Electronic Colloquium on Computational Complexity, Report No. 117 (2018)

Resolution with Counting: Lower Bounds over Different Moduli

Fedor Part* Iddo Tzameret†

June 23, 2018

Abstract

Resolution over linear equations (introduced in [RT08]) emerged recently as an important object of study. This refutation system, denoted $\mathrm{Res}(\mathrm{lin}_R)$, operates with disjunction of linear equations over a ring R.[1] On the one hand, the system captures a natural "minimal" extension of resolution in which efficient counting can be achieved; while on the other hand, as observed by, e.g., Krajíček [Kra17] (cf. [IS14, KO18, GK17]), when considered over prime fields, and specifically $\mathbb{F}_2$, super-polynomial lower bounds on (dag-like) $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$ is a first step towards the long-standing open problem of establishing constant-depth Frege with counting gates ($\mathsf{AC}^0[2]$-Frege) lower bounds.

In this work we develop new lower bound techniques for resolution over linear equations and extend existing ones to work over different rings. We obtain a host of new lower bounds, separations and upper bounds, while calibrating the relative strength of different sub-systems. We first establish, over fields of characteristic zero, exponential-size *dag-like* lower bounds against resolution over linear equations refutations of instances with large coefficients. Specifically, we demonstrate that the subset sum principle $\alpha_1 x_1 + \ldots + \alpha_n x_n = \beta$, for β not in the image of the linear form, requires refutations proportional to the size of the image. Moreover, for instances with small coefficients, we separate the tree and dag-like versions of $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$, when $\mathbb{F}$ is of characteristic zero, by employing the notion of immunity from Alekhnovich-Razborov [AR01], among other techniques.

We then study resolution over linear equations over different *finite* fields, extending the work of Itsykson and Sokolov [IS14] who developed tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$ lower bounds techniques. We obtain new lower bounds and separations as follows: (**i**) exponential-size lower bounds for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_p})$ Tseitin mod q formulas, for every pair of distinct primes p, q. As a corollary we obtain an exponential-size separation between tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_p})$ and tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_q})$; (**ii**) exponential-size lower bounds for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_p})$ refutations of random k-CNF formulas, for every prime p and constant k; and (**iii**) exponential-size lower bounds for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutations of the pigeonhole principle, for *every* field $\mathbb{F}$.

*Department of Computer Science, Royal Holloway, University of London. Fjodor.Part.2012@live.rhul.ac.uk

†Department of Computer Science, Royal Holloway, University of London. Iddo.Tzameret@rhul.ac.uk

[1] Variables in $\mathrm{Res}(\mathrm{lin}_R)$ range over 0-1 values, i.e., the Boolean axioms $(x_i = 0) \vee (x_i = 1)$, for all variables x_i, are part of the system.

ISSN 1433-8092

Contents

1 Introduction

1.1 Background

The resolution refutation system is among the most prominent and well-studied propositional proof systems, and for good reasons: it is a natural and simple refutation system, that, at least in practice, is capable of being easily automatized. Furthermore, while being non-trivial, it is simple enough to succumb to many lower bound techniques.

Formally, a resolution refutation of an unsatisfiable CNF formula is a sequence of clauses $D_1, \dots, D_l = \emptyset$, where $\emptyset$ is the empty clause, such that each D_i is either a clause of the CNF or is derived from previous clauses $D_j, D_k, j \le k < i$ by means of applying the following *resolution rule*: from the clauses $C \vee x$ and $D \vee \neg x$ derive $C \vee D$.

The *tree-like* version of resolution, where every occurrence of a clause in the refutation is used at most once as a premise of a rule, is of particular importance, since it helps us to understand certain kind of satisfiability algorithms known as DPLL algorithms (cf. [Nor15]). DPLL algorithms are simple recursive algorithms for solving SAT that are the basis of successful contemporary SAT-solvers. The transcript of a run of DPLL on an unsatisfiable formula is a decision tree, which can be interpreted as a tree-like resolution refutation. Thus, lower bounds on the size of tree-like resolution refutations imply lower bounds on the run-time of DPLL algorithms.

In contrast to the apparent practical success of SAT-solvers, a variety of hard instances that require exponential-size refutations have been found for resolution during the years. Many classes of such hard instances are based on principles expressing some sort of counting. One famous example is the *pigeonhole principle*, denoted PHP^m_n, expressing that there is no (total) injective map from a set with cardinality m to a set with cardinality n if $m > n$ [Hak85]. Another important example is *Tseitin tautologies*, denoted TS_G, expressing that the sum of the degrees of vertices in a graph G must be even [Tse68].

Since such counting tautologies are a source of hard instances for resolution, it is useful to study extensions of resolution that can efficiently count, so to speak. This is important firstly, because such systems may become the basis of more efficient SAT-solvers and secondly, in order to extend the frontiers of lower bound techniques against stronger and stronger propositional proof systems. Indeed, there are quite a few works dedicated to the study of weak systems operating with De Morgan formulas with counting connectives; these are variations of resolution that operate with disjunctions of certain arithmetic expressions.

One such extension of resolution was introduced by Raz and Tzameret [RT08] under the name *resolution over linear equations* in which literals are replaced by linear equations. Specifically, the system R(lin), which operates with disjunctions of linear equations over $\mathbb{Z}$ was studied in [RT08]. This work demonstrated the power of resolution with counting over the integers, and specifically provided polynomial upper bounds for the pigeonhole principle and the Tseitin formulas, as well as other basic counting formulas. It also established exponential lower bounds for a subsystem of R(lin), denoted $\mathrm{R}^0(\mathrm{lin})$. Subsequently, Itsykson and Sokolov [IS14] studied resolution over linear equations over $\mathbb{F}_2$, denoted $\mathrm{Res}(\oplus)$. They demonstrated the power of resolution with counting mod 2 as well as its limitations by means of several upper and tree-like lower bounds. Moreover, [IS14] introduces DPLL algorithms, which can "branch" on arbitrary linear forms over $\mathbb{F}_2$, as well as parity decision trees, and showed a correspondence between parity decision trees and tree-like $\mathrm{Res}(\oplus)$ refutations. In both [RT08] and [IS14] the dag-like lower bound question for resolution over linear equations remained open.

As it happens, resolution over linear equations, holds a special place in the theory of proof complexity: it can be viewed as a natural "minimal" subsystem of important propositional proof systems, as we now explain. Resolution operates with clauses, which are De Morgan formulas ($\neg$, unbounded fan-in $\vee$ and $\wedge$) of a particular kind, namely, of depth 1. Thus, from the perspective of the theory of proof complexity, resolution is a fairly weak version of the propositional-calculus, where the latter operates with arbitrary De Morgan formulas. Under a natural and general definition, propositional-calculus systems go under the name *Frege systems*: they can be (axiomatic) Hilbert-style systems or sequent-calculus style systems. The task of proving lower bounds for general Frege systems is notoriously hard: no nontrivial lower bounds are known to date. Basically, the strongest fragment of Frege systems, for which lower bounds are known are AC^0-Frege systems, which are Frege proofs operating with constant-depth formulas. For example, both PHP^m_n and TS_G do not admit sub-exponential proofs in AC^0-Frege [Ajt88, PBI93, KPW95, BS02]. However, if we extend the De Morgan language with counting connectives such as unbounded fan-in mod p ($\mathsf{AC}^0[p]$-Frege) or threshold gates (TC^0-Frege), then we step again into the darkness: proving super-polynomial lower bounds for these systems is a long-standing open problem on what can be characterized as the "frontiers" of proof complexity. In this sense, resolution over linear equations over prime fields and over the integers is interesting as a first step towards $\mathsf{AC}^0[p]$-Frege lower and TC^0-Frege lower bounds, respectively. Works by Krajíček [Kra17], Garlik-Kołodziejczyk [GK17] and Krajíček-Oliveira [KO18] had suggested possible approaches to attack dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$ lower bounds.

1.2 Our Results

In this paper we continue the study of the power of resolution over linear equations, while extending it to different rings R, denoted $\mathrm{Res}(\mathrm{lin}_R)$, both finite and infinite. We prove a host of new lower bounds, separations and upper bounds for resolution over linear equations, including dag-like refutations. We focus mainly on finite fields $\mathbb{F}_q$, for different primes q, and fields of characteristic 0, most importantly the rational numbers $\mathbb{Q}$. Using our notation, R(lin) from [RT08] is simply $\mathrm{Res}(\mathrm{lin}_{\mathbb{Z}})$ and $\mathrm{Res}(\oplus)$ from [IS14] is $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$.

The refutation system $\mathrm{Res}(\mathrm{lin}_R)$ is defined as follows (see [RT08]). The proof-lines of $\mathrm{Res}(\mathrm{lin}_R)$ are ***linear clauses***, that is, disjunctions of linear equations. More formally, they are disjunctions of the form:

$$\left(\sum\nolimits_{i=0}^{n} a_{1i}x_i + b_1 = 0\right) \vee \cdots \vee \left(\sum\nolimits_{i=0}^{n} a_{ki}x_i + b_k = 0\right),$$

where k is some number (the *width* of the clause), and $a_{ji}, b_j \in R$. The *resolution rule* is the following:

$$\text{from } (C \vee f = 0) \text{ and } (D \vee g = 0) \text{ derive } (C \vee D \vee (\alpha f + \beta g) = 0),$$

where $\alpha, \beta \in R$, and C, D some linear clauses. A $\mathrm{Res}(\mathrm{lin}_R)$ *refutation* of an unsatisfiable over 0-1 set of linear clauses $C_1, \dots, C_m$ is a sequence of proof-lines, where each proof-line is either C_i, for $i \in [m]$, a boolean axiom $(x_i = 0 \vee x_i = 1)$ for a some variable x_i, or was derived from previous proof-lines by the above resolution rule, or by the *weakening rule* that allows to extend clauses with arbitrary disjuncts, or a *simplification rule* allowing to discard false constant linear forms (e.g., $1 = 0$) from a linear clause. The last proof-line in a refutation is the empty clause (standing for the truth value false).

We are interested in the following questions:

(Q1) For a given ring R, what kind of counting can be efficiently performed in $\text{Res}(\text{lin}_R)$ and tree-like $\text{Res}(\text{lin}_R)$?

(Q2) Can dag-like $\text{Res}(\text{lin}_R)$ be separated from tree-like $\text{Res}(\text{lin}_R)$?

(Q3) Can tree-like systems for different rings R be separated?

