18.10 Addendum: Arbitrary number of pigeons

Razborov's idea is to use a more subtle concept of "width" of clauses, tailor made for this particular CNF formula.

Theorem 18.22 *For every $m \geq n+1$, every resolution refutation proof of PHP_n^m has size at least $2^{\Omega(n^{1/4})}$.*

Recall that PHP_n^m denotes the AND of the following clauses (axioms):

- Pigeon Axioms: each of the m pigeon sits in at least one of n holes:

$$x_{i,1} \vee x_{i,2} \vee \cdots \vee x_{i,n} \quad \text{for all } i = 1, \ldots, m.$$

- Hole Axioms: no two pigeons sit is one hole:

$$\neg x_{i_1,j} \vee \neg x_{i_2,j} \quad \text{for all } i_1 \neq i_2 \text{ and } j = 1, \ldots, n.$$

A special feature of this CNF is that any resolution refutation of the set all its axioms (clauses) can be transformed to a monotone refutation of its pigeon axioms without any increase in the size of a derivation. To define monotone refutations, let $X_{i,j}$ be the OR of all but the ith variable in the jth column:

$$X_{i,j} = x_{1,j} \vee \cdots \vee x_{i-1,j} \vee x_{i+1,j} \vee \cdots \vee x_{m,j}\,.$$

By a *monotone refutation* of PHP_n^m we will mean a derivation of an empty clause from pigeon axioms and using the following *monotone resolution rule*:

$$\frac{A \vee x_{i,j} \qquad B \vee X_{i,j}}{A \vee B}\,.$$

Such a derivation can be obtained from the original (non-monotone) derivation by replacing each negated variable $\neg x_{i,j}$ by the OR of variables $X_{i,j}$. Note that, in general, this rule *is not* sound: there are assignments satisfying both assumptions but falsifying the conclusion. Still, the rule *is* sound if we consider only assignments satisfying all hole axioms; we call such assignments *legal*. That is, an assignment α is legal if it sends no two pigeons to the same hole (no column has more than one 1).

The following fact reduces the lower bounds problem for PHP_n^m to its monotone version.

Lemma 18.23 *If PHP_n^m has a resolution refutation of size S, then the set of pigeon axioms of PHP_n^m also has a monotone refutation of size at most S.*

Proof: Given a resolution refutation proof for PHP_n^m, just replace all occurrences of a negated variable $\neg x_{i,j}$ by the OR $X_{i,j} = \bigvee_{k \neq i} x_{k,j}$. It can be shown (do this!) that the resulting sequence of monotone clauses is a monotone refutation of the pigeon axioms. □

For the proof in the case when m is arbitrarily large, it will be convenient to increase the power of refutations by allowing a larger set of monotone derivation rules:

$$\frac{C_0 \vee X_{I_0,j} \qquad C_1 \vee X_{I_1,j}}{C}\,,$$

where $X_{I,j} = \bigvee_{i\in I} x_{i,j}$, $I_0 \cap I_1 = \emptyset$ and $C_0 \vee C_1 \le C$. From now on, by a *monotone refutation* of PHP_n^m we will mean a refutation of pigeon axioms using any of these rules. Note that these rules are still sound with respect to all legal truth assignments, that is, assignments sending no two pigeons to one hole: if such an assignment satisfies both clauses $C_0 \vee X_{I_0,j}$ and $C_1 \vee X_{I_1,j}$ then, due to the condition $I_0 \cap I_1 = \emptyset$, it must also satisfy at least one of the clauses C_0 or C_1, and hence, the clause C as well.

