Electronic Colloquium on Computational Complexity, Report No. 37 (2020)

Failure of Feasible Disjunction Property for k-DNF Resolution and NP-hardness of Automating It

Michal Garlík*

Dept. Ciències de la Computació
Universitat Politècnica de Catalunya
C. Jordi Girona, 1-3
08034 Barcelona, Spain
email: mgarlik@cs.upc.edu

March 18, 2020

Abstract

We show that for every integer $k \geq 2$, the Res(k) propositional proof system does not have the weak feasible disjunction property. Next, we generalize a recent result of Atserias and Müller [3] to Res(k). We show that if NP is not included in P (resp. QP, SUBEXP) then for every integer $k \geq 1$, Res(k) is not automatable in polynomial (resp. quasi-polynomial, subexponential) time.

1 Introduction

Following Pudlák [17], a proof system P has *weak feasible disjunction property* if there exists a polynomial p such that if a formula $A \vee B$, in which A and B do not share variables, has a P proof of length t, then either A or B has a P proof of length $p(t)$. We deal with refutation systems in this paper, which for the preceding definition amounts to replacing in it '$\vee$' by '$\wedge$' and 'proof' by 'refutation'. It is known and easy to see that resolution has the weak feasible disjunction property. Resolution also has feasible interpolation, a prominent concept in proof complexity introduced by Krajíček [9, 10]. A refutation system P has *feasible interpolation* if there is a polynomial p and an algorithm that when given as input a refutation Π of size r of a CNF $A(\overline{x}, \overline{y}) \wedge B(\overline{x}, \overline{z})$, where $\overline{y}, \overline{x}, \overline{z}$ are disjoint sets of propositional variables, and a truth assignment σ to the variables $\overline{x}$ outputs in time $p(r)$ a value $i \in \{0, 1\}$ such that if $i = 0$ then $A \restriction \sigma$ is unsatisfiable and if $i = 1$ then $B \restriction \sigma$ is unsatisfiable. Here $F \restriction \sigma$ denotes the formula obtained from F by an application of a partial truth assignment σ to the variables of F that are in the domain of σ.

*Funded by European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, grant agreement ERC-2014-CoG 648276 (AUTAR).

ISSN 1433-8092

Pudlák [17] comments that so far the weak feasible disjunction property has been observed in all proof systems that were shown to have feasible interpolation. This is because known feasible interpolation algorithms, like those in Chapter 17.7 in [13], actually construct a refutation of one of the conjuncts.

A proof system P is *polynomially bounded* if there is a polynomial p such that any tautology of size r has a P proof of size $p(r)$. A fundamental problem in proof complexity is to show that no polynomially bounded proof system exists. This is equivalent to establishing NP $\neq$ coNP, as observed by Cook and Reckhow [6]. There is a potentially useful observation by Krajíček [12] that for the purpose of proving that some proof system P is not polynomially bounded we may assume without a loss of generality that P admits the weak feasible disjunction property. This readily follows from the fact that if a disjunction of two formulas that do not share variables is a tautology, then one of the disjuncts is.

A propositional version of the negation of the *reflection principle* for a proof system P is a conjunction of a propositional formula expressing that '$\overline{z}$ is a satisfying assignment of formula $\overline{x}$ of length r' and a propositional formula expressing that '$\overline{y}$ is a P refutation of length t of formula $\overline{x}$ of length r'. Here P, t, r are fixed parameters and $\overline{x}, \overline{y}, \overline{z}$ are disjoint sets of variables. When we plug in for the common variables $\overline{x}$ some formula F of length r, we denote the conjunction by $\mathrm{SAT}^F \wedge \mathrm{REF}^F_{P,t}$, and we call the second conjunct a *P refutation statement for F*. We need to define one very mild requirement on a proof system in order to state a result from [17] about the weak feasible disjunction property that is the main source of motivation for this paper. We say that P is *closed under restrictions* if there is a polynomial p such that whenever F has a P proof of length t and σ is a partial truth assignment to the variables of F, then there is a P proof of $F\lceil\sigma$ of length at most $p(t)$.

There is a proposition proved in [17] saying that if a proof system P has the weak feasible disjunction property, has polynomial-size proofs of the reflection principle for P, is closed under restrictions, and has the property that given a P proof of $\neg\mathrm{SAT}^{\neg F}$ there is at most polynomially longer P proof of F, then for every formula F and every integer t which is at least the size of F, either there is a P proof of F of length $t^{O(1)}$, or there is a $t^{O(1)}$ long P proof of $\neg\mathrm{REF}^{\neg F}_{P,t}$. Pudlák comments that the conclusion of this proposition seems unlikely (and therefore it seems unlikely that a proof system satisfying the remaining three reasonable properties has the weak feasible disjunction property). He concludes that the weak feasible disjunction property is very unlikely to occur unless the system is very weak. Motivated to find and emphasize the contrast between resolution and Res(2) (see Section 2) in this respect, we show the following theorem.

Theorem 1. *For every integer $k \geq 2$, $\mathrm{Res}(k)$ does not have the weak feasible disjunction property. Moreover, there are families $\{A_n\}_{n\geq 1}$ and $\{B_{n,k}\}_{n\geq 1, k\geq 1}$ of CNFs, where A_n has size $n^{O(1)}$, $B_{n,k}$ has size $n^{O(k)}$, and A_n and $B_{n,k}$ do not share any variables, such that all the following hold:*

(i) There exists $\alpha > 0$ and an integer n_1 such that for every $k \geq 1$ and $n \geq n_1$, any Res(k) refutations of A_n has size greater than 2^{n^α}.

(ii) For every $k \geq 1$ there is $\beta > 0$ and an integer n_2 such that for every $n \geq n_2$, any $\mathrm{Res}(k)$ refutation of $B_{n,k}$ has size greater than $2^{\beta n}$.

(iii) For all integers $n \geq 1$ and $k \geq 1$, $A_n \wedge B_{n,k}$ has a $\mathrm{Res}(2)$ refutation of size $O(k^2 n^{7k+7})$.

The idea is to employ a reflection, but instead of the reflection principle for $\mathrm{Res}(k)$, which would correspond to the hypothesis of Pudlák's proposition above, we work with the reflection principle for resolution and make it harder by the relativization technique of Dantchev and Riis [7]. More precisely, we replace in the reflection principle the resolution refutation statement by its k-fold relativization. Most of this paper is then concerned with proving length lower bounds on $\mathrm{Res}(k)$ refutations of a version of the k-fold relativization of $\mathrm{REF}^F_{\mathrm{Res},t}$ for every unsatisfiable CNF F (Theorem 23). This lower bound will be used to prove item (ii) above, but since it works for every unsatisfiable F, item (i) will be easy to get choosing F to be hard enough for $\mathrm{Res}(k)$. The upper bound, item (iii), generalizes upper bounds for similar formulas [2, 3, 8], which all build on an idea from [17].

To prove Theorem 23, the mentioned main lower bound, we develop a switching lemma in the spirit of [18] but respecting the functional properties of the formula $\mathrm{REF}^F_{\mathrm{Res},t}$. This will come at a cost of worse parameters in the switching lemma and its narrowed applicability in terms of random restrictions it works for.

Our second result is a generalization of conditional non-automatability results for resolution [3] to the systems $\mathrm{Res}(k)$. Following [5, 2] and [3], we say that a refutation system P is *automatable in time* $T : \mathbb{N} \to \mathbb{N}$ if there is an algorithm that when given as input an unsatisfiable CNF F of size r outputs a P refutation of F in time $T(r + s_P(F))$, where $s_P(F)$ is the length of a shortest P refutation of F. If the function T is a polynomial, then P is simply called *automatable*. A refutation system P is *weakly automatable* if there is a refutation system Q, a polynomial p, and an algorithm that when given as input an unsatisfiable CNF F of size r outputs a Q refutation of F in time $p(r + s_P(F))$. It is known that feasible interpolation is implied by weak automatability in refutation systems that are closed under restrictions (see Theorem 3 in [2]).

First negative automatability results were obtained by Krajíček and Pudlák [14] who showed that Extended Frege systems do not have feasible interpolation assuming that RSA is secure against P/poly. Bonet et al. [5, 4] showed that Frege systems and constant-depth Frege systems do not have feasible interpolation assuming the Diffie-Hellman key exchange procedure is secure against polynomial and subexponential size circuits, respectively. All these proof systems are closed under restrictions, hence these results conditionally rule out weak automatability and automatability. As for resolution, before a recent breakthrough by Atserias and Müller [3] who showed that resolution is not automatable unless P = NP, it was known by a result of Alekhnovich and Razborov [1] that resolution is not automatable unless W[P] = FPT. Here W[P] is the class of parametrized problems that are fixed-parameter reducible to the problem of deciding if a monotone circuit C has a satisfying assignment of Hamming weight k. We refer an interested reader to the introduction section of [3] for more on the history of the automatability problem.

Let QP denote the class of problems decidable in quasi-polynomial time $2^{(\log n)^{O(1)}}$, and let SUBEXP denote the class of problems decidable in subexponential time $2^{n^{o(1)}}$. We show the following theorem, which was proved for $k = 1$ in [3].

Theorem 2. 1. *If* $\mathrm{NP} \not\subseteq \mathrm{P}$ *then for every integer* $k \geq 1$, $\mathrm{Res}(k)$ *is not automatable in polynomial time.*

2. *If* $\mathrm{NP} \not\subseteq \mathrm{QP}$ *then for every integer* $k \geq 1$, $\mathrm{Res}(k)$ *is not automatable in quasi-polynomial time.*

3. *If* $\mathrm{NP} \not\subseteq \mathrm{SUBEXP}$ *then for every integer* $k \geq 1$, $\mathrm{Res}(k)$ *is not automatable in subexponential time.*

The basic idea of the proof is the same as in [3]: to map every formula F to a resolution refutation statement for F, and show that if F is satisfiable then the refutation statement has a polynomial-length $\mathrm{Res}(k)$ refutation, and if F is unsatisfiable then the refutation statement requires long $\mathrm{Res}(k)$ refutations. An automating algorithm that finds short refutations quickly enough can then be used to distinguishing between the two situations, and hence to solve SAT. We thus need to show strong lower bounds on the length of $\mathrm{Res}(k)$ refutations of a version of resolution refutation statements. For this we use the already discussed Theorem 23 once more.

2 Preliminaries

For an integer s, the set $\{1, \ldots, s\}$ is denoted by $[s]$. We write $\mathrm{dom}(\sigma), \mathrm{im}(\sigma)$ for the domain and image of a function σ. Two functions σ, τ are *compatible* if $\sigma \cup \tau$ is a function. If x is a propositional variable, the *positive literal of* x, denoted by x^1, is x, and the *negative literal of* x, denoted by x^0, is $\neg x$. A *clause* is a set of literals. A clause is written as a disjunction of its elements. A *term* is a set of literals, and is written as a conjunction of the literals. A *CNF* is a set of clauses, written as a conjunction of the clauses. A *k-CNF* is a CNF whose every clause has at most k literals. A *DNF* is a set of terms, written as a disjunction of the terms. A *k-DNF* is a DNF whose every term has at most k literals. We will identify 1-DNFs with clauses. A clause is *non-tautological* if it does not contain both the positive and negative literal of the same variable. A clause C is a *weakening* of a clause D if $D \subseteq C$. A clause D is the *resolvent of* clauses C_1 and C_2 *on* a variable x if $x \in C_1, \neg x \in C_2$ and $D = (C_1 \setminus \{x\}) \cup (C_2 \setminus \{\neg x\})$. If E is a weakening of the resolvent of C_1 and C_2 on x, we say that E is obtained by the *resolution rule* from C_1 and C_2, and we call C_1 and C_2 the *premises* of the rule.

Let F be a CNF and C a clause. A *resolution derivation of C from F* is a sequence of clauses $\Pi = (C_1, \ldots, C_s)$ such that $C_s = C$ and for all $u \in [s]$, C_u is a weakening of a clause in F, or there are $v, w \in [u-1]$ such that C_u is obtained by the resolution rule from C_v and C_w. A *resolution refutation of F* is a resolution derivation of the empty clause from F. The *length* of a resolution derivation $\Pi = (C_1, \ldots, C_s)$ is s. For $u \in [s]$, the *height of u in* Π is the maximum h such that there is a subsequence $(C_{u_1}, \ldots, C_{u_h})$ of Π in which $u_h = u$ and for each $i \in [h-1]$, C_{u_i} is a premise of a resolution rule by which $C_{u_{i+1}}$ is obtained in Π. The *height* of Π is the maximum height of u in Π for $u \in [s]$.

