SIAM J. COMPUT.
Vol. 49, No. 1, pp. 98–118
© 2020 Society for Industrial and Applied Mathematics

SUPERCRITICAL SPACE-WIDTH TRADE-OFFS FOR RESOLUTION*

CHRISTOPH BERKHOLZ† AND JAKOB NORDSTRÖM‡

Abstract. We show that there are CNF formulas which can be refuted in resolution in both small space and small width, but for which any small-width proof must have space exceeding by far the linear worst-case upper bound. This significantly strengthens the space-width trade-offs in [E. Ben-Sasson, *SIAM J. Comput.*, 38 (2009), pp. 2511–2525], and provides one more example of trade-offs in the "supercritical" regime above worst case recently identified by [A.A. Razborov, *J. ACM*, 63 (2016), 16]. We obtain our results by using Razborov's new hardness condensation technique and combining it with the space lower bounds in [E. Ben-Sasson and J. Nordström, *Short proofs may be spacious: An optimal separation of space and length in resolution*, in Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS '08), 2008, pp. 709–718].

Key words. proof complexity, resolution, space, width, trade-offs, supercritical

AMS subject classifications. 03B70, 03D15, 03F20, 68Q15, 68Q17, 68T15

DOI. 10.1137/16M1109072

1. Introduction.

Propositional proof complexity studies the problem of how to provide concise, polynomial-time checkable certificates that formulas in conjunctive normal form (CNF) are unsatisfiable. Research in this area was initiated in [22] as a way of attacking the problem of showing that $\mathsf{NP} \neq \mathsf{coNP}$, and hence $\mathsf{P} \neq \mathsf{NP}$, and it is therefore natural that the main focus has been on proving upper and lower bounds on proof length/size. More recently, however, other complexity measures have also been investigated, and this study has revealed a rich and often surprising web of connections.

1.1. Resolution length, width, and space.

Arguably the most thoroughly studied proof system in proof complexity is *resolution*, which appeared in [17] and began to be investigated in connection with automated theorem proving in the 1960s [23, 24, 35]. Because of its simplicity—there is only one derivation rule—and because all lines in a proof are clauses, this proof system is well suited for proof search, and it lies at the heart of current state-of-the-art SAT solvers based on so-called *conflict-driven clause learning* [4, 28, 29].

It is not hard to show that any unsatisfiable CNF formula over n variables can be proven unsatisfiable, or *refuted*, by a resolution refutation containing at most $\exp(\mathrm{O}(n))$ clauses, and this holds even in the restricted setting of *tree-like resolution*, where each intermediate clause in the refutation has to be rederived from scratch

*Received by the editors December 21, 2016; accepted for publication (in revised form) July 30, 2018; published electronically February 4, 2020. A preliminary version of this work appeared in *Proceedings of the 43rd International Colloquium on Automata, Languages and Programming (ICALP '16)*.

https://doi.org/10.1137/16M1109072

Funding: The first author was supported by a fellowship within the Postdoc-Programme of the German Academic Exchange Service (DAAD). The second author was supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013) ERC grant agreement 279611 and by Swedish Research Council grants 621-2010-4797, 621-2012-5645, and 2016-00782.

†Humboldt-Universität zu Berlin, Berlin D-10099, Germany (berkholz@informatik.hu-berlin.de, http://www.informatik.hu-berlin.de/~berkhoch).

‡Department of Computer Science, University of Copenhagen, 2100 Copenhagen, Denmark, and School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, 10044 Stockholm, Sweden (jn@di.ku.dk, http://www.csc.kth.se/~jakobn/).

every time it is used. In the breakthrough paper [26], Haken obtained a length lower bound on the form $\exp\bigl(\Omega\bigl(n^{\delta}\bigr)\bigr)$ for general resolution refutations of so-called pigeonhole principle formulas, and this result was later followed by truly exponential lower bounds $\exp(\Omega(n))$ for other formula families in [6, 20, 37] and many other papers.

In a seminal work [12], Ben-Sasson and Wigderson identified *width*, measured as the largest size of any clause appearing in a refutation, as another interesting complexity measure for resolution. Clearly, any unsatisfiable CNF formula over n variables can be refuted in width at most n. Moreover, any refutation in width w need never be longer than $n^{\mathrm{O}(w)}$, since this is an upper bound on the number of distinct clauses of width w (and this naive counting argument is essentially tight [3]). What Ben-Sasson and Wigderson showed is that strong enough *lower* bounds on width also imply lower bounds on length; in particular, that linear $\Omega(n)$ width lower bounds imply exponential $\exp(\Omega(n))$ length lower bounds for CNF formulas of bounded width. This connection can be used to rederive almost all currently known resolution length lower bounds.

Motivated by questions in SAT solving, where efficient memory management is a major concern, a more recent line of research in proof complexity has examined a third complexity measure on proofs, namely *space*. This study was initiated by Esteban and Torán [25], who defined the *(clause) space* of a resolution proof as the maximal number of clauses needed to be kept in memory during verification of the proof,[1] a definition that was generalized to other proof systems by Alekhnovich et al. [1]. It should be noted that although the original impetus for investigating proof space came from the applied SAT solving side, space complexity is, of course, a well-studied measure in its own right in computational complexity, and the study of space in proof complexity has turned out to be of intrinsic interest in that it has uncovered intriguing connections to proof length and width. It can be shown that a CNF formula over n variables can always be refuted in space $n + \mathrm{O}(1)$ even in tree-like resolution [25], although the refutation thus obtained might have exponential length. Linear space lower bounds matching the worst-case upper bound up to constant factors were obtained for a number of formula families in [1, 9, 25].

The space lower bounds obtained in the papers just discussed turned out to match closely known lower bounds on width, and in a strikingly simple and beautiful result Atserias and Dalmau [2] showed that in fact the resolution width of refuting a k-CNF formula F is a lower bound on the clause space required,[2] minus an additive term k adjusting for the largest width of any clause in F. This allows to recover the space lower bounds mentioned above as immediate consequences of width lower bounds shown in [12]. Furthermore, it follows from [2] that for $k = \mathrm{O}(1)$ any k-CNF formula that can be refuted by keeping only a constant number of clauses in memory can also be refuted in polynomial length and constant width. These connections go only in one direction, however—in the sequence of papers [30, 32, 10] it was shown that there are formula families that have high space complexity although they possess refutations in linear length and constant width.

[1] For completeness, we want to mention that for resolution there is also a *total space* measure counting the total number of literals in memory (with repetitions), which has been studied in [1, 13, 18, 19]. In this paper, however, "space" will always mean "clause space" in the sense of [25] unless otherwise stated.

[2] Note that this is a nontrivial connection since a lower bound on width, i.e., the *number of literals* in a clause, is shown to imply essentially the same lower bound on the *number of clauses* needed.

1.2. Resolution trade-offs. As was discussed above, a resolution proof in sufficiently small width will by necessity also be short, whereas the linear worst-case upper bound on space is achieved by a proof in exponential length. It is natural to ask, therefore, whether for a given formula F there exists a single resolution refutation of F that can simultaneously optimize these different complexity measures. The question of trade-offs between proof complexity measures was first raised by Ben-Sasson [8], who gave a strong negative answer for space versus width. More precisely, what was shown in [8] is that there are formulas which are refutable separately in constant width and in constant space, but for which any resolution proof minimizing one of the measures must exhibit almost worst-case linear behavior with respect to the other measure.

A question that arises in the context of SAT solving is whether it is possible to simultaneously optimize size and space (corresponding to running time and memory usage). In addition to the space-width trade-offs discussed above, in [8] Ben-Sasson also proved a size-space trade-off for the subsystem tree-like resolution, and building on [8, 10] it was shown in [11] for general resolution that there are formulas which have refutations in linear length and also in small space, but for which any space-efficient refutation must have superpolynomial or even exponential length. Beame, Beck, and Impagliazzo [5] extended the range of parameters of the trade-offs further by exhibiting formulas over n variables refutable in length polynomial in n where bringing the space down to linear, or even just shaving a constant factor off the exponent in the polynomial space bound that follows immediately from the length bound, incurs a superpolynomial penalty in proof length, a result that was generalized and strengthened in [7].

Turning finally to the relation between length and width, what was shown in [12] is that a short resolution refutation can be converted to a refutation of small width, but the way this conversion is done in [12] blows up the length exponentially. Thapen [36] proved that this is inherent by exhibiting formulas refutable in small width and small length, but for which any small-width refutation has to have exponential length. For the restricted case of tree-like resolution, Razborov [34] recently showed that there are formulas refutable in small width for which any tree-like refutation even doing slightly better than the trivial linear upper bound with respect to width must by necessity have doubly exponential length.

We want to emphasize an intriguing property of the trade-off results in [5, 7, 34] that was highlighted by Razborov, and that sets these results apart from the other trade-offs surveyed above. Namely, for most trade-off results between complexity measures it is the case that the trade-off plays out in the region between the worst-case upper bounds for the measures, where as one measure decreases the other measure has to approach its critical worst-case value. However, the short resolution proofs in [5, 7] require space even polynomially larger than the worst-case upper bound, and the small-width tree-like proofs in [34] require proofs exponentially larger than the exponential upper bound for tree-like length. To underscore the dramatic nature of such trade-off results, Razborov refers to them as *ultimate* in the preliminary version [33] of [34]. In this paper, we will instead use the term *supercritical trade-offs*, which we feel better describes the behavior that one of the complexity measures is pushed up into the supercritical regime above worst case when the other measure is decreased.

1.3. Our contribution. Answering Razborov's call in [34] for more examples of the type of trade-offs discussed above, in this paper we prove a supercritical trade-

off between space and width in resolution. As already observed, any refutation in width w of a CNF formula over n variables in general resolution need not contain more than $O(n^w)$ clauses, which is also a trivial upper bound on the space complexity of such a refutation. Our main result is that this bound is essentially tight, and is also somewhat robust. Namely, we show that there are n-variable formulas that can be refuted in width w, but for which any refutation in width even up to almost a multiplicative logarithmic factor larger than this requires space $n^{\Omega(w)}$.

THEOREM 1.1. *For any constant $\varepsilon > 0$ and any nondecreasing function $\ell(n)$, $6 \leq \ell(n) \leq n^{\frac{1}{2}-\epsilon}$, there is a family $\{F_n\}_{n\in\mathbb{N}}$ of n-variable CNF formulas which can be refuted in resolution in width $\ell(n)$ but for which any resolution refutation in width $o(\ell(n)\log n)$ requires clause space at least $n^{\Omega(\ell(n))}$.*

1.4. Techniques. In one sentence, we obtain our results by using Razborov's hardness condensation technique in [34] and combining it with the space lower bounds in [10].

In slightly more detail, our starting point are the so-called *pebbling formulas* defined in [12]. These formulas are refutable in constant width, but it was observed in [8] that space lower bounds for pebble games on directed acyclic graphs (DAGs) carry over to lower bounds on the *variable space* (the *number of variables* kept simultaneously in memory) of resolution refutations for the corresponding pebbling formulas. It was shown in [10] that substituting every variable in such formulas by an exclusive or of two new variables and expanding out to CNF produces a new family of formulas which are still refutable in constant width but for which the variable space lower bounds have been amplified to clause space lower bounds.

