comput complexity **3** (1993), 97–140 1016-3328/93/020097-44 $1.50+0.20

EXPONENTIAL LOWER BOUNDS FOR THE PIGEONHOLE PRINCIPLE

TONIANN PITASSI, PAUL BEAME
AND RUSSELL IMPAGLIAZZO

Abstract. In this paper we prove an exponential lower bound on the size of bounded-depth Frege proofs for the pigeonhole principle (PHP). We also obtain an $\Omega(\log\log n)$-depth lower bound for any polynomial-sized Frege proof of the pigeonhole principle. Our theorem nearly completes the search for the exact complexity of the PHP, as S. Buss has constructed polynomial-size, $\log n$-depth Frege proofs for the PHP. The main lemma in our proof can be viewed as a general Håstad-style Switching Lemma for restrictions that are partial matchings. Our lower bounds for the pigeonhole principle improve on previous superpolynomial lower bounds.

Key words. Complexity of propositional proof systems, lower bounds.
Subject classifications. 68Q99, 03F20, 68R05

1. Introduction

The main result of this paper is an exponential bound on the size of bounded-depth Frege proofs for the pigeonhole principle. Before describing the proof, we will discuss several motivations for studying lower bounds for Frege and bounded-depth Frege systems.

The complexity of Frege proofs has been studied extensively by many people in the last 20 years, beginning with an early paper by Tseitin (1968). Later, Haken (1985) proved that any Resolution proof of the pigeonhole principle must have exponential size. The next major breakthrough was made by Ajtai (1988) who used nonstandard model theory to prove that any constant-depth Frege proof of the pigeonhole principle must have superpolynomial-size. Because Resolution is a particular depth-2 Frege system, Ajtai's proof yields a superpolynomial lower bound for Resolution as a special case. More recently, Bellantoni *et al.* (1992) obtained a new proof of Ajtai's theorem which eliminates the use of nonstandard models. While their techniques were more direct and more accessible, their improved bound was still barely superpolynomial.

Our new exponential lower bound has several interesting consequences. As a corollary, we show that any polynomial-sized Frege proof of the pigeonhole principle must have depth $\Omega(\log\log n)$. Our theorem nearly completes the search for the exact complexity of the pigeonhole principle, as Sam Buss (1987) has constructed polynomial-sized, logarithmic depth Frege proofs for the pigeonhole principle.

Constant-depth lower bounds are related to the power of weak systems of arithmetic (see Buss 1987, Paris *et al.* 1988). This relationship together with our exponential lower bound for the propositional pigeonhole principle shows that relativized Bounded Arithmetic, $S_2(f)$, cannot prove the pigeonhole principle for f.

To see why this question is of interest in logic, consider the following two proof sketches that every non-zero residue modulo a prime has an inverse. Let p be a prime, and let $0 < a \leq p-1$. Then if we consider the map $F_a : \{0, ...p-1\} \rightarrow \{0, ..p-1\}$ defined by $F_a(b) = ab$ mod p, it is easy to see that F_a is $1-1$. Therefore, (using the pigeonhole principle), it must also be onto, and so 1 must be in the image. Therefore, there exists a number b, $0 < b \leq p-1$, such that $ab = 1$ mod p. In the second proof, we would prove by induction on the length of numbers a, b that Euclid's Algorithm for extended gcd finds integers c, d so that $ca + db = gcd(a, b)$. Then applying this algorithm to a and p, we get $ca + dp = 1$, so $ca = 1$ mod p.

Both of the above proofs are simple, and only use basic facts of arithmetic. Both are constructive in the sense of intuitionistic logic. However, the first is combinatorially "non-constructive" in that it is based on a counting argument which yields no better way of finding the proven object than via exhaustive search. The second has "algorithmic content", and yields a good method for finding the object proven to exist. In this case, a counting argument was not necessary, and could be replaced by a more constructive computational argument. Our result can be interpreted as saying that there is no *generic procedure* for converting a counting argument involving exponentially large but finite sets into an argument which only involves concepts in the polynomial-time hierarchy (relative to the object being counted). Thus, in general, one cannot automatically convert such an argument into a more algorithmic one, although in any particular case, this might be possible using special properties of the sets being counted.

In contrast with this negative result, Paris, Wilkie and Woods (1988) showed that the weak pigeonhole principle, $WPHP_n$, is provable in $S_2(f)$. (PHP_n states that there is no 1-1 map from $[n+1]$ to $[n]$, while $WPHP_n$ states that there is no 1-1 map from $[2n]$ to $[n]$.) As a corollary, they show that $WPHP_n$

has quasi-polynomial size, constant-depth Frege proofs.

It is not hard to extend our results to weakenings of the pigeonhole principle that state the nonexistence of 1-1 mappings from sets of size $n + c$ to n (the lower bound is only minimally affected by c.) However, it is still an open problem whether $WPHP_n$ has constant-depth proofs of polynomial size. We can also extend our result quite easily to another weak version of the pigeonhole principle, which states that there is no 1-1 *and onto* map from $[n + 1]$ to $[n]$.

In all of the lower bounds for propositional proof systems mentioned above, *size* refers to the total number of *symbols* in the proof. Another measure of complexity is *line-size*, which counts the total number of formulas, or lines, in the proof. Clearly, a lower bound on line-size implies a lower bound on symbol size. For constant-depth Frege proofs, S. Buss (1993) has recently shown that superpolynomial lower bounds for symbol-size imply superpolynomial lower bounds for line-size. Therefore, our result also holds for line-size.

Our lower bound is a proof by induction on the depth of the Frege proof. The method used to reduce the depth from d to $d - 1$ is a new variation on the the bottom-up method of restrictions, first described in (Furst et al., 1984), and later improved by Yao (1985), Håstad (1987) and others. The key combinatorial lemma used in our proof is a new switching lemma, similar in spirit to the switching lemma of Håstad .

Håstad's switching lemma states that with high probability, a random restriction allows us to re-write an OR of small ANDs as an AND of small ORs. A major drawback of this lemma and related ones is that they only apply when there is very little dependency between variables in the underlying probability distribution of restrictions. There are many graph-based problems where the dependency between variables is too great to apply Håstad's Lemma, and there is no known reduction from a known hard problem in AC^0 to one of these problems. One graph-based problem for which a Håstad-style switching lemma has been shown is that of deciding whether or not a graph contains a clique on a small number of nodes (see Lynch 1986, Beame 1990). However, the restrictions needed in that case still have very limited dependency.

In this paper, we prove a new switching lemma which applies to restrictions for which there is a great deal of dependency, namely those that represent partial matchings. A key feature that makes this more difficult is that after our restrictions are applied, the converted formula is only equivalent to the original one for certain classes of assignments.

Below we will outline an overview of the proof. In Section 2, we give some preliminary definitions concerning our random restrictions. In Section 3, we discuss 1-1 decision trees, and state the main combinatorial lemma. In Sections

4 and 5, we present the exponential lower bound for the pigeonhole principle. Finally, in Section 6, we prove the main combinatorial lemma. This result has also been obtained independently by Krajíček, Pudlák and Woods (1992).

1.1. Proof Overview. A Frege proof is a sequence of propositional formulas, each of which is either an axiom instance or follows from previous formulas by one of a fixed set of inference rules. The pigeonhole principle can be expressed by a class of propositional formulas, $\{\mathrm{PHP}_n : n \in N\}$, where PHP_n asserts that there is no 1-1 mapping from a set D_0 of size $n+1$ to a set D_1 of size n. We encode PHP_n using $(n+1)n$ propositional variables, $\{P_{ij} : i \in D_0 \wedge j \in D_1\}$, where D_0 and D_1 are disjoint sets such that $|D_0| = n+1$ and $|D_1| = n$. Intuitively, $P_{ij} = 1$ iff i is mapped to j. Since our proof system will be a refutation system, we are concerned with the statement $\neg\mathrm{PHP}_n$, which can be written as the conjunction of the following *pigeonhole clauses*:

$$\begin{gathered} \bigvee\{P_{ij} : j \in D_1\},\ i \in D_0; \\ \bigvee\{\neg P_{ik}, \neg P_{jk}\},\ i \neq j,\ i,j \in D_0,\ k \in D_1. \end{gathered}$$

In a refutation, one starts with the negated pigeonhole principle, $\neg PHP_n$, and then derives $\bigvee\{\}$, i.e. False. Since we will be working in a Frege system over the basis OR and NOT, we will begin with $\neg PHP_n$, written as a depth-4 formula over the basis OR, and NOT.

As in the paper by Bellantoni, Pitassi and Urquhart (1992), we proceed by induction on the depth of the Frege proof. Assume that we have a small, depth d Frege proof of the pigeonhole principle. Without loss of generality, we also assume that each formula in the proof consists of ORs and NOTs, except for the bottom two levels which are ORs of small ANDs. Applying a random restriction to each formula in the refutation, we can simplify the bottom levels so that each occurrence of negation at depth 3 of each formula is replaced by the "pseudo complement". This allows us to reduce the depth of each formula to $d-1$, but now each depth $d-1$ formula only approximates the original depth d formula on the reduced domain. Due to this approximation, instead of obtaining a depth $d-1$ refutation of the pigeonhole principle (on the reduced domain) which is completely sound, we obtain a depth $d-1$ *approximate* refutation which is only approximately sound.

An approximate refutation is a Frege refutation where each inference is sound with respect to a large subset of all critical truth assignments. In contrast, an inference in a regular Frege refutation is sound with respect to all truth assignments. The approximation is obtained by a new method which will be described in the next section. The key property of the approximation is that the pseudo-complement has the property that it is identical to the actual complement on a large fraction of the assignments that are maximally 1-1, namely the critical truth assignments.

We repeat the restriction argument $d - 2$ times to obtain an approximate depth-2 Frege refutation of the pigeonhole principle, i.e., a refutation in which each formula is an OR of small ANDs. We then apply a separate base case argument which shows that there can be no good approximation to a Frege proof of small size and with this special form.

2. Random Restrictions and Map Disjunctions

Let D consist of two disjoint sets, D_0 and D_1. (When we refer to D as $D_0 \cup D_1$, we mean the disjoint union.) The *variables over* $D = D_0 \cup D_1$ are $\{P_{ij} : i \in D_0,\ j \in D_1\}$. Pictorially, D_0 can be thought of as a set of indices representing the pigeons, D_1 as a set of indices representing the holes, and for each $i \in D_0$, and $j \in D_1$, the edge between i and j is labelled by the variable P_{ij}.

A truth assignment φ over D is any total assignment of $\{0, 1\}$ to the variables over D. Let $D' = D'_0 \cup D'_1$, $D'_0 \subseteq D_0$, and $D'_1 \subseteq D_1$. A truth assignment φ over D is 1-1 over D' if for all $i \in D'_0$ there is a unique $j \in D_1$ such that $P_{ij} = 1$ and for all $j \in D'_1$ there is a unique $i \in D_0$ such that $P_{ij} = 1$.

2.1. Random $1-1$ Restrictions. We will now define a probability space of partial 1-1 truth assignments on D, where $D = D_0 \cup D_1$, and $|D_0| = |D_1| = n$. The probability space $\mathcal{R}_p^D$ is the set of all triplets $\rho =< S_0, S_1, \pi >$, where $S_0 \subseteq D_0$, $S_1 \subseteq D_1$ and $|S_0| = |S_1|$. The set S_0 is chosen as follows. For each $x \in D_0$, choose $x \in S_0$ with probability p and $x \notin S_0$ with probability $1 - p$. After all elements, S_0, in D_0 have been selected, the set S_1 is obtained by selecting exactly $|S_0|$ elements of D_1 uniformly and at random. The third component in the triple, π, is a uniformly chosen bijection from $D_0 \setminus S_0$ to $D_1 \setminus S_1$. The triplet $< S_0, S_1, \pi >$ will sometimes be referred to as $< S, \pi >$, where $S = S_0 \cup S_1$.

Alternatively, the probability space $\mathcal{R}_p^D$ can be generated by a second experiment described here. First, select a random subset of D_0 of size k, and a random subset of D_1 of size k, where k is chosen according to the binomial distribution, $B(n, p)$. The chosen subsets will be S_0 and S_1, respectively. Then select a uniformly chosen bijection from $D_0 \setminus S_0$ to $D_1 \setminus S_1$.

Every $\rho =< S, \pi >$ in $\mathcal{R}_p^D$ determines a unique mapping of the variables over D to the set $\{0, 1, *\}$, as follows. If $i \in S_0$ and $j \in S_1$, then $r(P_{ij}) = *$; if $i \notin S_0$ and $j \notin S_1$ and $\pi(i) = j$, then $r(P_{ij}) = 1$; otherwise $r(P_{ij}) = 0$. We will call this mapping of the variables determined by ρ the *variable restriction* induced by ρ.

Conversely, every 1-1 variable restriction is generated by a unique $\rho \in \mathcal{R}_p^D$. Thus, the distribution $\mathcal{R}_p^D$ defines a probability distribution of 1-1 variable restrictions. If r is a random restriction obtained by choosing a random ρ according to $\mathcal{R}_p^D$, we will refer to both the restriction and the random partial 1-1 function by ρ.

A restriction, $\rho =< S_0, S_1, \pi >$, applied to the variables over the domain D creates a subdomain, $D' \subseteq D = S_0 \cup S_1$. Namely, the subdomain D' induced by ρ is the maximal subset such that the underlying variables of D' are set to $*$ by ρ. We will also refer to the subdomain, D', by $D\lceil_\rho$.

Given a boolean function F, and an element $\rho \in \mathcal{R}_p^D$, we will denote by $F\lceil_\rho$ the function that we obtain by doing the substitutions prescribed by ρ. $F\lceil_\rho$ will be a function of the variables which were given the value $*$ by ρ (i.e., the variables over $D\lceil_\rho$).

It will be convenient to describe the probability space of restrictions in terms of another, slightly different distribution. The probability space $\mathcal{P}_p^D$ is the set of all pairs $< S_0, S_1, \pi >$, where π is a randomly chosen bijection from D_0 into D_1, and S_0 and S_1 are subsets of D_0 and D_1, respectively. The set S_0 is chosen as follows. For each $x \in D_0$, choose $x \in S_0$ with probability p and $x \notin S_0$ with probability $1 - p$. The set S_1 is then taken to be $\pi(S_0)$.

Equivalently, we can generate the probability space $\mathcal{P}_p^D$ by first choosing a complete bijection, π, from D_0 to D_1, and then choosing a k-subset, S_0, of D_0 uniformly, at random where k is chosen according to the binomial distribution, $B(n,p)$. The set S_1 is then defined to be $\pi(S_0)$.

We will now describe a third and final experiment which generates the distribution $\mathcal{P}_p^D$. Initially, we define $S = D_0$ and $T = D_1$. Choose a random $x \in S$ and a random $y \in T$. Set $\pi(x) = y$. Then choose both $x \in S_0$ and $y \in S_1$ with probability p, or both $x \notin S_0$ and $y \notin S_1$ with probability $1 - p$. Repeat this procedure on the smaller sets $S = S \setminus x$, and $T = T \setminus y$, until $S = \emptyset$ and $T = \emptyset$. At the end of the procedure, we will have completely determined a bijection, π, as well as sets S_0 and S_1. It can be checked that the probability of choosing a particular $\rho =< S_0, S_1, \pi >$ is the same in all of the above experiments.

Each $\rho =< S_0, S_1, \pi >$ in $\mathcal{P}_p^D$ determines a unique restriction of the variables over D: if $i \in S_0$ and $j \in S_1$, then $r(P_{ij}) = *$; if $i \notin S_0$ and $j \notin S_1$ and $\pi(i) = j$, then $r(P_{ij}) = 1$; otherwise $r(P_{ij}) = 0$.

Note, however, that each 1-1 restriction of the underlying variables can be generated by many $\rho \in \mathcal{P}_p^D$. This can be seen by noting that there is a one-to-many mapping from $\rho \in \mathcal{R}_p^D$ to $\rho \in \mathcal{P}_p^D$. The following lemma states that the experiments $\mathcal{R}_p^D$ and $\mathcal{P}_p^D$ each define the same distribution of restrictions.

LEMMA 1. *The distributions $\mathcal{R}_p^D$ and $\mathcal{P}_p^D$ define the same probability distributions over 1-1 variable restrictions.*

PROOF. For each element $\rho =< S_0, S_1, \pi > \in \mathcal{R}_p^D$, there is an associated unique set of elements $\rho' =< S'_0, S'_1, \pi' >$ from $\mathcal{P}_p^D$, which yields the same assignment to the variables P_{ij}. Namely, an element $\rho' =< S'_0, S'_1, \pi' > \in \mathcal{P}_p^D$ is associated with $\rho =< S_0, S_1, \pi > \in \mathcal{R}_p^D$ if the bijection, π' on $D \setminus (S_0 \cup S_1)$ is identical to π and $S'_1 = S_1$, and $S'_0 = S'_1$. Each element of $\mathcal{R}_p^D$ is associated with the same number of elements from $\mathcal{P}_p^D$; further, the probability over $\mathcal{R}$ of choosing a particular element, ρ, is equal to the probability over $\mathcal{P}$ of choosing an element in the set associated with ρ. Thus, the induced probability distributions on the setting of the variables P_{ij} are identical. □

In what follows, a *restriction* denotes a particular element from one of the two distributions, $\mathcal{R}_p^D$, or $\mathcal{P}_p^D$, and a *variable restriction* denotes a particular assignment of 0,1,* to the underlying variables. As mentioned above, a restriction from either $\mathcal{R}_p^D$ or $\mathcal{P}_p^D$ induces a corresponding variable restriction.

