A note about k-DNF resolution

Massimo Lauria

Dipartimento di Scienze Statistiche — Sapienza Università di Roma, Italy
massimo.lauria@uniroma1.it

Abstract

In this note we show two results about k-DNF resolution. First we prove that resolution extended with parities of size k is strictly weaker than k-DNF resolution. Then we show the connection between tree-like k-DNF resolution and narrow dag-like resolution. This latter result is clearly implicit in [Krajíček, 1994] but this direct proof is focused on resolution and provides information about refutation width.

Keywords: Proof Complexity, k-DNF, resolution

1. Introduction

The field of *proof complexity* studies the length of refutations, i.e., proofs of unsatisfiability. If unsatisfiable formulas with no short refutations exist then NP is different from co-NP, which immediately implies that P≠NP [13]. While this may be quite hard to prove, we can tackle the simpler task of proving that a refutation must be long when written according to a specific *proof system*, i.e., a language in which proofs are *efficiently verifiable*. The most studied proof system in literature is resolution [9, 30]. Relatively simple, resolution still captures many important algorithms that check satisfiability (i.e., SAT solvers). Indeed DPLL algorithms on unsatisfiable formulas compute tree-shaped resolution refutations [15, 14, 6, 8], and the core of the most efficient SAT solvers nowadays is based on *conflict driven clause learning (CDCL)* [5, 25, 26], another technique that produces resolution proofs [29, 3, 16]. The length of resolution refutations has been studied in great details, and thus gives information on the running time of these algorithms. The seminal paper [20] proves the first super-polynomial legth lower bound for resolution, on the pigeonhole principle formula. Lower bounds for many more formulas followed once general lower bound techniques were discovered [23, 7].

The connection between resolution and SAT solvers led to the development of the notion of proof space to model memory usage (a very precious resource in concrete applications) [17, 1]. The most effective way to analyze it is to study of memory allocations in pebbling games [28]. While it is hard to know how much memory is needed to win these games in general [19, 12], it is possible to use and combine well understood cases to get space lower bounds and length vs space trade-offs results [27]. Another complexity parameter for resolution proofs is the "width", i.e., the maximum size among proof lines. While this parameter is not directly related to the performance of

SAT solvers, it turns out that it is an excellent proxy for the study of proof length [7] or, to a lesser degree, space [2, 18].

Beyond the study of resolution, a large amount of work has been done to prove lower bounds for proof systems as strong as possible, and to determine, given two proof systems, whether they are equivalent, incomparable or one stronger than the other. In this note we focus on two extensions of resolution where proof lines are disjunction of either parities (the $\mathsf{Res}(\oplus_k)$ proof system) or conjunctions (the $\mathsf{Res}(k)$ proof system) of arity k. $\mathsf{Res}(k)$ has been studied first in [21]. In [31] they show that for every constant k there is a formula that requires refutations of exponential length in $\mathsf{Res}(k)$ but has refutations of polynomial length in $\mathsf{Res}(k+1)$, hence they show a hierarchy of systems with $\mathsf{Res}(1)$ (i.e., resolution) at the bottom.

Our contribution. The formula separating $\mathsf{Res}(k)$ from $\mathsf{Res}(k+1)$ is actually easy to refute in $\mathsf{Res}(\oplus_{k+1})$, which is a subsystem of $\mathsf{Res}(k+1)$. It is natural to explore the relation between $\mathsf{Res}(\oplus_k)$ and $\mathsf{Res}(k)$. The first contribution of this note is to show that $\mathsf{Res}(\oplus_k)$ is actually strictly weaker than $\mathsf{Res}(k)$. The second contribution of this note is the essential equivalence between "small" tree-like proofs in $\mathsf{Res}(k)$ and "narrow" proofs in resolution. This is interesting because in general $\mathsf{Res}(k)$ is stronger than resolution. A version of this claim was proved already for bounded-depth Frege proof systems [22], but our proof focuses on resolution and can take in account the width of the proofs. Indeed this new proof has been useful to some papers in bounded arithmetic [11, 4], which cite a preliminary unpublished version of this note [24].

This note is organized as follows. In Section 2 we give the necessary definitions and notation. In Section 3 we show the first of our result, namely that $\mathsf{Res}(\oplus_k)$ is a strictly weaker proof system than $\mathsf{Res}(k)$. In Section 4 we give a proof of the correspondence between small tree-like $\mathsf{Res}(k)$ refutations with resolution refutations of small width.

2. Preliminaries

A literal is an occurrence of either a boolean variable or its negation. A clause is a disjunction of literals. A CNF formula F over a set of n variables is a conjunction of clauses. The *width* of F is the largest number of literals in a clause of F. If the width of F is at most w then F is a w-CNF. We denote as $\mathsf{Vars}(F)$ the set of variables of F. When we denote the i-th clause of F as C_i, we write F as $\bigwedge_i C_i$.