A *partial assignment* to the variables $x_1, \ldots, x_n$ is a partial map from $\{x_1, \ldots, x_n\}$ to $\{0, 1\}$. Let σ be a partial assignment. The CNF $F \restriction \sigma$ is formed from F by removing every clause containing a literal satisfied by σ, and removing every literal falsified by σ from the remaining clauses. If $\Pi = (C_1, \ldots, C_s)$ is a sequence of clauses, $\Pi \restriction \sigma$ is formed from Π by the same operations. Note that if Π is a resolution refutation of F, then $\Pi \restriction \sigma$ is a resolution refutation of $F \restriction \sigma$.

The $\mathrm{Res}(k)$ refutation system is a generalization of resolution introduced by Krajíček [11][1]. Its lines are k-DNFs and it has the following inference rules (A, B are k-DNFs, $j \in [k]$, and $l, l_1, \ldots, l_j$ are literals):

[1] In [11] (see also Chapter 5.7 in [13]) more general fragments $\mathrm{R}(f)$ of DNF-resolution are introduced, where $f : \mathbb{N} \to \mathbb{N}$ is non-decreasing and a refutation Π is said to have $\mathrm{R}(f)$-*size* s if its lines are $f(s)$-DNFs and $|\Pi| \leq s$. In the present paper we work with constant functions f.

$$\frac{A \vee l_1 \qquad B \vee (l_2 \wedge \cdots \wedge l_j)}{A \vee B \vee (l_1 \wedge \cdots \wedge l_j)}\ \wedge\text{-introduction} \qquad\qquad \frac{}{x \vee \neg x}\ \text{Axiom}$$

$$\frac{A \vee (l_1 \wedge \cdots \wedge l_j) \qquad B \vee \neg l_1 \vee \cdots \vee \neg l_j}{A \vee B}\ \text{Cut} \qquad\qquad \frac{A}{A \vee B}\ \text{Weakening}$$

Let F be a CNF. A *Res*(k) *derivation from* F is a sequence of k-DNFs $(D_1, \ldots, D_s)$ so that each D_i either belongs to F or follows from the preceding lines by an application of one of the inference rules. A *Res*(k) *refutation of* F is a Res(k) derivation from F whose final line is the empty clause. The *length* of a Res(k) derivation $\Pi = (D_1, \ldots, D_s)$, denoted by $|\Pi|$, is s. The *size* of Π, denoted by size(Π), is the number of symbols in it.

3 Resolution Refutations of s Levels of t Clauses

Like in [8], it will be convenient to work with a variant of resolution in which the clauses forming a refutation are arranged in layers. All the definitions in this section are taken from [8].

Definition 3. Let F be a CNF of r clauses in n variables $x_1, \ldots, x_n$. We say that F has a *resolution refutation of s levels of t clauses* if there is a sequence of clauses $C_{i,j}$ indexed by all pairs $(i, j) \in [s] \times [t]$, such that each clause $C_{1,j}$ on the first level is a weakening of a clause in F, each clause $C_{i,j}$ on level $i \in [s] \setminus \{1\}$ is a weakening of the resolvent of two clauses from level $i - 1$ on a variable, and the clause $C_{s,t}$ is empty.

The following proposition says that insisting that the clauses are arranged in layers is not a very limiting requirement since this system quadratically simulates resolution and preserves the refutation height.

Proposition 4 ([8])**.** *If a $(n-1)$-CNF F in n variables has a resolution refutation of height h and length s, then F has a resolution refutation of h levels of $3s$ clauses.*

We now formalize refutation statements for this system in the same way as in [8]. Let n, r, s, t be integers. Let F be a CNF consisting of r clauses $C_1, \ldots, C_r$ in n variables $x_1, \ldots, x_n$. We define a propositional formula $\mathrm{REF}^F_{s,t}$ expressing that F has a resolution refutation of s levels of t clauses.

We first list the variables of $\mathrm{REF}^F_{s,t}$. *D-variables* $D(i, j, \ell, b)$, $i \in [s]$, $j \in [t], \ell \in [n], b \in \{0, 1\}$, encode clauses $C_{i,j}$ as follows: $D(i, j, \ell, 1)$ (resp. $D(i, j, \ell, 0)$) means that the literal x_ℓ (resp. $\neg x_\ell$) is in $C_{i,j}$. *L-variables* $L(i, j, j')$ (resp. *R-variables* $R(i, j, j')$), $i \in [s] \setminus \{1\}, j, j' \in [t]$, say that $C_{i-1,j'}$ is a premise of the resolution rule by which $C_{i,j}$ is obtained, and it is the premise containing the positive (resp. negative) literal of the resolved variable. *V-variables* $V(i, j, \ell)$, $i \in [s] \setminus \{1\}, j \in [t], \ell \in [n]$, say that $C_{i,j}$ is obtained by resolving on x_ℓ. *I-variables* $I(j, m)$, $j \in [t], m \in [r]$, say that $C_{1,j}$ is a weakening of C_m.

$\mathrm{REF}^F_{s,t}$ is the union of the following fifteen sets of clauses:

$$\neg I(j, m) \vee D(1, j, \ell, b) \qquad\qquad j \in [t], m \in [r], b \in \{0, 1\}, x^b_\ell \in C_m, \tag{1}$$

clause $C_{1,j}$ contains the literals of C_m assigned to it by $I(j, m)$,

$$\neg D(i, j, \ell, 1) \vee \neg D(i, j, \ell, 0) \qquad\qquad i \in [s], j \in [t], \ell \in [n], \tag{2}$$

no clause $C_{i,j}$ contains x_ℓ and $\neg x_\ell$ at the same time,

$$\neg L(i,j,j') \vee \neg V(i,j,\ell) \vee D(i-1,j',\ell,1) \qquad i\in[s]\backslash\{1\}, j,j'\in[t], \ell\in[n], \tag{3}$$

$$\neg R(i,j,j') \vee \neg V(i,j,\ell) \vee D(i-1,j',\ell,0) \qquad i\in[s]\backslash\{1\}, j,j'\in[t], \ell\in[n], \tag{4}$$

clause $C_{i-1,j'}$ used as the premise given by $L(i,j,j')$ (resp. $R(i,j,j')$) in resolving on x_ℓ must contain x_ℓ (resp. $\neg x_\ell$),

$$\begin{gathered}\neg L(i,j,j') \vee \neg V(i,j,\ell) \vee \neg D(i-1,j',\ell',b) \vee D(i,j,\ell',b) \\ i\in[s]\backslash\{1\}, j,j'\in[t], \ell,\ell'\in[n], b\in\{0,1\}, (\ell',b) \neq (\ell,1),\end{gathered} \tag{5}$$

$$\begin{gathered}\neg R(i,j,j') \vee \neg V(i,j,\ell) \vee \neg D(i-1,j',\ell',b) \vee D(i,j,\ell',b) \\ i\in[s]\backslash\{1\}, j,j'\in[t], \ell,\ell'\in[n], b\in\{0,1\}, (\ell',b) \neq (\ell,0),\end{gathered} \tag{6}$$

clause $C_{i,j}$ derived by resolving on x_ℓ must contain each literal different from x_ℓ (resp. $\neg x_\ell$) from the premise given by $L(i,j,j')$ (resp. $R(i,j,j')$),

$$\neg D(s,t,\ell,b) \qquad \ell\in[n], b\in\{0,1\}, \tag{7}$$

clause $C_{s,t}$ is empty,

$$V(i,j,1) \vee V(i,j,2) \vee \ldots \vee V(i,j,n) \qquad i\in[s]\backslash\{1\}, j\in[t], \tag{8}$$

$$I(j,1) \vee I(j,2) \vee \ldots \vee I(j,r) \qquad j\in[t], \tag{9}$$

$$L(i,j,1) \vee L(i,j,2) \vee \ldots \vee L(i,j,t) \qquad i\in[s]\backslash\{1\}, j\in[t], \tag{10}$$

$$R(i,j,1) \vee R(i,j,2) \vee \ldots \vee R(i,j,t) \qquad i\in[s]\backslash\{1\}, j\in[t], \tag{11}$$

$$\neg V(i,j,\ell) \vee \neg V(i,j,\ell') \qquad i\in[s]\backslash\{1\}, j\in[t], \ell,\ell'\in[n], \ell\neq\ell', \tag{12}$$

$$\neg I(j,m) \vee \neg I(j,m') \qquad j\in[t], m,m'\in[r], m\neq m', \tag{13}$$

$$\neg L(i,j,j') \vee \neg L(i,j,j'') \qquad i\in[s]\backslash\{1\}, j,j',j''\in[t], j'\neq j'', \tag{14}$$

$$\neg R(i,j,j') \vee \neg R(i,j,j'') \qquad i\in[s]\backslash\{1\}, j,j',j''\in[t], j'\neq j'', \tag{15}$$

the V, I, L, R-variables define functions with the required domains and ranges.

Definition 5. For $i \in [s], j, j' \in [t], \ell \in [n], b \in \{0,1\}, m \in [r]$, we say that (i,j) is the *home pair* of the variable $D(i,j,\ell,b)$, of the variables $R(i,j,j')$, $L(i,j,j')$, $V(i,j,\ell)$ if $i \neq 1$, and of the variable $I(j,m)$ if $i = 1$.

We write $V(i,j,\cdot)$ to stand for the set $\{V(i,j,\ell) : \ell \in [n]\}$. Similarly, we write $I(j,\cdot), L(i,j,\cdot)$, and $R(i,j,\cdot)$ to stand for the set $\{I(j,m) : m \in [r]\}$, $\{L(i,j,j') : j' \in [t]\}$, and $\{R(i,j,j') : j' \in [t]\}$, respectively. We denote by $D(i,j,\cdot,\cdot)$ the set $\{D(i,j,\ell,b) : \ell \in [n], b \in \{0,1\}\}$.

Let σ be a partial assignment. We say that $V(i,j,\cdot)$ is *set to* ℓ *by* σ if $\sigma(V(i,j,\ell)) = 1$ and for all $\ell' \in [n] \backslash \{\ell\}$, $\sigma(V(i,j,\ell')) = 0$. Similarly, we say that $I(j,\cdot)$ is *set to* m *by* σ if $\sigma(I(j,m)) = 1$ and for all $m' \in [r]\backslash\{m\}$ we have $\sigma(I(j,m')) = 0$. We say that $L(i,j,\cdot)$ (resp. $R(i,j,\cdot)$) is *set to* j' *by* σ if $\sigma(L(i,j,j')) = 1$ (resp. $\sigma(R(i,j,j')) = 1$) and for all $j'' \in [t]\backslash\{j'\}$, we have $\sigma(L(i,j,j'')) = 0$ (resp. $\sigma(R(i,j,j'')) = 0$). We say that $D(i,j,\cdot,\cdot)$ is *set to* a clause $C_{i,j}$ *by* σ if for all $\ell \in [n], b \in \{0,1\}$ we have $\sigma(D(i,j,\ell,b)) = 1$ if $x_\ell^b \in C_{i,j}$ and $\sigma(D(i,j,\ell,b)) = 0$ if $x_\ell^b \notin C_{i,j}$.

For $Y \in \{D(i,j,\cdot,\cdot), V(i,j,\cdot), I(j,\cdot), R(i,j,\cdot), L(i,j,\cdot)\}$, we say that Y is *set by* σ if Y is set to v by σ for some value v. We will often omit saying "by σ" if σ is clear from the context.

4 Reflection Principle for Resolution

We repeat the formulation of a version of the reflection principle from [8]. We express the negation of the reflection principle for resolution by a CNF in the form of a conjunction $\mathrm{SAT}^{n,r} \wedge \mathrm{REF}^{n,r}_{s,t}$. The only shared variables by the formulas $\mathrm{SAT}^{n,r}$ and $\mathrm{REF}^{n,r}_{s,t}$ encode a CNF with r clauses in n variables. The meaning of $\mathrm{SAT}^{n,r}$ is that the encoded CNF is satisfiable, while the meaning of $\mathrm{REF}^{n,r}_{s,t}$ is that the same CNF has a resolution refutation of s levels of t clauses. A formal definition is given next.

Formula $\mathrm{SAT}^{n,r}$ has the following variables. *C-variables* $C(m,\ell,b)$, $m \in [r], \ell \in [n], b \in \{0,1\}$, encode clauses C_m as follows: $C(m,\ell,1)$ (resp. $C(m,\ell,0)$) means that the literal x_ℓ (resp. $\neg x_\ell$) is in C_m. *T-variables* $T(\ell)$, $\ell \in [n]$, and $T(m,\ell,b)$, $m \in [r], \ell \in [n], b \in \{0,1\}$, encode that an assignment to variables $x_1,\dots,x_n$ satisfies the CNF $\{C_1,\dots,C_r\}$. The meaning of $T(\ell)$ is that the literal x_ℓ is satisfied by the assignment. The meaning of $T(m,\ell,1)$ (resp. $T(m,\ell,0)$) is that clause C_m is satisfied through the literal x_ℓ (resp. $\neg x_\ell$).