2.2. Random Pigeonhole Restrictions. The distributions defined above assumed that $|D_0| = |D_1|$. However, in the case of the variables underlying the pigeonhole principle, the domain actually has one more element than the range. We will now define a probability space of partial 1-1 functions on D, where now $D = D_0 \cup D_1$, and $|D_0| = |D_1| + 1$, and $0 < p < 1$. The probability space $\mathcal{Q}_p^D$ is the set of all quadruples $\rho =< i, S_0, S_1, \pi >$, where $i \in D_0$, $S_0 \subseteq D_0 \setminus \{i\}$, $S_1 \subseteq D_1$ and $|S_0| = |S_1|$. First, $i \in D_0$ is chosen uniformly and at random. Then S_0, S_1, and π are chosen from $D' = D_0 \setminus \{i\} \cup D_1$ in the same manner as in distribution $\mathcal{R}_p^{D'}$.

Again, the notation $Pr_p^{\mathcal{Q}}[A]$ denotes the probability that A occurs when ρ is drawn from $\mathcal{Q}_p^D$. We will need the following simple observation.

LEMMA 2. *Let $D = D_0 \cup D_1$, where $|D_0| = |D_1| + 1$. For $\rho \in \mathcal{Q}_p^D$, let $spare(\rho) = x$ denote the event that the pigeon which is unmapped by ρ is x—i.e., if $\rho =< i, S_0, S_1, \pi >$, then $i = x$. Then the distribution of variable restrictions over $D' = D_0 \setminus \{x\} \cup D_1$ induced by $\mathcal{Q}_p^D$, given that $spare(\rho) = x$, is equal to the distribution of variable restrictions over D' induced by $\mathcal{R}_p^{D'}$.*

For a Boolean formula F and an element $\rho \in \mathcal{R}_p^D$ ($\rho \in \mathcal{Q}_p^D$), F restricted by ρ will be denoted by $F\lceil_\rho$.

2.3. Map Disjunctions. A *map over* D is defined to be a conjunction of the form $\bigwedge \Gamma$, where Γ is a set of variables over D such that distinct variables in Γ have distinct left subscripts and distinct right subscripts. Maps describe bijections between subsets of D_0 and subsets of D_1. The *size* of a map $\bigwedge \Gamma$ is $|\Gamma|$; if the size of a map is bounded by t, it is said to be a t-map. An OR of maps is called a *map disjunction*. The *mapsize* of a map disjunction is the size of the largest map in the disjunction; if all the maps are of size at most t, then it will be called a *t-disjunction*.

Let σ be a map over D, with underlying variables $\{P_{x_1y_1}, P_{x_2y_2}, \ldots, P_{x_ky_k}\}$, where $x_i \in D_0$, and $y_i \in D_1$ for all i, $1 \leq i \leq k$. Then σ can also be viewed as a 1-1 variable restriction; namely, the restriction that maps x_i to y_i for all $1 \leq i \leq k$. We will sometimes refer to σ interchangeably as a restriction and as a map.

If Y is a map or a set of variables, then $v(Y)$ denotes the elements in $D_0 \cup D_1$ that are indexed by the variables in Y. Alternatively, if we view D_0 and D_1 as disjoint sets of vertices, and Y as a set of edges, then $v(Y)$ denotes the subset of vertices which are incident upon the edges of Y.

Let $K \subseteq D = D_0 \cup D_1$. Then $Proj_D[K]$ is the set of all minimal partial 1-1 maps over D which involve all of the elements of K.

Where it is convenient, we shall assume that an ordering is given for each of D_0 and D_1. Whenever we write a real, positive number where an integer is required, we mean the integer part of the number (floor). When we assert an inequality involving n, we shall often assume tacitly that n is sufficiently large.

3. 1-1 Decision Trees and the Switching Lemma

Decision trees are a very natural and simple model of computation where a boolean function is computed by a binary tree, whose nodes are labelled with bits of the input, and whose leaves are labelled either "accept" or "reject". If a node is labelled with a particular input bit, x_i, then the two edges out of that node are labelled x_i and $\overline{x}_i$. To compute the function on an input, α, we start at the root of the binary tree, and follow the path whose edge labels are consistent with α until we hit a leaf node, at which point we output the label of that leaf. The decision tree complexity of a boolean function, f, is defined to be the minimum height of all decision trees which compute f.

It turns out that Håstad's Switching Lemma (see Håstad 1987) can actually be stated in a stronger form in terms of decision trees. That is, the proof of Håstad's Switching Lemma shows that if f is an OR of small-sized ANDs, and ρ is a random restriction, then with very high probability, $f\lceil_\rho$ can be represented by a small-depth decision tree. It follows that $f\lceil_\rho$ can be written both as an OR of small ANDs, and as an AND of small ORs. In order to prove our switching lemma for pigeonhole restrictions, we will first introduce a new class of decision trees, called *1-1 decision trees.* We will then prove an analog of Håstad's Switching Lemma (stated in terms of decision trees), for 1-1 decision trees.

A *1-1 decision tree* over domain $D = D_0 \cup D_1$ is defined as follows. It is a rooted tree where each interior node v is labelled by a query $i \in D_0$ or $j \in D_1$ and each edge is labelled by some pair $[i, j]$ where $i \in D_0$ and $j \in D_1$. Leaves are labelled with either "0" or "1". For each interior node v labelled by $i \in D_0$ ($j \in D_1$), there is exactly one out-edge labelled $[i, j]$ for each $j \in D_1$ ($i \in D_0$) that does not appear in any edge label on the path from the root to v. The label of an interior node v may not appear in any edge label on the path from the root to v. Thus the set of edge labels on any path defines a map: if $[i, j]$ labels an edge on the path, then P_{ij} appears in the map.

A 1-1 decision tree T over D *represents* a boolean function f over (the variables of) D if for all leaf nodes $v \in T$, if we let σ be the map defined by the path in T from the root to v then for all truth assignments α over D that are 1-1 on $v(\sigma)$ and consistent with σ, $f(\alpha)$ is equal to the label of v. For a boolean function f over domain D, we define $d_D(f)$ to be the minimum height of all 1-1 decision trees representing f. If f cannot be represented by any 1-1 decision tree, then $d_D(f) = \infty$.

If ρ is a partial 1-1 restriction over D and T is a 1-1 decision tree over D, then define $T\lceil_\rho$ to be the decision tree obtained from T by removing all paths which have a label that has been set to "0" by ρ, and contracting all edges whose labels are set to "1" by ρ. (The node resulting from the edge contraction has the label of the child.)

LEMMA 3. *Let f be a boolean function over D and let T be a 1-1 decision tree representing f over D. If ρ is a partial 1-1 restriction over D, then $T\lceil_\rho$ is a 1-1 decision tree which represents $f\lceil_\rho$ over $D\lceil_\rho$.*

PROOF. We will show that for all root-to-leaf paths $p \in T\lceil_\rho$ with leaf label l_p, if σ is the map defined by p, then for all truth assignments α over $D\lceil_\rho$ consistent with σ, $f\lceil_\rho(\alpha) = l_p$. Fix a path $p \in T\lceil_\rho$ with leaf label l_p, and let σ be the map defined by p. Then by the definition of $T\lceil_\rho$, there exists a map σ' such that $\sigma'\lceil_\rho = \sigma$ and σ' corresponds to a path p' in T with leaf label l_p. Because T represents f, it follows that for all truth assignments α' over D which extend σ', $f(\alpha') = l_p$. Now fix a truth assignment α over $D\lceil_\rho$ consistent with σ. Then let α' be the truth assignment $\alpha\rho$ over D. Then because α' extends σ', we have $f\lceil_\rho (\alpha) = f(\alpha') = l_p$, as desired. □

For any decision tree T let T' denote the tree obtained from T by switching the 1's and 0's labelling the leaves of T. Note that if T represents f over D then T' represents $\neg f$. Given a 1-1 decision tree T over D of height d, we define a d-disjunction, $maps(T)$, over D as follows. The maps in $maps(T)$ correspond to the paths in T that end in a leaf labelled 1. For a particular path in T, the corresponding map in $maps(T)$ is defined to be the conjunction of the edge labels along the path. Notice that T represents $maps(T)$. Furthermore note that for any partial 1-1 restriction ρ over D, $maps(T\lceil_\rho) = maps(T)\lceil_\rho$. The lemma in the next section is actually a switching lemma in the sense of Håstad. That is, it will allow us to obtain a map disjunction that approximates the negation of f. This is obtained by representing f by a 1-1 decision tree, T, and then taking $maps(T')$ to be the map disjunction approximating the negation of f, where T' is the tree obtained from T by switching the 1's and 0's labelling the leaves.

We define the *complete 1-1 tree* for $K \subseteq D$ over D inductively as follows. If K consists of a single node $k \in D_0$ ($k \in D_1$), then label the root "$k \in D_0$" ("$k \in D_1$"), and create n (or $n+1$) edges adjacent to the root, labelled by $[k, j]$, for all $j \in D_1$ ($[j, k]$ for all $j \in D_0$). Otherwise, $K = K' \cup \{k\} \subseteq D$. Assume that we have created the complete tree for K'; we will now extend it to a complete tree for K. This is done by extending each leaf node v_l as follows. Let p_l be the path from the root to v_l. The edge labellings along p_l define a partial 1-1 map involving all elements of K'. If this partial map does not include k, then label v_l by k, and add new edges leading out of v_l, one for every possible mapping for k that results in a 1-1 map extending the partial 1-1 map along p_l. Otherwise, if k is involved in the partial 1-1 map, leave v_l unlabelled. Note that each path of the complete tree over K will be labelled by some $\sigma \in Proj_D[K]$.

LEMMA 4. *Let f be a boolean function over the variables P_{ij}, $i \in D_0$, $j \in D_1$, where $|D_0| = |D_1|$. For every $K \subseteq D_0 \cup D_1$, there exists a restriction $\sigma \in Proj_D[K]$ such that $d_D(f) \leq |\sigma| + d_{D\lceil_\sigma}(f\lceil_\sigma)$.*

PROOF. The proof is very similar to that of Beame and Håstad (1989). Fix $K \subseteq D$. We start with the complete 1-1 tree for K. As noted above, the paths of this tree correspond exactly to elements of $Proj_D[K]$. Let v_σ be the leaf node corresponding to the path labelled by $\sigma \in Proj_D[K]$. For each σ, we replace the leaf node v_σ by a subtree that is a 1-1 decision tree for $f\lceil_\sigma$ over

$D\lceil_\sigma$. The resulting tree clearly represents f over D, and has depth at most $\max_\sigma\{|\sigma| + d_{D\lceil_\sigma}(f\lceil_\sigma)\}$. □

If f is a map disjunction defined over a set D and ρ is a restriction on D then we will use the notation $\delta(f\lceil_\rho)$ for $d_{D\lceil_\rho}(f\lceil_\rho)$.

Let $D = D_0 \cup D_1$, $|D_0| = n+1$ and $|D_1| = n$. We will now state the main combinatorial lemma. This lemma states that with extremely high probability, after applying a random restriction to a t-disjunction, we can represent the resulting formula by a small-depth 1-1 decision tree.

LEMMA 5. (THE PIGEONHOLE SWITCHING LEMMA) *Let f be an r-disjunction over D. Choose ρ at random from Q_p^D. For $s \geq 0$, $p \leq 1/36$, and $pn \geq 8(s+r)^2$ we have*

$$Pr[\delta(f\lceil_\rho) \geq s+1] < n\alpha^s,$$

for any $\alpha > 0$ satisfying $(1 + \frac{36p^4n^3}{\alpha^2})^r \leq 2$.

Fact: The inequality $(1 + 36p^4n^3/\alpha^2)^r \leq 2$ holds when $\alpha = 8p^2n^{3/2}r^{1/2}$. This fact can be shown by taking logarithms of both sides, and using the inequality $\log(1+x) \leq x$.

We will postpone the proof of the pigeonhole switching lemma until the last section of this paper.

4. Critical Truth Assignments and Approximate Negation

For the pigeonhole variables, P_{ij}, $i \in D_0$, $j \in D_1$, where $size(D_0) = size(D_1)+1$, we will consider the class of truth assignments which are *maximally* 1-1. The set of *critical truth assignments* over D, CTA_D, is defined to be the class of all truth assignments over D which are 1-1 on all but one element of D_0: $CTA_D = \{\alpha \mid \exists x \in D_0$ such that α is 1-1 on $D_0 \setminus \{x\} \cup D_1$, and $\forall j \in D_1\ P_{xj} = 0\}$. Given a map disjunction f over the pigeonhole variables, we want to apply the above switching lemma in order to obtain a new map disjunction which approximates $\neg f$.

LEMMA 6. *Let $D = D_0 \cup D_1$ where $|D_0| = n+1$, $|D_1| = n$, and let T be a 1-1 decision tree of height k defined over the set D. The fraction of all critical truth assignments α over D that are consistent with some path in T is at least $1 - \frac{k}{|D_0|}$.*

PROOF. We prove this lemma by induction on the height k of T. Consider a randomly chosen critical truth assignment α over D.

If $k = 0$ then T is just a single node and the lemma is vacuously true.

Now suppose that the lemma is true for all trees of height at most k and suppose that T has height $k + 1$.

If the root of T is labelled by some $j \in D_1$ then α matches j with a unique $i \in D_0$. Let σ be the map consisting of P_{ij}. Then $T\lceil_\sigma$ is a 1-1 decision tree of height at most k defined over $D\lceil_\sigma$. Furthermore, the probability that α is consistent with some path in T is equal to the probability that it is consistent with some path in $T\lceil_\sigma$. By the induction hypothesis this is at least $1 - k/n \geq 1 - (k + 1)/(n + 1)$ as required.

If the root of T is labelled by some $i \in D_0$ then either $i = spare(\alpha)$, or $spare(\alpha) \neq i$ and α matches i with a unique $j \in D_0$. Let E be the event that α is not consistent with any path in T. Thus we have

$$\begin{aligned} Pr[E] \leq \; & Pr[spare(\alpha) = i] + \\ & Pr[spare(\alpha) \neq i] \times Pr[E \mid spare(\alpha) \neq i]. \end{aligned}$$

Since the induced distribution on $spare(\alpha)$ is uniform over D_0, $Pr[spare(\alpha) = i] = 1/(n + 1)$. Given that $spare(\alpha) \neq i$ we can argue, as in the case that the label was $j \in D_1$, that the probability of E is at most k/n. Thus we get a total probability of E of

$$\frac{1}{n+1} + \left(1 - \frac{1}{n+1}\right)\frac{k}{n} = \frac{1}{n+1} + \frac{k}{n+1} = \frac{k+1}{n+1}$$

as required. □

COROLLARY 7. *Let $D = D_0 \cup D_1$ where $|D_0| = n + 1$, $|D_1| = n$, and let T be a 1-1 decision tree of height k representing f over the set D; let T' be the 1-1 decision tree obtained from T by switching the 1's and 0's labelling the leaves of T. Then $maps(T')$ and $\neg f$ agree on at least a $(1 - \frac{k}{n+1})$ fraction of all critical truth assignments over D.*

5. Exponential Lower Bounds

5.1. Definitions. For concreteness, we will work with propositional proofs in a particular system, H, to be defined below. The crucial property of H that we will exploit is that each rule and axiom involves at most one negation. In

Excluded Middle Axiom: $A \vee \neg A$

Weakening Rule: $\frac{A}{(A \vee B)}$

Cut Rule: $\frac{(A \vee B),\ (\neg A \vee C)}{(B \vee C)}$

Merging Rule: $\frac{\bigvee(\{\bigvee \Gamma\} \cup \Delta)}{\bigvee(\Gamma \cup \Delta)}$

Unmerging Rule: $\frac{\bigvee(\Gamma \cup \Delta)}{\bigvee(\{\bigvee \Gamma\} \cup \Delta)}$

Figure 1: Rules of the system H

any Frege system, each rule and axiom involves at most a constant number of negations, and this also suffices to prove the lower bound.

The Frege refutation system that we will use is the system H described by Bellantoni, Pitassi and Urquhart (1992). H is slightly nonstandard in that it is formulated as a propositional proof system for unbounded fan-in formulas. More precisely, the formulas of H are unordered rooted trees defined inductively by the rules: (1) if x is a variable, then $\bigvee\{\bigwedge x\}$ is a formula; (2) if A is a formula then $\neg A$ is a formula; and (3) if Γ is a finite set of formulas, then $\bigvee \Gamma$ is a formula. Thus the system allows $\wedge$ only at the bottom level, and in fact requires $\wedge$'s there. This syntactic requirement simplifies the exposition. The rules of H are listed in figure 1. We will sometimes use the notation $A \vee B$ to mean $\vee\{A, B\}$.

