

Lower Bounds for Bounded Depth Frege Proofs via Pudlák-Buss Games

ELI BEN-SASSON

Technion

and

PRAHLADH HARSHA

The University of Texas at Austin

We present a simple proof of the bounded-depth Frege proof lower bounds of Pitassi et. al [*Computational Complexity*, 3(2):97-140, 1993] and Krajíček et. al [*J. Random Structures & Algorithms*, 7(1):15-39, 1995] for the pigeonhole principle. Our method uses the interpretation of proofs as two player games given by Pudlák and Buss. Our lower bound is conceptually simpler than previous ones, and relies on tools and intuition that are well-known in the context of computational complexity. This makes the lower bound of Pitassi et. al. and Krajíček et. al. accessible to the general computational complexity audience. We hope this new view will open new directions for research in proof complexity.

Categories and Subject Descriptors: F.2 [Theory of Computation]: Analysis of Algorithms and Problem Complexity

General Terms: Theory

Additional Key Words and Phrases: proof complexity, Frege proofs, pigeonhole principle, lower bounds

1. INTRODUCTION

In this paper we present an alternative proof of one of the most advanced theorems in proof complexity. Our method does not increase the strength of the claim itself (actually, it proves a slightly weaker claim), but it does improve the clarity of the proof. Our motivation is twofold. For those not versed in the intricacies of proof complexity, we offer a simple explanation of one of the deepest results in the field, using terminology that is familiar to the general public with moderate background in computational complexity. For “hard-core” proof-complexity researchers, we hope our presentation will help advance on other unsolved problems.

Author’s addresses: Eli Ben-Sasson, Computer Science Department, Technion – Israel Institute of Technology, Haifa, Israel. email: eli@cs.technion.ac.il and Prahladh Harsha, Department of Computer Science, The University of Texas at Austin, Texas, USA. email: prahladh@cs.utexas.edu.

Work done when the first author was at the Division of Engineering and Applied Sciences, Harvard University and the second author was at the Laboratory for Computer Science, Massachusetts Institute of Technology.

Permission to make digital/hard copy of all or part of this material without fee for personal or classroom use provided that the copies are not made or distributed for profit or commercial advantage, the ACM copyright/server notice, the title of the publication, and its date appear, and notice is given that copying is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or a fee.

© 20YY ACM 1529-3785/20YY/0700-0001 \$5.00

The pigeonhole principle is one of the simplest mathematical statements, and is used heavily in many mathematical proofs. It has special importance in discrete mathematics and combinatorics, where counting arguments carry the burden of many a proof. It is also the most extensively studied formula in proof complexity. In its simplest form, it claims that if $n+1$ pigeons sit in n pigeonholes, there must be a pigeonhole occupied by more than one pigeon. One of the major achievements of proof complexity has been to show that this simple claim is hard to prove in various proof systems such as resolution [Haken 1985], the polynomial calculus [Razborov 1998], and bounded depth Frege proofs [Ajtai 1994; Krajíček et al. 1995; Pitassi et al. 1993]. In this paper we give an alternative proof of the hardness of the pigeonhole principle for bounded depth Frege proofs.

1.1 Previous Results

The first super-polynomial lower bounds for the pigeonhole principle in bounded depth Frege were presented by Ajtai [Ajtai 1994]. This proof was simplified and improved by Bellantoni et. al. [Bellantoni et al. 1992]. The first exponential lower bounds were given by Pitassi et. al. [Pitassi et al. 1993] and independently by Krajíček et. al. [Krajíček et al. 1995]. Several extensions of this result have appeared over the years (see e.g. [Buss et al. 1996] and the recent [Buresh-Oppenheimer et al. 2005]). Our paper gives an alternative proof of the following exponential lower bound of [Pitassi et al. 1993] and [Krajíček et al. 1995].

THEOREM 1.1 [PITASSI ET AL. 1993; KRAJÍČEK ET AL. 1995]. *For any Frege system \mathcal{F} , and any integer d , there exists a constant $\delta > 0$ such that for any large enough n , the size of a depth d \mathcal{F} -proof of the pigeonhole principle of size n , is at least $\exp(n^\delta)$.*

Both lower bounds of [Pitassi et al. 1993] and [Krajíček et al. 1995] use a specially tailored *switching lemma*, and additionally a nonstandard interpretation of the lines of the proof. The complex switching lemma, combined with the non-intuitive interpretation, make the proofs extremely difficult to understand and explain. One successful line of research has led to the simplification of the switching lemma. This was initially done by Razborov [Razborov 1995], and then put in the context of the pigeonhole principle by Beame [Beame 1994]. Finally, Urquhart and Fu [Urquhart and Fu 1996] presented a complete simplified proof of the lower bound using the simpler proof of the switching lemma.

The other difficulty, which is the non-standard interpretation of the lines of a proof, has prevailed in all earlier proofs [Ajtai 1994; Bellantoni et al. 1992; Pitassi et al. 1993; Krajíček et al. 1995]. We believe that our alternative proof decreases this difficulty.

1.2 Constant Depth Circuits and Constant Depth Proofs

It is well-known that constant depth circuits are very inadequate for counting, and there is a standard technique to prove this - a switching lemma [Furst et al. 1984; Ajtai 1983; Håstad 1989]. The pigeonhole principle is a statement about counting, so it is natural to believe that it cannot be proved using reasoning that involves only constant depth circuits. A constant depth proof of the pigeonhole principle is merely a sequence of constant depth circuits. When seeking a lower bound for

constant depth proofs, the first thing a complexity researcher would do is hit all lines of a purported proof with a restriction, and use a switching lemma to argue that they are all transformed into simple functions, and hence cannot prove the pigeonhole principle. On second thought, there is a big problem with this approach. All lines of the proof, including the conclusion, are *tautologies*. Thus, even without a restriction, the functions computed by these circuits are the simplest possible - they all compute the constant 1 function ! It is clear that one needs to proceed differently. For this we need to gain a better understanding of a constant depth Frege proof.

1.3 Proofs as Games

The standard definition of Frege proofs can be found in any introductory book to mathematical logic. We use a different definition, introduced by Pudlák and Buss in [Pudlák and Buss 1994]. Under this definition, the proof of a tautology Φ is a two player game. Sam the Spoiler claims that he knows an assignment α setting Φ to 0. Pavel the Prover tries to expose his lie. Pavel is restricted to ask Sam only questions that have a yes/no answer. As a first solution, Pavel can ask the value of α on all variables and gates of Φ . Having done this, there will clearly be an inconsistency in one of the gates of Φ , and Sam's lie is exposed. The problem with this approach is that it requires a linear number of queries. A more efficient way is to present Sam with circuits, and ask for the value of these circuits on the input α . In this case, Pavel may save on the number of queries. Notice that Sam can lie in his answers to these queries, just as he was lying with respect to Φ . It is Pavel's role to decide whether to "believe" Sam's answer or try to expose the lie within a circuit. The beautiful observation of [Pudlák and Buss 1994] is that the minimal number of queries Pavel needs, is proportional to the logarithm of the minimal size Frege proof of Φ !!! Although stated for Frege proofs, their observation applies just as well to bounded depth proofs, with the following modification: The minimal number of constant depth queries needed by Pavel, is proportional to the logarithm of a minimal size constant depth Frege proof.

