Lower Bounds for the Weak Pigeonhole Principle Beyond Resolution

Albert Atserias* María Luisa Bonet†
Juan Luis Esteban‡

Departament de Llenguatges i Sistemes Informàtics
Universitat Politècnica de Catalunya
C/Jordi Girona Salgado, 1-3, Edif. C6.
08034 Barcelona - Spain.
Tel: +34 93 401 69 94
Fax: +34 93 401 70 14
{atserias,bonet,esteban}@lsi.upc.es.§

April 2, 2001

Abstract

We work with an extension of Resolution, called Res(2), that allows clauses with conjunctions of two literals. In this system there are rules to introduce and eliminate such conjunctions. We prove that the weak pigeonhole principle PHP_n^{cn} and random unsatisfiable CNF formulas require exponential-size proofs in this system. This is the strongest system beyond Resolution for which such lower bounds are known. As a consequence to the result about the weak pigeonhole principle, Res(log) is exponentially more powerful than Res(2). Also we prove that Resolution cannot polynomially simulate Res(2), and that Res(2) does not have feasible monotone interpolation solving an open problem posed by Krajíček.

*Supported by the CUR, Generalitat de Catalunya, through grant 1999FI 00532.
†Partially supported by MEC through grant PB98-0937-C04 (FRESCO project) and CICYT TIC 98-0410-C02-01.
‡Partially supported by MEC through grant PB98-0937-C04 (FRESCO project)
§Partially supported by ALCOM-FT, IST-99-14186.

1 Introduction

The pigeonhole principle, PHP_n^{n+1}, expresses that it is not possible to have a one-to-one mapping from $n+1$ pigeons to n holes. Since it can be formalized in propositional logic, it is natural to ask in which propositional proof systems such a principle can be proved in polynomial-size, with respect to the size of the encoding.

A fair amount of information is known about sizes of proofs of PHP_n^{n+1} in various proof systems. Haken [12] proved that this principle requires exponential-size proofs in Resolution. His proof techniques were later extended and simplified [4, 5]. Also Beame et al. [2] proved that PHP_n^{n+1} requires exponential-size proofs in bounded-depth Frege systems. Regarding upper bounds, Buss [8] gave polynomial-size proofs of PHP_n^{n+1} in unrestricted Frege systems.

The pigeonhole principle can be formulated in more general terms, allowing the number of pigeons to be greater than $n+1$. We call this principle weak pigeonhole principle, or PHP_n^m, when the number of pigeons m is at least $2n$. This simple principle is central to many mathematical arguments but quite often, it occurs implicitely only. See the introductions in [14, 16] for a nice discussion on this. The proof techniques of Haken where extended in [9] to prove that $\mathrm{PHP}_n^{n^{2-\epsilon}}$ requires exponential-size proofs in Resolution. A very intriguing and often studied open problem is to prove exponential-size lower bounds for Resolution proofs of PHP_n^m for any m. As a contrast, the techniques of [2] for proving lower bounds for the pigeonhole principle in bounded-depth Frege systems can only prove lower bounds for PHP_n^{n+c}, and it is again open whether lower bounds can be proved when the number of pigeons in greater than $n+c$. Regarding upper bounds, it is known that PHP_n^{2n} has quasipolynomial-size proofs in bounded-depth Frege [15, 14].

We work with the proof system $\mathrm{Res}(2)$, proposed by Krajíček [13], that can be viewed either as an extension of Resolution, or as a restriction of bounded-depth Frege. In this system the clauses do not only contain literals, but can also have conjunctions of two literals. The resolution rule gets modified to be able to eliminate a conjunction of two literals from a clause. We prove that PHP_n^{cn} (and in fact $\mathrm{PHP}_n^{n^{9/8-\epsilon}}$) requires exponential-size proofs in Res(2). This is, to our knowledge, the first lower bound proof for the weak pigeonhole principle in a subsystem of bounded-depth Frege that extends Resolution. We note that the quasipolynomial upper bound for bounded-depth Frege mentioned above can be carried over in depth-0.5 LK [14], which is equivalent to $\mathrm{Res}(\log)$ (the analogue of $\mathrm{Res}(2)$ when we allow conjunctions of up to polylog literals). As a consequence of our lower bound, there is an exponential separation between $\mathrm{Res}(2)$ and $\mathrm{Res}(\log)$.

We also consider the complexity of refuting random unsatisfiable k-CNF formulas. Chvátal and Szemerédi [10] proved them hard to refute in Resolution, and the results were improved by Beame, Karp, Pitassi and Saks [3]. Combining our techniques with those of [3], we also obtain an exponential-size lower bound for $\mathrm{Res}(2)$-refutations of random unsatisfiable k-CNF formulas with clause density near the threshold. Again, this is the strongest system beyond Resolution for

which such a lower bound is known. This result may be considered as a first step towards proving random k-CNF formulas hard for bounded-depth Frege.

Our techniques are based on the method of random restrictions. The main technical contribution of our work consists in proving that a relatively short random restriction kills all large formulas of a $\mathrm{Res}(2)$-refutation. We note that this task is trivial in the case of Resolution because a large clause is killed by setting a single literal to one. However, our formulas are disjunctions of conjunctions of two literals, and this task becomes much more involved. The difficulty is in the fact that we must keep the restriction short not to trivialize the initial clauses of the refutation. In other words, we overcome the main difficulty in trying to apply switching-like lemmas to prove lower bounds for the weak pigeonhole principle or random formulas.

Another important question to ask is whether $\mathrm{Res}(2)$ is more powerful than Resolution. Here we prove that Resolution cannot polynomially simulate $\mathrm{Res}(2)$, and therefore $\mathrm{Res}(2)$ is superpolynomially more efficient than Resolution. As a corollary, we see that $\mathrm{Res}(2)$ does not have feasible monotone interpolation, solving this way a conjecture of Krajíček [13].

Another motivation for working with the system $\mathrm{Res}(2)$ is to see how useful it can be in automated theorem proving. Given that it is more efficient than Resolution (at least there is a superpolynomial separation), it might be a good idea to try to find good heuristics to find proofs in $\mathrm{Res}(2)$ to be able to use it as a theorem prover.

2 Definitions and Overview of the Lower Bound Proof

A k-*term* is a conjunction of up to k literals. A k-*disjunction* is an (unbounded fan-in) disjunction of k-terms. If F is a k-disjunction, a 1-term of F is also called a *free-literal*. The refutation system $\mathrm{Res}(k)$, defined by Krajíček [13], works with k-disjunctions. There are three inference rules in $\mathrm{Res}(k)$: Weakening, $\wedge$-Introduction, and Cut

$$\frac{A}{A \vee \bigwedge_{i\in I} l_i} \qquad \frac{A \vee \bigwedge_{i\in I} l_i \qquad B \vee \bigwedge_{i\in J} l_i}{A \vee B \vee \bigwedge_{i\in I\cup J} l_i} \qquad \frac{A \vee \bigwedge_{i\in I} l_i \qquad B \vee \bigvee_{i\in I} \bar{l}_i}{A \vee B}$$

where A and B are k-disjunctions, I, J are sets of indices such that $|I \cup J| \leq k$, and the l_i's are literals. As usual, if l is a literal, $\bar{l}$ denotes its negation. Observe that $\mathrm{Res}(1)$ coincides with Resolution with the Weakening rule. The size of a $\mathrm{Res}(k)$-refutation is the number of symbols in it. Mainly, we will work with $\mathrm{Res}(2)$.

As we mentioned in the introduction, our arguments are based on random restrictions. In general terms, what we do is the following. Given an unsatisfiable CNF formula F, and an alleged small $\mathrm{Res}(2)$-refutation P of F, we apply a random restriction ρ, from a suitable distribution, and we get a refutation $P|_\rho$ of $F|_\rho$. The distribution on restrictions that we choose will satisfy the following two properties:

(i) $F|_\rho$ satisfies certain expansion properties,

(ii) Every 2-disjunction in $P|_\rho$ is short (measured by the number of literals that occur).

The argument will be complete since these two conditions will be shown to be contradictory.

As a contrast with the lower bound arguments for Resolution, the most difficult part of our proof is showing that property (ii) is satisfied. The conjunctions make this task more involved. In order to overcome this, we split the restriction into two parts $\rho = \rho_1\rho_2$. Then, the main contribution is showing that every large clause in $P|_{\rho_1}$ contains many free literals. That allows us show, by a standard argument, that no large clause remains in $P|_{\rho_1\rho_2}$.

For the sake of clarity of exposition, we explain this outline again in the particular case of the Weak Pigeonhole Principle. Let $G = (U \cup V, E)$ be a bipartite graph on the sets U and V of cardinality m and n respectively, where $m > n$. The G-PHP_n^m, defined by Ben-Sasson and Wigderson [5], states that there is no matching of U into V. For every edge $(u, v) \in E$, let $x_{u,v}$ be a propositional variable meaning that u is mapped to v. The principle is then formalized as the conjunction of the following set of clauses:

$$x_{u,v_1} \vee \cdots \vee x_{u,v_r} \qquad u \in U,\ N_G(u) = \{v_1, \ldots, v_r\} \tag{1}$$

$$\bar{x}_{u,v} \vee \bar{x}_{u',v} \qquad v \in V,\ u, u' \in N_G(v),\ u \neq u'. \tag{2}$$

Here, $N_G(w)$ denotes the set of neighbors of w in G. Observe that if G is the complete bipartite graph K_n^m, then G-PHP_n^m coincides with the usual pigeonhole principle PHP_n^m. It is easy to see that a lower bound for the size of $\mathrm{Res}(2)$-refutations of G-PHP_n^m implies the same lower bound for the size of $\mathrm{Res}(2)$-refutations of PHP_n^m.

Ben-Sasson and Wigderson proved that whenever G is expanding in a sense defined next, every Resolution refutation of G-PHP_n^m must contain a clause with many literals. We observe that this result is not unique to Resolution and holds in a more general setting. Before we state the precise result, let us recall the definition of expansion:

Definition 1 [5] *Let $G = (U \cup V, E)$ be a bipartite graph where $|U| = m$, and $|V| = n$. For $U' \subset U$, the* boundary *of U, denoted by $\partial U'$, is the set of vertices in V that have exactly one neighbor in U'; that is, $\partial U' = \{v \in V : |N(v) \cap U'| = 1\}$. We say that G is (m, n, r, f)-*expanding *if every subset $U' \subseteq U$ of size at most r is such that $|\partial U'| \geq f \cdot |U'|$.*

The proof of the following statement is the same as in [5] for Resolution.