Note that the system H is not complete in the usual sense because there are no rules for AND. However, H is complete for formulas over the basis OR and NOT.

If we begin with a bounded-depth Frege proof over the (unbounded fan-in) boolean basis AND, OR, and NOT, in a different Frege system than the one specified above, then we can convert the proof into a bounded-depth proof in H by using the ideas in the simulation result of Cook and Reckhow (1979). We would like to point out that although the simulation preserves constant-depth, the depth after the conversion may increase by a constant factor. Thus, the actual constants in our lower bound are sensitive to the particular Frege system that one starts with.

In this paper, a depth d formula will be an unbounded fan-in boolean tree, consisting of $d-2$ levels of unbounded fan-in OR and NOT gates, followed by the bottom two levels which are map-disjunctions. Note that any depth d formula in a proof in H consists of $d-2$ levels of unbounded fan-in OR and NOT gates, followed by two levels of 1-disjunctions. The *size* of a formula, is the number of occurrences of $\vee$ and $\neg$ in the formula; the size of a Frege proof (or a sequence of formulas) is the sum of the sizes of the formulas occurring in the proof (or sequence). The depth of a Frege proof (or a sequence of formulas) is the maximum depth of the formulas in the proof (or sequence). A Frege refutation of $A_1 \wedge A_2 \wedge ... \wedge A_k$ can be viewed as a directed acyclic graph, where each node in the graph is a formula of the proof. The leaves of the graph are the formulas A_i, the root of the graph is the empty (false) formula, and two formulas, A and B are parents of another formula C if C follows by some inference rule from A and B. A Frege refutation has *height h* if the directed acyclic graph which describes the proof has height no greater than h.

In our lower bound, we will begin with a proof of depth d in the system H. We will then apply a transformation to the original depth d proof to obtain a new sequence of formulas, where now each formula will have depth at most $d-1$. Because each formula of the original proof has been modified, the resulting sequence of formulas will not have the syntactic form required by the rules of H; thus, the new sequence of formulas will not be a *syntactically proper* proof in H. On the other hand, the resulting sequence of formulas will still approximate a proof, in the *semantic* sense. More precisely, we will maintain the property that each formula in the new sequence will still follow from the same previous formulas by a γ-*sound* inference. An inference $(f_1, f_2) \rightarrow f$ over D is γ-sound if there exists a subset, S, of all critical truth assignments, CTA_D, of size at least $\gamma|CTA_D|$, such that for each assignment $s \in S$, if s makes both f_1 and f_2 true, then s also makes f true. (Note that an axiom can be viewed as a rule with one precedent, the "true" formula.) This notion of soundness is more general than the usual notion of soundness. In particular, all inferences in the original proof in H are 1-sound. At the other extreme, an example of a completely unsound (0-sound) inference is: $[(1,1) \rightarrow 0]$.

5.2. Reducing the Depth. In this section we show how a proof of depth d is converted into one of depth $d-1$ while preserving approximate soundness. Let P be a sequence of formulas over D, $|D_0| = n+1$, $|D_1| = n$, each of depth at most d ($d > 2$) and let $\rho \in \mathcal{Q}_p^D$. Suppose that ρ leaves exactly those variables in $D' \subseteq D$ unset, where $|D'_0| = n'+1$, $|D'_1| = n'$. P is converted into a sequence of depth $d-1$ formulas over D' in the following three steps.

(1) Apply ρ to each formula of P, obtaining $P\lceil_\rho$.

(2) Let $G_0 ... G_m$ be the distinct map disjunctions appearing in formulas of $P\lceil_\rho$. Represent each G_i by some 1-1 decision tree T_i over D'. Let T_i' denote the 1-1 decision tree representing $\neg G_i$, obtained from T_i by switching the 0's and the 1's at the leaves. Define the *pseudo-complement* of G_i, $c_{D'}(G_i) = maps(T_i')$. Replace each occurrence of $\neg G_i$ by $c_{D'}(G_i)$, uniformly throughout $P\lceil_\rho$.

(3) For each formula of $P\lceil_\rho$, merge together OR gates appearing at heights 2 and 3.

When P' is obtained by applying the conversion process to P with ρ, we say that P' is P converted by ρ. If f is a formula of P of depth d, then f converted by ρ (in P') will have depth $d-1$. This is because step (2) ensures that all gates at levels 2 and 3 will be OR gates, and step (3) merges these two adjacent levels of OR gates.

DEFINITION 8. *A sequence of formulas over $D = D_0 \cup D_1$ is a (n, d, t, γ, S)-approximate refutation if: $|D_0| = n+1$, $|D_1| = n$, each formula has depth at most d, each map disjunction has mapsize at most t, the total size of all formulas in the proof is at most S, each inference is γ-sound, and the proof was obtained from a (1-sound) proof in H of the pigeonhole principle over a larger universe, by applying the above conversion process (to the sequence of formulas in the proof) a finite number of times.*

The following lemma shows that if we choose the right restrictions, then successive applications of the above conversion process result in an approximately sound refutation. The main idea behind the proof of this lemma is that while each formula may not be approximated well at all (since every negation is approximated, and there may be many negations in each formula), each inference will still remain approximately sound because each rule and axiom of H involves at most one negation.

LEMMA 9. (CONVERSION LEMMA) *Let P^0 be a refutation in H of PHP_n over D, of depth d and size S. Let $k+1 \leq d-2$. Let $\rho = \rho^0, \rho^1, \rho^2, ..., \rho^k$ be a sequence of restrictions such that D^{i+1} is the set of unset variables of ρ^i, $D^{k+1} \subseteq D^k \subseteq ... \subseteq D^1 \subseteq D$. Also, let $|D_1^i| = n_i$, and $|D_0^i| = n_i + 1$. Let $P^1, P^2, ..., P^{k+1}$ be a sequence of approximate proofs where P^{i+1} is equal to P^i converted by ρ^i. Suppose also that for every $i \leq k$, every map disjunction in $P^i\lceil_{\rho^i}$ is represented by a 1-1 decision tree over D^{i+1} of depth at most*

t. Let $\gamma_i = 1 - \frac{t}{n_i+1}$. If for all i, $1 \leq i \leq k$, P^i is a refutation which is $(n_i, d-i, t, \gamma_i, S)$-approximate, then P^{k+1} is a refutation of $PHP_{n_{k+1}}$ which is $(n_{i+1}, d-(k+1), t, \gamma_{k+1}, S)$-approximate.

PROOF. The conversion process, applied to any sequence of formulas of depth d, yields a new sequence of formulas of depth $d-1$ and size at most S. Applying the conversion process $k+1$ times to a refutation in H thus yields an approximate proof of depth $d-(k+1)$ and size at most S. Because ρ leaves exactly those variables in D^{k+1} unset, where $|D_0^{k+1}| = n_{k+1}+1$ and $|D_1^{k+1}| = n_{k+1}$, it follows that P^{k+1} is an approximate proof of $PHP_{n_{k+1}}$ over D^{k+1}. Also, by assumption, for every map disjunction G in $P^k\lceil_{\rho^k}$, G is represented by a 1-1 decision tree over D^{k+1} of depth at most t, and therefore, $mapsize[c_{D^{k+1}}(G)] \leq t$. Thus, the conversion process results in an approximate proof P^{k+1} where every map disjunction has mapsize at most t.

It is left to show that every inference in P^{k+1} is γ_{k+1}-sound. Fix a particular formula f^0 in P^0. Let f^i be the formula which results from f^0 after i conversion steps – f^i is the corresponding formula in P^i. We want to show that f^{k+1} follows from a γ_{k+1}-sound inference. There are five cases to consider: either f^0 is an application of the Excluded Middle Axiom, or f^0 follows from the Cut Rule, or f^0 follows from either the Weakening, Merging or Unmerging Rule. If f^0 follows from either the Weakening, Merging or Unmerging Rule, then because these rules do not involve any negations, it is easy to see that the corresponding inference in P^{k+1} is 1-sound. Now assume that f^0 follows from the cut rule; the other case (when f^0 follows from the Excluded Middle Axiom) is handled similarly. Let $f^0 = B \vee C$, where f^0 follows from $g^0 = A \vee B$ and $h^0 = \neg A \vee C$. Then for all proofs P^i, $1 \leq i \leq k$, g^i and h^i are the two formulas in P^i which approximately imply f^i. The inference $(g^k, h^k) \rightarrow f^k$ has one of two forms, depending on the depth of g^k and h^k.

(1) If the inference has the form $(A' \vee B'), (\neg A' \vee C') \rightarrow (B' \vee C')$ then there are two cases to consider. If A' has depth greater than 2, then the new inference will have the same form since the negation in front of A' will not yet be converted; hence the new inference will be 1-sound. On the other hand, if A' is a map disjunction, then $\neg A'\lceil_{\rho^k}$ will be replaced by $c_{D^{k+1}}(A'\lceil_{\rho^k})$. Because $A'\lceil_{\rho^k}$ is represented by a depth-t 1-1 decision tree T, by Corollary 7 we know that $c_{D^{k+1}}(A'\lceil_{\rho^k}) = maps(T')$ will equal $\neg A'\lceil_{\rho^k}$ for at least $1 - \frac{t}{n_{k+1}+1}$ of the critical truth assignments over D^{k+1} Hence this inference will be γ_{k+1}-sound.

(2) Otherwise, some previous f^i, $i < k+1$, which follows from g^i and h^i, has the form $(A' \vee B'), (c_{D^i}(A') \vee C') \rightarrow (B' \vee C')$. Let $\rho' = \rho^i\rho^{i+1}..\rho^k$. The

inference $(g^{k+1}, h^{k+1}) \to f^{k+1}$ thus has the form $(A'\upharpoonright_{\rho'} \vee B'\upharpoonright_{\rho'}), (c_{D^i}(A')\upharpoonright_{\rho'} \vee C'\upharpoonright_{\rho'}) \to (B'\upharpoonright_{\rho'} \vee C'\upharpoonright_{\rho'})$. By assumption, we know that A' was represented by a 1-1 decision tree T of depth at most t, and $c_{D^i}(A') = maps(T')$. Thus, by Lemma 3, $T\upharpoonright_{\rho'}$ is a 1-1 decision tree over D^{k+1} which represents $A'\upharpoonright_{\rho'}$. Also, $c_{D^i}(A')\upharpoonright_{\rho'}$ is equal to $maps(T'\upharpoonright_{\rho'})$. Therefore, by Corollary 7, this implies that $c_{D^i}(A')\upharpoonright_{\rho'}$ equals $\neg A'\upharpoonright_{\rho'}$ for at least a fraction $1 - \frac{t}{n_{k+1}+1}$ of the critical truth assignments over D^{k+1}, and hence the new inference will be γ_{k+1}-sound.

5.3. The Lower Bound.

THEOREM 10. (LOWER BOUND ON SIZE) *There exists a constant c such that for sufficiently large n, any Frege refutation of PHP_n of depth d must have size at least* $\exp\left[\Omega\left(n^{6^{-(d+c)}}\right)\right]$.

COROLLARY 11. (LOWER BOUND ON DEPTH) *For sufficiently large n, any Frege refutation of PHP_n of polynomial-size must have depth at least* $\Omega(\log\log n)$.

Theorem 10 will be proven by induction on d, the depth of the Frege refutation. To facilitate the proof of the base case, we would like to restrict attention to Frege proofs that are in *tree form*. A Frege refutation, P, is *tree-like* if the refutation, when viewed as a directed, acyclic graph, is a tree. In other words, each intermediate formula is used no more than once. The following lemma originally due to Krajíček (1991) and later improved by Bonet and Buss (1992) states that any Frege refutation can be efficiently converted into an equivalent, tree-like refutation.

LEMMA 12. (TREE LEMMA) *Any Frege refutation of size S and depth d can be transformed into another tree-like Frege refutation of size $O(S^2)$ and depth $d + o(1)$.*

By the above Tree Lemma, Theorem 10 is a corollary to the following theorem.

THEOREM 13. (LOWER BOUND FOR TREE-LIKE FREGE REFUTATIONS) *For sufficiently large n, any tree-like Frege refutation of PHP_n of depth d must have size at least* $S = \exp\left(n^{6^{-(d+1)}}\right)$.

PROOF. The proof is by induction on d. Suppose that there is such a refutation, P, of PHP_n in our system H, of size S, depth d, and height at most $\log S$. Recall that each formula in any proof in H consists of $d-2$ levels of ORs and NOTs, followed by the 2 bottom levels, which are 1-disjunctions. Let $t = \frac{1}{4}\log(Sn)$. Define $\lambda(n) = n^{1/6}$. If λ^i is the i-fold composition of λ with itself, then $\lambda^i(n) = n^{6^{-i}}$. If $S < \exp\left(n^{6^{-(d+1)}}\right)$, then for sufficiently large n, $t < 1/4\lambda^d(n)$.

Because the system H is sound, and each map disjunction has mapsize 1, P is a refutation which is (n_0, d, t, γ_0, S)-approximate, where $n_0 = n$ and $\gamma_0 = \left(1 - \frac{t}{n_0+1}\right)$. Applying the Induction and Base lemmas below, we show that that for sufficiently large n_0, that there is no approximate proof of PHP_{n_0} which is (n_0, d, t, γ_0, S)-approximate. □

Suppose that $n_i \geq \lambda(n_{i-1}) \geq \lambda^2(n_{i-2}) \geq \ldots \geq \lambda^i(n_0)$, for all $0 \leq i \leq d-2$. Let $p_i = \lambda(n_i)/n_i$ and $\gamma_i = \left(1 - \frac{t}{n_i+1}\right)$ for all such i.

LEMMA 14. (INDUCTION LEMMA) *Let P^0 be a refutation in H of PHP_{n_0} over D^0 of depth d and size S. Let $\rho^0, \rho^1, \ldots, \rho^{i-1}$ be a sequence of restrictions, $i \leq d-3$, such that for $0 \leq k \leq i-1$, ρ^k leaves all variables over D^{k+1} unset, $|D_0^k| = n_k + 1$, $|D_1^k| = n_k$. Let $P^1, \ldots, P^i$ be a sequence of approximate proofs where for $1 \leq k \leq i-1$, P^k is equal to P^{k-1} converted by ρ^{k-1}; furthermore, P^k is $(n_k, d-k, t, \gamma_i, S)$-approximate, where t, γ_k, and n_k are as above. Then there is a restriction ρ^i such that P^i converted by ρ^i is an approximate refutation of $PHP_{n_{i+1}}$ which is $(n_{i+1}, d-(i+1), t, \gamma_{i+1}, S)$-approximate, where t, γ_{i+1} and n_{i+1} are as above.*

PROOF. Let D be the domain of the formulas in P^i. Since $t \leq \frac{1}{4}\lambda^d(n)$, and $p_i n_i \geq \lambda^{d-1}(n)$ for all $i \leq d-2$, we have $8(t+t)^2 \leq 8(\frac{1}{2}\lambda^d(n))^2 \leq \lambda^{d-1}(n) \leq p_i n_i$. Therefore, we can apply the Pigeonhole Switching Lemma for ρ drawn at random from $Q_{p_i}^D$ to each distinct map disjunction in P^i. For each map disjunction f in P^i, for a randomly chosen $\rho \in Q_{p_i}^D$, the probability that $f\lceil_\rho$ cannot be represented by a 1-1 decision tree over D^{i+1} of depth at most t is at most $n\alpha^{t-1}$ $(n = n_0 \geq n_i)$. Because the size of P^i is at most S, there are at most S map disjunctions in P^i, and therefore, for a randomly chosen ρ, the probability that some map disjunction $f\lceil_\rho$ in P^i cannot be represented by a 1-1 decision tree over D^{i+1} is at most $Sn\alpha^{t-1}$, where $0 < \alpha < 8p_i^2 n_i^{3/2} t^{1/2}$. Since $p_i = \lambda(n_i)/n_i$,

$$\alpha < \frac{8p_i^2 n_i^2 t^{1/2}}{n_i^{1/2}} = \frac{8(\lambda(n_i))^2 t^{1/2}}{n_i^{1/2}} = \frac{8n_i^{1/3} t^{1/2}}{n_i^{1/2}} \leq \frac{1}{2^9}.$$

It follows that since $t = \frac{1}{4}\log(Sn)$, $Sn\alpha^{t-1} \leq Sn\alpha^{t/2} = Sn(\frac{1}{2})^{9/8\log(Sn)}$, which is no greater than $1/6$.

The expected number of stars after applying the restriction ρ is $n_i p_i = \lambda(n_i)$. Since the number of stars is binomially distributed, for sufficiently large n_0, a random ρ leaves at least the expected number of stars with probability greater than $1/3$. (See, for example, Beame and Håstad 1989, Lemma 4.1.) Thus, there exists a restriction, ρ, leaving n_{i+1} stars, $n_{i+1} \geq \lambda(n_i)$, such that every map disjunction in $P^i\lceil_\rho$ is represented by a 1-1 decision tree over D^{i+1} of depth at most t. Therefore, we can apply the Conversion Lemma which states that P^i converted by ρ results in a new proof P^{i+1} of $PHP_{n_{i+1}}$, which is $(n_{i+1}, d-(i+1), t, \gamma_{i+1}, S)$-approximate. □

LEMMA 15. (BASE CASE LEMMA) *P_{d-1} cannot be an $(n_{d-1}, 1, t_{d-1}, \gamma_{d-1}, S)$-approximate refutation of $\neg PHP_{n_{d-1}}$, where $\gamma = 1 - \frac{t_{d-1}}{n_{d-1}+1}$, $n_{d-1} = n^{6^{-(d-1)}}$, and $t_{d-1} = (\log S)^{d-1}$.*

We will need the following proposition in order to prove the Base Case Lemma.

