Information and Computation **176**, 136–152 (2002)
doi:10.1006/inco.2002.3114

Lower Bounds for the Weak Pigeonhole Principle and Random Formulas beyond Resolution[1]

Albert Atserias,[2] Maria Luisa Bonet,[3] and Juan Luis Esteban[4]

Departament de Llenguatges i Sistemes Informàtics Universitat Politècnica de Catalunya, C/Jordi Girona Salgado, 1-3, Edif. C6. E08034 Barcelona, Spain
E-mail: atserias@lsi.upc.es, bonet@lsi.upc.es, esteban@lsi.upc.es

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. © 2002 Elsevier Science (USA)

Key Words: proof complexity; resolution; lower bounds; random restrictions; martingales; monotone interpolation.

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 implicitly 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. Recently, Pitassi and Raz [16] proved regular Resolution exponential lower bounds for PHP_n^m for any m. Finally Raz [18], and Razborov [19] simplifying [18], proved exponential lower bounds for the same principle in Resolution. 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 open whether lower bounds can be proved when the number of pigeons is greater than $n + c$. Regarding upper bounds, it is known that PHP_n^{2n} has quasipolynomial-size proofs in bounded-depth Frege [14, 15].

We work with the proof system 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 $\mathrm{PHP}_n^{n^{9/8-\epsilon}}$)

[1] Partially supported by ALCOM-FT, IST-99-14186.

[2] Supported by the CUR, Generalitat de Catalunya, through Grant 1999FI 00532.

[3] Partially supported by MEC through Grant PB98-0937-C04 (FRESCO project) and CICYT TIC 98-0410-C02-01.

[4] Partially supported by MEC through Grant PB98-0937-C04 (FRESCO project).

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 Res(log) (the analogue of Res(2) when we allow conjunctions of up to polylog literals). As a consequence of our lower bound, there is an exponential separation between Res(2) and 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 *et al.* [3]. Combining our techniques with those of [3], we also obtain an exponential-size lower bound for 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, combined with concentration bounds on martingales. The use of such bounds is to our knowledge novel in the field. The main technical contribution of our work consists in proving that a relatively short random restriction kills all large formulas of a 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, otherwise the initial clauses of the refutation would become trivial. 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 Res(2) is more powerful than Resolution. Here we prove that Resolution cannot polynomially simulate Res(2), and therefore Res(2) is superpolynomially more efficient than Resolution. As a corollary, we see that Res(2) does not have feasible monotone interpolation, proving this way a conjecture of Krajíček [13].

Another motivation for working with the system 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 Res(2) to be able to use it as a theorem prover.

2. DEFINITIONS AND OVERVIEW OF THE ARGUMENT

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 Res(k), defined by Krajíček [13], works with k-disjunctions. There are three inference rules in Res(k): Weakening, $\wedge$-Introduction, and Cut

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

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 Res(1) coincides with Resolution with the Weakening rule. The size of a Res(k)-refutation is the number of symbols in it. Mainly, we will work with 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 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^m_n, 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^m_n, then G-PHP^m_n coincides with the usual pigeonhole principle PHP^m_n. It is easy to see that a lower bound for the size of Res(2)-refutations of G-PHP^m_n implies the same lower bound for the size of Res(2)-refutations of PHP^m_n.

Ben-Sasson and Wigderson proved that whenever G is expanding in a sense defined next, every Resolution refutation of G-PHP^m_n 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 2.1 [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 2.1 [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^m_n 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-$\mathrm{PHP}^{cn'}_{n'}$ requires exponential-size Res(2)-refutations as follows. Assume, for contradiction, that Π is a small refutation of G-$\mathrm{PHP}^{cn'}_{n'}$. 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^m_n where all 2-disjunctions have less than $rf/2$ literals and $G(\rho)$ is (m, n, r, f)-expanding. This contradicts Theorem 2.1.