A *derivation* (or *proof*) of length τ of a formula D from the CNF formula F is a sequence of proof lines $D_1, \ldots, D_\tau$ such that $D_\tau = D$ and such that any formula in the sequence is either one of the C_i, an *axiom* of the proof system, or it is *inferred* from previous formulas in the sequence using an inference rule of the proof system. A proof of D from F is also denoted as $F \vdash D$. A *refutation* of F is a derivation of the empty disjunction, often denoted as $\bot$, from it. The axioms, inference rules and formulas allowed in a derivation depend on the proof system considered. For all the proof systems in this note the proof lines can only contain variables in $\mathsf{Vars}(F)$. In *resolution* (Res) every proof line is a clause, the inference rules are

$$\text{(Cut)}\ \frac{A \vee x \qquad B \vee \neg x}{A \vee B} \qquad \text{(Weakening)}\ \frac{A}{A \vee B}\ , \tag{1}$$

where A and B are clauses, and the axioms have the form $x \vee \neg x$ for $x \in \mathsf{Vars}(F)$. The *width* of a Res derivation is the maximum number of literals among all its clauses. $\mathsf{Res}(k)$ is an extension of resolution where proof lines are k-DNFs (i.e., disjunctions of conjunctions of at most k literals). Inference rules are

$$\text{(Cut)}\ \frac{A \vee l_1 l_2 \cdots l_s \qquad B \vee \neg l_1 \vee \neg l_2 \vee \ldots \vee \neg l_s}{A \vee B} \qquad \text{(Weakening)}\ \frac{A}{A \vee B}\ , \tag{2}$$

where A and B are k-DNFs and $1 \leq s \leq k$. Furthermore $\mathsf{Res}(k)$ has an axiom

$$l_1 l_2 \cdots l_s \vee \neg l_1 \vee \neg l_2 \vee \ldots \vee \neg l_s \tag{3}$$

for every s literals $l_1, l_2, \cdots, l_s$, with $1 \leq s \leq k$. It is easy to see that $\mathsf{Res}(1)$ is Res.

It is useful to consider proofs with *tree-like* structure, i.e., any line in the proof can be used as the premise of an inference step at most once. If another occurrence of the same formula is needed, it has to be re-derived from scratch. Indeed such a proof can be seen as a rooted tree where axioms and initial clauses are at the leaf nodes, and the root is the final clause to be derived. A proof system is called *dag-like* is no constraint on the structure of the proof is imposed (indeed the refutation looks like a directed acyclic graph). We denote tree-like resolution as Res^* and tree-like $\mathsf{Res}(k)$ as $\mathsf{Res}^*(k)$.

Fact 1. $\mathsf{Res}^*(k)$ *has implicational completeness, i.e., if F logically implies the k-DNF formula D and* $\mathsf{Vars}(D) \subseteq \mathsf{Vars}(F)$*, then there is a* $\mathsf{Res}^*(k)$ *proof $F \vdash D$.*

In this note we will also consider the relatively non-standard proof system $\mathsf{Res}(\oplus_k)$ where each proof line is a disjunction of terms of the form $x_{i_1} \oplus x_{i_2} \oplus \cdots \oplus x_{i_s} = b$ where $1 \leq s \leq k$ and $b \in \{0, 1\}$. We name $\vee\oplus_k$*-type formula* this type of formula. In this system we use $x_i = 1$ and $x_i = 0$ to represent positive and negative literals, respectively. Axioms are formulas $(x_{i_1} \oplus \cdots \oplus x_{i_s} = 0) \vee (x_{i_1} \oplus \cdots \oplus x_{i_s} = 1)$ and the inference rules are

$$\text{(Cut)}\ \frac{A \vee \bigoplus_{i \in I} x_i = b_1 \qquad B \vee \bigoplus_{i \in J} x_i = b_2}{A \vee B \vee \bigoplus_{i \in I \Delta J} x_i = b_1 \oplus b_2} \qquad \text{(Weakening)}\ \frac{A}{A \vee B}\ , \tag{4}$$

where $|I|$ and $|J|$ are at most k and the equations $0 = 1$ are removed from the result. When k is constant $\mathsf{Res}(k)$ efficiently simulates $\mathsf{Res}(\oplus_k)$ because each parity can be represented as a k-DNF formula of length $\leq 2^{k-1}$ and implicational completeness allows the simulation of a cut in $O(2^k)$ steps.

Given a (partial) assignment $\rho\colon \{x_i\}_{i \in I \subseteq [n]} \to \{0, 1\}$ and a formula F we denote as $F\restriction_\rho$ the formula obtained by assigning into F the variables in the domain of ρ and applying the standard simplifications.