In order to be able to do some non-trivial counting in tree-like versions of resolution over linear equations we define a semantic version of the system as follows:

Tree-like $\text{Res}(\text{lin}_R)$ **with semantic weakening.** The system $\text{Res}_{sw}(\text{lin}_R)$ is obtained from $\text{Res}(\text{lin}_R)$ by replacing the weakening and the simplification rules, as well as the boolean axioms, with the *semantic weakening* rule (the symbol $\models$ will denote in this work semantic implication ***with respect to 0-1 assignments***):

$$\frac{C}{D}\,(C \models D)\,.$$

Let $k = char(R)$ be the characteristic of the ring R. In case $k \notin \{1, 2, 3\}$, deciding whether an R-linear clause D is a tautology (that is, holds for every 0-1 assignment to its variables) is at least as hard as deciding whether a 3-DNF is a tautology (because over characteristic $k \notin \{1, 2, 3\}$ linear equations can express conjunction of three conjuncts). For this reason $\text{Res}_{sw}(\text{lin}_R)$ proofs cannot be checked in polynomial time and thus $\text{Res}_{sw}(\text{lin}_R)$ is not a Cook-Reckhow proof system unless $\mathsf{P} = \mathsf{coNP}$ (namely, the correctness of proofs in the system cannot necessarily be checked in polynomial-time, as required by a Cook-Reckhow propositional proof system [CR79]).

The reason for studying $\text{Res}_{sw}(\text{lin}_R)$ is mainly the following: Let Γ be an arbitrary set of tautological R-linear clauses. Then, lower bounds for tree-like $\text{Res}_{sw}(\text{lin}_R)$ imply lower bounds for tree-like $\text{Res}(\text{lin}_R)$ with formulas in Γ as axioms. For example, in case $\mathbb{F}$ is a field of characteristic 0, the possibility to do counting in tree-like $\text{Res}(\text{lin}_{\mathbb{F}})$ is quite limited. For instance, we show that $2x_1 + \ldots + 2x_n = 1$ requires an exponential-size in n refutation (Corollary 25). On the other hand, such contradictions *do* admit short tree-like $\text{Res}(\text{lin}_{\mathbb{F}})$ refutations in the presence of the following *generalized boolean axioms* (which is a tautological linear clause):

$$\mathsf{Im}(f) := \bigvee\nolimits_{A \in im_2(f)} (f = A), \tag{1}$$

where $im_2(f)$ is the image of f under 0-1 assignments. Similar to the way the Boolean axioms $(x_i = 0) \vee (x_i = 1)$ state that the possible value of a variable is either zero or one, the $\mathsf{Im}(f)$ axiom states all the possible values that the linear form f can take. If a lower bound holds for tree-like $\text{Res}_{sw}(\text{lin}_{\mathbb{F}})$ it also holds, in particular, for tree-like $\text{Res}(\text{lin}_{\mathbb{F}})$ with the axioms $\mathsf{Im}(f)$, and this makes tree-like $\text{Res}_{sw}(\text{lin}_{\mathbb{F}})$ a useful system, for which lower bounds against are sufficiently interesting.

Lower bounds and separations in characteristic zero. First, we show that over any field $\mathbb{F}$, whenever $\alpha_1x_1+\ldots+\alpha_nx_n+\beta=0$ is unsatisfiable (over 0-1 assignments), it requires dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutations proportional to the image of the linear form (under 0-1 assignments). Note that $\alpha_1x_1+\ldots+\alpha_nx_n+\beta=0$ expresses the *subset sum principle*: $\alpha_1x_1+\ldots+\alpha_nx_n=-\beta$ iff there is a subset of the integral coefficients α_i whose sum is precisely $-\beta$. Our result implies an exponential-lower bound for dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutations, for large enough coefficients, as follows:

Theorem (Theorem 22; Dag-like lower bound)**.** *If $\mathbb{F}$ is a field of characteristic zero, then* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ *refutations of $x_1+2x_2+\ldots+2^nx_n+1=0$ are of size $2^{\Omega(n)}$.*

This lower bound is proved by showing (see Lemma 21) that every (dag- or tree-like) refutation of a subset sum principle of the form $f+\beta=0$ can be transformed without much increase in size into a normal form refutation (in dag- or tree-like, resp.): a derivation of $\mathsf{Im}(f)$, combined with a successive use of resolution with $f+\beta=0$ to derive the empty clause. This then provides the desired lower bound whenever $\mathsf{Im}(f)$ is sufficiently large.

The idea behind the normal form transformation is as follows: given a refutation in which the only non-Boolean axiom is $f+\beta=0$, we *defer* all resolution steps using this axiom. Namely, we mimic the same refutation had we not used resolution with $f+\beta=0$. We show that in this case, each clause in the resulting refutation is essentially a weakening of the original clause, possibly weakened by (i.e., is a disjunction with) disjunct of the form $f+b=0$, for some constant b. This concludes the argument, since the last clause must be such a tautological weakening of the *empty clause*, but such a tautology ought to be a weakening of the subset sum principle itself (note that every proof-line in the transformation is a tautology (over 0-1 assignments), since the only axioms used throughout the derivation are the Boolean axioms).

Moreover, we prove an exponential-size $2^{\Omega(n)}$ lower bound on tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ refutations of the pigeonhole principle PHP^m_n *for every field* $\mathbb{F}$ (including finite fields). This extends a previous result by Itsykson and Sokolov [IS14] for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$. Together with the polynomial upper bound for PHP^m_n refutations in dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ for fields $\mathbb{F}$ of characteristic zero demonstrated in [RT08], our results establish a separation between dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ and tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ for characteristic zero fields.

Theorem (Theorem 26; Pigeonhole principle lower bounds)**.** *Let $\mathbb{F}$ be any field. Then every tree-like* $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ *refutation of $\neg\mathrm{PHP}^m_n$ has size $2^{\Omega\left(\frac{n-1}{2}\right)}$.*

Theorem (Theorem 15; Raz-Tzameret [RT08]; Short dag-like pigeonhole principle refutations)**.** *For every ring R of characteristic zero there exists a* $\mathrm{Res}(\mathrm{lin}_R)$ *refutation of $\neg\mathrm{PHP}^m_n$ of polynomial size.*

To prove this theorem, as well as some other lower bounds, we extend the Prover-Delayer game technique as originated in Pudlak-Impagliazzo [PI00] for resolution, and developed further by Itsykson-Sokolov [IS14] for $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$, to general rings, including characteristic zero rings. Specifically, to prove Theorem 26 we need to prove that Delayer's strategy from [IS14] is successful over any field. This argument is new, and uses a result of Alon-Füredi [AF93] about the hyperplane coverings of the hypercube.

We prove another separation between dag-like $\text{Res}(\text{lin}_{\mathbb{F}})$ and tree-like $\text{Res}_{sw}(\text{lin}_{\mathbb{F}})$, as follows. We define the *Image Avoidance principle* to be:

$$\mathsf{ImAv}(x_1 + \cdots + x_n) := \{\langle x_1 + \cdots + x_n \neq k\rangle\}_{k\in\{0,\ldots,n\}},$$

where $\langle x_1 + \cdots + x_n \neq k\rangle := \bigvee_{k'\in\{0,\ldots,n\},\ k\neq k'} x_1 + \cdots + x_n = k'$. In words, the Image Avoidance principle expresses the contradictory statement that for every $0 \leq i \leq n$, $x_1 + \ldots + x_n$ equals some element in $\{0, \ldots, n\} \setminus i$.

Theorem (Theorem 11)**.** *For every ring R and every linear form f, there are polynomial-size* $\text{Res}(\text{lin}_R)$ *refutations of* $\mathsf{ImAv}(f)$*.*

Theorem (Theorem 24)**.** *Let $f = \epsilon_1 x_1 + \cdots + \epsilon_n x_n$, where $\epsilon_i \in \{-1, 1\} \subset \mathbb{F}$, and let $\mathbb{F}$ be a field of characteristic zero. Then, the following hold:*

1. *Any tree-like* $\text{Res}_{sw}(\text{lin}_{\mathbb{F}})$ *refutation of* $\mathsf{ImAv}(f)$ *is of size at least* $2^{\frac{n}{4}}$*.*

2. *Any tree-like* $\text{Res}(\text{lin}_{\mathbb{F}})$ *derivation of the clause* $\mathsf{Im}(f)$ *is of size at least* $2^{\frac{n}{4}}$*.*

Together with the above mentioned normal form lemma (Lemma 21) that we establish for (both dag- and tree-like) refutations of $\mathsf{Im}(f)$, we get the following:

Corollary (Corollary 25)**.** *Let f and $\mathbb{F}$ be as in the previous theorem. Then the shortest tree-like* $\text{Res}(\text{lin}_{\mathbb{F}})$ *refutation of $f = n + 1$ is of size at least* $2^{\frac{n}{4}}$*.*

The lower bounds in Theorem 24 and Corollary 25 are novel applications of the Prover-Delayer game argument, together with the notion of immunity from Alekhnovich and Razborov [AR01], as we now explain briefly. Let f be a linear form as in Theorem 24. We consider two instances of the Prover-Delayer game: for $\mathsf{ImAv}(f)$ and for $\mathsf{Im}(f)$. A position in the games is determined by a *set Φ of linear non-equalities* of the form $g \neq 0$, which we think of as the set of non-equalities learned up to this point by Prover. For each of the two games we define Delayer's strategy in such a way that for Φ an end-game position, there is a satisfiable subset $\Phi' = \{g_1 \neq 0, \ldots, g_m \neq 0\} \subseteq \Phi$ such that $\Phi' \models f = A$ for some $A \in \mathbb{F}$, and Delayer earns at least $|\Phi'| = m$ coins. Because $\mathbb{F}$ is of characteristic zero, it follows that $f \equiv A + 1 \pmod 2 \models f \neq A \models g_1 \cdot \ldots \cdot g_m = 0$ and thus the $\frac{n}{4}$-immunity of $f \equiv A + 1 (\text{mod } 2)$ ([AR01]) implies $m \geq \frac{n}{4}$. To conclude, we use the standard argument that shows that if Delayer always earns $\frac{n}{4}$ coins, then the shortest proof is of size at least $2^{\frac{n}{4}}$.

Table 1 sums up our knowledge up to this point with respect to characteristic 0 fields.

	$\sum_{i=1}^{n} 2x_i = 1$	$\sum_{i=1}^{n} 2^i x_i = -1$	$\mathsf{ImAv}(\sum_{i=1}^{n} x_i)$	PHP_n^m	$\mathsf{Im}\left(\sum_{i=1}^{n} x_i\right)$
t-l $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$
t-l $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$	poly	poly	$2^{\Omega(n)}$	$2^{\Omega(n)}$	poly
$\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$	poly	$2^{\Omega(n)}$	poly	poly [RT08]	poly

Table 1: Lower bounds for fields of characteristic 0. The notation t-l $\mathrm{Res}(\mathrm{lin}_R)$ stands for tree-like $\mathrm{Res}(\mathrm{lin}_R)$. The rightmost column describes bounds on *derivations*, in contract to refutations.