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^m_n will require exponential size 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 3.1. *$\Pr[\forall v \in V : mp/2 < \deg_G(v) < 2mp] \geq 1 - 2ne^{-mp/8}$ when G is drawn from $\mathcal{G}(m, n, p)$.*

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

LEMMA 3.2. *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, $\Pr[G$ is $(m, n, \alpha m, f)$-expanding$] \geq 1/2$.*

Proof. Fix $U' \subseteq U$ of size $s \leq \alpha m$ and $v \in V$. Then, $\Pr[v \in \partial U'] = sp(1-p)^{s-1}$. Let $q = \Pr[v \in \partial U']$. 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 $\mathrm{E}[|\partial U'|] = nq$. By Chernoff bound, $\Pr[|\partial U'| \leq nq/2] \leq e^{-nq/8}$. 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 $\Pr[|\partial U'| < f \cdot |U'|] \leq \Pr[|\partial U'| \leq nq/2] \leq e^{-nq/8} \leq e^{-snp/24}$. Finally, we bound the probability that G is not $(m, n, \alpha m, f)$-expanding by

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

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 $\sum_{s=1}^{\infty} \frac{1}{4^s} \leq \frac{1}{2}$. ■

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 let $\mathrm{ran}(\rho) = \{v_1, \ldots, v_r\}$.

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

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

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 these definitions we are ready to prove:

LEMMA 3.3. *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, $\Pr[G(\rho)$ is $(m, n, \alpha m, f)$-expanding$] \geq 1/2$.*

Proof. Let A be the event that $G(\rho)$ is $(m, n, \alpha m, f)$-expanding. Let us define $S = \{R \subseteq \{1, \ldots, n+r\} : |R| = r\}$. Then, splitting by disjoint cases we have that $\Pr[A] = \sum_{R \in S} \Pr[A \mid \mathrm{ran}(\rho) = R] \Pr[\mathrm{ran}(\rho) = R]$. Replacing V by $V - R$, the proof that $\Pr[A \mid \mathrm{ran}(\rho) = R] \geq 1/2$ is the same as in Lemma 3.2. The result follows. ■

LEMMA 3.4. *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*

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

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

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

Proof. Let (G, ρ) be drawn from distribution $\mathcal{D}(m, n, p, r)$. By Lemma 3.3 we have $\Pr[G(\rho)$ is $(m, n, \alpha m, f)$-expanding$] \geq 1/2$. Moreover by Lemma 3.1 we have that $\Pr[\forall v \in V : mp/2 < \deg_G(v) < 2mp] \geq 1 - (n + r)e^{-mp/9} \geq 5/6$. 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 $\Pr[E(G, \rho)] \geq 1/3$. On the other hand, $\Pr[E(G, \rho)] = \sum_H \Pr[E(G, \rho) \mid G = H] \Pr[G = H]$ where H ranges over all bipartite graphs on m and $n + r$ nodes. Therefore, there exists some fixed H such that $\Pr[E(G, \rho) \mid G = H] \geq 1/3$. Moreover, $\Pr[E(G, \rho) \mid G = H]$ equals $\Pr[E(H, \pi)]$ 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. ■

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 Res(2)-refutations of G-PHP$_n^m$. We claim that every Res(2)-refutation of G-PHP$_n^m$ can be turned into a 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 that will be central to the argument.

DEFINITION 3.1. We say that $(u, v) \in E(G)$ *hits* C if either $x_{u,v}$ occurs positively in C or $x_{u',v}$ occurs negatively in C for some $u' \in N_G(v) - \{u\}$. 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.

Note that an equivalent definition for *hits* is that (u, v) sets some literal of C to true. Observe that if the literal is free, it knocks the 2-disjunction, and if the literal is part of a conjunction, it will locally create a free literal. Finally, notice that a bad choice may or may not be a hit.

LEMMA 3.5. *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 disappears when we apply (u, v) to C, otherwise (u, v) would be a bad choice.

(ii) The literal is $\bar{x}_{u',v}$, and it appears in a conjunction of the form $\bar{x}_{u',v} \wedge y$. 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 disappears when we apply (u, v).