We list the clauses of $\mathrm{SAT}^{n,r}$:

$$T(m,1,1) \vee T(m,1,0) \vee \ldots \vee T(m,n,1) \vee T(m,n,0) \qquad m \in [r], \tag{16}$$

$$\neg T(m,\ell,1) \vee T(\ell) \qquad m \in [r], \ell \in [n], \tag{17}$$

$$\neg T(m,\ell,0) \vee \neg T(\ell) \qquad m \in [r], \ell \in [n], \tag{18}$$

$$\neg T(m,\ell,b) \vee C(m,\ell,b) \qquad m \in [r], \ell \in [n], b \in \{0,1\}. \tag{19}$$

The meaning of (16) is that clause C_m is satisfied through at least one literal. Clauses (17) and (18) say that if C_m is satisfied through a literal, then the literal is satisfied. The meaning of (19) is that if C_m is satisfied through a literal, then it contains the literal.

Variables of $\mathrm{REF}^{n,r}_{s,t}$ are the variables $C(m,\ell,b)$ of $\mathrm{SAT}^{n,r}$ together with all the variables of $\mathrm{REF}^{F}_{s,t}$ for some (and every) F of r clauses in n variables. That is, $\mathrm{REF}^{n,r}_{s,t}$ has the following variables: $C(m,\ell,b)$ for $m \in [r], \ell \in [n], b \in \{0,1\}$; $D(i,j,\ell,b)$ for $i \in [s], j \in [t], \ell \in [n], b \in \{0,1\}$; $R(i,j,j')$ and $L(i,j,j')$ for $i \in [s]\backslash\{1\}, j,j' \in [t]$; $V(i,j,\ell)$ for $i \in [s]\backslash\{1\}, j \in [t], \ell \in [n]$; $I(j,m)$ for $j \in [t], m \in [r]$.

The clauses of $\mathrm{REF}^{n,r}_{s,t}$ are (2) - (15) of $\mathrm{REF}^{F}_{s,t}$ together with the following clauses (to replace clauses (1)):

$$\neg I(j,m) \vee \neg C(m,\ell,b) \vee D(1,j,\ell,b) \qquad j \in [t], m \in [r], \ell \in [n], b \in \{0,1\}, \tag{20}$$

saying that if clause $C_{1,j}$ is a weakening of clause C_m, then the former contains each literal of the latter. So the difference from (1) is that C_m is no longer a clause of some fixed formula F, but it is described by C-variables.

Proposition 6. *Let F be a CNF with r clauses $C_1,\dots,C_r$ in n variables $x_1,\dots,x_n$, and let γ_F be an assignment such that its domain are all C-variables and $\gamma_F(C(m,\ell,b)) = 1$ if $x_\ell^b \in C_m$ and $\gamma_F(C(m,\ell,b)) = 0$ if $x_\ell^b \notin C_m$. There is a substitution τ that maps the variables of* $\mathrm{SAT}^{n,r} \restriction \gamma_F$ *to* $\{0,1\} \cup \{x_\ell^b : \ell \in [n], b \in \{0,1\}\}$ *such that* $(\mathrm{SAT}^{n,r} \restriction \gamma_F) \restriction \tau$ *is F together with some tautological clauses in the variables $x_1,\dots,x_n$.*

Proof. Define τ as follows. If $\gamma_F(C(m,\ell,b)) = 0$, then $\tau(T(m,\ell,b)) = 0$. This deletes $T(m,\ell,b)$ from (16) and satisfies (19) together with either (17) (if $b=1$) or (18) (if $b=0$). If $\gamma_F(C(m,\ell,b)) = 1$, then (19) has been satisfied and we define $\tau(T(m,\ell,b)) = x_\ell^b$ and $\tau(T(\ell)) = x_\ell$. This choice turns (17) (if $b=1$) or (18) (if $b=0$) into a tautological clause and correctly substitutes the remaining literals of (16) to yield the clause C_m of F. □

5 The Upper Bounds

In this section we work with a stronger formulation of the negation of the reflection principle for resolution, expressed by a CNF formula $\mathrm{SAT}^{n,r} \wedge \mathrm{R}^k\mathrm{REF}^{n,r}_{s,t}$. The difference from the previous formulation $\mathrm{SAT}^{n,r} \wedge \mathrm{REF}^{n,r}_{s,t}$ is that we have replaced $\mathrm{REF}^{n,r}_{s,t}$ by its k-fold relativization $\mathrm{R}^k\mathrm{REF}^{n,r}_{s,t}$. The first-order logic notion of relativization of a first-order formula to a relation was put to use in propositional proof complexity by Dantchev and Riis [7].

We first describe the k-fold relativization of $\mathrm{REF}^F_{s,t}$, denoted by $\mathrm{R}^k\mathrm{REF}^F_{s,t}$. The variables of this CNF are those of $\mathrm{REF}^F_{s,t}$ together with new variables $S_u(i,j)$, $(i,j) \in [s] \times [t]$, $u \in [k]$. The meaning of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ is that those clauses $C_{i,j}$ (described by D-variables) for which $\bigwedge_{u\in[k]} S_u(i,j)$ is satisfied form a resolution refutation of F of s levels of at most t clauses. That is, only the selected clauses $C_{i,j}$ have to form a refutation, and nothing is asked of the clauses that are not selected. Formally, $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ is the union of the following sets of clauses:

$$\bigvee_{u\in[k]} \neg S_u(1,j) \vee \neg I(j,m) \vee D(1,j,\ell,b) \qquad j\in[t], m\in[r], b\in\{0,1\}, x^b_\ell \in C_m, \tag{21}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg D(i,j,\ell,1) \vee \neg D(i,j,\ell,0) \qquad i\in[s], j\in[t], \ell\in[n], \tag{22}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg L(i,j,j') \vee \neg V(i,j,\ell) \vee D(i-1,j',\ell,1) \qquad i\in[s]\setminus\{1\}, j,j'\in[t], \ell\in[n], \tag{23}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg R(i,j,j') \vee \neg V(i,j,\ell) \vee D(i-1,j',\ell,0) \qquad i\in[s]\setminus\{1\}, j,j'\in[t], \ell\in[n], \tag{24}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg L(i,j,j') \vee \neg V(i,j,\ell) \vee \neg D(i-1,j',\ell',b) \vee D(i,j,\ell',b) \qquad i\in[s]\setminus\{1\}, j,j'\in[t], \ell,\ell'\in[n], b\in\{0,1\}, (\ell',b)\neq(\ell,1), \tag{25}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg R(i,j,j') \vee \neg V(i,j,\ell) \vee \neg D(i-1,j',\ell',b) \vee D(i,j,\ell',b) \qquad i\in[s]\setminus\{1\}, j,j'\in[t], \ell,\ell'\in[n], b\in\{0,1\}, (\ell',b)\neq(\ell,0), \tag{26}$$

$$\bigvee_{u\in[k]} \neg S_u(s,t) \vee \neg D(s,t,\ell,b) \qquad \ell\in[n], b\in\{0,1\}, \tag{27}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \bigvee_{\ell\in[n]} V(i,j,\ell) \qquad i\in[s]\setminus\{1\}, j\in[t], \tag{28}$$

$$\bigvee_{u\in[k]} \neg S_u(1,j) \vee \bigvee_{m\in[r]} I(j,m) \qquad j\in[t], \tag{29}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \bigvee_{j'\in[t]} L(i,j,j') \qquad i\in[s]\setminus\{1\}, j\in[t], \tag{30}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \bigvee_{j'\in[t]} R(i,j,j') \qquad i\in[s]\setminus\{1\}, j\in[t], \tag{31}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg V(i,j,\ell) \vee \neg V(i,j,\ell') \qquad i\in[s]\setminus\{1\}, j\in[t], \ell,\ell'\in[n], \ell\neq\ell', \tag{32}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg I(j,m) \vee \neg I(j,m') \qquad j\in[t], m,m'\in[r], m\neq m', \tag{33}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg L(i,j,j') \vee \neg L(i,j,j'') \qquad i\in[s]\setminus\{1\}, j,j',j''\in[t], j'\neq j'', \tag{34}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg R(i,j,j') \vee \neg R(i,j,j'') \qquad i\in[s]\setminus\{1\}, j,j',j''\in[t], j'\neq j'', \tag{35}$$

$$S_u(s,t) \qquad u\in[k], \tag{36}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg L(i,j,j') \vee S_{u'}(i-1,j') \qquad i\in[s]\setminus\{1\}, j,j'\in[t], u'\in[k], \tag{37}$$

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg R(i,j,j') \vee S_{u'}(i-1,j') \qquad i\in[s]\setminus\{1\}, j,j'\in[t], u'\in[k]. \tag{38}$$

Clauses in (21) - (35) are just the clauses in (1) - (15) with the additional disjuncts $\bigvee_{u\in[k]} \neg S_u(i,j)$ with the corresponding (i,j). Clauses (36) together with (27) make sure that $C_{s,t}$ is empty. Clauses in (37) and (38) ensure that if $C_{i-1,j'}$ is not selected then it cannot be used as a premise.

It is immediate that the partial assignment that maps $S_u(i,j)$ to 1 for all $(i,j)\in[s]\times[t]$ and all $u\in[k]$ maps $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ to $\mathrm{REF}^F_{s,t}$.

We now define the formula $\mathrm{R}^k\mathrm{REF}^{n,r}_{s,t}$ by a change to $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ analogous to the change by which we obtained $\mathrm{REF}^{n,r}_{s,t}$ from $\mathrm{REF}^F_{s,t}$. That is, the clauses of $\mathrm{R}^k\mathrm{REF}^{n,r}_{s,t}$ are (22) - (38) of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ together with the following clauses (to replace (21)):

$$\bigvee_{u\in[k]} \neg S_u(1,j) \vee \neg I(j,m) \vee \neg C(m,\ell,b) \vee D(1,j,\ell,b) \qquad j\in[t], m\in[r], \ell\in[n], b\in\{0,1\}, \tag{39}$$

saying that if clause $C_{1,j}$ is selected and is a weakening of clause C_m (described by C-variables), then it contains each literal of C_m.

Theorem 7. *The negation of the reflection principle for resolution expressed by the formula* $\mathrm{SAT}^{n,r} \wedge \mathrm{R}^k\mathrm{REF}^{n,r}_{s,t}$ *has Res(2) refutations of size* $O(trn^2 + tr^2 + trnk + st^2n^3 + st^2n^2k + st^2nk^2 + st^3n)$.

Proof. By induction on $i\in[s]$ we derive for each $j\in[t]$ the formula

$$D_{i,j} := \bigvee_{u\in[k]} \neg S(i,j) \vee \bigvee_{\ell\in[n],b\in\{0,1\}} \left(D(i,j,\ell,b) \wedge T(\ell)^b\right). \tag{40}$$

Then, cutting $D_{s,t}$ with (27) for each $\ell\in[n]$ and $b\in\{0,1\}$, followed by k cuts with clauses (36), yields the empty clause.

Base case: $i=1$. For each $j\in[t], m\in[r], \ell\in[n], b\in\{0,1\}$, cut (19) with (39) to obtain $\bigvee_{u\in[k]} \neg S_u(1,j) \vee \neg I(j,m) \vee \neg T(m,\ell,b) \vee D(1,j,\ell,b)$. Applying $\wedge$-introduction to this and $\neg T(m,\ell,b) \vee T(\ell)^b$ (which is either (17) or (18)) yields

$$\bigvee_{u\in[k]} \neg S_u(1,j) \vee \neg I(j,m) \vee \neg T(m,\ell,b) \vee \left(D(1,j,\ell,b) \wedge T(\ell)^b\right). \tag{41}$$

Cutting (41) for each $\ell \in [n]$ and $b \in \{0,1\}$ with (16) gives $\neg I(j,m) \vee D_{1,j}$. Cutting this for each $m \in [r]$ with (29) yields $D_{1,j}$.

Induction step: Assume we have derived $D_{i-1,j'}$ for all $j' \in [t]$. We derive $D_{i,j}$ for each $j \in [t]$. Write P_1 in place of L and P_0 in place of R.