Theorem 1 [5] *Let $\mathcal{S}$ be a sound refutation system with all rules having fan-in at most two. Then, if G is (m, n, r, f)-expanding, every $\mathcal{S}$-refutation of G-PHP_n^m must contain a formula that involves at least $rf/2$ distinct literals.*

With these definitions, we are ready to outline the argument of the lower bound proof. In section 3.1, we will prove the existence of a bipartite graph $G = (U \cup V, E)$ with $|U| = cn'$ and $|V| = n'$ such that if we remove a small random subset of nodes from V, and the corresponding edges, the resulting graph is (m, n, r, f)-expanding for certain m, n, r and f. Then we will argue that G-$\mathrm{PHP}_{n'}^{cn'}$ requires exponential-size $\mathrm{Res}(2)$-refutations as follows. Assume, for contradiction, that Π is a small refutation of G-$\mathrm{PHP}_{n'}^{cn'}$. We say that a 2-disjunction in Π is large if it contains at least $d = rf/2$ distinct literals. We apply a random restriction ρ_1 to the refutation such that for every large C, either $C|_{\rho_1}$ contains many free literals, or the total number of literals in $C|_{\rho_1}$ is less than d. Then we extend ρ_1 to a new random restriction $\rho \supseteq \rho_1$ that knocks out all those large C such that $C|_{\rho_1}$ contains many free literals, ignoring those that are not free. After applying ρ, we obtain a refutation of $G(\rho)$-PHP_n^m where all 2-disjunctions have less than $rf/2$ literals and $G(\rho)$ is (m, n, r, f)-expanding. This contradicts Theorem 1.

3 Lower Bound for the Weak Pigeonhole Principle

3.1 Random Graphs and Restrictions

In this section we will prove the existence of a bipartite graph G as claimed in Section 2. The principle G-PHP_n^m will require exponential size $\mathrm{Res}(2)$-proofs.

Let $\mathcal{G}(m, n, p)$ denote the distribution on bipartite graphs on sets U and V of sizes m and n respectively, with edge probability p independently for each edge.

Lemma 1 *If G is drawn from $\mathcal{G}(m, n, p)$, then* $\Pr\left[\forall v \in V : mp/2 < \deg_G(v) < 2mp\right] \geq 1 - 2ne^{-\frac{mp}{8}}$.

Proof: Fix a vertex $v \in V$. Then, $\deg_G(v) \sim \mathrm{Bin}(m, p)$, so that $\mathrm{E}\left[\deg_G(v)\right] = mp$. By Chernoff bounds, $\Pr\left[\deg_G(v) \geq 2mp\right] \leq e^{-mp/3}$ and $\Pr\left[\deg_G(v) \leq mp/2\right] \leq e^{-mp/8}$. By a union bound, $\Pr\left[\exists v \in V : \deg_G(v) \leq mp/2 \vee \deg_G(v) \geq 2mp\right] \leq ne^{-mp/3} + ne^{-mp/8} \leq 2ne^{-mp/8}$, and so $\Pr\left[\forall v \in V : mp/2 < \deg_G(v) < 2mp\right] \geq 1 - 2ne^{-mp/8}$. □ (of lemma 1)

Lemma 2 *Let $m = kn$, $p = 48k\ln(m)/m$, $\alpha = 1/mp$ and $f = np/6$. Let G be drawn from $\mathcal{G}(m, n, p)$. Then,* $\Pr\left[G \text{ is } (m, n, \alpha m, f)\text{-expanding}\right] \geq 1/2$.

Proof: Fix $U' \subseteq U$ of size $s \leq \alpha m$, and $v \in V$. Then, $\Pr\left[v \in \partial U'\right] = sp(1-p)^{s-1}$. Let $q = \Pr\left[v \in \partial U'\right]$. Let X_v be the indicator random variable for the event that $v \in \partial U'$. Then, $|\partial U'| = \sum_{v \in V} X_v$. Observe that X_v and $X_{v'}$ are independent whenever $v \neq v'$. Hence, $|\partial U'| \sim \mathrm{Bin}(n, q)$, so that $\mathrm{E}\left[|\partial U'|\right] = nq$. By Chernoff bound, $\Pr\left[|\partial U'| \leq nq/2\right] \leq e^{-nq/8}$. On the other hand, $nq = nsp(1-p)^{s-1} \geq snp(1-p)^{\alpha m}$. Moreover, $(1-p)^{\alpha m} = (1-p)^{1/p}$ approaches $1/e$

for sufficiently large m. Therefore, $nq \geq snp/3$. It follows that $nq/2 \geq sf$ and $e^{-nq/8} \leq e^{-snp/24}$. We conclude that $\Pr\left[|\partial U'| < f \cdot |U'|\right] \leq \Pr\left[|\partial U'| \leq nq/2\right] \leq e^{-nq/8} \leq e^{-snp/24}$. Finally, we bound the probability that G is not $(m, n, \alpha m, f)$-expanding by

$$\sum_{s=1}^{\alpha m} \binom{m}{s} e^{-snp/24} \leq \sum_{s=1}^{\alpha m} m^s e^{-snp/24} \leq \sum_{s=1}^{\alpha m} (me^{-np/24})^s. \tag{3}$$

Recall that $p = 48k\ln(m)/m$ and $m = kn$. So $me^{-np/24} \leq me^{-2\ln(m)} = m^{-1} < 1/4$. Hence the sum in (3) is bounded by $\sum_{s=1}^{\infty} \frac{1}{4^s} \leq \frac{1}{2}$. □ (of lemma 2)

Let G be a fixed bipartite graph on $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$. A *restriction* (for G) is a sequence of pairs $\rho = ((u_1, v_1), \ldots, (u_r, v_r))$ such that $(u_i, v_i) \in E(G)$, and all v_i's are distinct. We let $R_r(G)$ be the set of restrictions of length r. We define a distribution $\mathcal{R}_r(G)$ on $R_r(G)$ as follows: Let $V_0 = \{1, \ldots, n\}$; for every $i \in \{1, \ldots, r\}$ in increasing order, choose a hole v_i uniformly at random in V_{i-1}, choose a pigeon u_i uniformly at random in $N_G(v_i)$, and let $V_i = V_{i-1} - \{v_i\}$. The final restriction is $((u_1, v_1), \ldots, (u_r, v_r))$.

We define a distribution $\mathcal{D}(m, n, p, r)$ on the set of pairs (G, ρ) with $\rho \in R_r(G)$: the graph G is drawn from $\mathcal{G}(m, n + r, p)$ first, and then ρ is drawn from $\mathcal{R}_r(G)$. In other words, if (H, π) is a fixed pair with $\pi \in R_r(H)$, then

$$\Pr\left[G = H \wedge \rho = \pi\right] = p^{e(H)}(1-p)^{m(n+r)-e(H)}|R_r(H)|^{-1}.$$

If G is a bipartite graph on the vertex sets $\{1, \ldots, m\}$ and $\{1, \ldots, n + r\}$, and ρ is a restriction $((u_1, v_1), \ldots, (u_r, v_r)) \in R_r(G)$, then $G(\rho)$ denotes the graph that results from deleting $v_1, \ldots, v_r$ from G, and renaming nodes in an order-preserving way. With this definitions we are ready to prove:

Lemma 3 *Let $m = kn$, $p = 48k\ln(m)/m$, $\alpha = 1/mp$ and $f = np/6$. Let (G, ρ) be drawn from $\mathcal{D}(m, n, p, r)$. Then, $\Pr\left[G(\rho)$ is $(m, n, \alpha m, f)$-expanding$\right] \geq 1/2$.*

Proof: Let A be the event that $G(\rho)$ is $(m, n, \alpha m, f)$-expanding. Let $S = \{R \subseteq \{1, \ldots, n + r\} : |R| = r\}$. Then, $\Pr\left[A\right] = \sum_{R \in S} \Pr\left[A \mid \mathrm{ran}(\rho) = R\right] \Pr\left[\mathrm{ran}(\rho) = R\right]$. The proof that $\Pr\left[A \mid \mathrm{ran}(\rho) = R\right] \geq 1/2$ is the same as the proof of Lemma 2 replacing V by $V - R$. The result follows. □ (of lemma 3)

Lemma 4 *Let $m = kn$, $p = 48k\ln(m)/m$, $\alpha = 1/mp$ and $f = np/6$. For every $r \leq n$, there exists a bipartite graph H on $\{1, \ldots, m\}$ and $\{1, \ldots, n+r\}$ such that the following two properties hold:*

(i) $mp/2 \leq \deg_H(v) \leq 2mp$ *for every* $v \in \{1, \ldots, n + r\}$,

(ii) $\Pr\left[H(\rho)$ *is* $(m, n, \alpha m, f)$*-expanding*$\right] \geq 1/3$,

when ρ is drawn from $\mathcal{R}_r(H)$.

Proof: Let (G,ρ) be drawn from $\mathcal{D}(m,n,p,r)$. We have $\Pr\left[G(\rho) \text{ is } (m,n,\alpha m,f)\text{-expanding}\right] \geq 1/2$ by Lemma 3. Moreover, $\Pr\left[\forall v \in V : mp/2 < \deg_G(v) < 2mp\right] \geq 1-(n+r)e^{-mp/9} \geq 5/6$ by Lemma 1. Let $E(G,\rho)$ be the event that $G(\rho)$ is expanding *and* every right-node in G has degree between $mp/2$ and $2mp$. Combining both equations above we have that $\Pr\left[E(G,\rho)\right] \geq 1/3$. On the other hand, $\Pr\left[E(G,\rho)\right] = \sum_H \Pr\left[E(G,\rho) \mid G = H\right]\Pr\left[G = H\right]$ where H ranges over all bipartite graphs on m and $n+r$ nodes. Therefore, there exists some fixed H such that $\Pr\left[E(G,\rho) \mid G = H\right] \geq 1/3$. Moreover, $\Pr\left[E(G,\rho) \mid G = H\right]$ equals $\Pr\left[E(H,\pi)\right]$ when π is drawn from $\mathcal{R}_r(H)$. Finally, since this probability is strictly positive, it must be the case that H satisfies property (i) in the lemma since it is independent of π. □ (of lemma 4)

3.2 The Lower Bound Argument

Before we state and prove our main theorem, we will give some definitions and lemmas.

Let us first give a normal form for $\mathrm{Res}(2)$-refutations of G-PHP_n^m. We claim that every $\mathrm{Res}(2)$-refutation of G-PHP_n^m can be turned into a $\mathrm{Res}(2)$-refutation of similar size in which no 2-term is of the form $x_{u,v} \wedge x_{u',v}$ with $u \neq u'$. To check this, observe that such a 2-term must have been introduced at some point by the rule of $\wedge$-introduction with, say, $A \vee x_{u,v}$ and $B \vee x_{u',v}$. Cutting them with the axiom $\bar{x}_{u,v} \vee \bar{x}_{u',v}$ we get $A \vee B$ that can be used to continue the proof because it subsumes $A \vee B \vee (x_{u,v} \wedge x_{u',v})$.

