Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 175 (2016)

Random resolution refutations

Pavel Pudlák and Neil Thapen*

September 20, 2018

Abstract

We study the *random resolution* refutation system defined in [Buss et al. 2014]. This attempts to capture the notion of a resolution refutation that may make mistakes but is correct most of the time. By proving the equivalence of several different definitions, we show that this concept is robust. On the other hand, if $\mathbf{P} \neq \mathbf{NP}$, then random resolution cannot be polynomially simulated by any proof system in which correctness of proofs is checkable in polynomial time.

We prove several upper and lower bounds on the width and size of random resolution refutations of explicit and random unsatisfiable CNF formulas. Our main result is a separation between polylogarithmic width random resolution and quasipolynomial size resolution, which solves the problem stated in [Buss et al. 2014]. We also prove exponential size lower bounds on random resolution refutations of the pigeonhole principle CNFs, and of a family of CNFs which have polynomial size refutations in constant depth Frege.

1 Introduction

The following system for refuting propositional CNFs was introduced in [6]. Let F be a CNF in variables $x_1, \ldots, x_n$ and let $0 < \varepsilon < 1$.

Definition 1.1 *An* ε-random resolution distribution, *or* ε-RR distribution, *of F is a probability distribution $\mathcal{D}$ on pairs $(B_i, \Pi_i)_{i \sim \mathcal{D}}$ such that*

1. *for each $i \in \mathcal{D}$, B_i is a CNF in variables $x_1, \ldots, x_n$ and Π_i is a resolution refutation of $F \wedge B_i$*
2. *for every $\alpha \in \{0,1\}^n$, $\Pr_{i \sim \mathcal{D}}[B_i$ is satisfied by $\alpha] \geq 1 - \varepsilon$.*

The size *and the* width *of $\mathcal{D}$ are defined respectively as the maximum size and maximum width of the refutations Π_i (if these maxima exist).*

*Institute of Mathematics, Czech Academy of Sciences. The authors were supported by the ERC Advanced Grant 339691 (FEALORA)

ISSN 1433-8092

We will see later that, without loss of generality, we can assume that the support of $\mathcal{D}$ is finite, in which case the maxima and minima exist (Lemma 2.2). This definition was first proposed by Stefan Dantchev. Its appearance in [6] is ultimately motivated by an open problem in bounded arithmetic, which we will explain in a moment, after we mention some basic properties and equivalent formulations.

It is sound as a refutational system, in the sense that if F has an ε-RR distribution then F is unsatisfiable. To see this, consider any assignment $\alpha \in \{0,1\}^n$. Since $\varepsilon < 1$, there is at least one pair (B_i, Π_i) such that α satisfies B_i and Π_i is a resolution refutation of $F \wedge B_i$. So α cannot also satisfy F, by the soundness of resolution. The system is also complete, since resolution is complete and we can take $\mathcal{D}$ to consist of a single pair (B, Π) where B is any tautology and Π is a (possibly exponential sized) resolution refutation of F.

On the other hand it is not a propositional proof system in the sense of Cook and Reckhow [10], because it is defined by a semantic condition that, as we will show, cannot be tested in polynomial time unless $\mathbf{P} = \mathbf{NP}$ (Proposition 3.3). Nevertheless it makes perfect sense to compare the complexity of proofs in it with proofs in the standard proof systems, in particular with resolution and bounded depth Frege. We prove some results in this direction in this work. Note also that the definition is particular to resolution, and we must take care if we try to generalize it. For example, if we instead define a random Frege distribution system, in which B and Π can contain arbitrary formulas, then we can trivially refute any unsatisfiable F by setting $B = \neg F$.

As with some concepts of probabilistic computation studied in computation complexity theory, one can use the linear programming duality, often referred to as *Yao's minimax principle*, to give an equivalent definition of the system based on probability distributions over inputs rather than over proofs (see Definition 2.5). This is very useful if one needs to prove lower bounds. Another essentially equivalent formulation is in terms of semantic resolution derivations. This means, roughly speaking, that instead of having an auxiliary formula that is satisfied with high probability, we consider semantic derivations with respect to a large subset of assignments (all except a fraction ε), where lines in the proof are clauses. In a sense, this captures better the intuitive idea of a proof with errors.

Let us also mention that while we tend to think of the error ε as something small, there is a simple amplification lemma that allows us to shrink the error at some cost in proof size. Thus, for the questions we are interested in, without loss of generality we can take $\varepsilon = \frac{1}{2}$.

Our main result is that the propositional translation of the *coloured polynomial local search* principle CPLS [20], which has polynomial size resolution refutations, does not have narrow 1/2-random resolution distributions. We also prove some size lower bounds. Previously, lower bounds on random resolution have only been known for treelike refutations [6] or for relatively small errors ε [17].

The proof is based on a lemma that looks like a rudimentary version of the switching lemmas used in propositional proof complexity (see the discussion at the start of Section 4). Although our results do not give a new separation in bounded arithmetic, we believe that they are interesting for other reasons. It has been conjectured that in order to separate higher fragments of bounded arithmetic, we only need switching lemmas for certain more complicated tautologies, similar to CPLS (see for example [23]). Nevertheless, all attempts in this direction have failed so far because of the complexity of the associated combinatorial problems. Our proof gives us some hope that eventually it will be possible to prove such lemmas.

Background in bounded arithmetic. A central problem in this area is to show, in the relativized setting where an undefined relation symbol has been added to the language, that the set $\forall\Sigma^b_1(T^i_2)$, consisting of all $\forall\Sigma^b_1$ consequences of the bounded arithmetic theory T^i_2, gets strictly bigger as i increases. It is known that $\forall\Sigma^b_1(T^2_2)$ is bigger than $\forall\Sigma^b_1(T^1_2)$ [8], but there are no higher separations known, and it cannot be ruled out that $\forall\Sigma^b_1(T^2_2)$ is already the same as $\forall\Sigma^b_1(T_2)$.

There is a more-or-less equivalent question in propositional proof complexity: show that the set of polylogarithmic width CNFs with quasipolynomial size refutations in depth i Frege strictly increases as i increases.[1] Here the system $R(\log)$ [16], which can be thought of as depth $\frac{1}{2}$ Frege, corresponds to the theory T^2_2 and corresponding separations are known. That is, we can separate $R(\log)$ from weaker fragments of constant depth Frege, but not from stronger ones (see for example [22]). Another essentially equivalent question is whether we can show oracle separations for a certain family of **TFNP** search problem classes [23]. Of course the corresponding problem in Boolean circuit complexity, that is, separation of the classes of functions computed by circuits of depth i, was solved a long time ago.

Since the problem of separating $\forall\Sigma^b_1(T^2_2)$ from $\forall\Sigma^b_1(T_2)$ has been notoriously open for many years, it was proposed in [6] to consider, in place of T^2_2, a theory of similar strength but of a rather different nature, namely Jeřábek's theory of approximate counting [13]. This theory, called APC_2 in [6], consists of $T^1_2+\mathrm{sWPHP}(\mathrm{PV}_2)$ where PV_2 is a formalization of the $\mathrm{FP}^{\mathrm{NP}}$ functions and $\mathrm{sWPHP}(\mathrm{PV}_2)$ expresses that no such function is a surjection from $[a]$ onto $[2a]$, for any a.

A separation of $\forall\Sigma^b_1(\mathrm{APC}_2)$ from $\forall\Sigma^b_1(T_2)$ is still open, but [6] does show separations for subtheories of APC_2. However, it leaves the following as an open problem: show an $\forall\Sigma^b_1$ separation between $T^1_2 + \mathrm{sWPHP}(\mathrm{PV}_1)$ and T_2, where $\mathrm{sWPHP}(\mathrm{PV}_1)$ is the surjective weak pigeonhole principle only for FP functions.

A reason why this problem was interesting was the following proposition, which reduces it to a natural-looking question about the complexity of propositional proofs.

Proposition 1.2 ([6]) *Suppose $T^1_2+\mathrm{sWPHP}(\mathrm{PV}_1) \vdash \forall n \exists y < t(n)\, \theta(n, y)$, where θ is sharply bounded, and everything is relativized. Then the propositional translation $\langle \forall y < t(n)\, \neg\theta(n, y)\rangle$ has a polylogarithmic width $1/q(n)$-RR distribution, for q any quasipolynomial.*

Hence one way to solve the open problem mentioned above would be to prove width lower bounds on random resolution, for any CNF which has quasipolynomial size constant depth

[1] We emphasize that we are interested in this question for quasipolynomial size proofs. This matches the natural question in bounded arithmetic, and a separation for polynomial size is known [11], using a padded pigeonhole principle $\mathrm{PHP}_{(\log n)^k}$ which has short proofs in some depth i, but is such that the exponential size lower bound for PHP in depth $i-1$ gives a quasipolynomial lower bound for the padded version.

Frege refutations. This is what we show here, for CPLS. However, the problem specifically about T^1_2 + sWPHP(PV$_1$) was solved in [3], without using Proposition 1.2 or proving a lower bound on random resolution — instead the proof used the properties of the formula sWPHP(PV$_1$) in an essential way.

Contents of this paper. In Section 2 we fix our notation, prove some basic facts about random resolution distributions and present some equivalent or almost equivalent definitions, in terms of refutations with respect to a distribution over assignments, semantic resolution refutations, and refutations with random extension clauses.

In Section 3 we prove some upper bounds. We show that random 3-CNFs (Proposition 3.2) and the retraction weak pigeonhole principle (Proposition 3.5) have narrow, and thus quasipolynomial size, 1/2-RR distributions, while they require exponential sized refutations in standard resolution. Using the upper bound on random 3-CNFs we show that, under a standard hypothesis, 1/2-RR is not simulated by any Cook-Reckhow proof system (in Proposition 3.3). We also discuss random resolution refutations with respect to the uniform distribution on assignments (defined below) and observe that this system is unreasonably strong.

In Section 4 we describe our approach to lower bounds, and then use it first to prove that the bit pigeonhole principle requires exponential size 1/2-RR distributions (Theorem 4.3).

In Section 5, using the same approach, we show our main result, that there is a family of logarithmic width CNFs which have polynomial size resolution refutations, but do not have polylogarithmic width 1/2-RR distributions (Theorem 5.10). This answers the open problem about random resolution posed in [6].

In Section 6 we adapt the argument from Section 5 to show that there is a family of logarithmic width CNFs which have polynomial size refutations in constant depth Frege (in fact in Res(2)), but require exponential size 1/2-RR distributions.

In Section 7 we prove a form of feasible interpolation for treelike random resolution (Theorem 7.1) and derive size lower bounds on treelike RR distributions for a family of 3-CNFs (Corollary 7.2).

In Section 8 we show that the usual formalization of the pigeonhole principle requires exponential size 1/2-RR distributions (Theorem 8.1).

Finally in Section 9 we briefly discuss "random" versions of stronger proof systems than resolution, and show that the argument of Section 8 can be extended to prove exponential size lower bounds on 1/2-random constant depth Frege distributions for the pigeonhole principle.

Acknowledgments. We would like to thank Pavel Hrubeš and Jan Krajíček for discussions and valuable suggestions. In particular, the proof of Proposition 3.6 generalizes an idea of Hrubeš. We are grateful to Michal Garlík for pointing out some errors in an earlier version of the paper.

2 Basic properties and alternative definitions

We first introduce some notation. We identify CNF formulas with sets of clauses. We use 0 (false) and 1 (true) to represent truth values. For a CNF formula F and an assignment α of truth values to its variables, we denote by $F[\alpha]$ the truth value to which the formula is evaluated by α. If ρ is a partial assignment, we denote by F^ρ the CNF formula obtained by substituting ρ into F and simplifying the formula (that is, deleting false literals and omitting true clauses).

The *width* of a clause is the number of literals it contains. The *width* and *size* of a refutation are respectively the width of its widest clause and the total number of clauses. A k-CNF is a CNF in which every clause has width at most k.

We will often use the notation

$$p_1 \wedge \cdots \wedge p_r \rightarrow q_1 \vee \cdots \vee q_s$$

to stand for the clause $\neg p_1 \vee \cdots \vee \neg p_r \vee q_1 \vee \cdots \vee q_s$, where $p_1, \ldots, p_r, q_1, \ldots, q_r$ can be any literals. In this notation the resolution rule can, for example, have the form: from $A \wedge p \rightarrow C$ and $A \wedge \neg p \rightarrow D$ conclude $A \rightarrow C \vee D$, where p is a literal, A is a conjunction of literals and C and D are clauses.

If p is a literal, we will sometimes write $p = 1$ instead of the literal p and $p = 0$ instead of the literal $\neg p$. Similarly we will write $p \neq 1$ or $p \neq 0$ to mean respectively $\neg p$ or p. If $p_1, \ldots, p_r$ are literals and $\beta \in \{0,1\}^r$ we write $\bar{p} = \beta$ to stand for the conjunction $\bigwedge_{1 \leq i \leq k} p_i = \beta_i$ where each conjunct is formally either p_i or its negation, as above; and $\bar{p} \neq \beta$ to stand for the disjunction $\bigvee_{1 \leq i \leq k} p_i \neq \beta_i$. The following observation will be useful:

Fact 2.1 *The CNF $\bigwedge\{\bar{p} \neq \beta : \beta \in \{0,1\}^r\}$ has a refutation of width r, using $2^r - 1$ resolution steps.*

We write $[n]$ for $\{0, \ldots, n-1\}$. When we formalize combinatorial principles as CNFs, if the principle involves a function $f : [n] \rightarrow [m]$ we will often formalize f by introducing variables for its "bit-graph". That is, for each $x < n$ we introduce $\log m$ variables $(f(x))_0, \ldots, (f(x))_{\log m - 1}$ representing the value of $f(x)$ in binary. For the sake of simplicity, in this situation we will assume that m is a power of 2. For $y < n$, we will write $f(x) = y$ to stand for the conjunction $\bigwedge_i (f(x))_i = \beta_i$, where $\beta \in \{0,1\}^{\log m}$ is y written in binary, and we will write $f(x) \neq y$ for the disjunction $\bigvee_i (f(x))_i \neq \beta_i$.

A standard example here is the bit pigeonhole principle BPHP_n (see Section 4 below). Suppose $n = 2^k$. Then BPHP_n is a contradictory propositional CNF asserting that f is an injection from $[n+1]$ to $[n]$. In our notation, it consists of clauses

$$f(x) \neq y \vee f(x') \neq y$$

for all $x < x' < n+1$ and all $y < n$. Each clause has width $2k$.

