Electronic Colloquium on Computational Complexity, Report No. 33 (2015)

An Ultimate Trade-Off in Propositional Proof Complexity

Alexander Razborov[*]

March 6, 2015

Abstract

We exhibit an unusually strong trade-off between resolution proof width and tree-like proof size. Namely, we show that for any parameter $k = k(n)$ there are unsatisfiable k-CNFs that possess refutations of width $O(k)$, but such that any tree-like refutation of width $n^{1-\epsilon}/k$ must necessarily have *double* exponential size $\exp(n^{\Omega(k)})$. Conceptually, this means that there exist contradictions that allow narrow refutations, but in order to keep the size of such a refutation even within a single exponent, it must necessarily use a high degree of parallelism. Viewed differently, every tree-like narrow refutation is exponentially worse not only than wide refutations of the same contradiction, but of any other contradiction with the same number of variables. This seems to significantly deviate from the established pattern of most, if not all, trade-off results in complexity theory.

Our construction and proof methods combine, in a non-trivial way, two previously known techniques: the hardness escalation method based on substitution formulas and expansion. This combination results in a *hardness compression* approach that strives to preserve hardness of a contradiction while significantly decreasing the number of its variables.

[*] University of Chicago, razborov@math.uchicago.edu; part of this work was done at Steklov Mathematical Institute, Moscow, and at Toyota Technological Institute, Chicago. Supported by the Russian Science Foundation, grant # 14-50-00005.

ISSN 1433-8092

1. Introduction

Propositional proof complexity is an area of study that has seen a rapid development over several last decades, in part due to being well-connected to a number of other disciplines. One of these connections that has seen a particularly steady growth in very recent years is the interplay between propositional proof complexity and practical SAT solving. This interplay is based on the somewhat philosophical observation that the execution log of a solver's run on an unsatisfiable SAT instance can be viewed as a mathematical proof of its unsatisfiability in a certain proof system. Equally important is the accompanying empirical observation that proof systems arising in this way tend to be very natural, and, moreover, precisely of the kind that have been theoretically studied in propositional proof complexity since its inception in the seminal paper by Cook and Reckhoff [CR79]. As a matter of fact, the SAT solvers that seem to completely dominate the landscape at the moment, like those based on conflict-driven clause learning, lead to just one *resolution* proof system dating back to the papers by Blake [Bla37] and [Rob65]. This somewhat explains the fact that resolution is by far the most studied system in proof complexity, and much of this study has concentrated on simple complexity measures for resolution proofs like size, width and space, and on relations existing between them. All these measures have natural counterparts in the world of practical SAT solving.

Our paper continues this line of research, and we exclusively deal with the resolution proof system. The first measure of interest to us is *width.* This measure is extremely natural and robust, and in fact it is not very specific to resolution. As is well-known, width w proofs can more instructively be viewed as semantic proofs operating with arbitrary Boolean expressions and equally arbitrary (sound) inference rules with the sole restriction that every line depends on at most w variables.

Ben-Sasson and Wigderson [BW01] showed that short proofs can be transformed into proofs of small width, while Atserias and Dalmau [AD08] did this for proofs that have small clause space. Thus, despite its deluding simplicity, the class of contradictions possessing small-width refutations is rich.

In this paper we are interested in (yet another) confirmation of this thesis "from the opposite side": there exist contradictions that do have small-width refutations, but the latter are highly complex and non-efficient, and any attempts to simplify them must necessarily lead to a dramatic blow-up in width. Before reviewing previous work in this direction and stating our own contribution, it will be convenient to fix some basic notation (we will remind exact definitions in Section 2). For a CNF contradiction τ_n in n variables, let $w(\tau_n \vdash 0)$ $[S(\tau_n \vdash 0)]$ be the minimum possible width [size, respectively] of any resolution refutation of τ_n. $S_T(\tau_n \vdash 0)$ is the minimum size with respect to tree-like refutations, and $w(\tau_n)$ is the maximum width of a clause in τ_n itself.

In this notation, the main results from [BW01] can be stated as follows:

$$\begin{aligned} w(\tau_n \vdash 0) &\leq O(\log S_T(\tau_n \vdash 0)) \\ w(\tau_n \vdash 0) &\leq O(n \cdot \log S(\tau_n \vdash 0))^{1/2} + w(\tau_n). \end{aligned}$$