1.4 Finding A Strategy for Sam

By the previous discussion, a lower bound on the proof size of Φ is reduced to finding a strategy for Sam that enables him to answer many queries without contradicting himself. We accomplish this by transforming Pavel's queries to simple functions that will be *locally consistent* with each other. For simplicity assume Φ is a DNF, i.e. an OR of terms, where each term is an AND of literals. Suppose each query of Pavel is a term of Φ . In this case Sam can start with no assignment in hand, and with each query he will set one more variable, in a way that will set the queried term to 0. Thus, if these assignments do not fix any other term to 1, Sam will be consistent for a long time. Assume instead Pavel asks an OR of several terms. Sam is forced to answer 0 (otherwise he contradicts the answer 0 given to Φ), but he does not need to extend his assignment, and only when Pavel ask about each of the terms must Sam extend his assignment. But by this time, Pavel has asked many queries, and recall that Sam's aim is to maximize the number of queries.

Our lower bound is simply a strategy for Sam. We use the observation that given n holes, we can comfortably fit n or less pigeons into them without any double-

occupancies. Thus, Sam’s strategy in response to each query will be to assign a small number of pigeons to holes, and use this added information to compute an answer to the queries. If each response assigns k pigeons, he will succeed in answering consistently for n/k rounds, giving a lower bound of $2^{n/k}$ on the minimal bounded depth Frege proof size.

More to the point, Sam transforms every query φ to a *small domain function* f_φ , which is a Boolean function that is fixed by an assignment of a small number of pigeons (hence its name). Sam starts with the assignment α being empty, and when asked the value of $\varphi(\alpha)$, tries to evaluate $f_\varphi(\alpha)$. If necessary, Sam fixes a few more pigeons (remember that f_φ has small domain), and replies $f_\varphi(\alpha)$.

The small domain functions are constructed bottom up. Each variable is transformed to a function that can be fixed by setting a single pigeon. Negation gates are easy to handle: if Sam answered φ by $f_\varphi(\alpha)$ then answering $\neg\varphi$ by $\neg f_\varphi(\alpha)$ is consistent, and does not increase the size of α . The only tricky case is that of an OR gate, and this is where the switching lemma comes in. Suppose $\psi = \vee(\varphi_1 \dots \varphi_k)$, and suppose we have inductively constructed small domain functions $f_{\varphi_1} \dots f_{\varphi_k}$ for $\varphi_1, \dots, \varphi_k$. The switching lemma we use says that after applying a random restriction, there exists a small domain function f_ψ that is consistent with the functions $f_{\varphi_1} \dots f_{\varphi_k}$, meaning

- (1) f_ψ can be fixed by setting only a small number of pigeons.
- (2) For any assignment α fixing f_ψ to 1 there is some function f_{φ_i} such that $f_{\varphi_i}(\alpha) = 1$.
- (3) Any assignment α setting f_ψ to 0 cannot be extended to an assignment setting some f_{φ_i} to 1.

Thus, our answers are consistent also with respect to OR gates. Finally, we show that under this transformation, the pigeonhole principle is mapped to the constant 0 function, having an empty domain. This means that on the first query which is the pigeonhole principle itself, Sam answers 0, and does not need to extend α at all.

Let us sum up. Given a small purported bounded depth proof π , in the form of a set of queries, Sam transforms them to locally consistent, small domain functions. Sam initializes α to be empty, and with each query φ extends α so that $f_\varphi(\alpha)$ is fixed. He answers φ by $f_\varphi(\alpha)$. Since α is extended every time by fixing only a small number of pigeons, and all answers are locally consistent, Sam can keep on for many rounds. By the basic theorem of [Pudlák and Buss 1994] this implies that the minimal proof size is of exponential size.

We conclude by making two comments. First, a disclaimer. The switching lemma and the transformation we use are very similar to those originally used in previous proofs, most notably that of [Urquhart and Fu 1996]. Having said that, notice that the lower bound we present transferred the problem from the realm of logic and propositional proofs to that of constructing locally consistent partial functions. This latter problem is more accessible to computational complexity techniques (such as the switching lemma). We hope that our general approach can be extended, using similar complexity techniques, to derive lower bounds for other formulae (e.g. random 3-CNFs) and other proof systems (e.g. bounded depth Frege with counting gates).

1.5 Paper Organization

After giving formal definitions of Pudlák-Buss proofs and the pigeonhole principle (Section 2), we present a general sufficient condition for obtaining lower bounds for Frege proofs (Section 3). This condition applies to any Frege system, and any tautology Φ . We then show how to obtain this condition in the special case of the pigeonhole principle and constant depth Frege (Section 4). We end by presenting the lower bound itself (Section 5).

2. PRELIMINARIES

2.1 Bounded Depth Frege - Definitions

We begin by recalling the standard definitions of Frege proof systems and bounded depth Frege systems, in particular. For simplicity our logical language will be restricted to constants 0 (representing *false*), and 1 (representing *true*), and connectives $\{\vee, \neg\}$ where \vee is allowed to have unbounded fan-in. We will use the connective \wedge as a shorthand for $\neg \vee \neg$, and $A \implies B$ as a shorthand for $\neg A \vee B$.

We work with the Frege system **H** described by Bellantoni, Pitassi and Urquhart [Bellantoni et al. 1992], which is the standard bounded-depth Frege system, that was used for proving the exponential lower bounds for the pigeonhole principle. Lines of a proof are unbounded fan-in formulae over $\{\neg, \vee\}$. Namely, the allowable formulae are defined inductively by the rules:

- (1) A variable x is a formula, and so is the literal $\neg x$.
- (2) If A is a formula, then so is $\neg A$.
- (3) If Γ is a finite set of formulae, then so is $\vee \Gamma$. We will use the notation $A \vee B$ to mean $\vee\{A, B\}$.

The *depth* of a literal is 0, the *depth* of a formula is the maximal number of alternations of connectives in it and the *size* of the formula is the number of occurrences of connectives. Under this convention, the depth of a clause (disjunction of literals) is 0, the depth of a conjunction of literals is 1, the depth of a DNF is 2, and of a CNF 3. A conjunction of literals has depth one larger than a disjunction due to the conversion from \wedge to $\{\neg, \vee\}$. For similar reasons, depth of a DNF is 2 while that of a CNF is 3. We denote by $d(\varphi)$ the depth of the formula φ .

We now list the rules of **H**, which form a complete proof system over the basis $\{\vee, \neg\}$. We use the notation $\frac{\varphi_1 \dots \varphi_k}{\psi}$ to denote that ψ can be derived from $\{\varphi_1 \dots \varphi_k\}$.

- (1) Excluded Middle axiom: $\frac{}{A \vee \neg A}$
- (2) Weakening Rule: $\frac{A}{A \vee B}$
- (3) Merging Rule: $\frac{\vee(\{\vee \Gamma\} \cup \Delta)}{\vee(\Gamma \cup \Delta)}$
- (4) Unmerging Rule: $\frac{\vee(\Gamma \cup \Delta)}{\vee(\{\vee \Gamma\} \cup \Delta)}$
- (5) Cut Rule: $\frac{(A \vee B), (\neg A \vee C)}{B \vee C}$

A depth d Frege proof of a formula φ is a sequence of depth d formulae $\pi = \{\varphi_1, \dots, \varphi_s\}$, the last one being φ , where each formula in the sequence is either an excluded middle axiom, or is derived from previous lines by one of the other rules

listed above. The *size* of a proof is the sum of the sizes of the formulae in the proof. The *depth* of the proof is the maximal depth of the formulae in the proof.