The result in [10] is one of several examples of how the method of *XOR substitution*, or *XORification*, has been used to amplify weak proof complexity lower bounds to much stronger lower bounds. In all of these applications distinct variables of the original formula are replaced by disjoint sets of new variables. The wonderfully simple (with hindsight) but powerful new idea in [34] is to instead do XOR substitution with overlapping sets of variables from a much smaller variable pool (but with exclusive ors of higher arity).

This recycling of variables has the consequence that hardness amplification as in [10] no longer works, since it crucially depends on the fact that all new substitution variables are distinct. What Razborov showed in [34] was essentially that if the pattern of overlapping variable substitutions is described by a strong enough bipartite expander, then locally there are enough distinct new variables to make tree-like amplification lower bounds as in [8] go through over a fairly wide range of the parameter space, yielding supercritical trade-offs between width and tree-like length. Since, in addition, the number of variables in the formula has decreased significantly, this can be viewed as a kind of *hardness condensation*.

We use Razborov's idea of XORification with recycled variables, but since we want to obtain results not for tree-like but for DAG-like resolution the technical details of our proofs are somewhat different. At a high level, we start with formulas over N variables that are refutable in constant width but require space $\Omega(N/\log N)$. We modify these formulas by applying w-wise XORification using a much smaller set of n variables, and then show that from any refutation in width $O(w)$ of this new, XORified formula it is possible to recover a refutation of the original formula with comparable space complexity. But this means that any small-width refutation of the XORified formula must have space complexity roughly $\Omega(N/\log N)$. Choosing parameters so that $N \approx n^w$ yields the bound stated in Theorem 1.1.

We should point out that compared to [34] we get significantly less robust trade-offs, which break down already for a multiplicative logarithmic increase in width. This is mainly due to the fact that we deal not with tree-like resolution as in [34], but with the much stronger general resolution proof system producing DAG-like proofs. We share with [34] the less desirable feature that although our formulas only have n variables they contain on the order of n^w clauses. Thus, measured in terms of formula size our space-width trade-offs do not improve on [8], and the width of our formulas is not constant but scales linearly with w. Still, since the number of variables provides a worst-case upper bound on space (independently of formula size), measured in terms of variables it seems fair to say that the trade-off result in Theorem 1.1 is fairly dramatic.

1.5. Organization of this paper. The rest of this paper is organized as follows. We start by reviewing some preliminaries in section 2. In section 3 we prove our main result assuming a hardness condensation lemma, and this lemma is then established in section 4.

We conclude in section 5 with a discussion of possible directions for future research. For completeness, proofs of some technical claims are provided in Appendix A.

2. Preliminaries. A *literal* over a Boolean variable x is either the variable x itself (a *positive literal*) or its negation $\overline{x}$ (a *negative literal*). We define $\overline{\overline{x}} = x$. A *clause* $C = a_1 \vee \cdots \vee a_k$ is a disjunction of literals over pairwise disjoint variables (without loss of generality we assume that there are no trivial clauses containing both a variable and its negation). A clause C' *subsumes* another clause C if every literal from C' also appears in C. A *k-clause* is a clause that contains at most k literals. A *CNF formula* $F = C_1 \wedge \cdots \wedge C_m$ is a conjunction of clauses, and F is a *k-CNF formula* if it consists of k-clauses. We write $\mathit{Vars}(F)$ to denote the set of variables appearing in a formula F. We think of clauses and CNF formulas as sets: the order of elements is irrelevant and there are no repetitions.

A *resolution refutation* $\pi : F \vdash \bot$ of an unsatisfiable CNF formula F, which can also be referred to as a *resolution proof* for (the unsatisfiability of) F, is an ordered sequence of clauses $\pi = (D_1, \ldots, D_\tau)$ such that $D_\tau = \bot$ is the empty clause containing no literals, and each clause D_i, $i \in [\tau] = \{1, \ldots, \tau\}$, is either one of the clauses in F (an *axiom*) or is derived from clauses D_j, D_k in π with $j, k < i$ by the *resolution rule*

$$\frac{B \vee x \qquad C \vee \overline{x}}{B \vee C} \ . \tag{2.1}$$

For technical reasons, it will also be convenient to permit a *weakening rule*

$$\frac{B}{B \vee C} \tag{2.2}$$

allowing to derive a strictly weaker clause from a clause already derived, although this rule is not essential.

With every resolution proof π we can associate a DAG G_π by having a sequence of vertices v_i on a line in order of increasing i, labeled by the clauses $D_i \in \pi$, and with directed edges (v_j, v_i) and (v_k, v_i) if the clause D_i was derived by resolution from D_j and D_k or an edge (v_j, v_i) if D_i was derived from D_j by weakening. Note that there might be several occurrences of a clause D in the proof π, and if so each occurrence gets its own vertex in G_π.

Now we can formally define the proof complexity measures discussed in section 1. The *length* $L(\pi)$ of a resolution proof π is the number of clauses in it (counted with

repetitions). The *width* $W(C)$ of a clause C is $|C|$, i.e., the number of literals, and the width $W(\pi)$ of a proof π is the size of a largest clause in π. The *(clause) space* at step i is the number of clauses C_j, $j < i$, with edges to clauses C_k, $k \geq i$ in G_π, plus 1 for the clause C_i derived at this step. Intuitively, space measures the number of clauses we need to keep in memory at step i, since they were derived before step i but are used to infer new clauses at or after step i. The space $Sp(\pi)$ of a proof π is the maximum space over all steps in π. Taking the minimum over all resolution refutations of a CNF formula F, we define the length, width, and space of refuting F, respectively, as $L(F \vdash \bot) = \min_{\pi:F\vdash\bot}\{L(\pi)\}$, $W(F \vdash \bot) = \min_{\pi:F\vdash\bot}\{W(\pi)\}$, and $Sp(F \vdash \bot) = \min_{\pi:F\vdash\bot}\{Sp(\pi)\}$. We remark that any applications of the weakening rule (2.2) can always be eliminated from a refutation without increasing the length, width, or space.

When reasoning about space, it is sometimes convenient to use a slightly different, but equivalent, description of resolution that makes explicit what clauses are in memory at each point in time. We say that a *configuration-style resolution refutation* is a sequence $(\mathbb{D}_0, \ldots, \mathbb{D}_\tau)$ of sets of clauses, or *configurations*, such that $\mathbb{D}_0 = \emptyset$, $\bot \in \mathbb{D}_\tau$, and for all $t \in [\tau]$ the configuration $\mathbb{D}_t$ is obtained from $\mathbb{D}_{t-1}$ by one of the following *derivation steps*:

Axiom download. $\mathbb{D}_t = \mathbb{D}_{t-1} \cup \{C\}$, where C is a clause $C \in F$.

Inference. $\mathbb{D}_t = \mathbb{D}_{t-1} \cup \{D\}$ for a clause D derived by resolution (2.1) or weakening (2.2) from clauses in $\mathbb{D}_{t-1}$.

Erasure. $\mathbb{D}_t = \mathbb{D}_{t-1} \setminus \mathbb{D}'$ for some $\mathbb{D}' \subseteq \mathbb{D}_{t-1}$.

The length of a configuration-style refutation $\pi = (\mathbb{D}_0, \ldots, \mathbb{D}_\tau)$ is the number of axiom downloads and inference steps, the width is the size of a largest clause, as before, and the space is $\max_{t\in[\tau]}\{|\mathbb{D}_t|\}$. Given a refutation as an ordered sequence of clauses $\pi = (D_1, \ldots, D_\tau)$, we can construct a configuration-style refutation in the same length, width, and space by deriving each clause D_i via an axiom download or inference step, and interleave with erasures of clauses C_j, $j < i$, as soon as these clauses have no edges to clauses C_k, $k \geq i$, in the associated DAG G_π. In the other direction, taking a configuration-style refutation and listing the sequence of axiom download and inference steps yields a standard resolution refutation in the same length, width, and space (assuming that clauses are erased as soon as possible). Thus, we can switch freely between these two ways of describing resolution refutations.

In this paper, it will be convenient for us to limit our attention to a (slightly nonstandard) restricted form of resolution refutations as described next. We define a *homogeneous resolution refutation* to be a refutation where every resolution rule application is of the form

$$\frac{C \vee x \qquad C \vee \overline{x}}{C} \ . \tag{2.3}$$

The requirement of homogeneity is essentially without loss of generality, since we need to insert at most two weakening steps before each application of the resolution rule, which increases the width by at most 1, and the weakened clauses can then immediately be forgotten. We state this observation formally for the record.

OBSERVATION 2.1. *If a CNF formula F has a standard resolution refutation without weakening steps in length L, width w, and (clause) space s, then it has a homogeneous refutation in length at most $3L$, width at most $w + 1$, and (clause) space at most $s + 2$.*

As already mentioned, a useful trick to obtain hard CNF formulas for different proof systems and complexity measures, which will play a key role also in this paper,

is *XORification*, i.e., substituting variables by exclusive ors of new variables and expanding out in the canonical way to obtain a new CNF formula. For example, the standard way to define binary XOR substitution for a positive literal x is

$$x[\oplus_2] = (x_1 \vee x_2) \wedge (\overline{x}_1 \vee \overline{x}_2) \tag{2.4}$$

for a negative literal $\overline{y}$ we have

$$\overline{y}[\oplus_2] = (y_1 \vee \overline{y}_2) \wedge (\overline{y}_1 \vee y_2) \ , \tag{2.5}$$

and applying binary XOR substitution to the clause $x \vee \overline{y}$ we obtain the CNF formula

$$\begin{aligned}(x \vee \overline{y})[\oplus_2] = x[\oplus_2] \vee \overline{y}[\oplus_2] = (x_1 \vee x_2 \vee y_1 \vee \overline{y}_2) \wedge (x_1 \vee x_2 \vee \overline{y}_1 \vee y_2) \\ \wedge\ (\overline{x}_1 \vee \overline{x}_2 \vee y_1 \vee \overline{y}_2) \wedge (\overline{x}_1 \vee \overline{x}_2 \vee \overline{y}_1 \vee y_2) \ ,\end{aligned} \tag{2.6}$$

where we use the notation $\mathbb{C} \vee \mathbb{D}$ for two set of clauses $\mathbb{C}$ and $\mathbb{D}$ to mean

$$\mathbb{C} \vee \mathbb{D} = \{C \vee D \mid C \in \mathbb{C},\, D \in \mathbb{D},\, C \vee D \text{ nontrivial}\} \ . \tag{2.7}$$

The XORification of a CNF formula F is the conjunction of all the formulas corresponding to the XORified clauses of F. We trust that the reader has no problems parsing this slightly informal definition by example or generalizing it to substitutions with XOR of arbitrary arity (but see, e.g., Definition 2.12 in [31] for a more rigorous treatment).

Usually, XORification is done in such a way that any two variables in the original formula are replaced by exclusive ors over disjoint sets of new variables. Razborov [34] observed that it can sometimes be useful to allow XORification with overlapping sets of variables. Let us define this concept more carefully.