We also note the trivial (brute-search) bound in the opposite direction:

$$S(\tau_n \vdash 0) \leq n^{O(w(\tau_n \vdash 0))}. \tag{1}$$

As it turned out, little else can be said in general about relations between these basic measures. Namely, Ben-Sasson, Impagliazzo and Wigderson [BIW04] gave an example of contradictions τ_n with

$$w(\tau_n \vdash 0) \leq O(1), \quad S_T(\tau_n \vdash 0) \geq \exp\left(\Omega(n/\log n)\right); \tag{2}$$

the same paper also proved a matching upper bound (for arbitrary refutations in n variables). In more recent development, Atserias, Lauria and Nordström [ALN14] have shown that $S(\tau_n \vdash 0)$ can be as large as $n^{\Omega(w(\tau_n \vdash 0))}$, that is in certain situations the trivial brute-search proof is essentially optimal. Ben-Sasson [Ben09] established a trade-off between width and tree-like resolution size. Namely, he constructed contradictions τ_n that have tree-like refutations of *either* constant width *or* polynomial size but such that

$$w(\Pi) \cdot \log |\Pi| \geq \Omega(n/\log n) \tag{3}$$

for any tree-like refutation Π of τ_n, $w(\Pi)$ and $|\Pi|$ being its width and size respectively.

Our main result, Theorem 2.4 can be viewed as a far-reaching generalization of the previous contributions (2), (3). For any parameter $k = k(n)$ we construct a sequence of k-CNF contradictions τ_n such that $w(\tau_n \vdash 0) \leq O(k)$ while

$$|\Pi| \geq \exp\left(n^{\Omega(k)}\right) \tag{4}$$

for any tree-like refutation Π of width $\leq n^{1-\epsilon}/k$. Thus, when k, say, is a sufficiently large constant our bound becomes *super-exponential* in n, and for (say) $k = n^{1/3}$ it becomes *double exponential.* This bound is much larger than $S_T(\tau_n^* \vdash 0)$ for *any* contradiction τ_n^* in n variables, and for that reason we did not attempt to prove an exponential lower bound on $S_T(\tau_n \vdash 0)$, although we suspect it should not be very hard.

We propose the name "ultimate trade-offs" for this kind of results, although we are not immediately aware of any other example existing in the literature, be it in computational, proof or any other complexity.

On less general level, our result is complementary to that of Atserias et al. [ALN14]. Namely, they proved that the obvious brute-search refutation of size $n^{O(w)}$ (cf. (1)) in general can not be shortened. What we prove is that if we additionally want to keep the width reasonably small, and keep the tree-like size sane (at most single exponential), we need a high degree of parallelism.

Our construction and the proof combine two very popular techniques in proof complexity: hardness escalation and expansion. The former method converts every contradiction τ_n into another contradiction $\widehat{\tau}_n$ so that relatively mild hardness properties of τ_n transfer to lower bounds for $\widehat{\tau}_n$ in stronger proof systems. So far this technique has been used in two main flavors: *substitution formulas* (see e.g. the survey [Nor13, Section 2.4]) and more recent *lifting formulas* introduced by Beame, Huynh and Pitassi [BHP10]. One common feature of both approaches is that the price one has to pay for improving hardness is a moderate *increase* in the number of variables.

In our work we change the gears on both these counts and are interested in hardness *preservation*[1] and *variable compression*, that is in (exponentially) decreasing the number of variables. These two conflicting goals are balanced using linear substitutions whose support sets need not necessarily be disjoint as long as they form a good (boundary) expander. While by now expanders is one of the most common techniques in proof complexity, we are not aware of its previous applications in a similar context.

[1] As a matter of fact, our construction also gives the same hardness amplification as ordinary substitution formulas with disjoint sets of variables. This observation, however, plays no role in our conclusions.

2. Preliminaries

In this section we give necessary definitions, state some useful facts and formulate, in Section 2.1, our main results.

A *literal* is either a Boolean variable x or its negation $\bar{x}$; we will sometimes use the uniform notation $x^\epsilon \overset{\text{def}}{=} \begin{cases} x & \text{if } \epsilon = 1 \\ \bar{x} & \text{if } \epsilon = 0. \end{cases}$. A *clause* C is either a a disjunction of literals or 1. The latter is a convenient technicality (e.g. with this convention the set of all clauses makes a lattice in which $\vee$ is the join operator etc.); 1 should be thought of as a placeholder for all trivial clauses. C is a *sub-clause of* D, also denoted by $C \leq D$ if either $D = 1$ or $C, D \neq 1$ and every literal appearing in C also appears in D. The empty clause C will be denoted by 0. The set of variables occurring in a clause C (either positively or negatively) will be denoted by $Vars(C)$ ($Vars(1) \overset{\text{def}}{=} \emptyset$). The *width* of a clause is defined as $w(C) \overset{\text{def}}{=} |Vars(C)|$.

A *CNF* τ is a conjunction of clauses, often identified with the set of clauses it is comprised of. A CNF is a k*-CNF* if all clauses in it have width at most k. Unsatisfiable CNFs are traditionally called *contradictions*. For CNFs τ, τ', $\tau \models \tau'$ is the *semantical implication* meaning that every truth assignment satisfying τ also satisfies τ'. Thus, τ is a contradiction if and only if $\tau \models 0$. Also, for clauses C and D, $C \leq D$ if and only if $C \models D$. The subscript n in τ_n always stands for the number of variables in the CNF τ_n.

