Notre Dame Journal of Formal Logic
Volume 37, Number 4, Fall 1996

Simplified Lower Bounds for Propositional Proofs

ALASDAIR URQUHART and XUDONG FU

Abstract This article presents a simplified proof of the result that bounded depth propositional proofs of the pigeonhole principle are exponentially large. The proof uses the new techniques for proving switching lemmas developed by Razborov and Beame. A similar result is also proved for some examples based on graphs.

1 Introduction

Substantial progress has been made recently in proving lower bounds on the complexity of propositional proofs. A decisive advance was made by Ajtai [1], who proved superpolynomial lower bounds on the size of bounded depth Frege proofs of tautologies encoding the pigeonhole principle. In later work, Bellantoni, Pitassi, and Urquhart [4] simplified Ajtai's proof, and subsequently Krajíček, Pudlák, and Woods, and Pitassi, Beame, and Impagliazzo in [3], [10], and [12] independently extended it to prove exponential lower bounds for the same set of tautologies.

All of the arguments establishing these lower bounds use a type of combinatorial argument known generically as a *switching lemma*. In its simplest form (see Furst, Saxe, and Sipser [7], Yao [17], and Håstad [9]) this type of argument shows that if a formula in conjunctive normal form is simplified by a random partial assignment of truth-values to its variables (a *random restriction*) then with high probability it can be written in disjunctive normal form, where the conjuncts are not too large. Thus a random partial assignment of truth-values allows us to switch efficiently between conjunctive and disjunctive normal form. In a similar way, the switching lemmas used in other cases show that the use of random restrictions allows a formula of bounded depth to be represented by a formula in disjunctive normal form composed of small terms.

Proofs of the switching lemmas used in papers on the complexity of proofs [1, 3, 4, 10, 12] are significantly more complex than the proofs in Boolean complexity theory on which they are modeled [7, 9, 17]. In these lemmas, the variables are not

Received October 17, 1995; revised May 12, 1996

independent (as they are in the simpler case), and this fact leads to rather complex and delicate arguments involving conditional probability. However, recent work of Razborov [13] and Beame [2] has led to drastic simplification of these proofs.

The main purpose of the present paper is to present a fully detailed version of the resulting simplified proof of exponential lower bounds for bounded depth Frege proofs of the pigeonhole tautologies. Its principal contribution is to show how the rather complicated combinatorial and probabilistic arguments for the lower bounds used in the earlier papers [10, 12] can be replaced by the much simpler counting techniques of Razborov and Beame.

The proof techniques used for this proof can also be used to give exponential lower bounds for tautologies based on graphs. (The original idea for these tautologies is due to Tseitin [15].) In a later section of the paper a proof of this lower bound is sketched for a family of tautologies based on complete graphs. The last section of the paper gives a few open problems that appear to require new ideas going beyond the techniques expounded here.

2 *Frege systems*

The proof systems we consider are those familiar from textbook presentations of logic, consisting of a finite number of axiom schemes and schematic rules. We call such a system a *Frege system*. (Strictly speaking, this is a misnomer, since Frege's original system [8] included a tacitly applied rule of substitution; the use of schematic rules to avoid the rule of substitution is a device of von Neumann [11].)

The language for propositional logic used here is based on binary disjunction $\vee$ and negation $\neg$; a conjunction $A \wedge B$ is treated as an abbreviation for the formula $\neg(\neg A \vee \neg B)$. In addition, we include the propositional constants 0 and 1 standing for "false" and "true" respectively. The set of propositional variables will be specified in the following section. If A is a formula and $p_1, \ldots, p_m$ a sequence of variables then we write $A[B_1/p_1, \ldots, B_m/p_m]$ for the formula resulting from A by substituting $B_1, \ldots, B_m$ for $p_1, \ldots, p_m$.

A *Frege rule* is defined to be a sequence of formulas written in the form $A_1, \ldots, A_k \vdash A_0$. In the case in which the sequence $A_1, \ldots, A_k$ is empty, the rule is referred to as an *axiom scheme*. The rule is *sound* if $A_1, \ldots, A_k \models A_0$, that is, if every truth-value assignment satisfying $A_1, \ldots, A_k$ also satisfies A_0. If $A_1, \ldots, A_k \vdash A_0$ is a Frege rule, then C_0 is *inferred from* $C_1, \ldots, C_k$ by this rule if there is a sequence of formulas $B_1, \ldots, B_m$ and variables $p_1, \ldots, p_m$ so that for all i, $0 \leq i \leq k$, $C_i = A_i[B_1/p_1, \ldots, B_m/p_m]$.

If $\mathcal{F}$ is a set of Frege rules and A a formula, then a *proof of A in $\mathcal{F}$ from* $A_1, \ldots, A_m$ is a finite sequence of formulas such that every formula in the sequence is one of $A_1, \ldots, A_m$ or inferred from earlier formulas in the sequence by a rule in $\mathcal{F}$, and the last formula is A. The formulas in the sequence are the *lines* in the proof.

If $\mathcal{F}$ is a set of Frege rules, then it is *implicationally complete* if whenever $A_1, \ldots, A_m \models A_0$ then there is a proof of A_0 in $\mathcal{F}$ from $A_1, \ldots, A_m$. A *Frege system* is defined to be a finite set of sound Frege rules that is implicationally complete.

Example 2.1 Shoenfield's system ([14], p. 21), in which the primitive connectives are $\vee$ and $\neg$ follows:

Excluded Middle:	$\vdash \neg p \vee p$;
Expansion rule:	$p \vdash q \vee p$;
Contraction rule:	$p \vee p \vdash p$;
Associative rule:	$p \vee (q \vee r) \vdash (p \vee q) \vee r$;
Cut rule:	$p \vee q, \neg p \vee r \vdash q \vee r$.

If Γ is a sequence of formulas then the *size* of Γ is the number of distinct subformulas in Γ. In particular, we define the size of a Frege rule $A_1, \dots, A_k \vdash A_0$ to be the size of the sequence $A_1, \dots, A_k, A_0$. For example, the size of the cut rule in Shoenfield's system is 7.

A formula can be represented by its formation tree in which the leaves are labeled with propositional variables or constants, and an interior node is labeled with $\vee$ if it is the parent of two nodes, and with $\neg$ if it is the parent of only one. A branch in the tree representing a formula, when traversed from the root to the leaf at the end of the branch is labeled with a block of operators of one kind (say $\neg$), followed by a block of the other kind (say $\vee$), ..., ending with a variable or constant. The *logical depth* of a branch is defined to be the number of blocks of operators labeling the branch. The *depth* of a formula is the maximum logical depth of the branches in its formation tree.

Example 2.2 The formula $(\neg p \vee \neg\neg 1) \vee \neg(\neg q \vee r)$ has depth 4.

The depth of a proof in a Frege system is the maximum depth of a line in the proof. The lower bound proved in this paper is for proofs of bounded depth, in which all formulas have depth bounded by a fixed constant.

3 Matchings and restrictions In this section we introduce a language for the propositional pigeonhole principle and a space of matchings that serve to define restrictions on the propositional variables in the language.

Let D, R be finite nonempty sets where $D \cap R = \varnothing$, and let $S = D \cup R$. We shall suppose that S is ordered, with all elements of D preceding elements in R, and refer to this ordering as ordering by *size*. (Later, a different ordering on a subset of S plays an important role.) A *matching* between D and R is a set of mutually disjoint unordered pairs $\{i, j\}$, where $i \in D$, $j \in R$ (that is to say, a matching in the complete bipartite graph $D \times R$). A matching *covers* a vertex i if $\{i, j\}$ belongs to the matching for some vertex j; a matching covers a set X if it covers all the vertices in X. If $X \subseteq S$, then $M(X)$ denotes the set of all matchings ρ such that ρ covers X, but no matching properly contained in ρ covers X. If π is a matching then we denote by $V(\pi)$ the set of vertices covered by π. A matching between D and R is *perfect* if it covers all of the vertices in $D \cup R$.

The pigeonhole principle states that if $|D| = n + 1$, $|R| = n$ then there is no perfect matching between D and R. To formalize this as a tautology in propositional logic we introduce propositional variables P_{ij} for $i \in D$, $j \in R$. The language built from these variables and the constants 0 and 1 using the connnectives $\vee$ and $\neg$ we shall refer to as $L(D, R)$; we also refer to the language as L_n in contexts where D, R are understood as the basic sets. The tautology $PHP(D, R)$ is the disjunction

$$\bigvee_{\substack{i \neq j \in D \\ k \in R}} (P_{ik} \wedge P_{jk}) \vee \bigvee_{\substack{i \neq j \in R \\ k \in D}} (P_{ki} \wedge P_{kj}) \vee \bigvee_{i \in D} \bigwedge_{k \in R} \neg P_{ik} \vee \bigvee_{k \in R} \bigwedge_{i \in D} \neg P_{ik}.$$

We shall also refer to this as PHP_n when the underlying sets are understood.

Let D, R be fixed, where $|D| = n+1$, $|R| = n$. The set of matchings between D and R we shall denote by M_n. A matching π determines a *restriction* ρ_π of the variables of L_n by the following definition. For a variable P_{ij}, if i or j is covered by π then $\rho_\pi(P_{ij}) = 1$ if $\{i, j\} \in \pi$, $\rho_\pi(P_{ij}) = 0$ if $\{i, j\} \notin \pi$; otherwise $\rho_\pi(P_{ij})$ is undefined. Since a matching uniquely determines and is determined by the corresponding restriction, we shall identify a matching with the restriction it determines, and refer to it according to context as a matching or a restriction. If ρ_1 and ρ_2 are two matchings in M_n, and $\rho_1 \cup \rho_2$ is also a matching, then we say that they are *compatible*. If ρ_1 and ρ_2 are compatible matchings, then their union will be written as $\rho_1\rho_2$. If ρ is a matching, then $D\restriction\rho = D \setminus V(\rho)$, $R\restriction\rho = R \setminus V(\rho)$ and $S\restriction\rho = S \setminus V(\rho)$. If M is a set of matchings, and ρ a matching, then $M\restriction\rho$ is defined to be $\{\rho' \setminus \rho : \rho' \in M,\ \rho' \text{ compatible with } \rho\}$.