Definition 2.2 (XORification with recycling [34]). *Let F be a CNF formula over the set of variables $u_1, \ldots, u_N$, and let $\mathcal{G} = (U \dot\cup V, E)$ be a bipartite graph with left vertex set $U = \{u_1, \ldots, u_N\}$ and right vertex set $V = \{v_1, \ldots, v_n\}$. Then for the variables u_i we define the XORified literals $u_i[\mathcal{G}] = \bigoplus_{v \in \mathcal{N}(u_i)} v$ and $\overline{u}_i[\mathcal{G}] = \neg \bigoplus_{v \in \mathcal{N}(u_i)} v$ (where $\mathcal{N}(u_i)$ denotes the neighbors in V of u_i), for clauses $C \in F$ we define $C[\mathcal{G}] = \bigvee_{a \in C} a[\mathcal{G}]$ expanded out to CNF as in* (2.6) *but with trivial clauses pruned away, and the* XORification of F with respect to $\mathcal{G}$ *is defined to be $F[\mathcal{G}] = \bigwedge_{C \in F} C[\mathcal{G}]$.*

Note that if F is an N-variable k-CNF with m clauses and $\mathcal{G} = (\{u_1, \ldots, u_N\} \dot\cup \{v_1, \ldots, v_n\}, E)$ is a bipartite graph of left degree d, then $F[\mathcal{G}]$ is an n-variable kd-CNF formula with at most $2^{k(d-1)}m$ clauses. We want to highlight that by definition we have the equality

$$(C \vee a)[\mathcal{G}] = C[\mathcal{G}] \vee a[\mathcal{G}] \tag{2.8}$$

(where we can view the expressions in (2.8) either as the Boolean functions computed by these formulas or as the corresponding clause sets but with trivial clauses removed), and this will be convenient to use in some of our technical arguments.

We conclude this section with two simple observations that will also be useful in what follows.

Observation 2.3. *If F has a (homogeneous) resolution refutation in width w and $\mathcal{G}$ has left degree bounded by d, then $F[\mathcal{G}]$ can be refuted in (homogeneous) resolution in width at most $2dw$.*

Observation 2.3 is not hard to show and follows, e.g., from the proof of Theorem 2 in [11] (strictly speaking, this theorem is for XORification *without* recycling, but recycling can only decrease the width). For the convenience of the reader, we include a proof below.

Proof. First, consider a homogeneous resolution refutation π of F in width w. We show that for every $C \in \pi$ we can construct a homogeneous derivation of the clause set $C[\mathcal{G}]$ from $F[\mathcal{G}]$ in width dw. For axioms there is nothing to prove, and the case of weakening steps is immediate. Suppose, therefore, that C is derived by resolving $C \vee u$ and $C \vee \overline{u}$. Note that by definition we have

$$u[\mathcal{G}] \cup \overline{u}[\mathcal{G}] = \{D \mid \mathit{Vars}(D) = \mathcal{N}(u)\} \ , \tag{2.9}$$

where the latter set denotes all $2^{|\mathcal{N}(u)|}$ clauses over the variables in $\mathcal{N}(u)$. Using (2.7) and (2.8) to unpack and repackage the notation, we obtain that

$$(C \vee u)[\mathcal{G}] \cup (C \vee \overline{u})[\mathcal{G}] = C[\mathcal{G}] \vee \{D \mid \mathit{Vars}(D) = \mathcal{N}(u)\} \tag{2.10}$$

(implicitly pruning away trivial clauses), from which it is straightforward to see that we can derive every clause in $C[\mathcal{G}]$ by a sequence of homogeneous resolution steps resolving away all the variables in $\mathcal{N}(u) \setminus \mathit{Vars}(C[\mathcal{G}])$ in (2.10).

For refutations that are not homogeneous, we first apply Observation 2.1 to translate them into a homogenous refutation of width at most $w + 1$ and then apply the argument above to obtain a refutation of width $d(w + 1)$. Hence, in both cases the width is bounded by $2dw$. ∎

Observation 2.4. *If F has a (homogeneous) resolution refutation π such that the associated DAG G_π has depth (i.e., longest path) s, then π can be carried out (in homogeneous resolution) in space $s + 2$ (possibly by repeating and/or reordering clauses in π).*

This second observation is essentially due to [25]. To see why this is true, note that the proof DAG G_π can be turned into a binary tree of the same depth by repeating vertices/clauses, and it is then straightforward to show that any tree-like proof DAG in depth s can be realized in space at most $s + 2$.

3. Proof of main theorem. In this section we present a proof of Theorem 1.1. The proof makes use of the following hardness condensation lemma, which will be established in the next section and which is the main technical contribution of the paper.

Lemma 3.1 (hardness condensation lemma). *For all $k \in \mathbb{N}^+$ and $\varepsilon > 0$ there exist $n_0 \in \mathbb{N}^+$ and $\delta > 0$ such that the following holds. Let ℓ and n be integers satisfying $n \geq n_0$ and $k \leq \ell \leq n^{\frac{1}{2}-\varepsilon}$, and suppose that F is an unsatisfiable CNF formula over $N = \lfloor n^{\delta\ell} \rfloor$ variables which requires width $W(F \vdash \bot) = k$ and (clause) space $Sp(F \vdash \bot) = s$ to be refuted in resolution.*

Then there is a bipartite graph $\mathcal{G} = (U \dot\cup V, E)$ with $|U| = N$ and $|V| = n$ such that the n-variable CNF formula $F[\mathcal{G}]$ has the following properties:

- *$F[\mathcal{G}]$ can be refuted in resolution in width ℓ.*
- *Any resolution refutation $\pi : F[\mathcal{G}] \vdash \bot$ in width $w \leq \ell \log n$ requires space $Sp(\pi) \geq (s - w - 3)2^{-w}$.*

We want to apply this lemma to formulas of low width complexity but high space complexity as stated next.

THEOREM 3.2 (see [10]). *There is a family $\{F_N\}_{N\in\mathbb{N}}$ of N-variable 6-CNF formulas of size $\Theta(N)$ which can be refuted in width $W(F_N \vdash \bot) = 6$ but require space $Sp(F_N \vdash \bot) = \Omega(N/\log N)$.*

Combining Lemma 3.1 and Theorem 3.2, we can prove our main result.

Proof of Theorem 1.1. Recall that we want to prove that for any constant $\varepsilon > 0$ and any nondecreasing function $\ell(n)$ such that $6 \le \ell(n) \le n^{\frac{1}{2}-\epsilon}$, there is a family $\{F_n\}_{n\in\mathbb{N}}$ of n-variable CNF formulas which have resolution refutations of width $\ell(n)$ but for which any refutation of width $o(\ell(n)\log n)$ requires clause space $n^{\Omega(\ell(n))}$.

From Theorem 3.2 we obtain constants $\varepsilon' > 0$ and $N_0 \in \mathbb{N}^+$ and a family of N-variable 6-CNF formulas F_N that require clause space $\varepsilon' N/\log N$ for all $N \ge N_0$. We want to apply hardness condensation as in Lemma 3.1 to these formulas. Let $\varepsilon > 0$ be given in Theorem 1.1 and fix $k = 6$. Plugging this into Lemma 3.1 provides constants $\delta > 0$ and $n_0 \in \mathbb{N}^+$, where in addition we choose n_0 large enough so that $\lfloor n_0^{\delta\ell(n_0)} \rfloor \ge N_0$ (this is always possible since $\delta\ell(n_0) \ge 6\delta > 0$).

For any $n \ge n_0$, set $N = \lfloor n^{\delta\ell(n)} \rfloor \ge N_0$ and let $\mathcal{G} = (U \dot\cup V, E)$ with $|U| = N$ and $|V| = n$ be a bipartite graph with properties as guaranteed by Lemma 3.1. Then the lemma says that $F_N[\mathcal{G}]$ is an n-variable formula which can be refuted in width ℓ, but for which every refutation of width $w \le \ell \log n$ requires clause space $(s - w - 3)2^{-w}$, where $s \ge \varepsilon' N/\log N = \varepsilon' \lfloor n^{\delta\ell(n)} \rfloor / \log \lfloor n^{\delta\ell(n)} \rfloor$ is the space lower bound for F_N. Choosing $w \le \frac{\delta}{2} \cdot \ell(n) \log n - 3$ (which we can do since the width bound is $o(\ell(n)\log n)$ in the statement of Theorem 1.1) and using that $\ell(n)$ is nondecreasing in n, the calculation

$$(s - w - 3)2^{-w} \ge \left(\varepsilon' \lfloor n^{\delta\ell(n)} \rfloor / \log \lfloor n^{\delta\ell(n)} \rfloor - \tfrac{\delta}{2}\ell(n)\log n\right) 2^{-\frac{\delta}{2}\ell(n)\log n} \ge \Omega\left(n^{\frac{\delta}{3}\ell(n)}\right) \tag{3.1}$$

yields the desired space lower bound. ∎

If one looks more closely at what is going on inside the proof of Theorem 1.1, where Lemma 3.1 and Theorem 3.2 come together, one can make a somewhat intriguing observation.

As discussed in the introduction, Theorem 3.2 is shown by using so-called pebbling formulas, which we now describe briefly. Given a DAG $\mathcal{D}$ with sources S and a unique sink z, and with all nonsources having fan-in 2, we let every vertex in $\mathcal{D}$ correspond to a variable and define the *pebbling formula* over $\mathcal{D}$, denoted $Peb_{\mathcal{D}}$, to consist of the following clauses:

- for all $s \in S$, the clause s;
- for all nonsource vertices v with predecessors u_1, u_2, the clause $\overline{u}_1 \vee \overline{u}_2 \vee v$;
- for the sink z, the clause $\overline{z}$.

Applying standard binary XOR substitution (without recycling of variables) as in (2.6) to these CNF formulas amplifies lower bounds on the number of variables in memory $VarSp(Peb_{\mathcal{D}} \vdash \bot)$ (which follow from properties of the chosen DAG $\mathcal{D}$) to lower bounds on the number of clauses $Sp(Peb_{\mathcal{D}}[\oplus_2] \vdash \bot)$ [10]. In Lemma 3.1 we then do another round of XOR substitution, this time with recycling, to decrease the number of variables while maintaining the space lower bound for small-width refutations. It is not entirely clear why we would need two separate rounds of XORification to achieve this result. In one sense, it would seem more satisfying to get a clean one-shot argument that just takes pebbling formulas and yields the supercritical trade-offs by only one round of XORification.

And in fact, if we are willing to accept a slightly weaker bound, we could make such a one-shot argument and apply substitution with recycling directly to the pebbling formulas. The reason for this is that one can actually prove a somewhat stronger version of hardness condensation than in Lemma 3.1, as we will see in section 4. There is no need to require that the original formula should have high space complexity unconditionally, but it suffices that the formula exhibits a strong trade-off between width and clause space. Since the number of clauses times the maximal width of any clause is an upper bound on the total number of distinct variables in memory, for any resolution refutation π we have the inequality $Sp(\pi) \cdot W(\pi) \geq VarSp(\pi)$. In [8] a variable space lower bound $VarSp(Peb_{\mathcal{D}} \vdash \bot) = \Omega(N/\log N)$ was presented (for appropriately chosen DAGs $\mathcal{D}$), implying that any width-w refutation requires clause space at least $\Omega(N/(w \log N))$. Since our hardness condensation step incurs a loss of a factor 2^{-w}, by starting with standard pebbling formulas and applying XORification with recycling directly we could obtain asymptotically similar bounds to those in Theorem 1.1 in one shot.