For each $\ell \in [n], b \in \{0,1\}, j' \in [t]$, cut $\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg D(i-1,j',\ell,1) \vee \neg D(i-1,j',\ell,0)$ (from (22)) with $\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg P_{1-b}(i,j,j') \vee \neg V(i,j,\ell) \vee D(i-1,j',\ell,1-b)$ (which is from (23) or (24)) to obtain $\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg P_{1-b}(i,j,j') \vee \neg V(i,j,\ell) \vee \neg D(i-1,j',\ell,b)$. Cut this with $D_{i-1,j'}$ to get

$$\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg P_{1-b}(i,j,j') \vee \neg V(i,j,\ell) \vee \left(D_{i-1,j'} \setminus \{D(i-1,j',\ell,b) \wedge T(\ell)^b\}\right). \quad (42)$$

Cutting (42) with $T(\ell) \vee \neg T(\ell)$ yields

$$\begin{aligned}&\bigvee_{u\in[k]} \neg S(i,j) \vee \neg P_{1-b}(i,j,j') \vee \neg V(i,j,\ell) \vee T(\ell)^{1-b} \\ &\qquad \vee \left(D_{i-1,j'} \setminus \{D(i-1,j',\ell,0) \wedge \neg T(\ell), D(i-1,j',\ell,1) \wedge T(\ell)\}\right).\end{aligned} \quad (43)$$

Next, for each $\ell' \in [n] \setminus \{\ell\}$ and $b' \in \{0,1\}$, apply $\wedge$-introduction to $T(\ell') \vee \neg T(\ell')$ and $\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg P_{1-b}(i,j,j') \vee \neg V(i,j,\ell) \vee \neg D(i-1,j',\ell',b') \vee D(i,j,\ell',b')$ (from (25) or (26)) to get

$$\begin{aligned}&\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg P_{1-b}(i,j,j') \vee \neg V(i,j,\ell) \vee \left(D(i,j,\ell',b') \wedge T(\ell')^{b'}\right) \\ &\qquad \vee \neg D(i-1,j',\ell',b') \vee T(\ell')^{1-b'}.\end{aligned} \quad (44)$$

Cutting (44), for each $\ell' \in [n] \setminus \{\ell\}$ and $b' \in \{0,1\}$, with (43) results, after a weakening, in

$$\bigvee_{u\in[k]} \neg S_u(i-1,j') \vee \neg P_{1-b}(i,j,j') \vee \neg V(i,j,\ell) \vee T(\ell)^{1-b} \vee D_{i,j}. \quad (45)$$

Cut (45), for each $u' \in [k]$, with $\bigvee_{u\in[k]} \neg S_u(i,j) \vee \neg P_{1-b}(i,j,j') \vee S_{u'}(i-1,j')$ (from (37) or (38)) to get

$$\neg P_{1-b}(i,j,j') \vee \neg V(i,j,\ell) \vee T(\ell)^{1-b} \vee D_{i,j}. \quad (46)$$

Recall that we have obtained (46) for each $\ell \in [n], b \in \{0,1\}, j' \in [t]$. Cutting (46), for each $j' \in [t]$, with $\bigvee_{u\in[k]} \neg S_u(i,j) \vee \bigvee_{j'\in[t]} P_{1-b}(i,j,j')$ (which is from (30) or (31)) yields $\neg V(i,j,\ell) \vee T(\ell)^{1-b} \vee D_{i,j}$. We have derived such clause for each $\ell \in [n], b \in \{0,1\}$, so a cut on $T(\ell)$ gives $\neg V(i,j,\ell) \vee D_{i,j}$, and cutting this, for each $\ell \in [n]$, with (28) yields $D_{i,j}$.

As for bounding the size of the refutation, the size of the base case is $O(t(rn^2 + r^2 + rnk))$, the total size of the induction steps is $O(st(n^3t + n^2tk + ntk^2 + nt^2))$, and the size of the finish is $O(n^2 + nk)$. Altogether, this is $O(trn^2 + tr^2 + trnk + st^2n^3 + st^2n^2k + st^2nk^2 + st^3n)$. □

6 The Lower Bounds

We need a modification of two results of Segerlind, Buss and Impagliazzo [18]. Namely, their switching lemma works with the usual notion of width of a clause, and we would like it to work with the notion of 'number of pairs mentioned' in the sense of Definition 10 below. This is because our random restrictions have to respect the functional properties of the formula $\mathrm{REF}^F_{s,t}$ (expressed by clauses (8) - (15)), and it is therefore convenient to require that they evaluate variables in groups determined by home pair. Consequently, we do not want to represent a k-DNF simplified by a random restriction by a standard decision tree like in [18], as such a tree would branch exponentially in t, which would prevent taking union bounds over the branches of shallow trees occurring in the proof of our switching lemma. To circumvent this problem, the decision trees we construct (called decision trees over $\mathrm{REF}^F_{s,t}$) ask queries like "What is the left premise of clause $C_{i,j}$?" rather than queries like "Is $L(i,j,j')$ true?". This makes their branching a bit more manageable (though still exponential in the number of variables of F), but there is a price to pay in terms of parameters of the switching lemma (Theorem 20) and its more complicated proof, which uses certain independence properties of our random restrictions. Also, such trees no longer represent formulas over all partial assignments, but only over assignments that do not violate the functionality axioms and evaluate variables in groups determined by home pair. Accordingly, we need to adapt to our different notions of width and representation a result in [18] which says that if the lines of a $\mathrm{Res}(k)$ refutation can be strongly represented by shallow decision trees, the refutation can be converted into a resolution refutation of a small width.

Our random restrictions (Definition 19) will be applied to k-DNFs in the variables of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ and they are defined in two stages, the first of which evaluates all the S-variables, thereby declaring some pairs (i,j) selected (when $\bigwedge_{u\in[k]} S_u(i,j)$ evaluates to 1), and in the second stage all variables with a home pair that was not selected are evaluated randomly and independently. The restricted formula is therefore in the variables of $\mathrm{REF}^F_{s,t}$, and the purpose of the switching lemma is to show that it can be represented by a shallow decision tree over $\mathrm{REF}^F_{s,t}$ with a high probability. We begin with a definition of these trees and the notion of representation. Please recall Definition 5 before reading the next one.

Definition 8. A *decision tree over* $\mathrm{REF}^F_{s,t}$ is a rooted tree T in which every internal node is labelled with a pair $(i,j) \in [s] \times [t]$. There are $2^{2n} \cdot r$ edges leaving each node labelled with $(1,j) \in \{1\} \times [t]$, and they are labelled with pairs $(C_{1,j}, m)$, where $C_{1,j}$ is a clause in variables $x_1, \ldots, x_n$, and $m \in [r]$. There are $2^{2n} \cdot nt^2$ edges leaving each node labelled with $(i,j) \in \{2,\ldots,s\} \times [t]$, and these edges are labelled with tuples $(C_{i,j}, \ell, j', j'')$, where $C_{i,j}$ is a clause in variables $x_1,\ldots,x_n$, $\ell \in [n]$, and $j', j'' \in [t]$. The leaves of T are labelled with either 0 or 1. No pair (i,j) is allowed to label two nodes on any path from the root to a leaf of T. For each node v of T, the path from the root to v is viewed as a partial assignment π_v that for each edge that is on the path, leaving a node with a label (i,j), evaluates the variables of $\mathrm{REF}^F_{s,t}$ with home pair (i,j) in the following way: If $i = 1$ and the label of the edge is $(C_{1,j}, m)$, then π_v sets $D(1,j,\cdot,\cdot)$ to $C_{1,j}$ and $I(j,\cdot)$ to m; otherwise $i \in [s] \setminus \{1\}$ and the label of the edge is some tuple $(C_{i,j}, \ell, j', j'')$, in which case π_v sets $D(i,j,\cdot,\cdot)$ to $C_{i,j}$, $V(i,j,\cdot)$ to ℓ, $L(i,j,\cdot)$ to j', and $R(i,j,\cdot)$ to j''. For $b \in \{0,1\}$, we let $\mathrm{Br}_b(T)$ stand for the set of paths (viewed as partial assignments) that lead from the root to a leaf labelled with b.

Definition 9. Let G be a DNF in the variables of $\mathrm{REF}^F_{s,t}$. We say that a decision tree T over $\mathrm{REF}^F_{s,t}$ *strongly represents* G if for every $\pi \in \mathrm{Br}_0(T)$, for every $q \in G$, $q \restriction \pi = 0$ and for every $\pi \in \mathrm{Br}_1(T)$, there exists $q \in G$, $q \restriction \pi = 1$. The *representation index-height of* G, $h_{\mathrm{i}}(G)$, is the minimum height of a decision tree over $\mathrm{REF}^F_{s,t}$ strongly representing G.

Definition 10. Let π be a partial assignment to the variables of $\mathrm{REF}^F_{s,t}$, and let E be a clause in the variables of $\mathrm{REF}^F_{s,t}$. We say that a pair $(i,j) \in [s] \times [t]$ is *mentioned in* π (resp. E) if it is the home pair of a variable in $\mathrm{dom}(\pi)$ (resp. a literal of which is in E).

Definition 11. A partial assignment π to the variables of $\mathrm{REF}^F_{s,t}$ is called *respectful* if for each $(i,j) \in [s] \times [t]$, either (i,j) is not mentioned in π, or $i \in [s] \setminus \{1\}$ and each of $D(i,j,\cdot,\cdot)$, $V(i,j,\cdot)$, $R(i,j,\cdot)$, $L(i,j,\cdot)$ is set by π, or $i = 1$ and both $D(1,j,\cdot,\cdot)$ and $I(j,\cdot)$ are set by π. In other words, respectful assignments are exactly the assignments of the form π_v where v is a node of a decision tree over $\mathrm{REF}^F_{s,t}$.

If T is a decision tree over $\mathrm{REF}^F_{s,t}$ and π is a respectful partial assignment, $T \restriction \pi$ is obtained as follows: for each node v of T with a label (i,j) that is mentioned in π, contract the edge whose label determines an assignment to the variables with home pair (i,j) that is a subset of π, and delete all other edges leaving v (and delete their associated subtrees).

Lemma 12. *Let T be a decision tree over $\mathrm{REF}^F_{s,t}$, let G be a DNF, and let π be a respectful partial assignment. If T strongly represents G, then $T \restriction \pi$ strongly represents $G \restriction \pi$.*

Proof. For a leaf v of $T \restriction \pi$ there is a unique leaf u of T such that $\pi_v = \pi_u \setminus \pi$, where π_u, π_v are defined as in Definition 8. Moreover, v has the same label as u, and π and π_u are compatible. Therefore, for a term $q \in G$ we have $q \restriction (\pi \cup \pi_u) = q \restriction (\pi \cup \pi_v) = (q \restriction \pi) \restriction \pi_v$. Also, for $b \in \{0,1\}$, if $q \restriction \pi_u = b$ then $q \restriction (\pi \cup \pi_u) = b$. □

In the other direction, we have the following lemma.

Lemma 13. *Let T be a decision tree over $\mathrm{REF}^F_{s,t}$, and let G be a DNF in the variables of $\mathrm{REF}^F_{s,t}$. For each leaf v of T, let T_v be a decision tree that strongly represents $G \restriction \pi_v$, where π_v is the path in T from the root to v. Moreover, assume that each label (i,j) of an internal node of T_v is a home pair of a variable of $G \restriction \pi_v$. Then the tree T' obtained by appending to each leaf v of T the tree T_v strongly represents G.*

Proof. This follows directly from the definitions. □

Definition 14. Let C be a clause in the variables of $\mathrm{REF}^F_{s,t}$. The *index-width* of C is the number of pairs $(i,j) \in [s] \times [t]$ that are mentioned in C. The index-width of a resolution derivation is the maximum index-width of a clause in the derivation.

The following theorem is an adaptation of [18, Theorem 5.1].

Theorem 15. *Let H be a CNF in the variables of $\mathrm{REF}^F_{s,t}$ whose every clause has index-width at most $h \geq 1$. If for some $k \geq 1$ there is a $\mathrm{Res}(k)$ refutation of H such that for each line G of the refutation, $h_{\mathrm{i}}(G) \leq h$, then there is a resolution refutation of H together with the functionality clauses (8) - (15) of $\mathrm{REF}^F_{s,t}$ such that the index-width of the refutation is at most $3h$.*

Proof. Denote Π the $\mathrm{Res}(k)$ refutation. For a line G in Π, let T_G be a decision tree over $\mathrm{REF}^F_{s,t}$ of minimum height that strongly represents G. We can assume that no node of T_G is labelled with a pair (i,j) that is not a home pair of any variable of G.

