all the side formulas in Δ.*

In order to prove Theorem 4.3, we will need a few lemmas. Here it is crucial that f be balanced.

Lemma 5.2. *Let $f(x_1, \ldots, x_m)$ be a balanced Boolean formula in m variables and suppose that f is (σ, ϵ)-hard for some set $\mathcal{C}$ of formulas that is closed with respect to restrictions. Let $B(x_1, \ldots, x_m)$ be a conjunction of formulas that are either in $\mathcal{C}$ or are negations of formulas in $\mathcal{C}$. Suppose that the size of B is no greater than $2^{\sigma(m)}$ and that a fraction of at least $2\epsilon(m)$ truth assignments satisfy B. Then, among all the assignments that satisfy B, at least $1/4$ satisfy f and at least $1/4$ falsify f.*

Proof. Let r be the fraction of truth assignments that satisfy B. Then $r \geq 2\epsilon(m)$. Suppose that among the truth assignments that satisfy B, there is a fraction s that satisfy f. We will show that $s \geq 1/4$. Because f is balanced, a similar proof shows that a $1/4$ fraction of all assignments satisfying B falsify f.

The assignments on which $\neg B$ computes f correctly are those that satisfy either $\neg B \wedge f$ or $B \wedge \neg f$. Those assignments represent a $(1/2 - rs) + r(1 - s)$ fraction of all assignments. Since f is (σ, ϵ)-hard with respect to $\mathcal{C}$, $\neg B$ computes f correctly on no more than a $1/2 + \epsilon(m)$ fraction of all assignments. Therefore,

$$\frac{1}{2} - rs + r(1 - s) \leq \frac{1}{2} + \epsilon(m)$$

Since $\epsilon(m) \leq r/2$, it follows that $s \geq 1/4$. □

Lemma 5.3. *Let f be a balanced Boolean formula in m variables and suppose that f is (σ, ϵ)-hard for some set $\mathcal{C}$ of formulas that is closed with respect to restrictions. Let S be a sequent of the form $\Lambda \longrightarrow \Delta, \Gamma$ where Γ contains at least one occurrence of f. Suppose all the side formulas (with respect to Γ) are in $\mathcal{C}$ and that their total size is at most $2^{\sigma(m)}$. Suppose that the fraction t of assignments that are critical for S is at least $4\epsilon(m)$. Then, for each truth value v, there is an assignment τ to the variables of f such that $f(\tau) = v$ and $S|_\tau$ has at least a fraction $t/4$ of assignments that are critical.*

Proof. Let W be the assignments to the variables of S other than $x_1, \ldots, x_m$, the variables of f. Each assignment in W has 2^m extensions to all the variables in S. Let W_1 be those assignments in W that have a fraction of at least $t/2$ extensions that are critical for S. Together, all the assignments $W - W_1$ can be extended to at most $1/2$ of all critical assignments. Therefore, at least $1/2$ of all critical assignments are extensions of assignments in W_1.

Consider an arbitrary $\sigma \in W_1$. Since $t/2 \geq 2\epsilon(m)$ and since $\mathcal{C}$ is closed under restrictions, we can apply Lemma 5.2 to $S|_\sigma$ to get that at least $1/4$ of the critical extensions of σ give f value v. Therefore, at least $1/8$ of all critical assignments give f value v. In other words, at least $t/8$ of all assignments to the variables of S are critical and give f value v.

On the other hand, among all the assignments to the variables of f, at most $1/2$ give f the value v. As a result, there is an assignment τ to the variables of f that sets f to v and has a fraction of at least $t/4$ extensions that are critical. This implies that at least $t/4$ of the assignments of $S|_\tau$ are critical. □

Lemma 5.4. *Let f be a balanced Boolean formula in m variables and suppose that f is (σ, ϵ)-hard for some set $\mathcal{C}$ of formulas that is closed with respect to restrictions. Let S be a sequent of the form $\Lambda \longrightarrow \Delta, \Gamma$ where Γ contains multiple occurrences*

of f over distinct sets of variables $x_1^i, \ldots, x_m^i$, for $1 \leq i \leq k$. Suppose all the side formulas (with respect to Γ) are in $\mathcal{C}$ and that their total size is at most $2^{\sigma(m)}$. Suppose that the fraction t of assignments that are critical for S is at least $4^k\epsilon(m)$. Then, for any k truth values $v_1, \ldots, v_k$, there is an assignment τ to all the variables x_j^i so that $f(x_1^i, \ldots, x_m^i)|_\tau = v_i$, for $1 \leq i \leq k$, and $S|_\tau$ has at least a fraction $t/4^k$ of assignments that are critical.

Proof. By induction on k using Lemma 5.3. □

We are now ready to prove our main theorem. For the sake of the inductive argument, we will prove a more general result.

Theorem 5.5. *Let $\mathcal{P}$ be either the $\mathbf{PK}^\star[r]$ or $\mathbf{PK}^\star$ proof systems, $f(x_1, \ldots, x_m)$ be a balanced Boolean formula that is (σ, ϵ)-hard for some set $\mathcal{C}$ for formulas that is closed with respect to subformulas and restriction. Let k, n be such that they satisfy Condition* (2) *of Theorem 4.3, i.e.,*

$$m > 3k + n^2, \qquad \sigma(m) \geq n, \qquad \epsilon(m) < \frac{1}{2^{n^2}4^k}$$

Suppose that $\longrightarrow \Gamma$ has the Statman property of order $r \leq n$ and that the number j of variables in $\longrightarrow \Gamma$ is no greater than k. Let T be a sequent of the form $\Lambda \longrightarrow \Delta, \Gamma(f)$ where Λ and Γ are in $\mathcal{C}$ and the total size of Λ and Γ is at most 2^r. Suppose that the fraction of all truth assignments to the variables of T that are critical (with respect to $\Gamma(f)$) is at least

$$\frac{1}{4^{k-j}2^{(r+1)+(r+2)+\cdots+n}}$$

(where the sum $(r+1)+(r+2)+\cdots+n$ is 0 if $r = n$). Then any $\mathcal{P}(\mathcal{C})$ proof of T must have size at least 2^r.

By letting $r = n$ and $T = S(f)$, we get our main theorem.

Proof of Theorem 5.5. First note that if T has no variables, then it is easy to show that r must be 1. In addition, since T contains at least one nonempty conjunction, T must have size at least 2. So we now assume that $j \geq 1$.

The proof is by induction on r.

Inductive basis: $r = 1$. The sequent T cannot be an axiom because $\Gamma(f)$ contains at least one conjunction that does not belong to $\mathcal{C}$ and therefore cannot appear on the left. This implies that the proof contains at least two sequents.

Induction step: $r \geq 2$. Suppose that the lower bound holds for $r - 1$. We prove it for r.

Let π be a proof of T and let α be the fraction of assignments that are critical for T. Consider the subtree π' of π that is obtained by starting at the root and following all paths in π until one of the following is reached:

- an axiom,
- a sequent derived by an instance of weakening that introduces the first occurrence of one of the formulas of $\Gamma(f)$, or
- a sequent derived by an AND-right rule that introduces one of the formulas of $\Gamma(f)$ (but not necessarily the first occurrence of that formula).