However, one can also argue that by combining Lemma 3.1 and Theorem 3.2 in the way done above one obtains a more modular proof, which shows that any formulas satisfying the conditions in Theorem 3.2 can be used for hardness condensation in a black-box fashion. This is why we chose to present the proof in this way.

4. Hardness condensation. Let us now establish the hardness condensation lemma. We prove a slightly stronger version of the lemma below, which obviously subsumes Lemma 3.1.

LEMMA 4.1 (hardness condensation lemma, strong version). *For all $k \in \mathbb{N}^+$ and $\varepsilon > 0$ there are $n_0 \in \mathbb{N}^+$ and $\delta > 0$ such that the following holds. Let ℓ and n be integers satisfying $n \geq n_0$ and $k \leq \ell \leq n^{\frac{1}{2}-\varepsilon}$ and suppose that F is an unsatisfiable CNF formula over $N = \lfloor n^{\delta\ell} \rfloor$ variables which requires width $W(F \vdash \bot) \leq k$ to be refuted in resolution.*

Then there is a bipartite graph $\mathcal{G} = (U \dot\cup V, E)$ with $|U| = N$ and $|V| = n$ such that the n-variable CNF formula $F[\mathcal{G}]$ has the following properties:

- *The XORified formula $F[\mathcal{G}]$ can be refuted in width ℓ.*
- *Any resolution refutation $\pi : F[\mathcal{G}] \vdash \bot$ of the XORified formula $F[\mathcal{G}]$ in width $w \leq \ell \log n$ requires space $Sp(\pi) \geq (s - w - 3)2^{-w}$, where s is the minimal space of any refutation $\pi' : F \vdash \bot$ of the original formula F in width at most w.*

Clearly, the key to obtaining Lemma 4.1 is to choose the right kind of graphs. As in [34], we use boundary expander graphs where the right-hand side is significantly smaller than the left-hand side. Let us start by giving a proper definition of these graphs and reviewing the properties that we need from them. Most of our discussion of boundary expanders can be recovered from [34], but since our setting of parameters is slightly different we give a self-contained presentation and also provide full proofs of all claims in Appendix A for completeness. We remark that there is also a significant overlap with [15] in our treatment of expander graphs below.

In what follows, we will let $\mathcal{G} = (U \dot\cup V, E)$ denote a bipartite graph with left vertices U and right vertices V. We write $\mathcal{N}^{\mathcal{G}}(U') = \{v \,|\, \{u, v\} \in E(\mathcal{G}), u \in U'\}$ to denote the set of right neighbors of a left vertex subset $U' \subseteq U$ (and vice versa for right vertex subsets), dropping the graph $\mathcal{G}$ from the notation when it is clear from context. For a single vertex v we will use the abbreviation $\mathcal{N}(v) = \mathcal{N}(\{v\})$.

DEFINITION 4.2 (boundary expander). *A bipartite graph $\mathcal{G} = (U \dot\cup V, E)$ is an $N \times n$ (r, c)-*boundary expander*, or* unique neighbor expander*, if $|U| = N$, $|V| = n$,*

and for every set $U' \subseteq U$, $|U'| \leq r$, it holds that $|\partial(U')| \geq c|U'|$, where $\partial(U') = \{v \in \mathcal{N}^{\mathcal{G}}(U') : |\mathcal{N}^{\mathcal{G}}(v) \cap U'| = 1\}$ is the boundary *or the set of* unique neighbors *of U'. An (r, d, c)-boundary expander is an (r, c)-boundary expander where additionally $|\mathcal{N}^{\mathcal{G}}(u)| \leq d$ for all $u \in U$, i.e., where the left degree is bounded by d.*

An important property of (r, c)-boundary expanders, which holds for arbitrarily small but positive expansion $c > 0$, is that any left vertex subset $U' \subseteq U$ of size $|U'| \leq r$ has a matching into V. In addition, this matching can be chosen in such a way that there is an ordering of the vertices in U' such that every vertex $u_i \in U'$ is matched to a vertex outside of the neighborhood of the preceding vertices $u_1, \ldots, u_{i-1}$. The proof of this fact uses what is sometimes referred to as a *peeling argument*, which we recapitulate below for the reader's convenience.

LEMMA 4.3 (peeling lemma). *Let $\mathcal{G} = (U \dot\cup V, E)$ be an (r, c)-boundary expander with $r \geq 1$ and $c > 0$. Then every left vertex subset $U' \subseteq U$ of size $|U'| = \ell \leq r$ can be ordered $U' = (u_1, \ldots, u_\ell)$ in such a way that there is a matching into an ordered right vertex subset $V' = (v_1, \ldots, v_\ell) \subseteq V$ for which $v_i \in \mathcal{N}(u_i) \setminus \mathcal{N}(\{u_1, \ldots, u_{i-1}\})$.*

Proof. The proof is by induction on ℓ. The base case $\ell = 1$ is immediate since $r \geq 1$ and $c > 0$ implies that no left vertex can be isolated. For the induction step, suppose the lemma holds for $\ell - 1$. To define the sequence $v_1, \ldots, v_\ell$ we first fix any $v_\ell \in \partial(U')$, which exists because $|\partial(U')| \geq c|U'| > 0$. Since v_ℓ is in the boundary of U' there exists a unique $u_\ell \in U'$ such that $|\mathcal{N}(v_\ell) \cap U'| = \{u_\ell\}$. Thus, for this pair (u_ℓ, v_ℓ) it holds that $v_\ell \in \mathcal{N}(u_\ell) \setminus \mathcal{N}(U' \setminus \{u_\ell\})$. By the induction hypothesis we can now find sequences $u_1, \ldots, u_{\ell-1}$ and $v_1, \ldots, v_{\ell-1}$ for $U' \setminus \{u_\ell\}$ such that $v_i \in \mathcal{N}(u_i) \setminus \mathcal{N}(\{u_1, \ldots, u_{i-1}\})$, to which we can append u_ℓ and v_ℓ at the end. The lemma follows by the induction principle. ∎

For a right vertex subset $V' \subseteq V$ in $\mathcal{G} = (U \dot\cup V, E)$ we define the *kernel* $\ker(V') \subseteq U$ of V' to be the set of all left vertices whose entire neighborhood is contained in V', i.e.,

$$\ker(V') = \{u \in U \mid \mathcal{N}(u) \subseteq V'\} \ . \tag{4.1}$$

We write $\mathcal{G} \setminus V'$ to denote the subgraph of $\mathcal{G}$ induced on $(U \setminus \ker(V')) \dot\cup (V \setminus V')$. In other words, we can think of $\mathcal{G} \setminus V'$ as being obtained from $\mathcal{G}$ by first deleting V' and afterwards all isolated vertices from U.

Even if $\mathcal{G}$ is a good expander, it can happen that $\mathcal{G} \setminus V'$ is no longer expanding. But another key property of boundary expanders is that for any small enough right vertex set V' we can always find a *closure* $\gamma(V') \supseteq V'$ with a small kernel on the left such that the subgraph $\mathcal{G} \setminus \gamma(V')$ has good boundary expansion. The next lemma is very similar to an analogous lemma in [34], but since our parameters are slightly different we provide a proof in Appendix A.

LEMMA 4.4. *Let $\mathcal{G}$ be an $(r, 2)$-boundary expander. Then for every $V' \subseteq V$ with $|V'| \leq r/2$ there exists a set of vertices $\gamma(V') \supseteq V'$ such that $|\ker(\gamma(V'))| \leq |V'|$ and the induced subgraph $\mathcal{G} \setminus \gamma(V')$ is an $(r/2, 1)$-boundary expander.*

We also need a lemma stating that there exist $N \times n$ $(r, d, 2)$-boundary expanders where the size n of the right-hand side is significantly smaller than the size $N = n^{\Theta(d)}$ of the left-hand side. The proof, which closely follows [34, Lemma 2.2], is a standard application of the probabilistic method, but is included in Appendix A for completeness.

LEMMA 4.5. *Fix constants $\varepsilon, \delta > 0$ and $d_0 \geq 2$ such that $\delta + \frac{1}{d_0} < \varepsilon/2$. Then there exists an $n_0 \in \mathbb{N}^+$ such that for all n, d, and r satisfying $n \geq n_0$, $d_0 \leq d \leq n^{1/2-\varepsilon}$, and $r \leq n^{1/2}$ there are $\lfloor n^{\delta d} \rfloor \times n$ $(r, d, 2)$-boundary expanders.*

After this review of boundary expanders and their properties we now come to the core argument of the paper, namely that space lower bounds are preserved for small-width resolution refutations when we apply XORification as in Definition 2.2 with respect to an $(r, 2)$-boundary expander. To get cleaner technical arguments in the proofs we will restrict our attention to homogeneous resolution refutations as in (2.3), which for our purposes is without loss of generality by Observation 2.1.

LEMMA 4.6 (main technical lemma). *Let F be an unsatisfiable CNF-formula, and let $\mathcal{G}$ be an $(r, 2)$-boundary expander, and suppose that $\pi : F[\mathcal{G}] \vdash \bot$ is a homogeneous resolution refutation in width $w \leq r/2$ of the XORified formula $F[\mathcal{G}]$. Then there is a homogeneous refutation $\pi' : F \vdash \bot$ of the original formula F in width at most w and space $Sp(\pi') \leq 2^w \cdot Sp(\pi) + w + 3$.*

Proof. Assume that $\pi = (\mathbb{C}_0, \mathbb{C}_1, \ldots, \mathbb{C}_\tau)$ is a configuration-style homogeneous resolution refutation of $F[\mathcal{G}]$ in width $W(\pi) = w \leq r/2$. We will show how to transform π into a refutation π' of the original formula F in width and space as claimed in the lemma. To help the reader navigate the proof, we remark that in what follows we will use the notational conventions that B and C denote clauses over $\mathit{Vars}(F[\mathcal{G}])$, D denotes a clause over $\mathit{Vars}(F)$, and A denotes an axiom clause from the original formula F before XORification.

Recall that for clauses $C \in F[\mathcal{G}]$ we have $\mathit{Vars}(C) \subseteq V$ by construction. For convenience, we will overload notation and write $\ker(C) = \ker(\mathit{Vars}(C))$, which is a subset of the variables U of the original formula F. Furthermore, for every clause $C \in \pi$ we fix $\gamma(C) := \gamma(\mathit{Vars}(C)) \subseteq V$ to be a minimal closure with properties as guaranteed by Lemma 4.4, i.e., such that $\left|\ker\big(\gamma(C)\big)\right| \leq |\mathit{Vars}(C)|$ and the induced subgraph $\mathcal{G} \setminus \gamma(C)$ is an $(r/2, 1)$-boundary expander. Note that such closures exist since all clauses $C \in \pi$ have width at most w, and by requiring minimality we ensure that $\gamma(\bot) = \emptyset$. It might be worth pointing out, though, that the definition of closure is purely existential—we have no control over how these closures are constructed, and, in particular, for two clauses B and C such that $B \subseteq C$ it does not necessarily hold that $\gamma(B) \subseteq \gamma(C)$.

