Space Complexity of Random Formulae in Resolution

Eli Ben-Sasson,[1]* Nicola Galesi[2]†
[1]*Institute of Computer Science, Hebrew University, Jerusalem, Israel; e-mail: elli@cs.huji.ac.il*
[2]*Universitat Politècnica de Catalaunya, Department Llenguatges i Sistemes Informátics, Barcelona 08034, Spain; e-mail: galesi@lsi.upc.es*

Received 11 December 2001; revised 3 September 2002; accepted 7 March 2003

DOI 10.1002/rsa.10089

ABSTRACT: We study the space complexity of refuting unsatisfiable random k-CNFs in the Resolution proof system. We prove that for $\Delta \geq 1$ and any $\epsilon > 0$, with high probability a random k-CNF over n variables and Δn clauses requires resolution clause space of $\Omega(n/\Delta^{1+\epsilon})$. For constant Δ, this gives us linear, optimal, lower bounds on the clause space. One consequence of this lower bound is the first lower bound for size of treelike resolution refutations of random 3-CNFs with clause density $\Delta \gg \sqrt{n}$. This bound is nearly tight. Specifically, we show that with high probability, a random 3-CNF with Δn clauses requires treelike refutation size of $\exp(\Omega(n/\Delta^{1+\epsilon}))$, for any $\epsilon > 0$. Our space lower bound is the consequence of three main contributions: (1) We introduce a 2-player Matching Game on bipartite graphs G to prove that there are no perfect matchings in G. (2) We reduce lower bounds for the clause space of a formula F in Resolution to lower bounds for the complexity of the game played on the bipartite graph $G(F)$ associated with F. (3) We prove that the complexity of the game is large whenever G is an expander graph. Finally, a simple probabilistic analysis shows that for a random formula F, with high probability $G(F)$ is an expander. We also extend our result to the case of G-PHP, a generalization of the Pigeonhole principle based on bipartite graphs G. © 2003 Wiley Periodicals, Inc. Random Struct. Alg., 23: 92–109, 2003

Correspondence to: E. Ben-Sasson
*Supported by the Clore Foundation Doctoral Scholarship.
†This work was done while the author was a member of the School of Mathematics at the Institute for Advanced Study of Princeton, supported by the NSF Grant No. CCR-9987845. Also partly supported by Spanish CICYT TIC2001-1577-C03-03 Project and by a Canadian CSERC grant.

1. INTRODUCTION

1.1. Proof Space Complexity

The importance of Proof Complexity comes from the close relationship between its fundamental questions and long-standing open problems in Complexity Theory. In its more general setting, a Propositional Proof System can be defined as a polynomial time computable function that is onto the set of tautologies [7]. In similarity with Circuit Complexity, we have very little knowledge of the properties of arbitrary proof systems, and thus usually we restrict our attention to some simple concrete proof systems. The system receiving most attention by far is the Resolution system. The attention arises from several reasons. Resolution has a single rule, that is relatively simple to analyze. Resolution is used heavily in practice for Automated Theorem Proving. In the last 15 years several fundamental works have analyzed the complexity of proofs in Resolution, showing that many tautologies require exponentially long refutations in Resolution [10, 13, 4–6].

As it is well known, the complexity of an algorithm is measured not only in terms of the running time but also in terms of the *memory consumption.* The *space* not only is a natural measure for the complexity of algorithms, but, as is the time measure, is also widely studied in Complexity Theory.

The proof complexity measure related to the *time* complexity measure is the *size* of a proof, that is, the number of *symbols* used in the proof or, when polynomially related, the number of formulas used. Recently [2, 9] introduced and studied a new complexity measure for propositional proof systems, analogous to the *space* complexity measure for circuits. For the Resolution system, Esteban and Torán [9] proposed to consider as measure for the Resolution space complexity, the number of different clauses that must be simultaneously available (that is, *kept in memory*) to obtain the empty clause. Alekhnovich et al. [2] generalized several aspects, of [9]. First of all they extended the definition of clause space complexity in a natural way to all important propositional proof systems such as Frege Systems or Polynomial Calculus. Moreover, to measure the memory content in a given moment during a proof, they also considered *the variable space,* that is, the overall number of variables used, as well as the total number of symbols needed, *the bit space.*

In spite of its recent introduction several non trivial upper and lower bounds for space complexity are already known. Torán in [12] gave lower bounds for clause space in Resolution. He considered two well-known tautologies, for which several lower bounds for the *size* are known: the PHP_n and the so-called Tseitin Tautologies. Alekhnovich et al. [2] devised a general technique to give nontrivial lower bounds for the clause space in Resolution, and in other proof systems. Using this method, they obtain non trivial clause space lower bounds in Resolution for class of formulas like PHP_n, GT_n, and CT_n.

1.2. Random CNFs

It is well known that in circuit complexity simple counting arguments show that a random function is hard to compute. In studying the complexity of a given proof system, it is natural to ask what is the proof complexity of a tautology taken at random. However, we don't have a definition of what is a random tautology. Still, in some cases, if we restrict our attention only to certain kinds of tautologies, we can deduce information on their random behavior. An easy calculation shows that for a high enough constant Δ, with high probability [i.e., with probability $1 - o(1)$] a random 3-CNF formula with n variables

and Δn clauses is unsatisfiable (Δ is called the *clause density*). Let us introduce the definition of a random CNF.

Definition 1.1 (Random CNFs). *Let $\mathbb{F}_m^{k,n}$ be the probability distribution obtained by selecting m clauses uniformly at random from the set of all $2^k \cdot \binom{n}{k}$ clauses of size k over n variables. $\mathcal{F} \sim \mathbb{F}_m^{k,n}$, means that $\mathcal{F}$ is selected at random from this distribution. A random k-CNF formula is a formula $\mathcal{F} \sim \mathbb{F}_m^{k,n}$.*

It is not hard to see that for any integer k, there is a constant Δ_0 such that with high probability, a random k-CNF with n variables and Δ_0 clauses is unsatisfiable. The question of the existence and value of a *satisfiability threshold constant* is an important open problem in combinatorics, and for more information on this subject (see, e.g., [1, 8, 11]).

The proof size of unsatisfiable random CNFs has been widely studied. Chvatál and Szmerédi in their seminal paper [6] showed that, with high probability, any random 3-CNF over n variables and Δn clauses for $\Delta = O(1)$, requires exponentially long Resolution proofs to be refuted. The importance of their work was in showing that, in fact, Resolution is a very weak proof system, because in some sense almost all unsatisfiable 3-CNF require exponential size proofs to be refuted. Their lower bound was later improved and simplified by Beame and Pitassi in [4] and finally improved up to a ratio $\Delta = o(n^{1/4})$ by Beame, Karp, Pitassi, and Saks in [3], and reformulated in terms of a general technique based on the *width* by Ben-Sasson and Wigderson in [5]. All these results, as well as the results presented in this paper, can be generalized to k-CNFs for arbitrary constant $k > 3$.