If A is a formula of L_n, and $\rho \in M_n$, then we denote by $A\restriction\rho$ the formula resulting from A by substituting for the variables in A the constants representing their value under ρ. That is to say, if P_{ij} is set to 1 or 0 by ρ, then we substitute 1 or 0 for P_{ij}, otherwise the variable is unchanged. If Γ is a set of formulas and $\rho \in M_n$ then $\Gamma\restriction\rho$ is $\{A\restriction\rho : A \in \Gamma\}$. The formula $A\restriction\rho$ can be simplified by eliminating the constants by the rules $\neg 0 \equiv 1$, $\neg 1 \equiv 0$, $(0 \vee A) \equiv A$, $(A \vee 0) \equiv A$, $(1 \vee A) \equiv 1$, $(A \vee 1) \equiv 1$. If a formula A can be simplified to a formula B using these rules, then we write $A \equiv B$.

The language L_n contains only binary disjunction. However, in the proofs that follow it is convenient to introduce an auxiliary language that uses unbounded conjunctions and disjunctions. We shall distinguish the order of the terms in such conjunctions and disjunctions.

Let A be an unbounded conjunction each of whose conjuncts is a variable of L_n or a constant. We shall say that A is a *matching term* if the set of pairs $\{i, j\}$ for P_{ij} a variable in A forms a matching. The *size* of a matching term is the cardinality $|\pi|$ of the matching π corresponding to it; the set of vertices $V(A)$ associated with a matching term A is the set of vertices mentioned in the variables in A, that is, the set $V(\pi)$. If π is a matching, then we shall write $\wedge\pi$ for the matching term that describes it, the conjunction containing the set of variables P_{ij} for $\{i, j\} \in \pi$ as conjuncts.

An unbounded disjunction of matching terms we shall call a *matching disjunction*; it is a *matching disjunction over S* if all the vertices mentioned in it are in S. If all of the matching terms in a matching disjunction have size bounded by r, then it is an *r-disjunction*.

Let A be a disjunction in the language L_n, and A_i, $i \in I$, those subformulas of A that are not disjunctions, but every subformula of A properly containing them is a disjunction. Then the *merged form of A* is the unbounded disjunction $\bigvee_{i\in I} A_i$.

4 Matching trees In the present section, we introduce decision trees in which the branches represent matchings. We assume that the space of matchings is the set M_n of matchings between D and R, where $|D| = n+1$, $|R| = n$, $S = D \cup R$. The leaves of all trees are assumed to be ordered left to right. The nodes lying immediately below a node in a tree are its *children*. The *depth of a tree T*, $|T|$, is the maximum length of a branch in T.

Definition 4.1 A *matching tree* over S is a tree T satisfying the following conditions.

1. The nodes of T other than the leaves are labeled with vertices in S.
2. If a node in T is labeled with a vertex $i \in S$, then the edges leading out of the node are labeled with distinct pairs of the form $\{i, j\}$ where $j \in R$ if $i \in D$, $j \in D$ if $i \in R$.
3. No node or edge label is repeated on a branch of T.
4. If p is a node of T then the edge labels on the path from the root of T to p determine a matching $\pi(p)$ between D and R.

We shall use the notation $Br(T)$ for the set of matchings determined by the branches of T, that is, $\{\pi(l) : l \text{ a leaf in } T\}$. If M is a set of matchings, then T is said to be *complete for M* if for any node p in T labeled with a vertex $i \in S$, the set of matchings $\{\pi(q) : q \text{ a child of } p\}$ consists of all matchings in M of the form $\pi(p) \cup \{\{i, j\}\}$. If the space of matchings is M_n, we shall use the abbreviation "complete" instead of "complete for M_n".

Definition 4.2 Let X be a set of nodes in S. The *full matching tree T for X over S* is constructed as follows. If p is a node in T such that $\pi(p)$ does not cover X, then p is labeled with the first node i in X not covered by $\pi(p)$, and the set $\{\pi(q) : q \text{ a child of } p\}$ consists of all matchings in S of the form $\pi(p) \cup \{\{i, j\}\}$, for $j \in S$.

If T is the full matching tree for X over S, then $Br(T) = M(X)$. Every full matching tree for a subset of S is complete, but not every complete matching tree is a full matching tree for some subset of S.

Lemma 4.3 *Let T be a complete matching tree over the space $S = D \cup R$, $|D| = n + 1$, $|R| = n$, and ρ a matching in M_n such that $|\rho| + |T| \leq n$. Then there is a $\pi \in Br(T)$ such that $\pi \cup \rho \in M_n$.*

Proof: We show that by successively choosing nodes in T starting at the root we can find a branch in T so that the required π labels the chosen path. Let us suppose that the nodes have been chosen as far as a node p that is not a leaf. By assumption, $\rho \cup \pi(p) \in M_n$; since $|\rho| + |T| \leq n$, $|\rho \cup \pi(p)| < n$. Let i be the vertex in S labeling node p; there exists at least one matching extending $\rho \cup \pi(p)$ that covers i. Since T is complete, at least one edge below p is labeled with a pair that extends $\rho \cup \pi(p)$ to a matching in M_n. Then we can extend the path by choosing the node at the end of this edge. □

If the leaves of a matching tree T are each labeled with 0 or 1, then it is a *matching decision tree*. We define for $i = 0, 1$,

$$Br_i(T) = \{\pi(l) : l \text{ is a leaf of } T \text{ labeled } i\}.$$

If T is a matching decision tree, then T^c is the matching decision tree that results by changing the leaf labels of T from 0 to 1 and 1 to 0, while $Disj(T)$ is the unbounded disjunction $\bigvee\{\wedge\pi : \pi \in Br_1(T)\}$.

Lemma 4.4 *If T is a matching decision tree, and ρ extends a matching $\pi(l) \in Br(T)$, then $Disj(T)\restriction\rho \equiv 0$ or 1 according to whether l is labeled 0 or 1.*

Proof: If l is labeled 1, then since ρ extends $\pi(l)$, the term $\wedge\pi(l)$ is set to 1 by ρ, so that $Disj(T)\upharpoonright\rho \equiv 1$. If l is labeled 0, then we need to establish that for any leaf l' labeled 1, $\wedge\pi(l')\upharpoonright\rho \equiv 0$. Let p be the node at which the branches ending in l and l' diverge. If i is the vertex in S labeling p, then $\pi(l)$ and $\pi(l')$ must disagree on the vertex matched with i. Thus $\wedge\pi(l')\upharpoonright\rho \equiv 0$, showing that $Disj(T)\upharpoonright\rho \equiv 0$. □

Definition 4.5 Let $F = C_1 \vee \cdots \vee C_m$ be a matching disjunction over S. The *canonical matching decision tree for F over S*, $Tree_S(F)$, is defined inductively as follows.

1. If $F \equiv 0$ then $Tree_S(F)$ is a single node labeled 0; if $F \equiv 1$ then $Tree_S(F)$ is a single node labeled 1;
2. If $F \not\equiv 0$, $F \not\equiv 1$, let C be the first matching term in F such that $C \not\equiv 0$. Then $Tree_S(F)$ is constructed as follows.
   (a) Construct the full matching tree for $V(C)$ over S.
   (b) Replace each leaf l of the full matching tree for $V(C)$ by the canonical matching decision tree $Tree_{S\upharpoonright\pi(l)}(F\upharpoonright\pi(l))$.

In a canonical matching decision tree, certain nodes are singled out as *boundary nodes* and are specified by induction as follows: the boundary nodes of $Tree_S(F)$ are the root of $Tree_S(F)$ together with all the boundary nodes in the trees $Tree_{S\upharpoonright\pi(l)}(F\upharpoonright\pi(l))$ that form subtrees of $Tree_S(F)$ by clause 2b of the previous definition.

Example 4.6 Let $D = \{1, 2, 3, 4, 5\}$, $R = \{6, 7, 8, 9\}$, $S = D \cup R$, let F be the matching disjunction $(P_{17} \wedge P_{38}) \vee (P_{16} \wedge P_{27}) \vee (P_{56} \wedge P_{49}) \vee (P_{16} \wedge P_{59})$, and ρ the mapping $\{1 \mapsto 6\}$. Figure 1 shows the canonical tree $Tree_{S\upharpoonright\rho}(F\upharpoonright\rho)$. The filled-in nodes in the diagram represent boundary nodes in the canonical tree.

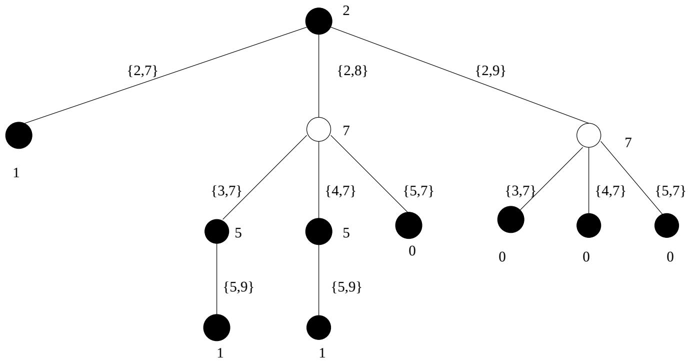

Figure 1: A canonical tree

If F is a matching disjunction, and T a matching decision tree, then we say that T *represents* F if for every $\pi(l) \in Br(T)$, $F\upharpoonright\pi(l) \equiv 1$ if l is labeled 1, and $F\upharpoonright\pi(l) \equiv 0$ if l is labeled 0. By construction, the canonical matching decision tree $Tree_S(F)$ represents F.

Definition 4.7 Let T be a complete matching decision tree and ρ a restriction. Then the tree $T\upharpoonright\rho$ that results from T by applying the restriction ρ is defined inductively as follows.

1. If T consists of a single node, then $T \restriction \rho$ is T.
2. If T consists of more than one node, and the root of T is labeled with the vertex i, then
   (a) if for some j, $\{i, j\} \in \rho$, then $T \restriction \rho$ is the decision tree $T' \restriction \rho$, where T' is the subtree attached to the root by the edge labeled with $\{i, j\}$;
   (b) if $i \notin V(\rho)$, then $T \restriction \rho$ has as its root a vertex labeled i, and as immediate subtrees all subtrees of the form $T' \restriction \rho$, where T' is attached to the root of T by an edge labeled $\{i, k\}$, where $k \notin V(\rho)$—the same pair labels the edge attaching $T' \restriction \rho$ to the root of $T \restriction \rho$.

In the following sections, restrictions are constructed by a process of successive extension. The following lemma guarantees that these extensions preserve certain relations.