Notes on the definition. There are different but equivalent ways to define these proof system. Furthermore, since in this note we only focus on *refutation of CNFs*, for both Res and $\mathsf{Res}^*(k)$ we can omit the weakening rule without any proof length penalty. Furthermore there is no need of axioms in resolution refutations.

3. Bounded parity vs bounded conjunction

In this section we prove that for every constant k the proof system $\mathsf{Res}(\oplus_k)$ is strictly weaker than $\mathsf{Res}(k)$. The *graph ordering principle (GOP)* is the base for this separation, as in [31]. For an undirected finite graph G, formula $\mathsf{GOP}(G)$ states that there is a partial order of its vertices so that every vertex has a smaller neighbor. This is unsatisfiable because partially ordered finite sets must have minimal elements. We omit unnecessary details about the formula. It suffices to know that it has $\binom{|V(G)|}{2}$ variables and $O(|V(G)|^3)$ clauses of width at most $\max(3, d)$ where d is the degree of G. This formula is interesting for proof complexity because is an extremal case for the well known length-width relation in resolution [7]. Namely, there is a family of graphs $\{G_t\}_{t\in\mathbb{N}}$ with $\Theta(t)$ vertices and constant degree d such that $\mathsf{GOP}(G_t)$ has a resolution refutation of length $O(t^3)$ but requires width $\Omega(t)$ to be refuted in resolution [32, 10].

We start with $\mathsf{GOP}(G)$ and build a hard formula for $\mathsf{Res}(\oplus_k)$ from it. For each $x_i \in \mathsf{Vars}(\mathsf{GOP}(G))$ we substitute the literals as follows

$$x_i \to y_{i,1}y_{i,2}\cdots y_{i,k} \bigvee z_{i,1}z_{i,2}\cdots z_{i,k} \tag{5a}$$

$$\neg x_i \to \bigvee_{1\le j,j'\le k} \neg y_{i,j} \wedge \neg z_{i,j'} \tag{5b}$$

and we denotes as X_i the set of new variables $\{y_{i,1},\ldots,y_{i,k},z_{i,1},\ldots,z_{i,k}\}$. We can take each of the clauses in $\mathsf{GOP}(G)$, apply the substitution and turn the resulting k-DNF into an equivalent set of clauses of width at most $2k\max(3,d)$. We name the resulting formula $\mathsf{GOP^k}(G)$. For constant d and k, $\mathsf{GOP^k}(G)$ is a CNF of constant width with $O(|V(G)|^2)$ variables and $O(|V(G)|^3)$ clauses. We are now ready to state the main result of this section.

Theorem 2. *Let k be given. There are $d > 0$, $\epsilon_k > 0$, and a family of $\{G_t\}_{t\in\mathbb{N}}$ of graphs with $\Theta(t)$ vertices and degree d, such that $\mathsf{GOP^k}(G_t)$ has a polynomial length refutations in $\mathsf{Res}(k)$ but requires refutations of length $2^{\Omega(\epsilon_k t)}$ in $\mathsf{Res}(\oplus_k)$.*

For the rest of the section we give the proof of this theorem. First we adapt the short resolution refutation of $\mathsf{GOP}(G_t)$ to get a short $\mathsf{Res}(k)$ refutation of $\mathsf{GOP^k}(G_t)$.

Lemma 3. *Let F be a CNF of constant width with a resolution refutation of length τ. Let F^k be obtained by applying substitution (5) to F and then by encoding all the resulting k-DNFs as equivalent sets of clauses. For any fixed constant k it holds that F^k has a $\mathsf{Res}(k)$ refutation of length $O(\tau)$.*

Proof. We apply substitution (5) to each proof line in the refutation of F to obtain a sequence Γ of k-DNF formulas. Γ does not form a refutation of F^k yet, but forms its backbone. Indeed we will derive each formula in Γ with a constant number of steps using F^k and earlier formulas in Γ as premises. This will conclude the proof.

For a clause C over the variables of F we denote as C' the k-DNF obtained from C after substitution. For $x_i \in \mathsf{Vars}(F)$ we denote as P_i and N_i the formulas (5a) and (5b), respectively. In the following we often use implicational completeness of $\mathsf{Res}(k)$, stated in Fact 1.

To simulate resolution weakening $A \vdash A \vee B$ we use $\mathsf{Res}(k)$ weakening $A' \vdash A' \vee B'$. After substitution, resolution axioms $x_i \vee \neg x_i$ becomes $P_i \vee N_i$. This is a constant size tautology and by implicational completeness it is derivable in a constant number of steps. If the original derivation uses a clause $C \in F$ then by construction F^k contains a constant number of clauses on a constant number of variables which are equivalent to C'. By implicational completeness we can derive C' from them in constant length.