1.3. RESULTS

Lower bounds for the clause space of unsatisfiable random 3-CNF didn't follow from any of the techniques devised in the previous works on space complexity [2, 9, 12]. In fact, this was left as an open problem in both [12] and [2].

In this paper we study the clause space complexity of refuting unsatisfiable random CNF in Resolution. Our main result is the following.

Theorem 1.2. *For any integer $k \geq 3$, any constant $\epsilon > 0$ and any $\Delta \geq 1$, with high probability refuting a random k-CNF with n variables and $\Delta \cdot n$ clauses requires Clause Space $\Omega(n/\Delta^{1+\epsilon})$.*

For instance, setting Δ to be a constant we get linear lower bounds.[1] Since [9] showed that the clause space of any formula is at most $n + 1$, this lower bound is optimal up to a multiplicative constant.

Corollary 1.3 [Constant Δ]. *For $k \geq 3$ and any constant $\Delta \geq 1$, a random k-CNF with n variables and Δn clauses requires with high probability $\Omega(n)$ clause space to refute.*

[1] We define the clause space of $\mathcal{C}$ to be ∞ whenever $\mathcal{C}$ is satisfiable. Thus the theorem is valid for values Δ for which whp a random CNF is satisfiable.

Another interesting corollary is for large clause density, which we state for concreteness for $k = 3$.

Corollary 1.4. *For any constant $\delta > 0$, and any $1 > \epsilon > 0$ a random 3-CNF with n variables and $n^{2-\delta}$ clauses requires whp clause space $\Omega(n^{\delta(1-\epsilon)})$ to refute.*

1.4. Techniques

A 3-CNF induces a natural bipartite graph, between the set of clauses and the set of variables, where each clause is connected to the three variables appearing in it. A matching in the graph $G(F)$ associated with F naturally defines a partial assignment, which satisfies part of the formula F.

In order to prove clause space lower bounds on F, we define a 2-player game (*The Matching Game*) to be played on the bipartite graph $G(F)$, associated with F. The aim of the first player is to prove that there is no perfect matching. The second player is an opponent who instead tries to force a perfect matching. The first player should complete his task "remembering" as few as possible of his moves in the game.

We prove two main properties: First, that the clause space of refuting a 3-CNF can be reduced to the natural complexity measure for the game (i.e., the minimal number of moves the first player needs to remember to win); then we prove that when the graph is an expander, the first player needs to remember a large number of moves, where the size is correlated to the expansion parameters of the graph.

As pointed out above, space lower bounds for resolution were already known. While for Tseitin formulas only an ad hoc lower bound proof was given (see [2, 12]). Alekhnovich et al. in [2] introduced a technique to prove space lower bounds only for those contradictions including *wide* clauses. In this work we devise a general technique based on the Matching Game, which correlates the expansion factor of the graph $G(F)$ with the minimal space to refute F in resolution. First, this allows us to overcome the problem of requiring large clauses among the initial ones. Moreover, our direct connection to the expansion of a graph reflects the analogous case of *size* lower bounds for Resolution, where the expansion is correlated with the minimal *width* (see [5]). This makes our characterization of clause space quite general. Indeed, we can extend our result also to the case of another tautology based on bipartite graphs G, the *G-PHP* introduced by [5]. On these contradictions previous techniques failed since initial clauses are of constant size.

1.5. Lower Bounds for Treelike Resolution Size

The lower bound of Theorem 1.2, which applies in the case of 3 CNFs for clause density greater than $\sqrt{n}$, is in contrast to the known lower bounds for *size* of general resolution proofs within this density range: Our best size lower bounds for random 3-CNFs become trivial when the clause density reaches $\sqrt{n}$. Actually, Theorem 1.2 gives us the first lower bounds for treelike resolution proof size, for clause density beyond $\sqrt{n}$. A treelike resolution proof is a proof in which every derived clause can be used at most once in resolution inference. We state and prove the theorem for general k.

Theorem 1.5. *For any $k \geq 3$, any $\epsilon > 0$, and any $\Delta \geq 1$, with high probability refuting a random k-CNF with n variables and $\Delta \cdot n$ clauses requires treelike resolution size of* $\exp(\Omega(n/\Delta^{1+\epsilon}))$.

Proof. Follows immediately from Theorem 1.2 and the following theorem of [9]. ∎

Theorem 1.6 [9]. *Let ϕ be an unsatisfiable CNF formula with tree-like resolution refutations of size S. Then ϕ has a resolution refutation of space $\lceil \log S \rceil + 1$.* ∎

For the sake of completeness, we state the explicit result for 3-CNFs with high clause density.

Theorem 1.7. *For any $\delta > 0$ and any $1 > \epsilon > 0$, refuting a random 3-CNF with n variables and $n^{2-\delta}$ clauses requires treelike refutation size of* $\exp(\Omega(n^{\delta(1-\epsilon)}))$.

We end this subsection by pointing out that for general resolution proofs of random 3-CNFs, we do not have any lower bounds for $\Delta \geq n^{1/4}$. This is in contrast with the space and the minimal treelike size, for which we have nearly optimal lower bounds.

1.6. Tightness of Lower Bounds

Both lower bounds, for clause space as well as for treelike resolution size, are nearly tight, by the following upper bound of [3].

Theorem 1.8 [3]. *Let $k \geq 3$, and $\Delta > \theta_k$. With high probability the size of a minimal treelike resolution refutation of a random k-CNF with n variables and Δn clauses, is $2^{O(n/\Delta^{1/(k-2)})} n^{O(1)}$.*

This upper bound is achieved by a satisfiability algorithm called *ordered DLL,* which produces a specific treelike resolution proof. Moreover, [3] proved a matching lower bound for this algorithm. Our lower bound (Theorem 1.5), that applies to all treelike resolution proofs, matches their upper bound up to a polynomial factor. By Theorem 1.6, the upper bound of [3] implies a clause space upper bound of $O(n/\Delta^{1/(k-2)})\log n$, which matches our lower bound, up to a polynomial factor. Recently, Zito [14] proved an even tighter upper bound on the clause space of $O(n/\Delta^{1/(k-2)})$.

1.7. Paper Organization

In Section 2 we give some preliminary definitions. Section 3 is dedicated to the definition of the Matching Game and its relationship with the clause space. In Section 4 we prove a lower bound for the Matching Game played on graphs from the class of (r, c)-bipartite expanders. In Section 5 we prove a Lemma studying under which parameters $r = r(n)$ and $c = c(n)$, a random k-CNF over n variables defines an (r, c)-bipartite expander. This result joint with results from previous sections gives the lower bound for resolution. Finally in Section 6 we show that the Matching Game applies also to the *G-PHP*.