PROPOSITION 16. *After applying the conversion procedure, the formula, $\neg PHP_n$, gets converted to* 1.

PROOF. For any restriction $\rho \in Q_p^{D^*}$ where ρ leaves the variables underlying $D \subseteq D^*$ unset, $|D_1| = n$, when we apply ρ to the formula $\neg PHP_{n^*}$, we obtain a new sequence of clauses, $\neg PHP_n$, where $\neg PHP_n$ is the negated pigeonhole principle over D. We will now show that when we replace each negation in $\neg PHP_n$ by the "pseudocomplement", we obtain the formula "1". PHP_n consists of the disjunction of the following formulas: $\neg(P_{i1} \vee P_{i2} \vee \ldots \vee P_{in})$, $i \leq n+1$, and $\neg(\neg P_{ik} \vee \neg P_{jk})$, $i, j \leq n+1, i \neq j, k \leq n$. Let f be a map disjunction and let T_f be a decision tree representing f, obtained as in the proof of the switching lemma. That is, we query all variables in the first term, f_1, and then proceed inductively on $f\lceil_\sigma$, where $\sigma \in Proj_D(v(f_1))$. Then $\neg f$, will be converted to the pseudocomplement of f, $maps(T_f^c)$. The decision tree representing the map disjunction $f = (P_{i1} \vee P_{i2} \vee \ldots \vee P_{in})$ has 1's labelling each leaf. Therefore, for all i, the formula $\neg(P_{i1} \vee \ldots \vee P_{in})$ is converted to 0. Similarly, it can be shown that $\neg(\neg P_{ik} \vee \neg P_{jk})$ is converted to 0. Therefore, $\neg PHP_n$ is converted to 1. □

PROOF. [Proof of Base Case Lemma] Recall that a γ_{d-1}-sound proof of $\neg PHP_{n_{d-1}}$ has the property that each inference is sound with respect to at least the fraction γ_{d-1} of the total number of critical truth assignments. The idea is to hit the proof with another restriction of size no greater than $O(t_{d-1} \log S)$ to obtain an approximate refutation of the pigeonhole principle on a smaller universe of size $2n' + 1$, with an inference of the form $[(1,1) \rightarrow 0]$. Such an inference is 0-sound. But this contradicts the fact that a γ_{d-1}-sound proof of $PHP_{n_{d-1}}$, when hit by a small restriction leaving a universe of size $2n' + 1$, should yield a $(1 - \frac{t_{d-1}}{n'+1})$-sound proof of $PHP_{n'}$.

Let T be a subtree of the original proof tree, and let $|T|$ denote the number of formulas (nodes) in T. The following lemma will allow us to find the small restriction which will force a 0-sound inference.

LEMMA 17. *Let P be a tree-like refutation of $\neg PHP_n$ of size S. Let T be an approximate proof over $D = D_0 \cup D_1$, $|D_0| = n + 1$, obtained by applying the inductive argument $d - 1$ times to P. (I.e., T is a γ-approximate proof of size $S' \leq S$, where the root formula is 0, all leaf formulas are not 0, all formulas are depth-t decision trees, and $\gamma = 1 - \frac{t}{n+1}$.) Then there exists a restriction of size $O(t \log |T|)$ such that $T\lceil_\rho$ has an inference $(1,1) \rightarrow 0$, and $|T|$ is the number of formulas in T.*

PROOF. The proof is by induction on $|T|$, the number of formulas in T. The base case is when $|T| = 3$. In this case, T consists of two leaf formulas, l_1 and l_2 which are t-disjunctions not equal to 0, and one root formula which is 0. Since each leaf formula is a t-disjunction, we can force one of them, say l_1, to 1 by setting t variables. It is left to argue that the other leaf formula, l_2, is not forced to 0, and hence by setting an additional t variables, both leaf formulas can be forced to 1. The formula l_2 is either a converted excluded middle axiom, or the converted formula $\neg PHP$. By the above proposition, $\neg PHP$ is 1, and therefore setting t variables will not force $\neg PHP$ to 0; similarly by the Conversion Lemma, setting t variables will not force a converted excluded middle axiom to 0. Therefore l_2 cannot be forced to 0 by setting t additional variables and hence we can force both l_1 and l_2 to 1 by setting at most $2t$ variables.

Now assume the inductive hypothesis for all T, $|T| \leq S' - 1$. Let T be an approximate proof in tree form, of size S', and satisfying the above properties. Because the proof is in tree form, there exists a partition of T into two subtrees, T_L and T_R, such that number of formulas (nodes) of both T_L and T_R are between $S'/3$ and $2S'/3$. Assume without loss of generality that the root formula of T_R is also the root formula of T, and let let r_L be the root formula of T_L. If r_L is 0, then T_R is a subtree with root formula 0, and leaf formulas which are not 0, so we can continue inductively on the subtree T_R. If r_L is 1, then T_L is a subtree with root formula 0, and leaf formulas which are not 0, so we can continue inductively on the subtree T_L. The last case is when r_L is a t-disjunction, which is not 0 or 1. In this case, we can force r_L to 1 by setting t variables; call this restriction ρ. The tree T_R will then be a subtree with root formula 0. Further, the leaf formulas are not 0 because setting t variables cannot force any converted excluded middle axiom, or the converted formula $\neg PHP$ to zero (by the Conversion Lemma, and the above proposition). Also by the Conversion Lemma, the new proof, $T\lceil_\rho$ is γ-sound, where $\gamma = 1 - \frac{t}{n'+1}$, $n' = n - t$, and the new proof is over a universe of size $2n' + 1$. We can therefore apply the inductive hypothesis which states that there exists a setting of at most $t\log(2S'/3)$ variables which forces an inference $(1, 1 \to 0)$ in $T_R\lceil_\rho$. Combining these two restrictions, we have forced a completely unsound inference by setting $t + t\log(2S'/3) \leq O(t\log S)$ variables, and the proof is complete. □

Since each leaf formula in P_{d-1} is either a converted instance of an excluded middle axiom or the converted pigeonhole formula, by the Conversion Lemma and the above proposition, each leaf formula of P_{d-1} is not 0. Also, because the original root formula is 0, the root formula of P_{d-1} is also 0. Therefore we can apply the above lemma which states that we can force an inference $1, 1 \to 0$ by setting at most $O(t\log S)$ variables. By the Conversion Lemma, we should now have an approximate refutation of $\neg PHP_{n'}$, where $n' = n - O(t\log S) \geq \frac{n}{2}$, which is $(1 - \frac{t}{n'+1})$-sound. Because $\frac{t}{n'+1} \leq 1/2$, we know that each inference in the approximate refutation is greater than 1/2-sound, and hence we have reached a contradiction. □

6. Proof of the Switching Lemma

In this section, we will prove the following Switching Lemma.

LEMMA 18. (SWITCHING LEMMA) *Let f be an r-disjunction over $D = D_0 \cup D_1$, $|D_0| = |D_1| = n$. Choose ρ at random from $\mathcal{R}_p^D$. For $s \geq 0$, $p \leq 1/36$, and $pn \geq 8(s+r)^2$ we have*

$$Pr[\delta(f\lceil_\rho) \geq s] < \alpha^s,$$

for any $\alpha > 0$ satisfying $(1 + 36p^4n^3/\alpha^2)^r \leq 2$.

Recall that the Pigeonhole Switching Lemma that is needed for the exponential lower bound (Lemma 5) is slightly different from the above Switching Lemma. Namely, in the above lemma, $|D_0| = |D_1| = m$, whereas in the Pigeonhole Switching Lemma, $|D_0| = n + 1$, and $|D_1| = n$. We obtain the Pigeonhole

Switching Lemma from the above lemma as an easy corollary; this will be proven at the end of this section.

Before starting the proof of the Switching Lemma, we will first need to understand our distribution of restrictions, given particular conditions. We describe these conditional distributions first, and then proceed with the proof.

Conditional Distributions

Let $\rho =< S_0, S_1, \pi >$ be a random 1-1 restriction of the underlying variables, and let Y be a map over D. Let Y_0 denote $v(Y) \cap D_0$, and let Y_1 denote $v(Y) \cap D_1$. Then $\rho(Y) = *$ denotes the event that all variables over $v(Y)$ are set to * by ρ, or equivalently, that $v(Y) \subseteq S$. In the proof of the Switching Lemma below, it will be necessary to understand the distribution of variable restrictions induced by $\mathcal{P}_p^D$ (or equivalently, by $\mathcal{R}_p^D$), given that $\rho(Y) = *$. In this section, it will be more convenient to work in the distribution $\mathcal{P}$.

Before examining that distribution, we will first define a new probability space. Let $\mathcal{T}_{p,i}^D$ be the set of all triplets $< S_0, S_1, \pi >$ generated as follows. First, choose i disjoint pairs, $< x, \pi(x) >$, where $x \in D_0$ and $\pi(x) \in D_1$. Let the i chosen elements of D_0 be S_0', and let the chosen elements of D_1 be S_1'. Add S_0' to S_0, and S_1' to S_1. Set these subsets aside, and let the remaining elements of D be D'. ($D_0' = D_0 \setminus S_0'$, and $D_1' = D_1 \setminus S_1'$). Now select the rest of the bijection, and the remaining elements of S_0 and S_1 according to $\mathcal{P}_p^{D'}$. Note that an alternative experiment resulting in the same probability space of variable restrictions is: choose the sets S_0 and S_1 by selecting k-subsets from D_0 and D_1 uniformly and at random, where k is selected from the shifted binomial distribution $B(n-i,p)+i$. (The notation $B(m,p)+k$ means select q according to $B(m,p)$ and then add k.) Then select a random bijection, π, from $D_0 \setminus S_0$ to $D_1 \setminus S_1$.

If $\mathcal{K}$ is a distribution of variable restrictions defined on the domain $D \setminus v(Y)$, then $< \mathcal{K}, Y = * >$ denotes the distribution of variable restrictions on the domain D, where each $\rho \in \mathcal{K}$ over $D \setminus v(Y)$, is extended to a restriction over D, by selecting all variables in $v(Y)$ to be set to *.

We are now ready to examine the distribution of variable restrictions induced by $\mathcal{P}_p^D$, given that $\rho(Y) = *$. Let $V = i$ denote the event that exactly i elements of Y_0 are mapped outside of Y_1 by π. (Note that this implies that exactly i elements of Y_1 are also mapped outside of Y_0 by π.) Then the distribution $\mathcal{P}_p^D$ given that $\rho(Y) = *$ can be partitioned into subdistributions: $\mathcal{P}_{p,i}^D$, given that $\rho(Y) = *$ and $V = i$, where i ranges from 0 to $|Y|$. Now let us take a look at the distribution $\mathcal{P}_{p,i}^D$, given that $\rho(Y)$ and $V = i$, for a particular i. We claim that the distribution of restrictions induced by $\mathcal{P}_{p,i}^D$, given that $\rho(Y) = *$ and $V = i$ is equivalent to the distribution of restrictions induced by

$< \mathcal{T}_{p,i}^{D\setminus v(Y)}, Y = * >$.

Let $\sigma \in Proj_D[Y]$ be a particular minimal partial 1-1 map over D which involves all of the elements of Y. We can further divide the distribution $\mathcal{P}_p^D$ given $\rho(Y) = *$ and $V = i$, based on the particular value of σ consistent with π. Note that because $V = i$, we have restricted our attention to those σ's where exactly i elements of $v(Y_0)$ are mapped outside of $v(Y)$; we will refer to this subset of all σ's by $Proj_{D,i}[Y]$. For a particular $\sigma_i \in Proj_{D,i}[Y]$, the event that π is consistent with σ_i will be denoted by σ_i. For a fixed $\sigma_i \in Proj_{D,i}[Y]$, the distribution of variable restrictions given that $\rho(Y) = *$, σ_i and $V = i$, is equivalent to the distribution of variable restrictions of $< \mathcal{P}_p^{D-v(\sigma_i)}, \sigma_i = * >$. This can be seen by viewing $\rho \in \mathcal{P}_p^D$ as being chosen from the third experiment, and observing that the conditions $(\rho(Y) = *, \sigma_i, V = i)$ completely fix π on $v(\sigma_i)$, and also $S_1 \cap v(\sigma_i)$, and $S_0 \cap v(\sigma_i)$.

Note that each $\sigma_i \in Proj_{D,i}[Y]$ is equally likely; i.e., $\forall \sigma_i^1, \sigma_i^2 \in Proj_{D,i}[Y]$, $Pr[\sigma_i^1 \mid V = i \wedge \rho(Y) = *] = Pr[\sigma_i^2 \mid V = i \wedge \rho(Y) = *]$. Therefore, as we range over the possible $\sigma_i \in Proj_{D,i}[Y]$, the set $\sigma_i(Y)$ is a set of i domain elements, and i range elements, each chosen randomly from $D \setminus v(Y)$. Thus, the distribution $\mathcal{P}_p^D$, given that $\rho(Y) = *$ and $V = i$, is equivalent to the distribution generated by $< \mathcal{T}_{p,i}^{D-v(Y)}, Y = * >$.

The distribution of variable restrictions given that $\rho(Y) = *$ can thus be generated as follows. First, select S_0' and S_1' by choosing k-subsets of $D_1 \setminus v(Y)$, and k-subsets of $D_0 \setminus v(Y)$ uniformly at random, where k is chosen as follows. Select a category i, $0 \leq i \leq |Y_0|$, where category j is chosen with probability $Pr[V = j \mid \rho(Y) = *]$. Then select k' according to the shifted binomial, $B(|D_0| - |Y_0| - i, p) + i$. Let $k = k' + i$. The selection of category i corresponds to deciding how many elements of $v(Y)$ are mapped outside of $v(Y)$. The selection of k' corresponds to the distribution $\mathcal{T}_{p,i}^{D-v(Y)}$. The set S_0 will be $S_0' \cup Y_0$, and the set S_1 will be $S_1' \cup Y_1$. Lastly, choose a random bijection between $D_0 \setminus S_0$ and $D_1 \setminus S_1$.

It will also be necessary to understand the variable distribution, given that $\rho(Y) = 1$. Let $Y = P_{u_1 v_1} P_{u_2 v_2} .. P_{u_k v_k}$; then the event $\rho(Y) = 1$ means that $v(Y_0) \cap S_0 = \emptyset$, and $v(Y_1) \cap S_1 = \emptyset$, and $\pi(u_1) = v_1$, $\pi(u_2) = v_2$, ..., $\pi(u_k) = v_k$. Again, it will be slightly more convenient to work with the distribution $\mathcal{P}_p^D$. Luckily, this conditional distribution is much simpler than the one described above where $\rho(Y) = *$. If we view $\rho \in \mathcal{P}_p^D$ as being chosen from the third experiment, then the condition $\rho(Y) = 1$ completely fixes π on $v(Y)$, and $S \cap v(Y)$. Therefore, the conditional distribution given that $\rho(Y) = 1$ can be shown to be equivalent to the distribution $< \mathcal{P}_p^{D'}, Y = 1 >$, where $D' =$

$D - v(Y)$, and $Y = 1$ denotes the extension of the restrictions in $\mathcal{P}_p^{D'}$ to D by selecting $\rho(Y) = 1$.

Switching Lemma Proof

Recall that f is a disjunction of maps. The proof of the Switching Lemma, like that of Håstad, proceeds by induction on the number of maps in f. We work along the terms one by one: if ρ falsifies a particular map, then we are left with essentially the same problem as before; if ρ does not falsify the map then, it is much more likely that ρ satisfies the map (and thus ensures that the whole formula is set to true) than ρ leaves any variable in the map unset. There are significant complications however in dealing with our partial 1-1 restrictions as opposed to fully independent ones. Once we know that a variable (edge) is unset we have information that biases incident variables towards being unset. Furthermore there is the subtler problem that having some variables set to 0 may bias other variables towards being unset. Both of these complicate the application of the inductive argument in the case that a given map is not falsified. We handle the first problem by considering not only all possible assignments to the unset variables in the map (as in Håstad's proof) but also to all variables that are incident to those unset variables. We get around the second problem by showing that, although setting variables to 0 may make a given variable more likely to be unset, it cannot bias the total number of unset variables to be larger and this turns out to be sufficient for our purposes.

We will obtain the Switching Lemma from the somewhat stronger Lemma 22 by setting $F = 0$ and $Q = \emptyset$, but first we prove a couple of technical lemmas.