Lower bounds and separations in finite fields. Apart from the lower bounds for the pigeonhole principle that hold both for infinite and finite fields as discussed above, we prove a separation between tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_{p^k}})$ (resp. tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}_{p^k}})$) and tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_{q^l}})$ (resp. tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}_{q^l}})$) for every pair of distinct primes $p \neq q$ and every $k, l \in \mathbb{N} \setminus \{0\}$. The separating instances are mod p Tseitin formulas $\mathrm{TS}_{G,\sigma}^{(p)}$, which are reformulations of the standard Tseitin graph formulas TS_G for counting mod p. Furthermore, we establish an exponential lower bound for tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}_{p^k}})$ on random k-CNFs.[2]

We now explain the general lower bounds argument, followed by the precise results. The lower bounds for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ for finite fields $\mathbb{F}$ are obtained via a variant of the size-width relation for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ together with a translation to polynomial calculus over the field $\mathbb{F}$, denoted $PC_{\mathbb{F}}$ [CEI96], such that $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ proofs of width ω are translated to $PC_{\mathbb{F}}$ proofs of degree ω (the *width* ω of a clause is defined to be the total number of disjuncts in a clause). This establishes the lower bounds for the size of tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ proofs via lower bounds on $PC_{\mathbb{F}}$ degrees.

We show that

$$\omega_0(\phi \vdash \bot) = O\left(\omega_0(\phi) + \log S_{\text{t-l Res}(\mathrm{lin}_R)}(\phi \vdash \bot)\right),$$

where ω_0 is what we call the *principal width*, which counts the number of linear equations in clauses after the identification of those defining parallel hyperplanes, and $S_{\text{t-l Res}(\mathrm{lin}_R)}(\phi \vdash \bot)$ denotes the minimal size of a tree-like $\mathrm{Res}(\mathrm{lin}_R)$ refutation of ϕ.

Specifically, over finite fields the following upper and lower bounds provide exponential separations:

Theorem (Theorem 32; Size-width relation)**.** *Assume ϕ is an unsatisfiable CNF formula. The following relation between principal width and size holds for tree-like* $\mathrm{Res}_{sw}(\mathrm{lin}_R)$*: $S(\phi \vdash \bot) = 2^{\Omega(\omega_0(\phi \vdash \bot) - \omega_0(\phi))}$. If R is a finite ring, then the same relation holds for the (standard) width of a clause ω.*

[2] We thank Dmitry Itsykson for telling us about the lower bound for random k-CNF for the case of tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$, that was proved by Garlik and Kołodziejczyk using size-width relations (unpublished note). Our result extends Garlik and Kołodziejczyk's result to all finite fields. Similar to their result, we use a size-width argument and simulation by the polynomial calculus to establish the lower bound.

This extends to every field a result by Garlik-Kołodziejczyk [GK17, Theorem 14] who showed a size-width relation for a system denoted tree-like $\mathrm{PK}^{\mathrm{id}}_{O(1)}(\oplus)$, which is a system extending tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$ by allowing arbitrary constant-depth De Morgan formulas as inputs to $\oplus$ (XOR gates) (though note that our result does not deal with *arbitrary* constant-depth formulas).

Theorem (Theorem 33)**.** *Let $\mathbb{F}$ be a field and π be a $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutation of an unsatisfiable CNF formula ϕ. Then, there exists a $PC_{\mathbb{F}}$ refutation π' of (the arithmetization of) ϕ of degree $\omega(\pi)$.*

Corollary (Corollary 34; Tseitin mod p lower bounds)**.** *For any fixed prime p there exists a constant $d_0 = d_0(p)$ such that the following holds. If $d \geq d_0$, G is a d-regular directed graph satisfying certain expansion properties, and $\mathbb{F}$ is a finite field such that $char(\mathbb{F}) \neq p$, then every tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ refutation of the Tseitin mod p formula $\neg\mathrm{TS}^{(p)}_{G,\sigma}$ has size $2^{\Omega(dn)}$.*

Corollary (Corollary 35; Random k-CNF formulas lower bounds)**.** *Let ϕ be a randomly generated k-CNF with clause-variable ratio Δ, and where $\Delta = \Delta(n)$ is such that $\Delta = o\left(n^{\frac{k-2}{2}}\right)$, and let $\mathbb{F}$ be a finite field. Then, every tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ refutation of ϕ has size $2^{\Omega\left(\frac{n}{\Delta^{2/(k-2)}\cdot\log\Delta}\right)}$ with probability $1 - o(1)$.*

The tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ upper bounds for mod p Tseitin formulas in the case $char(\mathbb{F}) = p$ stem from the following proposition:

Proposition (Proposition 12; Upper bounds on unsatisfiable linear systems)**.** *Let R be a ring and assume that the linear system $A\overline{x} = \overline{b}$, where A is a $k \times n$ matrix over R, has no solutions (over R). Let ϕ be a CNF formula encoding the linear system $A\overline{x} = \overline{b}$. Then, there exists tree-like $\mathrm{Res}(\mathrm{lin}_R)$ refutations of ϕ of size polynomial in the sum of sizes of encodings of all coefficients in A.*

Table 2 shows the results for $\mathrm{Res}(\mathrm{lin}_R)$ over finite fields.

Nondeterministic linear decision trees. There is well-known size preserving (up to a constant factor) correspondence between tree-like resolution refutations for unsatisfiable formulas ϕ and decision trees, which solve the following problem: given an assignment ρ for the variables of ϕ, determine which clause $C \in \phi$ is falsified by querying values of the variables under the assignment ρ. In Itsykson-Sokolov [IS14] this correspondence was generalized to tree-like $\mathrm{Res}(\oplus)$ refutations and parity decision trees. In this paper we extend the correspondence to a correspondence between tree-like $\mathrm{Res}(\mathrm{lin}_R)$ (and tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$) derivations to *nondeterministic linear decision trees* (NLDT).

NLDTs for an unsatisfiable set of linear clauses ϕ are binary rooted trees, where every edge is labeled with a non-equality $f \neq 0$ for a linear form f and every leaf is labeled with a linear clause $C \in \phi$, which is violated by the non-equalities on the path from the root to the leaf.

	$A\overline{x} = \overline{b}$	$\mathrm{TS}^{(-)}_{G,\sigma}$	$\mathrm{TS}^{(q)}_{G,\sigma}$	random k-CNF	PHP^m_n
t-l $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_{p^k}})$	?	poly	$2^{\Omega(dn)}$	$2^{\Omega\left(\frac{n}{\Delta^{2/(k-2)}\cdot\log\Delta}\right)}$	$2^{\Omega(n)}$
t-l $\mathrm{Res}(\oplus)$	poly [IS14]	poly [IS14]	$2^{\Omega(dn)}$	$2^{\Omega\left(\frac{n}{\Delta^{2/(k-2)}\cdot\log\Delta}\right)}$	$2^{\Omega(n)}$ [IS14]
t-l $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}_{p^k}})$	poly	poly	?	?	$2^{\Omega(n)}$

Table 2: Lower bounds over finite fields. Here G is d-regular graph and Δ is the clause density (number of clauses divided by the number of variables), $A\overline{x} = \overline{b}$ stands for a linear system over $\mathbb{F}_{p^k}$ that has no 0-1 solutions in the first and the third rows, and in the second row the linear system $A\overline{x} = \overline{b}$ is over $\mathbb{F}_2$. The notation $\mathrm{TS}^{(-)}_{G,\sigma}$ stands for $\mathrm{TS}^{(p)}_{G,\sigma}$ in the first and the third rows and for $\mathrm{TS}^{(2)}_{G,\sigma}$ in the second raw. t-l $\mathrm{Res}(\mathrm{lin}_R)$ stands for tree-like $\mathrm{Res}(\mathrm{lin}_R)$, and $p \neq q$ are primes (in the second raw we assume $q \neq 2$). Circled "?" denotes an open problem. The results marked with [IS14] were proved in the corresponding paper. All other results are from the current work.

Theorem (Theorem 17)**.** *If ϕ is an unsatisfiable CNF formula, then every tree-like* $\mathrm{Res}(\mathrm{lin}_R)$ *or tree-like* $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ *refutation can be transformed to an NLDT for ϕ of the same size up to a constant factor, and vise versa.*

2 Preliminaries

2.1 Notation

Denote by $[n]$ the set $\{1, \dots, n\}$. We use $x_1, x_2, \dots$ to denote variables, both propositional and algebraic. Let f be a linear form (equivalently, an affine function) over a ring R, that is, a function of the form $\sum_{i=1}^n a_i x_i + a_0$ with $a_i \in R$. We sometimes refer to a linear form as a *hyperplane*, since a linear form determines a hyperplane. We denote by $im_2(f)$ the image of f under 0-1 assignments to its variables; $\langle f \neq A \rangle := \bigvee_{A \neq B \in im_2(f)} (f = B)$, where $A \in R$.

For ϕ a set of clauses or linear clauses, $vars(\phi)$ denotes the set of variables occurring in ϕ and let Vars denote the set of *all* variables.

Let A be a matrix over a ring. We introduce the notation $Ax \doteqdot b$ for a system of linear non-equalities, where a ***non-equality*** means $\neq$ (note the difference between $Ax \doteqdot b$, which stands for $A_i \cdot x \neq b_i$, for *all* rows A_i in A, and $Ax \neq b$, which stands for $A_i \cdot x \neq b_i$, for *some* row A_i in A).

If f is a linear form over R and A is a matrix over R, denote by $|f|$ the sum of sizes of encodings of coefficients in f and by $|A|$ the sum of sizes of encodings of elements in A.

If $C = (\bigvee_{i\in[m]} f_i = 0)$ is a linear clause (i.e., a disjunction of linear equations), denote by $\neg C$ the *set* of non-equalities $\{f_i \neq 0\}_{i\in[m]}$. Conversely, if $\Phi = \{f_i \neq 0\}_{i\in[n]}$ is a set of non-equalities, denote $\neg\Phi := \bigvee_{i\in[m]} f_i = 0$.

If ϕ is a set of linear clauses over a ring R and D is a linear clause over R, denote by $\bigwedge_{C\in\phi} C \models D$ and $\bigwedge_{C\in\phi} C \models_R D$ semantic entailment over 0-1 and R-valued assignments respectively.

Let l be a linear form not containing the variable x. If C is a linear clause, denote by $C\restriction_{x\leftarrow l}$ the linear clause, which is obtained from C by substituting l for x everywhere in C. If $\phi = \{C_i\}_{i\in I}$ is a set of clauses, denote $\phi\restriction_{x\leftarrow l} := \{C_i\restriction_{x\leftarrow l}\}_{i\in I}$. We define a *linear substitution* ρ to be a sequence $(x_1 \leftarrow l_1, \ldots, x_n \leftarrow l_n)$ such that linear forms l_i does not depend on x_i. For a clause or a set of clauses ϕ we define $\phi\restriction_\rho := (\ldots((\phi\restriction_{x_1\leftarrow l_1})\restriction_{x_2\leftarrow l_2})\ldots)\restriction_{x_n\leftarrow l_n}$.