2. DEFINITIONS

Let V be finite set of Boolean variables. A *literal* l is either a variable $x \in V$ or its negation $\bar{x}$. A *clause* is a disjunction (eventually empty) of literals. A *CNF formula* is a conjunction of clauses, and it will be convenient to see it as a set of clauses. We use calligraphic letters

(e.g., $\mathcal{F}$, $\mathcal{C}$) for denoting CNF formulas, and capital letters for denoting clauses. A 3-CNF formula is a CNF formula in which all the clauses have exactly 3 literals.

For $\mathcal{F}$ a formula, $Vars(\mathcal{F})$ is the set of variables appearing in $\mathcal{F}$. A *restriction* on $\mathcal{F}$ is a partial function $\rho : Vars(\mathcal{F}) \to \{0, 1\}$. $\mathcal{F}_\rho$ denotes the CNF formula obtained from $\mathcal{F}$ after applying ρ in the standard way: If a literal l is set to 1 by ρ, then all clauses C of $\mathcal{F}$ such that $l \in C$ disappear in $\mathcal{F}|_o$; all clauses C in $\mathcal{F}$ such that $C = \bar{l} \vee D$ become D in $\mathcal{F}_\rho$. We say $\rho(x) = \star$ when $x \notin Domain(\rho)$. The size of a restriction, $|\rho|$, is $|Domain(\rho)|$.

2.1. Clause Space in Resolution

Resolution is a refutation proof system for unsatisfiable CNF formulas based on the following propositional *resolution rule:*

$$\frac{D_1 \cup \{x\} \qquad D_2 \cup \{\bar{x}\}}{D_1 \cup D_2}.$$

We define a *space complexity measure* following the definitions of [2]. Let $[n]$ be the set $\{1, \ldots, n\}$.

Definition 2.1. *A* configuration *is a set of clauses. A* refutation π *of a CNF* $\mathcal{F}$*, is a sequence of configurations* $\mathcal{C}_0, \ldots, \mathcal{C}_s$ *such that* $\mathcal{C}_0 = \emptyset$, $\mathcal{C}_s = \{\blacksquare\}$ (*the empty clause*), *and, for all* $t \in [s]$, $\mathcal{C}_t$ *is obtained from* $\mathcal{C}_{t-1}$ *by one of the following rules*:

AXIOM DOWNLOAD $\mathcal{C}_t := \mathcal{C}_{t-1} \cup \{C\}$ *for some clause* $C \in \mathcal{F}$;

MEMORY ERASING $\mathcal{C}_t := \mathcal{C}_{t-1} - \{C\}$ *for some clause* $C \in \mathcal{C}_{t-1}$;

INFERENCE ADDING $\mathcal{C}_t := \mathcal{C}_{t-1} \cup \{C\}$, *for some C obtained by a single application of the resolution rule to two clauses in* $\mathcal{C}_{t-1}$.

The following definitions define the measure for the resolution space: the *clause space.* Let $\pi_F \vdash F$ denote that π_F is a resolution derivation (in the form of sequence of configurations) of F.

Definition 2.2 (Clause Space). *For* $\mathcal{C}$ *a set of clauses,* $|\mathcal{C}|$ *is the number of clauses in* $\mathcal{C}$*. The space of a set of configurations* $\pi = \{\mathcal{C}_0, \ldots, \mathcal{C}_s\}$ *is the maximal number of clauses in a configuration of* π*. The* clause *space of refuting an unsatisfiable CNF* $\mathcal{F}$*, denoted CSpace*($\mathcal{F}$)*, is the minimal space of a resolution refutation of* $\mathcal{F}$*, if a refutation exists. If no refutation exists* (*i.e.,* $\mathcal{F}$ *is satisfiable*)*, then CSpace*($\mathcal{F}$) $\stackrel{\text{def}}{=} \infty$.

Definition 2.3 (Width). $|C|$ [*also denoted by w*(*C*)*—the* width *of C*] *is the number of literals in the clause C. The width of a set of clauses* $\mathcal{F}$ *is the width of the largest clause in* $\mathcal{F}$*. The width of a resolution refutation of* $\mathcal{F}$ *is the width of the largest clause in the refutation. Finally, the width of refuting an unsatisfiable set of clauses* $\mathcal{F}$*, denoted by* $w(\vdash\mathcal{F})$ *is the minimal width taken over all refutations of* $\mathcal{F}$.

3. THE MATCHING GAME

We wish to reduce the space required to refute a CNF formula to a natural combinatorial game played on a bipartite graph. We shall prove lower bounds for this game whenever the bipartite graph is an expander.

Definition 3.1 (Bipartite Expanders). *A bipartite graph* $G = \langle V \cup U, E \rangle$ *is called an* (r, c)-bipartite expander *if*

$$\forall\ V' \subset V |V'| \leq r, \qquad |N(V')| \geq (1 + c)|V'|,$$

where $N(V')$ *is the set of neighbors of* V'.

3.1. Proving That There Is No Perfect Matching

For $G = \langle (V \cup U), E \rangle$ a bipartite graph, if $|V| > |U|$, then there is no matching of V into U. We wish to prove this claim, using "limited space." For this purpose let us define a two-player game. The players are Pete (Prover) and Dana (Disprover). Pete tries to prove that there is no matching from V to U, and Dana tries to prove that such a matching exists. Pete has k fingers, numbered $\{1, \ldots, k\}$, and Dana has k fingers, numbered identically. We start with all vertices of G uncovered, and on each round one of the following occurs:

1. Pete Places a finger j on some uncovered $v \in V$, and Dana must answer by placing her finger j on some uncovered $u \in U$ that is a neighbor of v.
2. Pete removes a finger j from a covered $v \in V$, and Dana answers by removing her finger j from its covered neighbor $u \in U$.

Notice that the set of fingers placed on the graph corresponds naturally to a partial matching in G: Each v covered by a finger j of Pete is matched to the u that is covered by Dana's finger j in reply. Pete wins the game when he places a finger on some vertex such that all its neighbors are already covered by Dana. If Pete cannot win the game, then Dana wins. We define $MSpace(G)$ (Matching Space) to be the minimal number of fingers that Pete needs in order to win the game. Clearly $MSpace(G) \leq |U| + 1$.

3.2. Reducing Clause Space to Matching Space

Definition 3.2. *For* $\mathcal{C}$ *a CNF formula, define* $G(\mathcal{C})$ *to be the following bipartite graph*:

1. V *is the set of clauses.*
2. U *is the set of variables.*
3. $(C, x) \in E(G)$ *iff the variable* x *appears in the clause* C (*we do not care whether* x *appears as a positive or negative literal*).

The main claim of this section is:

Theorem 3.3. *CSpace*($\mathcal{C}$) $\geq$ *MSpace*($G(\mathcal{C})$).