For the case of a cut from $A \vee x_i$ and $B \vee \neg x_i$ to $A \vee B$ we need to derive $A' \vee B'$ from $A' \vee P_i$ and $B' \vee P_i$. Observe that N_i logically implies the clauses $\neg y_{i,1} \vee \neg y_{i,2} \vee \cdots \vee \neg y_{i,k}$ and $\neg z_{i,1} \vee \neg z_{i,2} \vee \cdots \vee \neg z_{i,k}$. Therefore $B' \vee N_i \vdash B' \vee \neg y_{i,1} \vee \neg y_{i,2} \vee \cdots \vee \neg y_{i,k}$ and $B' \vee \neg z_{i,1} \vee \neg z_{i,2} \vee \cdots \vee \neg z_{i,k}$ in the constant length. Applying cuts between these two formulas and $A' \vee P_i$, we get $A' \vee B'$. □

Corollary 4. *Let k be fixed. If G is a graph of constant degree, $\mathsf{GOP^k}(G)$ has a $\mathsf{Res}(k)$ refutation of length $O(|V(G)|^3)$.*

Now we show that a $\mathsf{Res}(\oplus_k)$ refutation Γ for $\mathsf{GOP^k}(G_t)$ needs $\exp(\Omega(t))$ steps. We use a random restriction argument: we hit Γ with a random partial assignment ρ so $\mathsf{GOP^k}(G_t)\restriction_\rho = \mathsf{GOP}(G_t)$ and so that every line of Γ gets transformed into a CNF of small width with high probability. If the refutation is short then with non-zero probability ρ turns the whole a refutation into a resolution refutation for $\mathsf{GOP^k}(G_t)$ of small width, which is impossibile by [10].

The random partial assignment ρ is sampled as follows: for every i pick one variable uniformly at random from X_i. Let's assume we picked $y_{i,1}$, since all other cases look the same by symmetry. Set $y_{i,2}, \ldots, y_{i,k}$ to true and set $z_{i,1}, \ldots, z_{i,k}$ to an assignment picked uniformly at random among $\{0,1\}^k \setminus \{1^k\}$. Formulas $P_i\restriction_\rho$ and $N_i\restriction_\rho$ are equal to $y_{i,1}$ and $\neg y_{i,1}$ respectively, therefore $\mathsf{GOP^k}(G_t)\restriction_\rho$ is isomorphic to $\mathsf{GOP}(G_t)$.

The next goal is to check how much the random assignments simplify $\vee\oplus_k$-type formulas, and for that we need the concepts of *decision tree* and *strong representation*.

Decision trees. A decision tree T for a formula F is a rooted full binary tree where each internal node is labeled by a variable of F and has labels 0 and 1 on the left and right edge going out from it, respectively. The leaves of T are labeled by either 0 or 1. Each leaf ℓ in T corresponds to a root-to-leaf path. No variable is allowed to label two nodes on any such path. Partial assignment ρ_ℓ assigns each variable occurring on that path to the value labeling the corrisponding outgoing edge. The height of the tree $h(T)$ is the length of the longest root-to-leaf path in it.

A decision tree strongly represents a k-DNF D when (a) for every leaf labeled by 0, assignment ρ_ℓ sets to zero all terms of D; (b) for every leaf labeled by 1, assignment ρ_ℓ sets to one at least one term of D.

Definition 5. *Let D be a k-DNF formula. A* cover set *of F is a set of variables that intersects the set of variables of each term of F. The* cover number *$c(F)$ is the size of a smallest cover set for F. The* decision tree height *of F, denoted as $h(F)$ is the height $h(T)$ of the smallest tree T that strongly represents F. We extend the definitions to an $\vee\oplus_k$-type formula F: let F' be the k-DNFs obtained representing each parity in F as a k-DNF, then $c(F) := c(F')$ and $h(F) := h(F')$.*

It turns out that the random partial assignment described above is likely to kill $\vee\oplus_k$-type formulas with no small cover sets.

Lemma 6. *Let $k \geq 2$ be given. Let F be a $\vee\oplus_k$-type formula in the variables of* $\mathsf{GOP}^{\mathsf{k}}(G)$ *and let ρ be the random partial assignment described above. There is a constant $\delta_k > 0$, so that $\Pr_\rho[F\restriction_\rho \neq 1] < 2^{-\delta_k c(F)}$.*

Proof. We first show that the parity Q of up to k variables from X_i gets assigned to either 0 or 1 with probability at least $1/6$ each. Fix $Y_i = \{y_{i,1}, \dots, y_{i,k}\}$ and $Z_i = \{z_{i,1}, \dots, z_{i,k}\}$. Exactly one variable from X_i survives the partial assignment ρ. If it is one from Y_i then a non-empty parity of variables in Z_i gets any specific boolean value with probability at least $\frac{2^{k-1}-1}{2^k-1} \geq 1/3$. The same happens swapping Y_i and Z_i. Thus if Q only contains variables in one set among Y_i and Z_i then we have both $Q\restriction_\rho = 0$ and $Q\restriction_\rho = 1$ with probability at least $1/6$; otherwise $Q = Q_y \oplus Q_z$ where Q_y and Q_z are non-empty parities of variables in Y_i and Z_i, respectively. There are at most k veriables in Q, so with probability $1/2$ both Q_y and Q_z become constant and one among Q_y and Q_z is set to a specific value b with probability at least $1/3$.