2.2 Propositional Proof Systems

A *clause* is an expression of the form $l_1 \vee \cdots \vee l_k$, where l_i is a literal, where a *literal* is a propositional variable x or its negation $\neg x$. A formula is in *Conjunctive Normal Form* (CNF) if it is a conjunction of clauses. A CNF can thus be defined simply as a set of clauses. The choice of a reasonable binary encoding of sets of clauses allows us to define the language UNSAT $\subset \{0,1\}^*$ of unsatisfiable propositional formulas in CNF. We sometimes interpret an element in UNSAT as a formula and sometimes as a set of clauses. Dually, a formula is in *Disjunctive Normal Form* (DNF) if it is a disjunction of conjunctions of literals and TAUT is the language of tautological propositional formulas in DNF. There is a bijection between TAUT and UNSAT, which preserves the size of the formula, given by negation.

A formula is in k-CNF (resp. k-DNF) if it is in CNF (resp. DNF) and every clause (resp. conjunct) has at most k literals. k-UNSAT (resp. k-TAUT) is the language of unsatisfiable (resp. tautological) formulas in k-CNF (resp. k-DNF).

Definition 1 (Cook-Reckhow propositional proof system [CR79])**.** *A* propositional proof system Π *is a polynomial time computable onto function* $\Pi : \{0,1\}^* \to TAUT$.

Π-proofs of $\phi \in$ TAUT are elements in $\Pi^{-1}(\phi)$. Definition 1 can be generalized to arbitrary languages: proof system for a language L is polynomial time computable onto function $\Pi : \{0,1\}^* \to L$. In particular, a *refutation system* Π is a proof system for UNSAT. Post-composition with negation turns a propositional proof system into a refutation system and vise versa.

Denote by $S(\pi)$, and alternatively by $|\pi|$, the size of the binary encoding of a proof π in a proof system Π. For $\phi \in$ UNSAT and a refutation system Π denote by $S_\Pi(\phi \vdash \bot)$ (we sometimes omit the subscript Π when it is clear from the context) the minimal size of a Π-refutation of ϕ.

The *resolution* system (which we denote also by Res) is a refutation system, based on the following rule, allowing to derive new clauses from given ones:

$$\frac{C \vee x \qquad D \vee \neg x}{C \vee D} \quad \text{(Resolution rule).}$$

A *resolution derivation* of a clause D from a set of clauses ϕ is a sequence of clauses $(D_1, \ldots, D_s \equiv D)$ such that for every $1 \le i \le s$ either $D_i \in \phi$ or D_i is obtained from previous clauses by applying the resolution rule. A *resolution refutation* of $\phi \in$ UNSAT is a resolution derivation of the empty clause from ϕ, which stands for the truth value False.

A resolution derivation is *tree-like* if every clause in it is used at most once as a premise of a rule. Accordingly, *tree-like resolution* is the resolution system allowing only tree-like refutations.

Let $\mathbb{F}$ be a field. A *polynomial calculus* [CEI96] derivation of a polynomial $q \in \mathbb{F}[x_1, \ldots, x_n]$ from a set of polynomials $\mathcal{P} \subseteq \mathbb{F}[x_1, \ldots, x_n]$ is a sequence $(p_1, \ldots, p_s), p_i \in \mathbb{F}[x_1, \ldots, x_n]$ such that for every $1 \le i \le s$ either $p_i = x_j^2 - x_j$, $p_i \in \mathcal{P}$ or p_i is obtained from previous polynomials by applying one of the following rules:

$$\frac{f \qquad g}{\alpha f + \beta g} \quad (\alpha, \beta \in \mathbb{F}, f, g \in \mathbb{F}[x_1, \ldots, x_n]) \qquad \frac{f}{x \cdot f} \quad (f \in \mathbb{F}[x_1, \ldots, x_n]) \,.$$

A polynomial calculus refutation of $\mathcal{P} \subseteq \mathbb{F}[x_1, \ldots, x_n]$ is a derivation of 1. The degree $d(\pi)$ of a polynomial calculus derivation π is the maximal total degree of a polynomial appearing in it. This defines the proof system $PC_{\mathbb{F}}$ for the language of unsatisfiable systems of polynomial equations over $\mathbb{F}$. It can be turned into a proof system for k-UNSAT via *arithmetization of clauses* as follows: $(x_1 \vee \ldots \vee x_k \vee \neg y_1 \vee \ldots \vee \neg y_l)$ is represented as $(1 - x_1) \cdot \ldots \cdot (1 - x_k) \cdot y_1 \cdot \ldots \cdot y_l = 0$.

2.3 Hard Instances

2.3.1 Pigeonhole Principle

The *pigeonhole principle* states that there is no injective mapping from the set $[m]$ to the set $[n]$ for $m > n$. Elements of the former and the latter sets are referred to as *pigeons* and *holes*, respectively. The CNF formula, denoted PHP_n^m, encoding the negation of this principle is defined as follows. Let the set of propositional variables $\{x_{i,j}\}_{i \in [m], j \in [n]}$ correspond to the mapping from $[m]$ to $[n]$, that is, $x_{i,j} = 1$ iff the i^{th} pigeon is mapped to the j^{th} hole. Then $\neg\mathrm{PHP}_n^m := \mathsf{Pigeons}_n^m \cup \mathsf{Holes}_n^m \in$ UNSAT, where $\mathsf{Pigeons}_n^m = \{\bigvee_{j \in [n]} x_{i,j}\}_{i \in [m]}$ are axioms for pigeons and $\mathsf{Holes}_n^m = \{\neg x_{i,j} \vee \neg x_{i',j}\}_{i \neq i' \in [m], j \in [n]}$ are axioms for holes.

Weaker (namely, easier to refute) versions of $\neg\mathrm{PHP}_n^m$ are obtained by augmenting it with the *functionality* axioms $\mathrm{Func}_n^m := \{\neg x_{i,j} \vee \neg x_{i,j'}\}_{i \in [m], j \neq j' \in [n]}$ ($\neg\mathrm{FPHP}_n^m$) or the *surjectivity* axioms $\mathrm{Surj}_n^m := \{\bigvee_{i \in [m]} x_{i,j}\}_{j \in [n]}$ ($\neg$onto-PHP_n^m).

2.3.2 Mod p Tseitin Formulas

We use the version given in [AR01] (which is different from the one in [BGIP01, RT08]). Let $G = (V, E)$ be a directed d-regular graph. We assign to every edge $(u, v) \in E$ a corresponding variable $x_{(u,v)}$. Let $\sigma : V \to \mathbb{F}_p$. The *Tseitin mod p formulas* $\neg\mathrm{TS}_{G,\sigma}^{(p)}$ are the CNF encoding of the following equations for all $u \in V$:

$$\sum_{(u,v)\in E} x_{(u,v)} - \sum_{(v,u)\in E} x_{(v,u)} \equiv \sigma(u) \mod p\,. \tag{2}$$

Note that we use the standard encoding of boolean functions as CNF formulas and the number of clauses, required to encode these equations is $O(2^d|V|)$. $\neg\mathrm{TS}^{(p)}_{G,\sigma}$ is unsatisfiable if and only if $\sum_{u\in V}\sigma(u) \not\equiv 0 \mod p$. To see this, note that if we sum (2) over all nodes $u \in V$ we obtain precisely $\sum_{u\in V}\sigma(u)$ which is different from $0 \mod p$; but on the other hand, in this sum over all nodes $u \in V$ each edge $(u,v) \in E$ appears once with a positive sign as an outgoing edge from u and with a negative sign as an incoming edge to v, meaning the the total sum is 0, which is a contradiction.

In particular, $\neg\mathrm{TS}^{(2)}_{G,\sigma}$ are the classical Tseitin formulas [Tse68] and $\mathrm{TS}^{(2)}_{G,1}$, where 1 is the constant function $v \mapsto 1$ (for all $v \in V$), expresses the fact that the sum of total degrees (incoming + outgoing) of the vertices is even.

The proof complexity of Tseitin tautologies depends on the properties of the graph G. For example, if G is just a union of K_{d+1} (the complete graphs on $d+1$ vertices), then they are easy to prove. On the other hand, they are known to be hard for some proof systems if G satisfies certain expansion properties.

Let $G = (V,E)$ be an *undirected* graph. For $U, U' \subseteq V$ define $e(U,U') := \{(u,u') \in E \mid u \in U, u' \in U'\}$. Consider the following measure of expansion for $r \geq 1$:

$$c_E(r,G) := \min_{|U|\leq r} \frac{e(U, V\backslash U)}{|U|}$$

G is (r,d,c)-expander if G is d-regular and $c_E(r,G) \geq c$. There are explicit constructions of good expanders. For example:

Proposition 1 (Lubotzky *et. al* [LPS88]). *For any d, there exists an explicit construction of d-regular graph G, called Ramanujan graph, which is $(r, d, d(1-\frac{r}{n})-2\sqrt{d-1})$-expander for any $r \geq 1$.*

Proposition 2 (Alekhnovich-Razborov [AR01]). *For any fixed prime p there exists a constant $d_0 = d_0(p)$ such that the following holds. If $d \geq d_0$, G is a d-regular Ramanujan graph on n vertices (augmented with arbitrary orientation of its edges) and char$(\mathbb{F}) \neq p$, then for every function σ such that $\neg TS^{(p)}_{G,\sigma} \in UNSAT$ every $PC_{\mathbb{F}}$ refutation of $\neg TS^{(p)}_{G,\sigma}$ has degree $\Omega(dn)$.*

2.3.3 Random k-CNFs

A random k-CNF is a formula $\phi \sim \mathcal{F}^{n,\Delta}_k$ with n variables that is generated by picking randomly and independently $\Delta \cdot n$ clauses from the set of all $\binom{n}{k} \cdot 2^k$ clauses.

Proposition 3 (Alekhnovich-Razborov [AR01]). *Let $\phi \sim \mathcal{F}^{n,\Delta}_k$, $k \geq 3$ and $\Delta = \Delta(n)$ is such that $\Delta = o\left(n^{\frac{k-2}{2}}\right)$. Then every $PC_{\mathbb{F}}$ refutation of ϕ has degree $\Omega\left(\frac{n}{\Delta^{2/(k-2)}\cdot\log\Delta}\right)$ with probability $1-o(1)$ for any field $\mathbb{F}$.*

3 Resolution with Linear Equations over General Rings

In this section we define and outline some basic properties of systems that are extensions of resolution, where clauses are disjunctions of linear equations over a ring R: $(\sum_{i=0}^{n} a_{1i}x_i + b_1 = 0) \vee \cdots \vee (\sum_{i=0}^{n} a_{ki}x_i + b_k = 0)$. Disjunctions of this form are called *linear clauses*.