For its proof, we will be using the following locality lemma from [2]. We give the proof of the lemma for the sake of completeness.

Lemma 3.4 (Locality Lemma) [2]. *Let* ρ *be a restriction and* $\mathcal{C}$ *be a set of clauses, such that* $(\bigwedge_{C\in\mathcal{C}} C|_\rho) \equiv 1$. *Then there exists a subrestriction* ρ' *of* ρ, *such that* $(\bigwedge_{C\in\mathcal{C}} C|_{\rho'}) \equiv 1$, *and* $|\rho'| \leq |\mathcal{C}|$.

Proof. For each clause $C \in \mathcal{C}$, there exists some literal $\ell \in C$ such that $\ell_{|_\rho} = 1$. For each C, fix ℓ_C to be one such literal (arbitrarily). Set

$$\rho'(x) = \begin{cases} \rho(x), & \exists C \in \mathcal{C}, \ell_C \text{ is a literal over } x, \\ \star, & \text{otherwise.} \end{cases}$$

Notice that if ℓ_C and ℓ_D are literals over the same variable x, then $\ell_C = \ell_D$, because $\ell_C|_\rho = \ell_D|_\rho = 1$, so ρ' is well defined. By definition, $\mathcal{C}|_{\rho'} = 1$, and $|\rho'| \leq |\mathcal{C}|$. ∎

Proof of Theorem 3.3. For $m = \{(C_{i_1}, x_{i_1}), \ldots, (C_{i_k}, x_{i_k})\}$ a partial matching in $G(\mathcal{C})$ of size k, define $\rho(m)$ to be the restriction of size k that sets the variable x_{i_j} to the value that satisfies the clause C_{i_j}, for $j = 1, \ldots, k$, and leaves all other variables unassigned.

Assume Dana has a winning strategy when the matching game is played on $G(\mathcal{C})$ using k fingers. We will use this strategy to show that every set of clauses derivable in clause space k is satisfiable. Let $\mathcal{C}_0, \ldots, \mathcal{C}_l$ be a derivation from $\mathcal{C}$, of space k. We construct inductively a sequence of partial matchings in $G(\mathcal{C})$, $m_0, \ldots, m_\ell$ $m_t \subset E$, $t = 0, \ldots, \ell$, that maintains the following properties for all $t = 0, \ldots, \ell$:

1. m_t is the matching obtained by playing the matching game with k fingers for $t \leq t$ rounds.
2. $|m_t| \leq |\mathcal{C}_t|$.
3. $\mathcal{C}_t|_{\rho(m_t)} = 1$.

m_0 is the empty matching. For the induction step, we prove the claim according to the type of rule used at time t:

1. **Axiom Download:** If the new axiom C is satisfied by the restriction, we do nothing, and clearly all properties are maintained. Otherwise, the number of fingers Pete has at time $t-1$ is at most $|\mathcal{C}_{t-1}| \leq k-1$, and thus, when Pete places a finger on C in $G(\mathcal{C})$, Dana can respond by placing a finger on some uncovered x appearing in C. We set $m_t = m_{t-1} \cup \{(C, x)\}$. All properties are maintained: The new matching corresponds to playing one more round of the matching game. The memory size and the matching size are both incremented by 1. $\mathcal{C}_t|_{\rho(m_t)} = 1$, because $\mathcal{C}_{t-1}|_{\rho(m_{t-1})} = 1$, and $C|_{\rho(\{(C,x)\})} = 1$.
2. **Inference:** Set $m_t = m_{t-1}$. By the soundness of resolution, we know that $\mathcal{C}_{t-1}| = \mathcal{C}_t$ and hence $\mathcal{C}_t|_{\rho(m_t)} = \mathcal{C}_t|_{\rho(m_{t-1})} = 1$. Since $|\mathcal{C}_t| > |\mathcal{C}_{t-1}|$ it also follows that $|m_t| \leq |\mathcal{C}_t|$.

3. **Memory Erasure:** $\mathcal{C}_t = \mathcal{C}_{t-1} - C'$ for some clause C'. Any assignment satisfying $\mathcal{C}_{t-1}$ also satisfies $\mathcal{C}_t$. Hence m_{t-1} satisfies $\mathcal{C}_t$. We define m_t as the subrestriction of m_{t-1} obtained applying the Locality Lemma (Lemma 3.4) to m_t and $\mathcal{C}_t$. We then have that $|m_t| \leq |\mathcal{C}_t|$ and that m_t satisfies $\mathcal{C}_t$. Moreover, since $\rho(m)$ is clearly in 1–1 correspondence with m, for any m, then m_t is also a submatching of m_{t-1}, which implies the first property.

Thus we get that the clause space required to refute $\mathcal{C}$ is at least $MSpace(G(\mathcal{C}))$ and the theorem is proved. ∎

4. LOWER BOUND ON THE MATCHING GAME

We shall now prove lower bounds on the Matching Space when G is a good expander. In the proof we extensively use Hall's theorem:

Theorem 4.1 (Hall's Matching Theorem). *For a bipartite graph $G = \langle V \cup U, E\rangle$, there exists a perfect matching of $V' \subseteq V$ into U iff $\forall\ V'' \subseteq V'$, $|N(V'')| \geq |V''|$.*

We call V' *minimal unmatchable* into U' if V' is unmatchable into U', and any proper subset of V' is matchable into U'. By Hall's theorem, this occurs iff $|N(V')| < |V'|$ but for all $V'' \subset V' |N(V'')| \geq |V''|$.

Theorem 4.2 (Matching Space Lower Bound). *If a graph $G = \langle (V \cup U), E\rangle$ is an (r, c)-bipartite expander, then*

$$MSpace(G) > \frac{c \cdot r}{2 + c}$$

Proof. Suppose the game is played until, at time T, Pete wins. Thus, at time $0 \leq t < T$ the set of covered vertices corresponds to a partial matching in G. For any $0 \leq t < T$, let $E_t \subseteq E$ be the matching at time t, let $s_t = |E_t|$ be the matching space at time t, and let V_t (resp. U_t) be the set of uncovered vertices in V (resp. U). We define a strategy for Dana.

Dana's Strategy: Answer trying to maintain the property:

$$\forall\ V' \subseteq V_t, \quad |V'| \leq r - s_t, \qquad V' \text{ can be matched into } U_t. \tag{1}$$

At $t = 0$, $s_t = 0$, and the property holds, by the definition of expansion and Hall's theorem. Let t be the first time property (1) does not hold. We claim that $s_t \geq \frac{cr}{2+c}$. The proof of this claim is divided into two cases, according to the step taken at time t.

1. **Pete removes a finger from v:** Let u be the vertex matched to v at time $t - 1$ [i.e., $(v, u) \in E_{t-1}$, $V_t = V_{t-1} \cup \{v\}$, $U_t = U_{t-1} \cup \{u\}$, and $s_t = s_{t-1} - 1$.] There exists some $V' \subset V_t$ of size at most $r - s_t$ that is minimal unmatchable into U_t.