An important notion in what follows will be that of *simultaneous falsifiability*, where we say that two CNF formulas F and G are *simultaneously falsifiable* if there is a truth value assignment that at the same time falsifies both F and G. To transform the resolution refutation π of $F[\mathcal{G}]$ into a refutation π' of F we let $\mathbb{D}_t$ be obtained from $\mathbb{C}_t$ by replacing every clause $C \in \mathbb{C}_t$ by the set of clauses

$$\mathcal{G}^{-1}(C) := \{D \mid \mathit{Vars}(D) = \ker(\gamma(C));\ D[\mathcal{G}] \text{ and } C \text{ simultaneously falsifiable}\} \tag{4.2}$$

and defining

$$\mathbb{D}_t := \textstyle\bigcup_{C \in \mathbb{C}_t} \mathcal{G}^{-1}(C) \tag{4.3}$$

(where the notation $\mathcal{G}^{-1}(C)$ is chosen to suggest that this is in some intuitive sense the "inverse operation" of XORification with respect to $\mathcal{G}$). Every clause in $D \in \mathcal{G}^{-1}(C)$ has width at most w, because

$$|\mathit{Vars}(D)| = |\ker(\gamma(C))| \leq W(C) \leq w\ , \tag{4.4}$$

where the first equality is by the definition in (4.2), the following inequality is guaranteed by Lemma 4.4, and the final inequality is by assumption. Furthermore, we have

$|\mathcal{G}^{-1}(C)| \leq 2^w$, since all clauses in $\mathcal{G}^{-1}(C)$ are over the same set of variables and each variable appears positively or negatively in every clause, and hence

$$\left|\mathbb{D}_t\right| \leq 2^w \cdot \left|\mathbb{C}_t\right| \leq 2^w \cdot Sp(\pi) \ . \tag{4.5}$$

We want to argue that the sequence $\left(\mathbb{D}_0, \mathbb{D}_1, \ldots, \mathbb{D}_\tau\right)$ is the "backbone" of a resolution refutation π' of F, by which we mean that for every t it holds that $\mathbb{D}_{t+1}$ can be derived from $\mathbb{D}_t$ by a sequence of intermediate steps without affecting any proof complexity measure too much.

To make this claim formal, we first observe that for $\mathbb{C}_0 = \emptyset$ we obviously get $\mathbb{D}_0 = \emptyset$ by (4.3). Moreover, it holds that $\mathcal{G}^{-1}(\bot) = \{\bot\}$ and hence $\bot \in \mathbb{D}_\tau$, since the unique minimal closure of the empty set is the empty set itself. We want to show that for every $0 \leq t < \tau$ the configuration $\mathbb{D}_{t+1}$ can be obtained from $\mathbb{D}_t$ by a resolution derivation $(\mathbb{D}_t = \mathbb{D}_{t,0}, \mathbb{D}_{t,1}, \mathbb{D}_{t,2}, \ldots, \mathbb{D}_{t,j_t-1}, \mathbb{D}_{t,j_t} = \mathbb{D}_{t+1})$, where the space of every intermediate configuration is bounded by $\max\{Sp(\mathbb{D}_t), Sp(\mathbb{D}_{t+1})\} + w + 3$.

If $\mathbb{C}_{t+1}$ is obtained from $\mathbb{C}_t$ by erasing a clause C, then $\mathbb{D}_{t+1}$ can be obtained from $\mathbb{D}_t$ by erasing all clauses $\mathcal{G}^{-1}(C) \setminus \mathbb{D}_{t+1}$. Suppose that $\mathbb{C}_{t+1}$ is obtained from $\mathbb{C}_t$ by downloading an axiom $C \in F[\mathcal{G}]$. We claim that every clause in $\mathcal{G}^{-1}(C)$ is either an axiom or a weakening of an axiom from F. By the definition of $F[\mathcal{G}]$, every axiom $C \in F[\mathcal{G}]$ is a clause in the CNF formula $A[\mathcal{G}]$ for some original axiom $A \in F$. Fix any axiom $A \in F$ such that $C \in A[\mathcal{G}]$. Then for all $D \in \mathcal{G}^{-1}(C)$ it holds by (4.2) that $\mathit{Vars}(D) = \ker(\gamma(C)) \supseteq \ker(C) \supseteq \mathit{Vars}(A)$ and that there is an assignment falsifying both $D[\mathcal{G}]$ and C. To see that this implies that A subsumes D, suppose that there is a variable x appearing positively in A such that $\overline{x} \in D$. Any truth value assignment falsifying $D[\mathcal{G}]$ must falsify $a[\mathcal{G}]$ for all literals $a \in D$, and hence in particular $\overline{x}[\mathcal{G}]$. This means that $x[\mathcal{G}]$ is satisfied by the same assignment, and then so is all of the formula $A[\mathcal{G}]$ including C, since $A[\mathcal{G}] = \bigvee_{a \in A} a[\mathcal{G}]$. But this is a contradiction to the simultaneous falsifiability of $D[\mathcal{G}]$ and C, and so not only does it hold that $\mathit{Vars}(A) \subseteq \mathit{Vars}(D)$ but A is in fact a subclause of D as claimed. From this we see that we can add the clauses $\mathcal{G}^{-1}(C)$ to $\mathbb{D}_t$ using axiom download and weakening. After applying a weakening step we immediately delete the old clause. Hence, the additional weakening might increase the space by at most one. It follows that the space of the intermediate configurations need never exceed $Sp(\mathbb{D}_{t+1}) + 1$.

It remains to argue that $\mathbb{D}_{t+1}$ can be derived from $\mathbb{D}_t$ when $\mathbb{C}_{t+1}$ is obtained from $\mathbb{C}_t$ by an inference step adding a new clause $C \in \mathbb{C}_{t+1} \setminus \mathbb{C}_t$. This is stated in the following two claims regarding applications of the resolution and weakening rules.

Claim 4.7. *Every clause $D \in \mathcal{G}^{-1}(C)$ can be derived from $\mathcal{G}^{-1}(C \vee x) \cup \mathcal{G}^{-1}(C \vee \overline{x})$ by a homogeneous resolution derivation of width w and depth $w + 1$.*

Claim 4.8. *For any two clauses B and C with $B \subseteq C$ it holds that every clause $D \in \mathcal{G}^{-1}(C)$ can be derived from $\mathcal{G}^{-1}(B)$ by a homogeneous derivation of width w and depth $w + 1$.*

Taking these two claims on faith for now, let us see how they allow us to conclude the proof of the lemma. Since the depth of a refutation provides an upper bound on the clause space by Observation 2.4, it follows that in both cases we can derive all clauses in the clause set $\mathcal{G}^{-1}(C)$ one by one by using additional space $w + 3$ to perform the derivations in depth $w + 1$.

This shows that F has a homogeneous resolution refutation π' of width w and clause space $Sp(\pi') \leq 2^w Sp(\pi) + w + 3$, which establishes the lemma. □

We proceed to establish Claims 4.7 and 4.8.

Proof of Claim 4.7. Recall that by Lemma 4.4 the subgraph $\mathcal{G}_C := \mathcal{G} \setminus \gamma(C)$ is an $(r/2, 1)$-boundary expander and that for $\ker(\gamma(C \vee x)) = \ker(\gamma(C \vee \overline{x}))$ we have $|\ker(\gamma(C \vee x))| \leq W(C \vee x) \leq w \leq r/2$. Therefore, we can apply Lemma 4.3 to the set $K = \ker(\gamma(C \vee x)) \setminus \ker(\gamma(C))$ to obtain an ordering $u_1, \ldots, u_\ell$ of K satisfying

$$\mathcal{N}^{\mathcal{G}_C}(u_i) \setminus \mathcal{N}^{\mathcal{G}_C}(\{u_1, \ldots, u_{i-1}\}) \neq \emptyset \ . \tag{4.6}$$

For $0 \leq i \leq \ell$ we let

$$K^i := \big(\ker(\gamma(C)) \cap \ker(\gamma(C \vee x))\big) \cup \{u_j \mid 1 \leq j \leq i\} \tag{4.7}$$

so that $K^\ell = \ker(\gamma(C \vee x))$ and $K^0 \subseteq \ker(\gamma(C))$, and define

$$\mathbb{D}^i := \{D \mid \mathit{Vars}(D) = K^i;\ D[\mathcal{G}] \text{ and } C \text{ are simultaneously falsifiable}\} \ . \tag{4.8}$$

Observe that

$$\begin{aligned} & \mathcal{G}^{-1}(C \vee x) \cup \mathcal{G}^{-1}(C \vee \overline{x}) \\ = \ & \{D \mid \mathit{Vars}(D) = \ker(\gamma(C \vee x)); D[\mathcal{G}] \ \&\ C \vee x \text{ simultaneously falsifiable}\} \\ & \cup \{D \mid \mathit{Vars}(D) = \ker(\gamma(C \vee \overline{x})); D[\mathcal{G}] \ \&\ C \vee \overline{x} \text{ simultaneously falsifiable}\} \\ = \ & \{D \mid \mathit{Vars}(D) = \ker(\gamma(C \vee x));\ D[\mathcal{G}] \ \&\ C \text{ simultaneously falsifiable}\} \\ = \ & \mathbb{D}^\ell \end{aligned} \tag{4.9}$$

and that every clause in $\mathcal{G}^{-1}(C)$ is subsumed by a clause in $\mathbb{D}^0$ since $K^0 \subseteq \ker(\gamma(C))$. Thus, we are done if we can derive all clauses in $\mathbb{D}^0$ from the clauses in $\mathbb{D}^\ell$.

We do so inductively: for $i = \ell, \ell - 1, \ldots, 2, 1$ we can obtain any clause $D \in \mathbb{D}^{i-1}$ by an application of the homogeneous resolution rule to the clauses $D \vee u_i$ and $D \vee \overline{u}_i$, which we claim are both available in $\mathbb{D}^i$. What remains to show is that $D \in \mathbb{D}^{i-1}$ indeed implies that $\{D \vee u_i, D \vee \overline{u}_i\} \subseteq \mathbb{D}^i$. To argue this, note that by the definition of $\mathbb{D}^{i-1}$ in (4.8) there is a (partial) truth value assignment α that simultaneously falsifies $D[\mathcal{G}]$ and C. By (4.6) we are guaranteed that $\mathcal{N}^{\mathcal{G}_C}(u_i) \setminus \mathit{Vars}(D[\mathcal{G}]) = \mathcal{N}^{\mathcal{G}_C}(u_i) \setminus \mathcal{N}^{\mathcal{G}}(K^{i-1})$ has a nonempty intersection with $V \setminus \gamma(C)$, the right-hand side of the expander $\mathcal{G}_C$. Hence, we can extend α and set the variables in $\mathcal{N}^{\mathcal{G}_C}(u_i) \setminus \big(\mathit{Vars}(D[\mathcal{G}]) \cup \mathit{Vars}(C)\big) \supseteq \mathcal{N}^{\mathcal{G}_C}(K^i) \setminus \mathcal{N}^{\mathcal{G}_C}(K^{i-1}) \neq \emptyset$ to appropriate values so that the parity $\bigoplus_{v \in \mathcal{N}(u_i)} \alpha(v)$ is even and thus $(D \vee u_i)[\mathcal{G}] = D[\mathcal{G}] \vee u_i[\mathcal{G}]$ is falsified, and we do so without assigning any variables in C, which therefore remains falsified. In an analogous fashion, by instead ensuring that the parity $\bigoplus_{v \in \mathcal{N}(u_i)} \alpha(v)$ is odd we get a falsifying assignment for $(D \vee \overline{u}_i)[\mathcal{G}] \vee C$. Hence, by (4.8) it holds that $D \vee u_i$ and $D \vee \overline{u}_i$ both appear in $\mathbb{D}^i$.

