# A note on propositional proof complexity of some Ramsey-type statements

Jan Krajíček*

Charles University and Academy of Sciences, Prague

**Abstract**

Any valid Ramsey statement $n \longrightarrow (k)^2_2$ can be encoded into a DNF formula $\mathsf{RAM}(n, k)$ of size $O(n^k)$ and with terms of size $\binom{k}{2}$.

Let $r_k$ be the minimal $n$ for which the statement holds. We prove that $\mathsf{RAM}(r_k, k)$ requires exponential size constant depth Frege systems, answering a problem of Krishnamurthy and Moll [15].

As a consequence of Pudlák's work in bounded arithmetic [19] it is known that there are quasi-polynomial size constant depth Frege proofs of $\mathsf{RAM}(4^k, k)$, but the proof complexity of these formulas in resolution $R$ or in its extension $R(\log)$ is unknown. We define two relativizations of the Ramsey statement that still have quasi-polynomial size constant depth Frege proofs but for which we establish exponential lower bound for $R$.

The complexity of proving various Ramsey-type combinatorial statements is well studied in connection with Peano arithmetic or systems of second order arithmetic, or even with set theory. The foremost example is the Paris-Harrington extension of finite Ramsey theorem, see [17].

However, even sooner Krishnamurthy and Moll [15] proposed Ramsey theorem as a source of hard *propositional* tautologies (we discuss this more below). In this paper we continue the study of *propositional complexity* of Ramsey theorem, motivated by the following problem:

*Supported in part by grants IAA100190902, AV0Z10190503, MSM0021620839, LC505 (Eduard Čech Center) and by a grant from the John Templeton Foundation.

- *Find a sequence of formulas in DNF (preferably with narrow terms) that have short constant depth Frege proofs (in DeMorgan language) but require long proofs in $R(\log)$.*

Proof system $R(\log)$, introduced in [11], operates with clauses (i.e. disjunctions) formed not only of literals but also of terms (i.e. conjunctions of literals) via natural inference rules. The size[1] of an $R(\log)$ proof is the minimal $s$ such that the proof has at most $s$ symbols and all terms have size at most $\log(s)$.

The problem has several facets and interesting consequences. We do not require the sequence of formulas to be uniform in any way: the existence of any sequence of formulas with the required properties implies (via a known technique using the relation of reflection principles to simulations among proof systems, cf.[10]) the existence of a first-order principle that translates into a sequence of DNF formulas with the same properties. The existence of such DNF formulas implies also a non-conservativity result for bounded arithmetic $T_2(\alpha)$ over $T^2_2(\alpha)$, either a non-$\forall\Sigma^b_2(\alpha)$ conservativity or even a non-$\forall\Sigma^b_1(\alpha)$ conservativity, if the terms in the formulas are narrow. Furthermore, several of the so called no-gaps theorems [3, 20] would then yield the same non-conservativity of even $T^3_2(\alpha)$ over $T^2_2(\alpha)$, and that $R(\log)$ does not simulate the next higher fragment of constant depth Frege systems (of the so called $\Sigma$-depth 1 of [9, 10]). Finally, such a proof complexity separation can be usually turned into a non-reducibility result for corresponding $\mathcal{NP}$ search problems, cf.[4, 7, 8, 14].

For these consequences of the solution of the problem to hold we need to interpret the qualifications *narrow*, *short* and *long* as follows: narrow term should mean *poly-logarithmic* in the size of the formula, short proof in constant depth Frege systems should mean *quasi-polynomial* in the size of the formula, and long $R(\log)$-proof should mean *not quasi-polynomially bounded*. In fact, one may expect that there should be such formulas with terms of constant size, having polynomial size constant depth Frege proofs and requiring exponential size $R(\log)$-proofs.