Because we deal with propositional refutation systems, rather than proof systems, for us the natural translation into propositional logic of a true first order principle, such as the pigeonhole principle PHP, is a family of unsatisfiable CNFs that we want to refute, rather than a family of tautologous DNFs that we want to prove. Therefore we will use the same name, PHP, for both this family of CNFs and the original principle. It should be clear from the context which is meant, and the propositional version will often be written with a size parameter, for example as PHP_n.

2.1 Random resolution distributions

In the rest of this section, let F be a CNF in variables $x_1, \dots, x_n$ and let $0 < \varepsilon < 1$.

Our definition of the size of an ε-RR distribution above does not take into account the size of the sample space (that is, of the set of pairs (B, Π) appearing in $\mathcal{D}$). We show now that the size of the sample space can be bounded, at the cost of slightly increasing the error ε.

Lemma 2.2 *If F has an ε-RR distribution, then it also has a 2ε-RR distribution of the same size and width, in which the sample space has size $O(n/\varepsilon)$.*

Proof. Given the probability distribution $(B_i, \Pi_i)_{i \sim \mathcal{D}}$, consider m random samples $i_1, \dots, i_m$. By the Chernoff bound, for each α the probability that $B_{i_j}[\alpha] \neq 1$ for more than a fraction 2ε of the m samples is at most $\mathrm{e}^{-\varepsilon m/3}$. The probability that this happens for *some* α is at most 2^n times larger. Hence if $m > \frac{3 \ln 2}{\varepsilon} n$, there is nonzero probability that this does not happen for any α. We pick a set of samples for which it does not happen, and let the new distribution $\mathcal{D}'$ be given by the uniform distribution on $\{i_1, \dots, i_m\}$. ■

We next show an "amplification" result, that we can decrease the error ε at the cost of increasing the width and size.

Lemma 2.3 *Suppose F has an ε-RR distribution of width w and size s. Then for every $k \geq 1$ it also has an ε^k-RR distribution of width at most kw and size $O(s^k)$.*

Proof. Given $(B_i, \Pi_i)_{i \sim \mathcal{D}}$, take the distribution $(B_{i_1 \dots i_k}, \Pi_{i_1 \dots i_k})_{(i_1 \dots i_k) \sim \mathcal{D}^k}$, where $B_{i_1 \dots i_k}$ is the CNF formula obtained from $\bigvee_{\ell=1}^k B_{i_\ell}$ by applying the distributive law and $\Pi_{i_1 \dots i_k}$ is a refutation composed of the refutations $\Pi_{i_1}, \dots, \Pi_{i_k}$. In more detail, every clause C of $B_{i_1 \dots i_k}$ has the form $C_1 \vee \dots \vee C_k$ where each C_ℓ is a clause of B_{i_ℓ}. For every fixed tuple $C_2, \dots, C_k$ we can use Π_{i_1} to derive $C' := C_2 \vee \dots \vee C_k$ from C and F. Then for every fixed tuple $C_3, \dots, C_k$ we can use Π_{i_2} to derive $C_3 \vee \dots \vee C_k$ from C' and F, and so on until we derive the empty clause. ■

Corollary 2.4 *Let "small" mean either "polylogarithmic width" or "quasipolynomial size". Then for any quasipolynomial $q(n)$, F has a small $1/2$-RR distribution if and only if F has a small $1/q(n)$-RR distribution.*

The system we are most interested in is *narrow random resolution*. This is 1/2-RR with polylogarithmic width (and hence quasipolynomial size) distributions.

2.2 Random resolution refutations

Definition 2.5 *Let Δ be a probability distribution on $\{0,1\}^n$. An* (ε, Δ)-random resolution refutation*, or* (ε, Δ)-RR refutation*, of F is a pair (B, Π) such that*

1. *B is a CNF in variables $x_1, \dots, x_n$ and Π is a resolution refutation of $F \wedge B$*
2. *$\Pr_{\alpha \sim \Delta}[B[\alpha] = 1] \geq 1 - \varepsilon$.*

Note that this definition is, in general, not sound. In particular, let F be any (nonempty) CNF whatsoever. Let C be any clause from F and let α be an assignment which falsifies C. Let Δ be the distribution that puts all its weight on the single assignment α, and let B be the CNF $\bigwedge_i x_i = x_i[\alpha]$. Then B is true with probability 1 over Δ, and we can easily derive the empty clause from $F \wedge B$, since B contains the negation of each literal in C as a singleton clause. See also Section 3.4 below.

However, if an (ε, Δ)-RR refutation exists for *all* distributions Δ, then this is equivalent to the existence of an ε-RR distribution, as follows.

Proposition 2.6 *The following are equivalent.*

1. *F has an ε-RR distribution of width w and size s.*
2. *F has an (ε, Δ)-RR refutation of width w and size s for every distribution Δ on $\{0,1\}^n$.*

Proof. This is an immediate consequence of the minimax theorem. Consider a zero-sum game between two players, called the Prover and the Adversary, in which the Prover picks a pair (B, Π) such that B is a CNF and Π is a refutation of $F \wedge B$ of width w and size s, and the Adversary picks an assignment α. The payoff is $B[\alpha]$, that is, the Prover gets 1 if α satisfies B and 0 otherwise.

Then condition 1 says that the Prover has a mixed strategy to achieve a payoff of at least $1 - \varepsilon$, and condition 2 says that the Adversary does not have a mixed strategy to achieve a payoff less than $1 - \varepsilon$. By the minimax theorem these statements are equivalent. ■

Our main use of RR refutations will be to prove lower bounds on RR distributions, by carefully choosing a suitable Δ. However Proposition 2.6 as written will not be quite enough for us, as in it Δ is a distribution on *total* assignments, whereas we will, as is usual in lower bound arguments, need to use distributions on *partial* assignments, or in other words, *random restrictions*. For this reason we will need the following slightly more general versions of the definition and proposition above.

Definition 2.7 *Let $\mathcal{R}$ be a distribution of partial assignments to the variables $x_1, \dots, x_n$. An* $(\varepsilon, \mathcal{R})$-RR refutation *of F is a pair (B, Π) such that*

1. *B is a CNF and Π is a resolution refutation of $F \wedge B$*
2. *$\Pr_{\rho \sim \mathcal{R}}[B^\rho = 0] \leq \varepsilon$.*

Proposition 2.8 *The following are equivalent.*

1. *F has an ε-RR distribution of width w and size s.*

2. *F has an $(\varepsilon, \mathcal{R})$-RR refutation of width w and size s for every distribution $\mathcal{R}$ over partial assignments.*

Proof. Suppose condition 1 holds and let $\mathcal{R}$ be any distribution over partial assignments. Define a distribution Δ over total assignments as follows: choose $\rho \sim \mathcal{R}$ at random, then extend it to a total assignment by setting all unset variables to 0. Let (B, Π) be the (ε, Δ)-RR refutation of F given by Proposition 2.6, so that $\Pr_{\alpha \sim \Delta}[B[\alpha] = 0] \leq \varepsilon$. By the construction of Δ it follows that $\Pr_{\rho \sim \mathcal{R}}[B^\rho = 0] \leq \varepsilon$, and thus (B, Π) is also a $(\varepsilon, \mathcal{R})$-RR refutation of F.

The other direction is immediate from Proposition 2.6. ■

2.3 Semantic resolution refutations

Semantic derivations were introduced in [15]. We will use the special case defined by clauses.

Definition 2.9 *Let $\mathcal{A} \subseteq \{0,1\}^n$ be a nonempty set of truth assignments. We say that a formula C is a* semantic consequence over *$\mathcal{A}$ of formulas $C_1, \dots, C_r$, written $C_1, \dots, C_r \models^{\mathcal{A}} C$, if every assignment in $\mathcal{A}$ that satisfies $C_1, \dots, C_r$ also satisfies C.*

A degree d semantic resolution refutation of F over $\mathcal{A}$ *is a sequence* Π *of* clauses*, ending with the empty clause, in which every clause either belongs to F or is a semantic consequence over $\mathcal{A}$ of at most d earlier clauses.*

For the sake of simplicity we will only consider degree 2 semantic resolution refutations, which we will simply call *semantic resolution refutations.*

Definition 2.10 *Let Δ be a probability distribution on $\{0,1\}^n$. An* (ε, Δ)-semantic refutation *of F is a pair $(\mathcal{A}, \Pi)$ such that*

1. Π *is a semantic refutation of F over $\mathcal{A}$, and*

2. $\Pr_{\alpha \sim \Delta}[\alpha \in \mathcal{A}] \geq 1 - \varepsilon$.

One can also define an *ε-semantic resolution distribution* and prove that it is equivalent to the existence of (ε, Δ)-semantic resolution refutations for all distributions Δ, in the same manner as our proof in Proposition 2.6 that an ε-RR distribution is equivalent to the existence of (ε, Δ)-RR refutations for all distributions Δ. We leave this easy exercise to the reader.

Proposition 2.11 *If F has an (ε, Δ)-RR refutation of width w and size s, then it also has an (ε, Δ)-semantic resolution refutation of width $\leq w$ and size $\leq s$.*

In the opposite direction, if F has an (ε, Δ)-semantic refutation of width w and size s, then it also has an (ε, Δ)-RR refutation of width $O(w)$ and size at most $O(sw^2)$.

Proof. The first part follows immediately by letting $\mathcal{A}$ be the set of assignments that satisfy the auxiliary CNF B from the RR refutation. For the second part, let $(\mathcal{A}, \Pi)$ be the (ε, Δ)-semantic refutation of F. We may assume, by adding dummy premises as necessary, that every semantic consequence step in Π has exactly two premises.

Suppose C_1, C_2 and C are clauses such that $C_1, C_2 \vDash^{\mathcal{A}} C$. We claim that for all literals $x \in C_1$ and $y \in C_2$ we have $\vDash^{\mathcal{A}} x \wedge y \to C$. For otherwise there would exist some $\alpha \in \mathcal{A}$ with $C \cup \{\neg x, \neg y\}[\alpha] = 0$, which implies that $C[\alpha] = 0$ and $x[\alpha] = y[\alpha] = 1$, whence $C_1[\alpha] = C_2[\alpha] = 1$, which contradicts the assumption. Let $B_{C_1,C_2,C}$ be the CNF

$$\bigwedge\{x \wedge y \to C : x \in C_1,\, y \in C_2\}$$

(the clauses of this CNF may contain repeated variables).

Let B be the conjunction of the CNFs $B_{C_1,C_2,C}$ over all semantic consequence steps $C_1, C_2 \vDash^{\mathcal{A}} C$ in the semantic refutation Π. By the claim, $B[\alpha] = 1$ for every $\alpha \in \mathcal{A}$. Hence $\Pr_{\alpha \sim \Delta}[B[\alpha] = 1] \geq \Pr_{\alpha \sim \Delta}[\alpha \in \mathcal{A}] \geq 1 - \varepsilon$.

It remains to construct a small resolution refutation of $F \wedge B$. We can derive C from C_1, C_2 and $B_{C_1,C_2,C}$ as follows. For each $y \in C_2$, we derive $y \to C$ by resolving C_1 with all clauses in the set $\{x \wedge y \to C : x \in C_1\}$ in turn. After $O(w^2)$ steps, we have derived every clause in the set $\{y \to C : y \in C_2\}$. Resolving all these clauses with C_2 in turn, another $O(w)$ steps, gives us C. We replace every semantic consequence step in Π with a derivation of this form. ■

2.4 Random extension clauses

The paper [5] considered the problem of proving lower bounds on bounded depth Frege proofs with connectives defined using counting modulo a prime, and reduced it to a problem about proving lower bounds on Nullstellensatz proofs containing certain specific low degree *extension polynomials.* These polynomials use additional variables r_j and have the following property: for every fixed assignment α to the original variables x_i, the extension polynomials are zero with high probability, if we fix the new variables r_j randomly.

Unfortunately, this property alone cannot be used for proving lower bounds. In [7] the authors showed that for every unsatisfiable set of low degree polynomials, there exist some extension polynomials that have the above property but are such that using them, one can derive a contradiction with a low degree proof.

In contrast to this, resolution with random extension clauses is a nontrivial concept. We will show that it is essentially equivalent to RR distributions.

Definition 2.12 *A* resolution refutation with ε-random extension clauses *is a pair* (B, Π) *such that*

1. *B is a CNF containing additional variables $r_1, \dots, r_\ell$ not appearing in F*

2. *Π is a resolution refutation of $F \wedge B$*

3. for every $\alpha \in \{0,1\}^n$, $\Pr_\beta[B[\alpha,\beta]=1] \geq 1-\varepsilon$ where the probability is with respect to an assignment β to the variables $\bar{r}$ chosen uniformly from $\{0,1\}^\ell$.

Proposition 2.13 *If F has a resolution refutation with ε-random extension clauses of width w and size s, then it has an ε-RR distribution of width w and size s.*

In the other direction, if F has an ε-RR distribution of width w and size s, then it also has a resolution refutation with 3ε-random extension clauses of width $w+\log(n/\varepsilon^2)+O(1)$ and size $O(sn/\varepsilon^2)$.

Proof. The first part follows by substituting a random assignment to the variables $\bar{r}$ into B.

For the second part, let $\ell = \log(n/\varepsilon^2)+c$, where c is a constant which we will specify later. By Lemma 2.2, there is a 2ε-RR distribution of F with width w and size s and with a sample space containing $m = O(n/\varepsilon)$ pairs (B_i, Π_i). Let the probabilities of the samples be $p_1,\ldots,p_m$. We will approximate these numbers by multiples of $2^{-\ell}$. Choose $Q_1,\ldots,Q_m$ so that each Q_i is $2^\ell p_i$ rounded down, or up, to an integer in such a way that $\sum_i Q_i = 2^\ell$. Let $p_i' = Q_i/2^\ell$. Then for every subset $I \subseteq [m]$ we have $\left|\sum_{i\in I} p_i - \sum_{i\in I} p_i'\right| < m/2^\ell \leq \varepsilon$, where the last inequality holds if we take c sufficiently large.

Distribute the strings $\beta \in \{0,1\}^\ell$ among the formulas B_i so that the fraction of strings assigned to B_i is p_i'. More formally, take a mapping $\iota : \{0,1\}^\ell \to [m]$ such that the preimage of each i has size $p_i' 2^\ell$. For $\beta \in \{0,1\}^\ell$ let B_β be the CNF

$$\bigwedge\{r_1 = \beta_1 \wedge \cdots \wedge r_\ell = \beta_\ell \to C : \; C \text{ is a clause in } B_{\iota(\beta)}\}.$$

This is logically equivalent to $\bar{r} = \beta \to B_{\iota(\beta)}$. Let B be the CNF $\bigwedge_{\beta\in\{0,1\}^\ell} B_\beta$.