18.10.1 Size versus pseudo-width of refutations

Suppose we have a monotone refutation proof $\mathcal{R}$ of the pigeon axioms

$$X_{i,[n]} = \bigvee_{j=1}^{n} x_{i,j}, \qquad i = 1,\dots,m.$$

To analyze the refutation $\mathcal{R}$, we are going to allow much more axioms. For this we fix two parameters. First set

$$\delta := \frac{n}{2\log_2 m}\,.$$

A *threshold string* is a string $d = (d_1,\dots,d_m)$ of positive integers with $\delta < d_i \le n$ for all i. Having such a string d, we will allow all clauses of the form

$$X_{i,J} = \bigvee_{j\in J}^{n} x_{i,j} \quad \text{with} \quad i \in [m] \quad \text{and} \quad |J| \ge d_i$$

be used as axioms; we call such axioms *d-axioms*. Note that every monotone refutation of PHP_n^m is a monotone refutation of the set of d-axioms for the threshold string $d = (n,\dots,n)$.

Allowing more axioms does not hurt us, since our goal is to prove a *lower* bound on the size of a refutation. The reason for introducing new axioms is that we can then "filter out" from the refutation proof all clauses containing at least one such axiom: we just replace each such clause by the corresponding axiom.

For this purpose we consider the *degree of freedom*

$$d_i(C) = \big|\{j : x_{i,j} \in C\}\big|\,.$$

of each pigeon i in a clause C. This is the number of holes offered by C to this pigeon. The clause C is "filtered out" from the proof, that is, can be replaced by an axiom, if $d_i(C) \geq d_i$ for at least one pigeon i.

The main concept of our analysis will be the following very special notion of the "width" of refutation proofs of PHP_n^m, tailor made for this particular CNF. Namely, define the *pseudo-width* $w_d(C)$ of a clause as the number

$$w_d(C) = \big|\{i : d_i(C) \geq d_i - \delta\}\big|$$

of pigeons who passed the filter only "narrowly:" their degree of freedom $d_i(C)$ in C is near to the threshold d_i. The pseudo-width of a refutation $\mathcal{R}$ is the maximum pseudo-width of a clause in it.

Our first task (Lemma 18.24 below) will be to show that if the thresholds d_i are chosen in a clever way, then in every clause $C \in \mathcal{R}$ passing the filter (that is, having $d_i(C) < d_i$ for all pigeons i) almost all, namely, at least $m - \mathcal{O}(\log|\mathcal{R}|)$ pigeons pass it *safely*: their "degree of freedom" in C is well below the corresponding threshold d_i, is $\leq d_i - \delta$. Thus, the number of pigeons who *narrowly* (= non-safely) pass the filter $(d_1, \ldots, d_m)$, and hence, the pseudo-width of the refutation $\mathcal{R}$, must be at most $\mathcal{O}(\log|\mathcal{R}|)$.

Lemma 18.24 *If PHP_n^m has a resolution refutation of size S then there exists a threshold string d such that some set of at most S d-axioms has a monotone refutation of size S and pseudo-width $\mathcal{O}(\log S)$.*

The second task is to show that the number of pigeons who *narrowly* passed the filter, and hence, the pseudowidth of a refutation, must be large.

Lemma 18.25 *For every threshold string d, every monotone refutation $\mathcal{R}$ of a set of S d-axioms requires pseudo-width at least $\Omega\left(n/\log^3 S\right)$.*

Note that these two lemmas already imply the theorem. Let S be the minimum size of a resolution refutation of PHP_n^m. By Lemma 18.24, there exists a threshold string d such that some set of at most S d-axioms has a monotone refutation $\mathcal{R}$ of size S and pseudo-width $\mathcal{O}(\log S)$. But, by Lemma 18.25, the pseudo-width of $\mathcal{R}$ must be $\Omega\left(n/\log^3 S\right)$. Thus, $\log S = \Omega(n^{1/4})$, as desired. So, it remains to prove these two lemmas.

18.10.2 Short proofs have small pseudo-width

To prove the Lemma 18.24, we have to somehow "filter out" clauses of large pseudo-width. For this we need the following combinatorial lemma which may be of independent interest; we will give its proof later in Section 18.10.4.