The *resolution proof system* operates with clauses and it consists of the only *resolution rule*

$$\frac{C \vee x \qquad D \vee \bar{x}}{C \vee D}. \tag{5}$$

A *tree-like*[2] *resolution proof* Π is a binary rooted tree in which all nodes all labelled by clauses, and such that the clause assigned to every internal node can be deduced from clauses sitting at its two children via a single application of the resolution rule. A *tree-like resolution proof of a clause* C *from a CNF* τ is a tree-like resolution proof Π in which all leaves are labelled by clauses from τ, and the root is labelled by a clause $\widetilde{C}$ such that $\widetilde{C} \leq C$ (the latter technicality is necessary since we did not include the weakening rule). A *refutation* of a contradiction is a proof of 0 from it. The *depth* $D(\Pi)$ of a proof Π is the height (the number of edges in the longest path) of its

[2] DAG-like proofs are not considered in this paper.

underlying tree, and its *size* $|\Pi|$ is the number of leaves. The *width* $w(\Pi)$ is the maximum width of a clause appearing in Π.

For a CNF τ and a clause C, we let $D(\tau \vdash C)$, $S_T(\tau \vdash C)$ and $w(\tau \vdash C)$ denote the minimum possible value of $D(\Pi)$, $|\Pi|$ and $w(\Pi)$, respectively, taken over all tree-like resolution proofs of C from τ.

The following result will be one of the starting points for our construction.

Proposition 2.1 ([BIW04]) *There exists an increasing sequence $\{\tau_n\}$ of 4-CNF refutations such that $w(\tau_n \vdash 0) \leq 6$, but $S_T(\tau_n \vdash 0) \geq \exp(\Omega(n/\log n))$.*

Let A be a $m \times n$ 0-1 matrix. For $i \in [m]$,[3] let $J_i(A) \overset{\text{def}}{=} \{j \in [n] \mid a_{ij} = 1\}$. For a clause E in the variables $\{y_1, \ldots, y_m\}$, by $E[A]$ we will denote the CNF obtained from E by the $\mathbb{F}_2$-linear substitution $y_i \to \bigoplus_{j \in J_i(A)} x_j$ $(i \in [m])$ followed by expanding the resulted Boolean function as a CNF in the straightforward way. The following easy observation will be important in what follows: for every clause C in $E[A]$,

$$Vars(C) = \bigcup_{y_i \in Vars(E)} \{x_j \mid j \in J_i(A)\} \tag{6}$$

(we do claim equality here). For a CNF $\tau = E_1 \wedge E_2 \wedge \ldots \wedge E_\ell$, we let $\tau[A] \overset{\text{def}}{=} E_1[A] \wedge \ldots \wedge E_\ell[A]$. If τ is a contradiction then evidently $\tau[A]$ is a contradiction, too. The converse need not be true in general, of course.

For $I \subseteq [m]$, the *boundary* of this set of rows is defined as

$$\partial_A(I) \overset{\text{def}}{=} \{j \in [n] \mid |\{i \in I \mid j \in J_i(A)\}| = 1\},$$

i.e., it is the set of columns that have *precisely* one 1 at their intersections with I. A is an (r,s,c)*-boundary expander*[4] if $|J_i(A)| \leq s$ for any $i \in [m]$ and $|\partial_A(I)| \geq c|I|$ for every set of rows $I \subseteq [m]$ with $|I| \leq r$. An (r,n,c)-boundary expander (i.e., a $m \times n$ matrix satisfying only the second of these conditions) will be simply called an (r,c)*-expander.*

[3] $[m] \overset{\text{def}}{=} \{1, \ldots, m\}$.

[4] In [ABRW04] such matrices were simply called expanders. But since beginning with (apparently) [AAT11] the research in proof complexity also made good use of ordinary vertex expanders and, as a consequence, tended to differentiate between boundary and ordinary expansion, we also adopt this terminological change. What we, however, keep in this paper is the matrix notation as we find it more instructive for many reasons.

For a set of columns $J \subseteq [n]$, we let

$$\mathrm{Ker}(J) \stackrel{\mathrm{def}}{=} \{i \in [m] \mid J_i(A) \subseteq J\}$$

be the set of rows completely contained in J. Let $A \setminus J$ be the sub-matrix of A obtained by removing all columns in J and all rows in $\mathrm{Ker}(J)$.

We need two properties of boundary expanders whose analogues were used, in one or another form, in almost all their applications in proof complexity. The first one, proven by a simple probabilistic argument, says that good expanders exist.