Clearly, any substitution β to the variables $\bar{r}$ in B produces the formula $B_{\iota(\beta)}$. Hence for every assignment α to $\bar{x}$ the probability that a random assignment β to $\bar{r}$ falsifies F is, by construction, at most 3ε. Hence B satisfies the third part of Definition 2.12.

For the second part of Definition 2.12, we must construct a small refutation of $F \wedge B$. First, for each β, we use the refutation $\Pi_{\iota(\beta)}$ to derive the clause $\bar{r} \neq \beta$ from B_β. We then apply Fact 2.1. The resulting refutation has size $O(s2^\ell) = O(sn/\varepsilon^2)$ and width $w+\ell = w+\log(n/\varepsilon^2)+O(1)$, as required. ■

3 Upper bounds

3.1 Random 3-CNFs

We will show that random 3-CNFs with sufficiently high density have small RR distributions, while as is well-known, they only have exponentially large resolution refutations [9]. We will first prove a lemma.

Lemma 3.1 *Let $F := C_1 \wedge \ldots \wedge C_m$ be a k-CNF formula such that for every assignment α the number of clauses that are satisfied by α is $\leq \delta m$ for some constant $0 < \delta < 1$. Then F has a δ-RR distribution of width k and size $2k$ which can be constructed in polynomial time.*

Proof. Define a distribution of auxiliary CNFs B_i as follows. Choose $i \in \{1, \ldots, m\}$ uniformly at random. The clause C_i is a disjunction $y_1 \vee \ldots \vee y_l$ of l literals, $l \leq k$; we set B_i simply to be the conjunction $\neg y_1 \wedge \ldots \wedge \neg y_l$. Then we can derive the empty clause from C_i and B_i in just l resolution steps, and for any fixed α, the probability that random B_i is satisfied by α is $\geq 1 - \delta$. ■

Proposition 3.2 *A random* 3*-CNF with* n *variables and* $64n$ *clauses has a* 1/2*-RR distribution of constant width and constant size with probability exponentially close to 1.*

Proof. Let $m = 64n$. Let $C_1, \ldots, C_m$ be randomly chosen 3-clauses. We claim that the probability that there exists an assignment α that satisfies more than $\frac{15}{16}m$ clauses is exponentially small. To prove this claim, let α be a fixed truth assignment. Let N_α denote the number of clauses that are not satisfied by α. The expectation of N_α is $m/8$. By the Chernoff bound,

$$\Pr[N_\alpha \leq \tfrac{m}{16}] \leq \mathrm{e}^{-(\frac{1}{2})^2 \cdot \frac{1}{2} \cdot \frac{m}{8}} = \mathrm{e}^{-\frac{m}{64}}.$$

Thus the probability that there is any α for which $N_\alpha \leq \frac{m}{16}$ is, by the union bound, at most $2^n \cdot \mathrm{e}^{-\frac{m}{64}} = \mathrm{e}^{(\ln 2 - 1)n}$, which is exponentially small.

By Lemma 3.1, it follows that a random 3-CNF with n variables and $64n$ clauses has a 15/16-RR distribution of width 3 and size 6, with probability $1 - o(1)$. By Lemma 2.3, such 3-CNFs also have 1/2-RR distributions of constant width and size. ■

3.2 The PCP Theorem and random resolution proofs

We now address the natural question of whether random resolution can be presented as a standard propositional proof system in the sense of Cook and Reckhow [10], or at least whether it can be polynomially simulated by such a system. Because we want to compare other systems with random resolution, we adapt the definition to refutation systems — this makes no difference to the result, since any proof system can be considered as a refutation system and vice versa. The essential property of Cook and Reckhow's definition is that one can test the correctness of refutations in polynomial time, that is, that the binary relation "Π is a refutation of F" is decidable in deterministic polynomial time. The other two properties, soundness and completeness, are satisfied by random resolution.

In order to state our question formally, we must say which object we choose to represent a refutation in random resolution, and what polynomial simulation means. We will consider 1/2-RR distributions in which all samples have the same weight. Such a distribution can be written down simply as a list of pairs (B_i, Π_i), and by Lemma 2.2 we do not lose anything important if we only consider 1/2-RR distributions in this form.

Polynomial simulation of refutation systems can be defined also in this case, where correctness may not be decidable in polynomial time, in essentially the same way as for standard refutation systems. Let R_1 and R_2 be binary relations defining two refutation systems in the general sense described above. Then f *is a polynomial simulation of* R_1 *by* R_2 if f is computable in polynomial time and satisfies the condition

$$R_1(\Pi, F) \to R_2(f(\Pi, F), F).$$

We will prove that random resolution cannot be polynomially simulated by a standard refutation system unless $\mathbf{P} = \mathbf{NP}$. This is an easy corollary of the PCP Theorem, an equivalent version of which is (see Theorem 11.9 in [2]):

PCP Theorem *There exists a polynomial time computable function* g *and a constant* $\delta < 1$ *such that for every CNF formula* F*,* $g(F)$ *is a 3-CNF formula such that*

1. *if* F *is satisfiable, then* $g(F)$ *is also satisfiable*

2. *if* F *is unsatisfiable, then every assignment satisfies at most a fraction* δ *of the clauses of* $g(F)$*.*

Proposition 3.3 *If* $\mathbf{P} \neq \mathbf{NP}$*, then* $1/2$*-RR cannot be polynomially simulated by any Cook-Reckhow refutation system.*

Proof. By Lemma 3.1, the PCP Theorem implies that if F is unsatisfiable, then $g(F)$ has a δ-RR distribution constructable in polynomial time. The error δ can be reduced to $1/2$ by Lemma 2.3 and, again, this can be done in polynomial time. Let h denote the polynomial time computable function h that from a given unsatisfiable CNF formula F produces a $1/2$-RR distribution that refutes $g(F)$.

Suppose that $1/2$-RR can be polynomially simulated by a refutation system given by a polynomial time binary relation R and let f be the simulation. Then we can test whether F is satisfiable by computing $R(f(h(F), g(F)), g(F))$. ■

In particular, $1/2$-RR is not itself a Cook-Reckhow refutation system if $\mathbf{P} \neq \mathbf{NP}$.

3.3 The retraction weak pigeonhole principle

We can also separate narrow random resolution from resolution using an explicit sequence of CNFs. The *retraction weak pigeonhole principle* (see [8, 13]), which we denote rWPHP_n, asserts that there is no pair of functions $f : [2n] \to [n]$ and $g : [n] \to [2n]$ such that $g(f(x)) = x$ for all $x < n$. In particular, if rWPHP fails for f and g, then f is an injection and g is a surjection, so both the injective and surjective forms of the usual weak pigeonhole principle fail.

We formalize this as a propositional contradiction, which talks about f and g via their bit-graphs. So suppose $n = 2^k$. Then rWPHP_n is a CNF with variables $(f(x))_i$ for $x < 2n$ and $i < k$, for the ith bit of the value of $f(x)$, and variables $(g(y))_i$ for $y < n$ and $i < k + 1$, for the ith bit of the value of $g(y)$. It consists of the clauses

$$f(x) = y \to (g(y))_i = x_i \quad \text{for } x < 2n,\ y < n \text{ and } i < k + 1$$

where x_i is the ith bit of x. This is logically equivalent to $f(x) = y \rightarrow g(y) = x$ for every x and y.

Proposition 3.4 *Every resolution refutation of* rWPHP_n *requires width at least* n *and exponential size.*

Proof. It is easy to show the width lower bound by a Prover-Adversary argument, in which the Adversary maintains a partial matching between $[2n]$ and $[n]$ of size up to n.

Let m be the total number of variables in rWPHP_n and let ℓ be its initial width. Then $m \leq O(n \log n)$ and $\ell \leq O(\log n)$. By a well-known result of Ben-Sasson and Wigderson [4], if rWPHP_n has a refutation of size s then it has one of width $\ell + \sqrt{m \log s}$. A straightforward calculation shows that our width lower bound implies an exponential size lower bound. ■

Proposition 3.5 *The formulas* rWPHP_n *have narrow* $1/2$*-RR distributions.*

Proof. Choose $x < 2n$ uniformly at random and let B_x be the $(k+1)$-CNF

$$\bigwedge_{y<n} g(y) \neq x.$$

In any total assignment this is true with probability at least $1/2$.

For the narrow refutation, first for each $y < n$ resolve the clause $g(y) \neq x$ from B_x with the clauses $f(x) \neq y \vee (g(y))_i = x_i$ for $i = 0, \dots, k$ from rWPHP_n. The result is the clause $f(x) \neq y$. Once we have all clauses $f(x) \neq y$ we can derive the empty clause with a narrow resolution derivation, using Fact 2.1. ■

3.4 RR refutations over the uniform distribution

As discussed in Section 2.2, the definition of (ε, Δ)-RR refutations is unreasonably strong if we are able to fix the distribution Δ. We show in this section that the distribution does not have to be unnatural for this to happen.

Let $\mathbf{U}_n$ be the uniform distribution on $\{0,1\}^n$. We show first that, for every constant k, every unsatisfiable k-CNF has a small size and width random resolution refutation with respect to $\mathbf{U}_n$. In fact we prove something slightly more general, that this is true even for k-CNFs which are satisfied with some small probability (that is, which have a small set of satisfying assignments). Note that it follows that the $(\varepsilon, \mathbf{U}_n)$-RR refutation system is not sound, because such formulas are not contradictions.

Proposition 3.6 *For every* $k \in \mathbb{N}$ *and* $\varepsilon > 0$*, there exist* $s \in \mathbb{N}$ *and* $\delta > 0$ *such that for every* k*-CNF formula* F *that is satisfied with probability* $\leq \delta$ *in* $\mathbf{U}_n$*, there exists an* $(\varepsilon, \mathbf{U}_n)$*-RR refutation of size* $\leq s$*.*

Proof. The proof is by induction on k. First suppose $k = 1$. We put $\delta = \varepsilon$. The 1-CNF F is just a conjunction of literals. If F has two complementary literals we can derive the empty clause in one step. So suppose that F has m literals, with no pair of complementary ones, and is satisfied with probability $\leq \delta$. Then $m \geq -\log_2 \delta$. Let the auxiliary CNF B be a single clause consisting of $\lceil -\log_2 \delta \rceil$ negated literals from F. Clearly, B has the desired properties.

Now suppose that the proposition is true for k and let ε be given. Let ℓ be the least integer such that $(1 - 2^{-(k+1)})^\ell \leq \varepsilon$ and let $r := (\ell - 1)(k + 1)$. Let $\delta > 0$ be the constant given by the inductive assumption for k and $\varepsilon 2^{-r}$. Let F be a $(k + 1)$-CNF that is satisfied with probability $\leq \delta 2^{-r}$. Now there are two cases.

First, suppose that F has ℓ disjoint clauses (meaning that no two clauses share a common variable, negated or not negated). Then let Γ be a conjunction of such a set of clauses. Then Γ is satisfied with probability exactly $(1 - 2^{-(k+1)})^\ell \leq \varepsilon$. Let B be $\neg\Gamma$ written as a CNF. Then B has at most $(k+1)^\ell$ clauses, each of size ℓ, and B is satisfied with probability $\geq 1-\varepsilon$. Since ℓ is a constant, we also have a constant size refutation of $\Gamma \wedge B$, and hence of $F \wedge B$, since Γ is a subset of the clauses of F.

Otherwise, there exists a set of variables X of size r such that every clause of F contains a variable from X. Consider any assignment σ to the variables X. Then F^σ is a k-CNF and by the assumption of the proposition, F^σ is satisfied with probability at most δ. By the inductive hypothesis there exists a formula B_σ such that $F^\sigma \wedge B_\sigma$ has a constant size refutation and B_σ is satisfied with probability $\geq 1 - \varepsilon 2^{-r}$. Hence $F \wedge \bigwedge_\sigma B_\sigma$ has a constant size refutation and $\bigwedge_\sigma B_\sigma$ is satisfied with probability at least $1 - \varepsilon$, where the conjunction is over all assignments σ to the variables X. ∎

One can prove a similar proposition for other parameters, either by applying Lemma 2.3 or by modifying the proof above. For example, for every constant k, every unsatisfiable k-CNF F has an $(\varepsilon, \mathbf{U}_n)$-RR refutation with polylogarithmic width and ε^{-1} quasipolynomial.

We next show that any CNF with a small resolution refutation has a narrow RR refutation with respect to $\mathbf{U}_n$. When proving size lower bounds, the first step often is to reduce the width of a refutation by applying a random restriction. The proposition shows that in the case of RR refutations there is another possibility, if the distribution is uniform or close to uniform: we can reduce the width by increasing the error. We do something like this in the proof of Theorem 4.3 below.

Proposition 3.7 *Suppose F has a resolution refutation of size s. Then for every $r \geq 1$, F also has a $(2^{-r}, \mathbf{U}_n)$-RR refutation of size $O(sw^2)$ and width $O(w)$, where $w = r + \lceil \log s \rceil$.*

Proof. Let Π be a refutation of F of size s. Say that a clause is *wide* if it has width at least w. Each wide clause is falsified by at most $2^{n-w} \leq 2^{n-r}/s$ total assignments. Hence if we let Z be the set of assignments which falsify any wide clause in Π, then $|Z| \leq 2^{n-r}$. Define $\mathcal{A} = \{0,1\}^n \setminus Z$. Then every assignment in $\mathcal{A}$ satisfies every wide clause in Π and $\Pr_{\alpha \sim \mathbf{U}_n}[\alpha \in \mathcal{A}] \geq 1 - 2^{-r}$. Now delete all wide clauses from Π (including possibly initial clauses from F) and let Σ be the resulting sequence of clauses. Then $(\mathcal{A}, \Sigma)$ is a $(2^{-r}, \mathbf{U}_n)$-semantic resolution refutation of F. Hence, by Proposition 2.11, there exists a $(2^{-r}, \mathbf{U}_n)$-RR refutation of F of size $O(sw^2)$ and width $O(w)$. ∎

It is clear that this proposition can be generalized in at least two ways. First, we may weaken the assumption that F has a small resolution refutation to the assumption that F has an ε-RR refutation, with some error ε that will be added to the parameter in the narrow proof. Second, we can use an arbitrary distribution Δ instead of $\mathbf{U}_n$, with a suitably modified concept of the width. Namely, we could define the width of a clause C with respect to Δ by

$$\mathrm{width}_\Delta(C) := \log(\Pr_{\alpha\sim\Delta}[C[\alpha] = 0])^{-1}.$$

4 Lower bounds for the bit pigeonhole principle

In this section and in Section 8 we prove size lower bounds on versions of the pigeonhole principle, by first using a random restriction to reduce to a width lower bound. In Section 5, on the coloured polynomial local search principle, we only prove a width lower bound, as CPLS already has small resolution refutations; however in Section 6 we use this width bound to prove a size lower bound on a related principle.