Claim 4.3. $|V'| = r - s_t$.

Proof. By definition, $|V'| \leq r - s_t$. We prove that $|V'| \geq r - s_t$. By the minimality of t, every set of size $(r - s_t) - 1 = r - s_{t-1}$ is matchable into U_t because it is even matchable into $U_{t-1} \subset U_t$. Therefore, $|U_t| \geq r - s_t - 1$. Since by definition V' is minimally unmatchable into U_t, then any proper subset V'' is matchable into U_t. Therefore, $|V''| \geq r - s_t - 1$. But since $|V'| \geq |V''| + 1$, hence $|V'| \geq r - s_t$. ∎

Let us calculate the size of the set of neighbors of V' in the original graph G. On the one hand, $|V'| \leq r$, and hence by the definition of expansion:

$$|N(V')| \geq (1 + c)|V'|. \tag{2}$$

On the other hand, V' is *minimal* unmatchable into U_t, and hence by Hall's matching theorem $|V'| > |N(V') \cap U_t|$. The only other possible neighbors of V' in G are in the matching E_t, which has size s_t. Thus we get

$$|V'| + s_t > |N(V')|. \tag{3}$$

Combining the two inequalities and setting $|V'| = r - s_t$ we get

$$|V'| + s_t > (1 + c)|V'| \Rightarrow \tag{4}$$

$$s_t > c \cdot |V'| = c \cdot (r - s_t) \Rightarrow \tag{5}$$

$$s_t > \frac{c \cdot r}{1 + c} > \frac{c \cdot r}{2 + c}. \tag{6}$$

Case 1 is proven.

2. **Pete places a finger on** v**:** Let $u_1, \ldots, u_d$ be the neighbors of v in U_{t-1} (for some $d > 0$). For any choice u_i that Dana makes, there is some $V^i \subset V_t$, $|V^i| \leq r - s_t$ that is minimally unmatchable into $U_{t-1} \backslash \{u_i\}$.

Claim 4.4. *There is no matching of* $\tilde{V} = \cup_{i=1}^{d} V^i \cup \{v\}$ *into* U_{t-1}.

Proof. Assume for the sake of contradiction that $\tilde{V}$ is matchable into U_{t-1}, and fix such a matching. Let $U^i \subset U_{t-1}$ be the image of V^i under this matching. Recalling that V^i is minimal unmatchable into $U_{t-1} \backslash \{u_i\}$, we conclude that $u_i \in U^i$. This is true for any $i = 1, \ldots, d$, and hence all neighbors of v are already taken by the matching on $\cup_{i=1}^{d} V^i$. Thus v cannot be matched, contradiction. ∎

By the claim, and the minimality of t, $|\tilde{V}| > r - s_{t-1}$, and since $s_t = s_{t-1} + 1$, we get $|\cup_{i=1}^{d} V^i| > r - s_t$.

Claim 4.5. *For* $i = 1, \ldots, d$, $|N(V^i) \cap U_{t-1}| = |V^i|$.

Proof. $|V^i| \leq r - s_t < r - s_{t-1}$, so V^i is matchable into U_{t-1}, and hence $|N(V^i) \cap U_{t-1}| \geq |V^i|$. V^i is minimal unmatchable into $U_{t-1}\backslash\{u_i\}$, so $|N(V^i) \cap U_{t-1}\backslash\{u_i\}| < |V^i|$, and hence $|N(V^i) \cap U_{t-1}| \leq |V^i|$. ∎

For all i, $|V^i| \leq r - s_t$, whereas, $|\cup_{i=1}^{d} V^i| > r - s_t$. There must exist some subset $I \subseteq [d]$ such that $\frac{r-s_t}{2} < |\cup_{i\in I} V_i| \leq r - s_t$. Fix such an I, and denote $V' = \cup_{i\in I} V^i$.

Claim 4.6. *For $G = \langle V \cup U, E\rangle$ a bipartite graph, let $V^1, \ldots, V^d$ be subsets of V such that for all $i \in [d]$*

(a) $|N(V^i)| = |V^i|$,
(b) V^i *is matchable into* U, *i.e., for each* $V'' \subseteq V^i$, $|N(V'')| \geq |V''|$.

Then $|N(\cup_{i\in[d]} V^i)| \leq |\cup_{i\in[d]} V^i|$.

Proof. By induction on d. For $d = 1$ the claim is simply the condition (a). Let $V^d_{NEW} = V^d\backslash(\cup_{i=1}^{d-1} V^i)$ be the set of "new" vertices added by V^d, and $V^d_{OLD} = V^d\backslash V^d_{NEW} \cdot V^d_{OLD}$ is a subset of V^d, so by property (b):

$$|N(V^d_{OLD})| \geq |V^d_{OLD}| \tag{7}$$

By property (a) we get

$$\begin{aligned} |V^d| = |N(V^d)| &= |N(V^d_{OLD})| + |N(V^d_{NEW})\backslash N(V^d_{OLD})| \\ &\geq |N(V^d_{OLD})| + |N(V^d_{NEW})\backslash N(\cup_{i=1}^{d-1} V^i)| \\ &\geq |V^d_{OLD}| + |N(V^d_{NEW})\backslash N(\cup_{i=1}^{d-1} V^i)|. \end{aligned}$$

The second inequality follows from the fact that $V^d_{OLD} \subseteq \cup_{i=1}^{d-1} V^i$ and the third from (7). Thus

$$|N(V^d_{NEW})\backslash N(\cup_{i=1}^{d-1} V^i)| \leq |V^d| - |V^d_{OLD}| = |V^d_{NEW}|. \tag{8}$$

We are ready to finish the proof:

$$|N(\cup_{i\in[d]} V^i)| = |N(\cup_{i=1}^{d-1} V^i)| + |(N(V^d_{NEW})\backslash N(\cup_{i=1}^{d-1} V^i))| \tag{9}$$

$$\leq |\cup_{i=1}^{d-1} V^i| + |V^d_{NEW}| \tag{10}$$

$$= |\cup_{i=1}^{d} V^i|. \tag{11}$$

The inequality (10) follows from the inductive hypothesis and (8). The claim is proved. ∎

Let us calculate once again the size of the set of neighbors of V' in G. Look at the bipartite graph G' induced by $V' = \cup_{i \in I} V^i$ and U_{t-1}. Each V^i is matchable into U_{t-1}, as shown in the proof of claim 4.5. Together with Claim 4.5 we conclude that conditions (a), (b) of Claim 4.6 apply to $\{V^i\}_{i \in I}$ in G'. By Claim 4.6 we get

$$|V'| \geq |N(V') \cap U_{t-1}|,$$

which, by definition of U_{t-1} and s_{t-1}, gives

$$s_{t-1} + |V'| \geq |N(V')|.$$

On the other hand, the expansion property and the fact $|V'| \leq r$ give us

$$|N(V')| \geq (1 + c)|V'|.$$

Combining the two together, we get

$$|V'| + s_{t-1} \geq (1 + c)|V'| \Rightarrow (s_t > s_t - 1),$$

$$s_t > c|V'| \Rightarrow \left(\left| V' \right| \geq \frac{r - s_t}{2} \right),$$

$$s_t > \frac{c \cdot r}{2 + c}.$$

Case 2 is proven, and with it the theorem. ∎

5. PROOF OF THE THEOREM 1.2

In this section we complete the proof of the main Theorem 1.2. The proof follows from the expansion properties of a random $G(\mathcal{F})$. The following lemma is essentially the same as in [3, 5], building on the original analysis of [6]. After completing the proof of Theorem 1.2, the rest of the section is devoted to the proof of the lemma, which we present here for the sake of completeness.