Let C be a 2-disjunction, and let $(u,v) \in E(G)$. We let $C|_{(u,v)}$ be the result of assigning $x_{u,v} = 1$ and $x_{u',v} = 0$ for every $u' \in N_G(v) - \{u\}$ to C, and simplifying as much as possible. This includes replacing subformulas of the form $l \vee (l \wedge l')$ by l, and subformulas of the form $\bar{l} \vee (l \wedge l')$ by $\bar{l} \vee l'$ in some specified order; here l and l' are literals. Given a restriction $\rho = ((u_1,v_1),\ldots,(u_r,v_r))$, we let $C|_\rho$ be the result of applying $(u_1,v_1),\ldots,(u_r,v_r)$ successively in this order. For every $i \in \{1,\ldots,r\}$, we let $\rho_i = ((u_1,v_1),\ldots,(u_i,v_i))$.

Let us now study in more detail the result of applying a pair of a restriction to a 2-disjunction. First we give some definitions. We say that a pair $(u,v) \in E(G)$ *hits* C if either $x_{u,v}$ occurs positively in C, or $x_{u',v}$ occurs negatively in C for some $u' \in N_G(v) - \{u\}$. Equivalently, (u,v) hits C if it sets some literal of C to 1. If the literal is free, it knocks out the 2-disjunction. If the literal is part of a conjunction, it will locally create a free literal. In general, we say that $(u,v) \in E(G)$ *knocks* C if $C|_{(u,v)} \equiv 1$. We say that $(u,v) \in E(G)$ is a *bad choice* for C if it does not knock it and there exists $u' \in N_G(v) - \{u\}$ such that (u',v) knocks C. A bad choice may or may not be a hit.

Lemma 5 *Let C be a simplified 2-disjunction, and $(u,v) \in E(G)$. If (u,v) hits C and is not a knock or a bad choice, then $C|_{(u,v)}$ has more free literals than C.*

Proof: First notice that the literals that (u, v) sets to 1 are in a conjunction, otherwise (u, v) is a knock. Such literals can appear positive or negative. We will discuss the two cases:

(i) The literal is $x_{u,v}$ and appears in a conjunction of the form $x_{u,v} \wedge y$. The pair (u, v) does not set y to 1 otherwise we would have a knock. Also, it does not set it to 0 either, otherwise $y = x_{u',v}$ and such a conjunction is not allowed in the normal form. On the other hand, y does not appear free because C is a simplified 2-disjunction. Finally no free literal of C desapears when we apply (u, v) to C, otherwise (u, v) would be a bad choice.

(ii) The literal is $\overline{x}_{u',v}$, and it appears in a conjunction of the form $\overline{x}_{u',v} \wedge y$. Because (u, v) is not a knock, it does not set y to 1. Also, (u, v) does not set y to 0 either, otherwise it would be a bad choice, given that the indegree of v is 3 or more. As in the previous case and for the same reasons, y does not appear free in C, and no free literal of C desapears when we apply (u, v).

The lemma follows. □

Theorem 2 *Let $c > 1$ be a constant. For all sufficiently large n, every $\mathrm{Res}(2)$-refutation of PHP_n^{cn} has size at least $e^{n/(\log n)^{14}}$.*

Proof: Let $k = c+1$, $r = n/c$, $n' = n + r$, and $m = kn = cn'$. Let $G = (U \cup V, E)$ with $|U| = m$ and $|V| = n + r$ be the bipartite graph of Lemma 4. We show that every $\mathrm{Res}(2)$-refutation of G-PHP has size at least $e^{n/(\log n)^{14}}$. This will imply the Theorem since a $\mathrm{Res}(2)$-refutation of $\mathrm{PHP}_{n'}^{cn'}$ gives a $\mathrm{Res}(2)$-refutation of G-PHP of no bigger size. Let us assume, for contradiction, that G-PHP has a $\mathrm{Res}(2)$-refutation Π of size $S < e^{n/(\log n)^{14}}$.

We will use the following concepts. We say that C is *large* if it contains at least $d = n/12$ distinct literals; otherwise, C is *small*. We say that C is *wide* if it contains at least $s = n/(\log n)^5$ free literals; otherwise, C is *narrow*.

In all probabilities that follow, ρ is drawn from the distribution $\mathcal{R}_r(G)$. Our main goal is to prove that the probability that a fixed 2-disjunction C of Π remains large is exponentially small; that is, we aim for a proof that

$$\Pr\left[C|_\rho \text{ is large}\right] \leq e^{-n/(\log n)^{13}}. \tag{4}$$

This will suffice because then $\Pr\left[\exists C \in \Pi : C|_\rho \text{ is large}\right] \leq Se^{-n/(\log n)^{13}} < 1/3$, and also $\Pr\left[G(\rho) \text{ not } (m, n, \alpha m, f)\text{-expanding}\right] \leq 2/3$ by Lemma 4. This means that there exists a restriction $\rho \in R_r(G)$ such that $G(\rho)$ is $(m, n, \alpha m, f)$-expanding and every 2-disjunction in $\Pi|_\rho$ has less than $d = \alpha m f/2$ literals. This is a contradiction with Theorem 1.

For $i \in \{1, \ldots, r\}$, let A_i be the event that $C|_{\rho_i}$ is large, and let B_i be the event that $C|_{\rho_i}$ is narrow. Recall that $\rho_i = ((u_1, v_1), \ldots, (u_i, v_i))$. Then,

$$\begin{aligned}
\Pr\left[C|_\rho \text{ is large}\right] &\leq \Pr\left[A_r \wedge \bigvee_{j \geq r/2} B_j\right] + \Pr\left[A_r \wedge \bigwedge_{j \geq r/2} \overline{B_j}\right] \leq \\
&\leq \sum_{j=r/2}^{r} \Pr\left[A_j \wedge B_j\right] + \Pr\left[A_r \wedge \bigwedge_{j \geq r/2} \overline{B_j}\right].
\end{aligned}$$

We will show that every term in this expression is exponentially small. The bound on terms of the form $\Pr\left[A_j \wedge B_j\right]$ will be proven in Lemma 7. For the last term, we use an argument similar in spirit to the one by Beame and Pitassi [4]:

Lemma 6 $\Pr\left[A_r \wedge \bigwedge_{j \geq r/2} \overline{B_j}\right] \leq e^{-n/(\log n)^8}$.

Proof: Let S_i be the indicator random variable for the event that (u_i, v_i) knocks $C|_{\rho_{i-1}}$. Then,

$$\begin{aligned}
\Pr\left[A_r \wedge \bigwedge_{j \geq r/2} \overline{B_j}\right] &\leq \Pr\left[\bigwedge_{i > r/2} S_i = 0 \wedge \bigwedge_{j \geq r/2} \overline{B_j}\right] = \\
&= \prod_{i > r/2} \Pr\left[S_i = 0 \wedge \bigwedge_{j \geq r/2} \overline{B_j} \;\middle|\; \bigwedge_{r/2 < j < i} S_j = 0\right] \leq \\
&\leq \prod_{i > r/2} \Pr\left[S_i = 0 \wedge \overline{B_{i-1}} \;\middle|\; \bigwedge_{r/2 < j < i} S_j = 0\right] \leq \\
&\leq \prod_{i > r/2} \Pr\left[S_i = 0 \;\middle|\; \overline{B_{i-1}} \wedge \bigwedge_{r/2 < j < i} S_j = 0\right].
\end{aligned}$$

Fix $i \in \{r/2 + 1, \ldots, r\}$ and let H be the set of holes that occur in a free literal of $C|_{\rho_{i-1}}$. Given that $\overline{B_{i-1}}$ holds, $C|_{\rho_{i-1}}$ is wide which means that there are at least s free literals. Therefore $|H| \geq s/2\Delta$, where $\Delta = 2mp$ is an upper bound on the right-degree of G. Moreover, every $v \in H$ gives a possible knock, and different holes give different knocks. The reason is the following: if $x_{u,v}$ is a free literal, then (u, v) is a knock; and if $\bar{x}_{u,v}$ is a free literal, then (u', v) is a knock for every $u' \in N_G(v) - \{u\}$, which is non-empty since the right-degree of G is at least two. Therefore,

$$\Pr\left[S_i = 1 \;\middle|\; \overline{B_{i-1}} \wedge \bigwedge_{r/2 < j < i} S_j = 0\right] \geq \frac{|H|}{\Delta(n + r - i + 1)} \geq \frac{s}{3\Delta^2 n}.$$

Therefore,

$$\Pr\left[A_r \wedge \bigwedge_{j \geq r/2} \overline{B_j}\right] \leq \left(1 - \frac{s}{3\Delta^2 n}\right)^{r/2} \leq e^{-\frac{sr}{6\Delta^2 n}} \leq e^{-n/(\log n)^8}.$$

□ (of lemma 6)

Lemma 7 *Let j be such that $r/2 \leq j \leq r$. Then,* $\Pr\left[A_j \wedge B_j\right] \leq e^{-n/(\log n)^{11}}$.

Proof: Recall that A_j is the event that $C|_{\rho_j}$ is large, and B_j is the event that $C|_{\rho_j}$ is narrow. We let S_i be the indicator random variable for the event that (u_i, v_i) hits $C|_{\rho_{i-1}}$, where $\rho_{i-1} = ((u_1, v_1), \ldots, (u_{i-1}, v_{i-1}))$. Let $S = \sum_{i=1}^{j} S_i$. Then, for every h,

$$\begin{aligned}\Pr\left[A_j \wedge B_j\right] &= \Pr\left[A_j \wedge B_j \wedge S < h\right] + \Pr\left[A_j \wedge B_j \wedge S \geq h\right] \leq \\ &\leq \Pr\left[A_j \wedge S < h\right] + \Pr\left[A_j \wedge B_j \wedge S \geq h\right].\end{aligned}$$

We show that each term in this expression is exponentially small. More precisely, we show that $\Pr\left[A_j \wedge S < h\right] \leq e^{-n/(\log n)^3}$ and $\Pr\left[A_j \wedge B_j \wedge S \geq h\right] \leq e^{-n/(\log n)^{10}}$ which is clearly enough to prove Lemma 7.

Claim 1 *Let $h = n/(\log n)^4$. Then,* $\Pr\left[A_j \wedge S < h\right] \leq e^{-n/(\log n)^3}$.