Before going into details of our lower bound proofs, we outline the basic structure that the proofs will follow. After that we present the simplest of our three main lower bounds on RR distributions.

4.1 Lower bound strategy

To prove width lower bounds on a 1/2-RR distribution, we use Proposition 2.8 to convert the distribution into a $(1/2, \mathcal{R})$-RR refutation (B, Π) with respect to a distribution $\mathcal{R}$ on partial assignments (we will use the terms "restriction" and "partial assignment" interchangeably). The crucial thing is to choose the distribution $\mathcal{R}$ carefully.

The ideal would be that there are many restrictions ρ from $\mathcal{R}$ which make the auxiliary formula B true, thus making it vanish and leaving us with a resolution refutation for which we already have a lower bound. To this end we use a sort of rudimentary version of the switching lemma, which we call a *fixing lemma* (a different lemma in each case, because it depends on the distribution $\mathcal{R}$). Intuitively this shows that, with reasonably high probability, ρ fixes the value of B to either true or false. From the definition of a $(1/2, \mathcal{R})$-RR refutation we know that $B^\rho = 0$ with probability at most a half, so we can conclude that many restrictions ρ make B true.

However, in practice it is not possible to achieve the ideal that ρ makes B true. Instead we only ask that the restricted formula B^ρ cannot be falsified by any "legal" extension $\sigma \supseteq \rho$. What counts as a legal extension depends on F — for example, for the pigeonhole principle it will be a partial assignment that represents a partial matching. The definition is chosen so so that we can both prove the fixing lemma and then prove a width lower bound on Π by an adversary argument, in which the adversary only works with legal extensions of ρ.

So a generic fixing lemma has the following form. We first define suitable notions of random and legal restrictions. We say that a CNF B is *fixed* by a restriction ρ if either ρ falsifies B, or no legal $\sigma \supseteq \rho$ falsifies B. Then the lemma states that every CNF is, with high probability, fixed by a random ρ.

The proof of such a lemma should, in principle, be a special case of a proof of a switching lemma, since we are essentially switching a CNF to a decision tree of height 0, or to a trivial DNF. However in the one case we consider in which a switching lemma is known, for the (non-bit) pigeonhole principle, we do not use it directly, but rather prove our own fixing lemma. One reason is that the usual lemma works with *syntactic* transformations of formulas and does not seem to guarantee that our *semantic* condition on B, that B is satisfied with high probability, is preserved. For the CPLS formula in the next section, there is unlikely to be any traditional switching lemma. This is because, understood very broadly, such a lemma would imply strong size lower bounds on CPLS in constant depth Frege, while we know that CPLS already has polynomial size refutations in resolution.

4.2 Lower bounds for BPHP_n

Let $n = 2^k$. As already described, BPHP_n is a contradictory CNF asserting that a function f is an injection from $[n+1]$ to $[n]$. It has variables $(f(x))_j$ for each $x < n+1$ and $j < k$, for the jth bit of the value of $f(x)$, and consists of clauses

$$f(x) \neq y \vee f(x') \neq y$$

for all $x < x' < n+1$ and all $y < n$.

In our proof, we will only consider partial assignments in which, for every x, either all or none of the variables $(f(x))_j$ are set. We identify such assignments with the corresponding partial functions from $n+1$ pigeons to n holes.

Given a probability p, define the distribution $\mathcal{R}_p$ of partial injections ρ from $[n+1]$ into $[n]$ as follows: choose the domain of ρ by putting each pigeon into the domain independently at random with probability $1-p$, then choose uniformly at random from all possible partial injections with this domain (if all $n+1$ pigeons get put into the domain, we just take ρ to be empty). For the rest of the proof, set $p = n^{-2/3}$ and $w = n^{1/4}$.

Lemma 4.1 (fixing lemma) *Say that a CNF B is* fixed *by a restriction ρ if either ρ falsifies B, or no partial injection $\sigma \supseteq \rho$ falsifies B. Then for sufficiently large n, for any w-CNF B,*

$$\Pr[B \textit{ is not fixed by } \rho] \leq n^{-1/13}.$$

Proof. Let S be the set of $\rho \in \mathcal{R}_p$ which do not fix B. Fix an ordering of the clauses of B. Consider any $\rho \in S$. No clause in B is falsified by ρ, but there must be at least one clause which is falsified in some partial injection $\sigma \supseteq \rho$. Let C be the first such clause and let σ be such an extension of ρ falsifying it. The literals in C appear in some fixed order. Let x be the first pigeon mentioned in C which is not in the domain of ρ, and let $i < w$ be the position in C at which the first variable from pigeon x appears. Let σ' be σ restricted to the pigeons in the domain of ρ together with pigeon x, that is, $\sigma' = \rho \cup \{\langle x, \sigma(x)\rangle\}$.

Define a function θ on S by $\theta : \rho \mapsto (\sigma', i)$, where σ' and i are chosen as above. Then θ is an injection, because we can first recover C from $\theta(\rho)$ as the first clause of B which is falsified in some extension of σ' to a partial injection; then we can recover x as the pigeon associated with the variable at position i in C; and finally we can recover ρ from σ' by unsetting pigeon x.

Let τ be any restriction which sets $m > 0$ pigeons. The the probability of τ, that is, the probability that a randomly sampled restriction is equal to τ, is

$$\Pr[\tau] = (1-p)^m p^{n+1-m} \frac{(n-m)!}{n!}.$$

Hence if ρ sets m pigeons then $\Pr[\sigma']/\Pr[\rho] = (1-p)/p(n-m)$. By the Chernoff bound, for a random ρ the number $n-m$ of unset pigeons is smaller than $2pn$ with exponentially high probability in n. Let S_{bad} be the set of restrictions for which this bound fails, so that $\Pr[\sigma']/\Pr[\rho] > (1-p)/2p^2 n > 1/4p^2 n$ for $\rho \in S \setminus S_{\mathrm{bad}}$. Partition $S \setminus S_{\mathrm{bad}}$ into subsets $S_0, \ldots, S_{w-1}$ according to the second component i of θ. On each S_i, the first component θ_1 of θ is an injection from $\mathcal{R}_p$ to $\mathcal{R}_p$ which increases probability by at least $1/4p^2 n$. Therefore

$$\Pr[\theta_1[S_i]] = \sum_{\rho \in S_i} \Pr[\theta_1(\rho)] > \frac{1}{4p^2 n} \sum_{\rho \in S_i} \Pr[\rho] = \frac{1}{4p^2 n} \Pr[S_i].$$

Since $\Pr[\theta_1[S_i]] \leq 1$ we can conclude that $\Pr[S_i] < 4p^2 n$, and hence that $\Pr[S \setminus S_{\mathrm{bad}}] < 4p^2 nw = 4n^{-1/12}$. Since $\Pr[S_{\mathrm{bad}}]$ is exponentially small, the result follows. ■

Theorem 4.2 BPHP_n *does not have* $1/2$*-RR distributions of width* $w = n^{1/4}$.

Proof. We will show that BPHP_n has no $(1/2, \mathcal{R}_p)$-RR refutation with this width. Suppose for a contradiction that there is such a refutation (B, Π), where B is the auxiliary w-CNF which is false in $\mathcal{R}_p$ with probability at most $1/2$.

We claim that $\Pr[\text{no partial injection } \sigma \supseteq \rho \text{ falsifies } B] \geq 1/3$. This is because, by Lemma 4.1, a random $\rho \in \mathcal{R}_p$ fixes B with probability close to 1. Together with the condition on B, this implies that with probability at least $1/3$, ρ fixes B but does not falsify B, giving the claim. Using the claim and the Chernoff bound we choose one such restriction ρ which also leaves at least $pn/2 = n^{1/3}/2$ holes free.

Now consider any clause C in the refutation Π. Suppose we have a partial injection $\sigma \supseteq \rho$ that falsifies C, and suppose that C is derived by resolution from clauses $D \vee v$ and $E \vee \neg v$, where v is a variable $(f(x))_j$ for some $x < n+1$ and $j < k$. Since $|C| \leq n^{1/4}$ we can find $\sigma' \subseteq \sigma$ that falsifies C and sets at most $n^{1/4}$ pigeons not set in ρ. Hence we can find a free hole to assign to pigeon x, thus extending σ' to a partial injection which falsifies either $D \vee v$ or $E \vee \neg v$.

In this way, working inductively up through the refutation, we can find a partial injection $\sigma \supseteq \rho$ which falsifies some initial clause. But this is a contradiction, since a partial injection cannot falsify any clause from BPHP_n, and by our choice of ρ a partial injection extending ρ cannot falsify any clause from B. ■

We show size lower bounds by combining the argument of Theorem 4.2 with a standard application of random restrictions to remove clauses mentioning many pigeons from Π.

Theorem 4.3 BPHP_n *has no* $1/2$*-RR distribution of subexponential size, that is, of size* $< 2^{n^\delta}$ *for some fixed* $\delta > 0$*.*

Proof. Suppose (B, Π) is a subexponential size $(1/2, \mathcal{R}_p)$-RR refutation of BPHP_n. Let $p = n^{-2/3}$ and $w = n^{1/4}$ as above.

Let C be any clause that mentions at least w pigeons, and choose literals $v_1, \ldots, v_w$ from C such that v_i comes from the ith pigeon mentioned by C. Consider ρ chosen at random from $\mathcal{R}_p$. Then for any i we have that ρ satisfies v_i with probability $(1-p)/2$. However these events are not completely independent for different i, since ρ is constrained to be a partial injection. We claim that the probability that ρ does not satisfy any v_i is bounded above by $(2/3)^w$. To see this, first note that we can view ρ as being chosen step-by-step, starting with the pigeons associated with the literals $v_1, \ldots, v_w$ in that order, choosing in turn whether each pigeon is put in the domain of ρ and, if so, which hole it goes to. When the ith pigeon is considered there are m holes available, for some m with $n - w < m \leq n$, and of these at least $n/2 - w$ will satisfy v_i. Hence the probability v_i is satisfied, given the choices already made for previous pigeons, is at least $(1-p)(n/2-w)/(n-w) > 1/3$ for sufficiently large n. Hence, for large n, C is satisfied with probability $\geq 1 - (2/3)^w$. By the union bound, with exponential high probability ρ satisfies every clause in Π that mentions w or more pigeons.

Now the arguments of Lemma 4.1 and Theorem 4.2 go through, with some tweaks. Firstly, they still work if we replace the *width* of a clause C with the *number of pigeons mentioned* in C. In particular Lemma 4.1 works if we only assume that each clause in B mentions no more than w pigeons – we just need to use the index i to record which of those w pigeons σ extends ρ by, rather than recording the position of the relevant literal. Secondly, in the proof of Lemma 4.1 we ignored a set S_{bad} of restrictions ρ with a certain undesirable property, and it was safe to do this because the probability of S_{bad} was exponentially small. We now add to S_{bad} the set of restrictions ρ which do not satisfy every clause that mentions w or more pigeons, and can then safely assume that every clause C that we consider in the proof mentions fewer than w pigeons, to give the required bound on the index i. ■

5 A separation of resolution from narrow RR

The *coloured polynomial local search* principle (CPLS) was introduced in [20]. The propositional version of it was studied in [24]. We refer to those two papers for more on the principle, and only remark here that it is a good candidate for proving separations of this kind because it is in some sense "complete" among narrow CNFs with short resolution refutations [20], while at the same time its combinatorial structure is simple enough that we are able to come up with useful random restrictions. We take our definitions from [24].

The principle has parameters a, b, c (for the lower bound we will eventually take $b = a$ and $c = a^{1/7}$). Consider a levelled directed graph whose nodes consist of all pairs (i, x) from $[a] \times [b]$. We refer to (i, x) as *node x on level i*. If $i < a - 1$, this node has a single neighbour in the graph, node $f_i(x)$ on level $i + 1$. Every node in the graph is coloured with some set of colours from $[c]$. CPLS expresses that the following three sentences cannot all be true at once.

1. Node 0 on level 0 has no colours.
2. For every node x on every level $i < a - 1$, if the neighbour $f_i(x)$ of x on level $i + 1$ has any colour y, then x also has colour y.
3. Every node x on the bottom level $a - 1$ has at least one colour, $u(x)$.

We will express this principle as a family of propositional contradictions. Let a be any natural number and let b and c be powers of two. We will define a CNF formula $\mathrm{CPLS}_{a,b,c}$, in the following propositional variables.

- For each $i < a$, $x < b$ and $y < c$, there is a variable $G_i(x, y)$, expressing whether colour y is present at node (i, x).
- For each $i < a$, $x < b$ and $j < \log b$, there is a variable $(f_i(x))_j$, standing for the jth bit of the value of $f_i(x)$.
- For each $x < b$ and $j < \log c$, there is a variable $(u(x))_j$, standing for the jth bit of the value of $u(x)$.

Definition 5.1 *The formula* CPLS *consists of the following three sets of clauses, which we will call Axioms 1, 2 and 3:*

Axiom 1. For each $y < c$, the clause

$$\neg G_0(0, y)$$

Axiom 2. For each $i < a - 1$, each pair $x, x' < b$ and each $y < c$, the clause

$$f_i(x) = x' \wedge G_{i+1}(x', y) \rightarrow G_i(x, y)$$

Axiom 3. For each $x < b$ and each $y < c$, the clause

$$u(x) = y \rightarrow G_{a-1}(x, y).$$

Note that we do not require f_i to be one-to-one. However, our lower bound proof will also work for the contradiction that additionally contains the clauses $f_i(x) \neq y \vee f_i(x') \neq y$ for all $x \neq x'$ and all y.

Proposition 5.2 $\mathrm{CPLS}_{a,b,c}$ *has polynomial size resolution refutations.*

Proof. For $i < a$, let M_i be the set of clauses $\{\bigvee_{y<c} G_i(x, y) : x < b\}$ expressing that every node at level i has a colour. We can derive M_{a-1} from Axiom 3, using Fact 2.1. Then repeatedly using Axiom 2 and Fact 2.1 we can derive M_{a-2}, M_{a-3}, etc. Once we have M_0 we can derive a contradiction from Axiom 1. For more detail see [24]. ■

We define a class of partial assignments which we will use in our lower bound argument. We first define the important notion of a *path* in a partial assignment.