Lemma 4.8 *Let T be a matching decision tree and ρ a restriction.*

1. $Disj(T) \restriction \rho \equiv Disj(T \restriction \rho)$.
2. *If T is complete for M_n, then $T \restriction \rho$ is complete for $M_n \restriction \rho$.*
3. $(T \restriction \rho)^c = T^c \restriction \rho$.
4. *If l is a leaf in $T \restriction \rho$, then there is a leaf l' in T bearing the same label as l so that $\pi(l') \subseteq \pi(l) \cup \rho$.*
5. *If T represents a matching disjunction F, then $T \restriction \rho$ represents $F \restriction \rho$.*

Proof: The first four parts of the lemma are proved by induction on the depth of the tree T. The fifth part follows from the fourth. □

5 Evaluations In this section, we introduce the basic concept of a k-evaluation: a k-evaluation can be considered as a kind of nonstandard truth-definition for a set of formulas. The notion of k-evaluation is due to Krajíček, Pudlák, and Woods [10]. The definition of k-evaluation used here differs from that of [10]; in that paper a more general definition is used in which formulas are assigned sets of restrictions rather than complete decision trees.

Definition 5.1 Let Γ be a set of formulas of L_n, closed under subformulas, where $S = D \cup R$, $|D| = n + 1$, $|R| = n$. Let $k > 0$. A k-evaluation T is an assignment of complete matching decision trees $T(A)$ to formulas $A \in L_n$ so that

1. $T(A)$ has depth $\leq k$;
2. $T(1)$ is the tree with a single node labeled 1, and $T(0)$ is the tree with a single node labeled 0;
3. $T(P_{ij})$ is the full matching tree for $\{i, j\}$ over S, with a leaf l labeled 1 if $\pi(l)$ contains $\{i, j\}$, otherwise 0;
4. $T(\neg A) = T(A)^c$;
5. if A is a disjunction, and $\bigvee_{i \in I} A_i$ is the merged form of A then $T(A)$ represents $\bigvee_{i \in I} Disj(T(A_i))$.

If T is a k-evaluation for a set of formulas Γ, then the set of matchings $Br(T(A))$ can be considered as a space of truth-value assignments for A; thus if $T(A)$ has all its leaves labeled 1, we can think of A as a kind of "tautology" relative to this space.

However, in contrast to the classical notion of tautology, this notion is not preserved under classically sound inferences (this fact is the key to the lower bound argument).

Example 5.2 Let $D = \{1, 2, 3\}$ and $R = \{4, 5\}$, and let $\Gamma = \{P_{14} \vee P_{15}, \neg P_{15} \vee \neg P_{25}, P_{14} \vee \neg P_{25}\}$. Then there is a 2-evaluation for Γ so that the first two formulas in Γ have 1 on all their leaves, but the third formula does not, although it is a logical consequence of the first two.

The following lemma shows that examples like this do not exist if the depth of a k-evaluation is small enough relative to the size of the inference rules of the proof system.

Lemma 5.3 *Let $\mathcal{F}$ be a Frege system in which the size of the rules is bounded by f, and $\mathcal{P}$ a proof in $\mathcal{F}$ in the language $L(D, R)$, where $S = D \cup R$, $|R| = n$. If T is a k-evaluation for all the formulas in $\mathcal{P}$ and $k \leq n/f$, then for any line A in $\mathcal{P}$,*

$$\forall \pi (\pi \in Br(T(A)) \Rightarrow Disj(T(A)) \upharpoonright \pi \equiv 1),$$

that is, $T(A)$ has all of its leaves labeled 1.

Proof: The lemma is proved by induction on the number of lines in the proof $\mathcal{P}$. Let

$$\frac{A_1(B_1/p_1, \ldots, B_m/p_m), \ldots, A_k(B_1/p_1, \ldots, B_m/p_m)}{A_0(B_1/p_1, \ldots, B_m/p_m)}$$

be an instance of a rule of $\mathcal{F}$, and assume that the lemma holds for all of the premises of the inference. Let Γ be the set of formulas $A(B_1/p_1, \ldots, B_m/p_m)$, where $A(p_1, \ldots, p_m)$ is a subformula of some A_i. By assumption, $|\Gamma| \leq f$; let $M = \{\pi_1 \cup \cdots \cup \pi_j \in M_n : \pi_i \in Br(T(C_i))\}$, where $\Gamma = \{C_1, \ldots, C_j\}$. By Lemma 4.3, if $\pi_i \in Br(T(C_i))$, then there is a $\pi \in M_n$ so that $\pi_i \subseteq \pi$. Let us abbreviate $Disj(T(A))$ as $D(A)$. Then for $\pi \in M$ and $A, B \in \Gamma$,

1. $D(A) \upharpoonright \pi \equiv 0$ or $D(A) \upharpoonright \pi \equiv 1$;
2. $D(0) \upharpoonright \pi \equiv 0$ and $D(1) \upharpoonright \pi \equiv 1$;
3. if $\neg A \in \Gamma$ then $D(\neg A) \upharpoonright \pi \equiv 1 \Longleftrightarrow D(A) \upharpoonright \pi \equiv 0$;
4. if $(A \vee B) \in \Gamma$ then $D(A \vee B) \upharpoonright \pi \equiv 1 \Longleftrightarrow D(A) \upharpoonright \pi \equiv 1$ or $D(B) \upharpoonright \pi \equiv 1$.

These equivalences follow from the definition of a k-evaluation and from Lemmas 4.3 and 4.4.

For any $\pi \in M$, define an assignment V_π of truth-values to the formulas in Γ by setting $V_\pi(C_i) = 1$ if $D(C_i) \upharpoonright \pi \equiv 1$, $V_\pi(C_i) \equiv 0$ if $D(C_i) \upharpoonright \pi \equiv 0$. The list of equivalences above shows that V_π respects the rules of classical logic. By Lemma 4.4, the premises of the inference are all assigned the value 1 by V_π; since the rule of inference is sound, the conclusion of the inference is also assigned 1 by V_π. Now let $\sigma \in Br(T(A_0(B_1/p_1, \ldots, B_m/p_m)))$. There is a $\pi \in M$ extending σ, so $V_\pi(A_0(B_1/p_1, \ldots, B_m/p_m)) = 1$, equivalently, $D(A_0(B_1/p_1, \ldots, B_m/p_m)) \upharpoonright \sigma \equiv 1$, concluding the proof of the lemma. □

The next lemma shows that, relative to a k-evaluation, $k < n$, the pigeonhole tautology PHP_n is a “contradiction.”

Lemma 5.4 *Let $D \cup R = S$, $|D| = n + 1$, $|R| = n$, $PHP_n = PHP(D, R)$. If T is a k-evaluation for a set of formulas containing PHP_n, $k < n - 1$, then all the leaves of $T(PHP_n)$ are labeled 0.*

Proof: The reduced form of PHP_n is the join of the following sequence of formulas:

1. $(P_{ik} \wedge P_{jk})$, where $i \neq j \in D$, $k \in R$;
2. $(P_{ki} \wedge P_{kj})$, where $i \neq j \in R$, $k \in D$;
3. $\bigwedge_{k \in R} \neg P_{ik}$, for $i \in D$;
4. $\bigwedge_{i \in D} \neg P_{ik}$, for $k \in R$.

By the definition of k-evaluation and Lemma 4.4, it is sufficient to prove for any formula A in the above list that the leaves of $T(A)$ are all labeled with 0.

For a formula of the first kind, this amounts to showing that the leaves of $T(\neg P_{ik} \vee \neg P_{jk})$ are all labeled 1. By Definition 5.1, $T(\neg P_{ik} \vee \neg P_{jk})$ represents $Disj(T(\neg P_{ik}) \vee Disj(T(\neg P_{jk}))$, that is, the matching disjunction containing all terms of the form $(P_{iq} \wedge P_{rk})$, $q \neq k$, $r \neq i$, and all terms of the form $(P_{jq} \wedge P_{rk})$, $q \neq k$, $r \neq j$. Let l be a leaf of $T(\neg P_{ik} \vee \neg P_{jk})$. Since $|\pi(l)| < n - 1$, there is a restriction extending $\pi(l)$ that sets one of these terms to 1. It follows that $\pi(l)$ must set the disjunction to 1, so that l bears the label 1. The proof for formulas of the second kind proceeds similarly.

For formulas of the third kind, we are required to show that the tree $T = T(\bigvee_{k \in R} \neg\neg P_{ik})$ has all its leaves labeled 1. By Definition 5.1, T represents the matching disjunction $\bigvee\{P_{ik} : k \in R\}$. Let l be a leaf of T. Since $|\pi(l)| < n - 1$, there is an extension π of $\pi(l)$ where $i \in V(\pi)$, so that $\bigvee\{P_{ik} : k \in R\}\restriction\pi \equiv 1$, hence $\bigvee\{P_{ik} : k \in R\}\restriction\pi(l) \equiv 1$, showing that l must be labeled 1. For formulas of the fourth kind, a symmetrical argument holds. □

If T is a k-evaluation of a set of formulas Γ in L_n and $\rho \in M_n$, then $T\restriction\rho$ is defined to be the assignment of trees to formulas in $\Gamma\restriction\rho$ given by the definition: $T(0)$ is the tree with a single node labeled 0, while $(T\restriction\rho)(A\restriction\rho) = T(A)\restriction\rho$ if $A\restriction\rho$ is not the constant 0. It follows from this definition that $Disj((T\restriction\rho)(A\restriction\rho)) = Disj(T(A)\restriction\rho)$ for any formula A.

Lemma 5.5 *Let T be a k-evaluation of a set of formulas Γ in L_n. If $\rho \in M_n$, then $T\restriction\rho$ is a k-evaluation of $\Gamma\restriction\rho$.*

Proof: By induction on the complexity of a formula $A \in \Gamma$. If A is a constant or a propositional variable, then the lemma is immediate. If A is a negated formula, then the lemma follows by the third part of Lemma 4.8.

Finally, let A be a disjunction and $\bigvee_{i \in I} A_i$ the merged form of A. By assumption, $T(A)$ represents $\bigvee_{i \in I} Disj(T(A_i))$, so by Lemma 4.8, $T(A)\restriction\rho$ represents

$$\bigvee_{i \in I} Disj(T(A_i))\restriction\rho \equiv \bigvee_{i \in I} Disj(T(A_i)\restriction\rho).$$

Hence, by the remark following the definition of $T\restriction\rho$, the matching decision tree $T(A)\restriction\rho$ represents $\bigvee_{i \in I} Disj((T\restriction\rho)(A_i\restriction\rho))$, completing the proof. □

6 *The switching lemma* In this section we prove the appropriate switching lemma by mapping the "bad" restrictions (those that result in a matching tree of large depth) into a small set. We begin by defining a set of sequences used in defining the mapping.