The DNF formulas that were proposed in [3] (in the language of bounded arithmetic) as suitable candidates for the problem formalize a Ramsey statement. Consider a valid Ramsey statement

$$n \longrightarrow (k)^2_2$$

[1]Some later authors have used a more naive definition, see the end of the introduction.

expressing that every graph (tacitly undirected) with vertices $[n] = \{1, \ldots, n\}$ contains a homogeneous subgraph, a clique or an independent set, of size at least $k$. This can be encoded into a propositional tautology $\mathsf{RAM}(n, k)$ in a DNF form as follows. The formula is built from atoms $x_e$, one for each of potential $\binom{n}{2}$ edges $e \in [n]^{(2)}$ (the set of unordered pairs of different elements of $[n]$), and for each subset $A \subseteq [n]$ of size $k$ contains two terms

$$\mathsf{Cli}(A) \ := \ \bigwedge_{e \subseteq A} x_e$$

and

$$\mathsf{Ind}(A) \ := \ \bigwedge_{e \subseteq A} \neg x_e$$

as disjuncts. Hence $\mathsf{RAM}(n, k)$ has $2\binom{n}{k}$ disjuncts each of size $\binom{k}{2}$.

Denote by $r_k$ the minimal $n$ for which the statement is valid. It is known that $2^{k/2} < r_k < 4^k$, cf.[6]. These critical formulas $\mathsf{RAM}(r_k, k)$ were first considered as candidate hard formulas for resolution in [15], where the authors established an $r_k/2$ width lower bound and an exponential lower bound for the Davis-Putnam procedure. We note in Section 1 that the method from [11], relating proof complexity of Ramsey statements to that of the so called (weak) pigeonhole principle ($\mathsf{PHP}$), can be straightforwardly modified to show that these formulas are too hard for our purposes: they require exponential size constant depth Frege proofs. No bounds were known for the formulas previously[2].

Fortunately if we replace the optimal parameter $r_k$ by $4^k$ the proof complexity decreases dramatically: by [19] it is known that $\mathsf{RAM}(4^k, k)$ have quasi-polynomial size (i.e. $2^{k^{O(1)}}$) constant-depth Frege proofs (the estimate to the depth resulting from the argument in [19] has been optimally counted in [1]). In the direction of lower bounds we know that $\mathsf{RAM}(4^k, k)$ requires width of resolution proofs at least $\frac{1}{2}4^{k/4}$ and that $R^*(\log)$-proofs (i.e. tree-like $R(\log)$-proofs) require exponential size, cf. [11]. Moreover, it is known that a lower bound for resolution proofs of $\mathsf{RAM}(4^k, k)$ would follow from a lower bound for $R(2)$-proofs of the weak PHP with $n^4$ pigeons and $n$ holes, cf.[11].

Thus our original problem can be reduced to:

- *Show that formulas* $\mathsf{RAM}(4^k, k)$ *require long (at least more than of a quasi-polynomial size)* $R(\log)$*-proofs.*

[2] I pointed out this lower bound in various talks but never wrote it up. I use this occasion to give a finitary version of the original model-theoretic argument of [11].

No lower bound is known for $R$ either. In this paper we do not prove the lower bound for $R(\log)$ but we make nevertheless a progress: We shall define two relativizations of the formulas, to be denoted $\mathsf{RAM}^U(n,k)$ and $\mathsf{RAM}^f(n,k)$, and we show that while they are still easy for constant depth Frege systems, they both require exponential size $R$-proofs.

Let us remark on other combinatorial principles sometimes mentioned in the connection with the problem. First, the reader may wonder why we base the tautologies on the simplest case of finite Ramsey theorem rather than on some enhanced version unprovable even in strong theories. The reason is that these statements have a $\Pi^0_2$ form $\forall x \exists y A(x,y)$ and their unprovability is deduced from an enormous growth of a bound $N_n$ to $y$ in terms of the instance $x := n$. But the propositional formula encoding the $n$-th instance will have the length at least $N_n$ and hence even a simple proof system (often $R$ or $R(\log)$) can prove the tautology in polynomial size because the qualification polynomial means "polynomial in $N_n$".