Definition 5.3 *A* path *in a partial assignment β is a sequence of nodes $(i, x_0), \ldots, (i+k, x_k)$ of maximal length such that $f_{i+j}(x_j) = x_{j+1}$ in β for each $j \in [0, k)$.*

A path may consist of only one node. Thus every node is on some unique path (as long as all functions f_i are partial injections).

We now define three types of restrictions. *Legal restrictions* are a general class of partial assignments that do not falsify $\mathrm{CPLS}_{a,b,c}$ and have some additional nice properties. *Random restrictions* are essentially a probability distribution on legal restrictions, although with exponentially small probability a random restriction may fail to be legal (the zero path may reach the bottom). *Good restrictions* are in a sense typical random restrictions; formally, they are those for which certain parameters, such as the lengths of paths, are close to the mean value.

Definition 5.4 (legal restriction) *A* legal restriction *is a partial assignment β with the following properties.*

L1. At every node (i, x), either all variables $(f_i(x))_j$ are set, or none are. At every level i, the variables that are set define f_i as a partial injection.

L2. For every node (i, x) on the path π beginning at $(0, 0)$, we have $G_i(x, y) = 0$ for every colour y. Furthermore π does not reach all the way to the bottom level $a-1$. We call π the zero path*.*

L3. Every other path π is either starred *or* coloured*, where*

- *(a) if π is starred, then for every node (i, x) on π, no colour $G_i(x, y)$ is set*
- *(b) if π is coloured, then there is some single colour y such that for every node (i, x) on π, $G_i(x, y) = 1$ and $G_i(x, y') = 0$ for all colours $y' \neq y$.*

L4. For every node $(a-1, x)$ on the bottom level, let π be the path containing $(a-1, x)$. If π is starred then all variables $(u(x))_j$ are unset. If π is coloured then $u(x) = y$ where y is the unique colour such that $G_i(x, y) = 1$.

We state two obvious lemmas, without proof.

Lemma 5.5 *No legal restriction falsifies* $\mathrm{CPLS}_{a,b,c}$.

Lemma 5.6 *Let $\rho \subseteq \sigma$ be legal restrictions. Let z be a variable that is not set by ρ, but is set in σ. Then there exists a unique minimal legal extension $\rho \subset \sigma' \subseteq \sigma$ that sets z and extends ρ in one of the following two ways.*

1. *It changes some starred path into a coloured path.*
2. *It sets some value $f_i(x)$ that is not set in ρ. This necessarily means connecting two paths. Either at least one of these is starred, or they both have the same colouring. The resulting path inherits the colouring of one, or both paths. If they are both starred, then so is the resulting path.*

In either case if a node $(a-1, x)$ which was previously on a starred path is now on a coloured path, then σ' also sets $u(x) = y$ where y is the colour of the path.

We now define a distribution of random restrictions for which we will be able to prove a form of switching lemma.

Definition 5.7 (random restriction) *Fix parameters $0 < p, q < 1$. Let $\mathcal{R}_{p,q}$ be the distribution of random restrictions chosen as follows.*

R1. For each pair $i < a$ and $x < b$, with probability $(1-p)$ include (i, x) in a set Z. For each $i < a$, choose f_i uniformly at random from the partial injections from the domain $\{x < b : (i, x) \in Z\}$ into b.

R2. Colour the path beginning at $(0, 0)$ so that $G_i(x, y) = 0$ for all nodes (i, x) on that path.

R3. For every other path π, with probability $(1-q)$ colour *π randomly with one colour. That is, choose uniformly at random a colour y and, for every node (i, x) on π, set $G_i(x, y) = 1$ and set $G_i(x, y') = 0$ for all $y' \neq y$.*

R4. Finally consider each node $(a-1, x)$ on the bottom level. It is on some path π. If π was coloured at step R3, then set $u(x) = y$ where y is the unique colour assigned to π (that is, $G_{a-1}(x, y) = 1$). Otherwise leave $u(x)$ undefined.

Abusing notation, we will also use $\mathcal{R}_{p,q}$ to denote the set of restrictions which have nonzero probability in $\mathcal{R}_{p,q}$. Note that this contains all legal restrictions.

For the rest of this section we will fix parameters as follows.

$$a = b = n,\ c = \lfloor n^{1/7} \rfloor,\ p = n^{-4/7},\ q = n^{-2/7},\ w = \lfloor n^{1/8} \rfloor \qquad (1)$$

where b and c are powers of 2. We will use the well-known Chernoff bound several times to show that the probability of an event is *exponentially small* (or *exponentially high*), by which we mean that the probability is less than $\exp(-n^{\varepsilon})$ (or more than $1 - \exp(-n^{\varepsilon})$) for some constant $\varepsilon > 0$.

Given a restriction σ, we will say that a node $(i + 1, x)$ is *free* if it is not in the range of the partial function f_i defined by σ. We will say that a node is zero, starred or coloured if the path containing it is respectively zero, starred or coloured.

Lemma 5.8 (good restrictions) *We say that a restriction $\rho \in \mathcal{R}_{p,q}$ is* good *if all of the following hold.*

G1. No path in ρ is longer than $n^{5/7}$. (It follows that ρ is legal.)

G2. At all levels i, there are at most $2np = 2n^{3/7}$ nodes (i,x) for which $f_i(x)$ is undefined.

G3. At all levels $i \geq 1$, there are least $npq/2 = n^{1/7}/2$ free, starred nodes.

Otherwise ρ is bad. *Then the probability that $\rho \sim \mathcal{R}_{p,q}$ is bad is exponentially small.*

Proof. For item G1, for a fixed vertex (i,x) the probability that there is a path of length ℓ starting at (i,x) is at most $(1-p)^\ell$. Using the union bound we can bound the probability, for $\ell = n^{5/7}$, by

$$n^2(1-p)^\ell \approx n^2 \exp(-p\ell) = n^2 \exp(-n^{1/7}).$$

Items G2 and G3 follow immediately from the Chernoff bound. ■

In the following lemma and proof probabilities, expectations etc. are over $\rho \sim \mathcal{R}_{p,q}$.

Lemma 5.9 (fixing lemma) *Say that a CNF B is* fixed *by a restriction ρ if either ρ falsifies B, or no legal $\sigma \supseteq \rho$ falsifies B. Then for sufficiently large n, for any w-CNF B,*

$$\Pr[B \textit{ is not fixed by } \rho] \leq n^{-1/57}.$$

Proof. Let S denote the set of good restrictions $\rho \in \mathcal{R}_{p,q}$ which do not fix B. By Lemma 5.8, the probability that a random ρ is bad is exponentially small. So to bound the probability that B is not fixed it is enough to show that $\Pr[\rho \in S]$ is small. To estimate this probability, we will construct a mapping $\theta : S \to \mathcal{R}_{p,q}$. Fix an ordering of the clauses of B. For $\rho \in S$, we have that B is not falsified by ρ but is falsified by some legal extension of ρ. We use this to define $\theta(\rho)$ as follows.

Let C be the first clause of B that is falsified by any legal extension $\sigma \supseteq \rho$, and fix one such σ. Let z be the first variable of C that is not fixed by ρ — such a z must exist, because C is not falsified by ρ. Let σ' be the minimal legal extension $\rho \subset \sigma' \subseteq \sigma$ that fixes z, as given by Lemma 5.6. We put $\theta(\rho) := \sigma'$.

Let us compare the probabilities of ρ and $\theta(\rho)$. For $i = 0, \ldots, n-2$ let Z_i be the set of nodes (i,x) for which $f_i(x)$ is defined in ρ. Let $m_i = |Z_i|$ and $m = \sum_i m_i$. Let r be the number of coloured and s the number of starred paths in ρ. Then the probability of ρ is

$$\Pr[\rho] := (1-p)^m p^{n^2-m} \cdot \prod_{i=0}^{n-2} \frac{(n-m_i)!}{n!} \cdot \left(\frac{1-q}{c}\right)^r q^s$$

where the three parts of the product calculate respectively the probability of this choice for the domain of the functions f_i, this choice for the values of the f_i, and this choice of a way of colouring paths. According to Lemma 5.6, there are two possible ways in which σ' extends ρ.

1. Some starred path in ρ is coloured in σ'. Then going from ρ to σ' increases r by one, decreases s by one and leaves the other parameters the same. Hence

$$\frac{\Pr[\theta(\rho)]}{\Pr[\rho]} = \frac{1-q}{c} \cdot \frac{1}{q} \geq \frac{1}{2n^{1/7}} \cdot n^{2/7} = \frac{1}{2}n^{1/7}.$$

2. Some value $f_i(x)$ undefined in ρ is set in σ'. Then m_i and m increase by one. For the other parameters, there are two cases.

   (a) If this connects a starred path to the zero path or a coloured path, or if it connects two starred paths, then s decreases by one and r is unchanged. So

   $$\frac{\Pr[\theta(\rho)]}{\Pr[\rho]} = \frac{1-p}{p} \cdot \frac{1}{n-m_i} \cdot \frac{1}{q} \geq \frac{n^{4/7}}{2} \cdot \frac{1}{2n^{3/7}} \cdot n^{2/7} = \frac{1}{4}n^{3/7}$$

   where we have $n - m_i \leq 2n^{3/7}$ by Lemma 5.8, since ρ is good.

   (b) If it connects two coloured paths, then r decreases by one and s is unchanged. So

   $$\frac{\Pr[\theta(\rho)]}{\Pr[\rho]} = \frac{1-p}{p} \cdot \frac{1}{n-m_i} \cdot \frac{c}{1-q} \geq \frac{n^{4/7}}{2} \cdot \frac{1}{2n^{3/7}} \cdot n^{1/7} = \frac{1}{4}n^{2/7}.$$

The mapping θ is not one-to-one, but is at most $3w$-to-one. This is because we can recover ρ from $\theta(\rho)$ as follows. We find the clause C by taking the first clause in B which is falsified in some legal extension of $\theta(\rho)$. Then it suffices to know the position of the literal z in the clause C (a number less than w) and, if $\theta(\rho)$ was obtained by connecting two paths and the resulting path has a colour, to know whether this path inherited the colour from the first part, the second part, or from both parts.

Now partition S as $S_0, \ldots, S_{3w-1}$ where $S_i = \{\rho \in S : \rho \text{ is the } i\text{th preimage of } \theta(\rho)\}$. Then

$$\Pr[S_i] = \sum_{\rho \in S_i} \Pr[\rho] = \sum_{\rho \in S_i} \Pr[\theta(\rho)] \frac{\Pr[\rho]}{\Pr[\theta(\rho)]} \leq 2n^{-1/7} \sum_{\rho \in S_i} \Pr[\theta(\rho)] \leq 2n^{-1/7}$$

where we use that $\sum_{\rho \in S_i} \Pr[\theta(\rho)] \leq 1$, since θ is injective on S_i. It follows that $\Pr[S] \leq 6wn^{-1/7} \leq 6n^{1/8-1/7} = 6n^{-1/56}$, giving the required bound. ■

We can now prove our main result.

Theorem 5.10 *For all sufficiently large n, the formula* $\mathrm{CPLS}_{a,b,c}$ *with our parameters $a = b = n$ and $c = \lfloor n^{1/7} \rfloor$ does not have a 1/2-RR distribution of width $n^{1/8}$.*

Proof. Suppose the formula has a 1/2-RR distribution of width w. By Proposition 2.8 it also has $(1/2, \mathcal{R}_{p,q})$-RR refutation (B, Π) of width w. We will show that this implies $w \geq n^{1/8}$.

By the definition of a RR-refutation over partial assignments, $\Pr_{\rho \sim \mathcal{R}_{p,q}}[B^\rho = 0] \leq 1/2$. By Lemmas 5.8 and 5.9, with probability close to 1 a random ρ is good and fixes B. Therefore we can find a ρ which is good, fixes B and does not falsify B, from which it follows that no legal extension of ρ falsifies B. Fix such a ρ.

We will prove the width lower bound using the well-known Prover-Adversary game. By replacing all clauses with their negations and reversing the direction of the arrows, we can view Π as a strategy for the Prover in the following game: at each turn the Prover either asks the Adversary for the value of a variable, or forgets a variable from memory to free the space for re-use; he wins as soon as the assignment in his memory falsifies either a clause of CPLS or a clause of B (as these were the initial clauses of Π). We will show that to have a winning strategy the Prover must be able to remember at least $n^{1/8}$ variables simultaneously. The width lower bound follows immediately.

So suppose the Prover is limited to remembering at most w variables. The Adversary's strategy is to always have in mind a legal extension σ of ρ which satisfies the conjunction currently known by the Prover, and is small in a certain sense.

Define the *ρ-size* of a legal extension $\sigma \supseteq \rho$ as follows. Let m be the number of nodes (i, x) for which the edge $f_i(x)$ is defined in σ but not in ρ. Let σ^- be the smallest legal extension of ρ which contains all these edges; we could alternatively define σ^- by adding these edges to ρ, suitably colouring any starred paths that are now connected to coloured paths or the zero path, and extending u appropriately. Now σ and σ^- have the same paths. We let r be the number of paths starred in σ^- but coloured in σ. The ρ-size of σ is then $m+r$. The following claim generalises Lemma 5.6.

Claim *If D is a conjunction of size $\ell < c$ and a legal extension $\sigma \supseteq \rho$ satisfies D, then D is also satisfied by a legal extension $\sigma' \supseteq \rho$ with ρ-size at most ℓ.*

Proof of claim. Suppose that D mentions m variables of the form $(f_i(x))_j$ and r variables of the form $G_i(x, y)$ or $(u(x))_j$. We first extend ρ to σ^- by setting every edge $f_i(x)$ mentioned in D the same way it is set in σ and colouring as necessary to make this a legal restriction. This deals with the $(f_i(x))_j$ variables, and every path in σ^- is now a section of a path in σ.

We now extend σ^- to σ' by colouring some paths. For each remaining variable $G_i(x, y)$ or $u(x)$ not already set in σ^-, consider the path on which the corresponding node lies in σ. It cannot be a starred path, or the variable would not be set. If it is coloured, we colour the path it lies on in σ^- the same way as it is coloured in σ. All that remains is a set of variables of the form $G_i(x, y)$ where (i, x) lies on the zero path in σ but not on the (shorter) zero path in σ^-. Since σ satisfies D, these variables must appear negatively in D. Since $\ell < c$, there must be some colour y that is not mentioned in any of these variables. Hence we can colour the paths on which they lie with colour y, and this will satisfy D. By construction the ρ-size of σ' is at most $m + r = \ell$. This completes the proof of the claim. ■

Now suppose the Prover's current memory consists of a conjunction D of size $\ell \leq w = n^{1/8}$, and the Adversary knows a legal extension $\sigma \supseteq \rho$ which satisfies D. By the claim, we may assume that the ρ-size of σ is at most ℓ. There are now three cases in the Adversary's strategy, depending on what the Prover does.