For any respectful partial assignment π let C_π be the clause consisting of the following literals: $D(i,j,\ell,b)$ if and only if $\pi(D(i,j,\ell,b)) = 0$, $\neg D(i,j,\ell,b)$ if and only if $\pi(D(i,j,\ell,b)) = 1$, $\neg I(j,m)$ if and only if π sets $I(j,\cdot)$ to m, $\neg V(i,j,\ell)$ if and only if π

sets $V(i,j,\cdot)$ to ℓ, $\neg L(i,j,j')$ if and only if π sets $L(i,j,\cdot)$ to j', $\neg R(i,j,j')$ if and only if π sets $R(i,j,\cdot)$ to j'.

By induction on the lines of Π we show that for each line G of Π and for each $\pi \in \mathrm{Br}_0(T_G)$, there is a resolution derivation $\Pi_G(\pi)$ of C_π from H together with the clauses (8) - (15), such that the index-width of $\Pi_G(\pi)$ is at most $3h$. The theorem then follows from $\{C_\pi : \pi \in \mathrm{Br}_0(T_\emptyset)\} = \{C_\emptyset\} = \{\emptyset\}$.

Assume that G is an axiom $X \vee \neg X$. Then all the branches of T_G are labelled with 1, and so $\{C_\pi : \pi \in \mathrm{Br}_0(T_G)\} = \emptyset$.

Next assume that $G \in H$. Let $\pi \in \mathrm{Br}_0(T_G)$. Since G is a clause, the node labels of T_G are exactly the pairs (i,j) mentioned in G. Note that since $G \restriction \pi = 0$, for every (i,j) each literal of a variable in $D(i,j,\cdot,\cdot)$ that is in G is also in C_π. Suppose that π sets $V(i,j,\cdot)$ to $\ell \in [n]$. If there is a literal in G of a variable from $V(i,j,\cdot)$ such that the literal is not in C_π, then the literal must be $V(i,j,\ell')$ for some $\ell' \in [n]$ with $\ell' \neq \ell$. This follows from $G \restriction \pi = 0$ and $\neg V(i,j,\ell) \in C_\pi$. Such literals $V(i,j,\ell')$ can be removed from G by resolving with the clause $\neg V(i,j,\ell) \vee \neg V(i,j,\ell')$ from (12). Similarly, we remove from G the literals in $G \setminus C_\pi$ of I, L, R-variables by resolving with the corresponding clauses from (13), (14), (15), respectively. We have thus obtained a resolution derivation $\Pi_G(\pi)$ of C_π from $\{G\}$ together with the clauses (12) - (15). Because the index-width of G is at most h, the same is true for the clauses in $\Pi_G(\pi)$.

Now assume that line G in Π is inferred from previously derived lines $G_1, \ldots, G_d$ for $d \in [2]$. By the induction hypothesis, we have for each $c \in [d]$ and for each $\pi \in \mathrm{Br}_0(T_{G_c})$ a resolution derivation $\Pi_G(\pi)$ of C_π with the required properties. First construct a decision tree T as follows: if $d = 1$, T is T_{G_1}; if $d = 2$, append to each branch $\pi \in \mathrm{Br}_1(T_{G_1})$ the tree $T_{G_2} \restriction \pi$. Observe that for each $\pi \in \mathrm{Br}_0(T)$ there is $c \in [d]$ and $\pi' \in \mathrm{Br}_0(T_c)$ such that $\pi' \subseteq \pi$, and C_π is a weakening of $C_{\pi'}$. Also, the index-width of C_π is at most $2h$, because so is the height of T. For a node v of T define a partial assignment π_v as in Definition 8.

Let $\sigma \in \mathrm{Br}_0(T_G)$ be given. Inductively, from the leaves to the root of T, we show that if a node v of T is such that π_v is compatible with σ, then there is a resolution derivation $\Pi_G(\pi_v, \sigma)$ of $C_{\pi_v} \vee C_\sigma$ from H together with the clauses (8) - (15), such that the index-width of $\Pi_G(\pi_v, \sigma)$ is at most $3h$. When we reach the root of T, we will have obtained a derivation $\Pi_G(\emptyset, \sigma)$ of C_σ, and this is the derivation $\Pi_G(\sigma)$ we are after.

Assume that v is a leaf of T and π_v is compatible with σ. Then $\pi_v \in \mathrm{Br}_0(T)$. This can be seen as follows. It is easy to check that the rules of $\mathrm{Res}(k)$ have the property, called strong soundness, that any partial assignment that satisfies all premises of a rule also satisfies the conclusion of the rule. If we had $\pi_v \in \mathrm{Br}_1(T)$, then for each $c \in [d]$, π_v contains some $\pi_c \in \mathrm{Br}_1(T_{G_c})$, and so $G_c \restriction \pi_v = G_c \restriction \pi_c = 1$ because T_{G_c} strongly represents G_c. By strong soundness it follows that $G \restriction \pi_v = 1$. But this means that π_v cannot be compatible with σ, because σ falsifies every term of G. So indeed $\pi_v \in \mathrm{Br}_0(T)$. Further, we have that $C_{\pi_v} \vee C_\sigma$ is a weakening of C_{π_v}, which in turn is a weakening of $C_{\pi'}$ for some $\pi' \in \mathrm{Br}_0(T_c)$ and some $c \in [d]$ such that that $\pi' \subseteq \pi_v$, by the construction of T. By the inductive hypothesis we have a resolution refutation $\Pi_G(\pi')$ of $C_{\pi'}$ with the required properties. Because the index-width of C_{π_v} is at most $2h$, the index-width of $C_{\pi_v} \vee C_\sigma$ is at most $3h$. We have thus obtained a resolution derivation $\Pi_G(\pi_v, \sigma)$ of $C_{\pi_v} \vee C_\sigma$ with the required properties.

Now assume that v is labelled with a pair (i,j) and π_v is compatible with σ. We distinguish two cases. In the first case, assume that (i,j) is mentioned in σ. Then there is a child u of v such that $\pi_u \setminus \pi_v \subseteq \sigma$. Also, π_u is compatible with σ. By the induction hypothesis we therefore have a resolution refutation $\Pi_G(\pi_u, \sigma)$ of $C_{\pi_u} \vee C_\sigma$ with the required properties. Because $\pi_u \cup \sigma = \pi_v \cup \sigma$, we have $C_{\pi_u} \vee C_\sigma = C_{\pi_v} \vee C_\sigma$, and so we define $\Pi_G(\pi_v, \sigma)$ to be $\Pi_G(\pi_u, \sigma)$. In the second case, assume that (i, j) is not mentioned in σ. Then for each child u of v, π_u is compatible with σ. By the induction hypothesis, for each such u there is a resolution refutation $\Pi_G(\pi_u, \sigma)$ of $C_{\pi_u} \vee C_\sigma$ with the required properties. Notice that $C_{\pi_u} \vee C_\sigma = C_{\pi_u \setminus \pi_v} \vee C_{\pi_v} \vee C_\sigma$. We first construct a resolution refutation Π' of $\{C_{\pi_u \setminus \pi_v} : u \text{ is a child of } v\}$ together with the clauses (8) - (11) such that the index-width of Π' is 1. This is easy: since $\{C_{\pi_u \setminus \pi_v} : u \text{ is a child of } v\} = \{C_\alpha : \alpha \text{ is respectful and mentions just the pair } (i, j)\}$, we use (8), (10), (11) (resp. (9) if $i = 1$) to remove all the negated V, L, R-variables (resp. the negated I-variables) from the clauses C_α, and we refute the resulting clauses by a refutation in the form of a complete binary tree to resolve all the D-variables. Now, having Π', we define $\Pi_G(\pi_v, \sigma)$ as follows: add the literals of $C_{\pi_v} \vee C_\sigma$ to each clause of Π' other than an initial clause from (8), (10), (11), (9), and derive each initial clause $C_{\pi_u} \vee C_\sigma$ in the resulting derivation using the derivation $\Pi_G(\pi_u, \sigma)$. It is easy to see that $\Pi_G(\pi_v, \sigma)$ has the required properties. □

We now turn our attention to the formula $\mathrm{R}^k\mathrm{REF}^F_{s,t}$. Recall from its definition in Section 5 that its variables are those of $\mathrm{REF}^F_{s,t}$ together with variables $S_u(i, j)$, $(i, j) \in [s] \times [t]$, $u \in [k]$. In the following definition we extend the notion of home pair from Definition 5 to the S-variables, and we extend the notion of a pair being mentioned accordingly.

Definition 16. For $(i, j) \in [s] \times [t]$ and $u \in [k]$, the *home pair* of the variable $S_u(i, j)$ is (i, j).

We say that a pair (i, j) is *mentioned in* a clause E (resp. a partial assignment π; a term q) if it is a home pair of a variable a literal of which is in E (resp. which is in $\mathrm{dom}(\pi)$; a literal of which is in q).

Definition 17. Let $U \subseteq [s] \times [t]$ and let G be a DNF in the variables of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$. If for each term $q \in G$ there is $(i, j) \in U$ such that (i, j) is mentioned in q, then we say that U is an *index-cover* of G. The *index-covering number of* G, $c_\mathrm{i}(G)$, is the minimum cardinality of an index-cover of G.

Definition 18. For a set $U \subseteq [s] \times [t]$, denote by $\mathrm{Var}(U)$ the set of all variables of $\mathrm{REF}^F_{s,t}$ with home pair in U, that is,

$$\mathrm{Var}(U) := \bigcup_{(i,j)\in U} D(i, j, \cdot, \cdot) \cup \bigcup_{(i,j)\in U\setminus([1]\times[t])} (R(i, j, \cdot) \cup L(i, j, \cdot) \cup V(i, j, \cdot)) \cup \bigcup_{(1,j)\in U} I(j, \cdot).$$

Also, denote by $\mathrm{Var}_S(U)$ the set of all S-variables with home pair in U; in symbols, $\mathrm{Var}_S(U) := \{S_u(i, j) : u \in [k], (i, j) \in U\}$.

We generalize random restrictions from [3] to our case of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$.

Definition 19. A *random restriction* ρ_k is a partial assignment to the variables of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ given by the following experiment:

1. Independently for each $(i, j) \in [s] \times [t]$ and $u \in [k]$, map $S_u(i, j)$ to 0 or 1, each with probability $1/2$.

2. Let A be the set of those $(i, j) \in [s] \times [t]$ such that for every $u \in [k]$, $S_u(i, j)$ is mapped to 1.

3. Map independently each variable from $\mathrm{Var}(([s]\times[t])\setminus A)$ to 0 or 1, each with probability 1/2.

Theorem 20. *Suppose that $k \geq 1, a \geq 1$ are integers such that $k \geq a$. There is $\delta > 0$ and an integer $n_0 > 0$ such that if n, r, s, t are integers satisfying*

$$r \leq t \leq 2^{\delta n} \text{ and } n_0 \leq n, \tag{47}$$

and F is a CNF with r clauses in n variables, then for every a-DNF G in the variables of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ and every $w > 0$,

$$\Pr[h_{\mathrm{i}}(G \restriction \rho_k) > w] \leq 2^{-\frac{w}{n^{a-1}}\gamma(a)}, \tag{48}$$

where $\gamma(a) = \frac{(\log e)^a}{2^{a^2+3a-2}a!}$.

Proof. Denote the right hand side of the inequality (48) by $p_a(w)$. Let $k \geq 1$ be given and denote $\rho := \rho_k$. We prove the theorem by induction on a.

Base case: $a = 1$. G is a clause. If $c_{\mathrm{i}}(G) \leq w$, then $\Pr[h_{\mathrm{i}}(G \restriction \rho) > w] = 0$ because we can build a decision tree strongly representing $G \restriction \rho$ by querying the pairs from the smallest index-cover of G. If $c_{\mathrm{i}}(G) > w$, we have $\Pr[h_{\mathrm{i}}(G \restriction \rho) > w] \leq \Pr[G \restriction \rho \neq 1] \leq \big(1-(1-2^{-k})/2\big)^{c_{\mathrm{i}}(G)} \leq (1-1/4)^{c_{\mathrm{i}}(G)} \leq e^{-c_{\mathrm{i}}(G)/4} = 2^{-c_{\mathrm{i}}(G)\gamma(1)} \leq 2^{-w\gamma(1)}$.

Induction step: Assume the theorem holds for $a-1$, witnessed by $\delta(k, a-1)$ and $n_0(k, a-1)$. Find a positive $\delta(k,a) \leq \delta(k, a-1)$ and an integer $n_0(k,a) \geq n_0(k,a-1)$ such that

$$-\frac{\gamma(a-1)}{2}n + \left(2\log t + \log n + \frac{\gamma(a-1)}{n^{a-2}}\right)\cdot\frac{\gamma(a-1)}{4} \leq -\gamma(a) \tag{49}$$

holds for any n, r, t satisfying 47 with $\delta(k,a)$ and $n_0(k,a)$ in place of δ and n_0, respectively. Let G be an a-DNF, and let U be an index cover of G of size $c_{\mathrm{i}}(G)$. We distinguish two cases based on $c_{\mathrm{i}}(G)$.