Define $Code(r, s)$ to be the set of all sequences $\beta_1, \ldots, \beta_k$, where for each i, $\beta_i \in \{\uparrow, *\}^r \setminus \{*\}^r$ and there are exactly s occurrences of $\uparrow$ in the sequence.

Lemma 6.1 $|Code(r, s)| \leq (2r)^s$.

Proof: Given $(\beta_1, \ldots, \beta_k) \in Code(r, s)$, define a map f from $\{1, \ldots, s\}$ to $\{1, \ldots, r\} \times \{0, 1\}$ as follows. $f(1) = (1, 0)$, and for $i > 1$, $f(i) = (j, b)$, where the ith $\uparrow$ in $\beta_1, \ldots, \beta_k$ occurs in the jth place in some entry β_l, and b is 0 or 1 depending on whether the $(i-1)$st $\uparrow$ occurs in β_l or β_{l-1}.

It is easy to see that a sequence $(\beta_1, \ldots, \beta_k) \in Code(r, s)$ is uniquely determined by the map corresponding to it, so that this construction defines an injective mapping from $Code(r, s)$ into the set of all maps $f : \{1, \ldots, s\} \to \{1, \ldots, r\} \times \{0, 1\}$. □

For a given n, $|D| = n + 1$, $|R| = n$, $S = D \cup R$, we define two sets of restrictions. For $l \leq n$, let

$$M_n^l = \{\rho \in M_n : |R \restriction \rho| = l\},$$

and for $s > 0$, F a matching disjunction over S,

$$Bad_n^l(F, s) = \{\rho \in M_n^l : |Tree_{S\restriction\rho}(F \restriction \rho)| \geq s\}.$$

Lemma 6.2 *Let $F = C_1 \vee \cdots \vee C_m$ be an r-disjunction over $D \cup R = S$, where $|D| = n + 1$, $|R| = n$. Then there is a bijection from $Bad_n^l(F, s)$ into*

$$\bigcup_{s/2 \leq j \leq s} M_n^{l-j} \times Code(r, j) \times [2l + 1]^s.$$

Proof: Let $\rho \in Bad_n^l(F, s)$; choose π to be the matching determined by the leftmost path originating in the root of $Tree_{S\restriction\rho}(F \restriction \rho)$ that has length s.

Starting from F and π, we define three sequences by induction that are used to define the bijection G:

1. $D_1, \ldots, D_k$, a subsequence of $C_1, \ldots, C_m$;
2. $\sigma_1, \ldots, \sigma_k$, a sequence of restrictions $\sigma_i \subseteq \delta_i$, where $D_i = \wedge\, \delta_i$, and $\rho\sigma_1, \ldots, \sigma_i \in M_n$;
3. $\pi_1, \ldots, \pi_k$, a partition of π, where each π_i, $i < k$, satisfies the conditions (a) $\pi_i \in M(V(\sigma_i))$ and (b) the restriction $\rho\pi_1, \ldots, \pi_i$ labels a path in $Tree_{S\restriction\rho}$ $(F \restriction \rho)$ ending in a boundary node.

Suppose that the sequences have been defined as far as $\pi_{i-1}, D_{i-1}, \sigma_{i-1}$, that $\pi_1, \ldots, \pi_{i-1}$ and $\sigma_1, \ldots, \sigma_{i-1}$ satisfy the stated conditions and that $\pi_1, \ldots, \pi_{i-1} \neq \pi$. Since $\rho\pi_1, \ldots, \pi_{i-1}$ labels a path ending in a boundary node, it follows that there must be a term D in F so that $D \restriction \rho\pi_1, \ldots, \pi_{i-1} \not\equiv 1$, $D \restriction \rho\pi_1, \ldots, \pi_{i-1} \not\equiv 0$, for otherwise the path labeled by π would end at that node. Define D_i to be the first such term in F.

Define σ_i' to be the unique minimal matching so that

$$D_i \restriction \rho\pi_1, \ldots, \pi_{i-1}\sigma_i' \equiv 1,$$

and let π_i be the set of pairs in π that covers vertices in $V(\sigma_i')$. In defining σ_i two cases arise.

Case 1: $\pi_1, \ldots, \pi_i \neq \pi$. Define σ_i to be σ_i'. In this case, both σ_i and π_i label paths in the full matching tree for $V(\sigma_i)$ over $S \restriction \sigma\rho_1, \ldots, \pi_{i-1}$, showing that $\pi_i \in M(V(\sigma_i))$ and that $\pi_1, \ldots, \pi_i$ labels a path in $Tree_{S\restriction\rho}(F \restriction \rho)$ ending in a boundary node. Since the boundary node is not a leaf of the tree, it follows that $\pi_i \neq \sigma_i$, so

$$D_i \restriction \rho\pi_1, \ldots, \pi_i \equiv 0.$$

Case 2: $\pi_1, \ldots, \pi_i = \pi$, so that $\pi_i = \pi_k$. Let $p_1, \ldots, p_t$ be the pairs constituting π_k, listed in the order they appear on the path. Each p_j contains a vertex v_j that is the first vertex in $V(\sigma_i')$ not in $p_1 \cup \cdots \cup p_{j-1}$. Define q_j to be the pair in $V(\sigma_i')$ containing v_j, and let $\sigma_i = \{q_1, \ldots, q_t\}$. In this second case, it is not guaranteed that

$$D_k \restriction \rho\pi_1, \ldots, \pi_{k-1}\sigma_k \equiv 1,$$

but only

$$D_k \restriction \rho\pi_1, \ldots, \pi_{k-1}\sigma_k \not\equiv 0.$$

We need to verify that $\rho\sigma_1, \ldots, \sigma_i \in M_n$. By assumption, $\rho\sigma_1, \ldots, \sigma_{i-1} \in M_n$. Let us assume in addition that $\rho\sigma_1, \ldots, \sigma_i \notin M_n$, so that there are $a, b, c \in S$, $b \neq c$ where $\{a, b\} \in \rho\sigma_1, \ldots, \sigma_{i-1}$, $\{a, c\} \in \sigma_i$. Since $D_i \restriction \rho \not\equiv 0$, $\{a, b\} \notin \rho$, so that $\{a, b\} \in \sigma_j$ for some $j < i$. Since by assumption $\pi_j \in M(V(\sigma_j))$, and $D_i \restriction \rho\pi_1, \ldots, \pi_{i-1} \not\equiv 0$, it follows that $\{a, b\} \in \pi_j$. This contradicts the assumption that $a \in V(\sigma_i)$, showing that $\rho\sigma_1, \ldots, \sigma_i \in M_n$.

We note here a fact used later in proving that G is a bijection: for any $i \leq k$, the set of pairs $\rho\pi_1, \ldots, \pi_{i-1}\sigma_i, \ldots, \sigma_k$ is in M_n. If the set is not in M_n, then because $\rho\pi_1, \ldots, \pi_k, \rho\sigma_1, \ldots, \sigma_k \in M_n$, there must be p, q where $1 \leq p \leq i - 1 < q \leq k$ so that for some $a, b, c, b \neq c$, $\{a, b\} \in \pi_p$ and $\{a, c\} \in \sigma_q$. However, if $\{a, b\} \in \pi_p$ then $a \notin S \restriction \rho\pi_1, \ldots, \pi_{i-1}$, contradicting $\{a, c\} \in \sigma_q$.

Before defining the map G it is convenient to introduce a special ordering of the $2l + 1$ vertices unset by the restriction ρ. The new ordering is determined by the original ordering of the vertices and the sequence of restrictions $\sigma_1, \ldots, \sigma_k$. Let $\sigma = \sigma_1, \ldots, \sigma_k$. To avoid confusion between the original ordering and the new ordering, we shall refer to the original order as ordering by *size*, and the new order as ordering by *index* and we shall refer to the position of an element in the new ordering as its *index*. The index ordering is defined as follows: order the $2l + 1$ vertices unset by ρ so that the vertices set by σ are listed first according to the order $V(\sigma_1) < \cdots < V(\sigma_k)$, then by size within each set $V(\sigma_i)$; next, the remaining vertices unset by $\rho\sigma$ are listed by size, in the index positions $2j + 1, \ldots, 2l + 1$, where $j = |\sigma|$.

The map $G(\rho) = \langle G_1(\rho), G_2(\rho), G_3(\rho)\rangle$ is now defined as follows.

1. $G_1(\rho) = \rho\sigma$.

2. For $i = 1, \ldots, k$, let β_i be the vector of length r so that:

$$\beta_i(j) = \begin{cases} \uparrow & \text{if } \sigma_i \text{ sets the } j\text{th variable in } D_i \\ * & \text{otherwise.} \end{cases}$$

Then $G_2(\rho)$ is defined as the sequence $\langle \beta_1, \ldots, \beta_k \rangle$.

3. $G_3(\rho) \in [2l + 1]^s$ is defined as follows.
   (a) List the elements of π according to the index ordering, where for each pair in π the element with the lower index determines the position of the pair.
   (b) From the ordered list of the pairs in π, create a new list by recording for each pair the index of the element in the pair with the higher index. This new list is $G_3(\rho)$.

Let $j = |\rho|$. We need to show that $G(\rho) \in M_n^{l-j} \times Code(r, j) \times [2l + 1]^s$, where $s/2 \leq j \leq s$. The fact that $\rho\sigma \in M_n^{l-j}$ was proved above. The definition of σ_i ensures that $G_2(\rho) \in Code(r, j)$, and $G_3(\rho) \in [2l + 1]^s$ by definition. For $i < k$, $\pi_i \in M(V(\sigma_i))$, so that $|\sigma_i| \leq |\pi_i| \leq 2|\sigma_i|$, while for $i = k$, $|\sigma_i| = |\pi_i|$ holds by construction. Thus $|\pi|/2 \leq |\sigma| \leq |\pi|$, that is, $s/2 \leq j \leq s$.

It remains to show that G is a bijection. We prove this by showing how to reconstruct the restriction ρ from $G(\rho)$ by successively recovering the elements of the three sequences used in defining $G(\rho)$.