Lemma 2.2 *Let $n \to \infty$ and m, s, c be arbitrary integer parameters possibly depending on n such that $c \leq \frac{3}{4}s$ and*

$$r \leq o(n/s) \cdot m^{-\frac{2}{s-c}}. \tag{7}$$

Then for sufficiently large n there exist $m \times n$ (r, s, c)-boundary expanders.

The second property says that in every good expander, the class of small sets of rows whose removal leads to a relatively good expander is in a sense everywhere dense.

Lemma 2.3 *Let A be an $m \times n$ $(r, 2)$-boundary expander. Then for every $J \subseteq [n]$ with $|J| \leq r/4$ there exists $\widehat{J} \supseteq J$ such that $\left|\mathrm{Ker}\left(\widehat{J}\right)\right| \leq 2|J|$ and $A \setminus \widehat{J}$ is an $(r/2, 3/2)$-boundary expander.*

We, however, have not been able to recover these statements from the literature in a referrable form, and for this reason their simple proofs are included in the Appendix.

2.1. Main results

In this brief section we formulate our main results.

Theorem 2.4 *Let $k = k(n) \geq 4$ be any parameter, and let $\epsilon > 0$ be an arbitrary constant. Then there exists a sequence of k-CNF contradictions $\{\tau_n\}$ in n variables such that $w(\tau_n \vdash 0) \leq O(k)$ but for any tree-like refutation Π with $w(\Pi) \leq n^{1-\epsilon}/k$ we have the bound*

$$|\Pi| \geq \exp\left(n^{\Omega(k)}\right).$$

As we noted in Introduction, our main technique is hardness preservation, and since the corresponding statement might be of independent interest, we formulate it here as a separate result.

Theorem 2.5 *Let τ_m be an atbitrary contradiction in the variables $y_1, \ldots, y_m$, and let A be an $m \times n$ $(r, 2)$-boundary expander for some r. Then every tree-like refutation Π of $\tau_m[A]$ with $w(\Pi) \leq r/4$ must satisfy*

$$|\Pi| \geq 2^{2D(\tau_m \vdash 0)/r}.$$

As stated, this is also a hardness escalation result (from depth to tree-like size), but that part alone was known before [Urq11, Theorem 5.4], and one does not need expanders for that.

3. Proofs

In this section we prove Theorems 2.4 and 2.5, and we begin with the latter. We present our proof as a plain inductive argument since, in our view, it is often more instructive than various top-down approaches (cf. the recent remarkable simplification [FLM+14] of the Atserias-Dalmau bound that resulted from adopting this point of view).

Fix an $m \times n$ $(r, 2)$-boundary expander A, where r is an arbitrary parameter. Let us say that a set of columns J is *closed* if $A \setminus J$ is an $(r/2, 3/2)$-boundary expander (cf. Lemma 2.3). Fix now an arbitrary CNF τ_m (that need not necessarily be a contradiction) in the variables $y_1, \ldots, y_m$. We are going to prove the following.

Claim 3.1 *Assume that C is a clause in the variables $x_1, \ldots, x_n$ that possesses a tree-like proof Π from $\tau[A]$ with $w(\Pi) \leq r/4$. Let $J \subseteq [n]$ be an arbitrary closed set with $J \supseteq \{x_j \mid j \in Vars(C)\}$, and let E be any clause in y-variables with*

$$Vars(E) = \{y_i \mid i \in \mathrm{Ker}(J)\}$$

such that

$$E[A] \vee C \not\equiv 1, \tag{8}$$

that is there exists an assignment of x-variables simultaneously falsifying $E[A]$ and C. Then

$$D(\tau \vdash E) \leq \frac{r}{2} \cdot \log_2 |\Pi|.$$

Proof of Claim 3.1. Let C, Π, J and E satisfy the assumptions of our claim. The argument proceeds by induction on Π.

Base $|\Pi| = 1$, i.e. C contains a sub-clause $\widetilde{C}$ that appears in $\widetilde{E}[A]$ for some $\widetilde{E} \in \tau$.
Applying (6) to the clause $\widetilde{E}$, we see that $J \supseteq Vars(C) \supseteq Vars(\widetilde{C})$ implies $\left\{ i \in [m] \,\middle|\, y_i \in Vars(\widetilde{E}) \right\} \subseteq \mathrm{Ker}(J)$, that is $Vars(\widetilde{E}) \subseteq Vars(E)$. Also, E and $\widetilde{E}$ must be consistent since

$$(E \vee \widetilde{E})[A] = E[A] \vee \widetilde{E}[A] \models E[A] \vee \widetilde{C} \models E[A] \vee C,$$

and hence their inconsistency would have implied that $(E \vee \widetilde{E})[A] \equiv E[A] \vee C \equiv 1$ in contradiction with (8). Hence $\widetilde{E} \leq E$ and thus $D(\tau \vdash E) = 0$.