Case 1: $c_{\mathrm{i}}(G) > \frac{w}{n^{a-1}}\cdot\frac{\gamma(a-1)}{4}$. In this case we want to show that ρ satisfies G with a high probability. To this end, note that there are at least $c_{\mathrm{i}}(G)/a$ many terms in G that are index-independent, that is, for no two of them there is a pair $(i,j) \in [s]\times[t]$ mentioned by both. (If every such set of terms was smaller than $c_{\mathrm{i}}(G)/a$, take a maximal one and observe that the set of pairs mentioned by the terms forms an index-cover of G of cardinality smaller than $c_{\mathrm{i}}(G)$, a contradiction.) It is easy to see that each of these index-independent terms is satisfied by ρ with independent probability at least 2^{-2a}. Therefore,

$$\begin{aligned}\Pr[h_{\mathrm{i}}(G \restriction \rho) > w] &\leq \Pr[G \restriction \rho \neq 1] \leq \big(1-2^{-2a}\big)^{c_{\mathrm{i}}(G)/a} \leq 2^{-\frac{(\log e)}{a2^{2a}}c_{\mathrm{i}}(G)} \leq 2^{-\frac{(\log e)}{a2^{2a}}\cdot\frac{w}{n^{a-1}}\cdot\frac{\gamma(a-1)}{4}} \\ &= 2^{-\frac{w}{n^{a-1}}\gamma(a)}.\end{aligned}$$

This finishes the inductive step for Case 1.

Case 2: $c_{\mathrm{i}}(G) \leq \frac{w}{n^{a-1}}\cdot\frac{\gamma(a-1)}{4}$. Let $U' \subseteq U$, and let $\nu : \mathrm{Var}_S(U) \cup \mathrm{Var}(U \setminus U') \to \{0,1\}$ satisfy the following conditions:

(ν1) for each $(i,j) \in U'$ and each $u \in [k]$, $\nu(S_u(i,j)) = 1$,

(ν2) for each $(i,j) \in U \setminus U'$ there is $u \in [k]$ with $\nu(S_u(i,j)) = 0$.

We have

$$\begin{aligned}
&\Pr[h_{\mathrm{i}}(G{\restriction}\rho) > w \mid \rho{\restriction}\mathrm{dom}(\nu) = \nu] \\
&\le \Pr[\exists\pi : \mathrm{Var}(U') \to \{0,1\}, \pi \text{ is respectful} \wedge h_{\mathrm{i}}((G{\restriction}\pi){\restriction}\rho) > w - |U'| \mid \rho{\restriction}\mathrm{dom}(\nu) = \nu] \\
&\le \sum_{\substack{\pi:\mathrm{Var}(U')\to\{0,1\},\\ \pi \text{ is respectful}}} \Pr[h_{\mathrm{i}}((G{\restriction}\pi){\restriction}\rho) > w - |U'| \mid \rho{\restriction}\mathrm{dom}(\nu) = \nu] \\
&= \sum_{\substack{\pi:\mathrm{Var}(U')\to\{0,1\},\\ \pi \text{ is respectful}}} \Pr[h_{\mathrm{i}}(((G{\restriction}\pi){\restriction}\nu){\restriction}\rho) > w - |U'|] \\
&\le \left(t^2 n 2^{2n}\right)^{|U'|} p_{a-1}(w - |U'|).
\end{aligned}$$

Here the first inequality follows from Lemma 13 and from $(G{\restriction}\pi){\restriction}\rho = (G{\restriction}\rho){\restriction}\pi$ (since $\mathrm{dom}(\pi) \cap \mathrm{dom}(\rho) = \emptyset$). The second inequality is obtained by the union bound. The equality follows since the events $h_{\mathrm{i}}(((G \restriction \pi) \restriction \nu) \restriction \rho) > w - |U'|$ and $\rho \restriction \mathrm{dom}(\nu) = \nu$ are independent (by the definition of ρ). And the last inequality is by the induction hypothesis and by the upper bound $t^2 n 2^{2n} = \max\{t^2 n 2^{2n}, r 2^{2n}\}$ (recall that $t \ge r$) over $(i, j) \in [s] \times [t]$ on the number of respectful partial assignments mentioning exactly the pair (i, j).

Since the event $A \cap U = U'$ (where the random variable A is given by Definition 19) is the disjoint union of events $\rho \restriction \mathrm{dom}(\nu) = \nu$ over all ν satisfying conditions (ν1) and (ν2), the above calculation implies

$$\Pr[h_{\mathrm{i}}(G{\restriction}\rho) > w \mid A \cap U = U'] \le \left(t^2 n 2^{2n}\right)^{|U'|} p_{a-1}(w - |U'|). \tag{50}$$

Therefore,

$$\begin{aligned}
\Pr[h_{\mathrm{i}}(G{\restriction}\rho) > w] &= \sum_{U' \subseteq U} \Pr[h_{\mathrm{i}}(G{\restriction}\rho) > w \wedge A \cap U = U'] \\
&= \sum_{U' \subseteq U} \Pr[h_{\mathrm{i}}(G{\restriction}\rho) > w \mid A \cap U = U'] \cdot \Pr[A \cap U = U'] \\
&\le \sum_{U' \subseteq U} \left(t^2 n 2^{2n}\right)^{|U'|} p_{a-1}(w - |U'|) \cdot 2^{-k|U'|} \left(1 - 2^{-k}\right)^{|U \setminus U'|} \\
&= \sum_{q=0}^{c_{\mathrm{i}}(G)} \binom{c_{\mathrm{i}}(G)}{q} \left(t^2 n 2^{2n}\right)^q p_{a-1}(w - q) \cdot 2^{-kq} \left(1 - 2^{-k}\right)^{c_{\mathrm{i}}(G) - q} \\
&\le \left(t^2 n 2^{2n}\right)^{c_{\mathrm{i}}(G)} p_{a-1}(w - c_{\mathrm{i}}(G)).
\end{aligned} \tag{51}$$

Here the first inequality is by 50 and by the definition of ρ. The second inequality follows from $(t^2 n 2^{2n})^q\, p_{a-1}(w - q) \le (t^2 n 2^{2n})^{c_{\mathrm{i}}(G)}\, p_{a-1}(w - c_{\mathrm{i}}(G))$ for $q \le c_{\mathrm{i}}(G)$. From 51, using the definition of $p_{a-1}(w - c_{\mathrm{i}}(G))$ and the assumption $c_{\mathrm{i}}(G) \leq \frac{w}{n^{a-1}} \cdot \frac{\gamma(a-1)}{4}$, we get

$$\begin{aligned}
\log(\Pr[h_{\mathrm{i}}(G \restriction \rho) > w]) &\leq (2\log t + \log n + 2n)\, c_{\mathrm{i}}(G) - \frac{w - c_{\mathrm{i}}(G)}{n^{a-2}} \gamma(a-1) \\
&= \left(2\log t + \log n + 2n + \frac{\gamma(a-1)}{n^{a-2}}\right) c_{\mathrm{i}}(G) - \frac{w\gamma(a-1)}{n^{a-2}} \\
&\leq \left(2\log t + \log n + 2n + \frac{\gamma(a-1)}{n^{a-2}}\right) \frac{w}{n^{a-1}} \cdot \frac{\gamma(a-1)}{4} - \frac{w\gamma(a-1)}{n^{a-2}} \\
&= -\frac{w\gamma(a-1)}{2n^{a-2}} + \left(2\log t + \log n + \frac{\gamma(a-1)}{n^{a-2}}\right) \frac{w}{n^{a-1}} \cdot \frac{\gamma(a-1)}{4} \\
&\leq -\frac{w}{n^{a-1}} \gamma(a),
\end{aligned}$$

where the last inequality is equivalent to 49. This finishes the inductive step for Case 2, and the proof of the theorem. □

We now show an index-width lower bound on resolution refutations of $\mathrm{REF}^F_{s,t}$ for an unsatisfiable F. This was done in [3] for a non-layered version of the formula, of which our $\mathrm{REF}^F_{s,t}$ is a restriction, so the index-width lower bound we need does not immediately follow from that in [3]. We provide a simpler proof for $\mathrm{REF}^F_{s,t}$. First a definition.

Definition 21. A partial assignment σ to the variables of $\mathrm{REF}^F_{s,t}$ is called *admissible* if it satisfies all the following conditions.

(A1) For each $(i,j) \in [s] \times [t]$, $D(i,j,\cdot,\cdot)$ (resp. $V(i,j,\cdot)$, $I(j,\cdot)$, $L(i,j,\cdot)$, $R(i,j,\cdot)$) either is set to some clause (resp. some $\ell \in [n]$, some $m \in [r]$, some $j' \in [t]$, some $j' \in [t]$) by σ or contains no variable that is in $\mathrm{dom}(\sigma)$.

(A2) For each $(i,j) \in [s] \times [t]$, if $L(i,j,\cdot)$ or $R(i,j,\cdot)$ is set to some $j' \in [t]$, then both $D(i,j,\cdot,\cdot)$ and $D(i-1,j',\cdot,\cdot)$ are set.

(A3) For each $(i,j) \in ([s] \setminus \{1\}) \times [t]$, $D(i,j,\cdot,\cdot)$ is set if and only if $V(i,j,\cdot)$ is set. For each $j \in [t]$, $D(1,j,\cdot,\cdot)$ is set if and only if $I(j,\cdot)$ is set.

(A4) For each $(i,j) \in [s] \times [t]$, if $D(i,j,\cdot,\cdot)$ is set to a clause $C_{i,j}$, then $C_{i,j}$ is non-tautological and has at least $\min\{s-i, n\}$ many literals. If $D(i,j,\cdot,\cdot)$ is set to a clause $C_{i,j}$ with less than n literals and $V(i,j,\cdot)$ is set to some $\ell \in [n]$, then none of the literals of x_ℓ is in $C_{i,j}$.

(A5) If $D(s,t,\cdot,\cdot)$ is set, it is set to the empty clause.

(A6) For each $j \in [t]$, if $I(j,\cdot)$ is set, then σ satisfies all clauses in (1) with this j.

(A7) For each $(i,j) \in ([s] \setminus \{1\}) \times [t]$, if $L(i,j,\cdot)$ (resp. $R(i,j,\cdot)$) is set, then σ satisfies all clauses in (3) and (5) (resp. (4) and (6)) with this (i,j) (i.e., those clauses that contain the literal $\neg L(i,j,j')$ (resp. $\neg R(i,j,j')$) for some $j' \in [t]$).

Theorem 22. *Let $w > 0$. If n, r, s, t are integers satisfying*

$$2 \leq n + 1 \leq s, \quad 2w < t, \tag{52}$$

and F is an unsatisfiable CNF consisting of r clauses $C_1, \ldots, C_r$ in n variables $x_1, \ldots, x_n$, then any resolution refutation of $\mathrm{REF}^F_{s,t}$ has index-width greater than w.

Proof. Assume for a contradiction that there is a resolution refutation Π of $\mathrm{REF}^F_{s,t}$ of index-width at most w. We will show that if there is an admissible partial assignment falsifying a clause E in Π obtained by the resolution rule from E_0 and E_1, then there is an admissible partial assignment falsifying either E_0 or E_1. This immediately (by induction) leads to a contradiction, since the empty assignment is admissible and falsifies the last (empty) clause in Π, and, by definition, no partial admissible assignment falsifies any clause of $\mathrm{REF}^F_{s,t}$.

Let then σ be an admissible partial assignment falsifying a clause E in Π. Without loss of generality, assume that σ is a minimal (with respect to inclusion) admissible partial assignment with this property.

Let Q be the variable resolved on to obtain E from E_0 and E_1. If $Q \in \mathrm{dom}(\sigma)$, then σ already falsifies either E_0 or E_1. So assume that $Q \notin \mathrm{dom}(\sigma)$. We consider two cases.