Finally, to get from $\mathbb{D}^0$ to $\mathcal{G}^{-1}(C)$ we might need an extra weakening step as observed above. The total depth of the whole derivation is at most $\ell + 1 \leq w + 1$. □

Proof of Claim 4.8. Notice that if $\ker(\gamma(B)) \subseteq \ker(\gamma(C))$, this claim would be easy to establish, but as noted above we have no guarantee that this is the case. Instead, we apply a proof strategy similar to the one for the previous claim. We again have that $\mathcal{G}_C := \mathcal{G} \setminus \gamma(C)$ is an $(r/2, 1)$-boundary expander, so that we can apply the peeling lemma to the left-hand vertex set $\ker(\gamma(B)) \setminus \ker(\gamma(C))$ to obtain an ordering $u_1, \ldots, u_\ell$ of its vertices satisfying $\mathcal{N}^{\mathcal{G}_C}(u_i) \setminus \mathcal{N}^{\mathcal{G}_C}(\{u_1, \ldots, u_{i-1}\}) \neq \emptyset$. For $0 \leq i \leq \ell$ we let

$$K^i := \big(\ker(\gamma(C)) \cap \ker(\gamma(B))\big) \cup \{u_j \mid 1 \leq j \leq i\} \tag{4.10}$$

and as before define

$$\mathbb{D}^i := \{D \mid \mathit{Vars}(D) = K^i;\ D[\mathcal{G}] \text{ and } C \text{ are simultaneously falsifiable}\}\ . \tag{4.11}$$

Note that $\mathbb{D}^\ell \subseteq \mathcal{G}^{-1}(B)$, because if $D[\mathcal{G}]$ and C are simultaneously falsifiable, then $D[\mathcal{G}]$ and $B \subseteq C$ are certainly simultaneously falsifiable. Hence, we can obtain $\mathbb{D}^\ell$ from $\mathcal{G}^{-1}(B)$ by just erasing clauses. Once more, we apply the peeling argument in an inductive fashion and derive any $D \in \mathbb{D}^{i-1}$ from $D \vee u_i$ and $D \vee \overline{u}_i$ appearing in $\mathbb{D}^i$, exactly as in the proof of Claim 4.7. In the end, we can infer any clause in $\mathcal{G}^{-1}(C)$ from $\mathbb{D}^0$ because every clause in $\mathcal{G}^{-1}(C)$ can be seen to be a weakening of some clause in $\mathbb{D}^0$. ∎

We can now combine the construction in Lemma 4.6 with the existence of good boundary expanders in Lemma 4.5 to prove the hardness condensation in Lemma 4.1.

Proof of Lemma 4.1. Given $\varepsilon > 0$ and $k \in \mathbb{N}^+$ we choose $\delta := \frac{\varepsilon}{10k}$. Note that we can assume $\varepsilon \leq 1/2$ since otherwise the lemma is vacuous. Suppose ℓ and n are parameters such that $k \leq \ell \leq n^{\frac{1}{2}-\varepsilon}$ and let F be an unsatisfiable CNF formula over $N = \lfloor n^{\delta\ell} \rfloor$ variables that can be refuted in width k. To apply Lemma 4.5 we set $d_0 := \lfloor \frac{5}{\varepsilon} \rfloor \geq 10$. Noting that $d_0 = \frac{5}{\varepsilon} - \eta$ for $\eta < 1 \leq \frac{1}{2\varepsilon}$, we get that $d_0 \geq \frac{9}{2\varepsilon}$, and hence $\delta + \frac{1}{d_0} \leq \frac{\varepsilon}{10k} + \frac{2\varepsilon}{9} < \frac{\varepsilon}{2}$. We choose the degree of the expander to be $d := \lfloor \frac{\ell}{2k} \rfloor$ and set the size guarantee for expanding left vertex sets to $r := 2\ell \log n$. By the bound on ℓ we have $d \leq \ell \leq n^{\frac{1}{2}-\varepsilon}$. Furthermore, we adjust n_0 obtained from Lemma 4.5 to be large enough so that $r \leq 2n^{\frac{1}{2}-\varepsilon} \log n \leq n^{\frac{1}{2}}$ for all $n \geq n_0$.

Now we have two cases. The first, and interesting, case is when $d \geq d_0$ holds. Then Lemma 4.5 guarantees that there exists an $N \times n$ $(r, d, 2)$-boundary expander $\mathcal{G}$. Applying XORification with respect to $\mathcal{G}$, we obtain a CNF formula $F[\mathcal{G}]$ with n variables. By Observation 2.3 it holds that $F[\mathcal{G}]$ has a resolution refutation in width $2dk \leq \ell$. Now suppose that $\pi : F[\mathcal{G}] \vdash \bot$ is a refutation in width $w \leq \ell \log n = r/2$. Then Lemma 4.6 yields a refutation π' of the original formula F in width w and space $2^w \cdot \mathit{Sp}(\pi) + w + 3$, and so if this space has to be at least s we get that $\mathit{Sp}(\pi) \geq (s - w - 3)2^{-w}$.

The second case is when $d < d_0$. Then we do not actually need any XORification but can use the original formula. Formally, let $\mathcal{G} = (U \mathbin{\dot\cup} (V \cup V'), E)$ be a matching between two sets U and V of size $|U| = |V| = N$ plus some isolated vertices V' on the right-hand side such that $|V \cup V'| = n$. To check that this is well defined, we have to verify that $N \leq n$, which follows from the calculations $N = \lfloor n^{\delta\ell} \rfloor \leq \lfloor n^{\frac{\varepsilon}{10k} 2k(d+1)} \rfloor \leq \lfloor n^{\frac{\varepsilon}{10k} 2kd_0} \rfloor \leq \lfloor n^{\frac{\varepsilon}{10k} 2k\frac{5}{\varepsilon}} \rfloor = n$, where we used $d_0 \geq d + 1 \geq \frac{\ell}{2k}$. In this somewhat convoluted way we obtain $F[\mathcal{G}] = F$ (plus some left-over dummy variables) and we have $W(F[\mathcal{G}] \vdash \bot) = W(F \vdash \bot) = k \leq \ell$ as well as $\mathit{Sp}(\pi) \geq s \geq (s - w - 3)2^{-w}$. The lemma follows. ∎

5. Concluding remarks. In this paper we prove that there are CNF formulas over n variables exhibiting an $n^{\Omega(w)}$ clause space lower bound for resolution refutations in width w. This lower bound is optimal (up to constants in the exponent) as every refutation in width w can be carried out in length, and hence space, at most $n^{\mathrm{O}(w)}$. Our lower bounds do not only hold for the minimal refutation width w but remain valid for any refutations in width asymptotically smaller than $w \log n$. Measured in terms of the number of variables n, this is a major improvement over the previous space-width trade-off result in [8], and provides another example of trade-offs in the supercritical regime above worst-case recently identified in [34].

Regarding possible future research directions, a first open problem is whether the range of applicability can be extended even further so that the space lower bound holds true up to width $o(n)$. It is clear that the lower bound has to break down at some point, since if one is allowed maximal width n any formula can be refuted in clause space $n+2$ [25]. A supercritical trade-off on resolution proof depth over width ranging from w all the way up to $n^{1-\epsilon}/w$ was shown in [34], suggesting that the above goal might not be completely out of reach.

Another intriguing open problem is to prove space trade-offs that are superlinear not only in terms of the number of variables but measured also in formula size. Such lower bounds cannot be obtained by the techniques used in this paper, but they are likely to exist as the following argument shows (see [27] for a more detailed discussion). Suppose that every refutation in width $w(n)$ can be transformed into a refutation that has width $w(n)$ and clause space polynomial in the size of the formula. Then we can find such a refutation nondeterministically in polynomial space by keeping the current configuration in memory and guessing the inference steps. Thus, by Savitch's theorem, finding refutations of width $w(n)$ would be in deterministic PSPACE. On the other hand, it has been shown by the first author that the problem of finding resolution refutations of bounded width is EXPTIME-complete [14]. Hence, unless EXPTIME = PSPACE there are formulas where every refutation of minimal width needs clause space that is superpolynomial in the size of the formula.

Finally, it would be interesting to study if the supercritical trade-offs between clause space and width in resolution shown in this paper could be extended to similar trade-offs between monomial space and degree for polynomial calculus or polynomial calculus resolution as defined in [1, 21].

Appendix A. Proofs of two lemmas concerning boundary expanders.

In this appendix we give proofs for Lemmas 4.4 and 4.5. As already mentioned, most of this material appears in a similar form in [34] (although the exact parameters are slightly different), and there is also a substantial overlap with analogous technical lemmas in [16]. In fact, Lemma 4.4 is exactly as stated in in [16], but we present a proof below to give a self-contained exposition of our version of Razborov's hardness condensation technique adapted to general resolution.

LEMMA 4.4 (restated). *Let $\mathcal{G}$ be a bipartite $(r,2)$-boundary expander. Then for every right vertex set $V' \subseteq V$ of size $|V'| \leq r/2$ there exists a superset $\gamma(V') \supseteq V'$ such that $\left|\ker\left(\gamma\left(V'\right)\right)\right| \leq \left|V'\right|$ and the induced subgraph $\mathcal{G} \setminus \gamma(V')$ is an $(r/2,1)$-boundary expander.*

Proof. With assumptions as in the lemma, let $\mathcal{G} = (U \dot{\cup} V, E)$ be an $(r,2)$-boundary expander, and let $V' \subseteq V$ be a right vertex set of size $|V'| \leq r/2$. We will construct an increasing sequence of right vertex sets $V' = V_0 \subset V_1 \subset \cdots \subset V_\tau$ such that for $\gamma(V') = V_\tau$ it holds that $\mathcal{G} \setminus V_\tau$ is an $(r/2,1)$-boundary expander.