Inductive step $|\Pi| > 1$.
Assume that the last application of the resolution rule has the form

$$\frac{C_0 \vee x_j \qquad C_1 \vee \bar{x}_j}{C_0 \vee C_1}.$$

Fix arbitrarily an assignment α to $\{x_j \,|\, j \in J\}$ falsifying both $E[A]$ and $C_0 \vee C_1$ that exists by our assumption. Further analysis depends on whether $j \in J$ or not.

Case 1, $j \in J$.
This case is easy. Assume w.l.o.g. that $\alpha(x_j) = 0$. Note that $Vars(C_0 \vee x_j) \subseteq J$ and $\alpha(C_0 \vee x_j) = 0$. Thus we can apply the inductive assumption to the clause $C_0 \vee x_j$, the corresponding sub-proof Π_0 of Π and to the same J and E. We conclude that $D(\tau \vdash E) \leq r \cdot \log_2 |\Pi_0| \leq r \cdot \log_2 |\Pi|$.

Case 2, $j \notin J$.
One of the two sub-trees Π_0, Π_1 (say, Π_0) determined by the children of the root has size $\leq |\Pi|/2$, and we assume w.l.o.g. that it corresponds to the child labeled by $C_0 \vee x_j$. Since $w(C_0 \vee x_j) \leq r/4$ by our assumption, we can apply Lemma 2.3 to the set $J' \stackrel{\text{def}}{=} \{j' \,|\, x_{j'} \in Vars(C_0 \vee x_j)\}$. This will give us a closed $\widehat{J} \supseteq \{j' \,|\, x_{j'} \in Vars(C_0 \vee x_j)\}$ with $\left|\mathrm{Ker}\left(\widehat{J}\right)\right| \leq r/2$, and our first goal is to prove that *every* clause $\widehat{E}$ with $Vars\left(\widehat{E}\right) = \left\{ y_i \,\middle|\, i \in \mathrm{Ker}\left(\widehat{J}\right) \right\}$ *and consistent with* E satisfies the assumptions of Claim 3.1 with $C := C_0 \vee x_j$, $J := \widehat{J}$ and $E := \widehat{E}$. For that we only have to extend our original assignment α to the variables $\left\{ x_j \,\middle|\, j \in J \cup \widehat{J} \right\}$ in such a way that it will falsify

both $\widehat{E}[A]$ and $C_0 \vee x_j$. Since $C_0 \leq C_0 \vee C_1 \leq C$ is already falsified by α, the latter task can be achieved simply by setting additionally $\alpha(x_j) \stackrel{\mathrm{def}}{=} 0$ (recall that $j \notin J$). Also, every literal y_i^ϵ of a variable $y_i \in Vars(E) \cap Vars\left(\widehat{E}\right)$ maps to $y_i^\epsilon[A] = \bigoplus_{j \in J_i(A)} \oplus \bar{\epsilon}$ and, since $J_i(A) \subseteq J$, it has been already decided by α. As E and $\widehat{E}$ are consistent by our assumption (and α falsifies E), $\alpha(y_i[A])$ is actually $\bar{\epsilon}$. It only remains to show that α' can be extended in such a way that it sets all $y_i[A]$ for $i \in \mathrm{Ker}\left(\widehat{J}\right) \setminus \mathrm{Ker}(J)$ to fixed values predetermined to falsify the formula $\widehat{E}[A]$.

Let A' be the matrix obtained from $A \setminus J$ by additionally removing the column j from it. Since $A \setminus J$ is an $(r/2, 3/2)$-boundary expander, A' is an $(r/2, 1/2)$-boundary expander. Also, $\mathrm{Ker}\left(\widehat{J}\right) \setminus \mathrm{Ker}(J)$ is a set of rows of cardinality $\leq r/2$, therefore $\partial_{A'}(I) \neq \emptyset$ for every non-empty subset $I \subseteq \mathrm{Ker}\left(\widehat{J}\right) \setminus \mathrm{Ker}(J)$. This allows us to order, by reverse induction, the rows in $\mathrm{Ker}\left(\widehat{J}\right) \setminus \mathrm{Ker}(J)$ in such a way $\mathrm{Ker}\left(\widehat{J}\right) \setminus \mathrm{Ker}(J) = \{i_1, \ldots, i_\ell\}$ that for every $\nu \in [\ell]$ the set of rows $J_{i_\nu}(A') \setminus \bigcup_{\mu=1}^{\nu-1} J_{i_\mu}(A')$ is not empty; fix arbitrarily $j_\nu \in J_{i_\nu}(A') \setminus \bigcup_{\mu=1}^{\nu-1} J_{i_\mu}(A')$. Now, we first extend α' to $\{x_j \mid j \in (J \cup J') \setminus \{j_1, j_2, \ldots, j_\ell\}\}$ arbitrarily (say, by zeros) and then consecutively extend it to $x_{j_1}, \ldots, x_{j_\ell}$ so that the linear forms $\bigoplus_{j \in J_1(A)} x_j, \ldots, \bigoplus_{j \in J_\ell(A)} x_j$ are set to the right values.