Lemma 5.1. *For any integer $k \geq 3$ and any constant $0 < c < k - 2$, there exists a constant $\kappa = \kappa(k, c)$ such that for any $\Delta \geq 1$ the following holds. For $\mathcal{F} \sim \mathbb{F}^{k,n}_{\Delta \cdot n}$, with high probability $G(\mathcal{F})$ is a $(\kappa \cdot n \cdot \Delta^{-(1+c/k-2-c)}, c)$-bipartite expander.*

Proof of Theorem 1.2. Notice that the function $f(c) \stackrel{\text{def}}{=} \frac{1+c}{k-2-c}$ is monotonically decreasing from ∞ to 1 as c ranges from 0 to $\frac{k-3}{2}$. Thus, for any $\epsilon > 0$ one can select $0 < c < \frac{k-3}{2}$ such that $f(c) = 1 + \epsilon$. Fix such a c. By Lemma 5.1, with high probability $\mathcal{F} \sim \mathbb{F}^{k,n}_{\Delta \cdot n}$ is an $(\Omega(n \cdot \Delta^{-f(c)}), c)$-bipartite expander; i.e., it is an $(\Omega(n/\Delta^{1+\epsilon}), c)$-bipartite expander. By Theorem 4.2 the matching game played on $G(\mathcal{F})$ requires space $(\Omega(n/\Delta^{1+\epsilon}))$, and finally by Theorem 3.3, the clause space is at least $(\Omega(n/\Delta^{1+\epsilon}))$. ∎

Proof of Lemma 5.1. Fix k and c, and let $r = \kappa \cdot n \cdot \Delta^{-(1+c/k-2-c)}$, where $\kappa = \kappa(k, c)$ will be determined later. Let BAD be the event that $G(\mathcal{F})$ is not an (r, c)-bipartite expander, where $G(\mathcal{F})$ is the random bipartite graph associated with $\mathcal{F} \sim \mathbb{F}_{\Delta \cdot n}^{k,n}$, according to Definition 3.2.

BAD occurs only if there exists a set $V' \subseteq V$, with $1 \leq |V'| \leq r$, such that $|N(V')| < (1 + c)|V'|$. There are $\binom{\Delta n}{i}$ possible sets $V' \subseteq V$ of size i, and there are $\binom{n}{(1+c)i}$ possible small sets of neighbors of V'. For a given set V' of size i, and a given set U' of size $(1 + c)i$, the probability that $N(V') \subseteq U'$ is

$$P_i = \left(\frac{\binom{(1+c)i}{k}}{\binom{n}{k}} \right)^i \leq \left(\frac{(1 + c)i}{n} \right)^{ki}.$$

Fix $d = d(k, c) \stackrel{\text{def}}{=} e^{2+c} \cdot (1 + c)^{k-1-c}$. Let us bound the probability of BAD, using the well-known estimation $\binom{a}{b} \leq \left(\frac{ea}{b}\right)^b$:

$$\begin{aligned}
\Pr[BAD] &\leq \sum_{i=1}^{r} \binom{\Delta n}{i} \cdot \binom{n}{(1 + c)i} \cdot P_i \\
&\leq \sum_{i=1}^{r} \left(\frac{e\Delta n}{i} \right)^i \cdot \left(\frac{en}{(1 + c)i} \right)^{(1+c)i} \cdot \left(\frac{(1 + c)i}{n} \right)^{ki} \\
&= \sum_{i=1}^{r} \left[d \cdot \Delta \cdot \left(\frac{i}{n} \right)^{(k-2-c)} \right]^i. \qquad (12)
\end{aligned}$$

Set $\kappa(k, c) = \left(\frac{1}{2d}\right)^{(1/k-2-c)}$, and notice that

$$\left(\frac{r}{n} \right)^{k-2-c} = \frac{1}{2d\Delta^{1+c}}. \qquad (13)$$

Set $s \stackrel{\text{def}}{=} n^{(1/4)(k-2-c)}$. We split the proof into cases:

Case 1: $\Delta \geq s$.

$$\begin{aligned}
\Pr[BAD] &\leq \sum_{i=1}^{r} \left[d \cdot \Delta \cdot \left(\frac{i}{n} \right)^{(k-2-c)} \right]^i \qquad (14) \\
&\leq \sum_{i=1}^{r} \left[d \cdot \Delta \cdot \left(\frac{r}{n} \right)^{(k-2-c)} \right]^i \qquad (15) \\
&\leq \sum_{i=1}^{r} \left[\frac{1}{2} \cdot \Delta^{-c} \right]^i. \qquad (16)
\end{aligned}$$

The last inequality follows from (13). For any $0 \leq a \leq 1/2$ and any integer $m \geq 1$,

we use the bound

$$\sum_{i=1}^{m} a^i \le 2a. \tag{17}$$

Recalling $\Delta \ge 1$ and $c > 0$, we get via (17)

$$\sum_{i=1}^{r} \left[\frac{1}{2} \cdot \Delta^{-c}\right]^i \le \Delta^{-c} \le n^{-(c/4)(k-2-c)}. \tag{18}$$

The last inequality follows because $\Delta \ge s \ge n^{(1/4)(k-2-c)}$.