2.2 Pudlák-Buss Games

Our proof uses the equivalent elegant definition of Frege systems given by [Pudlák and Buss 1994]. A Pudlák-Buss proof of Φ is best thought of a two player game, very similar to a modern criminal trial. Pavel the Prover (or Prosecutor) wants to convince us that Φ is a tautology, whereas Sam the Spoiler tries to cheat us into believing this is not the case. The trial starts with Sam claiming he knows an assignment α such that $\Phi(\alpha) = 0$. The trial then proceeds in rounds. In round t , Pavel presents a Boolean formula φ_t , and Sam answers with a single bit b_t , which is the claimed value of $\varphi_t(\alpha)$. After several rounds, Pavel addresses the jury and presents an inconsistency in Sam's answers. The jury has very limited understanding of the mysteries of Boolean formulae, and only knows the definition of the basic Boolean gates (say $\{\neg, \vee\}$). Thus, for Pavel to convict Sam, he needs to present an *immediate contradiction*.

Definition 2.1 Immediate Contradiction. For B a set of Boolean gates, an *immediate contradiction* with respect to B is a set of formulae, $\psi, \varphi_1, \dots, \varphi_k$ and a set of bits a, b_1, \dots, b_k such that

- (1) ψ is $g(\varphi_1, \dots, \varphi_k)$, where $g \in B$.
- (2) Sam was asked the formulae $\psi, \varphi_1, \dots, \varphi_k$, and gave answers a, b_1, \dots, b_k to them respectively.
- (3) $a \neq g(b_1, \dots, b_k)$.

If a set of answers b_1, \dots, b_S to a set of queries $\varphi_1 \dots \varphi_S$ includes no immediate contradiction as a subset, we call these answers *locally consistent*.

Notice that for Condition 1 to hold, we need *syntactical* equivalence, i.e. φ has to be syntactically the same as $g(\varphi_1, \dots, \varphi_k)$. The *semantical* equivalence of the two is not enough. For instance, φ is not syntactically the same as $\neg\neg\varphi$ although they are semantically equivalent. This distinction between semantical and syntactical objects is at the heart of our lower bounds.

A proof of Φ is a set of queries that convicts Sam for any answers he gives to the queries. Naturally, Pavel's queries may depend on Sam's answers. Thus, a proof is a binary tree, called a *game tree*, where each internal node is labeled by a query of Pavel, and each edge is labeled by Sam's answer to that query. The root is labeled Φ and has a single edge labeled 0. We say that a game tree *convicts* Sam (on Φ) if every leaf ℓ is labeled by an immediate contradiction as described in Definition 2.1, where (reusing the notation of Definition 2.1) $\psi, \varphi_1, \dots, \varphi_k$ are labels of some nodes on the path leading to ℓ , and a, b_1, \dots, b_k are the edges leaving $\psi, \varphi_1, \dots, \varphi_k$, respectively, on the path leading to ℓ . We say a proof has depth d if all queries are depth d formulae and we define the *height* of the proof to be the length of longest path from the root to a leaf in the tree. Finally, the *size* of the proof is the number of nodes in it.

The following theorem was originally proved by [Pudlák and Buss 1994] for Frege proofs, but their proof applies directly to bounded Frege.

THEOREM 2.2 [PUDLÁK AND BUSS 1994], PROPOSITION 2. *For any Frege system \mathcal{F} there exist integers c, c' such that the following holds.*

—If Φ has a standard \mathcal{F} -proof of size S and maximal depth d , then Φ has a Pudlák-Buss proof of height $\log(S) + O(1)$ and depth $d + c$ and each query is of size at most S .

—Conversely, if Φ has a Pudlák-Buss proof of height r and depth d , then Φ has a standard \mathcal{F} -proof of size 2^r and depth at most $d + c'$.

Remarks:

- (1) The exact set of connectives B is not extremely important, just as it is not very important in the definition of a standard Frege system [Cook and Reckhow 1979]. We only need B to be a complete basis for Boolean functions. For simplicity in this paper we fix $B = \{\neg, \vee\}$ where \vee can have unbounded fan-in.
- (2) Theorem 2.2 holds also for the case of B having gates with unbounded fan-in. In this case one only need naturally extend the definition of immediate contradictions to the unbounded fan-in gates. For example, an immediate contradiction to an unbounded \vee -gate is either an answer of 1 to the output of the gate and 0's to *all* its inputs, or an answer of 0 to the output and 1 to one of its inputs.
- (3) The size of a line in a standard Frege proof of Φ is wlog polynomially bounded by the number of lines in the proof of Φ and the size of Φ itself [Krajíček 1995]. Thus we assume wlog that the size of each query (i.e. the size of the formula being queried) is polynomially bounded by the number of nodes in the proof tree, and the size of Φ .

For the sake of completeness, we include the proof of the first part, which is the part we need for our lower bound.

PROOF OF THEOREM 2.2 (PART I). Let $\varphi_1, \dots, \varphi_S$ be a standard Frege proof of $\Phi = \varphi_S$, where each formula has depth $\leq d$. We present a Pudlák-Buss proof for it over a set of Boolean gates B . We assume wlog that the AND gate (denoted \wedge) has a constant depth encoding in B .

After querying Φ and receiving the answer 0, Pavel queries $\wedge(\varphi_1, \dots, \varphi_S)$. If Sam answers 1, this immediately contradicts the answer to Φ . Otherwise, Pavel conducts a binary search to find the smallest i such that Sam answers $\wedge(\varphi_1, \dots, \varphi_i)$ by 1, and $\wedge(\varphi_1, \dots, \varphi_{i+1})$ by 0. Notice that this requires $\log(S)$ queries. If no such i exists, Sam answered φ_1 by 0. By the definition of a Frege system, φ_1 is an axiom, i.e. it is defined by a substitution to a *constant size* tautology (such as $A \vee \neg A$), and a constant number of queries reveals an immediate contradiction. If φ_{i+1} is an axiom of the Frege system, we find an immediate contradiction as in the previous case.

Otherwise, φ_{i+1} was derived by some derivation rule (e.g. the cut rule) from a constant number of previous formulae $\varphi_{i_1} \dots \varphi_{i_k}$. Pavel queries $\varphi_{i_1} \dots \varphi_{i_k}$ (a constant number of queries). If Sam answers any of these queries by 0, this immediately contradicts his answer to $\wedge(\varphi_1, \dots, \varphi_i)$. Assuming all Sam's answers are 1, an additional constant number of queries reveals an immediate contradiction, because a Frege rule is defined by a substitution to a *constant size* tautology.

Notice that the depth of each query is at most $d+c$, where c is the depth required for encoding an AND gate in the basis B . \square

For the second part of the proof of Theorem 2.2 we refer the interested reader to the original proof of [Pudlák and Buss 1994], noting that the transformation from Pudlák-Buss proofs to standard ones does not increase the depth by more than a constant.