The lemma follows. ■

THEOREM 3.1. *Let $c > 1$ be a constant. For all sufficiently large n, every* 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 $m = kn = cn'$. By Lemma 3.4 there exists a graph $G = (U \cup V, E)$ with $|U| = m$ and $|V| = n + r$ such that

$$\Pr_{\rho \sim \mathcal{R}_r(G)}[G(\rho) \text{ is } (m, n, \alpha m, f)\text{-expanding}] \geq 1/3,$$

for $p = 48k \ln(m)/m$, $\alpha = 1/mp$, and $f = np/6$, and moreover $mp/2 \leq \deg_G(v) \leq 2mp$ for every $v \in \{1, \dots, n + r\}$.

For such a graph G, we show that every Res(2)-refutation of G-PHP has size at least $e^{n/(\log n)^{14}}$. This will imply the theorem since a Res(2)-refutation of $\mathrm{PHP}^{cn'}_{n'}$ gives a Res(2)-refutation of G-PHP of no bigger size. Let us assume, for contradiction, that G-PHP has a 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

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

This will suffice because then $\Pr[\exists C \in \Pi : C|_\rho \text{ is large}] \leq Se^{-n/(\log n)^{13}} < 1/3$, and also $\Pr[G(\rho)$ not $(m, n, \alpha m, f)$-expanding$] \leq 2/3$. 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 2.1.

For $i \in \{1, \dots, r\}$, 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 $\rho_i = ((u_1, v_1), \dots, (u_i, v_i))$. Observe that A_r implies A_j for every $i \in \{1, \dots, r\}$. Then,

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

We will show that every term in this expression is exponentially small. The bound on terms of the form $\Pr[A_j \wedge B_j]$ will be proven in Lemma 3.7. For the last term, we use an argument similar in spirit to the one by Beame and Pitassi [4]:

LEMMA 3.6. $\Pr[A_r \wedge \bigwedge_{j \geq r/2} \overline{B_j}] \leq e^{-n/(\log n)^8}$.

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

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

Fix $i \in \{r/2+1, \ldots, r\}$ and let H be the set of holes that occurs 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 nonempty since the right-degree of G is at least two. Therefore,

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

Therefore,

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

■

LEMMA 3.7. *Let j be such that $r/2 \leq j \leq r$. Then,* $\Pr[A_j \wedge B_j] \leq e^{-n/(\log n)^{11}}$.

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 $S = \sum_{i=1}^{j} S_i$ and $h = n/(\log n)^4$. Then,

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

We show that each term in this expression is exponentially small. More precisely, we show that $\Pr[A_j \wedge S < h] \leq e^{-n/(\log n)^3}$ and $\Pr[A_j \wedge B_j \wedge S \geq h] \leq e^{-n/(\log n)^{10}}$ which is clearly enough to prove Lemma 3.7.

CLAIM 3.1. $\Pr[A_j \wedge S < h] \leq e^{-n/(\log n)^3}$.

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,

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

Fix $i \in \{1, \ldots, j\}$. Let H be the set of holes that occurs 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 3.6 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

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

■

Claim 3.2. $\Pr[A_j \wedge B_j \wedge S \geq h] \leq e^{-n/(\log n)^{10}}$.

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 3.5). 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 $\Pr[A \wedge B \wedge S \geq h] \leq \Pr[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b]$. 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.

Let P_i^t denote the random variable $\Pr[T_i = t \mid \rho_0, \ldots, \rho_{i-1}]$ where $t \in \{\mathrm{k}, \mathrm{b}, \mathrm{n}\}$ and $i \in \{1, \ldots, j\}$. 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

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

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

Proof. Fix $i \in \{1, \ldots, j\}$ and a restriction $\rho = ((u_1, v_1), \ldots, (u_r, v_r))$. Our goal is to show that $P_i^{\mathrm{k}}(\rho) \geq P_i^{\mathrm{b}}(\rho)/\Delta$. First we define the following three sets: 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. ■

To complete the proof of Claim 3.2 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, $\Pr[|X_n - X_0| \geq \lambda] \leq 2e^{-\lambda^2/n}$ for every $\lambda > 0$ [11]. Now,

$$\Pr[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b] = \Pr[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b \wedge X_j \geq b/2] + \Pr[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b \wedge X_j < b/2].$$