Suppose the Prover forgets a variable. Then the Adversary can apply the claim to shrink the legal extension to one of ρ-size at most $\ell - 1$.

Suppose the Prover queries a variable $(f_i(x))_j$. If this is already set in σ, reply with its value. Otherwise choose a free node $(i+1, x')$ on the next level down. This must exist by item G3 of the definition of a good restriction and the fact that the ρ-size of σ is less than $n^{1/7}/2$. Extend σ to σ' by setting $f_i(x) = x'$ and extending colourings appropriately. By item G1 of the definition of a good restriction and the bound on ρ-size, the zero path in σ' cannot reach all the way to the bottom level, so σ' is legal. Reply with the value of the variable in σ'.

Suppose the variable queries a variable $G_i(x, y)$ or $(u(x))_j$. If this is already set in σ, reply with its value. Otherwise, extend σ to σ' by colouring the corresponding path with some arbitrary colour y. Reply with the value of the variable in σ'.

Finally we observe that the Prover cannot win against this strategy, since no legal extension of ρ falsifies any clause from $\mathrm{CPLS}_{a,b,c}$ or from B. ■

6 A separation of constant-depth Frege from RR

We exhibit a narrow CNF which requires exponential size 1/2-RR distributions but which, unlike the pigeonhole principle, has polynomial size refutations in constant depth Frege, in fact in Res(2). Here Res(2), introduced in [16], is an extension of resolution in which clauses may contain conjunctions of pairs of literals.

The formula is CPLS^2, a variant of CPLS. For each i, x, y, instead of the single variable $G_i(x, y)$ it has two variables $G_i^0(x, y)$ and $G_i^1(x, y)$. To express that colour y is present at node (i, x) we now use the conjunction $G_i^0(x, y) \wedge G_i^1(x, y)$.

Definition 6.1 *The formula* $\mathrm{CPLS}^2_{a,b,c}$ *consists of the following three sets of clauses:*

Axiom 1'. For each $y < c$, the clause

$$\neg G_0^0(0, y) \vee \neg G_0^1(0, y)$$

Axiom 2'. For each $i < a - 1$, each pair $x, x' < b$ and each $y < c$, the two clauses

$$f_i(x) = x' \wedge G_{i+1}^0(x', y) \wedge G_{i+1}^1(x', y) \to G_i^0(x, y)$$

$$f_i(x) = x' \wedge G_{i+1}^0(x', y) \wedge G_{i+1}^1(x', y) \to G_i^1(x, y)$$

Axiom 3'. For each $x < b$ and each $y < c$, the two clauses

$$u(x) = y \to G_{a-1}^0(x, y)$$

$$u(x) = y \to G_{a-1}^1(x, y).$$

As before, these respectively express that node $(0,0)$ has no colours; that every colour present at node $(i+1, f_i(x))$ is also present at node (i,x); and that colour $u(x)$ is present at node $(a-1,x)$.

Proposition 6.2 $\mathrm{CPLS}^2_{a,b,c}$ *has polynomial size Res(2) refutations.*

Proof. Let F be the formula (a conjunction of 2-DNFs) obtained by taking $\mathrm{CPLS}_{a,b,c}$ and replacing every occurrence of the literal $G_i(x,y)$ with $G_i^0(x,y) \wedge G_i^1(x,y)$ and every occurrence of the literal $\neg G_i(x,y)$ with $\neg G_i^0(x,y) \vee \neg G_i^1(x,y)$. Then by making the same substitution into the resolution refutation of $\mathrm{CPLS}_{a,b,c}$ given by Proposition 5.2, we get a small Res(2) refutation of F. On the other hand, F is easily derivable from $\mathrm{CPLS}^2_{a,b,c}$ in Res(2). ∎

The idea of the lower bound proof is first to hit a small RR distribution with a random restriction which reduces its width and transforms CPLS^2 into CPLS, and then to reuse our width lower bound argument for CPLS. In fact, our initial random restriction will only give small width as measured with respect to positive literals $G_i(x,y)$. This is not an insignificant measure of width; recall that in the small resolution refutation of CPLS in Proposition 5.2, the wide clauses are of the form $\bigvee_y G_i(x,y)$. We will then show that there is no subexponential size 1/2-RR distribution for CPLS if all clauses may only contain a small number of positive literals $G_i(x,y)$, but any number of other literals.

Definition 6.3 *The* positive G-width *of a clause C is the number of positive literals of the form $G_i(x,y)$ which appear in C. As usual, the* positive G-width *of a refutation Π is the maximum of this value over all clauses in Π and the* positive G-width *of a distribution $(B_j,\Pi_j)_{j\sim\mathcal{D}}$ is the maximum over all refutations Π_j.*

We fix the same parameters (1) as in the previous section, that is $a = b = n$, $c = \lfloor n^{1/7} \rfloor$, $p = n^{-4/7}$, $q = n^{-2/7}$ and $w = \lfloor n^{1/8} \rfloor$, where b and c are powers of 2.

Lemma 6.4 *If* CPLS^2 *has a 1/2-RR distribution of size s, then the original formula* CPLS *has a 1/2-RR distribution of size $\leq O(s^2)$ with positive G-width $\leq O(\log s)$.*

Proof. Suppose that there exists a 1/2-RR distribution of size s for CPLS^2. By Lemmas 2.2 and 2.3 it follows that there is a 1/2-RR distribution $\Pi' = (B_j, \Pi'_j)_{j \in \mathcal{D}}$ for CPLS^2 of size $s' = O(s^2)$ in which the sample size $|\mathcal{D}|$ is of the order of the number of variables. In particular the total number of clauses appearing in Π' is polynomial in s.

Choose a random substitution θ as follows. Independently for each triple (i,x,y), choose $b \in \{0,1\}$ at random, then replace each occurrence of the variable $G_i^b(x,y)$ with the constant 1 and replace each occurrence of the variable $G_i^{1-b}(x,y)$ with the variable $G_i(x,y)$. We claim that $\Pi := (\theta(B_j), \theta(\Pi'_j))_{j\in\mathcal{D}}$ is a 1/2-RR distribution for CPLS. To see this, first observe that every clause of CPLS^2 under θ is either satisfied or simplifies into a clause of CPLS, hence $\theta(\Pi'_j)$ is a resolution refutation of CPLS $\wedge\, \theta(B_j)$ for each j. Now, for any assignment α to the variables of CPLS, define an assignment α' to the variables of CPLS^2 as

$$v[\alpha'] = \begin{cases} 1 & \text{if } v \text{ has the form } G_i^b(x,y) \\ G_i(x,y)[\alpha] & \text{if } v \text{ has the form } G_i^{1-b}(x,y) \\ v[\alpha] & \text{otherwise} \end{cases}$$

where $b \in \{0,1\}$ is the value chosen by θ for the triple (i,x,y). Then $\theta(B_j)[\alpha] = B_j[\alpha']$, so since $B_j[\alpha'] = 1$ with probability at least $1/2$ for $j \sim \mathcal{D}$, the same is true for $\theta(B_j)[\alpha]$.

Now let m be the total number of clauses in Π'. If a clause C has positive G-width larger than $\log m$ then either there is some triple (i,x,y) for which C contains both literals $G_i^0(x,y)$ and $G_i^1(x,y)$, or there are more than $\log m$ many triples (i,x,y) for which C contains at least one of the literals $G_i^0(x,y)$ or $G_i^1(x,y)$. In the first case $\theta(C)$ is always satisfied, and in the second case it is satisfied with probability $> 1-(1/2)^{\log m} = 1-1/m$. So by the union bound we may assume that with nonzero probability all such clauses in Π' are satisfied and can be removed, so that we can find a Π with positive G-width $O(\log s)$ as required. ■

Theorem 6.5 *For all sufficiently large n, the formula* CPLS$^2_{a,b,c}$ *with the parameters (1) above does not have a $1/2$-RR distribution of size $\leq 2^{n^{1/17}}$.*

Proof. Suppose that CPLS$^2_{a,b,c}$ has a 1/2-RR distribution of size $\leq 2^{n^{1/17}}$. If n is sufficiently large, then by Lemma 6.4, CPLS$_{a,b,c}$ has a 1/2-RR distribution with size $O(2^{2n^{1/17}})$ and with positive G-width $O(n^{1/17})$.

Let $\mathcal{R}'_{p,q}$ be the random restriction from Definition 5.7 in the previous section, with two changes. Firstly, we add a new step at the beginning of the construction of a restriction:

R0. For each colour $y < c$, with probability $1/2$ set $G_i(x,y) = 0$ for every node (i,x).

We call these colours *forbidden.* Secondly, when we choose random colours for paths at step R3 of Definition 5.7, we choose only from the colours which are not forbidden.

By an averaging argument (or one direction of Proposition 2.8) it follows that CPLS$_{a,b,c}$ has a $(1/2, \mathcal{R}'_{p,q})$-RR refutation (B,Π) which preserves the bounds on size and positive G-width. We would now like to repeat the proof of Theorem 5.10. However, that proof required a bound of $n^{1/8}$ on the width of all clauses in Π. This no longer holds, but we will show that with high probability all clauses wider than $n^{1/8}$ are satisfied by a random ρ.

Formally, in Lemma 5.8 we defined the notion of a *good* restriction ρ as one for which three well-behavedness conditions on paths and colourings hold, and showed that a random restriction is good with exponentially high probability. We now extend the notion by adding two extra conditions:

G4. No more than $\frac{2}{3}c$ colours are forbidden by ρ.

G5. Every clause in Π of width greater than $n^{1/8}$ is satisfied by ρ.

We will show that a random restriction satisfies these conditions with exponentially high probability (this is immediate for G4). This is then enough for the proof of Theorem 5.10 to go through; in fact, it would be enough to prove that the extra conditions are satisfied with probability $1-\varepsilon$ for a sufficiently small ε. Condition G5 is used in the two places where the proof of Theorem 5.10 uses the bound on width. Firstly, in the fixing lemma, we need the width bound for clauses which the restriction ρ does not satisfy; but we only need this for restrictions in $\rho \in S$, which only contains good restrictions, so G5 is enough. Secondly, we use the width bound to limit the Prover's memory in the Prover-Adversary argument; but at this point we have fixed one good restriction so may assume by G5 that all wide clauses have been satisfied and removed from the refutation. Condition G4 is used when we need to keep track of the number of available colours. It changes some calculations by a factor of 3, which does not affect anything important.

It remains to prove that G5 is satisfied with high probability. Let C be any clause of width $w \geq n^{1/8}$. We will show that the probability that C is not satisfied by ρ is exponentially smaller than $2^{-2n^{1/17}}$. By the assumption on the width of C, one of the following four things must be true.

1. C contains the literal $G_i(x,y)$ for at least $w/4$ distinct triples (i,x,y).

2. C contains the literal $\neg G_i(x,y)$ for at least $w/4$ distinct triples (i,x,y).

3. C contains a variable from $f_i(x)$ for at least $w/(4\log b)$ many distinct pairs (i,x).

4. C contains a variable from $u(x)$ for at least $w/(4\log c)$ many distinct values x.

For a sufficiently large n, Case 1 cannot happen, because we are given that the refutation has positive G-width $O(n^{1/17})$. We postpone Case 2 to the end of the proof.

For Case 3, recall that for each pair (i,x) the variables $(f_i(x))_j$ are all given values with probability $(1-p)$, and those values are unbiased coin tosses that are almost independent: they are only subject to the constraint that f_i is a partial injection. Hence it is not hard to show, as in the proof of Theorem 4.3, that in this case C is satisfied with probability at least $1-(2/3)^{w/(4\log b)} \geq 1-2^{-n^{1/9}}$ for large n.

In Case 4, the values of the variables $(u(x))_j$, as chosen by steps R3 and R4 of Definition 5.7, are again almost independent unbiased coin tosses; however this time they are biased by the fact that the value y of $u(x)$ must not be a forbidden colour. By the union bound over the $\log c$ values of j, the following is true with exponentially high probability, where we write y_j for the jth binary bit of y:

> Let Y be the set of colours which are not forbidden. Then for every $j < \log c$ the fraction of colours $y \in Y$ with $y_j = 0$ is in the interval $[\frac{1}{3}, \frac{2}{3}]$. (This implies the same condition for $y_j = 1$.)

Thus again C is satisfied with probability at least $1-\delta^{n^{1/8}}$ for some constant $\delta < 1$.

Finally for Case 2, we exploit the fact that ρ sets almost all variables $G_i(x,y)$ to 0. We consider two subcases. Let $m = \sqrt{w}/2 \geq n^{1/16}/2$.

2a. C contains a literal of the form $\neg G_i(x, y)$ for at least m distinct colours y

2b. C contains a literal of the form $\neg G_i(x, y)$ for at least m distinct pairs (i, x).

In Case 2a, since every colour is forbidden independently in ρ with probability 1/2, we have that C is satisfied with probability at least $1 - (1/2)^{n^{1/16}/2}$.

For Case 2b, we first prove a claim.

Claim *There exists $\gamma < 1$ such that given a set Z of $m \leq n/2$ nodes (i, x), the number of paths in ρ touched by any node in Z is at least $m/3$ with probability $\geq 1 - \gamma^m$.*

Proof of claim. Let Z be given. We will lower bound the probability by the probability of a simpler event: that in an $n \times n$ table, if we randomly choose n node-disjoint paths running all the way from the top row to the bottom row of the table, at least $m/3$ paths are touched by nodes in Z. Notice then that, without changing the probability, we can replace this event by one in which the paths are fixed to be the n columns, and the nodes Z are chosen randomly by picking a fixed number of nodes from each row.

So, suppose that the elements of this random set Z are chosen one by one. Suppose that we already have chosen i elements and we want to choose the $(i+1)$st element in a row j in which we already chose $k \leq i$ elements. Clearly, the number of the remaining nodes in row j that are in empty columns is at least $n/2$, because we assume $m \leq n/2$. Hence we will hit a new column with probability at least 1/2.