Now consider two parities Q_1 and Q_2 in F, we say that they are index-disjoint when either $\mathsf{Vars}(Q_1) \cap X_i = \varnothing$ or $\mathsf{Vars}(Q_2) \cap X_i = \varnothing$ for every $1 \leq i \leq n$. Formula F has at least $c(F)/2k^2$ parities that are pairwise index-disjoint, because each parity touches at most k sets of $2k$ variables. If we pick those variables for a maximal set of index-disjoint parities we get a cover set. On each of these parities ρ behaves independently and with probability at least $1/6^k$ sets it to the value that statisfies F. Hence

$$\Pr_\rho[F\restriction_\rho \neq 1] \leq \left(1 - \frac{1}{6^k}\right)^{c(F)/2k^2} < 2^{-\delta_k c(F)} \ . \tag{6}$$

□

Now we are going to use two results from [31]. The first one has been modified to deal with $\vee\oplus_k$-type formulas instead of k-DNFs. The proof is identical therefore we do not include it here. We just highlight the few minor differences.

Theorem 7 ([31], restated here for $\vee\oplus_k$-type formulas)**.** *Let k and s be posittive integers, and $\delta \in (0, 1]$, and let ρ a random partial assignments so that for every $\vee\oplus_k$-type formula G, $\Pr_\rho[G\restriction_\rho \neq 1] < 2^{-\delta c(F)}$. For every $\vee\oplus_k$-type formula F,*

$$\Pr_\rho[h(F\restriction_\rho) > 2s] < k2^{-\delta' s} \tag{7}$$

where $\delta' = 2(\delta/4)^k$.

Proof sketch. The original proof for k-DNFs is split between [31, Theorem 3.3] and [31, Corollary 3.4], and it goes by induction on k. The base case is when the k-CNF has large cover number. With high probability the partial assignment shrinks the formula to a small decision tree. Otherwise if the k-DNF has small cover set S, a decision tree queries just the variables in S and reduces to the case of $(k-1)$-DNF. For $\vee\oplus_k$-type formulas the argument is the same, but the inductive hypothesis is that with good probability the random partial assignment is able to set $\vee\oplus_i$-type formulas with large cover number to constant, for $2 \leq i \leq k$. □

The hypothesis in Theorem 7 is weaker than the one used in [31], and only allows to prove the theorem for $\vee\oplus_k$-type formulas. Indeed to separate $\mathsf{Res}(\oplus_k)$ from $\mathsf{Res}(k)$ we need a restriction that simplifies $\vee\oplus_k$-type formulas but not all k-DNFs. Once the partial assignment simplifies enough the lines in the refutation, the whole thing can be translated into resolution with narrow clauses.

Lemma 8 (Corollary 5.2 in [31])**.** *Let F be a CNF of width at most h. Let Γ be a $\mathsf{Res}(k)$ refutation of F, and let ρ be a partial assignment so that for every line D of Γ, $h(D\restriction_\rho) \leq h$. Then $F\restriction_\rho$ has a resolution refutation of width at most kh.*

Now we have all the ingredients to prove Lemma 9 which, together with Corollary 4 concludes the proof of Theorem 2.

Lemma 9. *Let k be given. There is integer $d > 0$, $\epsilon_k > 0$, and a family of graphs $\{G_t\}_{t\in\mathbb{N}}$ with $\Theta(t)$ vertices and degree d, such that formula $\mathsf{GOP}^\mathsf{k}(G_t)$ requires refutations of length $2^{\Omega(\epsilon_k t)}$ in $\mathsf{Res}(\oplus_k)$.*

Proof. Let Γ be a refutation of $\mathsf{GOP}^\mathsf{k}(G_t)$ in $\mathsf{Res}(\oplus_k)$. We take a random partial assignment ρ according to the distribution described above, and we apply it to every line in the refutation Γ. By the choice of G_t we know that there is a constant β so that any resolution refutation of $\mathsf{GOP}(G_t)$ must have width βt. By Lemma 6 the distribution of ρ satisfies the hypothesis of Theorem 7, which we apply setting $s = (\beta t - 1)/2k$. Hence for any specific line D in the refutation Γ we get that

$$\Pr_\rho[h(D\restriction_\rho) > (\beta t - 1)/k] < k2^{-(\delta/4)^k(\beta t-1)/k}\ , \tag{8}$$

which we can rewrite as

$$\Pr_\rho[h(D\restriction_\rho) > (\beta t - 1)/k] < 2^{-\epsilon_k t} \tag{9}$$

for some ϵ_k. If the length of Γ is less than $2^{\epsilon_k t}$ then by union bound there is a value of ρ that makes all lines in Γ to shrink to decision trees of height at most $(\beta t - 1)/k$. By Lemma 8 this allows resolution to refute $\mathsf{GOP}(G_t)$ in width $\beta t - 1$, which is a contradiction. Hence the length of Γ has to be at least $2^{\epsilon_k t}$. □