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

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

The first inequality follows from Subclaim 3.1, 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. This ends the proof of Claim 3.2 and Lemma 3.7. ■

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

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

This ends the proof of Theorem 3.1. ■

By a different setting of parameters, it is easy to see that the strongest lower bound for PHP_n^m is of the form $e^{n^9/(m \log m)^8 \log^3 n}$. 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 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 3.1. *There is an exponential separation between* Res(2) *and* 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 [3, 10]. 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 4.1. Let F be a *CNF* formula. We say that a literal l is *pure* in F if some clause of F contains l and no clause of F contains $\bar{l}$. For a real number n, F is *n-sparse* if $|F| \leq n|v(F)|$ where $v(F)$ is the set of variables appearing in F.

For $s \geq 1$ and $\epsilon \in (0, 1)$, the following properties are defined for 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.

THEOREM 4.1 (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*. We will consider two probability distributions.

- $\mathcal{A}_t$ chooses a permutation of the variables uniformly at random and 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 and 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 restrictions. We only must show that any restriction ρ has the same probability in both distributions of probability.

LEMMA 4.1. *For every $x_1, \ldots, x_r$ and $v_1, \ldots, v_r$,*

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

Proof. The probability $\Pr_{\rho\sim\mathcal{B}_t}[\rho = ((x_1, v_1), \ldots, (x_r, v_r))]$ 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 $\Pr_{\rho\sim\mathcal{A}_t}[\rho = ((x_1, v_1), \ldots, (x_r, v_r))]$ 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 and 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. ■

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

LEMMA 4.2. *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 $\rho \sim \mathcal{A}_t$.*

(i) *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.*
(ii) *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.*

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, $\Pr[|C|_\rho| \geq d] \leq 1/M$.

THEOREM 4.2. *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,*

$$\Pr_{F\sim\mathcal{F}}[res\, 2(F) < M/2] \leq 2 \Pr_{F\sim\mathcal{F},\rho\sim\mathcal{R}}[F|_\rho \textit{ does not satisfy } A(s)] + 2 \Pr_{F\sim\mathcal{F},\rho\sim\mathcal{R}}[F|_\rho \textit{ does not satisfy } B_\epsilon(s)],$$

where $res\, 2(F)$ is the minimum size of a Res(2)*-refutation of F.*

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

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

The first inequality follows by Theorem 4.1, 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

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

Then, $res\, 2(F) < M/2$ implies that $\Pr_\rho[F|_\rho$ satisfies $A(s) \wedge B_\epsilon(s)] < 1/2$, and so $p(F) > 1/2$. Therefore, $\Pr_F[res\, 2(F) < M/2] \leq \Pr_F[p(F) > 1/2] \leq 2E_F[p(F)]$ by Markov's inequality. The result follows. ■

4.2. The Lower Bound Argument

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

THEOREM 4.3. *Let $F \sim \mathcal{F}^{3,n}_{5n}$. Then almost surely* Res(2)*-refutations of F require size $2^{\Omega(n^{1/3}/(\log(n))^2)}$.*

Proof. Let $m = 5n$ and $k = 3$, 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 4.2. 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 $\Pr_F[res2(F) < M/2] < 2p(F)$ by Theorem 4.2. Since $p(F)$ is $o(1)$ according to Lemma 4.2, 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 $\rho = ((x_1, v_1), \ldots, (x_r, v_r))$.

The whole argument is very similar to the one for Theorem 3.1. Some definitions need to be adapted, however. 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. 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

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

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 $A_{|\rho|}$.