In part because $\mathcal{C}$ is closed with respect to subformulas, it is not hard to see that all the sequents in π' are of the form $\Lambda' \longrightarrow \Delta', \Gamma'(f)$, where all the formulas in Λ' and Δ' belong to $\mathcal{C}$ and where $\longrightarrow \Gamma'$ is similar to $\longrightarrow \Gamma$. In addition, all the partial truth assignments that are critical for T are preserved as we go up π' (with respect to the appropriate $\Gamma'(f)$). In particular, every assignment that is critical for T is critical for at least one leaf of π'.

If π' has size at least 2^r, then so does π, and we are done. Otherwise, there must be a leaf $L = (\Lambda_L \longrightarrow \Delta_L, \Gamma_L(f))$ of π' for which a fraction of at least $\alpha/2^r$ assignments are critical (with respect to $\Gamma_L(f)$).

This leaf L cannot be an axiom, for the same reason that T was not an axiom in the inductive basis.

So suppose that L is obtained from some sequent L' by a weakening that introduces the first occurrence of one of the formulas of $\Gamma_L(f)$:

$$\frac{L'}{L} \quad = \quad \frac{\Lambda_L \longrightarrow \Delta_L, \Gamma'_L(f)}{\Lambda_L \longrightarrow \Delta_L, \Gamma_L(f)}$$

where Γ'_L is just like Γ_L but with one formula missing. The sequent $\longrightarrow \Gamma_L$ has the Statman property because it is similar to $\longrightarrow \Gamma$. Therefore, $\longrightarrow \Gamma'_L$ is not a tautology and there is a truth assignment ρ' that falsifies this sequent. The total size of the side formulas in L' (with respect to $\Gamma'_L(f)$) is at most $2^r \leq 2^n \leq 2^{\sigma(m)}$. The fraction of assignments that are critical for L' is at least $\alpha/2^r > 4^j\epsilon(m)$ since $\epsilon(m) < 1/(2^{n^2}4^k)$. We can therefore apply Lemma 5.4: there is a partial truth assignment τ' that is consistent with ρ' and such that the number of critical assignments of $L'|_{\tau'}$ is at least $2^{jm}\alpha/(2^r 4^j)$. This number is at least 1 since $m > 3k + n^2$. The existence of such a critical assignment and the fact that τ' falsifies $\longrightarrow \Gamma'_L(f)$ implies that L' is not a tautology. Therefore, L could not have been derived from L' by weakening.

The only other possibility is that L is obtained by an AND-right rule that introduces one of the formulas of $\Gamma_L(f)$:

$$\frac{L' \qquad L''}{L} \quad = \quad \frac{\Lambda_L \longrightarrow \Delta_L, \Gamma'_L(f) \qquad \Lambda_L \longrightarrow \Delta_L, \Gamma''_L(f)}{\Lambda_L \longrightarrow \Delta_L, \Gamma_L(f)}$$

where each of $\Gamma'_L(f)$ and $\Gamma''_L(f)$ is just like $\Gamma_L(f)$ but with one of the formulas $\wedge(F)$ replaced by either A or $\wedge(F')$.

We now show every proof of L' contains at least 2^{r-1} sequents. Because $\longrightarrow \Gamma_L$ is similar to $\longrightarrow \Gamma$, it has the Statman property of order r and there is a partial truth assignment ρ' to the variables of $\longrightarrow \Gamma_L$ such that $(\longrightarrow \Gamma'_L)|_{\rho'} = (\longrightarrow \Psi')$ has the Statman property of order $r-1$. Let $j' \leq j$ be the number of variables of $\longrightarrow \Psi'$, which means that ρ' sets $j - j'$ variables in $\longrightarrow \Gamma_L$. As before, the total size of the side formulas in L' (with respect to $\Gamma'_L(f)$) is at most $2^{\sigma(m)}$ and the fraction of assignments that are critical for L' is at least $4^j\epsilon(m)$. So we can again apply Lemma 5.4: there is a partial truth assignment τ' that is consistent with ρ' and such that the fraction of assignments critical for

$$L'|_{\tau'} = (\Lambda_L|_{\tau'} \longrightarrow \Delta_L|_{\tau'}, \Psi'(f))$$

is at least

$$\frac{\alpha}{2^r 4^{j-j'}} = \frac{1}{4^{k-j'} 2^{r+(r+1)+\cdots+n}}$$

Since both $\mathcal{C}$ and $\mathcal{P}(\mathcal{C})$ are closed with respect to restrictions, the inductive hypothesis implies that every proof of either $L'|_{\tau'}$ or L' contains at least 2^{r-1} sequents.

The same argument can be used to show every proof of L'' also contains at least 2^{r-1} sequents. This implies that π contains at least 2^r sequents. □

6 Applications to Propositional Proof Systems

In this section, we apply our main theorem (Theorem 4.3) to obtain a variety of results concerning propositional proof systems. Most of these results are conditional on the circuit hardness results conjectured in Section 2.4. First, we obtain lower bounds for constant-depth $\mathbf{PK}^\star[r]$ and constant-depth $\mathbf{PK}^\star$ proofs. Second, we obtain separation results for constant-depth $\mathbf{PK}^\star[r]$ proofs that use different modular connectives. Third, we show that constant-depth $\mathbf{PK}^\star[r]$ proofs cannot p-simulate cut-free $\mathbf{PK}$ proofs. Finally, we prove a hierarchy theorem for constant-depth $\mathbf{PK}^\star[r]$ proofs.

6.1 Lower Bounds for d-$\mathbf{PK}^\star[r]$ and d-$\mathbf{PK}^\star$

Our first application of the main theorem is a conditional exponential lower bound for d-$\mathbf{PK}^\star[r]$. As mentioned earlier, as far as we know, this is the first known lower bound result for an extension of $\mathbf{AC}^0$-Frege under a complexity assumption seemingly weaker than $\mathbf{NP}$ not closed under complementation.

Theorem 6.1. *Suppose that Conjecture 2.14 is true and let f be a balanced polynomial-size AND-OR formula that is hard to approximate by depth-d $\mathbf{ACC}[r]$ circuits of size $2^{n^{1/(d+1)}}$. For sufficiently large n and for $m = (5n^2)^{d+1}$, any d-$\mathbf{PK}^\star[r]$ proof of either $\mathbf{PHP}_n(f_m)$ or $\mathbf{STATMAN}_n(f_m)$ has size at least $2^{N^{1/O(d)}}$, where N is the size of the tautology.*

The proof is essentially just a matter of verifying that the two tautologies satisfy the conditions of the main theorem.

Proof Sketch. Corollary 2.14 says that f is (σ, ϵ)-hard for $\mathbf{ACC}^0[r]$ circuits of depth d, where

$$\sigma(m) = m^{\frac{1}{d+1}}, \qquad \epsilon(m) = 1/2^{m^{\frac{1}{d+1}}}$$

For $\mathbf{PHP}_n$ we have $k = (n+1)n$, while for $\mathbf{STATMAN}_n$, $k = 2n$. Thus it is straightforward to verify that Condition (2) of Theorem 4.3 is satisfied. Theorem 6.1 then follows from Theorem 4.3 by using the fact that both $\mathbf{PHP}_n$ and $\mathbf{STATMAN}_n$ have the Statman property of order n. In particular, the size N of the sequent (i.e., either $\mathbf{PHP}_n(f_m)$ or $\mathbf{STATMAN}_n(f_m)$) is a polynomial in n and m, so $n = N^{1/O(d)}$, and hence the lower bound $2^n = 2^{N^{1/O(d)}}$. □

Note that in this lower bound result, m depends on d. This implies that we have different tautologies for each depth. We can prove a lower bound with a single tautology for every depth but the lower bound is slightly weaker.