Proof: Let $Y = \{(a_1, \ldots, a_j) \in \{0,1\}^j : \sum_{i=1}^{j} a_i < h\}$. Observe that A_j implies A_i for every $i \leq j$ because if $C|_{\rho_j}$ is large, so is $C|_{\rho_i}$ for every $i \leq j$. Then,

$$\begin{aligned}\Pr\left[A_j \wedge S < h\right] &= \Pr\left[\sum_{i=1}^{j} S_i < h \wedge A_j\right] = \sum_{\overline{a} \in Y} \Pr\left[\bigwedge_{i=1}^{j} S_i = a_i \wedge A_j\right] = \\ &= \sum_{\overline{a} \in Y} \prod_{i=1}^{j} \Pr\left[S_i = a_i \wedge A_j \;\middle|\; \bigwedge_{k=1}^{i-1} S_k = a_k\right] \leq \\ &\leq \sum_{\overline{a} \in Y} \prod_{i=1}^{j} \Pr\left[S_i = a_i \wedge A_{i-1} \;\middle|\; \bigwedge_{k=1}^{i-1} S_k = a_k\right] \leq \\ &\leq \sum_{\overline{a} \in Y} \prod_{i=1}^{j} \Pr\left[S_i = a_i \;\middle|\; A_{i-1} \wedge \bigwedge_{k=1}^{i-1} S_k = a_k\right].\end{aligned}$$

Fix $i \in \{1, \ldots, j\}$. Let H be the set of holes that occur in $C|_{\rho_{i-1}}$. We have $|H| \geq d/2\Delta$ given that A_{i-1} holds. Again, $\Delta = 2mp$ is an upper bound to the right-degree of G. Moreover, every $v \in H$ gives a possible hit, and different holes give different hits (the reason is the same as in Lemma 6 for knocks). Therefore,

$$\Pr\left[S_i = 1 \;\middle|\; A_{i-1} \wedge \bigwedge_{k=1}^{i-1} S_k = a_k\right] \geq \frac{|H|}{\Delta(n + r - i + 1)} \geq \frac{d}{3\Delta^2 n}.$$

Since there are at least $j - h$ zeros in $(a_1, \ldots, a_j)$, we obtain

$$\begin{aligned}\Pr\left[A_j \wedge S < h\right] &\leq \sum_{\overline{a} \in Y} \left(1 - \frac{d}{3\Delta^2 n}\right)^{j-h} \leq \sum_{i<h} \binom{j}{i} e^{-\frac{d(j-h)}{3\Delta^2 n}} \leq h j^h e^{-\frac{d(j-h)}{3\Delta^2 n}} \leq \\ &\leq \exp\left(-\frac{j-h}{36\Delta^2} + h\log(j) + \log(h)\right) \leq e^{-n/(\log n)^3}.\end{aligned}$$

□ (of claim 1)

Claim 2 $\Pr\left[A_j \wedge B_j \wedge S \geq h\right] \leq e^{-n/(\log n)^{10}}$.

Proof: During this proof we will drop the subindex j in A_j and B_j since it will always be the same. For every $i \in \{1, \ldots, r\}$, let $T_i \in \{\mathrm{k}, \mathrm{b}, \mathrm{n}\}$ be a random variable indicating whether (u_i, v_i) is a knock, a bad choice, or none of the previous respectively for $C|_{\rho_{i-1}}$. For $t \in \{\mathrm{k}, \mathrm{b}, \mathrm{n}\}$, let S_i^t be the indicator random variable for the event that $T_i = t$, and let $S^t = \sum_{i=1}^{j} S_i^t$. Thus, S^{k} is the number of knocks and S^{b} is the number of bad choices of ρ_j.

Fix $\rho = ((u_1, v_1), \ldots, (u_r, v_r))$ such that $A \wedge B \wedge S \geq h$ holds under ρ. Observe that (u_i, v_i) does not knock $C|_{\rho_{i-1}}$ for any $i \in \{1, \ldots, j\}$ since $C|_{\rho_j}$ must be large. Hence, $S^{\mathrm{k}} = 0$ under ρ. Let $b = (h - s)/(2\Delta + 1)$. We now claim that $S^{\mathrm{b}} \geq b$. Suppose for contradiction that the number of bad choices is less than b. Every bad choice (u_i, v_i) removes at most 2Δ free literals since at most those many literals about hole v_i may appear. Moreover, since there are no knocks, every hit (u_i, v_i) that is not a bad choice increases the number of free literals by at least one (see lemma 5). It follows that the number of free literals in $C|_{\rho_j}$ is at least $(S - S^{\mathrm{b}}) - 2\Delta S^{\mathrm{b}} > h - (2\Delta + 1)b = s$, a contradiction with the fact that B holds under ρ. We have proved that $\Pr\left[A \wedge B \wedge S \geq h\right] \leq \Pr\left[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b\right]$. The intuition behind why this probability is small is that every bad choice could have been a knock. This makes unlikely that ρ produces many bad choices and no knocks. In what follows, we will prove this intuition using martingales.

For $t \in \{\mathrm{k}, \mathrm{b}, \mathrm{n}\}$ and $i \in \{1, \ldots, j\}$, let P_i^t be the random variable $\Pr\left[T_i = t \mid \rho_0, \ldots, \rho_{i-1}\right]$. We define a martingale $X_0, \ldots, X_j$ with respect to $\rho_0, \ldots, \rho_j$ as follows: Let $X_0 = 0$, and $X_{i+1} = X_i + S_{i+1}^{\mathrm{b}} - P_{i+1}^{\mathrm{b}}$. Recall that S_{i+1}^{b} is the indicator random variable for the event that $T_{i+1} = \mathrm{b}$. So

$$\begin{aligned} \mathrm{E}\left[X_{i+1} \mid \rho_0, \ldots, \rho_i\right] &= (X_i + 1 - P_{i+1}^{\mathrm{b}}) \cdot P_{i+1}^{\mathrm{b}} + (X_i - P_{i+1}^{\mathrm{b}}) \cdot (1 - P_{i+1}^{\mathrm{b}}) = \\ &= (X_i - P_{i+1}^{\mathrm{b}})(P_{i+1}^{\mathrm{b}} + 1 - P_{i+1}^{\mathrm{b}}) + P_{i+1}^{\mathrm{b}} = X_i. \end{aligned}$$

Hence, $\{X_i\}_i$ is a martingale with respect to $\{\rho_i\}_i$. Observe also that $X_j = S^{\mathrm{b}} - \sum_{i=1}^{j} P_i^{\mathrm{b}}$. Similarly, we define $Y_0, \ldots, Y_j$ as follows: Let $Y_0 = 0$, and $Y_{i+1} = Y_i + S_{i+1}^{\mathrm{k}} - P_{i+1}^{\mathrm{k}}$. It is also easy to see that $\{Y_i\}_i$ is a martingale with respect to $\{\rho_i\}_i$. Again, $Y_j = S^{\mathrm{k}} - \sum_{i=1}^{j} P_i^{\mathrm{k}}$.

Subclaim 1 *$P_i^{\mathrm{k}}(\rho) \geq P_i^{\mathrm{b}}(\rho)/\Delta$ for every $\rho \in R_r(G)$ and $i \in \{1, \ldots, j\}$.*

Proof: Fix $i \in \{1, \ldots, j\}$ and $\rho = ((u_1, v_1), \ldots, (u_r, v_r))$. We want to show that $P_i^{\mathrm{k}}(\rho) \geq P_i^{\mathrm{b}}(\rho)/\Delta$. Define three sets as follows: let $Q = \{(u, v) \in E(G) : v \notin \{v_1, \ldots, v_{i-1}\}\}$, let Q^{k} be the set of knocks for $C|_{\rho_{i-1}}$ in Q, and let Q^{b} be the set of bad choices for $C|_{\rho_{i-1}}$ in Q. Observe that $P_i^{\mathrm{b}}(\rho) = |Q^{\mathrm{b}}| \cdot |Q|^{-1}$ and $P_i^{\mathrm{k}}(\rho) = |Q^{\mathrm{k}}| \cdot |Q|^{-1}$. On the other hand, every bad choice $(u, v) \in Q^{\mathrm{b}}$ gives a possible knock $(u', v) \in Q^{\mathrm{k}}$ by definition. Moreover, bad choices with different hole

components give different possible knocks. Grouping Q^{b} by holes, we have that $|Q^{\mathrm{k}}| \geq |Q^{\mathrm{b}}|/\Delta$. Consequently, $P_i^{\mathrm{k}}(\rho) \geq P_i^{\mathrm{b}}(\rho)/\Delta$ as required. □ (of subclaim 1)

To complete the proof of claim 2 we will need the following form of Azuma's Inequality: Let $X_0, \ldots, X_n$ be a martingale such that $|X_i - X_{i-1}| \leq 1$; then, $\Pr\left[|X_n - X_0| \geq \lambda\right] \leq 2e^{-\lambda^2/n}$ for every $\lambda > 0$ [11]. Now,

$$\begin{aligned}\Pr\left[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b\right] &= \Pr\left[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b \wedge X_j \geq b/2\right] + \\ &+ \Pr\left[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b \wedge X_j < b/2\right].\end{aligned}$$

The first summand is bounded by $\Pr\left[X_j \geq b/2\right] \leq 2e^{-b^2/4j}$ by Azuma's Inequality. The second summand is bounded by

$$\begin{aligned}\Pr\left[S^{\mathrm{k}} = 0 \wedge \textstyle\sum_{i=1}^{j} P_i^{\mathrm{b}} \geq b/2\right] &\leq \Pr\left[S^{\mathrm{k}} = 0 \wedge \textstyle\sum_{i=1}^{j} P_i^{\mathrm{k}} \geq b/2\Delta\right] \leq \\ &\leq \Pr\left[Y_j \leq -b/2\Delta\right] \leq 2e^{-b^2/4\Delta^2 j}.\end{aligned}$$

The first inequality follows from Subclaim 1, and the third follows from Azuma's Inequality again. The addition of the two summands is then bounded by $e^{-n/(\log n)^{10}}$ as required. □ (of claim 2 and lemma 7)

We are ready to complete the proof of our goal: equation (4). We have shown that

$$\Pr\left[C|_\rho \text{ large}\right] \leq \sum_{j=r/2}^{r} e^{-n/(\log n)^{11}} + e^{-n/(\log n)^{8}} \leq e^{-n/(\log n)^{13}}.$$

□ (of theorem 2)

By a different setting of parameters, it is easy to see that the strongest lower bound for PHP_n^m is of the form $e^{\frac{n^9}{(m\log m)^8 \log^3 n}}$. Namely, put $r = n/8$, $h = n^3/((m\log m)^2 \log n)$ and $s = h/2$ for that calculation. Therefore the best result is an exponential lower bound for $\mathrm{PHP}_n^{n^{9/8-\epsilon}}$.

We conclude this section with a separation result. Given that $\mathrm{Res}(\log)$ and depth-0.5 LK are polynomially equivalent, and given that PHP_n^{2n} has quasipolynomial-size proofs in depth-0.5 LK [14], we obtain:

Corollary 1 *There is an exponential separation between* $\mathrm{Res}(2)$ *and* $\mathrm{Res}(\log)$.

4 Lower Bound for Random CNF Formulas

4.1 Random Formulas and Restrictions

The model of random k-CNF formulas that we use is the one considered in [10, 3]. The distribution is denoted $\mathcal{F}_m^{k,n}$ and consists in choosing m clauses of exactly k literals independently with replacement. Most of the next definitions are taken and adapted from [3].