Remark. At the cost of making the proof a bit more complicated it is possible to improve this separation a little. Consider a more elaborate substitution $x_i \to \bigvee_{j=1}^k \bigwedge_{j'=1}^k y_{i,j,j'}$ and the corresponding random partial assignment that reduces it back to a single variable. It would be relatively straightforward to adapt Lemma 6 to parities of size $k^2/2$ and then show that $\mathsf{Res}(\oplus_{k^2/2})$ cannot simulate $\mathsf{Res}(k)$.

4. Tree-like k-DNF resolution vs narrow resolution

In this section we show that small $\mathsf{Res}^*(k)$ refutations and narrow Res refutations are quite close in power. As first step we show that a small tree-like $\mathsf{Res}(k)$ refutation can be converted to a narrow resolution refutation.

Lemma 10. *Let F be a w-CNF formula. If F has a $\mathsf{Res}^*(k)$ refutation of length τ, then F has a resolution refutation of width $k\lceil\log\tau\rceil + \max(k, w)$.*

Proof. To prove the lemma we show, by induction on the number of leaves L in the refutation, that there is a resolution refutation of width $W = k\lceil\log L\rceil + \max(k, w)$. This is sufficient since in a tree-like refutation we have $L < \tau$. When $L = 1$ the initial CNF contains the empty clause, and the result is trivial. Let us assume $L > 1$. The last line of the refutation is the empty DNF, hence it must be the result of a cut between conjunction $C = \bigwedge_{i=1}^{s} l_i$ and disjunction $D = \bigvee_{i=1}^{s} \neg l_i$, for a set of s literals.

The two proofs of C and D are disjoint because the refutation is tree-like, so the number of leaves in each proofs (L_C and L_D respectively) are such that $L_C + L_D = L$. Thus either L_C or L_D is less than or equal to $\frac{L}{2}$, and both are less than L. The proof is divided in two cases:

Case $L_C \le L/2$: the term $\bigwedge_{i=1}^{s} l_i$ has a $\mathsf{Res}^*(k)$ derivation with L_C leaves. By fixing $l_i = 0$ we get that $F\restriction_{l_i=0}$ has a Res^*k refutation with at most L_C leaves. By inductive hypothesis the same refutation can be done by resolution in width at most $k\lceil\log L_C\rceil + \max(k, w) \le k\lceil\log L\rceil - k + \max(k, w) \le W - 1$.

This gives a resolution proof of $F \vdash l_i$ of width W, for any $1 \le i \le s$. Using such literals we can deduce $F\restriction_{l_1=1,\ldots,l_s=1}$ from F in width k by removing any occurrences of literals $\neg l_i$. Since $F \vdash \bigvee_{i=1}^{s} \neg l_i$ in $\mathsf{Res}^*(k)$ with strictly less than L leaves, we can prove $F|_{l_1=1,\ldots,l_s=1} \vdash \bot$ in $\mathsf{Res}^*(k)$ with strictly less than L leaves. By inductive hypothesis this refutation can be done in resolution in width W. Composing the resolution proofs $F \vdash l_i$ for $1 \le i \le s$, the proof $F, l_1, \ldots l_s \vdash F\restriction_{l_1=1,\ldots,l_s=1}$ and the proof $F\restriction_{l_1=1,\ldots,l_s=1} \vdash \bot$, we get a resolution refutation of width W of F.

Case $L_D \le L/2$: we may assume that $s > 1$ because otherwise formulas C and D can be swapped and the reasoning for the previous case applies. From the $\mathsf{Res}^*(k)$ derivation $F \vdash \bigvee_{i=1}^{s} \neg l_i$ with L_D leaves we get a derivation $F\restriction_{l_1=1,\ldots,l_s=1} \vdash \bot$ with at most L_D leaves. By inductive hypothesis the same refutation can be done by resolution in width at most $k\lceil\log L_D\rceil + \max(k, w) \le k\lceil\log L\rceil - k + \max(k, w) \le W - k$. By weakening resolution proves $F \vdash \bigvee_{i=1}^{s} \neg l_i$ in width W.