Case 2: $\Delta < s$

We split the sum of Eq. (12) into two:

$$\Pr[BAD] \le \sum_{i=1}^{r} \left[d \cdot \Delta \cdot \left(\frac{i}{n}\right)^{(k-2-c)}\right]^i \tag{19}$$

$$\le \sum_{i=1}^{\sqrt{n}} \left[d \cdot \Delta \cdot \left(\frac{i}{n}\right)^{(k-2-c)}\right]^i \tag{20}$$

$$+ \sum_{i=\sqrt{n}}^{r} \left[d \cdot \Delta \cdot \left(\frac{i}{n}\right)^{(k-2-c)}\right]^i. \tag{21}$$

We bound the first sum by the geometric sum

$$\sum_{i=1}^{\sqrt{n}} \left[d \cdot \Delta \cdot \left(\frac{i}{n}\right)^{(k-2-c)}\right]^i < \sum_{i=1}^{\sqrt{n}} \left[d \cdot s \cdot \left(\frac{\sqrt{n}}{n}\right)^{(k-2-c)}\right]^i$$

$$= \sum_{i=1}^{\sqrt{n}} [d \cdot n^{(1/4)(k-2-c)} \cdot n^{-(1/2)(k-2-c)}]^i$$

$$= \sum_{i=1}^{\sqrt{n}} [d \cdot n^{-(1/4)(k-2-c)}]^i.$$

Recalling c, d are constants, and $k - 2 - c > 0$, for sufficiently large n we get via (17)

$$\sum_{i=1}^{\sqrt{n}} [d \cdot n^{-(1/4)(k-2-c)}]^i \le 2dn^{-(1/4)(k-2-c)}. \tag{22}$$

Now for the second sum:

$$\sum_{i=\sqrt{n}}^{r}\left[d\cdot\Delta\cdot\left(\frac{i}{n}\right)^{(k-2-c)}\right]^i \le \sum_{i=\sqrt{n}}^{r}\left[d\cdot\Delta\cdot\left(\frac{r}{n}\right)^{(k-2-c)}\right]^i$$

$$\le \sum_{i=\sqrt{n}}^{r}\left[\frac{d\cdot\Delta}{2\cdot d\cdot\Delta^{1+c}}\right]^i$$

$$\le \sum_{i=\sqrt{n}}^{r}\left(\frac{1}{2}\right)^i$$

$$\le n\cdot 2^{-\sqrt{n}}. \tag{23}$$

The second inequality uses (13), and the third one uses the assumptions $\Delta \ge 1$, $c > 0$. Collecting the upper bounds from (18), (22), (23) we get

$$\Pr[BAD] \le \max\{n^{-(c/4)(k-2-c)}, 2dn^{-(1/4)(k-2-c)} + n\cdot 2^{-\sqrt{n}}\} = o(1).$$

■

6. SPACE LOWER BOUNDS FOR *G-PHP*

An optimal space lower bound for refutations of the pigeonhole principle in Resolution was proved by [9], and even extended to the Polynomial Calculus by [2]. The Graph Pigeonhole Principle, *G-PHP*, was introduced by [5] as a generalization for which size lower bounds still apply. The idea is to restrict the number of holes that a pigeon may go to, according to some underlying graph G.

Definition 6.1 (*G-PHP*). *Let $G = \langle V \cup U, E\rangle$ be a bipartite graph, $|V| = m$, $|U| = n$. Assign each edge a distinct variable x_e. G-PHP is the conjunction of the following clauses*:

$$P_v \overset{\text{def}}{=} \bigvee_{v\in e} x_e \quad \text{for } v \in V,$$
$$H^u_{v,v'} \overset{\text{def}}{=} \bar{x}_e \vee \bar{x}_{e'} \quad \text{for } e = (v,u),\quad e' = (v',u),\quad v, v' \in V,\quad v \ne v',\quad u \in U.$$

We prove that clause space of refuting the *G-PHP* can be reduced to the matching game played on G, in a generalization of Theorem 3.3. This in turn implies a space lower bounds for *G-PHP*, whenever G is a bounded degree expander.

Denote by $\mathcal{H}$ the conjunction of $H^u_{v,v'}$, for all $e = (v,u)$, $e' = (v',u)$, $v, v' \in V$, $v \ne v'$, $u \in U$. For $\mathcal{C}$ a set of clauses over $Vars(G\text{-}PHP)$, and ρ a partial restriction that does not falsify $\mathcal{H}$, we say $\mathcal{C}|_\rho \equiv 1 (\mathrm{mod}\ \mathcal{H})$ if $\mathcal{H}| = \mathcal{C}|_\rho$, i.e., any assignment that satisfies $\mathcal{H}$, satisfies $\mathcal{C}|_\rho$ as well.

For $m = \{(v_{i_1}, u_{i_1}), \ldots, (v_{i_k}, u_{i_k})\}$ a partial matching in G of size k, we define its corresponding restriction to set 1's to all edges in m, 0's to all edges (v,u) such that u is in matched in m to some $v' \ne v$, and leave all other variables unassigned. Formally,

$$x_{(v,u)}|_{\rho(m)} = \begin{cases} 1, & (v,u) \in m \\ 0, & (v,u) \notin m \quad \text{and} \quad \exists\, v' \in V, (v',u) \in m, \\ \star, & \text{otherwise.} \end{cases}$$

As for Theorem 3.3, the main property in the reduction will be a Locality Lemma tailored for pigeon-hole matchings.

Lemma 6.2. *Let $\mathscr{C}$ be a set of clauses over Vars(G-PHP), satisfiable* (mod $\mathscr{H}$). *For all matchings m in G such that $\mathscr{C}|_{\rho(m)} \equiv 1$* (mod $\mathscr{H}$) *and for all clauses $C \in \mathscr{C}$, there exists an edge $e_C \in m$, such that $C|_{\rho(\{e_C\})} \equiv 1$* (mod $\mathscr{H}$)

Proof. Fix any $C \in \mathscr{C}$. Let m be a matching in G such that $\mathscr{C}|_{\rho(m)} \equiv 1$ (mod $\mathscr{H}$). We have that in particular $C|_{\rho(m)} \equiv 1$ (mod $\mathscr{H}$). Now look at C. If $\rho(m)$ makes C true fixing some positive literal x_e, then we fix $e_C = e$. Obviously $C|_\rho(\{e_C\}) \equiv 1$ (mod $\mathscr{H}$).

If, otherwise, $\rho(m)$ makes true C satisfying some negated literals $\bar{x}_e$ with $e = (v, u)$ for some $u \in U$ and $v \in V$, then, by definition of $\rho(m)$, there is a $v' \in V$ such that $e' = (v', u) \in m$. Fix $e_C = e'$. As before, this edge is good to prove our claim. ∎

Theorem 6.3. *For any graph G, CSpace(G-PHP) ≥ MSpace(G).*

Proof. Suppose by contradiction that Dana can win when the matching game over G is played with k fingers. Let $\mathscr{C}_0, \ldots, \mathscr{C}_\ell$ be a derivation from *G-PHP* (in the form of sequence of configurations) of space k. We construct inductively a sequence of matchings $m_0, \ldots, m_l$ that maintains the following properties for all $t = 0, \ldots, \ell$:

1. $|m_t| \leq |\mathscr{C}_t|$.
2. $\mathscr{C}_t|_{\rho(m_t)} \equiv 1$ (mod $\mathscr{H}$).