The rules of $\mathrm{Res}(\mathrm{lin}_R)$ are as follows (cf. [RT08]):

$$\text{(Resolution)} \quad \frac{C \vee f(\overline{x}) = 0 \qquad D \vee g(\overline{x}) = 0}{C \vee D \vee (\alpha f(\overline{x}) + \beta g(\overline{x})) = 0} \quad (\alpha, \beta \in R)$$

$$\text{(Simplification)} \, \frac{C \vee a = 0}{C} \, (0 \neq a \in R) \qquad \text{(Weakening)} \, \frac{C}{C \vee f(\overline{x}) = 0}$$

where $f(\overline{x}), g(\overline{x})$ are linear forms over R and C, D are linear clauses. The *Boolean axioms* are defined as follows:

$$x_i = 0 \vee x_i = 1, \text{ for } x_i \text{ a variable}$$

A $\mathrm{Res}(\mathrm{lin}_R)$ *derivation* of a linear clause D from a set of linear clauses ϕ is a sequence of linear clauses $(D_1, \ldots, D_s \equiv D)$ such that for every $1 \leq i \leq s$ either $D_i \in \phi$ or is a Boolean axiom or D_i is obtained from previous clauses by applying one of the rules above. A $\mathrm{Res}(\mathrm{lin}_R)$ *refutation* of an unsatisfiable set of linear clauses ϕ is a $\mathrm{Res}(\mathrm{lin}_R)$ derivation of the empty clause (which stands for false) from ϕ. The ***size*** of a $\mathrm{Res}(\mathrm{lin}_R)$ derivation is the total size of all the clauses in the derivation, where the size of a clause is defined to be the total number of occurrences of variables in it plus the total size of all the coefficient occurring in the clause. The size of a coefficient when using integers (or integers embedded in characteristic zero rings) will be the standard size of the binary representation of integers.

In this definition we assume that R is a non-trivial ($R \neq \mathbf{0}$) ring such that there are polynomial-time algorithms for addition, multiplication and taking additive inverses.

Along with size, we will be dealing with two complexity measures of derivations: *width* and *principal width.*

Definition 2. *A clause $C = (f_1 = 0 \vee \cdots \vee f_m = 0)$ has **width** $\omega(C) = m$ and **principal width** $\omega_0(C) = \left|\{f_i\}_{i\in[m]}/_\sim\right|$ where $\sim$ identifies R-linear forms $f_i = 0$ and $f_j = 0$ if they define parallel hyperplanes, that is, if $f_i = Af_j + B$ or $f_j = Af_i + B$ for some $A, B \in R$. For $\mu \in \{\omega, \omega_0\}$, the measure μ associated with a* $\mathrm{Res}(\mathrm{lin}_R)$ *derivation $\pi = (D_1, \ldots, D_s)$ is $\mu(\pi) := \max_{1\leq i\leq s} \mu(D_i)$. For $\phi \in$ UNSAT, denote by $\mu(\phi \vdash \perp)$ the minimal value of $\mu(\pi)$ over all* $\mathrm{Res}(\mathrm{lin}_R)$ *refutations π.*

Proposition 4. $\mathrm{Res}(\mathrm{lin}_R)$ *is sound and complete. It is also implicationally complete, that is if ϕ is a set of linear clauses and C is a linear clause such that $\phi \models C$, then there exists a* $\mathrm{Res}(\mathrm{lin}_R)$ *derivation of C from ϕ.*

Proof: The soundness can be checked by inspecting that each rule of $\text{Res}(\text{lin}_R)$ is sound. Implicational completeness (and thus completeness) follows from Proposition 18. $\square$

We now define two systems of resolution with linear equations over a ring, where some of the rules are semantic: $\text{Res}_{sw}(\text{lin}_R)$ and $\text{Sem-Res}(\text{lin}_R)$. $\text{Res}_{sw}(\text{lin}_R)$ is obtained from $\text{Res}(\text{lin}_R)$ by replacing the boolean axioms with $0 = 0$, discarding simplification rule and replacing the weakening rule with the following *semantic weakening rule*:

$$\text{(Semantic weakening)}\ \frac{C}{D}\ (C \models D)$$

The system $\text{Sem-Res}(\text{lin}_R)$ has no axioms except for $0 = 0$, and has only the following *semantic resolution rule*:

$$\text{(Semantic resolution)}\ \frac{C \qquad C'}{D}\ (C \wedge C' \models D)$$

It is easy to see that $\text{Res}(\text{lin}_R) \leq_p \text{Res}_{sw}(\text{lin}_R) \leq_p \text{Sem-Res}(\text{lin}_R)$, where $P \leq_p Q$ denotes that Q polynomially simulates P.

In contrast to the case $R = \mathbb{F}_2$ (see [IS14]), for rings R with $char(R) \notin \{1, 2, 3\}$ both $\text{Res}_{sw}(\text{lin}_R)$ and $\text{Sem-Res}(\text{lin}_R)$ are not Cook-Reckhow proof systems, unless $\mathsf{P} = \mathsf{NP}$:

Proposition 5. *The following decision problem is* coNP*-complete: given a linear clause over a ring* R *with* $char(R) \notin \{1, 2, 3\}$ *decide whether it is a tautology under 0-1 assignments.*

Proof: Consider a 3-DNF ϕ and encode every conjunct $(x_{i_1}^{\sigma_1} \wedge \cdots \wedge x_{i_k}^{\sigma_k}) \in \phi, 1 \leq k \leq 3, \sigma_i \in \{0, 1\}$ as the equation $(1 - 2\sigma_1)x_1 + \cdots + (1 - 2\sigma_k)x_k = k - (\sigma_1 + \cdots + \sigma_k)$, where $x^0 := x, x^1 := \neg x$. Then ϕ is tautological if and only if the disjunction of these linear equations is tautological (that is, for every 0-1 assignment to the variables at least one of the equations hold, when the equations are computed over a ring with characteristic zero or finite characteristic bigger than 3). $\square$

We leave it as an open question to determine the complexity of verifying a correct application of the semantic weakening in case $char(R) = 3$ or in case $char(R) = 2$ and $R \neq \mathbb{F}_2$. In the case $R = \mathbb{F}_2$ the negation of a clause is a system of linear equations and thus the existence of solutions for it can be checked in polynomial time. Therefore $\text{Res}_{sw}(\text{lin}_{\mathbb{F}_2})$ is a Cook-Reckhow propositional proof system. The definitions of $\text{Res}(\text{lin}_{\mathbb{F}_2})$, $\text{Res}_{sw}(\text{lin}_{\mathbb{F}_2})$ and $\text{Sem-Res}(\text{lin}_{\mathbb{F}_2})$ coincide with the definitions of syntactic $\text{Res}(\oplus)$, $\text{Res}(\oplus)$ and $\text{Res}_{\text{sem}}(\oplus)$ from [IS14], respectively[3]. As showed in [IS14], $\text{Res}(\text{lin}_{\mathbb{F}_2})$, $\text{Res}_{sw}(\text{lin}_{\mathbb{F}_2})$ and $\text{Sem-Res}(\text{lin}_{\mathbb{F}_2})$ are polynomially equivalent.

We now show that if $char(R) \notin \{1, 2, 3\}$, then $\text{Res}_{sw}(\text{lin}_R)$ is polynomially bounded as a proof system for 3-UNSAT (that is, admits polynomial-size refutation for every instance):

[3] There is, however, one minor difference in the formulation of syntactic $\text{Res}(\oplus)$ and $\text{Res}(\text{lin}_{\mathbb{F}_2})$: the former does not have the boolean axioms, but has an extra rule (*addition rule*).

Proposition 6. *If $char(R) \notin \{1,2,3\}$, then dag-like* $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ *and tree-like* Sem-Res(lin_R) are polynomially bounded (not necessarily Cook-Reckhow) propositionally proof systems for 3-UNSAT.

Proof: Let $\phi(x_1,\dots,x_n) = \{C_i\}_{i\in[m]} \in$ 3-UNSAT. Given $C = (x_{j_1}^{\sigma_1} \vee \dots \vee x_{j_k}^{\sigma_k})$ define $lin(\neg C) := ((2\sigma_1 - 1)x_{j_1} + \dots + (2\sigma_k - 1)x_{j_k} - (\sigma_1 + \dots + \sigma_k))$ where $\sigma_i \in \{0,1\}, j_l \in [n], x^0 := x, x^1 := \neg x$. The linear clause $lin(\neg\phi) := \bigvee_{i\in[m]} lin(\neg C_i) = 0$ is a tautology (under 0-1 assignments) and thus can be derived in $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ in a single step as a weakening of $0 = 0$ or resolving $0 = 0$ with $0 = 0$ in tree-like Sem-Res(lin_R).

In tree-like Sem-Res(lin_R) the disjunct $lin(\neg C_i) = 0$ can be eliminated from $lin(\neg\phi)$ by a single resolution with C_i, thus the empty clause is derived by a sequence of m resolutions of $lin(\neg\phi)$ with $C_1,\dots,C_m$.

Similarly, the disjuncts $lin(\neg C_i) = 0$ are eliminated from $lin(\neg\phi)$ in $\mathrm{Res}_{sw}(\mathrm{lin}_R)$, but with a few more steps. Let D_0 be the empty clause and $D_{s+1} := D_s \vee lin(\neg C_{s+1}) = 0, 0 \le s < m$. Assume D_{s+1} is derived and assume without loss of generality, that $C_{s+1} = (x_1 = 1 \vee \dots \vee x_k = 1)$ and thus $lin(\neg C_{s+1}) = (-x_1 - \dots - x_k)$. Derive D_s as follows. Resolve D_{s+1} with C_{s+1} on $lin(\neg C_{s+1}) + (x_k - 1)$ to get the clause $E_1 := D_s \vee (-x_1 - \dots - x_{k-1} - 1) = 0 \vee x_1 = 1 \vee \dots \vee x_{k-1} = 1$ and apply semantic weakening to get $E_1' := D_s \vee x_1 = 1 \vee \dots \vee x_{k-1} = 1$. Resolve D_{s+1} with E_1' on $lin(\neg C_{s+1}) + (x_{k-1} - 1)$ and apply semantic weakening to get the clause $E_2' := D_s \vee x_1 = 1 \vee \dots \vee x_{k-2} = 1$. After k steps the clause $D_s = E_k'$ can be derived. $\square$

The following proposition is straightforward, but useful as it allows, for example, to transfer results about Res($\mathrm{lin}_{\mathbb{Q}}$) to Res($\mathrm{lin}_{\mathbb{Z}}$).

Proposition 7. *If R is an integral domain and $Frac(R)$ is its field of fractions, then* Res(lin_R) *is equivalent to* Res($\mathrm{lin}_{Frac(R)}$) *and tree-like* Res(lin_R) *is equivalent to tree-like* Res($\mathrm{lin}_{Frac(R)}$).

Proof: Every proof in tree-like Res(lin_R) is also a proof in tree-like Res($\mathrm{lin}_{Frac(R)}$). To get the converse, just multiply every line by the least common multiple of all the coefficients in the tree-like Res($\mathrm{lin}_{Frac(R)}$) proof. $\square$

3.1 Basic Counting in Res(lin_R) and $\mathrm{Res}_{sw}(\mathrm{lin}_R)$

Here we introduce several unsatisfiable sets of linear clauses that express some counting principles, and serve to exemplify the ability of dag-like Res(lin_R), tree-like Res(lin_R) and tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ to reason about counting, for a ring R. We then summarize what we know about refutations of these instance in our different systems, proving along the way some upper bounds and stating some lower bounds proved in the sequel.