Case 1. Suppose that for some $(i,j) \in [s] \times [t]$, $Q \in D(i,j,\cdot,\cdot)$ or $Q \in V(i,j,\cdot)$ (resp. $Q \in I(j,\cdot)$ and $i = 1$). Note that by (A1), (A2), and (A3), no variable from $D(i,j,\cdot,\cdot) \cup V(i,j,\cdot) \cup L(i,j,\cdot) \cup R(i,j,\cdot)$ (resp. $D(1,j,\cdot,\cdot) \cup I(j,\cdot)$) is in $\mathrm{dom}(\sigma)$, and, moreoever, for any $j' \in [t]$, it is not the case that $L(i+1,j',\cdot)$ or $R(i+1,j',\cdot)$ is set to j by σ. Therefore, we can extend σ to a partial assignment σ' as follows. Set $D(i,j,\cdot,\cdot)$ to any non-tautological clause containing n literals, unless $(i,j) = (s,t)$, in which case set $D(i,j,\cdot,\cdot)$ to the empty clause. In case $i \geq 2$, set $V(i,j,\cdot)$ to an arbitrary value $\ell \in [n]$; in case $i = 1$, set $I(j,\cdot)$ to any $m \in [r]$ such that the clause C_m is a subset of the clause to which we have set $D(1,j,\cdot,\cdot)$. (Here we use that F is unsatisfiable.) It is straightforward to check that σ' is admissible. Since $Q \in \mathrm{dom}(\sigma')$, σ' falsifies $E \cup \{Q^{1-\sigma'(Q)}\}$, of which either E_0 or E_1 is a subset.

Case 2. Suppose that for some $(i,j) \in ([s] \setminus \{1\}) \times [t]$, $Q \in L(i,j,\cdot)$ (if $Q \in R(i,j,\cdot)$, we proceed in a completely analogous way). We may assume that $D(i,j,\cdot,\cdot)$ is set to some clause $C_{i,j}$ by σ and $V(i,j,\cdot)$ is set to some $\ell \in [n]$ by σ; if not, set them both as described in Case 1. We now concentrate on the level $i-1$. Since the index-width of E is at most w and σ is a minimal admissible partial assignment falsifying E,

$$|\{j' : D(i-1,j',\cdot,\cdot) \text{ is set by } \sigma\}| \leq 2w. \tag{53}$$

This is because $D(i-1,j',\cdot,\cdot)$ can be set by σ for two reasons: either $(i-1,j')$ is mentioned in E (which, together with (A2) and (A3), implies that $D(i-1,j',\cdot,\cdot)$ is set by σ) or there is some $j'' \in [t]$ such that a literal of a variable from $L(i,j'',\cdot)$ or $R(i,j'',\cdot)$ is in E (which forces σ to set $L(i,j'',\cdot)$ or $R(i,j'',\cdot)$, respectively, in order to falsify the literal) and σ happens to set $L(i,j'',\cdot)$ or $R(i,j'',\cdot)$, respectively, to j' (and therefore by (A2) $D(i-1,j',\cdot,\cdot)$ must be set by σ too).

We extend σ to a partial assignment σ' as follows. Set $L(i,j,\cdot)$ to any j' that is not from the set in (53). Such j' exists because $2w < t$. Thanks to that, set $D(i-1,j',\cdot,\cdot)$ to the clause $C_{i-1,j'} := (C_{i,j} \setminus \{\neg x_\ell\}) \cup \{x_\ell\}$, where $C_{i,j}$ and ℓ are as above. Finally, if $i \in \{3,\ldots,s\}$, then either $C_{i-1,j'}$ has less than n literals and we set $V(i-1,j',\cdot)$ to any $\ell' \in [n]$ such that no literal of $x_{\ell'}$ is in $C_{i-1,j'}$, or $C_{i-1,j'}$ has n literals, in which case we set $V(i-1,j',\cdot)$ arbitrarily. If $i = 2$, then by (A4), (52), and the definition of $C_{i-1,j'}$ we know that $C_{i-1,j'}$ has n literals, and we set $I(j',\cdot)$ to any $m \in [r]$ such that $C_m \subseteq C_{i-1,j'}$. (Here we use that F is unsatisfiable.) This finishes the definition of σ'.

It is again easy to verify that σ' is admissible. Because $Q \in \mathrm{dom}(\sigma')$, σ' falsifies $E \cup \{Q^{1-\sigma'(Q)}\}$, of which one of E_0, E_1 is a subset. □

We now put together all the results so far in this section to show a length lower bound on $\mathrm{Res}(k)$ refutations of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ with an unsatisfiable F.

Theorem 23. *Suppose $k \geq 1$ is an integer. There is $\delta > 0$ and an integer $n_0 > 0$ such that if n, r, s, t are integers satisfying*

$$n_0 \leq n, \quad n+1 \leq s \leq t, \quad r \leq t \leq 2^{\delta n}, \quad n^k \leq t, \tag{54}$$

and F is an unsatisfiable CNF consisting of r clauses $C_1, \ldots, C_r$ in n variables $x_1, \ldots, x_n$, then any Res(k) refutation of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ has length greater than $2^{\beta(k)\frac{t}{n^{k-1}}}$, where $\beta(k) := \frac{(\log e)^k}{2^{k^2+4k+4}k!}$.

Proof. Let $k \geq 1$ be given. Take δ and n_0 as given by Theorem 20 for $a = k$. If necessary, increase n_0 so that it satisfies

$$\beta(k) n_0 > k+1. \tag{55}$$

Let n, r, s, t be integers satisfying (54), and let F satisfy the hypothesis of the theorem. Assume for a contradiction that there is a Res(k) refutation Π of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ of length at most $2^{\beta(k)\frac{t}{n^{k-1}}}$.

Recall the random variable A from Definition 19. We have that with probability 2^{-k},

(a) $(s,t) \in A$.

By the Chernoff bound and the union bound, with probability at least $1 - se^{-t2^{-k}/8}$,

(b) for each $i \in [s]$ the cardinality of $A \cap (\{i\} \times [t])$ is at least $t/2^{k+1}$.

We have

$$se^{-t2^{-k}/8} = 2^{\log s - \frac{t \log e}{2^{k+3}}} \leq 2^{\log n_0 - \frac{n_0 \log e}{2^{k+3}}} < 2^{-(k+1)},$$

where we used $s \leq t$, $n_0 \leq s$ (from (54)), and (55).

By Theorem 20 and the union bound, with probability at least $1 - |\Pi| \cdot 2^{-\frac{t}{n^{k-1}2^{k+5}}\gamma(k)}$,

(c) for every line G in Π, $h_{\mathrm{i}}(G \restriction \rho_k) \leq t/2^{k+5}$.

We have

$$|\Pi| \cdot 2^{-\frac{t}{n^{k-1}2^{k+5}}\gamma(k)} \leq 2^{\beta(k)\frac{t}{n^{k-1}}} \cdot 2^{-\frac{t}{n^{k-1}2^{k+5}}\gamma(k)} = 2^{-\beta(k)\frac{t}{n^{k-1}}} \leq 2^{-\beta(k)n_0} < 2^{-(k+1)},$$

where we used $n^k \leq t$, $n_0 \leq n$ (from (54)), and (55).

It follows that there exists ρ_k such that (a), (b) and (c) hold. Fix any such ρ_k and denote it by ρ. We now restrict $\mathrm{R}^k\mathrm{REF}^F_{s,t} \restriction \rho$ some more before we apply Theorem 15.

For each level $i \in [s]$ select any $t' := \lfloor t/2^{k+1} \rfloor - 2$ home pairs (i,j) of variables of $\mathrm{R}^k\mathrm{REF}^F_{s,t} \restriction \rho$ (they exist thanks to (b)), making sure to include the pair (s,t) in the selection. Denote the set of selected pairs by B. Define a partial assignment $\nu : \mathrm{Var}(\mathrm{R}^k\mathrm{REF}^F_{s,t} \restriction \rho) \to \{0,1\}$ by mapping all the variables with not selected home pairs so that they form an arbitrary resolution derivation from F, that is, so that ν satisfies every clause of $\mathrm{R}^k\mathrm{REF}^F_{s,t} \restriction \rho$ that contains a literal of a variable in $\mathrm{dom}(\nu)$. (This derivation may require two clauses per level, which is why we selected only $\lfloor t/2^{k+1} \rfloor - 2$ on each level.) Note that ν is respectful. Hence by (c) and Lemma 12 we have that for any line G in $\Pi \restriction \rho$, $h_{\mathrm{i}}(G \restriction \nu) \leq t/2^{k+5}$.

Next, define a partial assignment λ as follows. For every $(i,j) \in B \setminus (\{1\} \times [t])$ and every $j' \in [t]$ such that $(i-1, j') \notin B$, map both $L(i,j,j')$ and $R(i,j,j')$ to 0. Let us verify that $((\mathrm{R}^k\mathrm{REF}^F_{s,t} \restriction \rho) \restriction \nu) \restriction \lambda$ is $\mathrm{REF}^F_{s,t'}$ up to a re-indexing of variables determined by a bijection that maps, for each $i \in [s]$, the elements of $B \cap (\{i\} \times [t])$ to $(i,1), \ldots, (i,t')$.

Thanks to (a), clauses (36) are satisfied by ρ. All clauses (37) and (38) are satisfied: if $(i,j) \in B$ and $(i-1,j') \notin B$, then the clause is satisfied by λ, otherwise it is satisfied by ρ or ν. Clauses (21) - (35) with $(i,j) \notin B$ are satisfied either by ρ (if $(i,j) \notin A$) or by ν. Clauses (21) - (35) with $(i,j) \in B$ become, after removing those clauses (23) - (26) that are satisfied by λ and after the re-indexing of variables, the clauses (1) - (15) with t replaced by t'. (Here notice that clauses (27) become (7) thanks to $(s,t) \in B$.) Hence $((\mathrm{R}^k\mathrm{REF}^F_{s,t} \restriction \rho) \restriction \nu) \restriction \lambda$ is indeed $\mathrm{REF}^F_{s,t'}$ up to the re-indexing of variables.

Let us now show that for a line G in $(\Pi \restriction \rho) \restriction \nu$ we have that $G \restriction \lambda$ is, after the re-indexing of variables, strongly represented by a decision tree over $\mathrm{REF}^F_{s,t'}$ of height at most $t/2^{k+5}$. As we already verified, $h_{\mathrm{i}}(G) \leq t/2^{k+5}$, and therefore there is a tree T over $\mathrm{REF}^F_{s,t}$ of minimum height which strongly represents G and whose height is at most $t/2^{k+5}$. Define a tree $T \restriction \lambda$ by deleting all edges (and the corresponding subtrees) in T whose label is of the form $(C_{i,j}, \ell, j', j'')$ with $(i-1,j') \notin B$ or $(i-1,j'') \notin B$. $T \restriction \lambda$ is, after relabelling its nodes and edges according to the re-indexing bijection, a decision tree over $\mathrm{REF}^F_{s,t'}$. With every branch π of $T \restriction \lambda$ we associate a partial assignment $\pi_{T\restriction\lambda} : \mathrm{Var}(((\mathrm{R}^k\mathrm{REF}^F_{s,t} \restriction \rho) \restriction \nu) \restriction \lambda) \to \{0,1\}$ defined via the re-indexing bijection and Definition 8, understanding the relabelled $T \restriction \lambda$ as a tree over $\mathrm{REF}^F_{s,t'}$. But every branch π of $T \restriction \lambda$ is also a branch of T, hence Definition 8 with T (which is a tree over $\mathrm{REF}^F_{s,t}$) says how π should be viewed as a partial assignment to $\mathrm{Var}(\mathrm{REF}^F_{s,t})$; let us denote the partial assignment by π_T for clarity. It is easy to see from the definitions that for every branch π in $T \restriction \lambda$, $\mathrm{dom}(\lambda) \cap \mathrm{dom}(\pi_{T\restriction\lambda}) = \emptyset$ and $\pi_T \subseteq \lambda \cup \pi_{T\restriction\lambda}$. It follows that $G \restriction \lambda$ is strongly represented by $T \restriction \lambda$. The tree $T \restriction \lambda$ has, of course, height at most $t/2^{k+5}$.

We can now apply Theorem 15 taking $\mathrm{REF}^F_{s,t'}$ (i.e., the re-indexed $((\mathrm{R}^k\mathrm{REF}^F_{s,t} \restriction \rho) \restriction \nu) \restriction \lambda)$ for H, t' for t, and $t/2^{k+5}$ for h, to obtain a resolution refutation of $\mathrm{REF}^F_{s,t'}$ of index-width at most $3t/2^{k+5}$.

But we have

$$2 \cdot 3t/2^{k+5} < t/2^{k+2} < \lfloor t/2^{k+1} \rfloor - 2 = t',$$

where the second inequality follows from 54 and 55. Therefore, we can use Theorem 22, taking $3t/2^{k+5}$ for w and t' for t, to conclude that any resolution refutation of $\mathrm{REF}^F_{s,t'}$ has index-width greater than $3t/2^{k+5}$. That is a contradiction. □

7 Proofs of Theorems 1 and 2

Proof of Theorem 1. Denote by F the well-known CNF $\neg\mathrm{PHP}^{n+1}_n$ called the negation of the pigeonhole principle, expressing that a multi-valued function from $n+1$ to n is injective. It consists of $r := n + 1 + (n^3 + n^2)/2$ clauses in $\widetilde{n} := (n+1)n$ variables.