The second remark concerns forms of the weak pigeonhole principle (with various settings of the parameters) that may seem simpler to work with than with Ramsey statement. Unfortunately the instances provable shortly in constant depth Frege systems do actually also have short proofs in $R(\log)$, cf.[16].

We use only standard concepts of proof complexity; the reader may find any relevant background in [10, 11]. Resolution is denoted $R$, its tree-like version $R^*$, and similarly for $R(\log)$. An $R$-proof means a proof in $R$, similarly for other proof systems. More details on the link to conservativity problems in bounded arithmetic can be found in [20]. Here we only remark on the definition of $R(\log)$. In [11] a system $R^+$ was defined, operating with clauses of terms via natural rules, and for a function $f$ on $\mathbf{N}$ one defined the $R(f)$-size of an $R^+$-proof: the minimal $s$ such that the proof has at most $s$ symbols and uses terms of size at most $f(s)$. Some later authors interpreted the definition as saying that terms have size at most $f(n)$, where $n$ is some canonical parameter of the formula, e.g. its number of variables. In this sense one can have an exponential lower bound for $R(\log)$-proofs while they use only terms of size $\log(n)$. Such a result says nothing about bounded arithmetic independence from $T^2_2(\alpha)$; to maintain the correspondence between proof systems and bounded arithmetic one has to use the original definition.

## 1 A lower bound for the critical parameter

Recall that the size of formulas $\mathsf{RAM}(n, k)$ is $O(n^k)$. In particular, the size of $\mathsf{RAM}(r_k, k)$ is at most $O(4^{k^2})$ due to the bound $r_k \leq 4^k$.

**Theorem 1.1** *For every $d \geq 2$ there is $\epsilon > 0$ such that for $k \geq 1$ every depth $d$ Frege proof of $\mathsf{RAM}(4^k, k)$ must have the size at least $2^{{r_k}^\epsilon}$.*

**Proof :**

We shall use the idea of an argument from [11]. Put $n := r_k - 1$, and let $p_{i,j}$, $i \in [n+1], j \in [n]$ be $(n+1)n$ atoms of the usual pigeonhole principle formula $\mathsf{PHP}_n$:

$$\bigvee_i \bigwedge_j \neg p_{i,j} \ \vee \bigvee_{i_1 \neq i_2, j} p_{i_1,j} \wedge p_{i_2,j} \ \vee \bigvee_{i, j_1 \neq j_2} p_{i,j_1} \wedge p_{i,j_2}$$

where $i, i_1, i_2 \in [n+1]$ and $j, j_1, j_2 \in [n]$.

Let $\pi$ be a depth $d$ size $s$ Frege proof of $\mathsf{RAM}(r_k, k)$ and let the variables of this formula be $x_e$, $e \in [r_k]^{(2)}$. By the definition of $r_k$ there exists a graph $G = ([n], E)$ that has no homogeneous subgraph of size $k$. Use it to define the following substitution for variables $x_e$ in terms of the variables of the $\mathsf{PHP}$ formula:

$$\sigma(x_{\{u,v\}}) \ := \bigvee_{\{i,j\} \in E} p_{u,i} \wedge p_{v,j} \ .$$

The following claim is established by induction on $t$.

**Claim 1:** *For any $t$ such that $1 \leq t \leq n$ and any size $t$ subset $A \subseteq [r_k]$ there are constant depth Frege proofs of size $n^{O(t)}$ of both formulas*

$$\sigma(\mathsf{Cli}(A)) \wedge \neg \mathsf{PHP}_n \ \longrightarrow \bigvee_{B \subseteq [n], |B| = t} \ \bigwedge_{\{i,j\} \subseteq B} E(i,j)$$

*and*

$$\sigma(\mathsf{Ind}(A)) \wedge \neg \mathsf{PHP}_n \ \longrightarrow \bigvee_{B \subseteq [n], |B| = t} \ \bigwedge_{\{i,j\} \subseteq B} \neg E(i,j)$$

Because $G$ has no homogeneous subset of size $k$, both conjunctions

$$\bigwedge_{\{i,j\} \subseteq B} E(i,j) \text{ and } \bigwedge_{\{i,j\} \subseteq B} \neg E(i,j)$$

have value 0 for $|B| = k$. This entails the next claim.