Our unsatisfiable instances are the following:

Linear systems: If $A = (B|b)$ is an $m \times (n+1)$ matrix over R, where the B sub-matrix consists of the first n columns, such that $B\overline{x} = b$ has no 0-1 solutions, then (B_i is the ith row in B):

$$\mathsf{LinSys}(A) := \{B_i \cdot \overline{x} = b_i\}_{i\in[m]} . \tag{3}$$

Subset Sum: Let f be a linear form over R such that $0 \notin im_2(f)$. Then,

$$\mathsf{SubSum}(f) := \{f = 0\} . \tag{4}$$

Image avoidance: Let f be a linear form over R and recall the notation $\langle f \neq A\rangle$ from Sec. 2.1. We define

$$\mathsf{ImAv}(f) := \{\langle f \neq A\rangle : A \in im_2(f)\} . \tag{5}$$

We also consider the following (tautological) generalization of the Boolean axiom $x = 0 \vee x = 1$.

Image axiom: For f a linear form, define

$$\mathsf{Im}(f) := \bigvee_{A\in im_2(f)} f = A . \tag{6}$$

The complexity of Res(lin$_R$) derivations of $\mathsf{Im}(f)$ clauses is related to the complexity of Res(lin$_R$) refutations of $\mathsf{SubSum}(f)$: we prove that out of any refutation of $\mathsf{SubSum}(f)$ a derivation of $\mathsf{Im}(f)$ of the same size (up to a constant) can be constructed (Lemma 21) and vise versa (proof of Proposition 9).

Dag-like Res(lin$_R$)

<u>Upper bounds.</u> For any given linear form f, $\mathsf{Im}(f)$ has a Res(lin$_R$)-derivation of polynomial-size (in the size of $\mathsf{Im}(f)$):

Proposition 8. *Let $f = \sum_{i=1}^{n} a_i x_i + b$ be a linear form over R. There exists a* Res(lin$_R$) *derivation of $\mathsf{Im}(f)$ of size polynomial in $|\mathsf{Im}(f)|$ and of principal width at most* 3.

Proof: We construct derivations of $\mathsf{Im}\left(\sum_{i=1}^{k} a_i x_i + b\right)$, $0 \le k \le n$, inductively on k.

Base case: $k = 0$. In this case $\mathsf{Im}(b)$ is just the axiom $b = b$ and thus derived in one step.

Induction step: Let $f_k := \sum_{i=1}^{k} a_i x_i + b$ and assume $\mathsf{Im}(f_k)$ was already derived. Derive $C_0 := \left(\bigvee_{A\in im_2(f_k)} f_k + a_{k+1}x_{k+1} = A\right) \vee x_{k+1} = 1$ from $\mathsf{Im}(f_k)$ by $|im_2(f_k)|$ many resolution applications with $x_{k+1} = 0 \vee x_{k+1} = 1$. Similarly derive $C_1 := \left(\bigvee_{A\in im_2(f_k)} f_k + a_{k+1}x_{k+1} = A + a_{k+1}\right) \vee x_{k+1} = 0$ and obtain $\mathsf{Im}(f_{k+1})$ by resolving C_0 with C_1 on x_{k+1}. The size of the derivation is $n \cdot |\mathsf{Im}(f)|$, and as there is no clause with more than 3 equations that determines non-parallel hyperplanes, hence the principal width of the derivation is at most 3. $\square$

Proposition 9. *For every linear form f such that $0 \notin im_2(f)$, the contradiction* $\mathsf{SubSum}(f)$ *admits* $\mathrm{Res}(\mathrm{lin}_R)$ *refutation of size polynomial in* $|\mathsf{Im}(f)|$.

Proof: First construct the shortest derivation of $\mathsf{Im}(f)$, and then by a sequence of $|im_2(f)|$ many application of the resolution rule with $f = 0$ derive the empty clause. By Proposition 8 the resulting refutation is of polynomial in $|\mathsf{Im}(f)|$ size. $\square$

Proposition 10. *Let f be a linear form over R, $a \in im_2(f)$ and $\phi = \{\langle f \neq b \rangle\}_{b \in im_2(f), b \neq a}$. Then there exists* $\mathrm{Res}(\mathrm{lin}_R)$ *derivation π of $f = a$ from ϕ, such that $S(\pi) = poly(|\phi|)$ and $\omega_0(\pi) \leq 3$.*

Proof: Let $A_1, \ldots, A_N = a$ be an enumeration of all the elements in $im_2(f)$. By Proposition 8 there exists a derivation of $\left(\bigvee_{i \geq 1} f = A_i\right)$ of principal width at most 3. For $1 < k < N$, we derive $C := \left(\bigvee_{i \geq k+1} f = A_i\right)$ from $\left(\bigvee_{i \geq k} f = A_i\right) = (C \vee f = A_k)$ and $\langle f \neq A_k \rangle = (C \vee f = A_1 \vee \cdots \vee f = A_{k-1})$ in $k - 1$ steps as follows: at the sth step we get $(C \vee f - f = A_s - A_k \vee f = A_{s+1} \vee \cdots \vee f = A_{k-1}) = (C \vee f = A_{s+1} \vee \cdots \vee f = A_{k-1})$ by resolving $C \vee f = A_s \vee \cdots \vee f = A_{k-1}$ with $C \vee f = A_k$. We thus obtain a derivation of principal width $\omega_0 \leq 3$ and of size $(1 + \cdots + (N-2))|f| = \frac{(N-1)(N-2)}{2}|f|$. $\square$

Corollary 11. *For every linear form f the contradiction* $\mathsf{ImAv}(f)$ *admits polynomial-size* $\mathrm{Res}(\mathrm{lin}_R)$ *refutations.*

Proof: Pick some $a \in im_2(f)$. By Proposition 10 there is a derivation of $f = a$ from $\mathsf{ImAv}(f)$ of polynomial size. This derivation can be extended to a refutation of $\mathsf{ImAv}(f)$ by a sequence of resolution rule applications of $f = a$ with $\langle f \neq a \rangle \in \mathsf{ImAv}(f)$. $\square$

The only $\mathrm{Res}(\mathrm{lin}_R)$ upper bounds for $\mathsf{LinSys}(A)$ we have so far are tree-like. So for $\mathsf{LinSys}(A)$ we refer the reader to the tree-like $\mathrm{Res}(\mathrm{lin}_R)$ upper bounds further in this section.

Lower bounds. In Sec. 5.2 we prove an exponential lower bound for $\mathsf{SubSum}(f)$ in case f is a linear form with large coefficients (Theorem 22).

Tree-like $\mathrm{Res}(\mathrm{lin}_R)$

Upper bounds. In case R is a finite ring, in Sec. 4 we prove that the clauses in $\mathsf{Im}(f)$ admit derivations of polynomial size (Theorem 19). Obviously, in that case (R is finite) any unsatisfiable R-linear equation $f = 0$ has at most $|R|$ variables and $\mathsf{SubSum}(f)$ are always refutable in constant size. In case R is a field of characteristic zero we prove a lower bound for $\mathsf{Im}(f)$, $\mathsf{SubSum}(f)$ and $\mathsf{ImAv}(f)$ for a specific f with small coefficients (see the lower bounds below).

In case a matrix $A = (B|b)$ with entries in a field $\mathbb{F}$ defines a system of equations $B\overline{x} = b$, that is unsatisfiable under arbitrary $\mathbb{F}$-valued assignments (not just under 0-1 assignments), we prove a polynomial upper bound for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutations of $\mathsf{LinSys}(A)$.

Proposition 12. *If a $m \times (n+1)$ matrix $A = (B|b)$ with entries in a field $\mathbb{F}$ is such that $B\overline{x} = b$ has no $\mathbb{F}$-valued solutions, then there exists tree-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ *refutation of* $\mathsf{LinSys}(A)$ *of linear size.*

Proof: It is a well-known fact from linear algebra that $B\overline{x} = b$ has no $\mathbb{F}$-valued solutions iff there exists $\alpha \in \mathbb{F}^m$ such that $\alpha^T B = 0$ and $\alpha^T b = 1$. Therefore, by $m-1$ resolutions of $B_1\overline{x} - b_1 = 0, \ldots, B_m\overline{x} - b_m = 0$ we can derive $-\alpha_1(B_1\overline{x} - b_1) - \ldots - \alpha_m(B_m\overline{x} - b_m) = 0$, which is $1 = 0$. □

<u>Lower bounds.</u> Let $\mathbb{F}$ be a field of characteristic zero. In Sec. 5.2 we prove tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ exponential-size lower bounds for derivations of $\mathsf{Im}(f)$ and refutations of $\mathsf{SubSum}(f)$ and $\mathsf{ImAv}(f)$ whenever f is of the form $f = \epsilon_1 x_1 + \ldots + \epsilon_n x_n - A$ for some $\epsilon_i \in \{-1, 1\}, A \in \mathbb{F}$ (Proposition 24 and Corollary 25).

Tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$

<u>Upper bounds.</u> Most of the instances above admit short derivations/refutations in tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$: $\mathsf{Im}(f)$ is semantic weakening of $0 = 0$ and thus derivable in one step; The empty clause is a semantic weakening of $\mathsf{SubSum}(f)$ and $\mathsf{LinSys}(A)$ and thus can be refuted via deriving $\bigvee_{i\in[m]}\langle A_i\overline{x} - b_i \neq 0\rangle$ as a semantic weakening of $0 = 0$ and resolving it with equalities in $\mathsf{LinSys}(A) = \{A_i\overline{x} - b_i = 0\}_{i\in[m]}$.
<u>Lower bounds.</u> In case $\mathbb{F}$ is a field of characteristic zero, $\mathsf{ImAv}(f)$ are hard even for tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ whenever f is of the form $f = \epsilon_1 x_1 + \ldots + \epsilon_n x_n - A$ for some $\epsilon_i \in \{-1, 1\}, A \in \mathbb{F}$ (Proposition 24).

3.2 CNF Upper Bounds for $\mathrm{Res}(\mathrm{lin}_R)$

In this section we outline two basic polynomial upper bounds, which we use to establish our separations in subsequent sections: short tree-like $\mathrm{Res}(\mathrm{lin}_R)$ refutations for CNF encodings of linear systems over a ring R, and short $\mathrm{Res}(\mathrm{lin}_R)$ refutations for $\neg\mathrm{PHP}^m_n$. Together with our lower bounds, these imply the separation between tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ and tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}'})$, where $\mathbb{F}, \mathbb{F}'$ are fields of positive characteristic such that $char(\mathbb{F}) \neq char(\mathbb{F}')$. The short refutation of the pigeonhole principle will imply a separation between dag-like and tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ for fields $\mathbb{F}$ of characteristic 0.