In conclusion, $\widehat{E}$ satisfies assumptions of Claim 3.1 with $C := C_0 \vee x_j$. Since this clause has a proof from $\tau[A]$ of width $\leq r/4$ and size $\leq |\Pi|/2$, $D(\tau \vdash E) \leq \frac{r}{2}(\log_2 |\Pi| - 1)$. This conclusion holds for an arbitrary clause $\widehat{E}$ in the variables $\left\{ y_i \,\middle|\, i \in \mathrm{Ker}\left(\widehat{J}\right) \right\}$ consistent with E. Now we resolve all these clauses in the brute-force way along all the variables $\left\{ y_i \,\middle|\, i \in \mathrm{Ker}\left(\widehat{J}\right) \setminus \mathrm{Ker}(J) \right\}$. Since the depth of this original proof is at most $r/2$, we get a proof of E in depth $\frac{r}{2} \log_2 |\Pi|$.

This completes the analysis in case 2 of the inductive step. Claim 3.1 is proved.■

Theorem 2.5 is now immediate. If τ is a contradiction and Π is a refutation of $\tau[A]$ with $w(\Pi) \leq r/4$, we simply apply Claim 3.1 with $C := 0$, $J := \emptyset$ and $E := 0$.

For Theorem 2.4, we are simply going to apply Theorem 2.5 to $\tau[A]$, where τ is the contradiction from Proposition 2.1 and A is the (random) matrix guaranteed by Lemma 2.2.

Proof of Theorem 2.4.

First of all we can assume that $k \geq 12$ since otherwise already the contradictions from Proposition 2.1 will do. Set $w := n^{1-\epsilon}/k$, $r := 4w$, $s := \lfloor k/4 \rfloor \geq 3$, $c := 2$ and choose the parameter m as the smallest value for which (7) is satisfied. Clearly, $m \geq (n/kw)^{\Omega(k)} \geq n^{\Omega(k)}$. If it turns out that $m \leq n^2$ then, as before, we simply take the 4-CNF contradiction τ_n provided by Proposition 2.1. Otherwise we take the formula τ_m provided by that proposition and compose it with an $m \times n$ $(r, s, 2)$-expander A guaranteed by Lemma 2.2.

Recall that $D(\tau_m \vdash 0) \geq \Omega(m/\log m)$. Hence Theorem 2.5 implies that every tree-like refutation Π of the k-CNF contradiction $\tau_m[A]$ with $w(\Pi) \leq w$ must have size at least

$$|\Pi| \geq \exp\left(\Omega\left(\frac{m}{r \log m}\right)\right) \geq \exp\left(\Omega\left(\frac{m}{n \log m}\right)\right) \overset{\text{since } m \geq n^2}{\geq} \exp(m^{\Omega(1)}) \geq \exp(n^{\Omega(k)}).$$

It only remains to remark that the width 6 refutation of τ_m stipulated by Proposition 2.1 can be converted into a width $O(k)$ refutation of $\tau_m[A]$ simply by applying the operator $E \mapsto E[A]$ to its lines.

4. Conclusion

This paper has exhibited a somewhat peculiar phenomenon. When we try to optimize in one resource (width in our case), the price we have to pay in another resource (size) increases exponentially with respect to straightforward unrestricted tree-like refutations for *all* contradictions of the same size, not only of the one we started with. This is very different from traditional trade-off results.

Although in the paragraph above we used the words like "proofs" and "contradictions", there is nothing in this paradigm that would be specific to proof complexity. Thus, the first question we would like to ask is this: do ultimate trade-offs exist elsewhere, or our example is singular? It looks like one natural place to look for ultimate trade-offs (given the abundance of traditional ones) is propositional space complexity.

Our result is a bit incomplete since the lower bound is double exponential only in the number of variables, not in the size of the contradiction. Does it necessarily have to be the case? Attempting a rigorous formulation (there are many other options to pinpoint this question), do there exist $O(1)$-contradictions τ_n for which any refutation of optimal, or, even better, nearly optimal, width requires tree-like size $\exp(n^{\omega(1)})$?

References

[AAT11] M. Alekhnovich, S. Arora, and I. Tourlakis. Toward strong non-approximability results in the Lovász-Schrijver hierarchy. *Computational Complexity*, (4):615–648, 2011.

[ABRW04] M. Alekhnovich, E. Ben-Sasson, A. Razborov, and A. Wigderson. Pseudorandom generators in propositional proof complexity. *SIAM Journal on Computing*, 34(1):67–88, 2004.