2.3 The Pigeonhole Principle

Fix sets D, R such that $D \cap R = \emptyset$, $|D| = n+1$, $|R| = n$, and denote $S = D \cup R$. Our set of connectives is $\{\neg, \vee\}$, so we use the notation $\bigwedge(\varphi_1, \dots, \varphi_k)$ as a shorthand for $\neg(\bigvee(\neg\varphi_1, \dots, \neg\varphi_k))$. The pigeonhole principle of size n , denoted PHP_n is the disjunction of the following four sets of formulae, over the variable set p_{ij} , $i \in D, j \in R$:

$$\begin{array}{ll} \neg \bigvee_{j \in R} p_{ij}, & i \in D; & p_{ik} \wedge p_{jk}, & i \neq j \in D, k \in R \\ \neg \bigvee_{i \in D} p_{ij}, & j \in R & p_{ij} \wedge p_{ik}, & i \in D, j \neq k \in R \end{array}$$

Each variable p_{ij} states whether pigeon i occupies pigeonhole j . It is to be noted that this version of the pigeonhole principle is called the onto and 1-to-1 version of the PHP. It is fine to work with this weaker version of PHP, as this only strengthens the corresponding lower bound obtained.

3. SAM'S STRATEGY

In this section, we present a general framework for proving lower bounds on proof size. By Theorem 2.2, a lower bound on the height of game-trees directly translates into a lower bound on the proof size of the related tautology. A lower bound on the height of the game-tree can be proved by demonstrating a strategy for Sam. Any strategy for Sam to escape being caught lying must satisfy the following two requirements.

- Answer the tautology Φ with 0.
- Answer Pavel's queries such that they are locally consistent.

A naive strategy to satisfy the second requirement would be to choose an assignment and answer Pavel's queries according to this assignment. However, no matter which assignment is chosen, this strategy fails to satisfy the first requirement since Φ is a tautology. We instead present an alternative strategy satisfying both the above requirements using partial functions.

First, for some notation. Let S be a set, $D \subseteq S$ and $f : D \rightarrow \{0, 1\}$ a function on D . The ordered pair (D, f) is then called a partial Boolean function on S . The set D is called the domain of f and is denoted by $\text{Dom}(f)$. For any set S , let Υ^S be the set of all partial Boolean functions defined on S . i.e.,

$$\Upsilon^S = \{(D, f) \mid D \subseteq S, f : D \rightarrow \{0, 1\}, \}$$

For any partial function (D, f) and $b \in \{0, 1\}$, let $f^{-1}(b) = \{x \in D \mid f(x) = b\}$.

For any game-tree \mathcal{T} , let $\Sigma_{\mathcal{T}}$ be the set of all formulae that occur in the game-tree \mathcal{T} . Any branch¹ of the game-tree is uniquely determined by the labels of the internal nodes and edges that occur along the branch. We will identify a branch with $((\varphi_1, b_1), \dots, (\varphi_s, b_s))$ where $\varphi_1, \dots, \varphi_s$ are the internal node labels and b_1, \dots, b_s the edge labels respectively starting from the root. (Notice that we do not include the label of the leaf in the branch since the leaf is uniquely specified by the label of the edge leading to it. Moreover, leaves in game-trees are labeled by immediate contradictions, not by queries.)

We now present a strategy for Sam using partial functions. Let \mathcal{T} be the game-tree for tautology Φ , proposed by Pavel that convicts Sam. Sam applies a transformation, mapping every formula $\varphi \in \Sigma_{\mathcal{T}}$ to a partial function $(D_{\varphi}, f_{\varphi})$, that satisfies the following three conditions.

- (1) $\forall x \in D_{\Phi}, f_{\Phi}(x) = 0$.
- (2) There exists a branch $((\varphi_1, b_1), \dots, (\varphi_s, b_s))$ in the game-tree \mathcal{T} such that

$$\bigcap_{i=1}^s (f_{\varphi_i})^{-1}(b_i) \neq \emptyset$$

- (3) For any subset $\Omega \subseteq \Sigma_{\mathcal{T}}$, if there exists a $x \in \bigcap_{\varphi \in \Omega} \text{Dom}(f_{\varphi})$, then the answers $(f_{\varphi}(x))_{\varphi \in \Omega}$ to the queries $(\varphi)_{\varphi \in \Omega}$ are locally consistent.

Condition 3 is the most important one of the above conditions as it ensures that Sam's answers to Pavel are locally consistent. We now prove that the existence of such a transformation provides Sam a strategy to answer Pavel without causing any immediate contradictions.

THEOREM 3.1. *Let Φ be a formula and \mathcal{T} a game-tree of height r for Φ . If there exists a set S and a transformation $\varphi \mapsto (D_{\varphi}, f_{\varphi})$, mapping every formula $\varphi \in \Sigma_{\mathcal{T}}$ to a partial function $(D_{\varphi}, f_{\varphi}) \in \Upsilon^S$, such that conditions 1, 2, and 3 are satisfied, then the game-tree \mathcal{T} does not convict Sam.*

PROOF. Let Φ be a formula and \mathcal{T} be a game-tree of height r for Φ . Suppose there exists a transformation $\varphi \mapsto (D_{\varphi}, f_{\varphi})$ as mentioned in the statement of the theorem. By condition 2, there exists a branch $((\varphi_1, b_1), \dots, (\varphi_s, b_s))$ in the game-tree \mathcal{T} satisfying

$$\bigcap_{i=1}^s (f_{\varphi_i})^{-1}(b_i) \neq \emptyset$$

Choose any $x \in \bigcap_{i=1}^s (f_{\varphi_i})^{-1}(b_i)$. Sam answers Pavel's queries $\varphi_1, \dots, \varphi_s$ along this branch with $f_{\varphi_1}(x) = b_1, \dots, f_{\varphi_s}(x) = b_s$ respectively. Note, Sam answers Pavel's first query $\varphi_1 = \Phi$ with $b_1 = 0$ since $f_{\Phi}(\Phi) = 0$ (by condition 1.) Since $x \in \bigcap_{i=1}^s \text{Dom}(f_{\varphi_i})$, we can conclude from condition 3 that Sam's responses to Pavel's queries along this branch are locally consistent. Hence, \mathcal{T} cannot be a game-tree that convicts Sam on Φ . \square

¹A branch is a path from the root to a leaf of the tree

4. COVERING PARTIAL FUNCTIONS AND K -TRANSFORMATIONS

In this section, we introduce *covering partial functions* and *k-transformations* which play the role of partial functions required for proving lower bounds on the proof-size for PHP_n . We will then show that a k -transformation mapping each formula to a covering partial function satisfies the three conditions of Sam's strategy, thus proving the lower bound. The covering partial functions and k -transformations introduced here are very similar to the matching decision trees and k -evaluations (introduced by Krajíček, Pudlák, and Woods [Krajíček et al. 1995]). In fact, all the steps in this part of the proof can be carried out using matching decision trees and k -evaluations. We, however, use covering partial functions and k -transformations since we believe these definitions are more natural to our proof setting.