In what follows we consider standard CNF encodings of linear equations $f = 0$ where the linear equations are considered as Boolean functions (i.e., functions from 0-1 assignments to $\{0, 1\}$); we do not use extension variable in these encodings.

Proposition 13. *Let $\mathbb{F}$ be a field and $A\overline{x} = b$ be a system of linear equations that has no solution over $\mathbb{F}$, where A is $k \times n$ matrix with entries in $\mathbb{F}$, and A_i denotes the ith row in A. Assume that ϕ_i is a CNF encoding of $A_i \cdot \overline{x} - b_i = 0$, for $i \in [k]$. Then, there exists a tree-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ *refutation of $\phi = \{\phi_i\}_{i\in[k]}$ of size polynomial in $|\phi| + \sum_{i\in[k]} \left|A_i \cdot \overline{x} - b_i = 0\right|$.*

Proof: The idea is to derive the actual linear system of equations from their CNF encoding, and then refute the linear system using a previous upper bound (Proposition 12).

If n_i is the number of variables in $A_i \cdot \overline{x} - b_i = 0$, then $|\phi_i| = \Theta(2^{n_i})$. By Proposition 18 proved in the sequel there exists a tree-like Res($\text{lin}_{\mathbb{F}}$) derivation of $A_i \cdot \overline{x} - b_i = 0$ from ϕ_i of size $O(2^{n_i}|A_i \cdot \overline{x} - b_i = 0|) = O(|\phi_i| \cdot \big|A_i \cdot \overline{x} - b_i = 0\big|)$.

By Proposition 12 there exists a tree-like Res($\text{lin}_{\mathbb{F}}$) refutation of $\{A_i \cdot \overline{x} - b_i = 0\}_{i\in[k]}$ of size $O\left(\sum_{i\in[k]} |A_i \cdot \overline{x} - b_i = 0|\right)$. The total size of the resulting refutation of ϕ is $O\left(\sum_{i\in[k]} |\phi_i| \cdot |A_i \cdot \overline{x} - b_i = 0|\right)$ and thus is $O\left(\left(\sum_{i\in[k]} |\phi_i| + \sum_{i\in[k]} |A_i \cdot \overline{x} - b_i = 0|\right)^2\right) = O\left(\left(|\phi| + \sum_{i\in[k]} |A_i \cdot \overline{x} - b_i = 0|\right)^2\right)$. □

As a corollary we get the polynomial upper bound for the Tseitin formulas (see Sec. 2.3.2 for the definition):

Theorem 14. *Let $G = (V, E)$ be a d-regular directed graph, p a prime number, $\sigma : V \to \mathbb{F}_p$ such that $\sum_{u\in V} \sigma(u) \not\equiv 0 \pmod p$, then $\neg\mathrm{TS}^{(p)}_{G,\sigma}$ admit tree-like* Res($\text{lin}_{\mathbb{F}_p}$) *refutations of polynomial size.*

Proof: $\neg\mathrm{TS}^{(p)}_{G,\sigma}$ is an unsatisfiable system of linear equations over $\mathbb{F}_p$ (note that no assignment of $\mathbb{F}$-elements to the variables in $\neg\mathrm{TS}^{(p)}_{G,\sigma}$ is satisfying, and so we do not need to use the (non-linear) Boolean axioms to get the unsatisfiability of the system of equations). Therefore, by Proposition 13 there exists a tree-like Res($\text{lin}_{\mathbb{F}_p}$) refutation of $\neg\mathrm{TS}^{(p)}_{G,\sigma}$ of polynomial size. □

Theorem 15 ([RT08])**.** *Let R be a ring such that $char(R) = 0$. There exists a* Res(lin_R) *refutation of $\neg PHP^m_n$ of polynomial size.*

Proof: This follows from the upper bound of [RT08] for Res($\text{lin}_{\mathbb{Z}}$) and the fact that any Res($\text{lin}_{\mathbb{Z}}$) proof can be interpreted as Res(lin_R) if R is of characteristic 0. □

4 Nondeterministic Linear Decision Trees

In this section we extend the classical correspondence between tree-like resolution refutations and decision trees (cf. [BKS04]) to tree-like Res(lin_R) and tree-like $\text{Res}_{sw}(\text{lin}_R)$. We define *nondeterministic linear decision trees* (NLDT), which generalize parity decision trees, proposed in [IS14] for $R = \mathbb{F}_2$, to arbitrary rings. We shall use these trees in the sequel to establish some of our upper and lower bounds.

Let ϕ be a set of linear clauses (that we wish to refute) and Φ a set of linear non-equalities over R (that we take as assumptions). Consider the following two decision problems:

- DP1 Assume $\Phi \models \neg\phi$. Given a satisfying Boolean assignment ρ to Φ, determine which clause $C \in \phi$ is violated by ρ by making queries of the form: which of $f|_\rho \neq 0$ or $g|_\rho \neq 0$ hold for linear forms f, g in case $f|_\rho + g|_\rho \neq 0$.
- DP2 Similar to DP1, only that we assume $\Phi \models_R \neg\phi$, and given R-valued assignment ρ, satisfying Φ, we ask to find a clause $C \in \phi$ falsified by ρ.

Below we define NLDTs of types $\mathrm{DT}_{sw}(R)$ and $\mathrm{DT}(R)$, which provide solutions to DP1 and DP2, respectively. The root of a tree is labeled with a system Φ, the edges in a tree are labeled with linear non-equalities of the form $f \neq 0$ and the leaves are labeled with clauses $C \in \phi$. Informally, at every node v there is a set Φ_v of all *learned* non-equalities, which is the union of Φ and the set of non-equalities along the path from the root to the node. If v is an internal node, two outgoing edges $f \neq 0$ and $g \neq 0$ define a query to be made at v, where $f + g \neq 0$ is a consequence of Φ_v. If v is a leaf, then $\Phi_v \cup \Phi$ contradicts a clause $C \in \phi$.

Starting from the root, based on the assignment ρ, we go along a path, from the root to a leaf, by choosing in each node to go along the left edge $f \neq 0$ or the right edge $g \neq 0$, depending on whether $f|_\rho \neq 0$ or $g|_\rho \neq 0$. Note that $f|_\rho \neq 0$ and $g|_\rho \neq 0$ may not be mutually exclusive, and this is why the decision made in each node may be *nondeterministic*.

Definition 3 (Nondeterministic linear decision tree NLDT; $\mathrm{DT}(R)$, $\mathrm{DT}_{sw}(R)$)**.** *Let ϕ be a set of linear clauses and Φ be a set of linear non-equalities over a ring R. A nondeterministic linear decision tree T of type $DT(R)$ and of type $DT_{sw}(R)$ for (ϕ, Φ) is a binary rooted tree, where every edge is labeled with some linear non-equality $f \neq 0$, in such a way that the conditions below hold. In what follows, for a node v, we denote by $\Phi_{r \rightsquigarrow v}$ the set of non-equalities along the path from the root r to v and by Φ_v the set $\Phi_{r \rightsquigarrow v} \cup \Phi$. We say that Φ_v is the set of* learned *non-equalities at v.*

1. *Let v be an internal node. Then v has two outgoing edges labeled by linear non-equalities $f_v \neq 0$ and $g_v \neq 0$, such that:*
   - *If $T \in DT(R)$, then $\alpha f_v + \beta g_v \neq 0 \in \Phi_v \cup \{a \neq 0 \mid a \in R \setminus 0\}$ for some $\alpha, \beta \in R$.*
   - *If $T \in DT_{sw}(R)$, then $\Phi_v \models \alpha f_v + \beta g_v \neq 0$ for some $\alpha, \beta \in R$.*
2. *A node v is a leaf if there is a linear clause $C \in \phi \cup \{0 = 0\}$ which is violated by Φ_v in the following sense:*
   - *If $T \in DT(R)$, then $\neg C \subseteq \Phi_v \cup \{a \neq 0 \mid a \in R \setminus 0\}$.*
   - *If $T \in DT_{sw}(R)$, then $\Phi_v \models \neg C$.*

In case Φ is empty, we sometimes simply write that the NLDT is for ϕ instead of $(\phi, \emptyset)$.

Assume $\Phi \models \neg\phi$. Then an NLDT for $(\phi \cup \{x = 0 \vee x = 1 \,|\, x \in vars(\phi)\}, \Phi)$ of type $\mathrm{DT}(R)$ can be converted into an NLDT of type $\mathrm{DT}_{sw}(R)$ for (ϕ, Φ) by truncating all maximal subtrees with all leaves from $\{x = 0 \vee x = 1 \,|\, x \in vars(\phi)\}$ and marking their roots with arbitrary clauses from ϕ.

Below we give several examples (and basic properties) of NLDTs.

Example 1 Let ϕ be a set of clauses, representing unsatisfiable CNF. Then any standard decision tree on Boolean variables is an NLDT for $\phi \cup \{x = 0 \vee x = 1 \,|\, x \in vars(\phi)\}$ of type DT(R), where a branching on the value of a variable x is realized by branching on $(1 - x) + x \neq 0$ to either $1 - x \neq 0$ or $x \neq 0$. This is illustrated by (the proof of) the following proposition:

Proposition 16. *If Φ is a set of linear non-equalities and ϕ is a set of linear clauses over R such that $\Phi \models \neg\phi$, then there exists a DT(R) tree for $(\phi \cup \{x = 0 \vee x = 1 \,|\, x \in vars(\phi \cup \{\neg\Phi\})\}, \Phi)$ of size $O(2^n|\Phi|)$, where $n = |vars(\phi \cup \{\neg\Phi\})|$.*

Proof: Let $vars(\phi \cup \{\neg\Phi\}) = \{x_1, \dots, x_n\}$ and fix an ordering on these variables. Construct a tree T_0 with 2^n nodes, that branches on $x_1, \dots, x_n$, in this order. Thus, in every leaf v of T_0 a total assignment to the variables is determined (i.e., $\Phi_v = \{x_i \neq \nu_i\}_{i\in[n]} \cup \Phi$ for some $\nu_i \in \{0, 1\}$). Since $\Phi \models \neg\phi$, this assignment violates either some clause $C = (f_1 = 0 \vee \dots \vee f_m = 0)$ in ϕ or some non-equality $g \neq 0$ in Φ. We augment T_0 to T by attaching a subtree to every leaf v of T_0 depending on whether the former or latter condition holds for v, as follows:
Case 1: $\{x_i \neq \nu_i\}_{i\in[n]} \models \neg C$. We attach a subtree to v that makes m sequences of branches as follows. If $f_i = a_1x_1 + \dots + a_nx_n + b$ then $a_1(1 - \nu_1) + \dots + a_n(1 - \nu_n) + b \neq 0$ holds and the ith sequence is the following sequence of "substitutions": $(a_1x_1 + a_2(1-\nu_2) + \dots + a_n(1-\nu_n)+b) + (a_1(1-\nu_1)-a_1x_1) \neq 0$ to $a_1x_1+a_2(1-\nu_2)+\dots+a_n(1-\nu_n)+b \neq 0$ and $a_1(1-\nu_1)-a_1x_1 \neq 0$, ..., $(a_1x_1 + \dots + a_{n-1}x_{n-1} + a_n(1-\nu_n)+b)+(a_n(1-\nu_n)-a_nx_n) \neq 0$ to $f_i \neq 0$ and $a_n(1 - \nu_n) - a_nx_n \neq 0$. All the right branches lead to nodes u such that $\{x_i \neq 0, x_i \neq 1\} \subseteq \Phi_u$ for some $i \in [n]$ and thus they satisfy the DT(R) leaf condition in Definition 3. Such a sequence indeed performs substitutions: the edge to the leftmost node is $f_i \neq 0$ and as we go upwards, we apply the substitutions $x_n \leftarrow 1 - \nu_n$, ..., $x_1 \leftarrow 1 - \nu_1$ to this non-equality.