At the beginning of the reconstruction process, we are given only the triple $G(\rho)$ and the r-disjunction F. From $G_1(\rho)$ we can find the set of vertices unset by $\rho\sigma$, and hence the indices of these vertices.

Let us suppose that the reconstruction process has been carried out as far as stage $i - 1$; at this stage we have found the terms $D_1, \ldots, D_{i-1}$, the restrictions $\sigma_1, \ldots, \sigma_{i-1}, \pi_1, \ldots, \pi_{i-1}$ and $\rho\pi_1, \ldots, \pi_{i-1}\sigma_i, \ldots, \sigma_k$. In addition, we have found the indices of all the vertices in $V(\sigma_1) \cup \cdots \cup V(\sigma_{i-1})$.

We now describe stage i of the reconstruction process. If C_j is a term in F that occurs earlier in F than D_i, then

$$C_j \restriction \rho\pi_1, \ldots, \pi_{i-1} \equiv 0,$$

hence

$$C_j \restriction \rho\pi_1, \ldots, \pi_{i-1}\sigma_i, \ldots, \sigma_k \equiv 0.$$

On the other hand, if $i < k$ then

$$D_i \restriction \rho\pi_1, \ldots, \pi_{i-1}\sigma_i \equiv 1,$$

while

$$D_i \restriction \rho\pi_1, \ldots, \pi_{k-1}\sigma_k \not\equiv 0.$$

Thus in either case,

$$D_i \restriction \rho\pi_1, \ldots, \pi_{i-1}\sigma_i, \ldots, \sigma_k \not\equiv 0,$$

so that D_i can be found as the first term in F not set to 0 by the restriction $\rho\pi_1, \ldots, \pi_{i-1}\ \sigma_i, \ldots, \sigma_k$. Having found D_i, we can consult the entry β_i in the sequence

$G_2(\rho) = \langle \beta_1, \ldots, \beta_k \rangle$ to find the variables in D_i that are set by σ_i, and hence find σ_i itself. We can now find the indices of the vertices in $V(\sigma_i)$ by extending the list of indices already compiled for $\sigma_1, \ldots, \sigma_{i-1}$. It remains to reconstruct π_i using $G_3(\rho)$. Every pair in π_i must contain at least one vertex in $V(\sigma_i)$, hence for every such pair we can find the vertex in the pair with lower index. The other vertex in the pair (with the higher index) must either be in $V(\sigma_i)$ or in the set of vertices unset by $\rho\sigma$. In either case, we can find the other vertex in the pair by consulting appropriate entries in $G_3(\rho)$—these entries follow immediately after the entries corresponding to pairs in $\pi_1, \ldots, \pi_{i-1}$. Thus we can reconstruct π_i. Lastly, by replacing σ_i by π_i, we can find the restriction $\rho\pi_1, \ldots, \pi_{i-1}\pi_i\sigma_{i+1}, \ldots, \sigma_k$.

Finally, having found all of $\sigma_1, \ldots, \sigma_k$, we can find ρ by removing all of the pairs in $\sigma_1, \ldots, \sigma_k$ from $\rho\sigma_1, \ldots, \sigma_k$. This completes the proof that G is a bijection. □

Example 6.3 To illustrate the definitions in the above proof, we continue Example 4.6. Given D, R, S, F, ρ as in that example, and setting $s = 3$, we have: $\pi = \{\{2, 8\}, \{3, 7\}, \{5, 9\}\}$, $D_1 = (P_{16} \wedge P_{27})$, $\sigma_1 = \{\{2, 7\}\}$, $\pi_1 = \{\{2, 8\}, \{3, 7\}\}$, $D_2 = (P_{16} \wedge P_{59})$, $\sigma_2 = \{\{5, 9\}\}$, $\pi_2 = \{\{5, 9\}\}$. Hence, $G_1(\rho) = \{\{1, 6\}, \{2, 7\}, \{5, 9\}\}$, $G_2(\rho) = \langle \uparrow *, \uparrow * \rangle$, $G_3(\rho) = \langle g, e, d \rangle$, where the index ordering is $\langle 2, 7, 5, 9, 3, 4, 8 \rangle$, with the corresponding indices a, b, c, d, e, f, g.

In the next lemma we use the notation $a^{\underline{m}}$ for the falling factorial power $a(a - 1), \ldots, (a - m + 1)$.

Lemma 6.4 (The switching lemma) *Let F be an r-disjunction over $D \cup R$, $|D| = n + 1$, $|R| = n$. Let $l \geq 10$, and set $p = l/n$. If $r \leq l$ and $p^4 n^3 \leq 1/10$ then*

$$\frac{|Bad_n^l(F, 2s)|}{|M_n^l|} \leq (11 p^4 n^3 r)^s.$$

Proof: By Lemma 6.2, it is sufficient to bound the ratio

$$\frac{\left| \bigcup_{s \leq j \leq 2s} M_n^{l-j} \times Code(r, j) \times [2l + 1]^{2s} \right|}{|M_n^l|}. \tag{1}$$

We begin by estimating the ratio $|M_n^{l-j}|/|M_n^l|$. A restriction in M_n^l is determined by the following process: pick l elements in R, then for each of the $n - l$ remaining vertices in R in turn choose the element of D with which it is matched. Thus

$$\begin{aligned} |M_n^l| &= \binom{n}{l} (n+1)^{\underline{n-l}} \\ &= \frac{n^{\underline{l}}\, (n+1)^{\underline{n-l}}}{l!}. \end{aligned} \tag{2}$$

Using (2) and the recursion equation $a^{\underline{m+n}} = a^{\underline{m}}\, (a - m)^{\underline{n}}$, we estimate

$$\begin{aligned} \frac{|M_n^{l-j}|}{|M_n^l|} &= \frac{n^{\underline{l-j}}\, (n+1)^{\underline{n-l+j}}\, l!}{(l-j)!\, n^{\underline{l}}\, (n+1)^{\underline{n-l}}} \\ &= \frac{(l+1)^{\underline{j}}\, l!}{(l-j)!\, (n-l+j)^{\underline{j}}} \end{aligned}$$

$$\begin{aligned} &= \frac{(l+1)^{\underline{j}}\, l^{\underline{j}}}{(n-l+j)^{\underline{j}}} \\ &\leq \left(\frac{l(l+1)}{(n-l)}\right)^{j}. \end{aligned} \tag{3}$$

Hence, using (3) and the estimate of $|Code(r, j)|$ from Lemma 6.1 we can bound the ratio (1) from above by the sum

$$\begin{aligned} &\sum_{s\leq j\leq 2s} \left(\frac{l(l+1)}{n-l}\right)^{j} (2r)^{j}(2l+1)^{2s} \\ &= (2l+1)^{2s} \sum_{s\leq j\leq 2s} \left(\frac{2l(l+1)r}{(n-l)}\right)^{j}. \end{aligned} \tag{4}$$

To bound this last sum, we begin by estimating

$$\begin{aligned} \frac{2l(l+1)r}{(n-l)} &\leq \frac{2l^2(l+1)}{(n-l)} \\ &= \frac{l^3}{n} \cdot \frac{2(1+1/l)}{(1-p)} \\ &\leq \frac{p^4n^3}{l} \cdot \frac{2.2}{0.9999} \\ &< 0.0221, \end{aligned} \tag{5}$$

using the inequalities $r \leq l$, $p^4n^3 \leq 1/10$ and $l \geq 10$. Hence the sum in (4) is bounded by the sum of a geometric series with ratio < 0.0221, so that it is less than 1.03 times its largest term. This provides us with the estimate

$$\begin{aligned} \frac{|Bad_n^l(F, 2s)|}{|M_n^l|} &\leq 1.03(2l+1)^{2s} \left(\frac{2l(l+1)r}{(n-l)}\right)^{s} \\ &= 1.03\left(\frac{2(2l+1)^2 l(l+1)r}{(n-l)}\right)^{s}. \end{aligned} \tag{6}$$

To put this inequality in more usable form, we bound the ratio in the *RHS*.

$$\begin{aligned} \frac{2(2l+1)^2 l(l+1)r}{(n-l)} &\leq \frac{8(l+1)^3 lr}{(n-l)} \\ &\leq \frac{l^4 r}{n} \cdot \frac{8(1+1/l)^3}{(1-p)} \\ &\leq \frac{10.65\, l^4 r}{n}. \end{aligned} \tag{7}$$

This last inequality together with (6) yields the bound

$$\begin{aligned} \frac{|Bad_n^l(F, 2s)|}{|M_n^l|} &\leq 1.03(10.65 p^4 n^3 r)^s \\ &< (11 p^4 n^3 r)^s, \end{aligned} \tag{8}$$

completing the proof of the lemma. □

7 ***Lower bounds for pigeonhole formulas*** In this section, we prove a lemma showing that if a set of bounded depth formulas of L_n is subjected to a random restriction, then, provided the set is not too large, the set of restricted formulas has associated decision trees of small depth. From this result the lower bound on the size of propositional proofs follows by earlier lemmas.

Lemma 7.1 *Let d be an integer, $0 < \epsilon < 1/5$, $0 < \delta < \epsilon^d$ and Γ a set of formulas of L_n of depth $\leq d$, closed under subformulas. If $|\Gamma| < 2^{n^\delta}$, $q = \lceil n^{\epsilon^d} \rceil$ and n is sufficiently large, then there exists $\rho \in M_n^q$ so that there is a $2n^\delta$-evaluation of $\Gamma \upharpoonright \rho$.*

Proof: The proof is by induction on d. For $d = 0$, the only formulas in Γ are propositional variables and constants. For any such formula A, $|Tree_S(A)| \leq 2$, so that we can set $\rho = \varnothing$.

Assume that the Lemma holds for d. Let Γ be a set of formulas of depth $d + 1$, closed under subformulas, $|\Gamma| < 2^{n^\delta}$, where $0 < \delta < \epsilon^{d+1}$. Let Δ be the set of formulas in Γ of depth $\leq d$. Since $0 < \delta < \epsilon^{d+1} < \epsilon^d$, by the induction hypothesis, there is $\rho \in M_n^q$, $q = \lceil n^{\epsilon^d} \rceil$, for which there is a $2n^\delta$-evaluation T of $\Delta \upharpoonright \rho$.