**Claim 2:** *For each $A \subseteq [r_k]$ there are constant depth Frege proofs of size $n^{O(k)}$ of both formulas*

$$\sigma(\mathsf{Cli}(A)) \longrightarrow \mathsf{PHP}_n$$

*and*

$$\sigma(\mathsf{Ind}(A)) \longrightarrow \mathsf{PHP}_n \ .$$

Combining the proof $\sigma(\pi)$ of $\sigma(\mathsf{RAM}(r_k, k))$ with the proofs from Claim 2, we get a constant depth proof of $\mathsf{PHP}_n$ of size at most

$$O(sn^2) \ + \ 2\binom{r_k}{k} n^{O(k)} \leq \ O(sn^2) \ + \ n^{O(\log(n))} \ .$$

It is known ([2, 13, 18]) that $\mathsf{PHP}_n$ requires constant depth Frege proofs of size $2^{n^\delta}$, $\delta$ depending on the depth. This entails the lower bound.

**q.e.d.**

# 2 Relativizations and constant depth Frege upper bounds

In this and the next section we concentrate on formula $\mathsf{RAM}(n, k)$ with the non-optimal parameter $n := 4^k$. From now on $n$ is fixed to denote this value.

We define two relativizations of formula $\mathsf{RAM}$ and show that they are still shortly provable in constant depth Frege systems. This will be complemented in the next section by exponential resolution lower bounds for both of them.

The first relativizations $\mathsf{RAM}^U$ is simpler to define but it appears less flexible for the hopeful attack on $R(\log)$ than the second relativization $\mathsf{RAM}^f$. The latter formula has also a trivial upper bound proof for constant depth Frege systems while for the former one has to check that the proof of the upper bound for $\mathsf{RAM}$ in [19] will work here as well.

Relativization in proof complexity appeared in [12] in connection with model-theoretic methods and lead to the question how relativization of first-order principles influences the proof complexity of their propositional translations. For resolution this was answered by a beautiful theorem of Dantchev and Riis [5]. Ramsey principle does not fall under the scope of this theorem but we shall be able to use the random restriction method from [5] on our formulas $\mathsf{RAM}^U$ and $\mathsf{RAM}^f$ nevertheless.

The first relativization $\mathsf{RAM}^U(n,k)$ formalizes the following principle:

- *Let $n = 4^k$, let $G = ([n], E)$ be any graph and let $U \subseteq [n]$ be arbitrary. Then either the induced subgraph with vertices $U$ or the induced subgraph with vertices $\bar{U} := [n] \setminus U$ contains a homogeneous subgraph of size $k-1$.*

At east one of the subgraphs has size $m \geq n/2$ and the validity of $\mathsf{RAM}^U(n,k)$ thus follows from the validity of the Ramsey relation $m \longrightarrow (\lfloor \frac{\log(m)}{2} \rfloor)^2_2$ as $k = \frac{\log(n)}{2}$.

**Definition 2.1** *Let $k \geq 2$, $n = 4^k$, and let $x_e$ and $u_i$ be atoms, where $e \in [n]^{(2)}$ and $i \in [n]$. Formula $\mathsf{RAM}^U(n,k)$ is the disjunction of the following $4\binom{n}{k-1}$ formulas:*

$$\bigwedge_{i \in A} u_i \ \wedge \ \mathsf{Cli}(A)$$

$$\bigwedge_{i \in A} \neg u_i \ \wedge \ \mathsf{Cli}(A)$$

$$\bigwedge_{i \in A} u_i \ \wedge \ \mathsf{Ind}(A)$$

$$\bigwedge_{i \in A} \neg u_i \ \wedge \ \mathsf{Ind}(A)$$

*where $A$ ranges over subsets of $[n]$ of size $k-1$.*

The size of formula $\mathsf{RAM}^U(n,k)$ is $O(k^2 n^k)$, its terms are narrow of size $k - 1 + \binom{k-1}{2} \leq k^2$.