Consecutive events are dependent, because the probability depends on the current configuration. Therefore we compare this process with another one. Consider a counter that is initially set to zero. At each step we increase the value of the counter by 1 with probability 1/2, and keep the value the same with probability 1/2. One can show by induction that the probability that the value in the counter is k after i steps is a lower bound on the probability that we have hit exactly k columns after i steps in the previous process. The value of the counter is, however, equivalent to the sum of independent $0 - 1$ random variables, and we can apply the Chernoff bound to it. This proves the claim. ■

Once we have the claim, we use the fact that given a path π in ρ and a colour y, the variable $G_i(x, y)$ is set to 0 in ρ at every node (i, x) on π with probability close to 1, and in particular greater than 1/2. This is independent for distinct paths, and hence C is satisfied with probability at least $1 - 2^{-m/3}$. Since $2^{-m/3} \leq 2^{-n^{1/16}/6}$ this proves that C is satisfied with the required probability.

Thus we have shown that condition G5 is satisfied by a random ρ with exponentially high probability. This reduces the proof of this theorem to the proof of Theorem 5.10. ■

7 Feasible interpolation

In this section we prove a form of feasible interpolation for RR distributions in which the resolution refutations are treelike. Feasible interpolation is one of the two main tools for proving lower bounds in propositional proof complexity (the other being random restrictions).

Therefore it is important to fully understand the limitations of this method. It seems that the standard form of feasible interpolation does not hold for random resolution, but as we will show, one can prove lower bounds using randomized communication protocols and obtain lower bounds at least for treelike proofs. The idea of the proof is not new; similar arguments have been used in other papers for different proof systems (see [12] for one of the first applications of this argument). Furthermore, in [17] Krajíček proved a lower bound on a stronger type of communication protocol, which enabled him to prove a lower bound for general (that is, dag-like) random resolution refutations. However his proofs seem only to work for a small error, so small that it cannot be amplified to get a nontrivial bound for constant $\varepsilon > 0$, or only for small auxiliary formulas B_i. In a follow-up article [18] he reduced the problem of proving lower bounds for dag-like communication protocols to *"circuits with local oracles"*, but the problem of proving sufficiently strong lower bounds for these computational devices remains open.

Theorem 7.1 *Let $\bar{x}, \bar{y}, \bar{z}, \bar{u}$ be disjoint tuples of variables with $\bar{x} = x_1 \dots x_n$ and $\bar{y} = y_1 \dots y_n$. Let $F(\bar{x}, \bar{z})$ and $G(\bar{y}, \bar{u})$ be CNF formulas in the variables shown. Suppose that*

$$F \wedge G \wedge \bigwedge_{j=1}^{n} (\neg x_j \vee \neg y_j)$$

is unsatisfiable and has an ε-RR distribution in which the refutations Π_i are treelike and have size at most s. Then there exists a randomized communication protocol P for two players with the following properties. When the players are given assignments α and β in $\{0,1\}^n$ such that $\exists \bar{z}.F[\alpha, \bar{z}] = 1$ and $\exists \bar{u}.G[\beta, \bar{u}] = 1$, then

1. *with probability $\geq 1 - \varepsilon$ the players find some j such that $\alpha_j = \beta_j = 1$*
2. *they use $O(\log s)$ communication bits.*

Proof. Given α and β such that $\exists \bar{z}.F[\alpha, \bar{z}] = 1$ and $\exists \bar{u}.G[\beta, \bar{u}] = 1$, each player picks some γ (respectively δ) such that $F[\alpha, \gamma] = 1$ ($G[\beta, \delta] = 1$). Then jointly they randomly pick some (B_i, Π_i) from the RR distribution. If the auxiliary formula B_i is not satisfied by the assignment $\alpha, \beta, \gamma, \delta$ then the protocol may fail, but this happens with probability at most ε. Otherwise, the following protocol succeeds in finding the bit j. The players pick a clause C in Π_i such that the subtree above C has size between $\frac{1}{3}s$ and $\frac{2}{3}s$. They exchange bits in order to find out whether C is falsified by $\alpha, \beta, \gamma, \delta$. If so, they continue with the subtree above C. Otherwise, they delete all clauses above C from Π_i and continue with the modified tree. Since the resulting stumps are satisfied, as are all clauses of F and G, the players eventually reach a clause $\neg x_j \vee \neg y_j$ that is falsified by α, β. Since the size of the tree decreases by a factor of at least $1/3$ at each step, the number of bits they use is $O(\log s)$. ■

We will show an application of Theorem 7.1 that gives an exponential size lower bound on treelike RR distributions. Note, however, that such a lower bound (for a different CNF, of nonconstant width) was already been proved in [6].

Recall that the disjointness function is defined for two n-bit strings by $D_n(x, y) = 1$ if and only if $x_j \wedge y_j = 0$ for all j. The probabilistic communication complexity of this function is $\Omega(n)$ [14]. Reduction of this function to the Karchmer-Wigderson games are known for several partial Boolean functions. For example, Raz and Wigderson [21] showed a reduction to the following problem:

> Let V be a set of $3m$ vertices. Player I is given a partial matching P (a set of independent edges) on V of size m. Player II is given a clique Q on V of size $2m + 1$. The goal is to find an edge $e \in P \cap Q$.

They prove that every probabilistic protocol for this game with error at most 1/3 needs $\Omega(m)$ communication bits.

This problem defines a partial Boolean function where the variables stand for the edges, the minterms are partial matchings of size m, and the maxterms are cliques of size $2m + 1$. The fact that every maxterm intersects every minterm can be stated as a tautology in the usual way. The negation of this tautology is an unsatisfiable CNF formula of the form used in the theorem above. The variables $\bar{x}$ and $\bar{y}$ stand for the edges, the variables $\bar{z}$ represent a one-to-one mapping from a set of size m into the set of edges, and the variables $\bar{u}$ represent a one-to-one mapping from an $2m+1$-element set into the set of vertices. Such a formalization gives a formula with clauses of nonconstant size, but using extension variables we can modify it to a 3-CNF. Thus we get:

Corollary 7.2 *There exists a sequence of unsatisfiable 3-CNF formulas of polynomial size such that all treelike $\frac{1}{3}$-RR distributions have size $2^{\Omega(\sqrt{n})}$, where n is the number of variables.*

8 Lower bounds for the pigeonhole principle

We now consider the usual formalization of the pigeonhole principle, rather than the bit-graph version. It will be convenient to distinguish pigeons from holes, so let U and V be disjoint sets of vertices with $|U| = n+1$ and $|V| = n$. The CNF formula PHP_n has variables p_{ij} for $i \in U$ and $j \in V$ and consists of clauses

1. $\bigvee_{j \in V} p_{ij}$ for all $i \in U$
2. $\neg p_{ij} \vee \neg p_{i'j}$ for all $i, i' \in U$ with $i \neq i'$ and all $j \in V$.

Theorem 8.1 PHP_n *has no 1/2-RR distribution of size less than* $2^{\Omega(n^{1/12})}$.

Before proving the theorem we need a couple of auxiliary results. We start by observing that any clause C in Π that contains a pair of literals $\neg p_{ij}, \neg p_{i'j}$, for some $i \neq i'$, can be replaced by the clause $\neg p_{ij} \vee \neg p_{i'j}$ of PHP_n that is at least as strong as C. This may require some simple modification of the proof, but in any case the resulting proof can only be smaller. Therefore, we can assume without loss of generality that every clause C in Π contains at most one negative literal $\neg p_{ij}$ for every $j \in V$, except for clauses of the form $\neg p_{ij} \vee \neg p_{i'j}$.

We will use random restrictions twice. First we apply a random restriction to reduce the width of B and then we apply it again to prove a fixing lemma. In this proof we will use a less standard concept of width, introduced by Ajtai in [1]. For a clause C, we define $w_{\mathrm{ec}}(C)$, the *edge covering width* or *ec-width* of C, to be the smallest size of a set $W \subseteq U \cup V$ that covers all edges mentioned in C. Formally,

$$w_{\mathrm{ec}}(C) := \min\{|W| \mid \forall i \in U, j \in V,\ (p_{ij} \in C \vee \neg p_{ij} \in C) \to (i \in W \vee j \in W)\}.$$

If Φ is a CNF formula, then $w_{\mathrm{ec}}(\Phi)$ is the maximum of the ec-widths of its clauses.

We will denote by $\mathcal{R}_m$ the set of partial matchings of size $n-m$ equipped with the uniform distribution. We will often identify $\rho \in \mathcal{R}_m$ with the following partial assignment:

1. set $p_{ij} = 1$ if $(i, j) \in \rho$,

2. set $p_{ij} = 0$ if $(i, j) \notin \rho$ and either i or j is in the domain of ρ,

3. the value of p_{ij} is undefined otherwise.

The set of vertices covered by the matching ρ will be called the *support of* ρ and denoted by $\mathrm{supp}(\rho)$.

Lemma 8.2 *There exist constants $c > 0$ and $0 < d < 1$ such that for every clause C and every $1 \leq \ell \leq n^{1/2}$,*

$$\Pr[w_{\mathrm{ec}}(C^\rho) > \ell] \leq d^\ell,$$

where the probability is over $\rho \sim \mathcal{R}_{\lfloor cn^{1/4} \rfloor}$.

Proof. We will use the following elementary estimate. Let $X \subseteq [n]$, $|X| = x$ and $y \in \mathbb{N}$ be fixed, and choose $Y \subseteq [n]$ with $|Y| = y$ at random. Then

$$\Pr[|X \cap Y| \geq \ell] \leq \left(\frac{\mathrm{e}xy}{n\ell}\right)^\ell. \tag{2}$$

Given a clause C, let $E^+ \subseteq U \times V$ be the bipartite graph determined by the positive literals of C. Let $E^- \subseteq U \times V$ be determined by the negative literals.

We will first prove that the width of the subclause of positive literals is small with high probability. Let

$$A := \{i \in U \ : \ \deg_{E^+}(i) \geq 2n^{1/2}\ell\}.$$

We consider two cases.

Case 1. Suppose $|A| \geq 2n^{1/2}$. We will show that in this case C is satisfied by ρ, and hence $w_{\mathrm{ec}}(C^\rho) = 0$, with high probability. Think of ρ being constructed by first selecting its domain $D \subseteq U$ and then gradually defining $\rho(i)$ for $i \in D$, where we begin with pigeons $i \in A \cap D$. Note that $|A \cap D| \geq 2n^{1/2} - cn^{1/4} \geq n^{1/2}$ provided that n is large enough.

Suppose $\rho(i)$ has been fixed for the first $k < n^{1/2}$ elements $i \in A \cap D$. We will estimate the probability that $(i, \rho(i)) \in E^+$ for the next element $i \in A \cap D$. Note that there are at least $2n^{1/2}\ell - k \geq n^{1/2}\ell$ neighbours of i in the graph E^+ that are not in the range of ρ as constructed so far. So the probability is at least $n^{1/2}\ell/n$. Hence the probability that C is *not* satisfied after the value of $\rho(j)$ is decided for the first $\lceil n^{1/2}\rceil$ elements $j \in A \cap D$ is

$$\leq \left(1 - \frac{n^{1/2}\ell}{n}\right)^{\lceil n^{1/2}\rceil} = (1 - o(1))\mathrm{e}^{-\ell}.$$

Case 2. Suppose $|A| < 2n^{1/2}$. Let $D \subseteq U$ and $R \subseteq V$ denote respectively the domain and range of ρ. Let H be the set of all E^+-neighbours of $U \setminus (A \cup D)$. Then $(A \setminus D) \cup (H \setminus R)$ covers all edges (i, j), for $p_{ij} \in C^\rho$, so for the lemma it is enough to show this set is small.

By applying (2) with $Y = U \setminus D$, we have

$$\Pr[|A \setminus D| \geq \ell/4] \leq \left(\frac{\mathrm{e} \cdot 2n^{1/2} \cdot cn^{1/4}}{(n+1)\ell/4}\right)^{\ell/4} \leq c_1^\ell,$$

for some constant $c_1 < 1$.

We have $|H| \leq cn^{1/4} \cdot 2n^{1/2}\ell = 2cn^{3/4}\ell$. Using (2) with $Y = V \setminus R$, we have

$$\Pr[|H \setminus R| \geq \ell/4] \leq \left(\frac{\mathrm{e} \cdot 2cn^{3/4}\ell \cdot cn^{1/4}}{n\ell/4}\right)^{\ell/4} = (4\mathrm{e}c^2)^{\ell/4},$$

which is exponentially small if $4\mathrm{e}c^2 < 1$.

To prove that the width measured by the negative literals is also small, recall that the degree of every $j \in V$ in E^- is at most 2. Hence the argument of Case 2 (with U and V switched) gives us the required bound.

Needless to say, if both the positive and the negative parts of C^ρ have ec-width $\leq \ell/2$, then $w_{ec}(C^\rho) \leq \ell$. ■

Say that a matching ρ fixes a CNF B if either $B^\rho = 0$, or there is no matching $\sigma \supseteq \rho$ such that $B^\sigma = 0$.

Lemma 8.3 (fixing lemma) *Let B be a CNF formula such that $w_{\mathrm{ec}}(B) \leq \ell$ and let $\ell < m < n$. Then*

$$\Pr_{\rho \sim \mathcal{R}_m}[\rho \textit{ does not fix } B] \leq \frac{\ell m(m-1)}{n-m+2}.$$

Proof. Let S be the set of all $\rho \in \mathcal{R}_m$ such that $B^\rho \neq 0$ and there exists $\sigma \supseteq \rho$ such that $|\sigma| < n$ and $B^\sigma = 0$. We will upper-bound $|S|$ by defining a one-to-one mapping $\theta : S \to \mathcal{R}_{m-1} \times [\ell]$.

Let orderings of the clauses of B, of partial matchings, of elements of $U \cup V$ and elements of $U \times V$ be fixed. For every clause C of B, pick a set W_C witnessing that $w_{\mathrm{ec}}(C) \leq \ell$. Let $\rho \in S$ be given.

Claim *There exists a clause C in B and a pair (i, j) disjoint with ρ such that $C^\rho \neq 0$, $C^{\sigma'} = 0$ for some $\sigma' \supseteq \rho \cup \{(i, j)\}$, and $i \in W_C$ or $j \in W_C$.*

Proof of claim. Take a clause C of B such that $C^\sigma = 0$ for some $\sigma \supseteq \rho$, $|\sigma| < n$. Since $B^\rho \neq 0$, and hence also $C^\rho \neq 0$, there exists a pair $(i, j) \in \sigma \setminus \rho$ and a literal $z \in C$ such that $z^\rho = z$ and $z^{\{(i,j)\}} = 0$. This can happen in several ways:

1. z is $\neg p_{ij}$, or
2. z is $p_{ij'}$ for some $j' \neq j$, or
3. z is $p_{i'j}$ for some $i' \neq i$.