LEMMA 19. *Let $D = D_0 \cup D_1$ such that $|D_0| = |D_1| = n$; let Y be a map over D, $|Y| = k$, and let Z be another map over D, where $|Z| = z$, and $v(Z) \cap v(Y) = \emptyset$. If $\frac{k+z}{n-k-z} \leq p$, then for ρ chosen at random from $\mathcal{R}_p^D$,*

$$Pr[\rho(Y) = * \mid \rho(Z) = *] \leq (5p^2)^k.$$

PROOF. Let v be a map of size one, and Q be a map of size q. Assume for now that for all such v and Q, $Pr[\rho(v) = * \mid \rho(Q) = *] \leq 5p^2$, whenever $q/(n-q) \leq p$. Let $Y = y^1y^2...y^k$. Then the probability that Y is set to $*$, given that $\rho(Z) = *$ is equal to:

$$Pr[\rho(y^1) = * \mid \rho(Z) = *] \cdot Pr[\rho(y^2) = * \mid \rho(Z \cup y^1) = *] \cdot .. \cdot \\ Pr[\rho(y^k) = * \mid \rho(Z \cup y^1 \cup .. \cup y^{k-1}) = *].$$

Because each term is of the form $Pr[\rho(v) = * \mid \rho(Q) = *]$, where $q = |Q| \leq k+z$ satisfy $\frac{q}{n-q} \leq p$, we can upper bound each term by $5p^2$, and therefore the whole quantity by $(5p^2)^k$.

It is left to show that $Pr[\rho(v) = * \mid \rho(Q) = *] \leq 5p^2$, where v is a map of size one.

Let Q_0 denote the underlying variables of Q in D_0, and let Q_1 denote the underlying variables of Q in D_1. Similarly, let v_0 denote the underlying variable of v in D_0, and v_1 be the underlying variable of v in D_1. It will be helpful to think of the distribution R_p^D as being generated in the following way. Initially, we begin with D_0 and D_1. Iteratively, we choose an element $x_0 \in D_0$. Then we choose $\pi(x_0)$, from the elements in D_1. Next, we select both x_0 and $\pi(x_0)$ to be set to * with probability p. Now, let D_0 be $D_0 - x_0$, and let D_1 be $D_1 - \pi(x_0)$, and continue on the smaller domain and range until we have matched up all elements from the domain and range. The pairs $(x_i, \pi(x_i))$ which were not chosen to be set to * are then set to 1, and we obtain the natural setting of the underlying variables in the obvious way.

Let $V = i$ denote the event that exactly i elements of Q_0 are mapped outside of Q_1 by the bijection π. Then the probability is a weighted average of the probabilities $Pr[\rho(v) = * \mid \rho(Q) = * \ \wedge \ V = i]$, for all i, $0 \leq i \leq q$. We will upper bound the above probability for a fixed value of i.

Let $\pi(Q_0)$ denote the range of the bijection on the elements in Q_0. We will further divide the above probability based on the four possible ways that v_0 and v_1 get mapped by π: (1) $\pi(v_0) \in Q_1$ and $v_1 \in \pi(Q_0)$; (2) $\pi(v_0) \in Q_1$ and $v_1 \notin \pi(Q_0)$; (3) $\pi(v_0) \notin Q_1$ and $v_1 \in \pi(Q_0)$; and (4) $\pi(v_0) \notin Q_1$ and $v_1 \notin \pi(Q_0)$. Let the above events be denoted by (1), (2), (3), and (4), respectively. Then the above probability is equal to:

$$\sum_{j=1}^{4} Pr[\rho(v) = * \mid (j) \ \wedge \ V = i \ \wedge \rho(Q) = *] Pr[(j) \mid V = i \ \wedge \ \rho(Q) = *].$$

We will now calculate the 8 quantities in the above summation.

1. $Pr[\rho(v) = * \mid (1) \ \wedge \ V = i \ \wedge \ \rho(Q) = *] = 1$. This holds because both v_0 and v_1 have been paired with elements of Q, which have been selected to be $*$.

2. $Pr[\rho(v) = * \mid (2) \ \wedge \ V = i \ \wedge \ \rho(Q) = *] = p$. This holds because v_0 has been paired with an element of Q and therefore is automatically set to *, and the probability that $\rho(v_1) = *$ is p.

3. $Pr[\rho(v) = * \mid (3) \ \wedge \ V = i \ \wedge \ \rho(Q) = *] = p$. This holds for the same reason as 2.

4. $Pr[\rho(v) = * \mid (4) \ \wedge \ V = i \ \wedge \ \rho(Q) = *] = \frac{p}{n-q-i} + \frac{p^2(n-q-i-1)}{n-q-i}$. The conditions $V = i$ and (4) tell us that the partial bijection involving elements

of Q does not involve v_0 or v_1. Thus, the remaining $n-q-i$ domain elements and $n-q-i$ range elements are selected according to the above experiment. With probability $1/(n-q-i)$, v_0 is mapped to v_1—in this case, v is set to * with probability p; otherwise, with probability $\frac{n-q-i-1}{n-q-i}$, v_0 and v_1 are mapped to other elements, and thus they are both set to $*$ with probability p^2.

5. $Pr[(1) \mid V=i \ \wedge\ \rho(Q)=*] = (\frac{i}{n-q})^2$.
6. $Pr[(2) \mid V=i \ \wedge\ \rho(Q)=*] = (\frac{i}{n-q})(\frac{n-q-i}{n-q})$.
7. $Pr[(3) \mid V=i \ \wedge\ \rho(Q)=*] = (\frac{i}{n-q})(\frac{n-q-i}{n-q})$.
8. $Pr[(4) \mid V=i \ \wedge\ \rho(Q)=*] = (\frac{n-q-i}{n-q})^2$.

Thus, the total probability of $Pr[\rho(v)=* \mid V=i \ \wedge\ \rho(Q)=*]$ is:

$$\begin{aligned}
& (\frac{i}{n-q})^2 + 2p(\frac{i}{n-q})(\frac{n-q-i}{n-q}) + (\frac{n-q-i}{n-q})^2(\frac{p+p^2(n-q-i-1)}{n-q-i}) \\
\leq\ & (\frac{q}{n-q})^2 + 2p(\frac{q}{n-q}) + \frac{p}{n-q} + p^2 \\
\leq\ & p^2 + 2p^2 + p^2 + p^2 \\
\leq\ & 5p^2.
\end{aligned}$$

The first inequality holds because $i \leq q$, and the second inequality holds because $q/(n-q) \leq p$. □

LEMMA 20. *Let Y and Q be maps over $D = D_0 \cup D_1$, $|Y| = k$, $|Q| = q$, and $|D_0| = |D_1| = n$, and let $v(Y) \cap v(Q) = \emptyset$. If $n - 2(k+q) \geq 6n/7$, then $Pr[\rho(Y) = 1 \mid \rho(Q) = *] \geq (\frac{6(1-p)}{7n})^k$.*

PROOF. As in the previous lemma, the probability $Pr[\rho(Y) = 1 \mid \rho(Q) = *]$ can be written as: $Pr[\rho(y_1) = 1 \mid \rho(Q) = *] \cdot \ldots \cdot Pr[\rho(y_k) = 1 \mid \rho(Q) = 1 \wedge \rho(y_1..y_{k-1}) = 1]$. We will show that when k and q satisfy $n-2(k+q) \geq 6n/7$, then for a given map of size one, $Pr[\rho(y) = 1 \mid \rho(Q) = * \ \wedge\ \rho(Y) = 1]$ is at least $\frac{6(1-p)}{7n}$. Therefore the probability $Pr[\rho(Y) = 1 \mid \rho(Q) = *]$ is at least $(\frac{6(1-p)}{7n})^k$.

The probability $Pr[\rho(y) = 1 \mid \rho(Q) = * \ \wedge\ \rho(Y) = 1]$ is a weighted average of the probabilities $Pr[\rho(y) = 1 \mid \rho(Q) = * \ \wedge\ \rho(Y) = 1 \ \wedge\ V = i]$. We will obtain a lower bound for each of these probabilities. For a fixed i, we will again break up the probability according to where y is mapped by π, as described by the above events, (1), (2), (3), and (4).

The only event which does not result in a probability of zero is $Pr[\rho(y) = 1 \mid (4) \wedge V = i \wedge \rho(Y) = 1 \wedge \rho(Q) = *]$. This probability is equal to $\frac{1-p}{n-k-q-i}$, because we must guarantee that y_0 gets mapped to y_1, and this happens with probability $\frac{1}{n-k-q-i}$, and also that these two get set to 1, which happens with probability $1-p$. The probability $Pr[(4) \mid V = i \ \wedge \rho(Y) = 1 \ \wedge \ \rho(Q) = *]$ is equal to $(\frac{n-k-q-i}{n-k-q})^2$. Thus the total probability is:

$$(\frac{1-p}{n-k-q-i})(\frac{n-k-q-i}{n-k-q})^2 \geq \frac{(1-p)(n-2(k+q))}{n^2} \geq \frac{6(1-p)}{7n}.$$

The last inequality holds because $n - 2(k+q) \geq 6n/7$. □

LEMMA 21. *Suppose that* $0 \leq \alpha_0 \leq \alpha_1 \leq ... \leq \alpha_n$, *and for all* $k \leq n$, $\sum_{j=k}^n a_j \leq \sum_{j=k}^n b_j$. *Then for all* $k \leq n$, $\sum_{j=k}^n \alpha_j a_j \leq \sum_{j=k}^n \alpha_j b_j$.

PROOF. The proof is by downward induction on k. For $k = n$, the lemma holds. Now assume that the lemma holds for k. Consider $\sum_{j=k-1}^n \alpha_j b_j$. Either $b_{k-1} \geq a_{k-1}$ or $b_{k-1} < a_{k-1}$. In the first case, by the induction hypothesis, we know that $\sum_{j=k}^n \alpha_j b_j \geq \sum_{j=k}^n \alpha_j a_j$, thus because $b_{k-1} \geq a_{k-1}$, we also have $\sum_{j=k-1}^n \alpha_j b_j \geq \sum_{j=k-1}^n \alpha_j a_j$. In the second case, let $\delta = a_{k-1} - b_{k-1}$. Because $\sum_{j=k-1}^n b_j \geq \sum_{j=k-1}^n a_j$, we have that $\sum_{j=k}^n b_j \geq \sum_{j=k}^n a_j + \delta$. Applying the inductive hypothesis, with a_k replaced by $a_k + \delta$, we have:

$$\begin{aligned}
& \sum_{j=k}^n \alpha_j b_j \geq \sum_{j=k+1}^n \alpha_j a_j \ + \ \alpha_k(a_k + \delta) \\
\Rightarrow \ & \sum_{j=k-1}^n \alpha_j b_j \geq \sum_{j=k}^n \alpha_j a_j + \alpha_k \delta + \alpha_{k-1} b_{k-1} \\
\Rightarrow \ & \sum_{j=k-1}^n \alpha_j b_j \geq \sum_{j=k}^n \alpha_j a_j + \alpha_k(a_{k-1} - b_{k-1}) + \alpha_{k-1} b_{k-1} \\
\Rightarrow \ & \sum_{j=k-1}^n \alpha_j b_j \geq \sum_{j=k}^n \alpha_j a_j + \alpha_{k-1}(a_{k-1} - b_{k-1} + b_{k-1}) \\
\Rightarrow \ & \sum_{j=k-1}^n \alpha_j b_j \geq \sum_{j=k-1}^n \alpha_j a_j.
\end{aligned}$$

□

LEMMA 22. *Let* Q *be an arbitrary map over* $D = D_0 \cup D_1$, $|D_0| = |D_1| = n$, $|Q| = q$, *let* f *be an* r*-disjunction over* $D' = D \backslash v(Q)$, *and let* F *be an arbitrary* *function over D, $F \neq 1$. Let ρ be a random restriction chosen according to $\mathcal{R}_p^D$. Then for $s \geq 0$, $p \leq 1/36$, and $pn \geq 8(q+s+r)^2$ we have*

$$Pr[\delta(f\restriction_\rho) \geq s \mid F\restriction_\rho = 0 \ \wedge \ \rho(Q) = *] \leq \alpha^s,$$

for any $\alpha > 0$ satisfying $(1 + 36p^4n^3/\alpha^2)^r \leq 2$.

PROOF. The proof proceeds by induction on the total number of maps in f.

Base Case. There are no maps in f. In this case f is identically 0 and therefore f is represented by the tree consisting of the single node labelled 0. Hence $\delta(f\restriction_\rho) = 0$ and the lemma holds.

Induction Step. Assume that the lemma holds for all map disjunctions with fewer maps than the map disjunction of f. We will write f as $f_1 \vee f_2 \vee ...$, where each f_i is a map of f. We will analyze the probability by considering separately the cases in which ρ does or does not force the map f_1 to be 0. The failure probability, the probability that $\delta(f\restriction_\rho) \geq s$, is an average of the failure probabilities of these two cases. Thus

$$\begin{aligned} & Pr[\delta(f\restriction_\rho) \geq s \mid F\restriction_\rho = 0 \ \wedge \ \rho(Q) = *] \\ & \leq \ \max(Pr[\delta(f\restriction_\rho) \geq s \mid F\restriction_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ f_1\restriction_\rho = 0], \\ & \qquad Pr[\delta(f\restriction_\rho) \geq s \mid F\restriction_\rho = 0 \ \wedge \ \rho(Q) = * \wedge \ f_1\restriction_\rho \neq 0]). \end{aligned}$$

The first term in the maximum is $Pr[\delta(f\restriction_\rho) \geq s \mid (F \vee f_1)\restriction_\rho = 0 \wedge \rho(Q) = *]$. Let f' be f with map f_1 removed; then $Pr[\delta(f\restriction_\rho) \geq s \mid (F \vee f_1)\restriction_\rho = 0 \ \wedge \ \rho(Q) = *] = Pr[\delta(f'\restriction_\rho) \geq s \mid (F \vee f_1)\restriction_\rho = 0 \wedge \rho(Q) = *]$. Because f' has one less map than f, this probability is no greater than α^s, by the inductive hypothesis.

Now we will estimate the second term in the maximum. Let T be the set of variables appearing in the first map, f_1. By hypothesis, $size(T) \leq r$. We will analyze the cases based on the subset Y of the variables in T to which ρ assigns $*$; we use the notation $*(\rho_T) = Y$ to denote the event that the variables in T which are assigned $*$ by ρ_T are exactly those in Y (where ρ_T denotes ρ restricted to T). Then

$$\begin{aligned} & Pr[\delta(f\restriction_\rho) \geq s \mid F\restriction_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ f_1\restriction_\rho \neq 0] \\ & = \ \sum_{Y \subseteq T} Pr[\delta(f\restriction_\rho) \geq s \ \wedge \ *(\rho_T) = Y \mid F\restriction_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ f_1\restriction_\rho \neq 0]. \end{aligned}$$

Consider the case in which $Y = \emptyset$. In this case the value of f_1 is forced to 1 by ρ. It follows that f is forced to 1 and hence $\delta(f) = 0$ so the term corresponding to $Y = \emptyset$ has probability 0. The sum then becomes

$$\sum_{\substack{Y \subseteq T, \\ Y \neq \emptyset}} Pr[\delta(f\restriction_\rho) \geq s \ \wedge \ *(\rho_T) = Y \mid F\restriction_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ f_1\restriction_\rho \neq 0],$$

which is equal to

$$\sum_{\substack{Y\subseteq T,\\ Y\neq\emptyset}} Pr[\delta(f\restriction_\rho)\geq s \mid F\restriction_\rho = 0 \ \wedge\ \rho(Q)=* \ \wedge\ f_1\restriction_\rho \neq 0 \ \wedge\ *(\rho_T)=Y] \quad (1)$$
$$\times\ Pr[*(\rho_T)=Y \mid F\restriction_\rho = 0 \ \wedge\ \rho(Q)=* \ \wedge\ f_1\restriction_\rho \neq 0]. \quad (2)$$

We will first bound the latter term, (2), in each of these products. Given that $f_1\restriction_\rho \neq 0$, the probability that $*(\rho_T) = Y$ is equal to the probability that $\rho(Y) = * \ \wedge\ \rho(T\setminus Y) = 1$. Thus term (2) equals

$$Pr[\rho(Y) = * \wedge \rho(T\setminus Y) = 1 \mid F\restriction_\rho = 0 \ \wedge\ \rho(Q) = * \ \wedge\ f_1\restriction_\rho \neq 0]$$
$$\leq Pr[\rho(Y) = * \mid F\restriction_\rho = 0 \ \wedge\ \rho(Q) = * \ \wedge\ \rho(T\setminus Y) = 1 \ \wedge \rho(Y) \neq 0],$$

where $\rho(Y) \neq 0$ means that no variable in Y is set to 0, and $\rho(Y) = *$ means that every variable in Y is set to $*$. Let F' be $F \vee G$ where $G\restriction_\rho = 0$ if and only if ρ sets all variables in $T\setminus Y$ to 1. ($G = \bigvee_{x\in T\setminus Y}\overline{x}$.) Then the above probability is equal to $Pr[\rho(Y) = * \mid F'\restriction_\rho = 0 \ \wedge\ \rho(Q) = * \ \wedge\ \rho(Y) \neq 0]$. In proving a bound on term (2), we will need the following proposition.

PROPOSITION 23. *Let* $|Y| = k$, $|Q| = q$. *When* $pn \geq 8(k+q)^2$, $Pr[\rho(Y) = * \mid F\restriction_\rho = 0 \ \wedge\ \rho(Q) = * \ \wedge\ \rho(Y) \neq 0] \ \leq\ Pr[\rho(Y) = * \mid \rho(Q) = * \ \wedge\ \rho(Y) \neq 0]$.