Lemma 18.26 (Pigeon filter lemma) *Let $R = \{r_{i,k}\}$ be an $m \times S$ matrix with integer entries. If S is sufficiently large, then there exists a sequence $r_1, \ldots, r_m$ of integers such that $r_i < \lfloor \log m \rfloor$ and for every column k at least one of the following two events happen:*

(i) $r_{i,k} \le r_i$ *for at least one row* i*;*
(ii) $r_{i,k} > r_i + 1$ *for all but at most* $\mathcal{O}(\log S)$ *rows* i*.*

Suppose now that PHP_n^m has a resolution refutation of size S. Then, by Lemma 18.23, the set of all m pigeon axioms has a monotone refutation of size S. Fix such a refutation $\mathcal{R}$ an consider an $m \times S$ matrix $R = \{r_{i,C}\}$ whose rows correspond to pigeons $i \in [m]$ and columns to clauses $C \in \mathcal{R}$ of this refutation. Define the entries of this matrix by

$$r_{i,C} := \left\lfloor \frac{n - d_i(C)}{\delta} \right\rfloor + 1.$$

Let $r_1, \ldots, r_m$ be a sequence of integers guaranteed by Lemma 18.26. Set

$$d_i := \lfloor n - \delta r_i \rfloor + 1\,,$$

and note that $d_i > \delta$ because $r_i < \log m$. Hence, $d = (d_1, \ldots, d_m)$ is a threshold string. This special choice of the entries $r_{i,C}$ as well as of the d_i guarantee us two properties (check this!):

(iii) If $r_{i,C} \le r_i$ then $d_i(C) \ge d_i$.
(v) If $d_i(C) \ge d_i - \delta$ then $r_{i,C} \le r_i + 1$.

Now take an arbitrary clause $C \in \mathcal{R}$. Our goal is to show that either C contains a d-axiom (and C can be replaced by that axiom which reduces its pseudo-width $w_d(C)$ to 1) or $w_d(C) = \mathcal{O}(\log S)$.

If the first case (i) in Lemma 18.26 takes place, then $r_{i,C} \le r_i$ for some pigeon i, and by (iii), $d_i(C) \ge d_i$. Hence, in this case C contains a subclause $X_{i,J}$ which is a d-axiom, and can be replaced by this axiom.

If the second case (ii) in Lemma 18.26 takes place, then the number of pigeons i for which $r_{i,C} \le r_i + 1$, and hence, by (v), the number of pigeons i for which $d_i(C) \ge d_i - \delta$ does not exceed $\mathcal{O}(\log S)$. Hence, in this case the pseudo-width $w_d(C)$ of C cannot exceed $\mathcal{O}(\log S)$. This completes the proof of Lemma 18.24. □

18.10.3 Pigeonhole proofs have large pseudo-width

We now prove Lemma 18.25. Let $d = (d_1, \ldots, d_m)$ be an integer vector with $\delta < d_i \le n$ for all i. Take an arbitrary set $\mathcal{A}$ of $|\mathcal{A}| \le S$ d-axioms, and set

$$w_0 := \frac{\epsilon \delta^2}{n \log |\mathcal{A}|}$$

where $\epsilon > 0$ is a sufficiently small constant. Take an arbitrary monotone refutation $\mathcal{R}$ of $\mathcal{A}$. We will show that $w_d(C) > w_0$ for at least one clause C in $\mathcal{R}$.

Suppose the opposite, i.e., that $w_d(C) \leq w_0$ for all clauses $C \in \mathcal{R}$. Our goal is to show that then the empty clause 0 does not belong to $\mathcal{R}$, i.e., that $\mathcal{R}$ is *not* a refutation of $\mathcal{A}$.