Theorem 6.2. *Let f be as in Theorem 6.1. Let $\alpha(n)$ be unbounded and nondecreasing. Let $m = n^{\alpha(n)}$. Then, for sufficiently large n, any d-$\mathbf{PK}^\star[r]$ proof of either $\mathbf{PHP}_n(f_m)$ or $\mathbf{STATMAN}_n(f_m)$ has size at least $2^{N^{1/O(\alpha(N))}}$, where N is the size of the tautology.*

This is no longer an exponential lower bound, but it is still very large and certainly much larger than quasipolynomial. For example, with $\alpha(n) = \log\log n$, we get a lower bound of $2^{N^{1/O(\log\log N)}}$.

We can also use our main theorem to obtain an *unconditional* exponential lower bound for d-$\mathbf{PK}^\star$. As mentioned earlier, constant-depth $\mathbf{PK}^\star$ and $\mathbf{AC}^0$-Frege are polynomially equivalent with respect to constant-depth tautologies and it is

already known that constant-depth $\mathbf{AC}^0$-Frege proofs of the Pigeonhole Principle have exponential size [1, 17, 24]. Therefore, constant-depth $\mathbf{PK}^\star$ proofs of the Pigeonhole Principle also have exponential size. Our main theorem provides a much simpler proof of this lower bound.

First note that for every p, the MOD_p function can be expressed by a polynomial-size AND-OR formula. Let MOD_p^b be the formula defined recursively as follows:

$$\begin{aligned}&\mathrm{MOD}_p^b(x_1,\ldots,x_m)\\ &= \bigvee_{a=0}^{p-1}\Big(\mathrm{MOD}_p^a(x_1,\ldots,x_{m/2})\wedge \mathrm{MOD}_p^{b-a}(x_{m/2+1},\ldots,x_m)\Big)\end{aligned}$$

Then MOD_p is simply MOD_p^0. Let $\mathrm{MOD}_{p,m}$ denote the MOD_p formula over m variables.

Theorem 6.3. *For sufficiently large n and for $m=(5n^2)^{d+1}$, any d-$\mathbf{PK}^\star$ proof of either $\mathbf{PHP}_n(\mathrm{MOD}_{2,m})$ or $\mathbf{STATMAN}_n(\mathrm{MOD}_{2,m})$ has size at least $2^{N^{1/O(d)}}$, where N is the size of the tautology.*

6.2 Separation Results for d-$\mathbf{PK}^\star[r]$ Proofs with Different Modular Connectives

The lower bound on the size of d-$\mathbf{PK}^\star[r]$ proofs of $\mathbf{PHP}(f_m)$ is interesting in part because it is a necessary step towards a lower bound on the size of d-$\mathbf{PK}^\star[r]$ proofs of $\mathbf{PHP}$. But by focusing on extensions of the Statman tautology, we can obtain separation results for the d-$\mathbf{PK}^\star[r]$ and d-$\mathbf{PK}^\star$ systems.

Theorem 6.4. *Let MOD_2 be the polynomial-size AND-OR formula described in the preceding subsection. Consider the tautology $\mathbf{STATMAN}_n(\mathrm{MOD}_{2,m})$ with $m=(5n^2)^{d+1}$. Let N denote the size of this sequent. Then the following holds:*

1. *$\mathbf{STATMAN}_n(\mathrm{MOD}_{2,m})$ has a 3-$\mathbf{PK}^\star[2]$ proof of size polynomial in N.*
2. *For sufficiently large n, any d-$\mathbf{PK}^\star$ proof of $\mathbf{STATMAN}_n(\mathrm{MOD}_{2,m})$ has size at least $2^{N^{1/O(d)}}$.*

Proof Sketch. The lower bound is from the previous subsection. A small 3-$\mathbf{PK}^\star[2]$ proof of $\mathbf{STATMAN}_n(\mathrm{MOD}_{2,m})$ can be constructed in two stages. First, prove (by a cut-free proof) that the AND-OR formula MOD_2 is equivalent to a formula consisting of a single $\oplus_2^0$ connective. Second, prove $\mathbf{STATMAN}_n(\oplus_2^0)$ by using the proof of Theorem 2.10, but now the cut formulas have depth 2, so the proof

is 2-$\mathbf{PK}^\star[2]$. Finally, prove $\mathbf{STATMAN}_n(\mathrm{MOD}_{2,m})$ from $\mathbf{STATMAN}_n(\oplus_2^0)$ by using the fact that MOD_2 is equivalent to $\oplus_2^0$. We need to cut on the formulas of $\mathbf{STATMAN}_n(\oplus_2^0)$, which are of depth 3. □

We can also prove a conditional separation of d-$\mathbf{PK}^\star[p]$ from d'-$\mathbf{PK}^\star[r]$ when p is a prime that does not divide r, for some d'. The separating sequents cannot mention the connectives $\oplus_r$ or $\oplus_p$. Therefore we need to use the formula MOD_p from the preceding section to express the polynomial-size $\mathbf{ACC}^0[p]$ function from Conjecture 2.13 as an AND-OR formula. The next theorem is proved in the same way as the previous one.

Theorem 6.5. *Suppose that p is a prime number that does not divide r and that Conjecture 2.13 is true. Let f be a polynomial-size AND-OR formula that expresses a balanced depth-k, polynomial-size $\mathbf{ACC}^0[p]$ function that is hard to approximate by depth-d $\mathbf{ACC}[r]$ circuits of size $2^{n^{1/(d+1)}}$. Consider the tautology $\mathbf{STATMAN}_n(f_m)$ with $m = (5n^2)^{d+1}$. Let N denote the size of this sequent. Then the following holds:*

1. *$\mathbf{STATMAN}_n(f_m)$ has a $(k+2)$-$\mathbf{PK}^\star[p]$ proof of size polynomial in N.*
2. *For sufficiently large n, any d-$\mathbf{PK}^\star[r]$ proof of $\mathbf{STATMAN}_n(f_m)$ has size at least $2^{N^{1/O(d)}}$.*

6.3 Tree-Like Versus Dag-Like Proofs

The lower bounds in the previous subsections are for the *tree-like* proof systems d-$\mathbf{PK}^\star[r]$ and d-$\mathbf{PK}^\star$. We would obviously like to extend these lower bounds to the corresponding dag-like systems. One way would be to show that the tree-like proofs can p-simulate the dag-like proofs. Our lower bounds for the tree-like systems would then immediately translate into lower bounds for the dag-like systems. And this is precisely the case with constant-depth Frege proofs: depth-$(d+1)$ tree-like $\mathbf{ACC}^0[r]$-Frege proofs can p-simulate depth-d dag-like $\mathbf{ACC}^0[r]$-Frege proofs, and similarly for $\mathbf{AC}^0$-Frege [14].

Unfortunately, we can combine our lower bounds and the cut-free $\mathbf{PK}$ proof of the Statman tautologies (Theorem 2.9) to show that d-$\mathbf{PK}^\star[r]$ proofs cannot even p-simulate cut-free $\mathbf{PK}$ proofs.