We now conclude arguing that resolution proves $F, \bigvee_{i=1}^{s} \neg l_i \vdash \bot$ in width W. To see that observe the $\mathsf{Res}^*(k)$ proof of $F \vdash \bigwedge_{i=1}^{s} l_i$: each occurrence of $\bigwedge_{i=1}^{s} l_i$ is introduced in the proof using the axiom $\bigwedge_{i=1}^{s} l_i \vee \bigvee_{i=1}^{s} \neg l_i$. Substitute such axiom with the new initial clause $\bigvee_{i=1}^{s} \neg l_i$. By an easy induction along the derivation, such transformation produces a $\mathsf{Res}^*(k)$ proof of $F, \bigvee_{i=1}^{s} \neg l_i \vdash \bot$ with $L_C < L$ leaves. By inductive hypothesis this implies a Res refutation of width W (notice that the initial width of the formula increases, but that is accounted in the definition of W). □

The following simple proof gives a transformation in the other direction.

Lemma 11. *Let F be any CNF. If F has a resolution refutation of width w and length τ, then F has a $\mathsf{Res}^*(w)$ refutation of length $O(\tau)$.*

Proof. Consider the resolution refutation $D_1, D_2, \ldots D_\tau$ of F, of width w. Without loss of generality we may assume that each line is either in F or have been derived by a cut (see the notes after the definition of resolution). We define the sequence of w-DNFs $E'_t = \bigvee_{i=1}^{t} \neg D_i$. By backward induction on t from $\tau - 1$ to 0 we are going to derive a w-DNF E_t such that the terms of E_t are a subset of the terms of E'_t. Since E'_0 is the empty DNF that would conclude the refutation.

For $t = \tau - 1$ notice that $E'_{\tau-1}$ contains axiom $x \vee \neg x$ for some $x \in \mathsf{Vars}(F)$.

Fix $t < \tau - 1$ and consider $D_{t+1} \in F$. Either $E_{t+1} \equiv \Delta$ or $E_{t+1} \equiv \Delta \vee \neg D_{t+1}$, where the terms of Δ are all contained in E'_t by inductive hypothesis. In both cases we fix $E_t = \Delta$, but only in the second case we need to derive it. To do that is sufficient to apply a cut between $\Delta \vee \neg D_{t+1}$ and initial clause D_{i+1}.

Fix $t < \tau - 1$ and consider D_a and D_b that have been used to derive D_{t+1}. For convenience we write as follows

$$D_a \equiv A \vee x \quad D_b \equiv B \vee \neg x \quad D_{t+1} \equiv A \vee B \quad E_{t+1} \equiv \Delta \vee (\neg A \wedge \neg B)$$

for some w-DNF Δ, some clauses A, B and some variable x. Terms of Δ are all contained in E'_t by inductive hypothesis. Employ the following tree-like deduction

$(\neg A \wedge \neg x) \vee A \vee x$	Axiom	(10a)
$(\neg B \wedge x) \vee B \vee \neg x$	Axiom	(10b)
$(\neg A \wedge \neg x) \vee (\neg B \wedge x) \vee A \vee B$	Cut between (10a) and (10b)	(10c)
$\Delta \vee (\neg A \wedge \neg B)$	E_{t+1} deduced by induction hypothesis	(10d)
$(\neg A \wedge \neg x) \vee (\neg B \wedge x) \vee \Delta$	Cut between (10c) and (10d)	(10e)

Notice that formula (10e) is a w-DNF, and its terms are contained in E'_t, therefore it is a valid choice for E_t. At each step E_t is derived using a single occurrence of formula E_{t+1}, which means that the whole refutation is tree-like and has $O(stdlength)$ proof lines. □

In the proof above the w-DNFs have at most τ terms each, hence the refutation has at most $O(\tau^2)$ terms. Putting together Lemma 10 and 11 we get the following result which we state in a form that is robust of useful for bounded arithmetic.

Theorem 12. *Let F be a k-CNF on n variables. F has a $\mathsf{Res}^*(\mathsf{polylog}(n))$ refutation of quasi-polynomial size if and only if has a Res refutation of poly-logarithmic width.*

Acknowledgments. The author thanks Neil Thapen for discussions on this topic.

References

[1] Michael Alekhnovich, Eli Ben-Sasson, Alexander A. Razborov, and Avi Wigderson. Space complexity in propositional calculus. *SIAM J. Comput.*, 31(4):1184–1211, 2002.

[2] Albert Atserias and Víctor Dalmau. A combinatorial characterization of resolution width. *J. Comput. Syst. Sci.*, 74(3):323–334, 2008.

[3] Albert Atserias, Johannes Klaus Fichte, and Marc Thurley. Clause-learning algorithms with many restarts and bounded-width resolution. *J. Artif. Intell. Res. (JAIR)*, 40:353–373, 2011.

[4] Albert Atserias and Neil Thapen. The ordering principle in a fragment of approximate counting. *ACM Transactions on Computational Logic (TOCL)*, 15(4):29, 2014.