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

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

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

We show now that $\Pr[A \mid |\rho| \geq t/2]$ 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,

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

We show that both terms in (9) are exponentially small. For every i such that $t/4 \leq i \leq t/2$, let K_i be the indicator random variable for the event that (x_i, v_i) is a knock. Then, as in the proof of Theorem 3.1, the second term in (9) is

$$\Pr\left[A \wedge \bigwedge_{j=t/4}^{t/2} \overline{B_j} \,\middle|\, D\right] \leq \prod_{i>t/4}^{t/2} \Pr\left[K_i = 0 \,\middle|\, \overline{B_{i-1}} \wedge \bigwedge_{j>t/4}^{i-1} K_j = 0 \wedge D\right] \leq e^{-\frac{n^{1/3}}{(\log(n))^3}}.$$

We will show that the first term in (9) is also exponentially small. First observe that

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

$$\leq \sum_{j=t/4}^{t/2} \Pr[A_j \wedge B_j \mid D]. \tag{11}$$

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

LEMMA 4.3. *If j is such that $t/4 \leq j \leq t/2$, then* $\Pr[A_j \wedge B_j \mid D] \leq e^{-n^{2/3}/(\log(n))^6}$.

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.

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

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.

CLAIM 4.3. $\Pr[A_j \wedge S < h \mid D] \leq e^{-n^{2/3}/(\log(n))^2}$.

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, as in the proof of Theorem 3.1,

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

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 $(a_1, \dots, a_j)$, we obtain

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

■

Now we will see what happens when the number of hits is big.

CLAIM 4.4. $\Pr[A_j \wedge B_j \wedge S \geq h \mid D] \leq e^{-n^{2/3}/(\log(n))^5}$.

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 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 $\Pr[A \wedge B \wedge S \geq h] \leq \Pr[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b]$. 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.

SUBCLAIM 4.2. $\Pr[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b] \leq e^{-n^{2/3}/(\log(n))^5}$.

Proof. For $t \in \{\mathrm{k}, \mathrm{b}, \mathrm{n}\}$ and $i \in \{1, \dots, j\}$, let P_i^t denote the random variable $\Pr[T_i = t \mid \rho_0, \dots, \rho_{i-1}]$. We define a martingale $X_0, \dots, X_j$ with respect to $\rho_0, \dots, \rho_j$ 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 $T_{i+1} = \mathrm{b}$. As in the proof of Theorem 3.1, it is easy to see that $\{X_i\}_i$ is indeed 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 $Y_0, \dots, Y_j$ 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, $Y_j = S^{\mathrm{k}} - \sum_{i=1}^{j} P_i^{\mathrm{k}}$.

In the next calculation we will use the fact that $P_i^{\mathrm{k}}(\rho) = P_i^{\mathrm{b}}(\rho)$ for every ρ and $i \in \{1, \dots, j\}$.

$$
\Pr[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b] = \Pr[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b \wedge X_j \geq b/2] + \Pr[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b \wedge X_j < b/2].
$$

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

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

by Azuma's inequality again. Therefore, the sum is bounded by $4e^{-t/32(\log(t))^4} \leq e^{-n^{2/3}/(\log(n))^5}$ as required. This ends the proof of Subclaim 4.2 and Claim 4.4. ■

With both sublemmas proved, so is Lemma 4.3. We are ready to complete the proof of our goal (7). We have shown that

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

This ends the proof of Theorem 4.3. ■

We give another proof of Subclaim 4.2 that does not require martingales.

Subclaim 4.3. $\Pr[S^{\mathrm{k}} = 0 \wedge S^{\mathrm{b}} \geq b] \leq 2^{-b}$

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 forms 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 preceding 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 a bad choice neither 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 produces 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 restrictions in F. So

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

■

5. SEPARATION BETWEEN RES(2) AND RESOLUTION

In this section we prove that Resolution cannot polynomially simulate Res(2). More precisely, we prove that a certain Clique-Coclique principle, as defined by Bonet *et al.* in [6], has polynomial-size Res(2)-refutations, but every Resolution 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:

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

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

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

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

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

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

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

THEOREM 5.1. *Let $k' < k \le n$. If $\mathrm{PHP}^k_{k'}$ has Resolution refutations of size S, then $\mathrm{CLIQUE}^n_{k,k'}$ has* Res(2)*-refutations of size Sn^c for some constant $c > 0$.*

Proof. We use the following 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:

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

We show in detail how to get clauses (1) from clauses (12) and (15), and just sketch how to get the initial clauses (2) and how to simulate a resolution step.