The second formalization $\mathsf{RAM}^f(n,k)$ formalizes the following principle:

- *Let $n = 4^k$, $G = ([n], E)$ be any graph and let $f : [n/4] \rightarrow [n]$ be an arbitrary injective function. Then the induced subgraph whose vertex set is the range $Rng(f)$ of $f$ contains a homogeneous subgraph of size $k-1$.*

This is valid as the induced subgraph has $4^{k-1}$ vertices.

**Definition 2.2** *Let $k \geq 2$, $n = 4^k$, and let $x_e$ and $f_{i,j}$ be atoms, where $e \in [n]^{(2)}$, $i \in [n/4]$, and $j \in [n]$. Formula $\mathsf{RAM}^f(n,k)$ is the disjunction of the following formulas:*

1. $\bigwedge_j \neg f_{i,j}$, *any* $i$,

2. $f_{i,j_1} \wedge f_{i,j_2}$, *any* $i$ *and* $j_1 \neq j_2$,

3. $f_{i_1,j} \wedge f_{i_2,j}$, *any* $i_1 \neq i_2$ *and* $j$,

4. $f_{i_1,j_1} \wedge \ldots \wedge f_{i_k,j_k} \wedge \mathsf{Cli}(\{j_1,\ldots,j_k\})$, *any ordered $k$-tuples of different elements* $i_1,\ldots,i_k \in [n/4]$ *and* $j_1,\ldots,j_k \in [n]$,

5. $f_{i_1,j_1} \wedge \ldots \wedge f_{i_k,j_k} \wedge \mathsf{Ind}(\{j_1,\ldots,j_k\})$, *any ordered $k$-tuples of different elements* $i_1,\ldots,i_k \in [n/4]$ *and* $j_1,\ldots,j_k \in [n]$,

*where* $i, i_1, i_2 \in [n/4]$ *and* $j, j_1, j_2 \in [n]$.

Note that $\mathsf{RAM}^f(n,k)$ has $n^{O(k)}$ terms which, due to item 1., are not narrow anymore, and total size $n^{O(k)}$ too.

**Theorem 2.3** *Let $k \geq 2$ and $n = 4^k$. Formulas $\mathsf{RAM}^U(n,k)$ and $\mathsf{RAM}^f(n,k)$ have both quasi-polynomial size (i.e. size $2^{k^{O(1)}}$) constant depth Frege proofs.*

**Proof :**

We start with the upper bound for $\mathsf{RAM}^f$. Assume for the sake of contradiction $\neg\mathsf{RAM}^f(n,k)$. Define a graph $H$ with vertices $[n/4]$ by pulling back the edges of $G$ via map $f$. As $f$ is injective $H$ is well-defined. As we assume $\neg\mathsf{RAM}^f(n,k)$, graph $H$ has no homogeneous subgraph of size $k-1$. But this can be brought to a contradiction in a constant depth Frege system: take a short proof of $\mathsf{RAM}(n/4, k-1)$ in the system (it exists by [19]) and substitute in it for the edge variables the definition of the edges of $H$. This will increase the depth by a constant and the size by a factor of $O(n^2)$.

For $\mathsf{RAM}^U(n,k)$ there does not seem to be such a simple proof by substitution into a known proof of $\mathsf{RAM}$ but it suffices to look how $\mathsf{RAM}(n,k)$ is proved in [19]. The argument there rests on the following construction. Given a graph with vertex set $V$ which has no homogeneous subgraph of size $\ell$, a mapping $F : V \to \{0,1\}^\ell$ is defined (by short constant depth formulas) that is injective. This is then brought into a contradiction with the weak $\mathsf{PHP}$ if $\ell$ is too small (i.e. $2^\ell \leq |V|/2$).