Theorem 6.6. *There is a tautology of size N that has polynomial-size cut-free $\mathbf{PK}$ proofs but requires d-$\mathbf{PK}^\star$ proofs of size at least $2^{N^{1/O(d)}}$, for sufficiently large N. If Conjecture 2.14 is true, then there is a tautology of size N that has polynomial-size cut-free $\mathbf{PK}$ proofs but requires d-$\mathbf{PK}^\star[r]$ proofs of size at least $2^{N^{1/O(d)}}$, for sufficiently large N.*

Proof Sketch. The sequent that separates cut-free $\mathbf{PK}$ and d-$\mathbf{PK}^\star$ is $\mathbf{STATMAN}_n(\mathrm{MOD}_{2,m})$ as in Theorem 6.4. A polynomial-size cut-free $\mathbf{PK}$ proof of this sequent is easily obtained by modifying the proof given in Theorem 2.9. The lower bound for d-$\mathbf{PK}^\star$ proofs is given in Theorem 6.4.

The second part is proved similarly. □

6.4 Hierarchy Theorems

It is known that the $\mathbf{AC}^0$-Frege hierarchy is infinite in the sense that $\mathbf{AC}^0$-Frege proofs of depth d cannot p-simulate $\mathbf{AC}^0$-Frege proofs of depth $d+1$ [13]. In this section, by combining our lower bounds with the 1-$\mathbf{PK}^\star$ proof of the Statman tautologies (Theorem 2.10), we show that the constant-depth $\mathbf{PK}^\star[r]$ hierarchy is also infinite, under the assumption that Conjecture 2.13 holds.

First, for every p, we show that the MOD_p function has (exponential-size) constant-depth AND-OR formulas. As explained earlier, this also shows that every $\mathbf{ACC}^0[p]$ function has a constant-depth AND-OR formula.

Lemma 6.7. *For each $d \geq 2$, there is an* AND-OR *formula $\mathrm{MOD}_{p,d,m}$ of depth d, size $mp^{(d-1)m^{1/(d-1)}}$ with an* OR *at the top that computes $\mathrm{MOD}_p(x_1,\ldots,x_m)$.*

Proof Sketch. Divide the input $\vec{x} = (x_1,\ldots,x_m)$ into $k = m^{1/(d-1)}$ blocks $\vec{y}_1,\ldots,\vec{y}_k$ each containing m/k variables. Then $\mathrm{MOD}_p(\vec{x})$ can be computed with a DNF formula of size kp^{k-1} from the various $\mathrm{MOD}_p^b(\vec{y}_j)$ with $b = 0,\ldots,p-1$ and $j = 1,\ldots,k$. (There are p^{k-1} terms, each of size k.) Then repeat recursively $d-1$ times, using either CNF or DNF formulas as appropriate, so that the total depth ends up being d and not $2(d-1)$. □

We use the formula $\mathrm{MOD}_{p,2d+4,m}$ from the lemma for the next theorem. The size of this formula is

$$mp^{(2d+3)m^{\frac{1}{2d+3}}}$$

We choose depth $2d+4$ because we must have $m = O(n^{2d+2})$. With this setting, the lower bound 2^n is still superpolynomial in the size of $\mathrm{MOD}_{p,2d+4,m}$ (and hence also superpolynomial in the size of the sequent). In the following theorem we assume that Conjecture 2.13 is true. We start with a balanced function f that is computable by a polynomial-size depth-k $\mathbf{ACC}^0[p]$ circuit that is hard to approximate by $\mathbf{ACC}^0[r]$ circuits, as stated by the conjecture. Each gate in the $\mathbf{ACC}^0[p]$ can be computed by a depth-$(2d+4)$ AND-OR formula, as shown by the lemma. So the whole circuit is computed by a depth-$k(2d+4)$ AND-OR formula.

Theorem 6.8. *Suppose that p is a prime that does not divide r and that Conjecture 2.13 is true. Let $f_{k(2d+4)}$ be the depth-$k(2d+4)$ AND-OR formula given by*

the lemma that expresses a balanced depth-k, polynomial-size $\mathbf{ACC}^0[p]$ function that is hard to approximate by depth-d $\mathbf{ACC}[r]$ circuits of size $2^{n^{1/(d+1)}}$. Consider the tautology $\mathbf{STATMAN}_n(f_{k(2d+4),m})$ with $m = (5n^2)^{d+1}$. Let N denote the size of this sequent. Then the following holds:

1. *$\mathbf{STATMAN}_n(f_{k(2d+4),m})$ has a $\mathbf{PK}^\star[r]$ proof of depth $k(2d+4)$ and size polynomial in N.*
2. *For sufficiently large n, any d-$\mathbf{PK}^\star[r]$ proof of $\mathbf{STATMAN}_n(f_{k(2d+4),m})$ has size at least $2^{(c\log N)^{1+1/(2d+2)}}$ (for some $c > 0$), which is superpolynomial in N.*

Proof. The lower bound follows from the main theorem (Theorem 4.3) just like the other lower bounds of this section. The upper bound is obtained by using the 1-$\mathbf{PK}^\star$ proof of the Statman tautologies (Theorem 2.10) and the fact that the cut formulas in that proof are all of the form $p_i \vee q_i$. When using this proof for $\mathbf{STATMAN}_n(f_{k(2d+4),m})$, the depth of these cuts becomes $k(2d+4)$. (Note that the sequent $\mathbf{STATMAN}_n(f_{k(2d+4),m})$ itself is of depth $k(2d+4)+2$.)

For the lower bound, as noted before the theorem, the size N of $\mathbf{STATMAN}_n(f_{k(2d+4),m})$ is a polynomial in n and

$$mp^{(2d+3)m^{\frac{1}{2d+3}}}$$

Therefore $\log N = O(n^{\frac{2d+2}{2d+3}})$, so $n > (c\log N)^{\frac{2d+3}{2d+2}}$ for some $c > 0$. □

At the beginning of this subsection, we mentioned that it was known that $\mathbf{AC}^0$-Frege proofs of depth d cannot p-simulate $\mathbf{AC}^0$-Frege proofs of depth $d+1$. The sequents that witness the separation are of depth at most d. (They must be for the $\mathbf{AC}^0$-Frege proofs of depth d to have any chance of proving them.) Based on the ideas in the proof of Theorem 2.4, this implies that these sequents show that d-$\mathbf{PK}^\star$ proofs cannot p-simulate $(d+1)$-$\mathbf{PK}^\star$ proofs.

By using our new lower bounds, we can give a simpler proof of the fact that the constant-depth $\mathbf{PK}^\star$ hierarchy is infinite. The reason we use depth $2d+4$ here is the same as for the previous theorem:

Theorem 6.9. *Let $\mathrm{MOD}_{2,2d+4}$ be one of the AND-OR formulas given by the Lemma 6.7. Consider the tautology $\mathbf{STATMAN}(\mathrm{MOD}_{2,2d+4,m})$ with $m = (5n^2)^{d+1}$. Let N denote the size of this sequent. Then the following hold:*

1. *$\mathbf{STATMAN}(\mathrm{MOD}_{2,2d+4,m})$ has a $\mathbf{PK}^\star$ proof of depth $2d+4$ and size polynomial in N.*

2. *For sufficiently large n, any d-$\mathbf{PK}^\star$ proof of $\mathbf{STATMAN}(\mathrm{MOD}_{2,2d+4,m})$ has size at least $2^{(c\log N)^{1+1/(2d+2)}}$ (for some $c > 0$), which is superpolynomial in N.*

6.5 Other Propositional Proof Systems

We briefly mention that our main theorem can be used to obtain results similar to those in this section but for other propositional proof systems. For example, the system $\mathbf{PTK}$ can be defined by adding threshold connectives and corresponding axioms and rules to the $\mathbf{PK}$ system. Constant-depth $\mathbf{PTK}$ proofs can then be defined like constant-depth $\mathbf{PK}$ proofs or constant-depth $\mathbf{PK}[r]$ proofs by limiting only the depth of the cuts. It is easy to verify that the proof of our main theorem applies to $\mathbf{PTK}^\star$ proofs and that the various applications of this section also apply to constant-depth $\mathbf{PTK}^\star$ proofs, assuming a conjecture similar to Conjecture 2.14, that there is a polynomial-size $\mathbf{NC}^1$ function that is hard to approximate by $\mathbf{TC}^0$ circuits.