Recall that each axiom in $\mathcal{A}$ has the form $X_{i,J} := \bigvee_{j \in J} x_{i,j}$ for some pigeon i and some set J of $|J| \geq d_i$ holes; $X_{i,J}$ is the axiom for the pigeon i. Let

$$\mathcal{A}_i = \{X_{i,J} \in \mathcal{A} : |J| \geq d_i\}$$

denote the set of all such axioms in $\mathcal{A}$, and let $\mathcal{A}_I := \bigcup_{i \in I} \mathcal{A}_i$. For a clause C in $\mathcal{R}$ let

$$\mathcal{A}_C = \bigcup_{i : d_i(C) \geq d_i - \delta} \mathcal{A}_i$$

denote the set of all axioms in $\mathcal{A}$ corresponding to pigeons that are "free enough" in the clause C.

As before, truth assignments are $m \times n$ $(0,1)$ matrices α. Such an assignment is legal if it satisfies all hole axioms, that is, if no column has more than one 1. Say that an assignment α is *critical* if it is legal and no row of α has more than

$$\ell := \left\lfloor \frac{\delta}{4w_0} \right\rfloor$$

1-entries. We say that a set $\mathcal{C}$ of clauses *implies* a clause C, and write $\mathcal{C} \models C$, if every critical assignment α satisfying all clauses of $\mathcal{C}$ also satisfies C.

Claim 18.27 For every clause C in $\mathcal{R}$ we have that $\mathcal{A}_C \models C$.

This already gives the desired contradiction, because $\delta < d_i$ for all i implies that $\mathcal{A}_0 = \emptyset$, and hence, that $\mathcal{A}_0 \not\models 0$. Thus, it remains to prove the lemma.

To prove Claim 18.27, we argue by induction on the number of steps in the derivation of C in $\mathcal{R}$. The case $C \in \mathcal{A}$ is obvious since then $C \in \mathcal{A}_C$.

For the inductive step suppose that $\mathcal{A}_A \models A$, $\mathcal{A}_B \models B$ and C is obtained from clauses A, B by a single application of the monotone refutation rule. Since the rule is sound with respect to all legal (and hence, also for all critical) truth assignments, we have that $\{A, B\} \models C$. Hence, if we take the set

$$I = \{i \mid d_i(A) \geq d_i - \delta \text{ or } d_i(B) \geq d_i - \delta\}$$

of pigeons of large degree of freedom in at least one of the clauses A or B, then $|I| \leq 2w_0$ and $\mathcal{A}_I \models C$. Let us choose a minimal $I \subseteq \{1, \ldots, m\}$ such that $\mathcal{A}_I \models C$; then still $|I| \leq 2w_0$. We will show that, in fact,

$$I \subseteq \{i \mid d_i(C) \geq d_i - \delta\}\,;$$

this will obviously imply $\mathcal{A}_I \subseteq \mathcal{A}_C$, and hence, $\mathcal{A}_C \models C$.

Assume the contrary, and pick an arbitrary $i_0 \in I$ with $d_i(C) < d_i - \delta$. Since I is minimal, we have that $\mathcal{A}_{I \setminus \{i_0\}} \not\models C$. Hence, there is a critical assignment $\alpha = (a_{i,j})$ which satisfies all clauses in $\mathcal{A}_{I \setminus \{i_0\}}$ but falsifies C. We

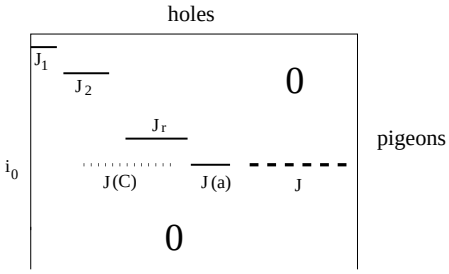

Fig. 18.7 $I \setminus \{i_0\} = \{1, \ldots, r\}$, $J_i = \{j \mid a_{i,j} = 1\}$, $J(a) = \{j \mid a_{i_0,j} = 1\}$ and $J(C) = \{j \mid x_{i_0,j} \in C\}$.