Notice that for a matching m, $\rho(m)$ does not falsify $\mathscr{H}$, and can be easily extended to an assignment that satisfies $\mathscr{H}$ (by setting all unassigned variables to 0). Thus, condition 2 implies that $\mathscr{C}_t$ is satisfiable for all $t = 0, \ldots, \ell$. This in turn gives a contradiction, since the last configuration must contain the empty clause.

For $t = 0$ we define m_0 the empty matching. For the induction step, we prove the claim according to the type of step taken:

1. **Axiom Download:** Suppose we download the axiom C. Observe that since $|\mathscr{C}| < k$, then by the first property and the fact we are downloading an axiom $|m_{t-1}| < k - 1$. If $C \in \mathscr{H}$, we set $m_t = m_{t-1}$, and it is easy to see that both conditions are maintained. If $C = P_v$, corresponding to Pete placing a finger on v in the graph G, then Dana can answers placing a finger on some $u \in N(v)$ (this is possible because $|m_{t-1}| < k$ and Dana wins when the game is played with k pebbles). We then set $m_t = m_{t-1} \cup \{(v, u)\}$, and, once again, it is easy to verify that both conditions are maintained.
2. **Inference:** By the soundness of resolution, $\mathscr{C}_{t-1} \models \mathscr{C}$, and hence $\mathscr{C}_t|_{\rho(m_{t-1})} \equiv 1$ (mod $\mathscr{H}$). Additionally, $|\mathscr{C}_t| = |\mathscr{C}_{t-1}| + 1$, so setting $m_t = m_{t-1}$ maintains both conditions.

3. **Memory Erasure:** Apply the previous Locality Lemma to $\rho(m_{t-1})$ and each clause of $\mathcal{C}_t$, we can build a matching of size at most $|\mathcal{C}_t|$ satisfying $\mathcal{C}_t \bmod \mathcal{H}$. ∎

Assembling together the previous theorem, Theorem 5.1, and the fact that there exist bounded degree $(\Omega(n), O(1))$-expanders $G = \langle V \cup U, E \rangle$, for $|V| = n + 1$ and $|U| = n$ (see, e.g., [5]) we obtain the following.

Theorem 6.4. *If $G = \langle V \cup U, E \rangle$, for $|V| = n + 1$ and $|U| = n$ is a bounded degree $(\Omega(n), O(1))$-expanders, then $CSpace(G\text{-}PHP) \geq \Omega(n)$.*

7. OPEN QUESTIONS

1. The *variable space* of a CNF formula $\mathcal{C}$ is $VSpace(\mathcal{C}) \stackrel{\text{def}}{=} \sum_{C \in \mathcal{C}} w(C)$, the variable space of a proof is the maximal variable space of a configuration in the proof, and the variable space of refuting a formula is the minimal variable space of a proof. For any $\mathcal{F}$ over n variables, $VSpace(\mathcal{F}) \leq n^2$, because $CSpace(\mathcal{F}) \leq n$. Alekhnovich et al. [2] proved $\Omega(n^2)$ lower bounds for a certain formula with initial width n. Can one find a 3-CNF for which $VSpace(\mathcal{F}) = \Omega(n^2)$? Is this true for a random 3-CNF with Δn clauses (constant Δ)? We believe the answer is positive.
2. What is the clause space complexity of refuting a random CNF formula in the Polynomial Calculus? We suspect that one should get essentially the same lower bounds as for resolution.
3. For many hard tautologies we get linear lower bounds on the *width* and on the *clause space.* This is true for Tseitin graph formulas, the Pigeonhole principle, and random formulas. Notice that width is also a space measure: It is the maximal space of a single clause in the proof. What is the relationship between the two measures? At least in one aspect width is "harder" than space. A width lower bound yields a size lower bound for treelike *and* general resolution, whereas a space lower bound yields a size lower bound only for treelike resolution. For this reason we conjecture that $CSpace(\mathcal{C}) \geq width(\mathcal{C})$. Is this true? Can one find a counterexample?
4. The following question was raised by Ron Lavi. One may view the graph matching game as an online problem. Let G be a fixed bipartite graph, with $|V| > |U|$. One receives "matching requests" online, and wishes to keep the set matched. The strategy we presented for Dana requires her to compute on each request the matching properties for an exponential number of subsets of V, and doing this in the trivial is inefficient. Can one find a polynomial time algorithm that would operate as well as Dana's strategy?

ACKNOWLEDGMENTS

We thank Avi Wigderson, Alexander A. Razborov, and Michael Alekhnovich for helpful discussions and Michele Zito for remarks about the satisfiability threshold. We thank the referees for helpful remarks that simplified the presentation, and for pointing out errors in a previous manuscript.

REFERENCES

[1] D. Achlioptas, Setting 2 variables at time yields a new lower bound for random 3-SAT, Proc 32th STOC, 2000, pp. 28–37.

[2] M. Alekhnovich, E. Ben-Sasson, A. A. Razborov, and A. Wigderson, Space complexity in propositional calculus, Proc 32nd STOC, 2000, pp. 358–367.

[3] P. Beame, R. Karp, T. Pitassi, and M. Saks, The efficiency of resolution and Davis Putnam procedures, Proc 30th Annu ACM Symp Theory Comput (STOC-98), 23–26 May 1998, ACM Press, New York, pp. 561–571.

[4] P. Beame and T. Pitassi, Simplified and improved resolution lower bound, FOCS'96, 1996, pp. 274–282.

[5] E. Ben-Sasson and A. Wigderson, Short proofs are narrow—resolution made simple, Proc 31st STOC, 1999, p. 517–526.

[6] V. Chvátal and E. Szemerédi, Many hard examples for resolutions, J ACM 35 (1988), 759–768.

[7] S. A. Cook and R. Reckhow, The relative efficiency of propositional proof systems, J Symbolic Logic 44 (1979), 36–50.

[8] O. Dubois, Y. Boufkhad, and J. Mandler, Typical random 3-SAT formulae and the satisfiability problem, 11-th SODA, 2000, pp. 126–127.

[9] J. L. Esteban and J. Toran, Space bounds for Resolution, Proc 16th STACS, 1999, pp. 530–539.

[10] A. Haken, The intractability of resolution. Theoret Comput Sci 35 (1985), 297–308.

[11] S. Janson, Y. C. Stamatiou, and M. Vamvakari, Bounding the unsatisfiability threshold for 3-SAT, Random Structure Algorithms 17(2) (2000), 118–116.

[12] J. Torán, Lower bounds for space in resolution, Proc CSL 1999, 1999, pp. 362–373.

[13] A. Urquhart, Hard examples for resolution, J ACM 34 (1987), 209–219.

[14] M. Zito, An upper bound on the space complexity of random formulae in resolution, Electron Coll Comput Complexity, Report TR01-079, 2001, available at www.eccc.uni-trier.de/eccc/.