In our case we apply this construction to both induced subgraphs of $G$ with the vertex sets $U$ and $\bar{U}$ respectively, getting two injective maps

$$F_1 : U \to \{0,1\}^{k-1} \text{ and } F_0 : \bar{U} \to \{0,1\}^{k-1} .$$

They combine to an injective map from $[n]$ into $\{0,1\}^k$ which is then brought to a contradiction with the weak $\mathsf{PHP}$ as before.

**q.e.d.**

# 3 Resolution lower bounds

The strategy of the lower bound argument is analogous to that of [5]: we show, employing a random restriction, that if either relativization $\mathsf{RAM}^U(n,k)$ or $\mathsf{RAM}^f(n,k)$ had a short $R$-proof then the unrelativized $\mathsf{RAM}(n,k-1)$ would have a narrow $R$-proof, contradicting the width lower bound from [11]. We start by recalling the latter, stating it in the form we need later.

**Theorem 3.1 ([11])** *Any $R$-proof of $\mathsf{RAM}(m,\ell)$ must have the width at least*

$$\frac{1}{2}2^{\ell/2} .$$

**Lemma 3.2** *Let $k \geq 2$ and $n = 4^k$. Assume that there is an $R$-proof of $\mathsf{RAM}^U(n,k)$ of size $s \leq 2^{n^{1/11}}$. Then $\mathsf{RAM}(n,k-1)$ has an $R$-proof of width at most $n^{1/5}$.*

**Proof :**

Let $\pi$ be a size $s$ $R$-proof of $\mathsf{RAM}^U(n,k)$. Substitute for all atoms $u_i$ a random value $\sigma(u_i) \in \{0,1\}$, independently and with probability $1/2$ of each value. Put $U := \{i \in [n] \mid \sigma(u_i) = 1\}$.

After $\sigma$ is chosen define a partial evaluation of variables $x_e$ as follows:

- If $e \subseteq U$ or $e \subseteq \bar{U}$ leave $x_e$ unassigned.
- Otherwise give $x_e$ randomly value 0 or 1, independently and with equal probability $1/2$.

Denote $\rho \supseteq \sigma$ the substitution thus defined.

A clause $C$ in $\pi$ has the form

$$v_1 \vee \ldots \vee v_s \vee \ell_{e_1} \vee \ldots \vee \ell_{e_t}$$

where $v_1, \ldots, v_s$ are literals $u_i$ or $\neg u_i$ and $\ell_e$ is literal $x_e$ or $\neg x_e$.

**Claim 1:** *Let $C$ be a clause as above. In the random process defining $\rho$ the probability that $\rho(C) \neq 1$ is at most $(3/4)^{\sqrt{t}/2}$.*

Consider the first edge $e_1$. Decide randomly the membership of its end-points in $U$. The probability that $\rho(\ell_{e_1}) \in \{0,1\}$ is $1/2$ and that it is not equal to 1 is $1/4$. Hence $3/4$ bounds the probability that $\rho(\ell_{e_1}) \neq 1$. In a general step take $e \in \{e_1, \ldots, e_t\}$ such that the membership in $U$ has not been decided for at least one end-point of $e$. Then again $\rho(\ell_e) \neq 1$ with probability at most $3/4$. In each step we decide about at most two points their membership in $U$ and $p$ points can cover up to $\binom{p}{2} \leq p^2$ edges. Hence this process can go on for at least $\sqrt{t}/2$ steps. This proves the claim.

**Claim 2:** *The probability that all clauses in $\pi$ not given value 1 by $\rho$ have the width less than $n^{1/5}$ is positive.*

The probability to fail to make true all clauses of width at least $n^{1/5}$ is bounded above by Claim 1 by

$$s \cdot (3/4)^{\frac{1}{2}n^{1/10}} \leq 2^{n^{1/11}} \cdot (3/4)^{\frac{1}{2}n^{1/10}}$$

which goes to 0.