7 Applications Beyond Propositional Proof Systems

In this section, we apply our general lower bound result and technique to obtain results that apply to other proof systems. First, we present a new proof of the non-finite axiomatizability of the bounded arithmetic theory $I\Delta_0(R)$. Then, we prove a conditional hierarchy theorem for the quantified propositional proof systems $\mathbf{G}_i^\star$.

7.1 Non-Finite Axiomatizability of $\mathbf{I\Delta}_0(R)$

We derive from Theorem 6.9 another proof of the fact that the relativized theory $\mathbf{I\Delta}_0(R)$ is not finitely axiomatizable (an earlier proof is given in [16]). We will present our argument for the two-sorted version of $\mathbf{I\Delta}_0(R)$, i.e., $\mathbf{\Sigma}_0^B(\mathbf{V}^0)$, the $\mathbf{\Sigma}_0^B$ consequences of $\mathbf{V}^0$. (The full theory $\mathbf{V}^0$ is associated with $\mathbf{AC}^0$ and serves as the base theory for the development in [8].)

The high-level idea of the proof is as follows. Suppose for a contradiction that $\mathbf{\Sigma}_0^B(\mathbf{V}^0)$ is finitely axiomatizable. Then by compactness it can be axiomatized by a finite set of induction axioms and probably some other basic axioms. Let $d \in \mathbb{N}$ be a common bound on the depth of all these axioms. By the Paris-Wilkie propositional translation, each theorem of $\mathbf{\Sigma}_0^B(\mathbf{V}^0)$ translates into a family of tautologies with polynomial-size d-$\mathbf{PK}^\star$ proofs. Now it is easy to see that the uniform version of the separating propositional sequents in Theorem 6.9 belongs to $\mathbf{\Sigma}_0^B(\mathbf{V}^0)$, and this gives a contradiction.

We refer to [8, Chapter 5] for basic definitions of $\mathbf{V}^0$. In short, there are two sorts of variables: the number variables $x, y, z, \ldots$ range over natural numbers $\mathbb{N}$, and the set (or string) variables $X, Y, Z, \ldots$ are meant to be finite subsets of $\mathbb{N}$. When presented as input to computing machines, set variables are given as binary strings while number variables are given in unary notation (and thus play only auxiliary role). The underlying language $\mathcal{L}_A^2$ is

$$\mathcal{L}_A^2 = [0, 1, +, \cdot, |X|;\ \leq, =_1, =_2, \in]$$

where $0, 1, +, \cdot, \leq, =_1$ are number functions and relations, $|X|$ is the length (with number value) of the string X which serves also as an upper bound for the elements of X, $\in$ is the membership relation, and $=_2$ is equality for sets. We often omit the subscripts in $=_1, =_2$, and also write $X(t)$ for $t \in X$ (we think of $X(i)$ as the i-th bit in the string representation of X). Note that the only string terms are string variables.

The bounded quantifiers are of the forms $\exists x \leq t$, $\forall x \leq t$, $\exists X \leq t$ and $\forall X \leq t$, where for the string quantifiers the bounding terms t bound the lengths of the string variables. $\mathbf{\Sigma}_0^B$ formulas are formulas with only bounded number quantifiers that may contain free string variables. The theory $\mathbf{V}^0$ is axiomatized by a set 2-$\mathbf{BASIC}$ of defining axioms for $\mathcal{L}_A^2$ together with comprehension axioms for $\mathbf{\Sigma}_0^B$ formulas, i.e., axioms of the form

$$\exists Y \leq b \forall y < t(Y(y) \leftrightarrow \varphi(y, Y))$$

for a $\mathbf{\Sigma}_0^B$ formula φ that might contain other free variables. It is known that $\mathbf{V}^0$ is a conservative extension of $\mathbf{I\Delta}_0$. Moreover, $\mathbf{V}^0$ is $\mathbf{\Sigma}_0^B$-conservative over the theory $\widetilde{\mathbf{V}}^0$ which is defined in the same way as $\mathbf{V}^0$ but with the comprehension axioms replaced by the induction axioms over $\mathbf{\Sigma}_0^B$ formulas:

$$[\varphi(0) \wedge \forall x,\ \varphi(x) \supset \varphi(x+1)] \supset \forall z \varphi(z)$$

(where $\varphi(z)$ is a $\mathbf{\Sigma}_0^B$ formula that may contain other free variables). In other words, $\widetilde{\mathbf{V}}^0$ can be axiomatized by the $\mathbf{\Sigma}_0^B$ consequences of $\mathbf{V}^0$.

We refer to [8, Chapter 7] for the translation of a first-order formulas into a family of propositional formulas. Basically, to translate a first-order formula $\varphi(\vec{x}, \vec{X})$ we give each number variable x a value $m \in \mathbb{N}$ and each string variable X a length n, and translate the bit $X(i)$ of X into a propositional variable p_i^X, for $0 \leq i \leq n-2$. (Other bits of X get constant values: $X(n-1) = \top$ and $X(i) = \bot$ for $i \geq n$.) The result is denoted by $\varphi(\vec{x}, \vec{X})[\vec{m}; \vec{n}]$.

It is known that $\mathbf{\Sigma}_0^B$ theorems of $\widetilde{\mathbf{V}}^0$ translate into families of tautologies that have polynomial size constant depth $\mathbf{PK}^\star$ proofs [8, Chapter 7]. This is done by

translating $\widetilde{\mathbf{V}}^0$ anchored (or free-cut free) proofs by translating their formulas as described above.

Now we give a uniform (i.e., first-order) version of the propositional sequents that separate d-$\mathbf{PK}^\star$ from $(2d+4)$-$\mathbf{PK}^\star$ (Theorem 6.9). This is constructed using the following uniform version of the propositional formula $\mathrm{MOD}_{2,d,m}$ from Lemma 6.7: Under the setting $a = 2^{m^{1/(d-1)}}$, $|X| = m+1$, the first-order formula $Parity_d(a, X)$ below translates into $\mathrm{MOD}_{2,d,m}$ (the parameter a is only to make sure that the length of X is not too large).

Lemma 7.1. *There is a constant $d_0 \in \mathbb{N}$ so that for every $d \geq d_0$, there is a $\mathbf{\Sigma}_0^B$ formula $Parity_d(a, X)$ of depth d (counting both quantifiers and Boolean connectives) such that for $m \in \mathbb{N}$, the propositional formula $Parity_d(a, X)[2^{m^{1/(d-1)}}; m+1]$ is $\mathrm{MOD}_{2,d,m}(p_0^X, p_1^X, \ldots, p_{m-1}^X)$.*

Proof. The formula $Parity_d(a, X)$ expresses the fact that there are u, v such that (u plays the role of m, and v plays the role of $m^{1/(d-1)}$)

(a) $u = v^{(d-1)}$ and $a = 2^{(d-1)v}$, and

(b) $|X| = u + 1$ and the string $X(0), X(1), \ldots, X(u-1)$ contains an odd number of $\top$.