Let A be a disjunction in Γ of depth $d + 1$, and $\bigvee_{i \in I} A_i$ its merged form. Let q' be $\lceil n^{\epsilon^{d+1}} \rceil$. In Lemma 6.4, set $D \longrightarrow D \upharpoonright \rho$, $R \longrightarrow R \upharpoonright \rho$, $n \longrightarrow \lceil n^{\epsilon^d} \rceil = q$, $l \longrightarrow \lceil n^{\epsilon^{d+1}} \rceil = q'$, $r \longrightarrow \lfloor 2n^\delta \rfloor$, $s \longrightarrow n^\delta$. For n sufficiently large, $p^4 \lceil n^{\epsilon^d} \rceil^3 < n^{-\epsilon^d/5} \leq 1/10$, where $p = l/n$, and $\lfloor 2n^\delta \rfloor \leq \lceil n^{\epsilon^{d+1}} \rceil$ since $\delta < \epsilon^{d+1}$. Thus the conditions for Lemma 6.4 hold, so that the ratio

$$\frac{|Bad_q^{q'}(\bigvee_{i \in I} Disj(T(A_i \upharpoonright \rho)), 2n^\delta)|}{|M_q^{q'}|}$$

is bounded by $(11n^{-\epsilon^d/5} \lfloor 2n^\delta \rfloor)^{n^\delta}$. Since $\delta < \epsilon^{d+1} < \epsilon^d/5$, for n sufficiently large, $11n^{-\epsilon^d/5} \lfloor 2n^\delta \rfloor < 1/2$, so that the above ratio is bounded by 2^{-n^δ}. It follows that there is a restriction $\rho' \in M_q^{q'}$ so that for every disjunction $A \in \Gamma$ of depth $d + 1$,

$$|Tree_{S \upharpoonright \rho\rho'}(\bigvee_{i \in I} Disj(T(A_i \upharpoonright \rho)) \upharpoonright \rho')| < 2n^\delta.$$

Set $\rho'' = \rho\rho'$; by construction, $\rho'' \in M_n^{q'}$. We wish to show that there is a $2n^\delta$-evaluation T'' of $\Gamma \upharpoonright \rho''$. By Lemma 5.5, $T' = T \upharpoonright \rho'$ is a $2n^\delta$-evaluation of $\Delta \upharpoonright \rho''$; we define T'' by extending T' to formulas of depth $d + 1$. For A a negated formula of depth $d + 1$, set $T''(A \upharpoonright \rho'') = (T'(A \upharpoonright \rho''))^c$. If A is a disjunction of depth $d + 1$, and $\bigvee_{i \in I} A_i$ its merged form of A, then set

$$T''(A \upharpoonright \rho'') = Tree_{S \upharpoonright \rho''}(\bigvee_{i \in I} Disj(T(A_i \upharpoonright \rho)) \upharpoonright \rho'),$$

where $\bigvee_{i \in I} A_i$ is the merged form of A. By definition, $T''(A \upharpoonright \rho'')$ represents $\bigvee_{i \in I} Disj(T(A_i \upharpoonright \rho)) \upharpoonright \rho'$. By Lemma 4.8 and the definition of T'',

$$\bigvee_{i \in I} Disj(T(A_i \upharpoonright \rho)) \upharpoonright \rho' \equiv \bigvee_{i \in I} Disj(T''(A_i)),$$

so that $T''(A)$ represents $\bigvee_{i \in I} Disj(T''(A_i))$. This completes the proof that T'' is a $2n^\delta$-evaluation of $\Gamma \upharpoonright \rho''$. □

Theorem 7.2 *Let $\mathcal{F}$ be a Frege system and $d > 4$. Then for sufficiently large n every depth d proof in $\mathcal{F}$ of PHP_n must have size at least 2^{n^δ}, for $\delta < (1/5)^d$.*

Proof: Let the rules of $\mathcal{F}$ have size bounded by f, $0 < \delta < (1/5)^d$, and let $A_1, \ldots, A_t$ be a proof in $\mathcal{F}$ of depth d and size $\leq 2^{n^\delta}$.

Choose ϵ so that $\epsilon < 1/5$, $\delta < \epsilon^d$. By Lemma 7.1, there exists $\rho \in M_n^q$, $q = \lceil n^{\epsilon^d} \rceil$, and a $2n^\delta$-evaluation T of $\Gamma\restriction\rho$, where Γ is the set of subformulas in the proof $A_1, \ldots, A_t$. Then $A_1\restriction\rho, \ldots, A_t\restriction\rho$ is a proof in $\mathcal{F}$ in the language $L(D\restriction\rho, R\restriction\rho)$.

Since $\delta < \epsilon^d$ and n is sufficiently large, $2n^\delta \leq n^{\epsilon^d}/f$, so by Lemma 5.3, for every step A_k in the proof, $T(A_k\restriction\rho)$ has all its leaves labeled 1. On the other hand, $PHP_n\restriction\rho \equiv PHP(D\restriction\rho, R\restriction\rho)$, so by Lemma 5.4, if PHP_n were the last line A_t of the proof, all the leaves of $T(PHP_n\restriction\rho)$ would be labeled 0. It follows that $A_1, \ldots, A_t$ cannot be a proof of PHP_n. Hence, any proof in $\mathcal{F}$ of PHP_n must have size at least 2^{n^δ}. □

The lower bound originally proved by Ajtai [1] is superpolynomial rather than exponential. The essential difference between Ajtai's original proof and the proof given here is as follows. The original proof has essentially the same structure as the present proof, but makes use of a more restricted class of complete decision trees. The class of trees appropriate to the original proof is the class of all full matching trees over a small set (with an appropriate sense of "small"). With this more restricted class of matching trees, it is not possible to prove exponential lower bounds. (For a proof, see the concluding section of [4].)

8 *Lower bounds for graph formulas* The techniques used in the proof of the main theorem above can be used for several other classes of formulas, for example, for classes of formulas based on other matching principles. In the present section, we illustrate this by sketching a proof of a lower bound for the case of tautologies based on a graphical construction; this class of tautologies was first defined by Tseitin [15].

Let G be a finite undirected graph, in which the vertices are labeled with 0 or 1, and the edges with distinct literals. Then a set of clauses $Clauses(G)$ associated with G is defined as follows. For each vertex $v \in G$, let $Clauses(v)$ be the set of clauses constituting the conjunctive normal form of the modulo 2 equation $p_1 \oplus \cdots \oplus p_k = c$, where $p_1, \ldots, p_k$ are the literals labeling the edges attached to v, and c is the label on v. Then $Clauses(G)$ is the union of all the clause sets $Clauses(v)$ for v a vertex in G. If the sum of the vertex labels of G is odd, then $Clauses(G)$ is inconsistent. (This follows from the fact that if we add the left-hand sides of all the modulo 2 equations associated with the vertices of G, the sum is zero, because each literal appears twice in the sum.)

The size of $Clauses(v)$ is exponential in the degree of v, so that if the graph G is of large degree, the size of $Clauses(G)$ can be exponential in the size of G. In the present section, we shall use sets of clauses associated with complete graphs. Cook [6] proposed a way to reduce the size of the associated sets of clauses by introducing extra variables. Let K_n be the complete graph on $n = 2m + 1$ vertices. Let the vertex set X of K_n be $\{0, 1, \ldots, n - 1\}$ and each edge $\{i, j\}$ be labeled with a variable P_{ij}. We introduce a set of extra variables $\{Q_0^i, Q_1^i, \ldots, Q_{n-3}^i\}$ for each vertex $i \in X$ as follows: we let $Q_0^i \equiv P_{i,i+1} \oplus P_{i,i+2}$ and $Q_j^i \equiv Q_{j-1}^i \oplus P_{i,i+j+2}$ $(1 \leq j \leq n - 3)$

where $+$ is modulo n addition. We define $Cl(i)$ to be the set of clauses comprising the conjunctive normal form of the set of above equations together with $Q^i_{n-3} \equiv 1$ which expresses the fact that the label on vertex i is 1. For any assignment of truth values to $P_{i,i+1}, \ldots, P_{i,i+n-1}$ satisfying $Cl(i)$, Q^i_j has the same value as $P_{i,i+1} \oplus \cdots \oplus P_{i,i+j+2}$. Let Cl_n be the union of all the sets $Cl(i)$ for $i \in X$ a vertex in K_n. Since n is odd, then Cl_n is contradictory. Finally, let $Graph_n$ be the tautology that results by negating all the clauses in Cl_n and then forming their disjunction; we will use $Graph_n$ in the form

$$\begin{gathered}
\bigvee_{i\in X}(\neg(Q^i_0 \vee P_{i,i+1} \vee \neg P_{i,i+2}) \vee \neg(Q^i_0 \vee \neg P_{i,i+1} \vee P_{i,i+2})\vee \\
\neg(\neg Q^i_0 \vee P_{i,i+1} \vee P_{i,i+2}) \vee \neg(\neg Q^i_0 \vee \neg P_{i,i+1} \vee \neg P_{i,i+2}))\vee \\
\bigvee_{i\in X} \bigvee_{j=1}^{n-3} \neg(Q^i_j \vee Q^i_{j-1} \vee \neg P_{i,i+j+2}) \vee \neg(Q^i_j \vee \neg Q^i_{j-1} \vee P_{i,i+j+2})\vee \\
\neg(\neg Q^i_j \vee Q^i_{j-1} \vee P_{i,i+j+2}) \vee \neg(\neg Q^i_j \vee \neg Q^i_{j-1} \vee \neg P_{i,i+j+2})) \vee \bigvee_{i\in X} \neg Q^i_{n-3}
\end{gathered}$$

where $\bigvee$ denotes repeated binary $\vee$. $Graph_n$ has size $O(n^2)$ and depth 4.

The restrictions used in the case of the graph clauses are determined by matchings, just as in the case of the pigeonhole formulas. The definitions of §3 above can be taken over with essentially no change; the only alteration required is that the definitions are to be taken as referring to the graph K_n rather than the complete bipartite graph $K(n+1,n)$. In particular, the concepts of matching terms and matching disjunction are defined just as in §3. The basic definitions and lemmas on matching decision trees in §4 can also be used here without alteration, except that the space of mappings is based on K_n rather than $K(n+1,n)$.

The definition of k-evaluation in §5 can be used as given, with added evaluations for the extension variables Q^i_j. We let $T(Q^i_j)$ be the full matching tree for $\{i\}$, with a leaf l labeled 1 if $\pi(l)$ is $\{i, i+\ell+2\}$ for $-1 \leq \ell \leq j$, otherwise l is labeled 0. The proof of Lemma 5.3 goes through exactly as before. The Lemma corresponding to Lemma 5.4 can be stated as follows.