It is interesting to note that, unlike earlier switching lemmas over other distributions, it is *not* true that setting some variables to $*$ can only increase the likelihood that the function is not forced to zero. The following counterexample illustrates that in our situation, it may be more likely to force F to 0, given that $\rho(Y) = *$. Let $D_0 = \{1,2\}$, $D_1 = \{a,b\}$, $Y = P_{2b}$, and $F = P_{1b}$. Then simple calculations show: $Pr[F\restriction_\rho = 0] = \frac{1}{2}(1-p^2)$; $Pr[F\restriction_\rho = 0 \wedge \rho(Y) = *] = \frac{1}{2}(1-p)p$; and $Pr[\rho(Y) = *] = \frac{1}{2}p(1+p)$. Thus, $Pr[F\restriction_\rho = 0 \mid \rho(Y) = *] = \frac{1-p}{1+p}$, and $Pr[F\restriction_\rho = 0] = \frac{1}{2}(1-p^2)$. Therefore, when $p < 1/4$, we have $Pr[F\restriction_\rho = 0] \leq Pr[F\restriction_\rho = 0 \mid \rho(Y) = *]$. Fortunately, Proposition 23 still holds, although the intuition is less obvious.

PROOF. [of Proposition 23.] As in previous proofs of Håstad's Lemma, we will prove Proposition 23 by showing that

$$Pr[F\restriction_\rho = 0 \mid \rho(Y) = * \ \wedge\ \rho(Q) = *] \ \leq\ Pr[F\restriction_\rho = 0 \mid \rho(Y) \neq 0 \ \wedge\ \rho(Q) = *].$$

For arbitrary events A, B, and C, $Pr[A \mid B\wedge C] \leq Pr[A \mid C] \leftrightarrow Pr[B \mid A\wedge C] \leq Pr[B \mid C]$. By applying this fact where A is the event $\rho(Y) = *$, B is the event $F\restriction_\rho = 0$, and C is the event $(\rho(Y) \neq 0 \ \wedge\ \rho(Q) = *)$, and then observing that the condition $\rho(Y) \neq 0 \ \wedge \ \rho(Y) = *$ is equivalent to the condition $\rho(Y) = *$, it follows that the above inequality implies the proposition.

Let $V = i$ denote the event that there are exactly i elements of $Y_0 \cup Q_0$ mapped outside of $Y_0 \cup Q_0$ by ρ. Then $Pr[F\lceil_\rho = 0 \mid \rho(Y) = * \ \wedge \ \rho(Q) = *]$ is equal to:

$$\sum_{i=0}^{k+q} Pr[F\lceil_\rho = 0 \mid \rho(Y) = * \wedge \rho(Q) = * \wedge V = i] \cdot Pr[V = i \mid \rho(Y) = * \wedge \rho(Q) = *].$$

Similarly, $Pr[F\lceil_\rho = 0 \mid \rho(Y) \neq 0 \ \wedge \ \rho(Q) = *]$ is equal to

$$\sum_{i=0}^{k+q} Pr[F\lceil_\rho = 0 | \rho(Y) \neq 0 \wedge \rho(Q) = * \wedge V = i] \cdot Pr[V = i \mid \rho(Y) \neq 0 \wedge \rho(Q) = *].$$

The proof proceeds in three steps:

(Step 1) For each i, $0 \leq i \leq k + q$,

$$\begin{aligned} &Pr[F\lceil_\rho = 0 \mid \rho(Y) = * \ \wedge \ \rho(Q) = * \ \wedge \ V = i] \\ &\quad \leq \ Pr[F\lceil_\rho = 0 \mid \rho(Y) \neq 0 \ \wedge \ \rho(Q) = * \ \wedge \ V = i]. \end{aligned}$$

(Step 2) For each i, $0 \leq i \leq k + q$,

$$\begin{aligned} &Pr[F\lceil_\rho = 0 \mid \rho(Y) = * \ \wedge \ \rho(Q) = * \ \wedge \ V = i] \\ &\quad \geq \ Pr[F\lceil_\rho = 0 \mid \rho(Y) = * \ \wedge \ \rho(Q) = * \ \wedge \ V = i + 1]. \end{aligned}$$

(Step 3) For all j, $0 \leq j \leq k + q$,

$$\sum_{i=0}^{j} Pr[V = i \mid \rho(Y) = * \ \wedge \ \rho(Q) = *] \leq \sum_{i=0}^{j} Pr[V = i \mid \rho(Y) \neq 0 \ \wedge \ \rho(Q) = *].$$

Then, by Lemma 21, the proposition follows.

Note that an easier proof would be to just show Step 1, and then show that for each i, $Pr[V = i \mid \rho(Y) = * \wedge \rho(Q) = *] \leq Pr[V = i \mid \rho(Y) \neq 0 \wedge \rho(Q) = *]$. Unfortunately, this is false. Instead, we are able to show that the aggregate sum is always smaller on the LHS- this is Step 3.

We will first prove Step 1. We will break up the collection of restrictions satisfying $(\rho(Y) \neq 0 \wedge \rho(Q) = * \wedge V = i)$ into equivalence classes as follows. Let $D_0' = D_0 \setminus (Y_0 \cup Q_0)$, $D_1' = D_1 \setminus (Y_1 \cup Q_1)$, and let $\rho^* = (B_0, B_1, \pi^*, B_0^*, B_1^*)$ be a partial restriction, defined as follows. B_0 is a subset of D_0' of size $|D_0'| - i$; B_1

is a subset of D_1' of size $|D_1'| - i$; π^* is a bijection from B_0 to B_1, and $B_0^* \subseteq B_0$, $B_1^* \subseteq B_1$. A restriction $\rho = (\pi, S_0, S_1) \in \mathcal{P}_p^D$ is in the equivalence class labelled by ρ^* if: (1) π is an extension of π^*; and (2) $S_0 \cap B_0 = B_0^*$, and $S_1 \cap B_1 = B_1^*$.

Note that each restriction which satisfies $V = i$ is consistent with exactly one ρ^*, and therefore the probability that $F\lceil_\rho = 0$, given that $(V = i \wedge \rho(Y) = * \wedge \rho(Q) = *)$ is equal to:

$$\sum_{\rho^*} \begin{array}{l} Pr[F\lceil_\rho = 0 \mid V = i \wedge \rho(Y) = * \wedge \rho(Q) = * \wedge \rho^*] \\ \times\ Pr[\rho^* \mid V = i \wedge \rho(Y) = * \wedge \rho(Q) = *], \end{array}$$

where ρ^* denotes the event that ρ is in the equivalence class labelled by ρ^*. Similarly, the probability that $F\lceil_\rho = 0$, given that $(V = i \wedge \rho(Y) \neq 0 \wedge \rho(Q) = *)$ is equal to:

$$\sum_{\rho^*} \begin{array}{l} Pr[F\lceil_\rho = 0 \mid V = i \wedge \rho(Y) \neq 0 \wedge \rho(Q) = * \wedge \rho^*] \\ \times\ Pr[\rho^* \mid V = i \wedge \rho(Y) \neq 0 \wedge \rho(Q) = *]. \end{array}$$

First, we will argue that for each ρ^*,

$$Pr[\rho^* \mid \rho(Y) = * \wedge \rho(Q) = * \wedge V = i] = Pr[\rho^* \mid \rho(Y) \neq 0 \wedge \rho(Q) = * \wedge V = i].$$

To see this, observe that given $V = i$, each choice of B_0, B_1 and π^* is equally likely, and independent of $(\rho(Y) = * \wedge \rho(Q) = *)$, and also of $(\rho(Y) \neq 0 \wedge \rho(Q) = *)$. Secondly, given B_0, B_1, π^*, $\rho(Y) \neq 0$, and $\rho(Q) = *$, the choice of B_0^*, and B_1^* is independent of whether $\rho(Y) = *$.

It is left to show that for each ρ^*,

$$\begin{array}{rl} & Pr[F\lceil_\rho = 0 \mid \rho^* \wedge \rho(Y) = * \wedge \rho(Q) = * \wedge V = i] \\ \leq & Pr[F\lceil_\rho = 0 \mid \rho^* \wedge \rho(Y) \neq 0 \wedge \rho(Q) = * \wedge V = i]. \end{array}$$

Note that when $(\rho^* \wedge \rho(Y) = * \wedge \rho(Q) = * \wedge V = i)$, the probability that $F\lceil_\rho = 0$ is either 0 or 1, because the underlying restriction to the variables is completely determined. Now consider the collection of ρ's, satisfying $(\rho(Y) = * \wedge \rho(Y) \neq 0 \wedge V = i)$, that lie in the same equivalence class as ρ^*. If $F\lceil_\rho$ is forced to zero, given that $(\rho(Y) = * \wedge \rho(Q) = * \wedge V = i \wedge \rho^*)$, then $F\lceil_\rho$ is also forced to 0 by all other ρ's such that $(\rho(Y) \neq 0 \wedge \rho(Q) = * \wedge V = i \wedge \rho^*)$. This holds because setting more variables to 1 or 0 continues to force F to zero. This completes the proof of Step 1.

The intuition behind Step 2 is simply that the larger $V = i$ is, the more stars there are in $D - Y$, and hence the less likely it is that F is forced to

0. To prove step (2), let $R = Q \cup Y$ and fix i. Then we want to prove $Pr[F\lceil_\rho = 0 \mid \rho(R) = * \wedge V = i] \leq Pr[F\lceil_\rho = 0 \mid \rho(R) = * \wedge V = i-1]$.

Let $D' = D \setminus v(R)$, and let $|D'_0| = |D'_1| = n - |R| = m$. The probability $Pr[F\lceil_\rho = 0 \mid \rho(R) = * \wedge V = i]$ can be divided, according to the exact number of *'s that are assigned to vertices in D'_0, and D'_1, by ρ:

$$\begin{aligned} & Pr[F\lceil_\rho = 0 \mid \rho(R) = * \wedge V = i] \\ = & \sum_{j=0}^{m} Pr[F\lceil_\rho = 0 \wedge \#(D') = j \mid \rho(R) = * \wedge V = i], \end{aligned}$$

where the event $\#(D') = i$ means that exactly i vertices in D'_0 are assigned *. (As a consequence, exactly i vertices in D'_1 are also assigned *.) A similar equation holds when $\rho(R) = *$ and $V = i - 1$:

$$\begin{aligned} & Pr[F\lceil_\rho = 0 \mid \rho(R) = * \wedge V = i-1] \\ = & \sum_{j=0}^{m} Pr[F\lceil_\rho = 0 \wedge \#(D') = j \mid \rho(R) = * \wedge V = i-1]. \end{aligned}$$

Note that when $j < i$, the probability $[F\lceil_\rho = 0 \wedge \#(D') = j \mid \rho(R) = * \wedge V = i]$ is zero; therefore, it is left to show:

$$\begin{aligned} & \sum_{j=i}^{m} Pr[F\lceil_\rho = 0 \mid \rho(R) = * \wedge V = i \wedge \#(D') = j] \\ & \quad \times Pr[\#(D') = j \mid \rho(R) = * \wedge V = i] \\ \leq & \sum_{j=i-1}^{m} Pr[F\lceil_\rho = 0 \mid \rho(R) = * \wedge V = i-1 \wedge \#(D') = j] \\ & \quad \times Pr[\#(D') = j \mid \rho(R) = * \wedge V = i-1]. \end{aligned}$$

First, note that the event $V = i - 1$, or $V = i$ is irrelevant, given that $\#(D') = j$ and $\rho(R) = *$. Therefore, we have

$$\begin{aligned} & Pr[F\lceil_\rho = 0 \mid V = i-1 \wedge \rho(R) = * \wedge \#(D') = j] \\ = & Pr[F\lceil_\rho = 0 \mid V = i \wedge \rho(R) = * \wedge \#(D') = j] \\ = & Pr[F\lceil_\rho = 0 \mid \rho(R) = * \wedge \#(D') = j]. \end{aligned}$$

Thus to complete Step 2, it remains to show:

$$\begin{aligned} & \sum_{j=i}^{m} Pr[F\lceil_\rho = 0 \mid \#(D') = j \wedge \rho(R) = *] \\ & \quad \times Pr[\#(D') = j \mid V = i \wedge \rho(R) = *] \end{aligned}$$

$$\begin{aligned} &= Pr[F\restriction_\rho = 0 \mid \#(D') = i-1 \wedge \rho(R) = *] \cdot 0 \\ &\quad + \sum_{j=i}^{m} Pr[F\restriction_\rho = 0 \mid \#(D') = j \wedge \rho(R) = *] \\ &\qquad \times Pr[\#(D') = j \mid V = i \wedge \rho(R) = *] \\ &\leq \sum_{j=i-1}^{m} Pr[F\restriction_\rho = 0 \mid \#(D') = j \wedge \rho(R) = *] \\ &\qquad \times Pr[\#(D') = j \mid V = i-1 \wedge \rho(R) = *]. \end{aligned}$$

By Lemma 21 , it suffices to show:

(Step 2a) $\forall k, i \leq k \leq m$,

$$\begin{aligned} &Pr[F\restriction_\rho = 0 \mid \#(D') = k-1 \wedge \rho(R) = *] \\ &\quad \geq Pr[F\restriction_\rho = 0 \mid \#(D') = k \wedge \rho(R) = *]. \end{aligned}$$

(Step 2b) $\forall j, i \leq j \leq m$,

$$\begin{aligned} &Pr[i \leq \#(D') \leq j \mid V = i \wedge \rho(R) = *] \\ &\quad \leq Pr[i-1 \leq \#(D') \leq j \mid V = i-1 \wedge \rho(R) = *]. \end{aligned}$$

We will first prove Step 2a. Let $\#(\rho)$ denote the exact number of $*$'s that are assigned to D_1 by ρ. Then the events $(\#(D') = k-1 \wedge \rho(R) = *)$ are equivalent to the events $(\#(\rho) = k-1+|R| \wedge \rho(R) = *)$. Therefore, the following proposition proves Step 2a.

PROPOSITION 24. *Let F be a boolean formula over $D = D_0 \cup D_1$, $|D_0| = |D_1| = n$. Then for all $k \leq n+1$, $Pr[F\restriction_\rho = 0 \mid \#(\rho) = k-1 \wedge \rho(Q) = *] \geq Pr[F\restriction_\rho = 0 \mid \#(\rho) = k \wedge \rho(Q) = *]$.*

PROOF. Recall that the distribution of restrictions given $\rho(Q) = *$ can be described as follows. Choose a category i, $0 \leq i \leq |Q|$, from some distribution. Then choose k' according to the shifted binomial distribution, $B(n-|Q|-i, p)+i$. Let $k = k'+i$. Choose a random set, S_0', of size k from $D_0 \setminus Q_0$, and a random set, S_1' of size k from $D_0 \setminus Q_1$. Let $S_0 = S_0' \cup Q_0$, and let $S_1 = S_1' \cup Q_1$. Then choose a random bijection, π, from $D_0 \setminus S_0$ to $D_1 \setminus S_1$. Therefore, the distribution of restrictions given $\rho(Q) = *$ and $\#(\rho) = k$ can be described by: Choose a random set, S_0' of size $|Q|-k$ from $D_0 \setminus Q_0$, and a random set S_1' of size k from $D_0 \setminus Q_1$. Let $S_0 = S_0' \cup Q_0$, and let $S_1 = S_1' \cup Q_1$. Then choose a random bijection, π, from $D_0 \setminus S_0$ to $D_1 \setminus S_1$.

Let A^k denote the subdistribution of restrictions given that $\rho(Q) = *$ and $\#(\rho) = k$, and let A^{k-1} denote the subdistribution of restrictions given that $\rho(Q) = *$ and $\#(\rho) = k-1$. We would like to show that the probability that $F\lceil_\rho = 0$ in A^k is no greater than the probability that $F\lceil_\rho = 0$ in A^{k-1}.

Let $m = n - |Q|$. Let $\rho^{k-1} =< S_0^{k-1}, S_1^{k-1}, \pi^{k-1} >\in A^{k-1}$, and let $\rho^k =< S_0^k, S_1^k, \pi^k >\in A^k$. Then we say that ρ^{k-1} and ρ^k *correspond* if there exists an $x \in S_0^k$, $y \in S_1^k$, such that $S_0^{k-1} \cup x = S_0^k$, $S_1^{k-1} \cup y = S_1^k$, and $\pi^k \cup < x, y >= \pi^{k-1}$. Note that whenever $\rho^k \in A^k$ forces F to 0, so do all of the elements of A^{k-1} which correspond. This is true because for every ρ^{k-1} which corresponds to ρ^k, all underlying variables are the same except for a few variables which are set to $*$ in ρ^k, and set to 0 or 1 in ρ^{k-1}; in other words, ρ^{k-1} is a further restriction of ρ^k. Now, because F is already forced to 0 by ρ^k, it must continue to be 0 as we set more variables. Thus, F is also forced to 0 by ρ^{k-1}.