Condition (a) is fulfilled by using the fact that the relation $y = 2^x$ can be expressed by a $\mathbf{\Sigma}_0^B$ formula (and the constant d_0 accounts for the depth of this formula). Condition (b) can be expressed by a $\mathbf{\Sigma}_0^B$ formula by the same arguments as for Lemma 6.7. □

Now we define the first-order version of Statman's sequent. For $d \geq d_0$, let

$$S_d = \qquad \longrightarrow \varphi_d(a, b, X, Y), \psi_d(a, b, X, Y)$$

where $\varphi_d(a, b, X, Y)$ and $\psi_d(a, b, X, Y)$ are defined (see below) so that $\varphi_d(a, b, X, Y)$ translates into

$$\big((\neg p_1 \wedge \neg q_1) \vee [\gamma_1 \wedge \neg p_2 \wedge \neg q_2] \vee \ldots \vee [\gamma_{n-1} \wedge \neg p_n \wedge \neg q_n]\big)(\mathrm{MOD}_{2,d,m})$$

and $\psi_d(a, b, X, Y)$ translates into

$$[(p_1 \vee q_1) \wedge \ldots \wedge (p_n \vee q_n)](\mathrm{MOD}_{2,d,m})$$

The translation of S_d is not exactly $\mathbf{STATMAN}_n(\mathrm{MOD}_{2,d,m})$ because here we put the first n formulas in $\mathbf{STATMAN}_n(\mathrm{MOD}_{2,d,m})$ in a disjunction. However we will be able to argue that it has the same lower bound 2^n as $\mathbf{STATMAN}_n(\mathrm{MOD}_{2,d,m})$.

Now we describe S_d in more detail. First, we want $\psi_d(a, b, X, Y)$ to translate into

$$\bigwedge_{i=1}^{n} (\mathrm{MOD}_{2,d,m}(x_1^i, x_2^i, \ldots, x_m^i) \vee \mathrm{MOD}_{2,d,m}(y_1^i, y_2^i, \ldots, y_m^i)) \tag{3}$$

For each i we have distinct sets of propositional variables $\overrightarrow{x^i}$ and $\overrightarrow{y^i}$. So we will view X and Y as arrays of strings by using the pairing function $\langle i, j\rangle$. The i-th row of X is denoted by $X^{[i]}$ and can be defined by a $\mathbf{\Sigma}_0^B$ formula. Thus define

$$\psi_d(a, b, X, Y) \equiv \forall i \leq b(Parity_d(a, X^{[i]}) \vee Parity_d(a, Y^{[i]}))$$

It can be verified that the propositional translation

$$\psi_d(a, b, X, Y)[2^{m^{1/(d-1)}}, n; \langle n, m+1\rangle, \langle n, m+1\rangle]$$

has the form (3) above.

The formula φ_d can be defined in the same way and we omit the details here. The next lemma follows from our discussion so far.

Lemma 7.2. *Under the setting $a = 2^{m^{1/(d-1)}}, b = n, |X| = |Y| = \langle n, m+1\rangle$ the sequent S_d translates into $\mathbf{STATMAN}'_n(\mathrm{MOD}_{2,d,m})$, where $\mathbf{STATMAN}'_n$ is the sequent (cf.* (1)*):*

$$\longrightarrow (\neg p_1 \wedge \neg q_1) \vee [\gamma_1 \wedge \neg p_2 \wedge \neg q_2] \vee \ldots \vee [\gamma_{n-1} \wedge \neg p_n \wedge \neg q_n], [(p_1 \vee q_1) \wedge \ldots \wedge (p_n \vee q_n)]$$

Now we argue that the sequent $\mathbf{STATMAN}'_n(\mathrm{MOD}_{2,2d+4,m})$ also requires large d-$\mathbf{PK}^\star$ proofs.

Lemma 7.3. *For $m = (5n^2)^{d+1}$, any d-$\mathbf{PK}^\star$ proof of the sequent $\mathbf{STATMAN}'_n(\mathrm{MOD}_{2,2d+4,m})$ has size at least $2^n/n$.*

Proof. For readability we argue that any cut-free proof of $\mathbf{STATMAN}'_n$ must have size at least $2^n/n$ by transforming proof of $\mathbf{STATMAN}'_n$ into proof of $\mathbf{STATMAN}_n$ with an increase of at most n multiplicative factor in size. The theorem can be proved by the same argument.

Any proof of $\mathbf{STATMAN}'_n$ can be transformed into a proof of $\mathbf{STATMAN}_n$ as follows. Simply replace every occurrence of

$$(\neg p_1 \wedge \neg q_1) \vee [\gamma_1 \wedge \neg p_2 \wedge \neg q_2] \vee \ldots \vee [\gamma_{n-1} \wedge \neg p_n \wedge \neg q_n] \tag{4}$$

by the list

$$(\neg p_1 \wedge \neg q_1),\ [\gamma_1 \wedge \neg p_2 \wedge \neg q_2],\ \ldots\ ,\ [\gamma_{n-1} \wedge \neg p_n \wedge \neg q_n]$$

Note that these can only appear in the antecedents. Now any contraction on (4) is simulated by n contractions on the formulas in the list. Also we do ignore the $\vee$-right rules that introduce disjunctions in (4). The proof of $\mathbf{STATMAN}_n$ obtained this way is of size at most n times the size of the original proof of $\mathbf{STATMAN}'_n$ (because of the increase in the number of contractions). Since $\mathbf{STATMAN}_n$ requires proof of size at least 2^n, $\mathbf{STATMAN}'_n$ requires proof of size at least $2^n/n$. □

We have used a "lazy" argument for the above lemma, which is sufficient for our application below. With a careful redefinition of Statman property (to allow disjunction) the lower bound proof of the main theorem (Theorem 4.3) goes through. It can then be seen that $\mathbf{STATMAN}'_n$ has this property, and hence requires d-$\mathbf{PK}^\star$ proof of size 2^n.

Theorem 7.4. *$\widetilde{\mathbf{V}}^0$ and $\mathbf{\Sigma}^B_0(\mathbf{V}^0)$ are not finitely axiomatizable.*

Proof. It suffices to show that $\widetilde{\mathbf{V}}^0$ is not finitely axiomatizable, because $\widetilde{\mathbf{V}}^0$ is axiomatized by the $\mathbf{\Sigma}^B_0$ consequences of $\mathbf{V}^0$. We follow the outline given at the beginning of this section.

Suppose for a contradiction that $\widetilde{\mathbf{V}}^0$ is axiomatized by a finite set S of formulas. Because $\widetilde{\mathbf{V}}^0$ can be axiomatized by 2-$\mathbf{BASIC}$ and the set of all induction axioms for $\mathbf{\Sigma}^B_0$ formulas, by compactness we can assume that S consists only of 2-$\mathbf{BASIC}$ and a finite set of induction axioms for $\mathbf{\Sigma}^B_0$ formulas. Let d_1 be a upper bound for the depth (counting both bounded number quantifiers and Boolean connectives) of the formulas in S. In other words, theorems of $\widetilde{\mathbf{V}}^0$ have $\widetilde{\mathbf{V}}^0$-proofs where cut formulas have depth at most d_1. (This follows by considering free-cut free, or anchored, proofs.)