Lemma 8.1 *Let $n = 2m+1$. If T is a k-evaluation for a set of formulas closed under subformulas and containing $Graph_n$, and $k < m$, then all the leaves of $T(Graph_n)$ are labeled 0.*

Proof: The merged form of $Graph_n$ is a disjunction of the negations of the formulas of following forms.

1. $Q^i_0 \vee P_{i,i+1} \vee \neg P_{i,i+2}$, for $i \in X$.
2. $Q^i_0 \vee \neg P_{i,i+1} \vee P_{i,i+2}$, for $i \in X$.
3. $\neg Q^i_0 \vee P_{i,i+1} \vee P_{i,i+2}$, for $i \in X$.
4. $\neg Q^i_0 \vee \neg P_{i,i+1} \vee \neg P_{i,i+2}$, for $i \in X$.
5. $\neg Q^i_j \vee \neg Q^i_{j-1} \vee \neg P_{i,i+j+2}$, where $1 \leq j \leq n-3$, $i \in X$.
6. $Q^i_j \vee Q^i_{j-1} \vee \neg P_{i,i+j+2}$, where $1 \leq j \leq n-3$, $i \in X$.
7. $Q^i_j \vee \neg Q^i_{j-1} \vee P_{i,i+j+2}$, where $1 \leq j \leq n-3$, $i \in X$.
8. $\neg Q^i_j \vee Q^i_{j-1} \vee P_{i,i+j+2}$, where $1 \leq j \leq n-3$, $i \in X$.
9. Q^i_{n-3}, for $i \in X$.

By the definition of k-evaluation and Lemma 4.4, it is sufficient to prove for any formula A in the above list that the leaves of $T(A)$ are all labeled with 1.

For a formula of the first kind, by Definition 5.1 and the evaluations for the extension variables, $T(Q^i_0 \vee P_{i,i+1} \vee \neg P_{i,i+2})$ represents $Disj(T(Q^i_0)) \vee Disj(T(P_{i,i+1}))$ $\vee Disj(T(\neg P_{i,i+2}))$. By definition, $Disj(T(Q^i_0)) = P_{i,i+1} \vee P_{i,i+2}$, $Disj(T(P_{i,i+1})) = P_{i,i+1}$ and $Disj(T(P_{i,i+2})) = P_{i,i+2}$. Hence

$$Disj(T(\neg P_{i,i+2})) = \bigvee\{(P_{ik} \wedge P_{l,i+2}) : k \neq l; k, l \notin \{i, i+2\}\}.$$

Thus $T(Q^i_0 \vee P_{i,i+1} \vee \neg P_{i,i+2})$ represents the matching disjunction

$$P_{i,i+1} \vee P_{i,i+2} \vee P_{i,i+1} \vee \bigvee\{(P_{ik} \wedge P_{l,i+2}) : k \neq l; k, l \notin \{i, i+2\}\}.$$

If l is a leaf of $T(Q^i_0 \vee P_{i,i+1} \vee \neg P_{i,i+2})$, then $|\pi(l)| < m$, so there is a restriction extending $\pi(l)$ that sets one of these terms to 1. It follows that $\pi(l)$ must set the disjunction to 1, so that l bears the label 1. The proof for formulas of the kinds from the second to the fourth proceeds similarly.

For a formula of the fifth kind, $T(\neg Q^i_j \vee \neg Q^i_{j-1} \vee \neg P_{i,i+j+2})$ represents $Disj(T(\neg Q^i_j)) \vee Disj(T(\neg Q^i_{j-1})) \vee Disj(T(\neg P_{i,i+j+2}))$. Since $Disj(T(Q^i_j)) = P_{i,i+1} \vee \cdots \vee P_{i,i+j+2}$, $Disj(T(\neg Q^i_j)) = P_{i,i+j+3} \vee \cdots \vee P_{i,i+n-1}$. For similar reasons, $Disj(T(\neg Q^i_{j-1})) = P_{i,i+j+2} \vee \cdots \vee P_{i,i+n-1}$. Hence $T(\neg Q^i_j \vee \neg Q^i_{j-1} \vee \neg P_{i,i+j+2})$ represents the matching disjunction

$$P_{i,i+j+2} \vee \cdots \vee P_{i,i+n-1} \vee \bigvee\{(P_{ik} \wedge P_{l,i+j+2}) : k \neq l; k, l \notin \{i, i+j+2\}\}.$$

Since $|\pi(l)| < m$ for l a leaf of $T(\neg Q^i_j \vee \neg Q^i_{j-1} \vee \neg P_{i,i+j+2})$, there is a restriction extending $\pi(l)$ that sets one of these terms to 1. It follows that $\pi(l)$ must set the disjunction to 1, so that l bears the label 1. The cases from the sixth kind to the eighth kind follow by exactly similar arguments.

For a formula of the ninth kind, it is true by the definition of the evaluation. □

Let M_n be the set of all partial 1-to-1 maps from X to X where X is the vertex set of the complete graph K_n, for $n = 2m + 1$. Let $V(h) = dom(h) \cup rng(h)$ for $h \in M_n$. For $l \leq m$ define

$$M^l_n = \{\rho \in M_n : V(\rho) = 2(m - l)\}.$$

The lemmas corresponding to the first two lemmas of §6 above carry through to the case of the graph formulas, with the single change that the lemma corresponding to Lemma 6.2 must be rephrased to refer to the new space of matchings. We can then state the switching lemma as follows.

Lemma 8.2 *Let $F = C_1 \vee \cdots \vee C_H$ be an r-disjunction over X. Then there is a 1-1 map from $Bad^l_n(F, s)$ into*

$$\bigcup_{s/2 \leq j \leq s} M^{l-j}_n \times Code(r, j) \times [2l + 1]^s.$$

Proof: The proof is similar to the proof of Lemma 6.2. We sketch the proof here. Let $\rho \in Bad_n^l(F, s)$ and let π be the map determined by the leftmost path originating in the root of $Tree_{S\restriction\rho}(F\restriction\rho)$ that has length s. We will construct the image of ρ by defining a partition $\pi_1, \pi_2, \dots, \pi_k$ of π.

Suppose that $\pi_1, \pi_2, \dots, \pi_{i-1} \subseteq \pi$ have already been defined and $\pi_1, \pi_2, \dots, \pi_{i-1} \neq \pi$. Then by the definition of $Tree_{S\restriction\rho}(F\restriction\rho)$ there is a term $C \in \{C_1, C_2, \dots, C_H\}$ such that $C\restriction\pi_1\pi_2, \dots, \pi_{i-1}\rho \not\equiv 0$. Then we let C_{ν_i} be the first such term. Let $K_i = V(C_{\nu_i}\restriction\pi_1\pi_2, \dots, \pi_{i-1}\rho)$ and let σ_i be the unique map that satisfies $C_{\nu_i}\restriction\pi_1\pi_2, \dots, \pi_{i-1}\rho$. Let π_i be the portion of π that touches K_i. Then two cases arise.

Case 1: If $\pi_1\pi_2, \dots, \pi_i \neq \pi$ then by the construction of $Tree_{S\restriction\rho}(F\restriction\rho)$, $\pi_1\pi_2, \dots, \pi_i$ touches all the vertices touched by σ_i. Thus $C_{\nu_i}\restriction\pi_1\pi_2, \dots, \pi_i\rho \equiv 0$.

Case 2: If $\pi_1\pi_2, \dots, \pi_i = \pi$, $i = k$. Then let $p_1, \dots, p_t$ be the pairs constituting π_k, listed in the order they appear on the path. Each p_j contains a vertex v_j that is the first vertex in $V(\sigma_i')$ not in $p_1 \cup \dots \cup p_{j-1}$. Define q_j to be the pair in $V(\sigma_i')$ containing v_j, and let $\sigma_i = \{q_1, \dots, q_t\}$.

For each σ_i we define a corresponding string β_i based on the fixed ordering of the variables in term C_{ν_i} by letting the jth component of β_i be $\uparrow$ if and only if the jth variable in C_{ν_i} is set by σ_i. Since C_{ν_i} is not empty, there is at least one $\uparrow$ in β_i. Thus $(\beta_1, \dots, \beta_k) \in Code(r, j)$. Clearly $\rho\sigma_1, \dots, \sigma_k \in M_n^{\ell-j}$. We let the image of ρ be $\langle\rho\sigma_1, \dots, \sigma_k, (\beta_1, \dots, \beta_k), \delta\rangle$ where $\delta \in [2\ell+1]^s$ encoding the relationship between σ_i and π_i. We number the $2j$ vertices in $V(\sigma)$ with $1, \dots, 2j$ in the order $V(\sigma_1) < V(\sigma_2) < \dots < V(\sigma_k)$ and the vertices unset by $\rho\sigma$ with $2j+1, \dots, 2\ell+1$. Then we list the pairs in π_i in the order of their smallest numbered elements in $V(\sigma_i)$. Thus we use the vector δ to store the numbers of the other vertices of the pairs in π.