By Claim 2 we can take a $\rho$ not leaving in $\rho(\pi)$ any clause wider than $n^{1/5}$. Let $|U| = m$. Assume without a loss of generality that $m \leq n/2$. The restricted proof $\rho(\pi)$ is a proof of a disjunction of two formulas

$$\mathsf{RAM}(m, k-1) \ \vee \ \mathsf{RAM}(n-m, k-1)$$

written in disjoint sets of variables. Identify $[m]$ with a subset of $[n-m]$ of size $m$, and consequently also the variables of $\mathsf{RAM}(m, k-1)$ with some variables of $\mathsf{RAM}(n-m, k-1)$, turning $\rho(\pi)$ into a proof of $\mathsf{RAM}(n-m, k-1)$, i.e. of $\mathsf{RAM}(n, k-1)$ too.

**q.e.d.**

Now we prove an analogous statement for the other relativization.

**Lemma 3.3** *Let $k \geq 2$ and $n = 4^k$. Assume that there is an R-proof of* $\mathsf{RAM}^f(n,k)$ *of size $s \leq 2^{n^{1/11}}$. Then* $\mathsf{RAM}(n,k-1)$ *has an R-proof of width at most $n^{1/5}$.*

**Proof :**

Let $\pi$ be a size $s$ $R$-proof of $\mathsf{RAM}^f(n,k)$. Assign first to each $i \in [n]$ a random value $\sigma(i) \in \{0,1\}$, independently and with equal probability $1/2$ of the values. Put $V = \{i \in [n] \mid \sigma(i) = 1\}$. Chernoff's bound implies that $\frac{n}{4} \leq |V| \leq \frac{3n}{4}$ with probability of failing at most $e^{-n^2/16}$.

Assuming $|V| \geq n/4$ proceed as follows. Take for $f$ an injective function from $[n/4]$ into $V$, selected from the set of all such functions in some canonical way, and evaluate variable $f_{i,j} := \rho(f_{i,j}) \in \{0,1\}$ accordingly.

Then extend $\rho$ to $\eta \supseteq \rho$ by randomly restricting some of the edge variables $x_e$ as follows:

- If $e \subseteq V$ leave $x_e$ unassigned.
- If $e \not\subseteq V$, assign $x_e$ value $\eta(x_e) \in \{0,1\}$, independently and uniformly at random.

A clause $D$ in $\pi$ has the form

$$g_1 \vee \ldots \vee g_s \vee \ell_{e_1} \vee \ldots \vee \ell_{e_t}$$

where $g_1, \ldots, g_s$ are literals $f_{i,j}$ or $\neg f_{i,j}$ and $\ell_e$ is literal $x_e$ or $\neg x_e$.

**Claim 1:** *Assume $|V| \leq 3n/4$. Let $D$ be a clause as above. In the random process defining $\eta$ the probability that $\eta(D) \neq 1$ is at most $(7/8)^{\sqrt{t}/2}$.*

The claim is proved analogously to Claim 1 in the proof of Lemma 3.2, noting that the assumption $|V| \leq 3n/4$ implies that $x_e$ is not assigned a value is at most $3/4$.

This yields the next claim as before.

**Claim 2:** *The probability that $n/4 \leq |V| \leq 3n/4$ and that all clauses in $\pi$ not given value 1 by $\eta$ have the width less than $n^{1/5}$ is positive.*

Take a restriction $\eta$ not leaving in $\eta(\pi)$ any clause wider than $n^{1/5}$. The proof is concluded by noting that $\eta(\pi)$ is a proof of $\mathsf{RAM}(m,k-1)$ where $m = |V|$, i.e. of $\mathsf{RAM}(n,k-1)$ as well.

**q.e.d.**

Theorem 3.1 imply together with Lemmas 3.2 and 3.3 the lower bound.

**Theorem 3.4** *Let $k \geq 2$ and $n = 4^k$. Then every R-proof of $\mathsf{RAM}^U(n,k)$ or of $\mathsf{RAM}^f(n,k)$ must have the size at least $\Omega(2^{n^{1/11}})$.*


# Acknowledgements:

I am indebted to Stefan Dantchev (Durham) and Neil Thapen (Prague) for comments on the draft of the paper. I thank for discussions on the topic to the current French Café circle: Phuong Nguyen, Pavel Pudlák, Neil Thapen and Iddo Tzameret.