may assume that $a_{i,j} = 0$ for all $i \notin I \setminus \{i_0\}$ and all j, because C is positive and none of such variables $x_{i,j}$ appears in $\mathcal{A}_{I\setminus\{i_0\}}$. Let now

$$J = \big\{j \mid x_{i_0,j} \notin C \text{ and } a_{i,j} = 0 \text{ for all } i \in I \setminus \{i_0\}\big\}$$

be the set of holes "permissible" for the pigeon i_0 (see Fig. 18.7): if we pick an arbitrary subset $J' \subseteq J$ of size $|J'| = \ell$ and change the assignment α by letting $a_{i_0,j} = 1$ iff $j \in J'$, then we will get a critical assignment α' which still satisfies all clauses in $\mathcal{A}_{I\setminus\{i_0\}}$ (we have not touched other pigeons) but falsifies C.

We want to show that J' can be chosen in such a way that this new assignment α' will also satisfy all clauses in $\mathcal{A}_{i_0}$; this will give the desired contradiction with $\mathcal{A}_I \models C$.

First, observe that the set J is large enough: since $d_{i_0}(C) < d_{i_0} - \delta$ and each row of α has at most ℓ 1-entries, we have that

$$|J| \geq n - \big(|I| \cdot \ell + d_{i_0}(C)\big) \geq n - \big(2w_0\ell + (d_{i_0} - \delta)\big) \geq n - d_{i_0} + \delta/2\,.$$

Now pick $\mathbf{J}$ uniformly and at random among all ℓ-element subsets of J, and let $\boldsymbol{\alpha}$ be the random assignment resulting from the assignment a by setting to 1 all $a_{i_0,j}$ with $j \in \mathbf{J}$. Take an arbitrary axiom $A \in \mathcal{A}_{i_0}$, and let $J_A = \{j \mid x_{i_0,j} \in A\}$ be the set of holes offered by the clause A to the pigeon i_0. Since $|J_A| \geq d_{i_0}$, by (18.10.3) we have

$$|J_A \cap J| \geq \delta/2\,.$$

Now we can apply Chernoff's inequality and conclude that

$$\mathrm{Prob}[A(\boldsymbol{\alpha}) = 1] = \mathrm{Prob}[J_A \cap \mathbf{J} \neq \emptyset] \geq 1 - e^{-\Omega(p\cdot\delta)} \geq 1 - e^{-\Omega(\delta\ell/n)}.$$

Since

$$\frac{\delta\ell}{n} \geq \frac{\delta^2}{4w_0 n} = \frac{\delta^2}{4n} \cdot \frac{n \cdot \log|\mathcal{A}|}{\epsilon\delta^2} = \frac{\log|\mathcal{A}|}{4\epsilon},$$

we obtain that

$$\mathrm{Prob}[A(\boldsymbol{\alpha}) = 1] \geq 1 - |\mathcal{A}|^{-2},$$

if the constant ϵ is sufficiently small. Since clearly, $|\mathcal{A}_{i_0}| < |\mathcal{A}_{i_0}|^2 \leq |\mathcal{A}|^2$, this implies that, for at least one choice α' of $\boldsymbol{\alpha}$, *all* axioms in $\mathcal{A}_{i_0}$ will be satisfied. Since (as we observed above) the assignment α' also satisfies all axioms in $\mathcal{A}_{I\setminus\{i_0\}}$ but falsifies C, we obtained a contradiction with $\mathcal{A}_I \models C$.