For the first part, if we expand clause (1) for a certain i we have:

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

We apply successively for $1 \le j \le k'$ 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 $1 \le j \le k'$ 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:

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

Now it is clear how to get (18).

Now we will sketch 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 $p_{j,1} \vee \cdots \vee p_{j,k'}$ as before. 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 $p_{j,1} \vee \cdots \vee p_{j,t-1} \vee p_{j,t+1} \vee \cdots \vee p_{j,k'}$. It is clear that $\bar{p}_{i,t} \vee \bar{p}_{j,t}$ is $A \vee B$. Now it is easy to see how to get $A \vee B$ from (21), (22), and (17).

It remains to sketch how to simulate a 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 the first one:

$$(x_{i,1} \wedge y_{j,1}) \vee (x_{i,2} \wedge y_{j,2}) \vee (x_{i,3} \wedge y_{j,3}) \vee \cdots \vee (x_{i,n} \wedge y_{j,n}) \vee A. \tag{23}$$

If we get clauses $\bar{x}_{i,l} \vee \bar{y}_{j,l} \vee B$ for $1 \leq l \leq n$, we cut them all with (23) and get $A \vee B$ as desired. To get, for example, $\bar{x}_{i,1} \vee \bar{y}_{j,1} \vee B$, we cut $\bar{p}_{i,j} \vee B$ with $\bar{y}_{j,1} \vee \bar{y}_{l,1}$, $l \neq j$ of course, and with clauses $\bar{x}_{i,1} \vee \bar{x}_{i,l}$, $l \neq 1$. ■

In order to show that the $\mathrm{CLIQUE}^n_{k,k'}$ principle requires superpolynomial size Resolution proofs, we will use two results. One is the monotone interpolation theorem (see [17]) for Resolution. This theorem allows us to build a monotone circuit computing a certain function (in this case the clique function), from a Resolution refutation of a set of clauses that expresses contradictory facts about the minterms and maxterms of the function. Also the size of the circuit is polynomial in the size of the proof. The other result is Theorem 5.2 from [1] which is stated below in order to carry out the calculations. This theorem establishes a lower bound on the size of monotone circuits that separate large cliques from small cocliques. The general argument of our result is by contradiction. Assume there is a short Resolution refutation of the clauses (12)–(17). By the monotone interpolation theorem there is a small monotone circuit separating large cliques from small cocliques, which is impossible by Theorem 5.2. Now comes the statement of the theorem:

Let $F(m, k, k')$ be the set of monotone functions that separate k-cliques from k'-cocliques on m nodes.

THEOREM 5.2 (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 5.3. *Let $k = \sqrt{m}$ and $k' = (\log m)^2/8 \log\log m$. Then* (i) $\mathrm{CLIQUE}^m_{k,k'}$ *has* 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 5.1, $\mathrm{CLIQUE}^m_{k,k'}$ has Res(2)-refutations of size polynomial in m. Regarding (ii), suppose for contradiction that $\mathrm{CLIQUE}^m_{k,k'}$ has a Resolution refutation of size $\exp(o((\log m)^2/\sqrt{\log\log m}))$. By the monotone interpolation theorem, we obtain a monotone circuit C separating k-cliques from k'-cocliques of size $\exp(o((\log m)^2/\sqrt{\log\log m}))$. Note that

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

Therefore, by Theorem 5.2, the size of C is at least

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

which is $\exp(\Omega((\log m)^2/\sqrt{\log\log m}))$. This is a contradiction. ■

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

COROLLARY 5.1. 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 Res(2). It is an open question whether the separation could be exponential or a quasipolynomial simulation of 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 $2^{n^{\epsilon}}$.

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 Res(2). We conjecture that Res(2) does not have a strong size–width trade-off. Notice that Res(log) does not have it. This is because (a) Res(log) is equivalent to depth-0.5 LK, (b) PHP_n^{2n} has quasipolynomial-size proofs in depth-0.5 LK [14], and (c) PHP_n^{2n} has $\Omega(n)$ width lower bounds for 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 Res(3), Res(4), ..., Res(log). It seems that some new ideas need to be developed to do that. This question is related to the optimality of the Res(log) upper bound for PHP_n^{2n}.