# References


[1] K. Aehlig and A. Beckmann, A remark on the induction needed to prove the Ramsey principle, unpublished manuscript, (2006).

[2] M. Ajtai, The complexity of the pigeonhole principle, in: *Proc. IEEE 29$^{th}$ Annual Symp. on Foundation of Computer Science*, (1988), pp. 346-355.

[3] M. Chiari and J. Krajíček, Lifting independence results in bounded arithmetic, *Archive for Mathematical Logic*, **38(2)**, (1999), pp.123-138.

[4] P. Beame, S. A. Cook, J. Edmonds, R. Impagliazzo, and T. Pitassi, The relative complexity of NP search problems. *J.of Computer and System Sciences*, **57**, (1998), pp.3-19.

[5] S. Dantchev and S. Riis, On Complexity gaps for Resolution-based proof systems, in: 12th Annual Conf. of the EACSL, Computer Sci Logic, LNCS **2903**, Springer, (2003), pp. 142-154.

[6] P. Erdös, Some remarks on the theory of graphs, *Bull. of the AMS*, **53**, (1947), pp.292-294.

[7] J. Hanika, *Search Problems and Bounded Arithmetic*, PhD Thesis, Charles University, Prague, 2004.

[8] J. Hanika, Herbrandizing search problems in Bounded Arithmetic, *Mathematical Logic Quarterly*, **50(6)**, (2004), pp.577-586.

[9] J. Krajíček, Lower bounds to the size of constant-depth propositional proofs, *Journal of Symbolic Logic*, **59(1)**, (1994), pp.73-86.

[10] J. Krajíček, *Bounded arithmetic, propositional logic, and complexity theory*, Encyclopedia of Mathematics and Its Applications, Vol. **60**, Cambridge University Press, (1995).

[11] J. Krajíček, On the weak pigeonhole principle, *Fundamenta Mathematicae*, Vol.**170(1-3)**, (2001), pp.123-140.

[12] J. Krajíček, Combinatorics of first order structures and propositional proof systems, *Archive for Mathematical Logic*, **43(4)**, (2004), pp.427-441.

[13] J. Krajíček, P. Pudlák,and A. Woods, An Exponential Lower Bound to the Size of Bounded Depth Frege Proofs of the Pigeonhole principle", *Random Structures and Algorithms*, **7(1)**, (1995), pp.15-39.

[14] J. Krajíček, A. Skelley and N. Thapen, NP search problems in low fragments of bounded arithmetic, *J. of Symbolic Logic*, **72(2)**, (2007), pp. 649-672.

[15] B. Krishnamurthy and R. N. Moll, Examples of hard tautologies in the propositional calculus, in: 13th ACM Symposium on Th. of Computing, (1981), pp.28-37.

[16] A. Maciel, T. Pitassi, and A. Woods, A new proof of the weak pigeonhole principle, *Journal of Computer and System Sciences archive*, **64(4)**, (2002), pp.843 - 872.

[17] J. Paris and L. Harrington, A Mathematical Incompleteness in Peano Arithmetic, in Handbook for Mathematical Logic (Ed. J. Barwise). Amsterdam, Netherlands: North-Holland, (1977).

[18] T. Pitassi, P. Beame, and R. Impagliazzo, Exponential lower bounds for the pigeonhole principle, *Computational complexity*, **3**, (1993), pp.97-308.

[19] P. Pudlák, Ramsey's Theorem in Bounded Arithmetic. In: Proc. Computer Science Logic'90, eds. E.Borger et.al., LNCS **553**, Springer-Verlag, (1991), pp.308-317.

[20] A. Skelley and N. Thapen, The provably total search problems of bounded arithmetic, submitted (preprint 2008).

**Mailing address:**
Department of Algebra
Faculty of Mathematics and Physics
Charles University
Sokolovská 83, Prague 8, CZ - 186 75
The Czech Republic