This completes the proof of Claim 18.27, and thus, the proof of Lemma 18.25. □

18.10.4 Proof of the pigeon filter lemma

The lemma is a direct consequence of the following property of randomly chosen numbers. Let m and S be positive integers. Set $t := \lfloor \log m \rfloor - 1$, and let $\boldsymbol{r}$ be a random variable taking its values in $[t] = \{1, \ldots, t\}$ with probabilities

$$\mathrm{Prob}[\boldsymbol{r} = t] = 2^{-(t+1)} \quad \text{and} \quad \mathrm{Prob}[\boldsymbol{r} = s] = 2^{-s} \quad \text{for each} \quad s = 1, 2, \ldots, t-1.$$

Claim 18.28 Let S be a positive integer, $x = (x_1, \ldots, x_m)$ an integer vector, and let $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_m$ be m independent copies of $\boldsymbol{r}$. Then with probability at least $1 - \mathcal{O}(S^{-2})$ at least one of the following two events happens:

A_x: $\boldsymbol{r}_i \geq x_i$ for at least one integer x_i;
B_x: $\boldsymbol{r}_i < x_i - 1$ for all but at most $\mathcal{O}(\log S)$ integers x_i.

Proof: Our goal is to show that at least one of $\mathrm{Prob}[\overline{A}_x]$ and $\mathrm{Prob}[\overline{B}_x]$ is at most $\mathcal{O}(S^{-2})$, implying that the desired sequence $r_1, \ldots, r_m$ satisfying both conditions of Lemma 18.26 exist with probability at least $1 - \mathcal{O}(S^{-2})$.

Define the "weight" of x as $W(x) := \sum_{i=1}^{m} 2^{-x_i}$. We consider two cases depending on whether $W(x) \geq 2 \ln S$ or not.

Case 1: $W(x) \geq 2 \ln S$. Let $I = \{i \mid x_i \leq t\}$ and note that

$$\sum_{i \notin I} 2^{-x_i} \leq m 2^{-(t+1)} \leq 2\,.$$

Therefore,

$$\sum_{i \in I} 2^{-x_i} \geq W(x) - 2 \geq 2 \ln S - 2\,.$$

On the other hand, for every $i \in I$ we have $\mathrm{Prob}[\boldsymbol{r}_i \geq x_i] \geq 2^{-x_i}$, and these events are independent. Since $\mathrm{Prob}[\boldsymbol{r}_i \geq x_i] = 0$ for all $i \notin I$, we have that in this case

$$\mathrm{Prob}[\overline{A}_x] = \mathrm{Prob}[\forall i : \boldsymbol{r}_i < x_i] = \prod_{i \in I} \left(1 - 2^{-x_i}\right) \leq \exp\left(-\sum_{i \in I} 2^{-x_i}\right) \leq e^2 S^{-2}\,,$$

where the last inequality follows from (18.10.4).

Case 2: $W(x) \le 2 \ln S$. We first show that $\mathrm{Prob}[\boldsymbol{r} \ge x_i - 1] \le 2^{2-x_i}$ for every i. Indeed, if $x_i > t$ then either $x_i = t+1$ and

$$\mathrm{Prob}[\boldsymbol{r} \ge x_i - 1] = \mathrm{Prob}[\boldsymbol{r} = t] = 2^{1-t} = 2^{2-x_i}\,,$$

or $x_i \ge t+2$ and $\mathrm{Prob}[\boldsymbol{r} \ge x_i - 1] = 0$. If $x_i \le t$ then

$$\mathrm{Prob}[\boldsymbol{r} \ge x_i - 1] = \sum_{s=x_i-1}^{t} 2^{-s} \le 2^{1-x_i} \sum_{j=0}^{\infty} 2^{-j} \le 2^{2-x_i}\,.$$

Hence, the expected number of i for which $\boldsymbol{r}_i \ge x_i - 1$ does not exceed

$$\sum_{i=1}^{m} 2^{2-x_i} = 4 \sum_{i=1}^{m} 2^{-x_i} = 4W(x) \le 8 \ln S\,.$$

Since the events "$\boldsymbol{r}_i \ge x_i - 1$" are independent, we may apply Chernoff's inequality and conclude that, for any sufficiently large constant c,

$$\mathrm{Prob}[\overline{B}_x] = \mathrm{Prob}[|\{i : \boldsymbol{r}_i \ge x_i - 1\}| \ge c \ln S] \le S^{-2}\,. \qquad \square$$