Let D, R be two fixed non-empty sets such that $D \cap R = \emptyset$, $|D| = n + 1$, $|R| = n$ and let $S = D \cup R$. A *matching* between D and R is defined as a set of mutually disjoint unordered pairs $\{i, j\}$, where $i \in D, j \in R$. M^S denotes the set of all matchings between D and R . For a matching π , let $|\pi|$, called the *size* of the matching, denote the number of ordered pairs $\{i, j\}$ in π . For any subset $N \subseteq M^S$, of matchings, let $|N|$ denote $\max_{\pi \in N} |\pi|$, i.e., the size of the largest matching in N . Two matching π and π' are said to be *incompatible* if $\pi \cup \pi'$ is not a matching (i.e., $\pi \cup \pi' \notin M^S$). A matching π is said to cover a vertex i if $\{i, j\} \in \pi$ for some $j \in S$. If π is a matching, then we denote by $V(\pi)$ the set of vertices covered by π . Given a subset of matchings $N \subseteq M^S$ and an unordered pair $\{i, j\}$, define

$$N + \{i, j\} = \{\pi \cup \{i, j\} \in M^S \mid \pi \in N, \pi \cup \{\{i, j\}\} \in M^S\}.$$

Informally, $N + \{i, j\}$ is the set of all matchings in N , compatible with the (singleton) matching $\{\{i, j\}\}$, extended by the ordered pair $\{i, j\}$. For any subset of matchings N , define

$$\text{Cover}(N) = \{\pi \in M^S \mid \exists \pi' \in N, \pi' \subseteq \pi\}.$$

In other words, $\text{Cover}(N)$ is the set of all possible extensions of matchings in N .

Definition 4.1. A set of matchings N is said to be a complete matching set² for $S = D \cup R$ if

- $N = \emptyset$ or,
- there exists $i \in D$ (or $j \in R$) such that

$$N = \bigcup (N_{i,j} + \{i, j\}),$$

where (i) the union is over all $j \in R$ (or $i \in D$) and (ii) each $N_{i,j}$ is a complete matching set for $S \setminus \{i, j\} = (D \setminus \{i\}) \cup (R \setminus \{j\})$.

It easily follows from the above definition, that any distinct pair of matchings in a complete matching set are incompatible. Furthermore, for every matching ρ , there exists a matching in the complete matching set that is compatible with it, if $|\rho|$ is not too large.

²It is to be noted that complete matching sets are identical to complete matching trees (see [Urquhart and Fu 1996]), merely stated in terms more convenient in our setting.

PROPOSITION 4.2 [URQUHART AND FU 1996, LEMMA 4.3]. *Let N be a complete matching set for S and ρ be a matching in M^S such that $|\rho| + |N| \leq n$. Then there exists a matching $\pi \in N$ such that $\pi \cup \rho \in M^S$.*

Definition 4.3. A covering partial function over $S(= D \cup R)$ is an ordered pair (N, f) such that

- N is a complete matching set for S ,
- $(\text{Cover}(N), f)$ is a partial function on M^S ,
- If $\pi, \pi' \in \text{Cover}(N)$ such that $\pi \subseteq \pi'$, then $f(\pi') = f(\pi)$.

The last condition states that f on $\text{Cover}(N)$ is defined by the value of f on the complete matching set N .

We now introduce k -transformations which play the role of the transformation Γ in Sam's strategy. We then need to show that such a transformation satisfies the three conditions mentioned in the earlier section. Lemma 4.5 and Lemma 4.6 imply conditions 3 and 1 respectively, while we defer the proof of condition 2 (Lemma 5.1) to the next section. The notion of k -transformations is very similar to k -evaluations that was introduced by Krajíček, Pudlák, and Woods [Krajíček et al. 1995].

If φ is a disjunction and $\varphi_i, i \in I$, those subformulae of φ that are not disjunctions, but every subformula of φ properly containing them is a disjunction, then the *merged form* of φ is defined as the unbounded disjunction $\bigvee_{i \in I} \varphi_i$.

Let (N, f) and $(N_j, f_j), j \in J$ be covering partial functions over S . We say that (N, f) *satisfies Disj* $[\bigcup_{j \in J} \{(N_j, f_j)\}]$ if for all $\pi \in \text{Cover}(N)$

- $f(\pi) = 1 \implies \exists j \in J, \pi \in \text{Cover}(N_j)$ and $f_j(\pi) = 1$.
- $f(\pi) = 0 \implies \forall j \in J$, either $\pi \in \text{Cover}(N_j)$ and $f_j(\pi) = 0$ or $\pi \notin \text{Cover}(N_j)$. (i.e., f_j is not defined on π .)

Definition 4.4. Let Σ be a set of formulae closed under subformulae. Let $k > 0$. A k -transformation T is a mapping of formulae $\varphi \in \Sigma$ to covering partial functions (N_φ, f_φ) over S satisfying the following properties.

- (1) For all φ , $|N_\varphi| \leq k$.
- (2) $N_0 = N_1 = \emptyset$.³
 $\forall \pi \in \text{Cover}(N_0), f_0(\pi) = 0$,
 $\forall \pi \in \text{Cover}(N_1), f_1(\pi) = 1$
- (3) $N_{p_{ij}} = \left\{ \{ \{i, j\} \} \right\} \cup \left\{ \{ \{i, r\}, \{s, j\} \} \mid r \in R \setminus \{j\}, s \in D \setminus \{i\} \right\}$,
 $f_{p_{ij}}(\pi) = 1$ if $\{i, j\} \in \pi$ and $f_{p_{ij}}(\pi) = 0$ otherwise.
- (4) [Negation Condition]
 $N_{\neg\varphi} = N_\varphi; f_{\neg\varphi}(\pi) = \neg f_\varphi(\pi), \forall \pi \in \text{Cover}(N_\varphi)$.
- (5) [Disjunction Condition] If φ is a disjunction and $\bigvee_{j \in J} \varphi_j$ is the merged form of φ , then (N_φ, f_φ) satisfies Disj $[\bigcup_{j \in J} \{(N_{\varphi_j}, f_{\varphi_j})\}]$.

It follows from the above definition that but for the disjunction condition it is trivial to construct k -transformations. The disjunction condition is the heart of the k -transformation and ensures that no immediate contradictions arise out of a

³Note if $N = \emptyset$, then $\text{Cover}(N) = M^S$

disjunction gate. We shall later show how to construct k -transformation with the disjunction condition. The definition of a k -transformation is tailored to answer queries locally consistently and answer the PHP_n with 0 as seen from the following two lemmata.

LEMMA 4.5. *Let Σ be a set of a formulae closed under subformulae. Let T be a k -transformation mapping formulae $\varphi \in \Sigma$, to covering partial functions (N_φ, f_φ) over S . If for some $\Omega \subset \Sigma$, there exists a $\pi \in \bigcap_{\varphi \in \Omega} \text{Dom}(f_\varphi)$, then the answers $(f_\varphi(\pi))_{\varphi \in \Omega}$ to the queries $(\varphi)_{\varphi \in \Omega}$ are locally consistent.*

LEMMA 4.6. *If T is k -transformation for a set of formulae containing PHP_n , $k < n - 1$, then $f_{PHP_n}(\pi) = 0$ for all $\pi \in \text{Cover}(N_{PHP_n})$.*

PROOF OF LEMMA 4.5. Let Σ, T , and π be as stated in the lemma. Since \neg and \vee are the only two gates allowed in our language, it suffices to consider the following two cases. (one for negation and the other for disjunction.)

[Negation] Let $\varphi, \neg\varphi \in \Sigma$. By definition of a k -transformation, $f_{\neg\varphi}(\pi) = \neg f_\varphi(\pi)$ for all $\pi \in \text{Dom}(f_\varphi) = \text{Cover}(N_\varphi)$. Thus, an immediate contradiction cannot arise at a \neg gate.

[Disjunction] Let $\varphi = \bigvee_{i \in I} \varphi_i$ for some I . We have two sub-cases here.