If $\mathcal{G} \setminus V_0$ is an $(r/2,1)$-boundary expander, then we can stop right away, but otherwise there must exist a left vertex set U_1 of size at most $r/2$ such that $\left|\partial^{\mathcal{G}\setminus V_0}(U_1)\right| \leq |U_1|$. Delete $\mathcal{N}^{\mathcal{G}\setminus V_0}(U_1)$ and $\ker\left(\mathcal{N}^{\mathcal{G}\setminus V_0}(U_1)\right)$ from $\mathcal{G} \setminus V_0$. If now the resulting graph is an $(r/2,1)$-boundary expander, then we are done, but otherwise we repeat this process and iteratively delete vertex sets that violate the expansion requirements. Formally, for $i \geq 1$ we let U_i be any left vertex set of size at most $r/2$ such that

$\left|\partial^{\mathcal{G}\setminus V_{i-1}}(U_i)\right| \leq |U_i|$, where we set

$$V_i := V_0 \cup \bigcup_{j=1}^{i} \mathcal{N}^{\mathcal{G}}(U_j) \tag{A.1}$$

(and where we note that what is deleted at the ith step is $\mathcal{N}^{\mathcal{G}\setminus V_{i-1}}(U_i)$ together with the kernel of this right vertex set $\ker\left(\mathcal{N}^{\mathcal{G}\setminus V_{i-1}}(U_i)\right)$, so that after the ith step all of $\mathcal{N}^{\mathcal{G}}(U_i)$ and $\ker(\mathcal{N}^{\mathcal{G}}(U_i))$ has been removed from the graph).

Since all sets U_i constructed above are nonempty, this process must terminate for some $i = \tau$ and the resulting graph $\mathcal{G} \setminus V_\tau$ is then an $(r/2, 1)$-boundary expander (if nothing else, an empty graph without vertices vacuously satisfies the expansion condition). However, we need to check that the condition $|\ker(V_\tau)| \leq |V_0|$ holds. This follows from the next claim.

Claim A.1. *Let $V_{-1} = U_0 = \emptyset$ and suppose that $i \geq 0$. Then for U_i and V_i as constructed above we have the following properties:*

1. *For all U' such that $\ker(V_{i-1}) \cup U_i \subseteq U' \subseteq \ker(V_i)$ it holds that $\left|\partial^{\mathcal{G}}(U') \setminus V_0\right| \leq |\ker(V_i)|$.*
2. *The kernel of V_i has size $|\ker(V_i)| \leq |V_0|$.*

We establish Claim A.1 by induction. For the base case $i = 0$, Property 1 holds since $U' \subseteq \ker(V_0)$ implies that $\partial^{\mathcal{G}}(U') \subseteq V_0$. For Property 2, suppose that $|\ker(V_0)| \leq r$. Then by the expansion of $\mathcal{G}$ we have that $2|\ker(V_0)| \leq |\partial^{\mathcal{G}}(\ker(V_0))|$, and in combination with $\partial^{\mathcal{G}}(\ker(V_0)) \subseteq V_0$ this implies $|\ker(V_0)| \leq \frac{1}{2}|V_0|$. If instead $|\ker(V_0)| > r$, then we can find a subset $U' \subseteq \ker(V_0)$ of size $|U'| = r$ for which it holds by expansion that $|\partial^{\mathcal{G}}(U')| \geq 2r$. But this is a contradiction since as argued above we should have $\left|\partial^{\mathcal{G}}(U')\right| \leq |V_0| \leq r/2$.

For the induction step, suppose that Properties 1 and 2 both hold for $i - 1$. Let us write $U^* = \ker(V_{i-1}) \cup U_i$ and consider any U' such that $U^* \subseteq U' \subseteq \ker(V_i)$. We claim that every vertex in $\partial^{\mathcal{G}}(U')$ is either in the boundary $\partial^{\mathcal{G}}(U^*)$ or is a member of V_0. To see why this is so, observe that since $U' \subseteq \ker(V_i)$ we have $\partial^{\mathcal{G}}(U') \subseteq V_i = V_0 \cup \bigcup_{j=1}^{i} \mathcal{N}^{\mathcal{G}}(U_j)$. Furthermore, note that $\bigcup_{j=1}^{i} U_j \subseteq U^* \subseteq U'$ holds (which is due to the fact that $\mathcal{N}(\ker(V')) \subseteq V'$ for any V'). Hence, for any $v \in \partial^{\mathcal{G}}(U') \setminus V_0$ it must be the case that $v \in \bigcup_{j=1}^{i} \mathcal{N}^{\mathcal{G}}(U_j)$, and so the unique neighbor of v on the left is contained in $\bigcup_{j=1}^{i} U_j$ and therefore also in U^*, implying that $v \in \partial(U^*)$. From this we can conclude that

$$\partial^{\mathcal{G}}(U') \setminus V_0 \subseteq \partial^{\mathcal{G}}(U^*) \setminus V_0 \tag{A.2}$$

(i.e., $U' \supseteq U^*$ does not add any new boundary elements outside of V_0 on top of what is already provided by U^*). We will use this to show that

$$\left|\partial^{\mathcal{G}}(U^*) \setminus V_0\right| = \left|\partial^{\mathcal{G}}(\ker(V_{i-1}) \cup U_i) \setminus V_0\right| \leq |\ker(V_i)| \tag{A.3}$$

in order to prove Property 1.

Towards this goal, observe that by construction it holds that every vertex in $V_{i-1} \setminus V_0$ has at least one neighbor in $\ker(V_{i-1})$. It follows that for the vertex set $U^* = \ker(V_{i-1}) \cup U_i$ all new boundary vertices in $\partial^{\mathcal{G}}(U^*) \setminus \partial^{\mathcal{G}}(\ker(V_{i-1}))$ are either from V_0 or from the boundary $\partial^{\mathcal{G}\setminus V_{i-1}}(U_i)$ of U_i that lies outside of V_{i-1}. Therefore, we have

$$\partial^{\mathcal{G}}(U^*) \setminus V_0 = \partial^{\mathcal{G}}\left(\ker(V_{i-1}) \cup U_i\right) \setminus V_0 \subseteq \left(\partial^{\mathcal{G}}(\ker(V_{i-1})) \setminus V_0\right) \dot\cup\, \partial^{\mathcal{G}\setminus V_{i-1}}(U_i) \ . \tag{A.4}$$

Since we have chosen U_i so that it does not satisfy the expansion condition we know that

$$\left|\partial^{\mathcal{G}\setminus V_{i-1}}(U_i)\right| \leq |U_i| \tag{A.5}$$

and by the inductive hypothesis for Property 1 it holds that

$$\left|\partial^{\mathcal{G}}(\ker(V_{i-1})) \setminus V_0\right| \leq |\ker(V_{i-1})| \ . \tag{A.6}$$

Combining (A.2) with (A.4)–(A.6) we conclude that

$$\begin{aligned}\left|\partial^{\mathcal{G}}(U') \setminus V_0\right| &\leq \left|\partial^{\mathcal{G}}(\ker(V_{i-1}) \cup U_i) \setminus V_0\right| \\ &\leq \left|\left(\partial^{\mathcal{G}}(\ker(V_{i-1})) \setminus V_0\right)\right| + \left|\partial^{\mathcal{G}\setminus V_{i-1}}(U_i)\right| \leq |\ker(V_{i-1})| + |U_i| \leq |\ker(V_i)| \ ,\end{aligned} \tag{A.7}$$

where the last inequality holds since $\ker(V_{i-1})$ and U_i are disjoint subsets of $\ker(V_i)$. This completes the inductive step for Property 1.

To show Property 2, let us first assume that $|\ker(V_i)| \leq r$. Then by the expansion properties of $\mathcal{G}$ together with Property 1 applied to the set $U' = \ker(V_i)$ we have

$$2|\ker(V_i)| \leq \left|\partial^{\mathcal{G}}(\ker(V_i))\right| \leq |V_0| + |\ker(V_i)| \ , \tag{A.8}$$

from which it follows that

$$|\ker(V_i)| \leq |V_0| \ . \tag{A.9}$$

If instead $|\ker(V_i)| > r$, then by the inductive hypothesis we know that $|\ker(V_{i-1})| \leq |V_0| \leq r/2$ and by construction we have $|U_i| \leq r/2$. Therefore, there must exist a vertex set U' of size r satisfying the condition $\ker(V_{i-1}) \cup U_i \subseteq U' \subseteq \ker(V_i)$ in Property 1. From the expansion properties of $\mathcal{G}$ we conclude that $|\partial(U')| \geq 2r$, which is a contradiction because for sets U' satisfying the conditions in Property 1 we derived (A.7), which implies that $|\partial(U')| \leq |V_0| + |\ker(V_{i-1})| + |U_i| \leq 3r/2$. The claim follows by the induction principle. ∎

LEMMA 4.5 (restated). *Fix constants $\varepsilon, \delta > 0$ and $d_0 \geq 2$ such that $\delta + \frac{1}{d_0} < \varepsilon/2$. Then there exists an $n_0 \in \mathbb{N}^+$ such that for all n, d, and r satisfying $n \geq n_0$, $d_0 \leq d \leq n^{1/2-\varepsilon}$, and $r \leq n^{1/2}$ there are $\lfloor n^{\delta d} \rfloor \times n$ $(r, d, 2)$-boundary expanders.*

Proof. Let U and V be two disjoint sets of vertices of size $|U| = N = \lfloor n^{\delta d} \rfloor$ and $|V| = n$. For every $u \in U$ we choose d times a neighbor $v \in V$ uniformly at random with repetitions. This gives us a bipartite graph $\mathcal{G} = (U \dot\cup V, E)$ of left-degree at most d. In what follows, we show that $\mathcal{G}$ is almost surely an $(r, d, 2)$-boundary expander as $n \to \infty$.

First, note that for every set $U' \subseteq U$ all neighbors $v \in \mathcal{N}(U') \setminus \partial(U')$ not in the boundary of U' have at least two neighbors in U'. Since there are at most $d|U'| - |\partial(U')|$ edges between U' and $\mathcal{N}(U') \setminus \partial(U')$, it follows that $|\mathcal{N}(U') \setminus \partial(U')| \leq (d|U'| - |\partial(U')|)/2$ and hence

$$\begin{aligned}|\mathcal{N}(U')| &= \left|\mathcal{N}(U') \cap \partial(U')\right| + \left|\mathcal{N}(U') \setminus \partial(U')\right| \\ &\leq |\partial(U')| + \frac{d|U'| - |\partial(U')|}{2} = \frac{d|U'| + |\partial(U')|}{2} \ .\end{aligned} \tag{A.10}$$