Let C^k denote the elements of A^k which force F to 0. For each ρ^k in A^k, there are k^2 elements in A^{k-1} which correspond, and conversely, for each $\rho^{k-1} \in A^{k-1}$, there are $m - k + 1$ elements of A^k which correspond. The probability that a random ρ^k over A^k forces F to 0 equals $\frac{|C^k|}{|A^k|}$; thus the probability that a random ρ^{k-1} over A^{k-1} forces F to 0 is at least $\frac{|C^k| \cdot k^2}{(m-k+1)|A^{k-1}|}$. Since $|A_\pi^{k-1}|$ is equal to $\frac{k^2 |A_\pi^k|}{m-k+1}$, the probability that F is forced to 0 over A^{k-1} is greater than or equal to the probability that F is forced to 0 over A^k. The completes the proof of Proposition 24. □

We will now prove Step 2b. First, note that $Pr[i \leq \#(D') \leq j \mid V = i \wedge \rho(R) = *]$ is equal to $Pr[\#(D') \leq j \mid V = i \wedge \rho(R) = *]$; similarly, $Pr[i-1 \leq \#(D') \leq j \mid V = i-1 \wedge \rho(R) = *]$ is equal to $Pr[\#(D') \leq j \mid V = i-1 \wedge \rho(R) = *]$. Recall that the distribution given $(V = i \wedge \rho(R) = *)$ can be described as follows. First, choose k at random, according to the binomial distribution, shifted by i: $B(m-i, p) + i$. Then randomly choose $S_0' \subseteq D_0'$, $S_1' \subseteq D_1'$, where $|S_0'| = |S_1'| = k$. Let $S_0 = S_0' \cup R_0$, and $S_1 = S_1' \cup R_1'$. Lastly, uniformly select a bijection π from $D_0 \setminus S_0$ to $D_1 \setminus S_1$. The distribution given $(V = i - 1 \wedge \rho(R) = *)$ can be described similarly, except that now k is chosen according to the binomial distribution, shifted by $i-1$: $B(m-i+1, p) + i - 1$. Therefore, the number of $*$'s in D_0', given $(V = i \wedge \rho(R) = *)$ is chosen according to $B(m-i, p) + i$, and the number of $*$'s in D_0', given $(V = i - 1 \wedge \rho(R) = *)$ is chosen according to $B(m-i+1, p) + i - 1$. Therefore, it is clear that $Pr[\#(D') \leq j \mid V = i \wedge \rho(R) = *] \leq Pr[\#(D') \leq j \mid V = i-1 \wedge \rho(R) = *]$. This completes Step 2.

We will now prove Step 3. We want to show that for all j, $0 \leq j \leq k+q$, $Pr[V \leq j \mid \rho(Y) = * \ \wedge \ \rho(Q) = *] \leq Pr[V \leq j \mid \rho(Y) \neq 0 \ \wedge \ \rho(Q) = *]$. The RHS of this inequality is a weighted sum of $Pr[V \leq j \mid \rho(Y') = 1 \ \wedge \ \rho(Y - Y') = * \ \wedge \ \rho(Q) = *]$, where Y' ranges over all subsets of Y. We want to show that when $Y' = \emptyset$, this probability is the smallest. If Y_1 and Y_2 are two equal-sized subsets of Y, then the above probabilities are equivalent; i.e., $Pr[V \leq j \mid \rho(Y_1) = 1 \ \wedge \ \rho(Y - Y_1) = * \ \wedge \ \rho(Q) = *] = Pr[V \leq j \mid \rho(Y_2) = 1 \ \wedge \ \rho(Y - Y_2) = * \ \wedge \ \rho(Q) = *]$. Therefore, it suffices to prove that $Pr[V \leq j \mid \rho(Y') = 1 \ \wedge \rho(Y - Y') = * \ \wedge \rho(Q) = *] \leq Pr[V \leq j \mid \rho(Y'') = 1 \ \wedge \rho(Y - Y'') = * \ \wedge \ \rho(Q) = *]$, where $Y'' = Y' \cup y$, and $y \in Y - Y'$.

Let $m = n - |Y'|$, and let $Q \cup (Y - Y') = Z$, and $Q \cup (Y - Y'') = Z'$. Then the above inequality is equivalent to showing, for all $j \leq |Z| = z$,

$$Pr_m[V \leq j \mid \rho(Z) = *] \leq Pr_{m-1}[V \leq j \mid \rho(Z') = *],$$

where the probability on the left is over a domain of size m, and the probability on the right is over a domain of size $m - 1$, and $|Z| = z$, $|Z'| = z - 1$.

When $j = z$, the inequality on the left side, $Pr_m[V \leq z \mid \rho(Z) = *]$ is 1. Similarly, when $j = z$, the probability on the right side is equal to $Pr_{m-1}[V \leq z \mid \rho(Z') = *]$ which is also 1. It is left to prove the inequality for $j \leq z - 1$.

We will first calculate the probability $Pr_m[V = i \wedge \rho(Z) = *]$, where $|Z| = z$. This probability is equal to $Pr_m[V = i] \cdot Pr_m[\rho(Z) = * \mid V = i]$. There are $\binom{z}{i}$ ways to pick the i domain elements of $v(Z)$ that will be mapped outside of Z, $\binom{m-z}{i}$ ways of picking the i range elements of $D_1/v(Z)$ that these elements will be mapped to, and $i!$ possible bijections between these two sets of elements. Similarly there are $\binom{z}{i}$ ways to pick the domain elements of $v(Z)$, $\binom{m-z}{i}$ ways of selecting the corresponding range elements, and $i!$ bijections between them. Of the remaining elements in $v(Z)$, there are $(z-i)!$ possible bijections between them, and of the remaining $m - z - i$ elements in D_0, there are $(m - z - i)!$ possible bijections between them. Given that $V = i$, the probability that $\rho(Z) = *$ is exactly p^{z+i}. Thus, the probability $Pr_m[V = i \wedge \rho(Q) = *]$ is equal to

$$S(i, z, m) = \frac{\binom{z}{i}^2 \binom{m-z}{i}^2 (i!)^2 (z-i)!(m-z-i)! p^{z+i}}{m!}.$$

The probability $Pr_m[V \leq j \mid \rho(Z) = *]$ is equal to $\frac{\sum_{i=0}^{j} Pr_m[V=i \wedge \rho(Z)=*]}{\sum_{i=0}^{z} Pr_m[V=i \wedge \rho(Z)=*]}$, which is equal to $\frac{\sum_{i=0}^{j} S(i,z,m)}{\sum_{i=0}^{z} S(i,z,m)}$. Similarly, the probability $Pr_{m-1}[V \leq j \mid \rho(Z') =$

$*]$, where $|Z'| = z-1$, is equal to $\frac{\sum_{i=0}^{j} S(i,z-1,m-1)}{\sum_{i=0}^{z-1} S(i,z-1,m-1)}$. Thus, we want to show:

$$\frac{\sum_{i=0}^{j} S(i,z,m)}{\sum_{i=0}^{z} S(i,z,m)} \leq \frac{\sum_{i=0}^{j} S(i,z-1,m-1)}{\sum_{i=0}^{z-1} S(i,z-1,m-1)}.$$

It is straightforward to show that $S(i,z-1,m-1) = S(i,z,m) \cdot \frac{(z-i)m}{z^2 p}$. Therefore,

$$\frac{\sum_{i=0}^{j} S(i,z-1,m-1)}{\sum_{i=0}^{z-1} S(i,z-1,m-1)} = \frac{\sum_{i=0}^{j}(z-i)S(i,z,m)}{\sum_{i=0}^{z-1}(z-i)S(i,z,m)} \geq \frac{\sum_{i=0}^{j} S(i,z,m)}{\sum_{i=0}^{z-1}(z-i)S(i,z,m)}.$$

We will show that $\sum_{i=0}^{z-1}(z-i)S(i,z,m) \leq \sum_{i=0}^{z} S(i,z,m)$, to complete the proof of Step 3. If we can show that $(z-i) \cdot S(i,z,m) \leq S(i+1,z,m)$, then the above inequality follows because then we have $\sum_{i=0}^{z-1}(z-i)S(i,z,m) \leq \sum_{i=1}^{z} S(i,z,m) \leq \sum_{i=0}^{z} S(i,z,m)$.

To see that $(z-i) \cdot S(i,z,m) \leq S(i+1,z,m)$, it is straightforward to show that $S(i+1,z,m) = S(i,z,m)\frac{(z-i)(m-z-i)p}{(i+1)^2}$. Using the fact that $pn \geq 8(k+q)^2$, and $i \leq z-1$, it can be shown that $\frac{(z-i)(m-z-i)p}{(i+1)^2} \geq \frac{(z-i)pm}{2(i+1)^2} \geq \frac{(z-i)pn}{4(i+1)^2} \geq \frac{(z-i)pn}{4z^2}$. Also because $pn \geq 8(k+q)^2 \geq 8z^2$, it follows that $\frac{pn}{4z^2} \geq 1$, and therefore $(z-i) \cdot S(i,z,m) \leq S(i+1,z,m)$. This completes the proof of Proposition 23. □

Because $|Y| \leq r$, $|Q| \leq q$, we have $8(r+q)^2 \leq 8(q+s+r)^2 \leq pn$. Therefore, we can apply Proposition 23 to show that term (2) is at most $Pr[\rho(Y) = * \mid \rho(Q) = * \ \wedge \ \rho(Y) \neq 0]$. Because $|Y| \leq r$, $|Q| = q$, and $pn \geq 8(q+s+r)^2$, it follows that $\frac{|Y|+|Q|}{n-|Y|-|Q|} \leq p$. Therefore we can apply Lemma 19 to obtain

$$Pr[\rho(Y) = * \mid \rho(Q) = *] \leq (5p^2)^{|Y|}.$$

Also, $Pr[\rho(Y) \neq 0 \mid \rho(Q) = *] \ \geq \ Pr[\rho(Y) = 1 \mid \rho(Q) = *]$. Now because $pn \geq 8(q+s+r)^2$, and $p \leq 1/36$, it also follows that $n - 2(|Y| + |Q|) \geq 6n/7$. Therefore, we can apply Lemma 20, to obtain $Pr[\rho(Y) = 1 \mid \rho(Q) = *] \geq (\frac{6(1-p)}{7n})^{|Y|}$. Then because $p \leq 1/36$, we have

$$Pr[\rho(Y) = * \mid \rho(Y) \neq 0 \ \wedge \ \rho(Q) = *] \ \leq \ (\frac{7 \cdot 5p^2 n}{6(1-p)})^{|Y|} \ \leq \ (6p^2 n)^{|Y|}.$$

Now we look at the first term, (1), in each product. Suppose that $2|Y| < s$. For each fixed Y, we will analyze the probability above by applying Lemma 4 with $K = v(Y)$. Recall that f is a map disjunction over $D' = D \setminus v(Q)$. By this lemma, if $\delta(f\upharpoonright_\rho) \geq s$ then there is some $\sigma \in Proj_{D'\upharpoonright_\rho}[v(Y)]$ such that $d_{(D'\upharpoonright_\rho)\upharpoonright_\sigma}((f\upharpoonright_\rho)\upharpoonright_\sigma) \geq s - |\sigma|$. To use this requires that we consider all maps in $Proj_{D'\upharpoonright_\rho}[v(Y)]$. One difficulty is that $D'\upharpoonright_\rho$ is itself a random variable dependent on ρ. We handle this by considering all maps σ in $Proj_{D'}[v(Y)]$ and including them only if $\rho(\sigma) = *$. For notational convenience let $P(D,Y) = Proj_{D'}[v(Y)]$. When $\rho(\sigma) = *$, $(f\upharpoonright_\rho)\upharpoonright_\sigma = (f\upharpoonright_\sigma)\upharpoonright_\rho$, and applying the definition of $\delta(f\upharpoonright_\rho)$, the probability denoted by term (1) is no greater than

$$\begin{aligned}
&\sum_{\sigma\in P(D,Y)} Pr[\delta((f\upharpoonright_\sigma)\upharpoonright_\rho) \geq s - |\sigma| \ \wedge \rho(\sigma) = * \mid \\
&\qquad\qquad F\upharpoonright_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ f_1\upharpoonright_\rho \neq 0 \ \wedge \ *(\rho_T) = Y] \\
&\leq \sum_{\sigma\in P(D,Y)} Pr[\delta((f\upharpoonright_\sigma)\upharpoonright_\rho) \geq s - |\sigma| \mid \\
&\qquad\qquad F\upharpoonright_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ f_1\upharpoonright_\rho \neq 0 \ \wedge \ *(\rho_T) = Y \ \wedge \ \rho(\sigma) = *] \\
&\qquad \times Pr[\rho(\sigma) = * \mid \\
&\qquad\qquad F\upharpoonright_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ f_1\upharpoonright_\rho \neq 0 \ \wedge \ *(\rho_T) = Y] \\
&= \sum_{\sigma\in P(D,Y)} Pr[\delta((f\upharpoonright_\sigma)\upharpoonright_\rho) \geq s - |\sigma| \mid \\
&\qquad\qquad F\upharpoonright_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ \rho(T\setminus Y) = 1 \ \wedge \ \rho(\sigma) = *] \\
&\qquad \times Pr[\rho(\sigma) = * \mid \\
&\qquad\qquad F\upharpoonright_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ \rho(T\setminus Y) = 1 \ \wedge \ \rho(Y) = *].
\end{aligned}$$

The last inequality above holds because the event $(f_1\upharpoonright_\rho \neq 0 \ \wedge \ *(\rho_T) = Y)$ is equivalent to the event $(\rho(Y) = * \ \wedge \ \rho(T\setminus Y) = 1)$, and the condition $\rho(Y) = *$ is implied by $\rho(\sigma) = *$. Recall that if Y is a map, $v(Y) \subseteq D'$ denotes the set of underlying vertices which are contained in the map. We will split up the map σ into two maps σ_1 and σ_2, where a variable $P_{ij} \in \sigma$ is in σ_1 if both $i \in v(Y)$ and $j \in v(Y)$. Otherwise, $P_{ij} \in \sigma_2$. Note that for every $\sigma \in Proj_{D'}[v(Y)]$, $0 \leq |\sigma_1| \leq |Y|$. We further divide the above probability into sums according to the size of σ_1 to get:

$$\begin{aligned}
\sum_{i=0}^{|Y|} \sum_{\substack{\sigma\in P(D,Y),\\ |\sigma_1| = |Y|-i}} & Pr[\delta((f\upharpoonright_\sigma)\upharpoonright_\rho) \geq s - |\sigma| \mid \\
&\qquad F\upharpoonright_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ \rho(T\setminus Y) = 1 \ \wedge \ \rho(\sigma) = *] && (3) \\
& \times Pr[\rho(\sigma) = * \mid \\
&\qquad F\upharpoonright_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ \rho(T\setminus Y) = 1 \ \wedge \ \rho(Y) = *] && (4)
\end{aligned}$$

For a fixed value of Y and $\sigma \in P(D,Y)$, we estimate the term (3). Let f' be f with the variables in $T\setminus Y$ set to 1. Let F' be $F \vee G$ where $G\upharpoonright_\rho = 0$ if and only if ρ sets all variables in $T\setminus Y$ to 1. ($G = \bigvee_{x\in(T\setminus Y)} \overline{x}$.) Then term (3) is equal to

$$Pr[\delta((f'\upharpoonright_\sigma)\upharpoonright_\rho) \geq s - |\sigma| \mid F'\upharpoonright_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ \rho(\sigma) = *].$$

First, notice that $f' \upharpoonright_\sigma$ is a map-disjunction which does not involve any variables which share vertices with $v(\sigma \cup Q)$. Also, $8(|Q| + |\sigma| + s - |\sigma| + r)^2 = 8(|Q| + s + r)^2$, and we know that $pn \geq 8(|Q| + s + r)^2$. Now if $\sigma = Y$, then f'_1 is satisfied by σ and $f\upharpoonright_\rho$ is the constant 1 and since $s > 2|Y|$ by assumption, this probability is $0 \leq \alpha^{s-|\sigma|}$. Otherwise, $\sigma \neq Y$, the map f'_1 is falsified by σ, so $f'\upharpoonright_\sigma$ has one fewer maps than the original f that we started with. Therefore, we can apply the inductive hypothesis where now Q is replaced by $Q \cup \sigma$, and D' is replaced by $D' \setminus v(\sigma) = D \setminus (v(Q) \cup v(\sigma))$, to upper bound the above quantity by $\alpha^{s-|\sigma|}$. Because $|\sigma| \leq 2|Y|$, for all σ, it follows that the above quantity is no greater than $\alpha^{s-2|Y|}$.