(true case) Let $\varphi \in \Sigma$ and $\varphi_j \in \Sigma$ for some $j \in I$ such that $f_{\varphi_j}(\pi) = 1$ and $f_\varphi(\pi) = 0$. By definition of a k -transformation, $f_\varphi(\pi) = 0$ implies for all $i \in I$, either $\pi \in \text{Cover}(N_{\varphi_j})$ and $f_{\varphi_i}(\pi) = 0$ or $\pi \notin \text{Cover}(N_{\varphi_j})$. This contradicts $f_{\varphi_j}(\pi) = 1$. Thus, there is no immediate contradiction in this case.

(false case) Let $\varphi \in \Sigma$ and $\varphi_j \in \Sigma$ for all $j \in I$ such that $f_{\varphi_j}(\pi) = 0$ for all $j \in I$ and $f_\varphi(\pi) = 1$. By definition of a k -transformation, $f_\varphi(\pi) = 1$ implies there exists $i \in I$ such that $f_{\varphi_i}(\pi) = 1$. This contradicts $f_{\varphi_j}(\pi) = 0$. Thus, there is no immediate contradiction in this case too.

Thus, the answers according to evaluation at π are locally consistent. \square

PROOF OF LEMMA 4.6. PHP_n is the the disjunction of formulae of the form $\neg\varphi$ where φ ranges over

$$\begin{aligned} \bigvee_{j \in R} p_{ij}, \quad i \in D; & \quad \neg p_{ik} \vee \neg p_{jk}, \quad i \neq j \in D, k \in R \\ \bigvee_{i \in D} p_{ij}, \quad j \in R & \quad \neg p_{ij} \vee \neg p_{ik}, \quad i \in D, j \neq k \in R \end{aligned}$$

From the definition of a k -transformation, we infer that it suffices for us to show that $f_\varphi(\pi) = 1, \forall \pi \in \text{Cover}(N_\varphi)$ for each of the above φ .

Let $i \in D$. Let $\varphi = \bigvee_{j \in R} p_{ij}$. Suppose $f_\varphi(\pi) = 0$ for some $\pi \in \text{Cover}(N_\varphi)$. Let $\pi \in N_\varphi$ such that $f_\varphi(\pi) = 0$. Combining the facts that $|N_\varphi| \leq k$, $\pi \in N_\varphi$ and $k < n - 1$, we obtain $|\pi| < n - 1$. Hence, there exists a $\pi' \in M^S$ such that $\pi \subseteq \pi'$ and π' covers i . Let $\{i, j\} \in \pi'$ for some $j \in R$. But then $f_{p_{ij}}(\pi') = 1$ while $f_\varphi(\pi') = f_\varphi(\pi) = 0$ contradicting the disjunction condition in the definition of a k -transformation. Hence, $f_\varphi(\pi) = 1, \forall \pi \in \text{Cover}(N_\varphi)$ for φ of the specified type.

Let $i \neq j \in D, k \in R$. Let $\varphi = \neg p_{ik} \vee \neg p_{jk}$. Suppose $f_\varphi(\pi) = 0$ for some $\pi \in \text{Cover}(N_\varphi)$. Let $\pi \in N_\varphi$ such that $f_\varphi(\pi) = 0$. As before, we have $|\pi| < n - 1$. Since π is a matching, either $\{i, k\} \notin \pi$ or $\{j, k\} \notin \pi$. Without loss of generality assume $\{i, k\} \notin \pi$. Since $|\pi| < n - 1$, there exists a $\pi' \in M^S$ such that

$\pi \subseteq \pi'$ and $\{i, r\}, \{s, k\} \in \pi'$ for some $r \neq k \in R$ and $s \neq i \in D$. We, now have $\pi' \in \text{Cover}(N_{p_{ik}})$ and $f_{p_{ik}}(\pi') = 0$. Hence, $f_{\neg p_{ik}}(\pi') = 1$. But, by assumption, $f_\varphi(\pi') = f_\varphi(\pi) = 0$ again contradicting the disjunction condition.

The other two types of formulae are proved similarly. \square

For any matching ρ , let $D \upharpoonright_\rho = D \setminus V(\rho)$, $R \upharpoonright_\rho = R \setminus V(\rho)$ and $S \upharpoonright_\rho = S \setminus V(\rho)$ (recall that $V(\rho)$ is the set of vertices covered by the matching ρ). For a set of matchings N , let $N \upharpoonright_\rho = \{\pi \setminus \rho \in M^S \mid \pi \in N, \pi \cup \rho \in M_S\}$. That is, $N \upharpoonright_\rho$ is the set of matchings in N , compatible with ρ , after removing the edges in ρ . It can easily be checked that if N is a complete matching set for S , then $N \upharpoonright_\rho$ is a complete matching set for $S \upharpoonright_\rho$. For (N, f) a covering partial function over S , define $f \upharpoonright_\rho : \text{Cover}(N \upharpoonright_\rho) \rightarrow \{0, 1\}$ as follows: $f \upharpoonright_\rho(\pi) = f(\pi \cup \rho)$ for all $\pi \in \text{Cover}(N \upharpoonright_\rho)$. It can easily be checked that $(N \upharpoonright_\rho, f \upharpoonright_\rho)$ is a covering partial function over $S \upharpoonright_\rho$. If T is a k -transformation mapping formulae, $\varphi \in \Sigma$ to covering partial functions (N_φ, S_φ) over S , then the transformation $T \upharpoonright_\rho$ mapping $\varphi \in \Sigma$ to the covering partial function $(N_\varphi \upharpoonright_\rho, f_\varphi \upharpoonright_\rho)$ over $S \upharpoonright_\rho$ is also a k -transformation.

We need to show that for any small set Σ of formulae, there exists a k -transformation mapping formulae, $\varphi \in \Sigma$, to covering partial functions (N_φ, f_φ) over S . Unfortunately, we won't be able to do exactly that, but we instead prove the following which is equally good. We show that for any small set Σ of formulae, there is a k -transformation mapping formulae, $\varphi \in \Sigma$, to covering partial functions (N_φ, f_φ) over $S \upharpoonright_\rho$ for some matching ρ . It can easily be checked that Lemma 4.5 holds even if the covering partial functions (N_φ, S_φ) are over $S \upharpoonright_\rho$ rather than over S .

As mentioned before, the difficulty in constructing a k -transformation lies in satisfying the disjunction condition. We use the following variant of the Switching Lemma to build covering partial functions satisfying the disjunction condition. This version of the Switching Lemma can be proved by methods similar to that in the Switching Lemma Primer [Beame 1994] and is in fact a restatement of the switching lemma of [Urquhart and Fu 1996] in our terminology.

LEMMA 4.7 [URQUHART AND FU 1996], LEMMA 6.4: SWITCHING LEMMA. *Let $(N_j, f_j), j \in J$ be covering partial functions over S such that $|N_j| \leq r$ for all $j \in J$. Let $l \geq 10$ and $p = l/n$. If $r \leq l$ and $p^4 n^3 \leq 1/10$, then for random $\rho \in M^S$ such that $|\rho| = n - l$, the event that “There exists a covering partial function (N, f) over $S \upharpoonright_\rho$ such that (N, f) satisfies $\text{Disj} [\bigcup_{j \in J} \{(N_j \upharpoonright_\rho, f_j \upharpoonright_\rho)\}]$ and $|N| < 2s$ ” holds with probability at least $1 - (11p^4 n^3 r)^s$*

In other words, with high probability a random restriction converts the disjunction into a formula that only depends on a complete matching set of small size ($2s$). Note that this is not the same as saying that the disjunction is converted to a formula that involves only $2s$ variables.