It is shown [8, Chapter 7] that $\widetilde{\mathbf{V}}^0$ proofs translate into polynomial size constant depth $\mathbf{PK}^\star$ proofs. Under the current hypothesis it can be seen that theorems of $\widetilde{\mathbf{V}}^0$ translate into families of tautologies with polynomial size d_1-$\mathbf{PK}^\star$ proofs.

By induction on b it can be seen that $\mathbf{V}^0$ proves S_d for any d. In particular, $\mathbf{V}^0$ proves S_{2d+4} where $d = \max\{d_0, d_1\}$ (d_0 is the constant from Lemma 7.1). Thus the translations of S_{2d+4} have polynomial-size d_1-$\mathbf{PK}^\star$ proofs, contradicts Theorem 6.9. □

7.2 Extension to QBF Proof Systems

We now consider the system $\mathbf{G}$ [15, 7] which is an extension of $\mathbf{PK}$ for quantified Boolean formulas. There are quantifiers $\exists, \forall$ with the following semantic interpretation.

$$\exists x A(x) \Leftrightarrow A(\bot) \vee A(\top), \qquad \forall x A(x) \Leftrightarrow A(\bot) \wedge A(\top)$$

Also, here we restrict the Boolean connectives $\vee$, $\wedge$ to have arity 2. Thus formulas are defined inductively as follows:

(i) atomic formulas are Boolean constants $\bot$ and $\top$, and Boolean variables p_i and x_i;

(ii) if A and B are formulas, then so are $(A \vee B)$, $(A \wedge B)$, $\neg A$, $\exists x_i A$ and $\forall x_i A$.

The structural rules, the cut rule and the introduction rules for $\neg$ (NEG-left and NEG-right) are as for $\mathbf{PK}$. The introduction rules for $\vee$, $\wedge$ and the quantifiers are listed below:

$$\frac{A, B, \Gamma \longrightarrow \Delta}{(A \wedge B), \Gamma \longrightarrow \Delta} \text{ AND-left} \qquad \frac{\Gamma \longrightarrow A, \Delta \qquad \Gamma \longrightarrow B, \Delta}{\Gamma \longrightarrow (A \wedge B), \Delta} \text{ AND-right}$$

$$\frac{A, \Gamma \longrightarrow \Delta \qquad B, \Gamma \longrightarrow \Delta}{(A \vee B), \Gamma \longrightarrow \Delta} \text{ OR-left} \qquad \frac{\Gamma \longrightarrow A, B, \Delta}{\Gamma \longrightarrow (A \vee B), \Delta} \text{ OR-right}$$

$$\frac{A(B), \Gamma \longrightarrow \Delta}{\forall x A(x), \Gamma \longrightarrow \Delta} \forall\text{-left} \qquad \frac{\Gamma \longrightarrow \Delta, A(p)}{\Gamma \longrightarrow \Delta, \forall x A(x)} \forall\text{-right}$$

$$\frac{A(p), \Gamma \longrightarrow \Delta}{\exists x A(x), \Gamma \longrightarrow \Delta} \exists\text{-left} \qquad \frac{\Gamma \longrightarrow \Delta, A(B)}{\Gamma \longrightarrow \Delta, \exists x A(x)} \exists\text{-right}$$

Restriction: In the rules $\forall$-right and $\exists$-left, p must not occur in the bottom sequent.

For $i \geq 0$, $\mathbf{\Sigma}_i^q$ (resp. $\mathbf{\Pi}_i^q$) is the set of formulas that have a prenex form where there are at most i alternations of quantifiers, with the outermost quantifier being $\exists$ (resp. $\forall$). In particular, $\mathbf{\Sigma}_0^q$ and $\mathbf{\Pi}_0^q$ both denote the set of quantifier-free propositional formulas. The system $\mathbf{G}_i$ is the subsystem of $\mathbf{G}$ in which all cut formulas belong to $\mathbf{\Sigma}_i^q \cup \mathbf{\Pi}_i^q$. $\mathbf{G}_i^\star$ denotes tree-like $\mathbf{G}_i$.

It is known that $\mathbf{G}_{i+1}^\star$ and $\mathbf{G}_i$ are p-equivalent for $\mathbf{\Sigma}_i^q \cup \mathbf{\Pi}_i^q$ formulas, and Perron [23] shows that $\mathbf{G}_i$ p-simulates $\mathbf{G}_{i+1}^\star$ for all quantified formulas. Here we will show that under some complexity theoretic assumption $\mathbf{G}_i^\star$ does not simulate cut-free $\mathbf{G}$. We need to show that $\mathbf{G}$ is closed under restrictions (Definition 2.7). First we extend Definition 2.5 to define restrictions of quantified formulas.

Definition 7.5. (Restriction of a quantified formula) *The restriction $f|_\rho$ of a quantified formula f is defined as in Definition 2.5 with the following additional case:*

6. *If $f = \exists x A$ and $A|_\rho$ does not contain any free occurrence of x, then $f|_\rho = A|_\rho$. Otherwise, $f|_\rho = \exists x(A|_\rho)$. Similarly for $f = \forall x A$.*

The result of applying restriction to a sequent of QBF is defined as in Definition 2.6.

Lemma 7.6. $\mathbf{G}$ *is closed under restrictions.*

Proof. We extend the proof of Lemma 2.8. The additional cases are the introduction rules for $\exists$ and $\vee$. Consider, for example, the case of $\exists$-right:

$$\frac{S_1}{S} = \frac{\Gamma \longrightarrow \Delta, A(B)}{\Gamma \longrightarrow \Delta, \exists x A(x)}$$

First, suppose that all free occurrences of x is deleted from the restriction A' of $A(x)$. Then B is also deleted from $(A(B))'$. By definition $(\exists x A(x))' = A'$. So in this case $S' = S'_1$, and no further derivation is required.

Now suppose that some free occurrence of x remains in $(A(x))'$. Then $(A(B))'$ has the form $A'(B')$, and S' can be obtained from S'_1 by the rule $\exists$-right with target formula B'. □

The following theorem is proved in the same way as the results in Section 6.

Theorem 7.7. *Let $i > j \geq 0$.*

(a) Suppose that there is a Boolean function f that is definable by a family f_m of QBF formulas and that is (σ, ϵ)-hard for $\mathbf{\Sigma}^q_j$ for some functions $\sigma(m)$, $\epsilon(m)$ satisfying the following condition: For sufficiently large n there is m that meets Condition (2) *of Theorem 4.3, and such that 2^n is superpolynomial in the size of f_m. Then $\mathbf{G}^\star_j$ does not simulate $\mathbf{G}^\star$ as well as cut-free $\mathbf{G}$.*

(b) Suppose that there is a function f as in (a) but now the family f_m defining f belongs to $\mathbf{\Sigma}^q_i$. Then $\mathbf{G}^\star_j$ does not simulate $\mathbf{G}^\star_i$.

It is known that $\mathbf{G}^\star_0$ p-simulates $\mathbf{G}_0$ for $\mathbf{\Sigma}^q_1$ formulas *in prenex form* [18]. It is still consistent with our knowledge that the hard formula for quantifier-free formulas in the hypothesis of the theorem belongs to $\mathbf{\Sigma}^q_1$. This is because the formulas in our separating sequent are not in prenex form (although they are in $\mathbf{\Sigma}^q_1$).