Define $A_n := \mathrm{SAT}^{\widetilde{n},r} \restriction \gamma_F$, where γ_F is as in Proposition 6.

Since by [15, 16] there exists $\alpha > 0$ and an integer n_1 such that for every $n \geq n_1$, $\neg\mathrm{PHP}^{n+1}_n$ has no Res(k) refutations of size at most 2^{n^α}, the same is true for A_n. This is because by Proposition 6 there is a substitution τ such that $A_n \restriction \tau$ is $\neg\mathrm{PHP}^{n+1}_n$ together with some tautological clauses, and if Π is a Res(k) refutation of A_n then $\Pi \restriction \tau$ is a Res(k) refutation of $A_n \restriction \tau$. This shows item (i).

Define $B_{n,k} := \mathrm{R}^k\mathrm{REF}^F_{s,t}$, where we set $s := \widetilde{n} + 1$ and $t := \widetilde{n}^k$.

Let $\delta > 0$ and integer n_0 witness Theorem 23. Set $n_2 \geq n_0$ so that the hypotheses (54) with $\widetilde{n}$ in place of n hold with our choice of r, s, t (as functions of $\widetilde{n}$) for all $\widetilde{n} \geq n_2$. By that theorem, for every $\widetilde{n} \geq n_2$, any Res(k) refutation of $B_{n,k}$ has size greater than $2^{\beta(k)\widetilde{n}}$. Item (ii) follows.

Note that $\mathrm{R}^k\mathrm{REF}^{\widetilde{n},r}_{s,t} \restriction \gamma_F$ is $\mathrm{R}^k\mathrm{REF}^{F}_{s,t}$, because γ_F turns the clauses (39) into (21) (and the clauses satisfied by γ_F are removed). By Theorem 7 there is a Res(2) refutation of $\mathrm{SAT}^{\widetilde{n},r} \wedge \mathrm{R}^k\mathrm{REF}^{\widetilde{n},r}_{s,t}$ of size $O(k^2\widetilde{n}^{3k+3})$. Hence the same holds true for $A_n \wedge B_{n,k} = \mathrm{SAT}^{\widetilde{n},r} \restriction \gamma_F \wedge \mathrm{R}^k\mathrm{REF}^{\widetilde{n},r}_{s,t} \restriction \gamma_F$. This gives item (iii). □

Theorem 2 follows immediately from the more general Theorem 24 below. A function $T : \mathbb{N} \to \mathbb{N}$ is called *time-constructible* if there is an algorithm that when given 1^n (the string of n many 1's) computes $1^{T(n)}$ in time $O(T(n))$. We call a function $T : \mathbb{N} \to \mathbb{N}$ *subexponential* if $T(n) \leq 2^{n^{o(1)}}$.

Theorem 24. *Let $T : \mathbb{N} \to \mathbb{N}$ be time-constructible, non-decreasing and subexponential. If there is an integer $k \geq 1$ such that* Res(k) *is automatable in time T, then there are $c_1, c_2, c_3, c_4 > 0$ and an algorithm that when given as input a 3-CNF F in n variables decides in time $c_3(T(c_1 n^{c_2 k}) + n^k)^{c_4}$ whether F is satisfiable.*

Proof. Assume that for some integer $k \geq 1$ the system Res(k) is automatable in time T satisfying the assumptions of the theorem. Set r, s and t as functions of n as follows: $r := \binom{2n}{3}$, $s := n + 1$, $t := n^{k+3}$.

By Theorem 7 there are integers $c_1, c_2 > 0$ such that $\mathrm{SAT}^{n,r} \wedge \mathrm{R}^k\mathrm{REF}^{n,r}_{s,t}$ has a Res(2) refutation Π of size at most $c_1 n^{c_2 k}$; if necessary, increase c_1 and c_2 so that the size of Π plus the size of the formula $\mathrm{R}^k\mathrm{REF}^{n,r}_{s,t}$ is at most $c_1 n^{c_2 k}$.

Let $\delta > 0$ and integer $n_0 > 0$ witness Theorem 23. Let $n_1 > n_0$ be such that for all $n \geq n_1$,

$$r \leq t \leq 2^{\delta n} \tag{56}$$

and

$$2^{\beta(k)\frac{t}{n^{k-1}}} > T(c_1 n^{c_2 k}), \tag{57}$$

where $\beta(k)$ is as in Theorem 23. Here we use that T is subexponential.

Define algorithm M as follows. Given as input a 3-CNF F in n variables, check if $n \geq n_1$. If $n < n_1$, use brute force to decide if F is satisfiable or not, and output the answer. If $n \geq n_1$, compute the formula $\mathrm{R}^k\mathrm{REF}^{F}_{s,t}$ and run the automating algorithm on this formula for up to $T(c_1 n^{c_2 k})$ steps. If the automating algorithm returns a Res(k) refutation of $\mathrm{R}^k\mathrm{REF}^{F}_{s,t}$, then output 'satisfiable'. Else output 'unsatisfiable'.

Since both computing $\mathrm{R}^k\mathrm{REF}^{F}_{s,t}$ from F and checking whether the output of the automating algorithm is a Res(k) refutation of $\mathrm{R}^k\mathrm{REF}^{F}_{s,t}$ are polynomial-time procedures, and since T is time-constructible, it follows that there are $c_3, c_4 > 0$ such that the running time of M is at most $c_3(T(c_1 n^{c_2 k}) + n^k)^{c_4}$. It suffices to show that M gives the correct answer on 3-CNFs F in $n \geq n_1$ variables such that each clause of F has exactly three literals. Let F be such a 3-CNF, and let r' be the number of its clauses. We have $r' \leq r = \binom{2n}{3}$.

Assume first that F is satisfiable. Let γ_F and τ be as in Proposition 6, and let ν be a satisfying assignment for F. We have

$$(((\mathrm{SAT}^{n,r'} \wedge \mathrm{R}^k\mathrm{REF}^{n,r'}_{s,t}) \restriction \gamma_F) \restriction \tau) \restriction \nu = ((\mathrm{SAT}^{n,r'} \restriction \gamma_F) \restriction \tau) \restriction \nu \wedge \mathrm{R}^k\mathrm{REF}^{n,r'}_{s,t} \restriction \gamma_F = \mathrm{R}^k\mathrm{REF}^{F}_{s,t},$$

because by Proposition 6, $(\mathrm{SAT}^{n,r'} \restriction \gamma_F) \restriction \tau$ is F together with some tautological clauses in the variables $x_1, \ldots, x_n$. Let Π' be the Res(2) refutation of $\mathrm{SAT}^{n,r'} \wedge \mathrm{R}^k\mathrm{REF}^{n,r'}_{s,t}$ given by Theorem 7. Then $\Pi'' := ((\Pi' \restriction \gamma_F) \restriction \tau) \restriction \nu$ is a Res(2) refutation of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ (note that it is actually a resolution refutation), and we have

$$\begin{aligned}
\mathrm{size}(\Pi'') + \mathrm{size}(\mathrm{R}^k\mathrm{REF}^F_{s,t}) &\leq \mathrm{size}(\Pi') + \mathrm{size}(\mathrm{R}^k\mathrm{REF}^{n,r'}_{s,t}) \\
&\leq \mathrm{size}(\Pi) + \mathrm{size}(\mathrm{R}^k\mathrm{REF}^{n,r}_{s,t}) \\
&\leq c_1 n^{c_2 k}.
\end{aligned}$$

Because T is non-decreasing, the automating algorithm finds within the allotted time $T(c_1 n^{c_2 k})$ a Res(k) refutation of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$, and M outputs 'satisfiable'.

Assume now that F is unsatisfiable. From our choices of r, s, t and n_1 and from (56) it follows that the hypotheses (54) of Theorem 23 are met for all $n \geq n_1$, and the same is true with r' in place of r. By that theorem, any Res(k) refutation of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ has size greater than $2^{\beta(k)\frac{t}{n^{k-1}}}$. Thanks to (57) this implies that the automating algorithm cannot output any Res(k) refutation of $\mathrm{R}^k\mathrm{REF}^F_{s,t}$ within the allotted time. M therefore outputs 'unsatisfiable'. □

8 Conclusion

We have shown that for every integer $k \geq 2$, the system Res(k) does not have the weak feasible disjunction property and, unless P = NP, it is not automatable. Because of the factor t/n^{k-1} that appears in the exponent of the lower bound in Theorem 23 and originates in the switching lemma (Theorem 20), we have not been able to extend the results to better than barely superconstant k. A more important open question is to rule out weak automatability of these systems assuming some standard hardness assumption.

Acknowledgement. I am grateful to Albert Atserias and Jan Krajíček for their valuable comments on an earlier version of the paper. I would like to thank Ilario Bonacina and Moritz Müller for several related conversations.

References

[1] Michael Alekhnovich and Alexander A. Razborov. Resolution is not automatizable unless W[P] is tractable. *SIAM Journal on Computing*, 38(4):1347–1363, 2008. `doi:10.1137/06066850X`.

[2] Albert Atserias and Maria Luisa Bonet. On the automatizability of resolution and related propositional proof systems. *Information and Computation*, 189(2):182–201, 2004.

[3] Albert Atserias and Moritz Müller. Automating resolution is NP-hard. In *60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 498–509. IEEE, 2019.

[4] Maria Luisa Bonet, Carlos Domingo, Ricard Gavaldà, Alexis Maciel, Toniann Pitassi, and Ran Raz. Non-automatizability of bounded-depth Frege proofs. *Computational Complexity*, 13(1-2):47–68, 2004. `doi:10.1007/s00037-004-0183-5`.

[5] Maria Luisa Bonet, Toniann Pitassi, and Ran Raz. On interpolation and automatization for Frege systems. *SIAM J. Comput.*, 29(6):1939–1967, 2000. `doi:10.1137/S0097539798353230`.

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

[7] Stefan Dantchev and Søren Riis. On relativisation and complexity gap for resolution-based proof systems. In M. Baaz and J. A. Makowsky, editors, *Computer Science Logic (CSL 2003)*, volume 2803 of *Lecture Notes in Computer Science*, pages 142–154. Springer, 2003.

[8] Michal Garlík. Resolution Lower Bounds for Refutation Statements. In Peter Rossmanith, Pinar Heggernes, and Joost-Pieter Katoen, editors, *44th International Symposium on Mathematical Foundations of Computer Science (MFCS 2019)*, volume 138 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 37:1–37:13. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2019. `doi:10.4230/LIPIcs.MFCS.2019.37`.

[9] Jan Krajíček. Lower bounds to the size of constant-depth propositional proofs. *The Journal of Symbolic Logic*, 59(1):73–86, 1994. `doi:10.2307/2275250`.

[10] 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. `doi:10.2307/2275541`.

[11] Jan Krajíček. On the weak pigeonhole principle. *Fundamenta Mathematicae*, 170(1-2):123–140, 2001. `doi:10.4064/fm170-1-8`.

[12] Jan Krajíček. On the proof complexity of the Nisan–Wigderson generator based on a hard NP $\cap$ coNP function. *Journal of Mathematical Logic*, 11(01):11–27, 2011. `doi:10.1142/S0219061311000979`.

[13] Jan Krajíček. *Proof Complexity.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2019. `doi:10.1017/9781108242066`.

[14] Jan Krajíček and Pavel Pudlák. Some consequences of cryptographical conjectures for S_2^1 and EF. *Information and Computation*, 140(1):82–94, 1998. `doi:10.1006/inco.1997.2674`.

[15] Jan Krajíček, Pavel Pudlák, and Alan Woods. An exponential lower bound to the size of bounded depth frege proofs of the pigeonhole principle. *Random Structures & Algorithms*, 7(1):15–39, 1995. `doi:10.1002/rsa.3240070103`.

[16] Toniann Pitassi, Paul Beame, and Russell Impagliazzo. Exponential lower bounds for the pigeonhole principle. *Computational Complexity*, 3(2):97–140, 1993. `doi:10.1007/BF01200117`.

[17] Pavel Pudlák. On reducibility and symmetry of disjoint NP-pairs. *Theoretical Computer Science*, 295:323–339, 2003.

[18] Nathan Segerlind, Samuel R. Buss, and Russel Impagliazzo. A switching lemma for small restrictions and lower bounds for k-DNF resolution. *SIAM Journal on Computing*, 33(5):1171–1200, 2004.