LEMMA 4.8. *Let d be an integer, $0 < \epsilon < 1/5, 0 < \delta < \epsilon^d$ and Σ a set of formulae of depth d , closed under subformulae. If $|\Sigma| < 2^{n^\delta}$, and n is sufficiently large, then there exists a matching $\rho \in M^S$ of size $n - n^{\epsilon^d}$ such that there is a $2n^\delta$ -transformation T mapping formulae $\varphi \in \Sigma$, to covering partial functions (N_φ, f_φ) over $S \upharpoonright_\rho$.*

PROOF. The proof is by induction on d . For $d = 1$, the only formulae in Σ are

propositional variables and constants. For any such formula φ , $|N_\varphi| \leq 2$, so we have a 2-transformation.

Suppose the lemma holds for d . Let Σ be a set of formulae of depth $d+1$, closed under subformulae such that $|\Sigma| < 2^{n^\delta}$ where $0 < \delta < \epsilon^{d+1}$. Let Δ be the set of formulae in Σ of depth at most d . Since $0 < \delta < \epsilon^{d+1} < \epsilon^d$, by the induction hypothesis, there exists a $\rho' \in M^S$ of size $n - n^{\epsilon^d}$, such that there is a $2n^\delta$ -transformation T' mapping formulae $\varphi \in \Delta$, to covering partial functions (N'_φ, f'_φ) over $S \upharpoonright_{\rho'}$. Let φ be any formula in Σ of depth $d+1$ that is a disjunction and let $\bigvee_{j \in J} \varphi_j$ be its merged form. We then apply the Switching Lemma with $S \rightarrow S \upharpoonright_{\rho'}$, $n \rightarrow n^{\epsilon^d}$, $l \rightarrow n^{\epsilon^{d+1}}$, $r \rightarrow 2n^\delta$, $s \rightarrow n^\delta$. For sufficiently large n , the conditions for the Switching Lemma are satisfied. Hence, with probability at least $1 - (11n^{4\epsilon^{d+1}} n^{-\epsilon^d} 2n^\delta)^{n^\delta}$, there exists a covering partial function (N_φ, f_φ) over $(S \upharpoonright_{\rho'}) \upharpoonright_{\rho''} = S \upharpoonright_{\rho' \cup \rho''}$ such that (N_φ, f_φ) satisfies $Disj [\bigcup_{j \in J} \{(N_{\varphi_j} \upharpoonright_{\rho''}, f_{\varphi_j} \upharpoonright_{\rho''})\}]$ and $|N_\varphi| < 2n^\delta$ over random choices of $\rho'' \in M^{S \upharpoonright_{\rho'}}$ such that $|\rho''| = n^{\epsilon^d} - n^{\epsilon^{d+1}}$. Since $\delta < \epsilon^{d+1} < \epsilon^d/5$, for sufficiently large n , $11n^{4\epsilon^{d+1}} n^{-\epsilon^d} 2n^\delta < 11n^{-\epsilon^d/5} 2n^\delta < 1/2$ the quantity $(11n^{4\epsilon^{d+1}} n^{-\epsilon^d} 2n^\delta)^{n^\delta}$ is bounded above by 2^{-n^δ} . Hence, the above probability is bounded below by $1 - 2^{-n^\delta}$. Since there exist no more than 2^{n^δ} disjunctions of depth $d+1$ in Σ , there exists a single $\rho'' \in M^{S \upharpoonright_{\rho'}}$ such that $|\rho''| = n^{\epsilon^d} - n^{\epsilon^{d+1}}$ and for all disjunctions φ of depth $d+1$ in Σ , there exists a covering partial function (N_φ, f_φ) over $S \upharpoonright_{\rho' \cup \rho''}$ such that (N_φ, f_φ) satisfies $Disj [\bigcup_{j \in J} \{(N_{\varphi_j} \upharpoonright_{\rho''}, f_{\varphi_j} \upharpoonright_{\rho''})\}]$ and $|N_\varphi| < 2n^\delta$.

Let $\rho = \rho' \cup \rho''$. Note $|\rho| = n - n^{\epsilon^{d+1}}$. We can now define the $2n^\delta$ -transformation T that maps $\varphi \in \Sigma$ to covering partial functions over $S \upharpoonright_\rho$ as follows: If φ is a disjunction of depth $d+1$, map φ to (N_φ, f_φ) as given above. If $\varphi \in \Delta$, then map it to $(N'_\varphi \upharpoonright_{\rho''}, f'_\varphi \upharpoonright_{\rho''})$ if T' maps φ to (N'_φ, f'_φ) . Finally, if φ is a negation of depth $d+1$, then map φ to $(N_\psi, \neg f_\psi)$ where $\varphi = \neg\psi$. Clearly, this transformation T is a $2n^\delta$ -transformation mapping formulae $\varphi \in \Sigma$, to covering partial functions (N_φ, f_φ) over $S \upharpoonright_\rho$ where $\rho \in M^S$ such that $|\rho| = n - n^{\epsilon^{d+1}}$. \square

5. LOWER BOUND FOR PHP_N

In this section, we use the k -transformations to demonstrate a strategy for Sam as indicated in Section 3. From this, we obtain a lower bound on the size of bounded depth Frege proofs of PHP_n .

LEMMA 5.1. *Let \mathcal{T} be a game-tree of height r for PHP_n . Let Σ be the set of all formulae that occur in \mathcal{T} and their subformulae. Let T be a k -transformation mapping formulae $\varphi \in \Sigma$ to covering partial functions (N_φ, f_φ) over $S \upharpoonright_\rho$ for some matching $\rho \in M^S$ of size $n - m$. If $kr \leq m$, then there exists a branch $((\varphi_1, b_1), \dots, (\varphi_s, b_s))$ in the game-tree \mathcal{T} such that*

$$\bigcap_{i=1}^s (f_{\varphi_i})^{-1}(b_i) \neq \emptyset$$

PROOF. Let $\mathcal{T}, \Sigma, \rho$ and T be as stated in the lemma. Consider the following procedure $WALK(\mathcal{T})$ that outputs one of the branches of \mathcal{T} .

WALK(\mathcal{T})

- (1) Set $\pi \leftarrow \emptyset$ and $i \leftarrow 1$.
- (2) Walk along the game-tree \mathcal{T} starting from the root node (labeled PHP_n) until a leaf is reached as follows:
 - (a) Set $\varphi_i \leftarrow$ label of current node.
 - (b) Choose a $\pi \in N_{\varphi_i}$ such that $\pi \cup \pi_i \in M^{S|\rho}$.
 - (c) Set $b_i \leftarrow f_{\varphi_i}(\pi_i)$ and $\pi \leftarrow \pi \cup \pi_i$.
 - (d) Walk along edge labeled b_i leading out of current node.
 - (e) Increment i .
- (3) Output $((\varphi_1, b_1), \dots, (\varphi_s, b_s))$.