8 Conclusion

In this paper we have presented a general method for taking a family of sequents that require large tree-like cut-free proofs, and "lifting" them in order to obtain a family of sequents that are hard for stronger classes of tree-like proof systems. An obvious open problem is to prove similar lower bounds without the tree-like restriction. While the methods used in this paper cannot be adapted straightforwardly, we nevertheless feel that our "lifting" approach should be generalizable to non-tree-like systems. For non-tree-like proofs, an obvious way to generalize our

argument would be to start with the basic lower bound technique for a dag-like cut-free system (i.e., resolution), rather than starting with the basic lower bound technique for tree-like cut-free proofs, and prove a similar result to our main theorem, where the size-width/bottleneck-counting technique used to obtain resolution lower bounds [10, 4] replaces the Statman lower bound method.

Secondly, we would like to develop a new, general-purpose method for obtaining $\mathbf{AC}^0$-Frege lower bounds for CNF formulas. For example, can we obtain a top-down strategy for the liar game formulation of $\mathbf{AC}^0$-Frege for the PHP? Toward this end, we would like to know whether inapproximability results are enough to prove lower bounds for CNF formulas. For example, can we reduce the $\mathbf{AC}^0$-Frege lower bound for some CNF formula to a natural hardness assumption about $\mathbf{AC}^0$, such as the inapproximability of parity by $\mathbf{AC}^0$ circuits? The only known proofs require structural information about $\mathbf{AC}^0$, such as the fact that under a special family of restrictions, an $\mathbf{AC}^0$ function reduces to a local function (a small-depth decision tree, or a function depending on only a constant number of variables).

Thirdly, in our last application we show that the G_i^* hierarchy does not collapse to G_1^* unless SAT can be approximated by polynomial-size circuits. In contrast, it has been known that the S_2^i hierarchy does not collapse to S_2^1 unless the polynomial hierarchy collapses. We would like to know how these assumptions compare to one another. In particular, do polynomial-size circuits approximating SAT imply the collapse of the polynomial-time hierarchy?

Acknowledgements

We would like to thank Alasdair Urquhart and Stephen Cook for many helpful comments and references. Phuong Nguyen would like to thank the referees of his LICS07 paper for helpful feedback. Toniann Pitassi would like to thank Maria Bonet for early conversations many years ago that led to this research.

References

[1] M. Ajtai. The complexity of the pigeonhole principle. In *Proc. 29th Ann. IEEE Symp. on Foundations of Computer Science*, pages 346–355, 1988.

[2] Paul Beame, Russell Impagliazzo, Jan Krajíček, Toniann Pitassi, and Pavel Pudlák. Lower Bounds on Hilbert's Nullstellensatz and Propositional Proofs. *Proceedings of the London Mathematical Society*, 73(3):1–26, 1996.

[3] M. Bonet, C. Domingo, R. Gavalda, A. Maciel and T. Pitassi. Non-automatizability of bounded-depth Frege proofs. *Computational Complexity*, 13:47–68, 2004.

[4] E. Ben-Sasson and A. Wigderson. Short proofs are narrow – resolution made simple. In *Proceedings of ACM Symposium on Theory of Computing*. 1999.

[5] S. Buss. Weak Formal Systems and Connections to Computational Complexity. Lecture notes, University of California, Berkeley. 1988.

[6] P. Clote and E. Kranakis. *Boolean Functions and Computation Models*. Springer, 2002.

[7] S. Cook and T. Morioka. Quantified Propositional Calculus and a Second-Order Theory for NC1. *Archive for Mathematical Logic*, 44(6):711–749, 2005.

[8] S. Cook and P. Nguyen. *Logical Foundations of Proof Complexity*. ASL Perspectives in Logic, Cambridge University Press, 2010.

[9] M. Furst and J.B. Saxe and M .Sipser. Parity, Circuits, and the Polynomial Time Hierarchy. *Mathematical Systems Theory*, 17(1):13–27, 1984.

[10] A. Haken. The intractability of resolution. *Theoretical Computer Science*, 39:297–308, 1985.

[11] J. Håstad. *Computational limitations for small depth circuits.* Phd thesis, Massachusetts Institute of Technology, 1986.

[12] A. Haken and S.A. Cook. An exponential lower bound for the size of monotone real circuits. *Journal of Computer and System Sciences*, 58:326–335, 1999.

[13] J. Krajíček. Lower bounds to the size of constant-depth propositional proofs. *Journal of Symbolic Logic*, 59(1):73–86, 1994.

[14] J. Krajíček. *Bounded Arithmetic, Propositional Logic and Computational Complexity.* Cambridge University Press, 1995.

[15] J. Krajíček and P. Pudlák. Quantified Propositional Calculi and Fragments of Bounded Arithmetic. *Zeitschrift f. Mathematkal Logik u. Grundlagen d. Mathematik*, 36:29–46, 1990.

[16] J. Krajíček, P. Pudlák and G. Takeuti. Bounded arithmetic and the polynomial hierarchy. *Annals of Pure and Applied Logic*, 52:143–153, 1991.

[17] J. Krajíček, P. Pudlák and A. Woods. Exponential lower bound to the size of bounded depth Frege proofs of the pigeonhole principle. *Random Structures and Algorithms*, 7:15–39, 1995.

[18] T. Morioka. Logical Approaches to the Complexity of Search Problems: Proof Complexity, Quantified Propositional Calculus, and Bounded Arithmetic. PhD thesis, University of Toronto, 2005.

[19] A. Maciel and T. Pitassi. A Conditional Lower Bound for a System of Constant-Depth Proofs with Modular Connectives. In *Proc. 21st Ann. IEEE Symp. on Logic in Computer Science*, 2006.

[20] A. Maciel, T. Pitassi and A. Woods. A New Proof of the Weak Pigeonhole Principle. *Journal of Computer Systems Sciences*, 64:843–872, 2002.

[21] P. Nguyen. Separating DAG-Like and Tree-Like Proof Systems. In *Proc. 22nd Ann. IEEE Symp. on Logic in Computer Science*, 2007.

[22] J.B. Paris, A.J. Wilkie and A.R. Woods. Provability of the pigeonhole principle and the existence of infinitely many primes. *Journal of Symbolic Logic*, 53:1235–1244, 1988.

[23] S. Perron. Power of Non-Uniformity in Proof Complexity. PhD thesis, University of Toronto, 2008.

[24] T. Pitassi, P. Beame and R. Impagliazzo. Exponential lower bounds for the pigeonhole principle. *Computational Complexity*, 97–140, 1993.

[25] P. Pudlák. Lower bounds for resolution and cutting plane proofs and monotone computations. *Journal of Symbolic Logic*, 62(3):981–998, September 1997.

[26] A.A. Razborov. Lower Bounds for the Size of Circuits with Bounded Depth with Basis $\{\wedge, \oplus\}$. *Mat. Zametki*, 41(4):598–607, 1987. English transl. in *Math. Notes of the Academy of Sci. of the USSR*, 41(4):333–338, 1987.

[27] R. Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proc. 19th Ann. ACM Symp. on Theory of Computing*, 1987.

[28] R. Smolensky. On Representations by Low-Degree Polynomials. In *Proc. 34th Ann. Symp. on Foundations of Computer Science*, 1993.

[29] R. Statman. Bounds for Proof-Search and Speed-up in the Predicate Calculus. *Annals of Mathematical Logic*, 15:225–287, 1978.

[30] A. Yao. Separating the Polynomial Hierarchy by Oracles: Part I. In *Proc. 26th Ann. Symp. on Foundations of Computer Science*, 1985.