In case 1 we take the pair (i, j). We do the same if $i \in W_C$ in case 2 or $j \in W_C$ in case 3. Otherwise $j' \in W_C$ in case 2 and $i' \in W_C$ in case 3. Consider case 2. Then j' may be in the range of σ or may not, but if it is not, then we can extend σ to a partial matching σ' that contains j' in its range, because $|\sigma| < n$. Thus either σ or σ' has a pair (i'', j') such that $j' \in W_C$ and, certainly, $C^{\sigma'} = C^{\sigma} = 0$. Case 3 is similar. This proves the claim. ■

Define $\theta : \mathcal{R}_m \to \mathcal{R}_{m-1}$ by $\theta(\rho) := (\rho \cup \{(i, j)\}, k)$, where we take the first clause C such that $C^\rho \neq 0$, $C^{\sigma'} = 0$ for some $\sigma' \supseteq \rho$ and (i, j) is the first pair in C satisfying the condition in the claim; k is the order of i in W_C, or the order of j in W_C if i is not in W_C.

We need to show that θ is one-to-one. Let $\rho' := \rho \cup \{(i, j)\}$. Given ρ', we can determine C, because it is the first clause of B such that $C^{\rho'}$ either is 0 or can be set to 0 by an extension of ρ'. The number k determines which pair $(x, y) \in \rho'$ with the property "$x \in W_C$ or $y \in W_C$" is (i, j).

Thus $|S|$ is at most $\ell \cdot |\mathcal{R}_{m-1}|$. Hence

$$\Pr_{\rho \sim \mathcal{R}_m}[\rho \in S] \leq \frac{\ell |\mathcal{R}_{m-1}|}{|\mathcal{R}_m|} = \frac{\ell \binom{n+1}{m-1} \binom{n}{m-2} (n + 2 - m)!}{\binom{n+1}{m} \binom{n}{m-1} (n + 1 - m)!} = \frac{\ell m(m-1)}{n - m + 2}.$$

■

Proof. (of Theorem 8.1) Now we can finish the proof of the theorem in a similar way as in the previous lower bounds. Let $\ell = \delta n^{1/12}$ for a sufficiently small constant $\delta > 0$. We will show that the $(1/4, \mathcal{M}_n)$-RR refutation Π of PHP_n has size $\geq \frac{1}{2}(1/d)^\ell$, where d is the constant from Lemma 8.2.

For a contradiction, suppose that the size of Π is $< \frac{1}{2}(1/d)^\ell$. Let $n_1 := \lfloor cn^{1/4} \rfloor$. For a random ρ from $\mathcal{R}_{n_1}$, the expected number of assignments from $\mathcal{M}_n$ that extend ρ and do not satisfy B^ρ is at most $1/4$ of all assignments from $\mathcal{M}_n$ that extend ρ. Hence, by Markov's inequality, with probability $\geq 1/2$ for a random ρ, B^ρ is satisfied by at least $1/2$ of such assignments.

By Lemma 8.2, for a random ρ from $\mathcal{R}_{n_1}$, all clauses of B^ρ have ec-width at most ℓ with probability $> 1/2$. Hence there exists $\rho \in \mathcal{R}_{n_1}$ such that both B^ρ is satisfied by $\geq 1/2$ of the assignments extending ρ and all clauses of B^ρ have ec-width $\leq \ell$. We take the refutation Π^ρ of $B^\rho \wedge \mathrm{PHP}_{n_1}$ obtained from Π by restricting by this particular ρ.

Next we apply Lemma 8.3 with n_1 instead of n, and $m = \eta n_1^{1/3}$, where $\eta > 0$ is a sufficiently small constant. The constants δ and η can, clearly, be chosen so that $3\ell < m$ and the estimate on the probability in Lemma 8.3 is positive. Let Π' be the refutation after applying the restriction from Lemma 8.3 to Π^ρ. In this refutation the auxiliary clauses are converted to clauses that cannot be falsified by any partial matching of size $< m$.

The rest of the proof is a standard adversary argument for proving a lower bound on the width. We traverse Π' from the empty clause at the bottom towards the initial clauses. For each clause we consider, we find a partial matching σ such that $\mathrm{supp}(\sigma) \supseteq W_C$ and $C^\sigma = 0$. Without loss of generality we can furthermore assume that $|\sigma| \leq \ell$, because C is falsified by the edges of σ that are incident with W_C. At the beginning we have $\sigma = \emptyset$. When going through a resolution step where C is derived from C_1 and C_2, we extend, if necessary, σ to σ' in an arbitrary way to ensure $W_{C_1}, W_{C_2} \subseteq \mathrm{supp}(\sigma')$. This is possible, because $|W_C| + |W_{C_1}| + |W_{C_2}| \leq 3\ell < m$. One of the clauses C_i must be falsified by σ' because the resolution rule is sound. We pick the falsified clause and, if necessary, reduce the size of the restriction to $\leq \ell$.

In this way we never reach an initial clause, because they cannot be falsified by partial matchings of size $< m$. Hence such a refutation does not exist and $|\Pi| \geq \frac{1}{2}(1/d)^\ell = \frac{1}{2}(1/d)^{\delta n^{1/12}}$. (Recall that $1/d$ is a constant greater than 1.) ■

9 Stronger refutation systems

We will briefly discuss the possibility of defining stronger random refutation systems. We speak about refutation systems, rather than proof systems, because we want to present then as generalizations random resolution; this is not essential and one can equivalently present these systems as proof systems with auxiliary formulas.

Given a refutation system P in which one can refute all unsatisfiable CNFs, we define *ε-random P distributions* and *ε-random P refutations* exactly as in Definitions 1.1 and 2.5, except that we replace resolution refutations with refutations in P. Note that the refuted formulas F and auxiliary formulas B are still CNFs, no matter whether the refutation system also uses more complex formulas. This is essential because otherwise we may be able to refute every formula trivially.[2]

To give an example of a nontrivial random refutation system, we consider the *random constant depth Frege system*. By Proposition 6.2 and Theorem 6.5, this system is stronger than random resolution, since it has polynomial size refutations of CPLS^2. We will show that it does not have subexponential size refutations of the pigeonhole principle.

Theorem 9.1 PHP_n *has no* $1/2$*-random constant depth Frege distribution of subexponential size.*

Proof. (sketch) To prove the theorem, we will combine the proof of Theorem 8.1 with a lower bound on constant depth Frege proofs.

We start by observing that Definition 2.7 and Proposition 2.8 can be adapted to the random constant depth Frege system, so it suffices to prove a lower bound on 1/2-random constant depth Frege refutations with respect to the same distributions on partial assignments as in the proof of Theorem 8.1, namely partial matchings of appropriate sizes.

[2] Certainly, some generalizations are possible; for example, cutting planes with auxiliary sets of inequalities, instead of CNFs, seem to be a nontrivial random refutation system.

Let such a refutation (B, Π) of PHP_n be given. In the first phase, we focus on the auxiliary CNF B. We apply random restrictions to reduce the ec-width of clauses in B and to "fix" it in the sense of Lemma 8.3. Note that once we "fix" B, the ec-width of B will play no role in the rest of the proof.

In the second phase, we apply random restrictions again, this time to define an *evaluation* of formulas in the refutation. The concept of evaluation was introduced in [19]; here we will use the simplified version of Urquhart and Fu [25]. An evaluation is an assignment of matching decision trees to all formulas of the refutation and their subformulas. (We will define matching decision trees below.) We say that a formula is evaluated true (false) if the tree assigned to it has all leaves labelled 1 (respectively 0). Evaluations have three basic properties:

1. the clauses of PHP_n are evaluated true,
2. Frege rules preserve the property of being evaluated true, and
3. the contradiction $\perp$ is evaluated false.

Hence the existence of an evaluation of a set of bounded depth formulas implies that the set is not a refutation of PHP_n. In the usual proof of the lower bound on bounded depth Frege refutations of PHP_n one shows, that given a set Φ of depth-d formulas of size $\leq 2^{n^{\delta_d}}$, for some constant $\delta_d > 0$, then after hitting Φ by a random restriction $\rho \in \mathcal{R}_m$, for a sufficiently small m, there exists an evaluation of formulas Φ^ρ.[3] Thus one gets a lower bound of $2^{n^{\delta_d}}$ on the size of depth-d Frege refutations of PHP_n.

In the case of random constant depth Frege refutations, the assumptions of the refutation moreover include B. It follows that for our lower bound, it suffices to prove that B is also evaluated true. But this is exactly what the fixing lemma guarantees.

We will now define matching decision trees and explain the last argument in more detail. A *matching decision tree* T is a finite labelled rooted tree where each node that is not a leaf is labelled by an element $v \in U \cup V$, the leaves are labelled by 0s and 1s, and the edges of T are labelled by pairs $(u, v) \in U \times V$ where u, or v is the label of the upper vertex of the labelled edge. The labeling furthermore must satisfy the following two conditions.

1. For every branch b of T, the labels of the edges of b are disjoint, i.e., they form a partial matching.
2. Let s be a node of T, let v be its label, and let ρ be the partial matching defined by the path from the root to s. If s is not a leaf, then the labels of the outgoing edges of s are all pairs (v, u) (or (u, v), depending on whether $v \in U$ or $v \in V$) that are disjoint with ρ.

To get some intuition, imagine that there is a bijection f between U and V (which, in reality, is impossible) and that T represents a subject querying f about the holes associated with

[3] The definition of Φ^ρ is more complicated than the one we used for CNFs; we refer the reader to [25].

pigeons and the pigeons associated with holes. After a certain number of queries, the subject decides to accept, or to reject f.

We say that a matching decision tree T *represents* a formula ϕ if the following condition is satisfied.

(*) For every branch b, if ρ is the partial matching defined by b and $a \in \{0, 1\}$ is the label of the leaf of b, then $\phi^\rho = a$.

Evaluations are defined by several conditions, but this property of evaluations is all we need to prove that all clauses of B are evaluated true. Indeed, suppose that B is not evaluated as true. Then the tree assigned to B has some branch with label 0. Thus, according to (*), we get some partial matching ρ such that $B^\rho = 0$. But the condition that Fixing Lemma 8.3 guarantees with high probability is that such a ρ does not exist and this is the condition that we ensured in the first phase of our proof. ■

References

[1] Miklós Ajtai. *The complexity of the pigeonhole principle.* Proc. 29th Annual Symp. on Foundations of Computer Science, pp. 346-55, 1988.

[2] Sanjeev Arora and Boaz Barak. *Computational Complexity. A modern Approach.* Cambridge, 2009.

[3] Albert Atserias and Neil Thapen. *The Ordering Principle in a Fragment of Approximate Counting.* ACM Transactions on Computational Logic 15:4, article 29, 2014.

[4] Eli Ben-Sasson and Avi Wigderson. *Short proofs are narrow - resolution made simple.* Journal of the ACM 48, pp. 149-169, 2001.

[5] Samuel Buss, Russell Impagliazzo, Jan Krajíček, Pavel Pudlák, Alexander Razborov and Jiří Sgall. *Proof complexity in algebraic systems and bounded depth Frege systems with modular counting.* Computational Complexity 6, pp. 256-298, 1996/1997.

[6] Samuel Buss, Leszek Aleksander Kołodziejczyk and Neil Thapen. *Fragments of approximate counting.* Journal of Symbolic Logic 79:2, pp. 496-525, 2014.

[7] Samuel Buss, Leszek Aleksander Kołodziejczyk and Konrad Zdanowski. *Collapsing Modular Counting in Bounded Arithmetic and Constant Depth Propositional Proofs.* Transactions of the American Mathematical Society 367, pp. 7517-7563, 2015.

[8] Mario Chiari and Jan Krajíček. *Witnessing functions in bounded arithmetic and search problems.* Journal of Symbolic Logic 63:3, pp. 1095-1115, 1998.

[9] Vašek Chvátal and Endre Szemerédi. *Many hard examples for resolution.* Journal of the ACM 35:4, pp. 759-768, 1988

[10] Stephen Cook and Robert Reckhow. *The relative efficiency of propositional proof systems.* Journal of Symbolic Logic 44:1, pp. 36-50, 1979.

[11] Russell Impagliazzo and Jan Krajíček. *A note on conservativity relations among bounded arithmetic theories.* Mathematical Logic Quarterly 48:3, pp. 375-7, 2002.

[12] Russel Impagliazzo, Tonian Pitassi and Alistair Urquhart. *Upper and lower bounds for tree-like cutting planes proofs.* Proc. Logic in Computer Science, pp. 220-228 1994.

[13] Emil Jeřábek. *On independence of variants of the weak pigeonhole principle.* Journal of Logic and Computation 17:3, pp. 587-604, 2007.

[14] Bala Kalyanasundaram and Georg Schnitger. *The Probabilistic Communication Complexity of Set Intersection.* Proc. Structure in Complexity Theory Conference, pp. 41-49, 1987.

[15] Jan Krajíček. *Interpolation theorems, lower bounds for proof systems, and independence results for bounded arithmetic.* Journal of Symbolic Logic 62:2, pp. 457-486, 1997.

[16] Jan Krajíček. *On the weak pigeonhole principle.* Fundamenta Mathematicae 170:1-3, pp. 123-140, 2001.

[17] Jan Krajíček. *A feasible interpolation for random resolution.* Logical Methods in Computer Science 13:1, 2017.

[18] Jan Krajíček. *Randomized feasible interpolation and monotone circuits with a local oracle.* Journal of Mathematical Logic, online ready, 2018.

[19] Jan Krajíček, Pavel Pudlák and Alan Woods. *Exponential lower bound to the size of bounded depth Frege proofs of the Pigeon Hole Principle.* Random Structures and Algorithms 7/1, pp.15-39, 1995

[20] Jan Krajíček, Alan Skelley and Neil Thapen. *NP search problems in low fragments of bounded arithmetic.* Journal of Symbolic Logic 72:2, pp. 649-672, 2007.

[21] Ran Raz and Avi Wigderson. *Monotone Circuits for Matching Require Linear Depth.* Journal of the ACM 39, pp. 736-744, 1992.

[22] Alexander Razborov. *Pseudorandom generators hard for k-DNF resolution and polynomial calculus resolution.* Annals of Mathematics 181:2, pp. 415-472, 2015.

[23] Alan Skelley and Neil Thapen. *The provably total search problems of bounded arithmetic.* Proceedings of the London Mathematical Society 103:1, pp. 106-138, 2011.

[24] Neil Thapen. *A tradeoff between length and width in resolution.* Theory of Computing 12, article 5, 2016.

[25] Alasdair Urquhart and Xudong Fu. *Simplified lower bounds for propositional proofs.* Notre Dame J. of Formal Logic 37:4, pp. 532-544, 1996

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il