Since \mathcal{T} is a game-tree for PHP_n , we have $\varphi_1 = PHP_n$ and $b_1 = 0$ for any branch in \mathcal{T} . By Lemma 4.6, $f_{PHP_n}(\pi) = 0$ for all $\pi \in \text{Cover}(PHP_n)$. Hence, WALK algorithm can choose any matching $\pi \in N_{PHP_n}$ at Step 2b in the first execution of the loop at Step 2. For latter executions of the loop, as long as $|\pi| + k \leq m$, Proposition 4.2 guarantees that a matching $\pi_i \in N_{\varphi_i}$ satisfying $\pi \cup \pi_i \in M^{S|\rho}$ can be chosen in Step 2b. As $|\pi|$ is extended at most r times, each time at most by k , and $rk \leq m$, the condition $|\pi| + k \leq m$ is true before each execution of Step 2b.

Let π be the matching at the final step of WALK algorithm. The branch $((\varphi_1, b_1), \dots, (\varphi_s, b_s))$ output by WALK satisfies $b_i = f_{\varphi_i}(\pi)$. Hence,

$$\pi \in \bigcap_{i=1}^s (f_{\varphi_i})^{-1}(b_i).$$

Thus, $\bigcap_{i=1}^s (f_{\varphi_i})^{-1}(b_i) \neq \emptyset$ \square

We are now ready to prove our main theorem.

THEOREM 5.2 [PITASSI ET AL. 1993; KRAJÍČEK ET AL. 1995]. *Let \mathcal{F} be a Frege system and let c be the constant that occurs in Theorem 2.2 corresponding to \mathcal{F} . Then for sufficiently large n , every depth d proof in \mathcal{F} of PHP_n must have size at least 2^{n^μ} , for $\mu < \frac{1}{2} \left(\frac{1}{5}\right)^{d+c}$.*

PROOF. Let $0 < \epsilon < \frac{1}{5}$ and $0 < \mu < \epsilon^{d+c}/2$. Suppose PHP_n has a depth d proof in \mathcal{F} of size 2^{n^μ} . By Theorem 2.2, we have that there exists a Pudlák-Buss game-tree \mathcal{T} of height n^μ consisting of formulae of size at most 2^{n^μ} and depth at most $d+c$ convicting Sam on PHP_n . Let Σ be the set of all formulae and their subformulae that occur in this game-tree \mathcal{T} . Clearly, $|\Sigma| \leq 2^{n^\mu} \cdot 2^{n^\mu} = 2^{2n^\mu}$. Choose δ such that $\mu < \delta < \epsilon^d/2$. Then, for sufficiently large n , we have $|\Sigma| < 2^{n^\delta}$. By Lemma 4.8, there exists a partial matching ρ of size $n - n^{\epsilon^d}$ such that Σ has a $2n^\delta$ -transformation T mapping formulae, $\varphi \in \Sigma$, to covering partial functions, (N_φ, f_φ) over $S|\rho$. By Lemma 4.6, we have that condition 1 (i.e., $\forall x \in \text{Dom}(f_{PHP_n}), f_{PHP_n}(x) = 0$) is satisfied since $2n^\delta < n^{\epsilon^d} - 1$ for sufficiently large n . Also as $2n^\delta \cdot n^\mu \leq n^{\epsilon^d}$ for sufficiently large n , the hypothesis for Lemma 5.1 is satisfied. Hence, the $2n^\delta$ -transformation satisfies condition 2. By Lemma 4.5, we have that condition 3 is also satisfied. Thus, T satisfies all the three conditions of Theorem 3.1. Hence, \mathcal{T} does not convict Sam contradicting our assumption. Thus, there exists no depth d proof of PHP_n in \mathcal{F} of size 2^{n^μ} . \square

ACKNOWLEDGMENT

We thank the anonymous referees for pointing an error in the earlier presentation of the paper.

REFERENCES

- AJTAL, M. 1983. Σ_1^1 -formulae on finite structures. *Annals of Pure and Applied Logic* 24, 1 (July), 1–48.
- AJTAL, M. 1994. The complexity of the pigeonhole principle. *Combinatorica* 14, 4, 417–433. (Preliminary Version in 29th FOCS, 1988).
- BEAME, P. 1994. A switching lemma primer. Tech. Rep. UW-CSE-95-07-01, University of Washington.
- BELLANTONI, S., PITASSI, T., AND URQUHART, A. 1992. Approximation and small-depth frege proofs. *SIAM J. Computing* 21, 6 (Dec.), 1161–1179.
- BURESH-OPPENHEIM, J., BEAME, P., PITASSI, T., RAZ, R., AND SABHARWAL, A. 2005. Bounded-depth Frege lower bounds for weaker pigeonhole principles. *SIAM J. Computing* 34, 2, 261–276.
- BUSS, S. R., IMPAGLIAZZO, R., KRAJÍČEK, J., PUDLÁK, P., RAZBOROV, A. A., AND SGALL, J. 1996. Proof complexity in algebraic systems and bounded depth frege systems with modular counting. *Computational Complexity* 6, 3, 256–298.
- COOK, S. A. AND RECKHOW, R. A. 1979. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic* 44, 1 (Mar.), 36–50. (Preliminary Version in 6th STOC, 1974).
- FURST, M. L., SAXE, J. B., AND SIPSER, M. 1984. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory* 17, 1, 13–27. (Preliminary Version in 22nd FOCS, 1981).
- HAKEN, A. 1985. The intractability of resolution. *Theoretical Comp. Science* 39, 2–3 (Aug.), 297–308.
- HÅSTAD, J. 1989. Almost optimal lower bounds for small depth circuits. In *Randomness and Computation*, S. Micali, Ed. Advances in Computing Research, vol. 5. JAI Press, 143–170.
- KRAJÍČEK, J. 1995. *Bounded Arithmetic, Propositional Logic, and Complexity Theory*. Encyclopedia of Mathematics and its Applications, vol. 60. Cambridge University Press.
- KRAJÍČEK, J., PUDLÁK, P., AND WOODS, A. 1995. Exponential lower bounds to the size of bounded depth frege proofs of the pigeonhole principle. *Journal of Random Structures and Algorithms* 7, 1 (Aug.), 15–39. (Preliminary Version in 24th STOC, 1992).
- PITASSI, T., BEAME, P., AND IMPAGLIAZZO, R. 1993. Exponential lower bounds for the pigeonhole principle. *Computational Complexity* 3, 2, 97–140. (Preliminary Version in 24th STOC, 1992).
- PUDLÁK, P. AND BUSS, S. R. 1994. How to lie without being (easily) convicted and the length of proofs in propositional calculus. In *Proc. 8th International Workshop, Conference for Computer Science Logic (CSL)* (25–30 Sept.), L. Pacholski and J. Tiuryn, Eds. LNCS, vol. 933. Springer, 151–162.
- RAZBOROV, A. A. 1995. Bounded arithmetic and lower bounds in boolean complexity. In *FEASMATH: Feasible Mathematics II: A Mathematical Sciences Institute Workshop*. Progress in Computer Science and Applied Logic, vol. 13. Birkhäuser, 344–387.
- RAZBOROV, A. A. 1998. Lower bounds for the polynomial calculus. *Computational Complexity* 7, 4 (Dec.), 291–324.
- URQUHART, A. AND FU, X. Fall 1996. Simplified lower bounds for propositional proofs. *Notre Dame Journal of Formal Logic* 37, 4, 523–544.

Received June 2008; revised June 2009; accepted June 2009