Definition 2 *For a real number n, a set of clauses $\mathcal{C}$ is* n-sparse *if $|\mathcal{C}| \leq n|v(\mathcal{C})|$ where $v(\mathcal{C})$ is the set of variables appearing in $\mathcal{C}$.*

Definition 3 *If $\mathcal{C}$ is a set of clauses and l is a literal, we say that l is* pure *in $\mathcal{C}$ if some clause of $\mathcal{C}$ contains l and no clause of $\mathcal{C}$ contains $\bar{l}$.*

Definition 4 *For $s \leq 1$ and $\epsilon \in (0,1)$, the following properties are defined for CNF formulas F:*

- *$A(s)$: Every set of $r \leq s$ clauses of F is 1-sparse.*
- *$B_\epsilon(s)$: For r such that $s/2 < r \leq s$, every subset of r clauses of F has at least ϵr pure literals.*

For a given refutation system $\mathcal{S}$, we say that an $\mathcal{S}$-refutation is k-bounded if all formulas of the refutation involve at most k distinct literals.

Proposition 1 [3] *Let $\mathcal{S}$ be a sound refutation system with all rules of fan-in at most two. Let $s > 0$ be an integer and F be a CNF formula. If properties $A(s)$ and $B_\epsilon(s)$ both hold for F, then F has no $\epsilon s/2$-bounded $\mathcal{S}$-refutation.*

A restriction is a sequence of pairs (x, v) where x is a variable and v is either $true$ or $false$. For a 2-disjunction C let $|C|$ be the number of distinct literals occurring in it. Let $\mathcal{R}$ be a probability distribution on restrictions. We say that $\mathcal{R}$ satisfies property $R(d, M)$ if and only if for every 2-disjunction C, $\Pr\left[|C|_\rho| \geq d\right] \leq 1/M$. We will consider two probability distributions.

- $\mathcal{A}_t$ chooses a permutation of the variables uniformly at random, then chooses each variable with probability t/n in the order of the permutation. The values assigned to the variables are chosen uniformly at random from $true$ and $false$.
- $\mathcal{B}_t$ chooses r, the length of the restriction, with a binomial distribution of parameters t/n and n. Then chooses uniformly at random any sequence of variables of length r without repetitions. The values assigned to the variables are chosen uniformly at random from $true$ and $false$.

We prove that $\mathcal{A}_t$ and $\mathcal{B}_t$ are the same distribution of probability. Obviously both distributions produce exactly the same restriccions. We only must show that any restriction ρ has the same probability in both distributions of probability.

Lemma 8 *For every $x_1, \ldots, x_r$ and $v_1, \ldots, v_r$,*

$$\Pr_{\rho \sim \mathcal{A}_t}\left[\rho = ((x_1, v_1), \ldots, (x_r, v_r))\right] = \Pr_{\rho \sim \mathcal{B}_t}\left[\rho = ((x_1, v_1), \ldots, (x_r, v_r))\right].$$

Proof: The probability $\Pr_{\rho\sim\mathcal{B}_t}\left[\rho = ((x_1,v_1),\ldots,(x_r,v_r))\right]$ is easy to find:

$$\binom{n}{r}\left(\frac{t}{n}\right)^r\left(1-\frac{t}{n}\right)^{n-r}\frac{1}{n2(n-1)2\ldots(n-r+1)2}. \tag{5}$$

The first part corresponds to the probability of choosing the value r from a binomial distribution. Remember that r is the length of the restriction. The rest of the expression is the probability of choosing the r correct pairs (x_i,v_i).

The probability $\Pr_{\rho\sim\mathcal{A}_t}\left[\rho = ((x_1,v_1),\ldots,(x_r,v_r))\right]$ is a little trickier. We will compute the probability of finding a permutation of the variables that is compatible with $(x_1,\ldots,x_r)$, that is, the variables $\{x_1,\ldots,x_r\}$ appear in that order. Then we multiply this probability by the probability of choosing the exact places where the variables in ρ are and choosing the right value for them:

$$\frac{\binom{n}{r}(n-r)!}{n!}\left(\frac{t}{n}\right)^r\left(1-\frac{t}{n}\right)^{n-r}\frac{1}{2^r}. \tag{6}$$

We first choose r places to put the variables in ρ, then we fill the gaps with the permutations of the other $n-r$ variables. These are the favorable cases, those that are compatible. With straightforward manipulations it is easy to see that (5) and (6) are equal. □ (of lemma 8)

The following is adapted from [3], with a minor change in the probability distribution.

Lemma 9 *For each integer $k \geq 3$ and $\epsilon > 0$, there are constants c_k, $c_{k,\epsilon}$, such that the following holds. Let m, n, s, t with $m = \Delta n$ for $\Delta \geq 1$. Let $F \sim \mathcal{F}_m^{k,n}$ and $\rho \sim \mathcal{A}_t$.*

- *If $t \leq c_k n/m^{1/k}$ and $s \leq c_k n/\Delta^{1/(k-2)}$, then $F|_\rho$ satisfies $A(s)$ with probability $1-o(1)$ in s.*
- *If $s,t \leq c_{k,\epsilon} n/\Delta^{2/(k-2-\epsilon)}$, then $F|_\rho$ satisfies $B_\epsilon(s)$ with probability $1-o(1)$ in s.*

Theorem 3 *Let $\mathcal{F}$ be a distribution over k-CNF formulas. Let $s, M \geq 1$ and $\epsilon > 0$ and let $\mathcal{R}$ be a distribution over restrictions that satisfies $R(\epsilon s/2, M)$. Then,*

$$\begin{aligned}\Pr_{F\sim\mathcal{F}}\left[res2(F) < M/2\right] &\leq 2\Pr_{F\sim\mathcal{F},\rho\sim\mathcal{R}}\left[F|_\rho \text{ does not satisfy } A(s)\right] + \\ &+ 2\Pr_{F\sim\mathcal{F},\rho\sim\mathcal{R}}\left[F|_\rho \text{ does not satisfy } B_\epsilon(s)\right],\end{aligned}$$

where $res2(F)$ is the minimum size of a $\mathrm{Res}(2)$-refutation of F.

Proof: For a fixed unsatisfiable k-CNF F, let P be a minimal-size $\mathrm{Res}(2)$-refutation of F. Let $\rho \sim \mathcal{R}$.

$$\begin{aligned}\Pr\left[F|_\rho \text{ satisfies } A(s)\wedge B_\epsilon(s)\right] &\leq \Pr\left[P|_\rho \text{ is not } \epsilon s/2 \text{ bounded}\right] \\ &\leq \Pr\left[\exists C\in P : |C|_\rho| > \epsilon s/2\right] \\ &\leq res2(F)\frac{1}{M}.\end{aligned}$$

The first inequality follows by Proposition 1, the second is immediate, and the third follows by union bound and the fact that $\mathcal{R}$ satisfies $R(\epsilon s/2, M)$.

To finish, let

$$p(F) = \Pr_{\rho}\left[F|_{\rho} \text{ does not satisfy } A(s)\right] + \Pr_{\rho}\left[F|_{\rho} \text{ does not satisfy } B_{\epsilon}(s)\right].$$

Then, $res2(F) < M/2$ implies that $\Pr_{\rho}\left[F|_{\rho} \text{ satisfies } A(s) \wedge B_{\epsilon}(s)\right] < 1/2$, and so $p(F) > 1/2$. Therefore, $\Pr_F\left[res2(F) < M/2\right] \leq \Pr_F\left[p(F) > 1/2\right] \leq 2E_F[p(F)]$ by Markov's inequality. The result follows. □ (of theorem 3)

4.2 The Lower Bound Argument

For simplicity, we only state the lower bound for the case $\mathcal{F}_{5n}^{3,n}$.

Theorem 4 *Let $F \sim \mathcal{F}_{5n}^{3,n}$. Then, $\mathrm{Res}(2)$-refutations of F require size $2^{\Omega(n^{1/3}/(\log(n))^2)}$ almost surely.*

Proof: Let $m = 5n$, $k = 3$, fix an arbitrary $\epsilon \in (0,1)$, and put $t = c_3 n/(5n)^{1/3} = c' n^{2/3}$ and $s = \min(c_3 n/5, c_{3,\epsilon} n/5^{2/1-\epsilon})$. Observe that these numbers satisfy the two hypothesis in Lemma 9. Let $M = 2^{n^{1/3}/(\log(n))^3}$. If we could prove that $\mathcal{B}_t$ satisfies property $R(\epsilon s/2, M)$, then $\Pr_F\left[res2(F) < M/2\right] < 2p(F)$ by Theorem 3. Since $p(F)$ is $o(1)$ according to Lemma 9, the Theorem would follow.

It remains to prove that $\mathcal{B}_t$ satisfies property $R(\epsilon s/2, M)$. In the following, we think of ρ as drawn from $\mathcal{B}_t$. We let $\rho = ((x_1, v_1), \ldots, (x_r, v_r))$.

A 2-disjunction is *large* if it contains at least $d = \epsilon s/2$ literals, otherwise it is *small*. A 2-disjunction is *wide* if it contains at least $w = t/2(\log(t))^2$ free literals, otherwise it is *narrow*. We say that (x_i, v_i) *knocks* a 2-disjunction if it makes it true. We say that (x_i, v_i) *hits* a 2-disjunction if it makes true a literal in it. Notice that every knock is a hit, but a hit might not be a knock. We say that (x_i, v_i) is a *bad choice* if it does not knock the 2-disjunction but could have knocked it just by giving the opposite value to the variable. For $i \leq r$, we let ρ_i be $((x_1, v_1), \ldots, (x_i, v_i))$. When possible we simplify 2-disjunctions: we substitute subformulas of the form $l \vee (l \wedge l')$ by l and subformulas of the form $\bar{l} \vee (l \wedge l')$ by $\bar{l} \vee l'$. We aim for a proof that

$$\Pr\left[C|_{\rho} \text{ is large}\right] \leq e^{-\frac{n^{1/3}}{(\log(n))^4}}, \tag{7}$$

where C is an arbitrary simplified 2-disjunction.