Finally, we note that exponential-size lower bounds for $\mathrm{PHP}_n^{n^{1+\epsilon}}$ in Res(k) implies lower bounds for $\mathrm{PHP}_n^{n^c}$ in Resolution for some c. In particular lower bounds for $\mathrm{PHP}_n^{n^{1.5}}$ in Res(2) imply lower bounds for $\mathrm{PHP}_n^{n^2}$ in Resolution.

REFERENCES

1. Alon, N., and Boppana, R. (1987), The monotone circuit complexity of boolean functions, *Combinatorica* **7**(1), 1–22.
2. Beame, P., Impagliazzo, R., Krajíček, J., Pitassi, T., Pudlák, P., and Woods, A. (1992), Exponential lower bounds for the pigeonhole principle, *in* "24th Annual ACM Symposium on Theory of Computing," pp. 200–220.
3. Beame, P., Karp, R., Pitassi, T., and Saks, M. (2002), The efficiency of resolution and Davis-Putnam procedures, *SIAM Journal on Computing* **31**(4), 1048–1075.
4. Beame, P., and Pitassi, T. (1996), Simplified and improved resolution lower bounds, *in* "37th Annual IEEE Symposium on Foundations of Computer Science," pp. 274–282.
5. Ben-Sasson, E., and Wigderson, A. (2001), Short proofs are narrow: Resolution made simple, *J. Assoc. Comput. Mach.* **48**(2), 149–169.
6. Bonet, M., Pitassi, T., and Raz, R. (1997), Lower bounds for cutting planes proofs with small coefficients, *J. Symbolic Logic* **62**(3), 708–728.
7. Buss, S., and Pitassi, T. (1997), Resolution and the weak pigeonhole principle, *in* "11th International Workshop on Computer Science Logic," Lecture Notes in Computer Science, Vol. 1414, pp. 149–156, Springer-Verlag, Berlin.
8. Buss, S. R. (1987), Polynomial size proofs of the propositional pigeonhole principle, *J. Symbolic Logic* **52**(4), 916–927.
9. Buss, S. R., and Turán, G. (1988), Resolution proofs on generalized pigeonhole principles, *Theoret. Comput. Sci.* **62**(3), 311–317.
10. Chvátal, V., and Szemerédi, E. (1988), Many hard examples for resolution, *J. Assoc. Comput. Mach.* **35**(4), 759–768.
11. Grimmet, G. R., and Stirzaker, D. R. (1982), "Probability and Random Processes," Oxford Science Publications, Oxford.
12. Haken, A. (1985), The intractability of resolution, *Theoret. Comput. Sci.* **39**(2/3), 297–308.
13. Krajíček, J. (2002), On the weak pigeonhole principle, *Fundamenta Mathematicæ* **170**(1–3), 123–140.
14. Maciel, A., Pitassi, T., and Woods, A. (2000), A new proof of the weak pigeonhole principle, *in* "32nd Annual ACM Symposium on Theory of Computing," pp. 368–377.
15. Paris, J. B., Wilkie, A. J., and Woods, A. R. (1988), Provability of the pigeonhole principle and the existence of infinitely many primes, *J. Symbolic Logic* **53**(4), 1235–1244.
16. Pitassi, T., and Raz, R. (2001), Regular resolution lower bounds for the weak pigeonhole principle, *in* "33rd Annual ACM Symposium on Theory of Computing," pp. 347–355.
17. Pudlák, P. (1997), Lower bounds for resolution and cutting plane proofs and monotone computations, *J. Symbolic Logic* **62**(3), 981–998.
18. Raz, R. (2002), Resolution lower bounds for the weak pigeonhole principle, "Electronic Colloquium on Computational Complexity," TR01-021. 34th Annual ACM Symposium on Theory of Computing, pp. 553–562.
19. Razborov, A. (2001), Improved resolution lower bounds for the weak pigeonhole principle, manuscript.