[AD08] A. Atserias and V. Dalmau. A combinatorial characterization of resolution width. *Journal of Computer and System Sciences*, 74(3):323–334, 2008.

[ALN14] A. Atserias, M. Lauria, and J. Nordström. Narrow proofs may be maximally long. In *Proceedings of the 29th IEEE Conference on Computational Complexity*, pages 286–297, 2014.

[Ben09] E. Ben-Sasson. Size-space tradeoffs for resolution. *SIAM Journal on Computing*, 38(6):2511–2525, 2009.

[BHP10] P. Beame, T. Huynh, and T. Pitassi. Hardness amplification in proof complexity. In *Proceedings of the 42nd Annual ACM Symposium on Theory of Computing,*, pages 87–96, 2010.

[BIW04] E. Ben-Sasson, R. Impagliazzo, and A. Wigderson. Near optimal separation of tree-like and general resolution. *Combinatorica*, 24(4):585–603, 2004.

[Bla37] A. Blake. *Canonical expressions in Boolean algebra*. PhD thesis, University of Chicago, 1937.

[BW01] E. Ben-Sasson and A. Wigderson. Short proofs are narrow - resolution made simple. *Journal of the ACM*, 48(2):149–169, 2001.

[CR79] S. A. Cook and A. R. Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 44(1):36–50, 1979.

[FLM$^+$14] Y. Filmus, M. Lauria, M. Mikša, J. Nordström, and M. Vinyals. From small space to small width in resolution. Technical Report TR14-081, Electronic Colloquium on Computational Complexity, 2014.

[Nor13] J. Nordström. Pebble games, proof complexity and time-space trade-offs. *Logical Methods in Computer Science*, 9:1–63, 2013.

[Rob65] J. A. Robinson. A machine-oriented logic based on the resolution principle. *Journal of the ACM*, 12(1):23–41, 1965.

[Urq11] A. Urquhart. The depth of resolution proofs. *Studia Logica*, 99:349–364, 2011.

A. Appendix

Here we give self-contained proofs of Lemmas 2.2 and 2.3.

Lemma 2.2. *Let $n \to \infty$ and m, s, c be arbitrary integer parameters possibly depending on n such that $c \leq \frac{3}{4}s$ and*

$$r \leq o(n/s) \cdot m^{-\frac{2}{s-c}}.$$

Then for sufficiently large n there exist $m \times n$ (r, s, c)-boundary expanders.

Proof. This lemma and its proof is identical to [ABRW04, Theorem 5.1], except that we relax some restrictions on the parameters. We construct a random $m \times n$ matrix $\boldsymbol{A}$ by picking independently in each row s random entries with repetitions (the latter feature is not crucial, but it does make calculations neater). That is, we let $J_i(\boldsymbol{A}) \overset{\text{def}}{=} \{\boldsymbol{j_{i1}}, \ldots, \boldsymbol{j_{is}}\}$, where $\{\boldsymbol{j_{i\nu}}\}$ $(i \in [m], \nu \in [s])$ is a collection of ms independent random $[n]$-valued variables.

Recall that a matrix A is an (ordinary) (r, s, c)-*expander* if, again, $|J_i(A)| \leq s$ for all $i \in [m]$, and for every $I \subseteq [m]$ with $|I| \leq r$ we have $|\bigcup_{i \in I} J_i(A)| \geq c \cdot |I|$. Thus, the only difference from boundary expanders consists in replacing $\partial_A(I)$ with $\bigcup_{i \in I} J_i(A)$.

Claim A.1 *Every $\left(r, s, \frac{s+c}{2}\right)$-expander is an (r, s, c)-boundary expander.*

Proof of Claim A.1. Since every column $j \in \bigcup_{i\in I} J_i(A) \setminus \partial_A(I)$ belongs to at least two sets $J_i(A)$ $(i \in I)$, we have the bound

$$\left|\bigcup_{i\in I} J_i(A)\right| \leq |\partial_A(I)| + \frac{1}{2}\left(\sum_{i\in I} |J_i(A)| - \partial_A(I)\right) \leq \frac{1}{2}(s|I| + |\partial_A(I)|).$$

On the other hand, $|\bigcup_{i\in I} J_i(A)| \geq \frac{s+c}{2}|I|$ since A is an $\left(r, s, \frac{s+c}{2}\right)$-expander. The required inequality $|\partial_A(I)| \geq c \cdot |I|$ follows.■

Thus, it remains to prove that $\boldsymbol{A}$ is an (r, s, c')-expander a.s., where $c' \stackrel{\text{def}}{=} \frac{c+s}{2}$. Let p_ℓ be the probability that any given ℓ rows of the matrix $\boldsymbol{A}$ violate the expansion property. Then

$$\mathbf{P}[\boldsymbol{A} \text{ is not a } (r, s, c')\text{-expander}] \leq \sum_{\ell=1}^{r} p_\ell \cdot m^\ell.$$