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

[6] Eli Ben-Sasson, Russell Impagliazzo, and Avi Wigderson. Near optimal separation of tree-like and general resolution. *Combinatorica*, 24(4):585–603, 2004.

[7] Eli Ben-Sasson and Avi Wigderson. Short proofs are narrow - resolution made simple. *J. ACM*, 48(2):149–169, 2001.

[8] Olaf Beyersdorff, Nicola Galesi, and Massimo Lauria. A characterization of tree-like resolution size. *Information Processing Letters*, 113(18):666–671, 2013.

[9] Archie Blake. *Canonical Expressions in Boolean Algebra*. PhD thesis, University of Chicago, 1938.

[10] Maria Luisa Bonet and Nicola Galesi. Optimality of size-width tradeoffs for resolution. *Computational Complexity*, 10(4):261–276, 2001.

[11] Samuel R Buss, Leszek Aleksander Kolodziejczyk, and Neil Thapen. Fragments of approximate counting. *J. Symb. Log.*, 79(2):496–525, 2014.

[12] Siu Man Chan, Massimo Lauria, Jakob Nordström, and Marc Vinyals. Hardness of approximation in pspace and separation results for pebble games. In *Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, FOCS '15, pages 466–485, 2015.

[13] Stephen A. Cook and Robert A. Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 44:36–50, 1979.

[14] Martin Davis, George Logemann, and Donald Loveland. A machine program for theorem-proving. *Commun. ACM*, 5:394–397, July 1962.

[15] Martin Davis and Hilary Putnam. A computing procedure for quantification theory. *J. ACM*, 7:201–215, July 1960.

[16] Jan Elffers, Jan Johannsen, Massimo Lauria, Thomas Magnard, Jakob Nordström, and Marc Vinyals. Trade-offs between time and memory in a tighter model of CDCL SAT solvers. In *Theory and Applications of Satisfiability Testing - SAT 2016*, pages 160–176, 2016.

[17] Juan Luis Esteban and Jacobo Torán. Space bounds for resolution. *Information and Computation*, 171(1):84–97, 2001. Preliminary versions of these results appeared in *STACS '99* and *CSL '99*.

[18] Yuval Filmus, Massimo Lauria, Mladen Miksa, Jakob Nordström, and Marc Vinyals. From small space to small width in resolution. In *31st International Symposium on Theoretical Aspects of Computer Science, STACS*, pages 300–311, 2014.

[19] John R Gilbert, Thomas Lengauer, and Robert Endre Tarjan. The pebbling problem is complete in polynomial space. *SIAM Journal on Computing*, 9(3):513–524, 1980.

[20] A. Haken. The intractability of resolution. *Theoretical Computer Science*, 39:297–308, 1985.

[21] Jan Krajíček. On the weak pigeonhole principle. *Fundamenta Mathematicae*, 170(1-3):123–140, 2001.

[22] Jan Krajíček. Lower bounds to the size of constant-depth propositional proofs. *Journal of Symbolic Logic*, 59(1):73–86, 1994.

[23] Jan Krajíček. Interpolation theorems, lower bounds for proof systems, and independence results for bounded arithmetic. *The Journal of Symbolic Logic*, 62(2):457–486, 1997.

[24] Massimo Lauria. Short res*(polylog) refutations if and only if narrow res refutations. *CoRR*, abs/1310.5714, 2013.

[25] João P. Marques-Silva and Karem A. Sakallah. GRASP: A search algorithm for propositional satisfiability. *IEEE Transactions on Computers*, 48(5):506–521, May 1999. Preliminary version in *ICCAD '96*.

[26] Matthew W. Moskewicz, Conor F. Madigan, Ying Zhao, Lintao Zhang, and Sharad Malik. Chaff: Engineering an efficient SAT solver. In *Proceedings of the 38th Design Automation Conference (DAC '01)*, pages 530–535, June 2001.

[27] Jakob Nordström. Pebble games, proof complexity and time-space trade-offs. *Logical Methods in Computer Science*, 9:15:1–15:63, September 2013.

[28] Michael S. Paterson and Carl E. Hewitt. Comparative schematology. In *Record of the Project MAC Conference on Concurrent Systems and Parallel Computation*, pages 119–127, 1970.

[29] Knot Pipatsrisawat and Adnan Darwiche. On the power of clause-learning sat solvers as resolution engines. *Artificial Intelligence*, 175(2):512 – 525, 2011.

[30] J. A. Robinson. A machine-oriented logic based on the resolution principle. *J. ACM*, 12:23–41, January 1965.

[31] Nathan Segerlind, Samuel R. Buss, and Russell Impagliazzo. A switching lemma for small restrictions and lower bounds for k-DNF resolution. *SIAM J. Comput.*, 33(5):1171–1200, 2004.

[32] Gunnar Stålmarck. Short resolution proofs for a sequence of tricky formulas. *Acta Informatica*, 33(3):277–280, May 1996.