Let A_i be the event that $C|_{\rho_i}$ contains at least d distinct literals. Let A be the event $A_{|\rho|}$.

$$\Pr\left[A\right] \;=\; \Pr\left[A \wedge |\rho| < t/2\right] + \Pr\left[A \wedge |\rho| \geq t/2\right]. \tag{8}$$

Obviously $\Pr\left[A \wedge |\rho| < t/2\right] \leq \Pr\left[|\rho| < t/2\right]$ which is smaller than $e^{-t/8}$ by Chernoff bounds, so

$$(8) \leq e^{-\frac{n^{2/3}}{\log(n)}} + \Pr\left[A \;\middle|\; |\rho| \geq t/2\right].$$

We show now that $\Pr\left[A \;\middle|\; |\rho| \geq t/2\right]$ is exponentially small. For every i such that $t/4 \leq i \leq t/2$, let B_i be the event that $C|_{\rho_i}$ is narrow, that is, it contains less than w free literals. Let D be the event that $|\rho| \geq t/2$. Then,

$$\Pr\left[A \;\middle|\; D\right] = \Pr\left[A \wedge \bigvee_{j=t/4}^{t/2} B_j \;\middle|\; D\right] + \Pr\left[A \wedge \bigwedge_{j=t/4}^{t/2} \overline{B_j} \;\middle|\; D\right]. \tag{9}$$

We show that both terms in (9) are exponentially small. For every i such that $t/4 \leq i \leq t/2$, let T_i be the indicator random variable for the event that (x_i, v_i) is a knock. Then the second term in (9) is

$$\begin{aligned}
\Pr\left[A \wedge \bigwedge_{j=t/4}^{t/2} \overline{B_j} \;\middle|\; D\right] &\leq \Pr\left[\bigwedge_{i>t/4}^{t/2} T_i = 0 \wedge \bigwedge_{j=t/4}^{t/2} \overline{B_j} \;\middle|\; D\right] = \\
&= \prod_{i>t/4}^{t/2} \Pr\left[T_i = 0 \;\wedge \bigwedge_{j=t/4}^{t/2} \overline{B_j} \;\middle|\; \bigwedge_{j>t/4}^{i-1} T_j = 0 \wedge D\right] \leq \\
&= \prod_{i>t/4}^{t/2} \Pr\left[T_i = 0 \wedge \overline{B_{i-1}} \;\middle|\; \bigwedge_{j>t/4}^{i-1} T_j = 0 \wedge D\right] \leq \\
&= \prod_{i>t/4}^{t/2} \Pr\left[T_i = 0 \;\middle|\; \overline{B_{i-1}} \wedge \bigwedge_{j>t/4}^{i-1} T_j = 0 \wedge D\right] \leq \\
&\leq \prod_{i>t/4}^{t/2} \left(1 - \frac{w}{2(n-i+1)}\right) \leq \left(1 - \frac{w}{2n}\right)^{t/4} \leq \\
&\leq e^{-\frac{wt}{8n}} = e^{-\frac{n^{1/3}}{(\log(n))^3}}.
\end{aligned}$$

The first term in (9) is also exponentially small. Observe that

$$\Pr\left[A \wedge \bigvee_{j=t/4}^{t/2} B_j \;\middle|\; D\right] = \Pr\left[\bigvee_{j=t/4}^{t/2} (A \wedge B_j) \;\middle|\; D\right] = \tag{10}$$

$$\leq \sum_{j=t/4}^{t/2} \Pr\left[A_j \wedge B_j \;\middle|\; D\right]. \tag{11}$$

The last inequality is true because A implies A_j for any $j \leq t/2$.

Lemma 10 *If j is such that $t/4 \leq j \leq t/2$, then* $\Pr\left[A_j \wedge B_j \mid D\right] \leq e^{-\frac{n^{2/3}}{(\log(n))^6}}$.

Proof: For every $i \leq j$ let S_i be the indicator random variable for the event that (x_i, v_i) hits $C|_{\rho_{i-1}}$, that is, that (x_i, v_i) gives value $true$ to a literal in $C|_{\rho_{i-1}}$. Let $S = \sum_{i=1}^{j} S_i$. We divide the calculation in two parts: what happens when the number of hits is less than a certain $h = t/(\log(t))^2$ and what happens otherwise.

$$\Pr\left[A_j \wedge B_j \mid D\right] \;=\; \Pr\left[A_j \wedge B_j \wedge S < h \mid D\right] + \Pr\left[A_j \wedge B_j \wedge S \geq h \mid D\right]$$

We start by the easiest part. The intuition is that if the 2-disjunction is large it would be extremely difficult to hit it only a few times.

Sublemma 1 $\Pr\left[A_j \wedge S < h \mid D\right] \leq e^{-\frac{n^{2/3}}{(\log(n))^2}}$.

Proof: Let $Y = \{(a_1, \ldots, a_j) \in \{0,1\}^j : \sum_{i=1}^{j} a_i < h\}$. Observe that A_j implies A_i for every $i \leq j$ because if $C|_{\rho_j}$ is large, so is $C|_{\rho_i}$. Then,

$$\begin{aligned}
\Pr\left[A_j \wedge S < h \mid D\right] &= \Pr\left[A_j \wedge \sum_{i=1}^{j} S_i < h \;\middle|\; D\right] = \\
&= \sum_{\overline{a} \in Y} \Pr\left[A_j \wedge \bigwedge_{i=1}^{j} S_i = a_i \;\middle|\; D\right] = \\
&= \sum_{\overline{a} \in Y} \prod_{i=1}^{j} \Pr\left[A_j \wedge S_i = a_i \;\middle|\; \bigwedge_{k=1}^{i-1} S_k = a_k \wedge D\right] \leq \\
&\leq \sum_{\overline{a} \in Y} \prod_{i=1}^{j} \Pr\left[A_{i-1} \wedge S_i = a_i \;\middle|\; \bigwedge_{k=1}^{i-1} S_k = a_k \wedge D\right] \leq \\
&\leq \sum_{\overline{a} \in Y} \prod_{i=1}^{j} \Pr\left[S_i = a_i \;\middle|\; A_{i-1} \wedge \bigwedge_{k=1}^{i-1} S_k = a_k \wedge D\right].
\end{aligned}$$

Fix $i \in \{1, \ldots, j\}$.

$$\Pr\left[S_i = 1 \;\middle|\; A_{i-1} \wedge \bigwedge_{k=1}^{i-1} S_k = a_k \wedge D\right] \geq \frac{d}{2(n-i+1)} \geq \frac{d}{2n}.$$

Since there are at least $j-h$ zeros in $(a_1,\ldots,a_j)$, we obtain

$$\begin{aligned}
\Pr\left[A_j \wedge S < h \mid D\right] &\leq \sum_{\overline{a}\in Y}\left(1-\frac{d}{2n}\right)^{j-h} \leq \\
&\leq \sum_{i<h}\binom{j}{i}e^{-\frac{d(j-h)}{2n}} \leq hj^h e^{-\frac{d(j-h)}{2n}} \leq \\
&\leq \exp\left(-\frac{d(j-h)}{2n}+h\log(j)+\log(h)\right) \leq \\
&\leq e^{-\frac{dt}{10n}+\frac{t}{\log(t)}} \leq e^{-\frac{n^{2/3}}{(\log(n))^2}}.
\end{aligned}$$

□ (of sublemma 1)

The last thing to do is to see what happens when the number of hits is big.

Sublemma 2 $\Pr\left[A_j \wedge B_j \wedge S \geq h \mid D\right] \leq e^{-\frac{n^{2/3}}{(\log(n))^5}}$.

Proof: For every $1 \leq i \leq t/2$, let $T_i \in \{\mathrm{k},\mathrm{b},\mathrm{n}\}$ be a random variable indicating whether (x_i,v_i) is a knock, a bad choice, or none of the previous respectively for $C|_{\rho_{i-1}}$. For $t \in \{\mathrm{k},\mathrm{b},\mathrm{n}\}$, let S_i^t be the indicator random variable for the event that $T_i = t$, and let $S^t = \sum_{i=1}^{j} S_i^t$. Thus, S^{k} is the number of knocks and S^{b} is the number of bad choices of ρ_j. For the rest of the proof we will skip the condition on D and the subindices from A and B. Fix ρ satisfying $A \wedge B \wedge S \geq h$. Note that the number of knocks is 0 because the 2-disjunction still exists, so $S^{\mathrm{k}} = 0$. Now let be $b = (h-w)/2$, we now claim that $S^{\mathrm{b}} \geq b$. Suppose for contradiction that the number of bad choices is less than b. Every bad choice (x_i,v_i) removes at most one free literal. Moreover, since there are no knocks, every hit (x_i,v_i) that is not a bad choice increases the number of free literals by at least one. The reason is that such a hit turns a conjunction into a free literal. Remember that we simplify the 2-disjunction when possible, and so the literal was not free before the hit (x_i,v_i) is applied. It follows then that the number of free literals in $C|_{\rho_j}$ is at least $(S-S^{\mathrm{b}})-S^{\mathrm{b}} > h-2b = w$, a contradiction with the fact that B holds under ρ.

So far we have proved that $\Pr\left[A \wedge B \wedge S \geq h\right] \leq \Pr\left[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b\right]$. The intuition behind why this probability is small is that every bad choice could have been a knock. This makes it unlikely that ρ produces many bad choices and no knocks. In what follows, we will prove this intuition using martingales.

Claim 3 $\Pr\left[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b\right] \leq e^{-\frac{n^{2/3}}{(\log(n))^5}}$.

Proof: For $t \in \{\mathrm{k},\mathrm{b},\mathrm{n}\}$ and $i \in \{1,\ldots,j\}$, let P_i^t denote the random variable $\Pr\left[T_i = t \mid \rho_0,\ldots,\rho_{i-1}\right]$. We define a martingale $X_0,\ldots,X_j$ with respect to $\rho_0,\ldots,\rho_j$ as follows: Let

$X_0 = 0$, and $X_{i+1} = X_i + S^{\mathrm{b}}_{i+1} - P^{\mathrm{b}}_{i+1}$. Recall that S^{b}_{i+1} is the indicator random variable for the event that $T_{i+1} = \mathrm{b}$. Observe that

$$\begin{aligned}\mathrm{E}\left[X_{i+1} \,\middle|\, \rho_0, \ldots, \rho_i\right] &= (X_i + 1 - P^{\mathrm{b}}_{i+1}) \cdot P^{\mathrm{b}}_{i+1} + (X_i - P^{\mathrm{b}}_{i+1}) \cdot (1 - P^{\mathrm{b}}_{i+1}) = \\ &= (X_i - P^{\mathrm{b}}_{i+1})(P^{\mathrm{b}}_{i+1} + 1 - P^{\mathrm{b}}_{i+1}) + P^{\mathrm{b}}_{i+1} = X_i.\end{aligned}$$

Hence, $\{X_i\}_i$ is a martingale with respect to $\{\rho_i\}_i$. Observe also that $X_j = S^{\mathrm{b}} - \sum_{i=1}^{j} P^{\mathrm{b}}_i$. Similarly, we define $Y_0, \ldots, Y_j$ as follows: Let $Y_0 = 0$, and $Y_{i+1} = Y_i + S^{\mathrm{k}}_{i+1} - P^{\mathrm{k}}_{i+1}$. It is also easy to see that $\{Y_i\}_i$ is a martingale with respect to $\{\rho_i\}_i$. Again, $Y_j = S^{\mathrm{k}} - \sum_{i=1}^{j} P^{\mathrm{k}}_i$.