On the other hand,

$$p_\ell = \mathbf{P}[|\{\boldsymbol{j_{i\nu}} \mid i \in I,\ \nu \in [s]\}| \leq c'\ell] \leq \binom{n}{c'\ell} \cdot \left(\frac{c'\ell}{n}\right)^{s\ell}$$
$$\leq O(1)^{c'\ell} \cdot \left(\frac{c'\ell}{n}\right)^{(s-c')\ell} \leq \{O((sl)/n)\}^{(s-c')\ell},$$

where for the last inequality we used that $c' \leq \frac{7}{8}c \leq \frac{7}{8}s$ and hence $c' \leq O(s-c)$. Thus,

$$\mathbf{P}[\boldsymbol{A} \text{ is not a } (r, s, c')\text{-expander}] \leq \sum_{\ell=1}^{r} \{O((sl)/n)\}^{(s-c')\ell} m^\ell \leq \sum_{\ell=1}^{r} \left(\{O((sr)/n)\}^{(s-c')\ell} m\right)^\ell,$$

and since $m(sr/n)^{s-c'} = m(sr/n)^{(s-c)/2} \leq o(1)$ by our assumption, Lemma 2.2 follows.■

Lemma 2.3. *Let A be an $m \times n$ $(r, 2)$-boundary expander. Then for every $J \subseteq [n]$ with $|J| \leq r/4$ there exists $\widehat{J} \supseteq J$ sich that $\left|\mathrm{Ker}\left(\widehat{J}\right)\right| \leq 2|J|$ and $A \setminus \widehat{J}$ is an $(r/2, 3/2)$-boundary expander.*

Proof. We define a strictly increasing sequence of sets of columns $J_0 \supset J_1 \supset \ldots \supset J_t \supset \ldots$ as follows. Let $J_0 \stackrel{\text{def}}{=} J$. For $t > 0$, we first let S_t be an arbitrary set of rows violating the $(r/2, 3/2)$-boundary expansion condition in $A \setminus J_{t-1}$ if such a set exists; otherwise, the construction terminates. Then we let

$$J_t \overset{\text{def}}{=} J_{t-1} \cup \bigcup_{i \in S_t} J_i(A).$$

Note that since the chain $J_0 \supset J_1 \supset \ldots \supset J_t \ldots$ is strictly increasing, the process does terminate at some point; let J_T be the final set in this chain. We claim that $\widehat{J} := J_T$ has the required properties, and the only thing that has to be checked is that $|\mathrm{Ker}(J_T)| \leq 2|J|$. For that we prove by induction on $t = 0, \ldots, T$ that $|\mathrm{Ker}(J_t)| \leq 2|J|$.

Base case $|\mathrm{Ker}(J)| \leq 2|J|$ immediately follows from the fact that A is an $(r, 2)$-boundary expander and $|J| \leq r/4$.

Inductive step. Assume that $|\mathrm{Ker}(J_{t-1})| \leq 2|J|$ for some $1 \leq t \leq T$, and let us prove that $|\mathrm{Ker}(J_t)| \leq 2|J|$.

Since $|S_t| \leq r/2$, $|\mathrm{Ker}(J_{t-1})| \leq 2|J| \leq r/2$ and $\mathrm{Ker}(J_{t-1}) \cup S_t \subseteq \mathrm{Ker}(J_t)$, we can choose a set of rows I such that $\mathrm{Ker}(J_{t-1}) \cup S_t \subseteq I \subseteq \mathrm{Ker}(J_t)$ and

$$|I| = \min(r, |\mathrm{Ker}(J_t)|). \tag{9}$$

Applying to I the expansion condition, we get

$$|\partial_A(I)| \geq 2|I|.$$

On the other hand, $I \subseteq \mathrm{Ker}(J_t)$ implies that

$$\partial_A(I) \subseteq J \cup \bigcup_{s=1}^{t} \partial_{A \setminus J_{s-1}}(S_s).$$

Since S_s's violate the $(r/2, 3/2)$-boundary expansion conditions in respective matrices, we conclude that

$$|\partial_A(I)| \leq |J| + \frac{3}{2} \sum_{s=1}^{t} |S_s| \leq |J| + \frac{3}{2}|I|,$$

where for the latter inequality we used the fact $I \supseteq S_1 \dot{\cup} S_2 \dot{\cup} \ldots \dot{\cup} S_t$. Comparing these two inequalities, we find that $|I| \leq 2|J| \leq r/2$. Now (9) implies that in fact $|I| = |\mathrm{Ker}(J_{t+1})|$ that completes the inductive step in the proof of Lemma 2.3.■