If $\mathcal{G}$ is not an $(r, d, 2)$-boundary expander, then there is a set U' of size $\ell \leq r$ that has a boundary $\partial(U')$ of size at most 2ℓ and from (A.10) it follows that $|\mathcal{N}(U')| \leq (1+d/2)\ell$. By a union bound argument we obtain

$$\Pr[\mathcal{G} \text{ is not an } (r, d, 2)\text{-boundary expander}] \tag{A.11a}$$

$$\leq \sum_{\ell=1}^{r} \sum_{U' \subseteq [N];\, |U'|=\ell} \Pr\left[|\partial(U')| \leq 2\ell\right] \tag{A.11b}$$

$$\leq \sum_{\ell=1}^{r} \binom{N}{\ell} \Pr\left[|\mathcal{N}(U')| \leq (1+d/2)\ell \text{ for some fixed } |U'| = \ell\right] \tag{A.11c}$$

$$\leq \sum_{\ell=1}^{r} \binom{N}{\ell} \binom{n}{(1+d/2)\ell} \left(\frac{(1+d/2)\ell}{n}\right)^{d\ell} \tag{A.11d}$$

$$\leq \sum_{\ell=1}^{r} N^{\ell} \left(\frac{en}{(1+d/2)\ell}\right)^{(1+d/2)\ell} ((1+d/2)\ell)^{d\ell}\, n^{-d\ell} \tag{A.11e}$$

$$= \sum_{\ell=1}^{r} N^{\ell} (en)^{(1+d/2)\ell} ((1+d/2)\ell)^{(d/2-1)\ell}\, n^{-d\ell} \tag{A.11f}$$

$$\leq \sum_{\ell=1}^{r} n^{\delta d\ell} (en)^{(1+d/2)\ell} ((1+d/2)\ell)^{(d/2-1)\ell}\, n^{-d\ell} \tag{A.11g}$$

$$= \sum_{\ell=1}^{r} n^{\delta d\ell} n^{\frac{\log e}{\log n}(1+d/2)\ell} n^{\frac{1}{\log n}\log\left((d/2+1)\ell\right)(d/2-1)\ell} n^{(-d/2+1)\ell} \tag{A.11h}$$

$$\leq \sum_{\ell=1}^{r} n^{\left(\frac{\log e}{\log n} d + \frac{1}{\log n}\log(dr)(d/2-1) - d/2 + 1 + \delta d\right)\ell} \tag{A.11i}$$

$$= \sum_{\ell=1}^{r} n^{\left(\frac{\log e}{\log n} + \frac{1}{\log n}\log(dr)(1/2-1/d) - 1/2 + 1/d + \delta\right) d\ell}\ , \tag{A.11j}$$

where to get from line (A.11d) to (A.11e) we used that $\binom{n}{k} \leq \left(\frac{en}{k}\right)^k$ for the Euler number e, from (A.11g) to (A.11h) we used that $n^{\log a/\log n} = a$, and from (A.11h) to (A.11i) that $d \geq d_0 \geq 2$ and $\ell \leq r$. In order to show that (A.11j) is bounded away from 1, it suffices to demonstrate that the expression

$$\frac{\log e}{\log n} + \frac{1}{\log n}\log(dr)(1/2 - 1/d) - 1/2 + 1/d + \delta \tag{A.12}$$

is negative and bounded away from zero. Set $\lambda = \varepsilon/2 - 1/d_0 - \delta > 0$ and choose $n_0 = 3^{2/\lambda}$. By the upper bounds on r and d it follows that

$$\log e/\log n + \log(dr)(1/2 - 1/d)/\log n - 1/2 + 1/d + \delta \tag{A.13a}$$

$$\leq \log e/\log n + (1/2)\log(n^{\frac{1}{2}-\varepsilon} n^{\frac{1}{2}})/\log n - 1/2 + 1/d + \delta \tag{A.13b}$$

$$= \log e/\log n - \varepsilon/2 + 1/d + \delta \tag{A.13c}$$

$$\leq \log e/\log n - \varepsilon/2 + 1/d_0 + \delta \tag{A.13d}$$

$$= \log e/\log n - \lambda \tag{A.13e}$$

$$\leq -\lambda/2\ , \tag{A.13f}$$

where the last inequality holds since $n \geq n_0 > e^{2/\lambda}$. It follows that the probability that $\mathcal{G}$ is not an $(r, d, 2)$-boundary expander is bounded by

$$\sum_{\ell=1}^{r} n^{\left(-\lambda/2\right) d\ell} \leq \sum_{\ell=1}^{r} n_0^{\left(-\lambda/2\right) d\ell} \leq \sum_{\ell=1}^{\infty} \left(\tfrac{1}{3}\right)^{d\ell} \leq \tfrac{1}{2}\ , \tag{A.14}$$

which establishes the lemma. ∎

Acknowledgments. We wish to thank Alexander Razborov for patiently explaining the hardness condensation technique in [34] during numerous and detailed discussions. We are also most grateful to the anonymous reviewers for very detailed comments that helped improve the manuscript considerably.

REFERENCES

[1] M. Alekhnovich, E. Ben-Sasson, A. A. Razborov, and A. Wigderson, *Space complexity in propositional calculus*, SIAM J. Comput., 31 (2002), pp. 1184–1211; preliminary version in STOC '00, https://doi.org/10.1137/S0097539700366735.

[2] A. Atserias and V. Dalmau, *A combinatorial characterization of resolution width*, J. Comput. System Sci., 74 (2008), pp. 323–334; preliminary version in CCC '03.

[3] A. Atserias, M. Lauria, and J. Nordström, *Narrow proofs may be maximally long*, ACM Trans. Comput. Logic, 17 (2016), 19; preliminary version in CCC '14.

[4] R. J. Bayardo, Jr. and R. Schrag, *Using CSP look-back techniques to solve real-world SAT instances*, in Proceedings of the 14th National Conference on Artificial Intelligence (AAAI '97), 1997, pp. 203–208.

[5] P. Beame, C. Beck, and R. Impagliazzo, *Time-space tradeoffs in resolution: Superpolynomial lower bounds for superlinear space*, SIAM J. Comput., 45 (2016), pp. 1612–1645; preliminary version in STOC '12, https://doi.org/10.1137/130914085.

[6] P. Beame, R. Karp, T. Pitassi, and M. Saks, *The efficiency of resolution and Davis–Putnam procedures*, SIAM J. Comput., 31 (2002), pp. 1048–1075; preliminary versions of these results appeared in FOCS '96 and STOC '98, https://doi.org/10.1137/S0097539700369156.

[7] C. Beck, J. Nordström, and B. Tang, *Some trade-off results for polynomial calculus*, in Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC '13), 2013, pp. 813–822.

[8] E. Ben-Sasson, *Size-space tradeoffs for resolution*, SIAM J. Comput., 38 (2009), pp. 2511–2525; preliminary version in STOC '02, https://doi.org/10.1137/080723880.

[9] E. Ben-Sasson and N. Galesi, *Space complexity of random formulae in resolution*, Random Structures Algorithms, 23 (2003), pp. 92–109; preliminary version in CCC '01.

[10] E. Ben-Sasson and J. Nordström, *Short proofs may be spacious: An optimal separation of space and length in resolution*, in Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS '08), 2008, pp. 709–718.

[11] E. Ben-Sasson and J. Nordström, *Understanding space in proof complexity: Separations and trade-offs via substitutions*, in Proceedings of the 2nd Symposium on Innovations in Computer Science (ICS '11), 2011, pp. 401–416.

[12] E. Ben-Sasson and A. Wigderson, *Short proofs are narrow—resolution made simple*, J. ACM, 48 (2001), pp. 149–169; preliminary version in STOC '99.

[13] P. Bennett, I. Bonacina, N. Galesi, T. Huynh, M. Molloy, and P. Wollan, *Space proof complexity for random* 3*-CNFs*, Inf. Comput., 255 (2017), pp. 165–176.

[14] C. Berkholz, *On the complexity of finding narrow proofs*, in Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS '12), 2012, pp. 351–360.

[15] C. Berkholz and J. Nordström, *Near-optimal lower bounds on quantifier depth and Weisfeiler-Leman refinement steps*, in Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science (LICS '16), 2016, pp. 267–276.

[16] C. Berkholz and J. Nordström, *Near-Optimal Lower Bounds on Quantifier Depth and Weisfeiler-Leman Refinement Steps*, Technical Report TR16-135, Electronic Colloquium on Computational Complexity (ECCC), 2016; preliminary version in LICS '16.

[17] A. Blake, *Canonical Expressions in Boolean Algebra*, Ph.D. thesis, University of Chicago, Chicago, IL, 1937.

[18] I. Bonacina, *Total space in resolution is at least width squared*, in Proceedings of the 43rd International Colloquium on Automata, Languages and Programming (ICALP '16), Leibniz International Proceedings in Informatics (LIPIcs) 55, 2016, 56.

[19] I. Bonacina, N. Galesi, and N. Thapen, *Total space in resolution*, in Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS '14), 2014, pp. 641–650.

[20] V. Chvátal and E. Szemerédi, *Many hard examples for resolution*, J. ACM, 35 (1988), pp. 759–768.

[21] M. Clegg, J. Edmonds, and R. Impagliazzo, *Using the Groebner basis algorithm to find proofs of unsatisfiability*, in Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC '96), 1996, pp. 174–183.

[22] S. A. Cook and R. Reckhow, *The relative efficiency of propositional proof systems*, J. Symbolic Logic, 44 (1979), pp. 36–50.
[23] M. Davis, G. Logemann, and D. Loveland, *A machine program for theorem proving*, Comm. ACM, 5 (1962), pp. 394–397.
[24] M. Davis and H. Putnam, *A computing procedure for quantification theory*, J. ACM, 7 (1960), pp. 201–215.
[25] J. L. Esteban and J. Torán, *Space bounds for resolution*, Inform. and Comput., 171 (2001), pp. 84–97, preliminary versions of these results appeared in STACS '99 and CSL '99.
[26] A. Haken, *The intractability of resolution*, Theoret. Comput. Sci., 39 (1985), pp. 297–308.
[27] A. Hertel, *Applications of Games to Propositional Proof Complexity*, Ph.D. thesis, University of Toronto, Toronto, ON, Canada, 2008, available at http://www.cs.utoronto.ca/~ahertel/.
[28] J. P. Marques-Silva and K. A. Sakallah, *GRASP: A search algorithm for propositional satisfiability*, IEEE Trans. Comput., 48 (1999), pp. 506–521, preliminary version in ICCAD '96.
[29] M. W. Moskewicz, C. F. Madigan, Y. Zhao, L. Zhang, and S. Malik, *Chaff: Engineering an efficient SAT solver*, in Proceedings of the 38th Design Automation Conference (DAC '01), 2001, pp. 530–535.
[30] J. Nordström, *Narrow proofs may be spacious: Separating space and width in resolution*, SIAM J. Comput., 39 (2009), pp. 59–121, preliminary version in STOC '06, https://doi.org/10.1137/060668250.
[31] J. Nordström, *Pebble games, proof complexity and time-space trade-offs*, Log. Methods Comput. Sci., 9 (2013), 3:15.
[32] J. Nordström and J. Håstad, *Towards an optimal separation of space and length in resolution*, Theory Comput., 9 (2013), pp. 471–557, preliminary version in STOC '08.
[33] A. A. Razborov, *An Ultimate Trade-Off in Propositional Proof Complexity*, Technical Report TR15-033, Electronic Colloquium on Computational Complexity (ECCC), 2015.
[34] A. A. Razborov, *A new kind of tradeoffs in propositional proof complexity*, J. ACM, 63 (2016), 16.
[35] J. A. Robinson, *A machine-oriented logic based on the resolution principle*, J. ACM, 12 (1965), pp. 23–41.
[36] N. Thapen, *A trade-off between length and width in resolution*, Theory Comput., 12 (2016), pp. 1–14.
[37] A. Urquhart, *Hard examples for resolution*, J. ACM, 34 (1987), pp. 209–219.