We will use the following form of Azuma's Inequality: Let $X_0, \ldots, X_n$ be a martingale such that $|X_i - X_{i-1}| \leq 1$; then, $\Pr\left[|X_n - X_0| \geq \lambda\right] \leq 2e^{-\lambda^2/n}$ for every $\lambda > 0$. In the next calculation we will also use the fact that $P^{\mathrm{k}}_i(\rho) = P^{\mathrm{b}}_i(\rho)$ for every ρ and $i \in \{1, \ldots, j\}$.

$$\begin{aligned}\Pr\left[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b\right] &= \Pr\left[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b \wedge X_j \geq b/2\right] + \\ &+ \Pr\left[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b \wedge X_j < b/2\right].\end{aligned}$$

The first summand is bounded by $\Pr\left[X_j \geq b/2\right] \leq 2e^{-b^2/4j}$ by Azuma's Inequality. The second summand is bounded by

$$\begin{aligned}\Pr\left[S^{\mathrm{k}} = 0 \wedge \textstyle\sum_{i=1}^{j} P^{\mathrm{b}}_i \geq b/2\right] &\leq \Pr\left[S^{\mathrm{k}} = 0 \wedge \textstyle\sum_{i=1}^{j} P^{\mathrm{k}}_i \geq b/2\right] \leq \\ &\leq \Pr\left[Y_j \leq -b/2\right] \leq \\ &\leq 2e^{-b^2/4j},\end{aligned}$$

by Azuma's Inequality again. Therefore, the sum is bounded by $4e^{-\frac{t}{32(\log(t))^4}} \leq e^{-\frac{n^{2/3}}{(\log(n))^5}}$ as required. □ (of claim 3 and sublemma 2).

With both sublemmas proved, so is Lemma 10. We are ready to complete the proof of our goal (7). We have shown that

$$\Pr\left[C|_\rho \text{ is large}\right] \leq e^{-\frac{n^{2/3}}{\log(n)}} + e^{-\frac{n^{1/3}}{(\log(n))^3}} + \sum_{i=t/4}^{t/2} e^{-\frac{n^{2/3}}{(\log(n))^6}} \leq e^{-\frac{n^{1/3}}{(\log(n))^4}}.$$

□ (of theorem 4)

We give another proof of claim 3 that does not require martingales.

Claim 4 $\Pr\left[S^{\mathrm{b}} = 0 \wedge S^{\mathrm{b}} \geq b\right] \leq 2^{-b}$

Proof: Let us call a restriction favorable if it has b or more bad choices and no knocks. By modifying a favorable restriction, we can get $2^b - 1$ restrictions with one knock or more just by changing the value of the variables that form the set of bad choices. Let us call these restrictions knock restrictions.

We will show now that no different favorable restrictions generate the same knock restrictions. Let us consider two favorable restrictions, say f_1 and f_2. Both restrictions must have the same variables in the same order, otherwise they cannot form the same knock restriction. Now, let us call x the first variable such that $f_1(x) \neq f_2(x)$. Let us suppose that x is a bad choice for f_1. This is impossible because f_1 and f_2 are equal up to the variable preceeding x, so if x is a bad choice for f_1, then x is a knock for f_2, so f_2 is not favorable. The same argument applies for f_2. If x is neither a bad choice for f_1 nor for f_2 then the value of x must coincide if we intend to build the same knock restriction, because we are only changing the value of variables that produce bad choices. We must conclude $f_1 = f_2$.

Now let us call F the set of favorable restrictions and K the set of knock restrictions generated by the restricions in F. So

$$\begin{aligned} \Pr\left[S^{\mathrm{b}} = 0 \wedge S^{\mathrm{b}} \geq b\right] &= \frac{\text{favorable}}{\text{possible}} \leq \frac{|F|}{|F| + |K|} = \\ &= \frac{1}{1 + |K|/|F|} = \frac{1}{1 + 2^b - 1} = \frac{1}{2^b} \end{aligned}$$

□ (of claim 4)

5 Separation between Res(2) and Resolution

In this section we prove that Resolution cannot polynomially simulate $\mathrm{Res}(2)$. More precisely, we prove that a certain Clique-Coclique principle, as defined by Bonet, Pitassi and Raz in [6], has polynomial-size $\mathrm{Res}(2)$-refutations, but every Res-refutation requires quasipolynomial size.

The Clique-Coclique principle that we use, $\mathrm{CLIQUE}^n_{k,k'}$, is the conjunction of the following set of clauses:

$$x_{i,1} \vee \cdots \vee x_{i,n} \qquad 1 \leq l \leq k \tag{12}$$

$$\bar{x}_{l,i} \vee \bar{x}_{l,j} \qquad 1 \leq l \leq k,\ 1 \leq i,j \leq n,\ i \neq j \tag{13}$$

$$\bar{x}_{l,i} \vee \bar{x}_{l',i} \qquad 1 \leq l,l' \leq k,\ 1 \leq i \leq n,\ l \neq l' \tag{14}$$

$$y_{1,i} \vee \cdots \vee y_{k',i} \qquad 1 \leq i \leq n \tag{15}$$

$$\bar{y}_{l,i} \vee \bar{y}_{l',i} \qquad 1 \leq l,l' \leq k',\ 1 \leq i \leq n,\ l \neq l' \tag{16}$$

$$\bar{x}_{l,i} \vee \bar{x}_{l',j} \vee \bar{y}_{t,i} \vee \bar{y}_{t,j} \qquad 1 \leq l,l' \leq k,\ 1 \leq t \leq k',\ 1 \leq i,j \leq n,\ l \neq l',\ i \neq j \tag{17}$$

We start with a reduction from $\mathrm{CLIQUE}^n_{k,k'}$ to $\mathrm{PHP}^k_{k'}$ that can be carried over in $\mathrm{Res}(2)$:

Theorem 5 *Let $k' < k \leq n$. If $\mathrm{PHP}^k_{k'}$ has Resolution refutations of size S, then $\mathrm{CLIQUE}^n_{k,k'}$ has $\mathrm{Res}(2)$-refutations of size Sn^c for some constant $c > 0$.*

Proof: We use the following $\mathrm{Res}(2)$-reduction to transform the formula $\mathrm{CLIQUE}^n_{k,k'}$ into $\mathrm{PHP}^k_{k'}$. The meaning of variable $p_{i,j}$ is that pigeon i sits in hole j. We perform the following substitutions:

$$p_{i,j} \equiv \bigvee_{l=1}^{n} (x_{i,l} \wedge y_{i,l}) \qquad\qquad \bar{p}_{i,j} \equiv \bigvee_{l=1, j' \neq j}^{n} (x_{i,l} \wedge y_{j',l})$$

First we show how to get clauses (1) from clauses (12) and (15). If we expand clause (1) for a certain i we have:

$$\begin{array}{c} (x_{i,1} \wedge y_{1,1}) \vee (x_{i,2} \wedge y_{1,2}) \vee (x_{i,3} \wedge y_{1,3}) \vee \cdots \vee (x_{i,n} \wedge y_{1,n}) \vee \\ (x_{i,1} \wedge y_{2,1}) \vee (x_{i,2} \wedge y_{2,2}) \vee (x_{i,3} \wedge y_{2,3}) \vee \cdots \vee (x_{i,n} \wedge y_{2,n}) \vee \\ (x_{i,1} \wedge y_{3,1}) \vee (x_{i,2} \wedge y_{3,2}) \vee (x_{i,3} \wedge y_{3,3}) \vee \cdots \vee (x_{i,n} \wedge y_{3,n}) \vee \\ \cdots \\ (x_{i,1} \wedge y_{k',1}) \vee (x_{i,2} \wedge y_{k',2}) \vee (x_{i,3} \wedge y_{k',3}) \vee \cdots \vee (x_{i,n} \wedge y_{k',n}) \end{array} \tag{18}$$

We apply successively for $1 \leq j \leq k'$ the $\wedge$-introduction rule to clauses $y_{1,1} \vee \cdots \vee y_{k',1}$ and $x_{i,1} \vee \cdots \vee x_{i,n}$ along variables $x_{i,1}$ and $y_{j,1}$ and get:

$$(x_{i,1} \wedge y_{1,1}) \vee (x_{i,1} \wedge y_{2,1}) \vee \cdots \vee (x_{i,1} \wedge y_{k',1}) \vee x_{i,2} \vee \cdots \vee x_{i,n} \tag{19}$$

Observe that the conjuctions in (19) form the first column in (18). To add the second column of (18) to (19) we apply successively for $1 \leq j \leq k'$ the $\wedge$-rule to clauses $y_{1,2} \vee \cdots \vee y_{k',2}$ and (19) along variables $x_{i,2}$ and $y_{j,2}$ and get:

$$\begin{array}{c} (x_{i,1} \wedge y_{1,1}) \vee (x_{i,1} \wedge y_{2,1}) \vee \cdots \vee (x_{i,1} \wedge y_{k',1}) \vee \\ (x_{i,2} \wedge y_{1,2}) \vee (x_{i,2} \wedge y_{2,2}) \vee \cdots \vee (x_{i,2} \wedge y_{k',2}) \vee x_{i,3} \vee \cdots \vee x_{i,n} \end{array} \tag{20}$$

Now it is clear how to get (18).

Now we will show how to get the initial clauses (2). Let us consider the clause $\bar{p}_{i,t} \vee \bar{p}_{j,t}$. We first generate $p_{i,1} \vee \cdots \vee p_{i,k'}$ and $p_{j,1} \vee \cdots \vee p_{j,k'}$. Let us rewrite them as:

$$(x_{i,1} \wedge y_{t,1}) \vee (x_{i,2} \wedge y_{t,2}) \vee (x_{i,3} \wedge y_{t,3}) \vee \cdots \vee (x_{i,n} \wedge y_{t,n}) \vee A \tag{21}$$

$$(x_{j,1} \wedge y_{t,1}) \vee (x_{j,2} \wedge y_{t,2}) \vee (x_{j,3} \wedge y_{t,3}) \vee \cdots \vee (x_{j,n} \wedge y_{t,n}) \vee B \tag{22}$$

where A is $p_{i,1} \vee \cdots \vee p_{i,t-1} \vee p_{i,t+1} \vee \cdots \vee p_{i,k'}$ and B is $p_{j,1} \vee \cdots \vee p_{j,t-1} \vee p_{j,t+1} \vee \cdots \vee p_{j,k'}$. For the sake of brevity we use $p_{i,j}$ as abbreviation of the 2-disjuction it denotes. It is clear that $\bar{p}_{i,t} \vee \bar{p}_{j,t}$ is $A \vee B$, that is:

$$p_{i,1} \vee \cdots \vee p_{i,t-1} \vee p_{i,t+1} \vee \cdots \vee p_{i,k'} \vee p_{j,1} \vee \cdots \vee p_{j,t-1} \vee p_{j,t+1} \vee \cdots \vee p_{j,k'}$$

Now we will get $A \vee B$ from (21), (22) and (17). We apply the cut rule to (22) and $\bar{x}_{i,1} \vee \bar{x}_{j,l} \vee \bar{y}_{t,1} \vee \bar{y}_{t,l}$ for $1 \leq l \leq n$, $l \neq 1$, and get:

$$\bar{x}_{i,1} \vee \bar{y}_{t,1} \vee (x_{j,1} \wedge y_{t,1}) \vee B \tag{23}$$

Solving it with $\bar{x}_{i,1} \vee \bar{x}_{j,1}$ we get $\bar{x}_{i,1} \vee \bar{y}_{t,1} \vee B$. Solving this clause with (21) we get

$$(x_{i,2} \wedge y_{t,2}) \vee \cdots \vee (x_{i,n} \wedge y_{t,n}) \vee A \vee B \tag{24}$$

Now we can get rid successively of $(x_{i,2} \wedge y_{t,2}), \ldots, (x_{i,n} \wedge y_{t,n})$ as we did with $(x_{i,1} \wedge y_{t,1})$.