Now we show the map is 1-1 by recovering ρ from its image. We do this by induction on i. Assume that we have already recovered $\pi_1, \pi_2, \dots, \pi_{i-1}, \sigma_1, \sigma_2, \dots, \sigma_{i-1}$. Then we know $\rho\pi_1, \dots, \pi_{i-1}\sigma_i, \dots, \sigma_k$. We can recover ν_i as the index of the first term of F that is not set to 0, since for $i < k$, $C_{\nu_i}\restriction\rho\pi_1\pi_2, \dots, \pi_{i-1}\sigma_i \equiv 1$, for $i = k$, $C_{\nu_i}\restriction\rho\pi_1\pi_2, \dots, \pi_{k-1}\sigma_k \not\equiv 0$ and $C_j\restriction\rho\pi_1\pi_2, \dots, \pi_{i-1}\sigma_i \equiv 0$ for all $j < \nu_i$. This is also true when $\sigma_{i+1}, \dots, \sigma_k$ is appended to the restriction. Once we obtain ν_i, we recover σ_i by checking C_{ν_i} and β_i. Then by examining the entries of δ associated with each of the vertices in $V(\sigma_i)$ we obtain π_i. After obtaining all the σ_i, we can recover ρ. □

Lemma 8.3 (The matching switching lemma) *Let F be an r-disjunction over X, $|X| = 2m+1$. If $r \leq l$, $p^4n^3 \leq 1/10$ and $l \geq 10$, then*

$$\frac{|Bad_n^l(F, 2s)|}{|M_n^l|} \leq (21p^4m^3r)^s.$$

Proof: By Lemma 8.2, it is sufficient to bound the ratio

$$\frac{\left|\bigcup_{s\leq j\leq 2s} M_n^{l-j} \times Code(r, j) \times [2l+1]^{2s}\right|}{|M_n^l|}. \tag{9}$$

We begin by estimating the ratio $|M_n^{l-j}|/|M_n^l|$. A restriction in M_n^l is determined by the following process: choose $2(m-l)$ elements from X, then choose a matching that matches these $2(m-l)$ elements. Thus

$$\begin{aligned} |M_n^l| &= \binom{2m+1}{2(m-l)} \frac{(2(m-l))!}{(m-l)!2^{m-l}} \\ &= \frac{(2m+1)!}{(m-l)!(2l+1)!2^{m-l}} \,. \end{aligned} \tag{10}$$

Using (10) we estimate

$$\frac{|M_n^{l-j}|}{|M_n^l|} = \frac{(2l+1)^{\underline{2j}}}{(m-l+j)^{\underline{j}}2^j} \leq \left(\frac{(2l+1)^2}{2(m-l)}\right)^j \,. \tag{11}$$

Hence, using (11) and the estimate of $|Code(r, j)|$ from Lemma 6.1 we can bound the ratio (9) from above by the sum

$$\begin{aligned} & \sum_{s \leq j \leq 2s} \left(\frac{(2l+1)^2}{2(m-l)}\right)^j (2r)^j (2l+1)^{2s} \\ = \;& (4l^2(1+1/2l)^2)^s \sum_{s \leq j \leq 2s} \left(\frac{4l^2(1+1/2l)^2 r}{m-l}\right)^j \\ \leq \;& (4l^2 e^{1/l})^s \sum_{s \leq j \leq 2s} \left(\frac{4l^2 e^{1/l} r}{m-l}\right)^j \,. \end{aligned} \tag{12}$$

Since $r \leq l$, $p^4 n^3 \leq 1/10$ and $l \geq 10$, for $p = l/m$ we have

$$\frac{4l^2 e^{1/l} r}{m-l} \leq \left(\frac{4e^{1/10}}{1-p}\right)\left(\frac{l^3}{m}\right) < 4.421\left(\frac{p^4 m^3}{m}\right) \leq 0.0421.$$

Hence the sum in (12) is bounded by the sum of a geometric series with ratio < 0.0421 so that it is less than 1.05 times its largest term. This provides us with the estimate

$$\begin{aligned} \frac{|Bad_n^l(F, 2s)|}{|M_n^l|} &\leq 1.05(4l^2 e^{1/10})^s \left(\frac{4l^2 e^{1/10} r}{m-l}\right)^s \\ &= \left(\frac{16.8e^{1/5} p^4 m^3 r}{1-p}\right)^s \\ &< (21 p^4 m^3 r)^s, \end{aligned}$$

and completes the proof of the lemma. □

Lemma 8.4 *Let d be an integer, $0 < \epsilon < 1/5$, $0 < \delta < \epsilon^d$ and Γ a set of formulas of L_n of depth $\leq d$, closed under subformulas. If $|\Gamma| < 2^{m^\delta}$, $q = \lceil m^{\epsilon^d} \rceil$ and m is sufficiently large, then there exists $\rho \in M_n^q$ so that there is a $2m^\delta$-evaluation of $\Gamma \restriction \rho$.*

Proof: The proof of the lemma is similar to that of Lemma 7.1; there is only a slight difference of the probability in the switching lemma. □

Theorem 8.5 *Let $\mathcal{F}$ be a Frege system and $d > 4$. Then for sufficiently large m every depth d proof in $\mathcal{F}$ of $Graph_n$ must have size at least 2^{m^δ}, for $\delta < (1/5)^d$.*

Proof: The proof follows the same argument as in the proof of Theorem 7.2. □

9 Open problems The techniques used to prove lower bounds expounded in the earlier sections of this paper are quite powerful, but there appear to be difficulties in extending them to more general situations. In this final section, a few open problems are stated that seem to require extensions of the methods used here.

In Urquhart [16], it is shown that there is a family $\{G_n\}$ of bipartite expander graphs of bounded degree so that the sets of clauses $Clauses(G_n)$ require exponentially long refutations in the resolution proof system. Let $Taut(G_n)$ be the corresponding tautologies formed by negating all the clauses in $Clauses(G_n)$ and then forming their disjunction.

Problem 9.1 *Do the tautologies $Taut(G_n)$ require proofs of superpolynomial size in a bounded depth Frege system?*

The problem in adapting the current methods to the case of the tautologies based on the graphs G_n lies in the fact that the application of a restriction to a graph in general simplifies the graph considerably. By contrast, in the case of the graphs $K_{n+1,n}$ and K_n, the application of a restriction results in a graph of the same type on a smaller vertex set.

In [5], Chvátal and Szemerédi generalized the argument of [16] to show that a random set of clauses, provided it is not too large, is both unsatisfiable and requires exponentially large resolution refutations. To be precise, Chvátal and Szemerédi define the random family of m clauses of size k over n variables to be a family of clauses defined by picking m samples with replacement from the family of all clauses of size k in n variables. Their theorem is then as follows.

Theorem 9.1 *For every choice of positive integers c and k such that $k \geq 3$ and $c2^{-k} \geq 0.7$, there is a positive number ϵ such that, with probability tending to 1 as n tends to infinity, the random family of cn clauses of size k over n variables is unsatisfiable and its resolution complexity is at least $(1+\epsilon)^n$.*

It seems likely that progress with the first problem would also allow a generalization similar to the preceding theorem.

Problem 9.2 *Can the theorem of Chvátal and Szemerédi be generalized to bounded depth Frege systems?*

Acknowledgments The authors thank Paul Beame, Stephen Cook, Peter Koellner, Quaid Morris, Toniann Pitassi, François Pitt, Charles Rackoff and Arnold Rosenbloom for their comments and suggestions. The authors also gratefully acknowledge the support of the National Sciences and Engineering Research Council of Canada.

REFERENCES

[1] Ajtai, M., “The complexity of the pigeonhole principle,” pp. 346–355 in *Proceedings of the 29th Annual IEEE Symposium on the Foundations of Computer Science*, IEEE Computer Society Press, 1988. Zbl 0811.03042 MR 96a:03065 1, 1, 7

[2] Beame, P. “A switching lemma primer,” preprint, 1993. 1

[3] Beame, P., R. Impagliazzo, J. Krajíček, T. Pitassi, P. Pudlák, and A. Woods, “Exponential lower bounds for the pigeonhole principle,” pp. 200–220 in *Proceedings of the 24th Annual ACM Symposium on the Theory of Computing*, ACM Press, 1992. Zbl 0784.03034 MR 94f:03019 1, 1

[4] Bellantoni, S., T. Pitassi, and A. Urquhart, “Approximation and small-depth Frege proofs,” *SIAM Journal of Computing*, vol. 21 (1992), pp. 1161–1179. Zbl 0762.03020 MR 93m:03095 1, 1, 7

[5] Chvátal, V., and E. Szemerédi, “Many hard examples for resolution,” *Journal of the Association for Computing Machinery*, vol. 35 (1988), pp. 759–768. Zbl 0712.03008 MR 91f:68182 9

[6] Cook, S. A., *Resolution Lower Bound for Complete Graph Clauses*, manuscript, University of Toronto, Toronto, 1993. 8

[7] Furst, M., J. B. Saxe, and M. Sipser, “Parity, circuits, and the polynomial-time hierarchy,” pp. 260–270 in *Proceedings of the 22nd Annual IEEE Symposium on the Foundations of Computer Science*, IEEE Computer Society Press, 1981. Zbl 0534.94008 MR 86e:68048 1, 1

[8] Frege, G., *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*, Nebert, Halle, 1879. 2

[9] Håstad, J. T., *Computational Limitations of Small-Depth Circuits*, MIT Press, Cambridge, 1987. 1, 1

[10] Krajíček, J., P. Pudlák, and A. Woods, “Exponential lower bound to the size of bounded depth Frege proofs of the pigeonhole principle,” *Random Structures and Algorithms*, vol. 7 (1995), pp. 15–39. Zbl 0843.03032 MR 96i:03053 1, 1, 1, 5, 5

[11] Von Neumann, J., “Zur Hilbertschen Beweistheorie,” *Mathematische Zeitschrift*, vol. 26 (1926), pp. 1–46. 2

[12] Pitassi. T., P. Beame, and R. Impagliazzo, “Exponential lower bounds for the pigeonhole principle,” *Computational Complexity*, vol. 3 (1993), pp. 97–140. Zbl 0784.03034 MR 94f:03019 1, 1, 1

[13] Razborov, A. A., “Bounded arithmetic and lower bounds in Boolean complexity,” pp. 344–386 in *Feasible Mathematics II*, edited by P. Clote and J. Remmel, Birkhäuser, Boston, 1995. Zbl 0838.03044 MR 96d:03057 1

[14] Shoenfield, J., *Mathematical Logic*, Addison-Wesley, Reading, 1967. Zbl 0965.03001 MR 2001h:03003 2.1

[15] Tseitin, G. S., “On the complexity of derivation in propositional calculus,” pp. 115–125 in *Studies in Constructive Mathematics and Mathematical Logic, Part 2*, edited by A. O. Slisenko, 1970 (reprinted in *Automation of Reasoning Vol. 2*, edited by J. Siekmann and G. Wrightson, Springer-Verlag, New York, 1983, pp. 466–483). Zbl 0567.03002 1, 8

[16] Urquhart, A., “Hard examples for resolution,” *Journal of the Association for Computing Machinery*, vol. 34 (1987), pp. 209–219. Zbl 0639.68093 MR 89e:68056 9, 9

[17] Yao, A., "Separating the polynomial-time hierarchy by oracles," pp. 1–10 in *Proceedings of the 26th Annual IEEE Symposium on the Foundations of Computer Science*, IEEE Computer Society Press, 1985. 1, 1

Department of Philosophy
University of Toronto
215 Huron Street
9th Floor
Toronto, Ontario M5S 1A1
CANADA
email: urquhart@ai.toronto.edu