Since the above calculation gives the same upper bound for term (3) for all values of σ, we can pull this quantity outside the sum to obtain:

$$\alpha^{s-2|Y|} \sum_{i=0}^{|Y|} \sum_{\substack{\sigma \in P(D,Y), \\ |\sigma_1| = |Y|-i}} \begin{array}{l} Pr[\,\rho(\sigma) = * \mid \\ \quad F\upharpoonright_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ \rho(T \setminus Y) = 1 \ \wedge \ \rho(Y) = *] \end{array} \quad (5)$$

Now we will estimate the inner summation for a fixed value of i. As above, we replace the condition $F \upharpoonright_\rho = 0 \ \wedge \rho(T \setminus Y) = 1$ by the single condition $F'\upharpoonright_\rho = 0$. Also, for a particular σ, the event $\rho(\sigma) = *$ is equivalent to the events $\rho(\sigma_1) = * \ \wedge \rho(\sigma_2) = *$. Because $\rho(\sigma_1) = *$ is implied by $\rho(Y) = *$, the inner summation is equivalent to

$$\sum_{\substack{\sigma \in P(D,Y), \\ |\sigma_1| = |Y|-i}} Pr[\rho(\sigma_2) = * \mid F'\upharpoonright_\rho = 0 \ \wedge \ \rho(Q) = * \wedge \rho(Y) = *].$$

We would like to remove the conditioning on $F' \upharpoonright_\rho = 0$ but we will not be able to remove this condition separately for each individual term, as we did in Proposition 23. Instead, we have to consider the terms in this sum in the aggregate rather than individually. Let N_i be the number of σ's such that $|\sigma_1| = |Y| - i$. Then the above probability can be rewritten as:

$$N_i \cdot Pr_{(\sigma_2,\rho)}[\rho(\sigma_2) = * \mid F'\upharpoonright_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ \rho(Y) = *],$$

where the above probability is over all pairs (σ_2, ρ), such that $|\sigma_1| = |Y| - i$. For each σ_2, let u be the set of vertices in σ_2 which are not contained in $v(Y)$. Note that the number of domain vertices of u equals the number of range vertices of

u and is equal to i. Also note that for σ_2 chosen at random, u is a uniformly distributed set over $D'' = D' \setminus v(Y) = D \setminus (v(Y) \cup v(Q))$ having these properties. Thus, the above probability can be written as:

$$N_i \cdot Pr_{(u,\rho)}[\rho(u) = * \mid F'\lceil_\rho = 0 \ \wedge \ \rho(Q) = * \ \wedge \ \rho(Y) = *],$$

where u is a set of size i, chosen uniformly from $D'' = D' \setminus v(Y)$, and ρ is chosen from $\mathcal{R}_p^D$.

Let $Q' = Q \cup Y$. This probability can be further divided according to $\#(\rho)$, the exact number of stars that are assigned to D_1 by ρ:

$$N_i \cdot \sum_{j=0}^{n} Pr_{(u,\rho)}[\rho(u) = * \mid F'\lceil_\rho = 0 \ \wedge \ \rho(Q') = * \wedge \ \#(\rho) = j]$$
$$\times \ Pr_{(u,\rho)}[\#(\rho) = j \mid F'\lceil_\rho = 0 \ \wedge \rho(Q') = *].$$

Given that $\#(\rho) = j$ and $\rho(Q') = *$, the exact number of *'s in D'' is completely determined and therefore, for a randomly chosen u, the event $\rho(u) = *$ is independent of $F'\lceil_\rho = 0$. Thus the above probability is equal to

$$N_i \cdot \sum_{j=0}^{n} Pr_{(u,\rho)}[\rho(u) = * \mid \#(\rho) = j \ \wedge \ \rho(Q') = *]$$
$$\times \ Pr[\#(\rho) = j \mid F'\lceil_\rho = 0 \ \wedge \ \rho(Q') = *],$$

where we have dropped the subscript on the probability in the second factor in each term since this probability only depends on ρ. For all $k \le n$, $\sum_{j \ge k} Pr[\#(\rho) = j \mid F'\lceil_\rho = 0 \ \wedge \ \rho(Q') = *]$ equals $Pr[\#(\rho) \ge k \mid F'\lceil_\rho = 0 \ \wedge \ \rho(Q') = *]$, because the events are disjoint. Similarly, $\sum_{j \ge k} Pr[\#(\rho) = j \mid \rho(Q') = *]$ equals $Pr[\#(\rho) \ge k \mid \rho(Q') = *]$.

PROPOSITION 25. *For all* k, $Pr[\#(\rho) \ge k \mid F'\lceil_\rho = 0 \wedge \rho(Q) = *] \le Pr[\#(\rho) \ge k \mid \rho(Q) = *]$.

PROOF. As in the proof of Proposition 23, we will prove this inequality by showing that for all k, $Pr[F'\lceil_\rho = 0 | \#(\rho) \ge k \ \wedge \ \rho(Q) = *] \le Pr[F'\lceil_\rho = 0 \mid \rho(Q) = *]$. Let $F(C) = Pr[F'\lceil_\rho = 0 \mid \rho(Q) = *]$. Then $F(C)$ is a weighted average of $F(A)$ and $F(B)$, where $F(A) = Pr[F'\lceil_\rho = 0 \mid \#(\rho) \ge k \ \wedge \ \rho(Q) = *]$ and $F(B) = Pr[F'\lceil_\rho = 0 \mid \#(\rho) < k \ \wedge \ \rho(Q) = *]$. We want to show that $F(A) \le F(B)$, and then it follows that $F(A) \le F(C)$, as desired. Let $F(i) = Pr[F'\lceil_\rho = 0 \mid \#(\rho) = i \ \wedge \ \rho(Q) = *]$. Then $F(A)$ is a weighted average of terms $\{F(i),\ k \leq i \leq n\}$, and $F(B)$ is a weighted average of terms $\{F(i),\ 1 \leq i < k\}$. Thus, it suffices to show that for all k, $F(k) \leq F(k-1)$, which follows from Proposition 24. □

Using Proposition 25 and noting that $Pr_{(u,\rho)}[\rho(u) = * \mid \#(\rho) = j \wedge \rho(Q') = *] \leq Pr_{(u,\rho)}[\rho(u) = * \mid \#(\rho) = j+1 \wedge \rho(Q') = *]$ for all $j \geq 0$, we can apply Lemma 21 with $\alpha_j = Pr_{(u,\rho)}[\rho(u) = * \mid \#(\rho) = j \wedge \rho(Q') = *]$, $a_j = Pr[\#(\rho) = j \mid F'\restriction_\rho = 0 \wedge \rho(Q') = *]$, and $b_j = Pr[\#(\rho) = j \mid \rho(Q') = *]$ to show that the above probability is no greater than

$$N_i \cdot \sum_{j=0}^{n} Pr_{(u,\rho)}[\rho(u) = * \mid \#(\rho) = j \wedge \rho(Q') = *] \cdot Pr[\#(\rho) = j \mid \rho(Q') = *]$$

which is equal to $N_i \cdot Pr_{(u,\rho)}[\rho(u) = * \mid \rho(Q') = *]$.

Since for each fixed value of $u \in V_i$, the probability that $\rho(u) = *$ is the same, the above probability is equal to $N_i \cdot Pr[\rho(u) = * \mid \rho(Q') = *]$. Using the fact that $|u| \leq r$ and $|Q'| \leq q + r$, we have that $\frac{q+2r}{n-(q+2r)} \leq p$ because $pn \geq 8(q+s+r)^2$. Therefore, we can apply Lemma 19 to conclude that for $u \in V_i$, $Pr[\rho(u) = * \mid \rho(Q') = *] \leq (5p^2)^i$.

Recall that N_i is equal to the number of σ's such that $|\sigma_1| = Y - i$. Let $m = n - |Q'| = n - |Q| - |Y|$. There are at most $\binom{|Y|}{i}^2 (|Y| - i)!$ choices of σ_1 with $|\sigma_1| = |Y| - i$ and for each such σ_1 there are at most $\left(\frac{(m)!}{(m-i)!}\right)^2$ choices of σ_2. Thus there are a total of at most $\binom{|Y|}{i}^2 (|Y| - i)! \left(\frac{(m)!}{(m-i)!}\right)^2$ choices of $\sigma \in P(D, Y)$ such that $|\sigma_1| = |Y| - i$.

Thus for all Y such that $2|Y| \leq s$, using the expression in (5), we have

$$\begin{aligned}
&Pr[\delta(f\restriction_\rho) \geq s \mid F\restriction_\rho = 0 \wedge f_1\restriction_\rho \neq 0 \wedge *(\rho_T) = Y] \\
&\leq \sum_{i=0}^{|Y|} \binom{|Y|}{i}^2 (|Y| - i)! \left(\frac{m!}{(m-i)!}\right)^2 \alpha^{s-2|Y|}(5p^2)^i \\
&\leq \alpha^{s-2|Y|} \sum_{i=0}^{|Y|} \binom{|Y|}{i}^2 (|Y| - i)!(5p^2n^2)^i \\
&\leq \alpha^{s-2|Y|} \sum_{i=0}^{|Y|} \binom{|Y|}{i} (5p^2n^2)^i (|Y|)^{|Y|-i} \\
&= \alpha^{s-2|Y|}|Y|^{|Y|} \sum_{i=0}^{|Y|} \binom{|Y|}{i} (\frac{5p^2n^2}{|Y|})^i \\
&= \alpha^{s-2|Y|}|Y|^{|Y|}(\frac{5p^2n^2}{|Y|} + 1)^{|Y|}
\end{aligned}$$

$$\begin{aligned}
&\leq \alpha^{s-2|Y|}|Y|^{|Y|}(\frac{6p^2n^2}{|Y|})^{|Y|} \\
&\leq \alpha^{s-2|Y|}(6p^2n^2)^{|Y|}.
\end{aligned}$$

For Y such that $2|Y| \geq s$ we cannot use the expansion in terms of (3) and (4) to estimate this probability. However in this case, since $\alpha \leq 1$ and $6p^2n^2 \geq 1$, $2\alpha^{s-2|Y|}(6p^2n^2)^{|Y|} \geq 1$ so it still is an upper bound on this probability.

Plugging in the bounds we have for the terms (1) and (2) we get

$$\begin{aligned}
&Pr[\delta(f\lceil_\rho) \geq s \mid F\lceil_\rho = 0 \ \wedge \ f_1\lceil_\rho \neq 0] \\
&\leq \sum_{\substack{Y\subseteq T,\\ Y\neq\emptyset}} \alpha^{s-2|Y|}(6p^2n^2)^{|Y|}(6p^2n)^{|Y|} \\
&= \alpha^s \sum_{\substack{Y\subseteq T,\\ Y\neq\emptyset}} \left(\frac{36p^4n^3}{\alpha^2}\right)^{|Y|} \\
&\leq \alpha^s \sum_{i=1}^{r} \binom{r}{i} \left(\frac{36p^4n^3}{\alpha^2}\right)^i \\
&= \alpha^s \left[\left(1+\frac{36p^4n^3}{\alpha^2}\right)^r - 1\right] \\
&\leq \alpha^s
\end{aligned}$$

The last inequality holds since α satisfies $(1+36p^4n^3/\alpha^2)^r \leq 2$. This completes the proof of Lemma 22. □

6.1. A Modified Switching Lemma. A difficulty in applying the Switching Lemma from this section is that the underlying distribution is $\mathcal{R}$, whereas for the pigeonhole formulas, we require that the restrictions be chosen from the distribution $\mathcal{Q}$, where there is one extra domain element. Although we can easily modify the proof of the Switching Lemma from this section to do this directly (and even show this for similar restrictions on pairs of sets whose size differs by any fixed constant c), we can derive this directly from Lemma 18 . We assume in this section that $D^* = D_0^* \cup D_1^*$, $|D_0^*| = n+1$ and $|D_1^*| = n$.

LEMMA 5 (THE PIGEONHOLE SWITCHING LEMMA) *Let f be an r-disjunction over D^*. Choose ρ at random from $\mathcal{Q}_p^{D^*}$. For $s \geq 0$, $p \leq 1/36$, and $pn \geq 8(s+r)^2$ we have*

$$Pr^{\mathcal{Q}}[\delta(f\lceil_\rho) \geq s+1] < n(\alpha^s),$$

for any $\alpha > 0$ satisfying $(1+\frac{36p^4n^3r}{\alpha^2})^r \leq 2$.

PROOF. Let $First(\rho) = j$ be the event that the first component of $\rho =< i, S_0, S_1, \pi >$, i, be equal to j. We will analyze the probability by considering separately the cases where $First(\rho) = j$, for all $j \in D_0^*$.

$$Pr^{\mathcal{Q}}[\delta(f\lceil_\rho) \geq s+1] \;\leq\; \max_{i \in D_0^*}\{Pr^{\mathcal{Q}}[\delta(f\lceil_\rho) \geq s+1 \mid First(\rho) = i]\}.$$

Fix a particular value of $i \in D_0$. For a particular $\sigma \in Proj_D[i]$, let $f_\sigma = f\lceil_\sigma$. Then by Lemma 4, the above probability is less than or equal to $\sum_{\sigma \in Proj_D[i]} Pr^{\mathcal{Q}}[\delta(f_\sigma\lceil_\rho) \geq s \mid First(\rho) = i]$. Fix a particular $\sigma \in Proj_D[i]$. Then by Lemma 2 the associated probability is no greater than $Pr^{\mathcal{R}}[\delta(f_\sigma\lceil_\rho) \geq s]$, where now the probability is over the distribution $\mathcal{R}_p^D$, where $D = D^* \backslash i$. We can now directly apply the Switching Lemma (Lemma 18) to obtain an upper bound of α^s for this probability. Because $|Proj_D[i]| \leq n$, the total probability is no greater than $n(\alpha^s)$. □

Acknowledgements

The results in this paper have appeared as part of the first author's Ph.D. thesis, under the supervision of Stephen Cook at the University of Toronto. The first author would like to take this opportunity to thank Stephen Cook for the profound impact that he had on this research, as well as for his careful reading and helpful comments on this work. We also wish to thank Alasdair Urquhart and Stephen Bellantoni for many valuable conversations. The second author was partially funded by NSF grants CCR-8858799 and CCR-8907960.

References

M. AJTAI, The complexity of the pigeonhole principle, in *Proc. 29th Ann. IEEE Symp. Foundations of Computer Science*, 1988, 346–355.

M. AJTAI, Σ_1^1-Formulae on finite structures, *Annals of Pure and Applied Logic*, **24** (1983), 1–48.

P. BEAME, Lower bounds for recognizing small cliques on CRCW PRAM's, *Discrete Applied Mathematics*, **29** (1990), 3–20.

P. BEAME, J. HÅSTAD, Optimal bounds for decision problems on the CRCW PRAM, *Journal of the ACM*, **36** (1989), 643–670.

P. BEAME, R. IMPLAGIAZZO, J. KRAJÍČEK, T. PITASSI, P. PUDLÁK, A. WOODS, Exponential lower bounds for the pigeonhole principle, *Proc. 24th Ann. ACM Symp. Theory of Computing*, 1992, 200–220.

S. BELLANTONI, T. PITASSI, A. URQUHART, Approximation and small-depth Frege proofs, *SIAM Journal of Computing*, **21** (1992), 1161-1179.

M. BONET AND S. BUSS, The deduction rule and linear and near-linear proof simulations, preprint 1992.

S. BUSS, Polynomial size proofs of the propositional pigeonhole principle, *Journal of Symbolic Logic*, **52** (1987), 916–927.

S. BUSS, Personal communication, 1993.

S. A. COOK AND R. RECKHOW, The relative efficiency of propositional proof systems, *Journal of Symbolic Logic*, **44** (1979), 36–50.

M. FURST, J. SAXE, M. SIPSER, Parity, circuits and the polynomial time hierarchy, *Mathematical Systems Theory*, **17** (1984) 13–27.

A. HAKEN, The intractability of Resolution, *Theoretical Computer Science* **39** (1985) 297–308.

J. HÅSTAD, *Computational limitations of small-depth circuits*, The MIT Press, Cambridge, Massachusetts, 1987.

J. KRAJÍČEK, Lower bounds to the size of constant-depth propositional proofs, preprint (1991).

J. KRAJÍČEK, P. PUDLÁK, A. WOODS, Exponential lower bounds to the size of bounded-depth Frege proofs of the pigeonhole principle, preprint (1991).

J. LYNCH, A depth-size tradeoff for Boolean circuits with unbounded fan-in, *Lecture Notes in Computer Science* **223** (1986), 234-248.

J. PARIS, A. WILKIE, A. WOODS, Provability of the pigeonhole principle and the existence of infinitely many primes, *Journal of Symbolic Logic*, **53** Number 4 (1988).

T. PITASSI, P. BEAME, R. IMPAGLIAZZO, Exponential lower bounds for the pigeonhole principle, University of Toronto TR 257/91, (1991).

G. S. TSEITIN, On the complexity of derivation in the propositional calculus, *Studies in Constructive Mathematics and Mathematical Logic*, Part II, A.O. Slisenko, 1968.

A. URQUHART, Hard examples for Resolution, *JACM*, **34** (1987), 209–219.

A. C. YAO, Separating the polynomial-time hierarchy by oracles, *Proc. 26th Ann. IEEE Symp. Foundations of Computer Science*, 1985, 1–10.

Manuscript received 11 May 1992

TONIANN PITASSI
Department of Computer Science
University of California at San Diego
La Jolla, CA
toni@cs.ucsd.edu

PAUL BEAME
Dept. of Computer Science & Engineering
University of Washington
Seattle, WA
beame@cs.washington.edu

Russell Impagliazzo
Department of Computer Science
University of California at San Diego
La Jolla, CA
russell@cs.ucsd.edu