It remains to show how to simulate a normal resolution step. We have $p_{i,j} \vee A$ and $\bar{p}_{i,j} \vee B$ and we want to get $A \vee B$. We expand both clauses:

$$(x_{i,1} \wedge y_{j,1}) \vee (x_{i,2} \wedge y_{j,2}) \vee (x_{i,3} \wedge y_{j,3}) \vee \cdots \vee (x_{i,n} \wedge y_{j,n}) \vee A \tag{25}$$

$$\begin{array}{c}
(x_{i,1} \wedge y_{1,1}) \vee (x_{i,2} \wedge y_{1,2}) \vee (x_{i,3} \wedge y_{1,3}) \vee \cdots \vee (x_{i,n} \wedge y_{1,n}) \vee \\
(x_{i,1} \wedge y_{2,1}) \vee (x_{i,2} \wedge y_{2,2}) \vee (x_{i,3} \wedge y_{2,3}) \vee \cdots \vee (x_{i,n} \wedge y_{2,n}) \vee \\
\cdots \\
(x_{i,1} \wedge y_{j-1,1}) \vee (x_{i,2} \wedge y_{j-1,2}) \vee (x_{i,3} \wedge y_{j-1,3}) \vee \cdots \vee (x_{i,n} \wedge y_{j-1,n}) \vee \\
(x_{i,1} \wedge y_{j+1,1}) \vee (x_{i,2} \wedge y_{j+1,2}) \vee (x_{i,3} \wedge y_{j+1,3}) \vee \cdots \vee (x_{i,n} \wedge y_{j+1,n}) \vee \\
\cdots \\
(x_{i,1} \wedge y_{k',1}) \vee (x_{i,2} \wedge y_{k',2}) \vee (x_{i,3} \wedge y_{k',3}) \vee \cdots \vee (x_{i,n} \wedge y_{k',n}) \vee B
\end{array} \tag{26}$$

If we get clauses $\bar{x}_{i,l} \vee \bar{y}_{j,l} \vee B$ for $1 \leq l \leq n$, we solve them all with (25) and get $A \vee B$ as desired. We will show how to get $\bar{x}_{i,1} \vee \bar{y}_{j,1} \vee B$. We solve (26) with $\bar{y}_{j,1} \vee \bar{y}_{l,1}$, $l \neq j$ of course. With these we get rid of the first column of (26) and we add a literal $\bar{y}_{j,1}$. We can get rid of the rest of columns by solving enough times with clauses $\bar{x}_{i,1} \vee \bar{x}_{i,l}$, $l \neq 1$, and we get $\bar{x}_{i,1} \vee \bar{y}_{j,1} \vee B$. □ (of theorem 5)

We will use the Monotone Interpolation Theorem for Resolution together with the following result of Alon and Boppana [1] establishing a lower bound to the size of monotone circuits that separate large cliques from small cocliques. In the following, $F(m, k, k')$ is the set of monotone functions that separate k-cliques from k'-cocliques on m nodes.

Theorem 6 [1] *If $f \in F(m, k, k')$ where $3 \leq k' \leq k$ and $k\sqrt{k'} \leq m/(8 \log m)$, then*

$$S^+(f) \geq \frac{1}{8} \left(\frac{m}{4k\sqrt{k'} \log m} \right)^{(\sqrt{k'}+1)/2},$$

where $S^+(f)$ is the monotone circuit size of f.

Theorem 7 *Let $k = \sqrt{m}$ and $k' = (\log m)^2/8 \log\log m$. Then, (i) $\mathrm{CLIQUE}^m_{k,k'}$ has $\mathrm{Res}(2)$-refutations of size polynomial in m, and (ii) every Resolution refutation of $\mathrm{CLIQUE}^m_{k,k'}$ has size at least $\exp(\Omega((\log m)^2/\sqrt{\log\log m}))$.*

Proof: Regarding (i), we have that $k' \log k' \leq \frac{1}{4}(\log m)^2$, and so $2^{\sqrt{k' \log k'}} \leq m^{1/2} = k$. On the other hand, Buss and Pitassi [7] proved that $\mathrm{PHP}^k_{k'}$ has Resolution refutations of size polynomial in k whenever $k \geq 2^{\sqrt{k' \log k'}}$. Therefore, by Theorem 5, $\mathrm{CLIQUE}^m_{k,k'}$ has $\mathrm{Res}(2)$-refutations of size polynomial in m. Regarding (ii), we apply the feasible monotone interpolation theorem for Resolution. We have

$$\frac{\log m}{3\sqrt{\log\log m}} \leq \sqrt{k'} \leq \log m.$$

Therefore, by Theorem 6, if $f \in F(m, k, k')$ is a monotone interpolant, then

$$S^+(f) \geq \frac{1}{8}\left(\frac{m}{4\sqrt{m}(\log m)^2}\right)^{\frac{\log m}{6\sqrt{\log\log m}}} \geq \frac{1}{8}\left(\frac{m}{m^{3/4}}\right)^{\frac{\log m}{6\sqrt{\log\log m}}},$$

which is $\exp(\Omega((\log m)^2/\sqrt{\log\log m}))$. □ (of theorem 7)

As a corollary, we solve an open problem posed by Krajíček [13].

Corollary 2 $\mathrm{Res}(2)$ *does not have the feasible monotone interpolation property.*

6 Discussion and Open Problems

In the paper we proved that there is a quasipolynomial separation between Resolution and $\mathrm{Res}(2)$. It is an open question whether the separation could be exponential, or a quasipolynomial simulation of $\mathrm{Res}(2)$ by Resolution exists. It is important to notice, that our lower bound for PHP would not follow from such a simulation. Indeed, the lower bound that would follow from that would be of the form 2^{n^ϵ}.

The previous separation was obtained using a lower bound for Resolution proved via the monotone interpolation theorem. It is open whether the separation (or a stronger one) could be obtained via the size-width trade-off [5] as a method for proving lower bounds for Resolution. It would also be interesting to see what would that mean in terms of possible size-width trade-offs for $\mathrm{Res}(2)$. We conjecture that $\mathrm{Res}(2)$ does not have a strong size-width trade-off. Notice that $\mathrm{Res}(\log)$ does not have it. This is because (a) $\mathrm{Res}(\log)$ is equivalent to depth-0.5 LK, (b) PHP^{2n}_n has quasipolynomial-size proofs in depth-0.5 LK [14], and (c) PHP^{2n}_n has $\Omega(n)$ width lower bounds for $\mathrm{Res}(\log)$.

In this paper we extended the width lower bound technique beyond Resolution. A very interesting open question is to see whether the technique can also be extended to give lower bounds for

$\mathrm{Res}(3), \mathrm{Res}(4), \ldots, \mathrm{Res}(\log)$. It seems that some new ideas need to be developed to do that. This question is related to the optimality of the $\mathrm{Res}(\log)$ upper bound for PHP_n^{2n}.

Finally, we note that exponential-size lower bounds for $\mathrm{PHP}_n^{n^{1+\epsilon}}$ in $\mathrm{Res}(k)$ implies lower bounds for $\mathrm{PHP}_n^{n^c}$ in Resolution for some c. This is a long-standing open question.

References

[1] N. Alon and R. Boppana. The monotone circuit complexity of boolean functions. *Combinatorica*, 7(1):1–22, 1987.

[2] P. Beame, R. Impagliazzo, J. Krajíček, T. Pitassi, P. Pudlák, and A. Woods. Exponential lower bounds for the pigeonhole principle. In *Proc. of the 24th Annual ACM STOC*, pages 200–220, 1992.

[3] P. Beame, R. Karp, T. Pitassi, and M. Saks. The efficiency of resolution and Davis-Putnam procedures. Submitted. Previous version in STOC'98, 1999.

[4] P. Beame and T. Pitassi. Simplified and improved resolution lower bounds. In *Proc. of the 37th Annual IEEE FOCS*, pages 274–282, 1996.

[5] E. Ben-Sasson and A. Wigderson. Short proofs are narrow: Resolution made simple. In *Proc. of the 31st Annual ACM STOC*, pages 517–527, 1999. Revised version (2000).

[6] M. Bonet, T. Pitassi, and R. Raz. Lower bounds for cutting planes proofs with small coefficients. *The Journal of Symbolic Logic*, 62(3):708–728, Sept. 1997.

[7] S. Buss and T. Pitassi. Resolution and the weak pigeonhole principle. In *CSL: 11th Workshop on Computer Science Logic*. LNCS, Springer-Verlag, 1997.

[8] S. R. Buss. Polynomial size proofs of the propositional pigeonhole principle. *The Journal of Symbolic Logic*, 52(4):916–927, Dec. 1987.

[9] S. R. Buss and G. Turán. Resolution proofs on generalized pigeonhole principles. *Theoretical Computer Science*, 62(3):311–317, Dec. 1988.

[10] V. Chvátal and E. Szemerédi. Many hard examples for resolution. *J. ACM*, 35(4):759–768, 1988.

[11] G. R. Grimmet and D. R. Stirzaker. *Probability and Random Processes*. Oxford Science Publications, 1982.

[12] A. Haken. The intractability of resolution. *Theoretical Computer Science*, 39(2-3):297–308, Aug. 1985.

[13] J. Krajíček. On the weak pigeonhole principle. To appear in Fundamenta Mathematicæ, 2000.

[14] A. Maciel, T. Pitassi, and A. Woods. A new proof of the weak pigeonhole principle. In *Proc. of the 32nd Annual ACM STOC*, pages 368–377, 2000.

[15] J. B. Paris, A. J. Wilkie, and A. R. Woods. Provability of the pigeonhole principle and the existence of infinitely many primes. *The Journal of Symbolic Logic*, 53(4):1235–1244, 1988.

[16] T. Pitassi and R. Raz. Regular resolution lower bounds for the weak pigeonhole principle. to appear in STOC'2001, 2001.