In the leftmost node w in the end of the mth sequence, $\{f_1 \neq 0, \dots, f_m \neq 0\} \subseteq \Phi_w$ holds and thus again C is violated at w in the sense of Definition 3 and therefore w is a legal DT(R)-leaf.
Case 2: $\{x_i \neq \nu_i\}_{i\in[n]} \models g = 0$, where $g \neq 0 \in \Phi_v$. Let $g = a_1x_1 + \dots + a_nx_n + b$. Attach to v a subtree that makes the following branches: $(a_1(1 - \nu_1) + a_2x_2 + \dots + a_nx_n + b) - (a_1(1-\nu_1)-a_1x_1) \neq 0$ to $(a_1(1-\nu_1)+a_2x_2+\dots+a_nx_n+b) \neq 0$ and $a_1(1-\nu_1)-a_1x_1 \neq 0$, ..., $(a_1(1 - \nu_1) + \dots + a_{n-1}(1 - \nu_{n-1}) + a_n(1 - \nu_n) + b) - (a_n(1 - \nu_n) - a_nx_n) \neq 0$ to $1 \neq 0$ and $a_1(1 - \nu_1) - a_1x_1 \neq 0$. All leaves of the subtree satisfy the condition for DT(R) leaves in Definition 3.

The tree T is a DT(R) tree for (ϕ, Φ). □

Example 2 Let ϕ be as in Example 1. *Parity decision trees*, as defined in [IS14], are NLDTs for ϕ of type DT$_{sw}(\mathbb{F}_2)$: branching on the value of an $\mathbb{F}_2$-linear form f is realized by branching from $(1 - f) + f \neq 0$ to $1 - f \neq 0$ and $f \neq 0$. And the converse also holds: a branching of $f + g \neq 0$ to $f \neq 0$ and $g \neq 0$, where, say, f is a non-constant $\mathbb{F}_2$-linear form, is equivalent to branching on the value of f.

Example 3 Let $\phi = \{f_1 = 0, \dots, f_m = 0\}$, where $f_1, \dots, f_m$ are R-linear forms such that $f_1 + \ldots + f_m = 1$. Then a polynomial-size NLDT of type DT(R) for ϕ makes the following branchings, where all right edges lead to a leaf: $(f_1 + \ldots + f_{m-1}) + f_m \neq 0$ (this is just $1 \neq 0$) to $f_1 + \ldots + f_{m-1} \neq 0$ and $f_m \neq 0$, ..., $f_1 + f_2 \neq 0$ to $f_1 \neq 0$ and $f_2 \neq 0$.

We now show the equivalence between NLDTs and tree-like Res(lin$_R$) proofs.

Theorem 17. *Let ϕ be a set of linear clauses over a ring R and Φ be a set of linear non-equalities over R. Then, there exist decision trees $DT(R)$ (resp. $DT_{sw}(R)$) for $(\phi \cup \{x = 0 \vee x = 1 \,|\, x \in vars(\phi)\}, \Phi)$ (resp. (ϕ, Φ)) of size s iff there exist tree-like* Res(lin$_R$) *(resp. tree-like* Res$_{sw}$(lin$_R$)*) derivations of the clause $\neg\Phi = \bigvee_{f \neq 0 \in \Phi} f = 0$ from ϕ of size $O(s)$.*

Proof: ($\Rightarrow$) Let T_ϕ be an NLDT of type DT(R) or DT$_{sw}$(R) for ϕ. We construct a tree-like Res(lin$_R$) or tree-like Res$_{sw}$(lin$_R$) derivation from T_ϕ, respectively, as follows. Consider the tree of clauses π_0, obtained from T_ϕ by replacing every vertex u with the clause $\neg\Phi_u$. This tree is not a valid tree-like derivation yet. We augment it to a valid derivation π by appropriate insertions of applications of weakening and simplification rules.

Case 1: If $\neg\Phi_u \in \pi_0$ is a leaf, then Φ_u violates a clause $D \in \phi \cup \{0 = 0\}$. By condition 2 in Definition 3, $\neg\Phi_u$ must be a weakening of D (syntactic for $T_\phi \in$ DT(R) and semantic for $T_\phi \in$ DT$_{sw}$(R)) and we add D as the only child of this node.

Case 2: Let $\neg\Phi_u \in \pi_0$ be an internal node with two outgoing edges labeled with $f_u \neq 0$ and $g_u \neq 0$.

If $T_\phi \in$ DT(R), then $\alpha f_u + \beta g_u \neq 0 \in \Phi_u \cup \{a \neq 0 \,|\, a \in R \setminus 0\}$. Apply resolution to $\neg\Phi_{l(u)} = (\neg\Phi_u \vee f_u = 0)$ and $\neg\Phi_{r(u)} = (\neg\Phi_u \vee g_u = 0)$ to derive $\neg\Phi_u \vee \alpha f_u + \beta g_u = 0$. In case $\alpha f_u + \beta g_u \neq 0 \in \Phi_u$ this clause coincides with $\neg\Phi_u$ and no additional steps are required. In case $\alpha f_u + \beta g_u \neq 0 \in \{a \neq 0 \,|\, a \in R \setminus 0\}$ insert an application of the simplification rule to get a derivation of $\neg\Phi_u$.

If $T_\phi \in$ DT$_{sw}$(R), $\Phi_u \models \alpha f_u + \beta g_u \neq 0$, we derive $\neg\Phi_u \vee \alpha f_u + \beta g_u = 0$ from $\neg\Phi_{l(u)} = (\neg\Phi_u \vee f_u = 0)$ and $\neg\Phi_{r(u)} = (\neg\Phi_u \vee g_u = 0)$ by an application of the resolution rule and then deriving $\neg\Phi_u$ by an application of the semantic weakening rule.

($\Leftarrow$) Conversely, assume π is a tree-like Res(lin$_R$) or a tree-like Res$_{sw}$(lin$_R$) derivation of a (possibly empty) clause $\mathcal{C}$ from ϕ. In what follows, when we say weakening we mean syntactic or semantic weakening depending on π being a tree-like Res(lin$_R$) or a tree-like Res$_{sw}$(lin$_R$) derivation, respectively.

Let the edges in the proof-tree of π be directed from conclusion to premises. We turn this proof-tree into a decision tree T_π for $(\phi, \neg\mathcal{C})$ as follows. Every node of outgoing degree 2 in the proof-tree π is a clause obtained from its children by a resolution rule. For each such node $C \vee D \vee (\alpha f + \beta g = 0)$ we label its outgoing edges to $C \vee f = 0$ and $D \vee g = 0$ with $f \neq 0$ and $g \neq 0$, respectively. We contract all unlabeled edges, which are precisely those corresponding to applications of weakening and simplification rules. If $C_1, \dots, C_k$ is a maximal (with respect to inclusion) sequence of weakening and simplification rule applications (the latter occur only in $\mathrm{Res}(\mathrm{lin}_R)$ derivations), then we contract it to C_k. In this way we obtain the tree T_π, where every edge is labeled with linear non-equality and every node u is labeled with a clause C_u such that if $f \neq 0$ and $g \neq 0$ are labels of edges to the left $l(u)$ and to the right $r(u)$ children respectively, then C_u is a weakening and a simplification (the latter again in case of $\mathrm{Res}(\mathrm{lin}_R)$) of the clause $C \vee D \vee \alpha f + \beta g = 0$ for some $\alpha, \beta \in R$, such that $C_{l(u)} = (C \vee f = 0)$, $C_{r(u)} = (D \vee g = 0)$.

We now prove that T_π is a valid decision tree of type $\mathrm{DT}(R)$ (respectively, $\mathrm{DT}_{sw}(R)$) if π is a tree-like $\mathrm{Res}(\mathrm{lin}_R)$ derivation (respectively, tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ derivation).

Case 1: Assume π is tree-like $\mathrm{Res}(\mathrm{lin}_R)$ derivation. We prove inductively that for every node u in T_π we have $\neg C_u \subseteq \Phi_u$.
Base case: u is the root r. We have $\Phi_r = \neg \mathcal{C} = \neg C_r$.
Induction step: For any other node u assume $\neg C_p \subseteq \Phi_p \cup \{a \neq 0 \,|\, a \in R \setminus 0\}$ holds for its parent node p. Let $f \neq 0$ be the label on the edge from p to u. Then $C_u = (C \vee f = 0)$ for some clause C and C_p must be of the form $(C \vee D)$ for some clause D, and hence $\neg C_u \subseteq \neg C \cup \{f \neq 0\} \subseteq \neg C_p \cup \{f \neq 0\} \subseteq \Phi_p \cup \{f \neq 0\} = \Phi_u$.

Now we show that T_π satisfies the conditions of Definition 3 for $\mathrm{DT}(R)$ trees.

- (Internal nodes) Let u be an internal node of T_π with outgoing edges labeled with $f \neq 0$ and $g \neq 0$. C_u must be both a weakening and a simplification of $(C \vee \alpha f + \beta g = 0)$ for some $\alpha, \beta \in R$ and a linear clause C. If $\alpha f + \beta g \neq 0 \in \{a \neq 0 \,|\, a \in R \setminus 0\}$, then the condition trivially holds, otherwise $\alpha f + \beta g = 0$ cannot be eliminated via simplification and thus $\alpha f + \beta g \neq 0 \in \neg C_u$ and $\neg C_u \subseteq \Phi_u$ imply $\alpha f + \beta g \neq 0 \in \Phi_u$ and the condition for internal nodes in Definition 3 is satisfied.

- (Leaves) Let u be a leaf of T_π. Then C_u must be both a weakening and a simplification of some clause C in $\phi \cup \{x = 0 \vee x = 1 \,|\, x \in vars(\phi)\} \cup \{0 = 0\}$, that is $C_u = (C \vee D)$ for some clause D. Therefore $\neg C_u \subseteq \Phi_u$ implies that C is falsified by Φ_u.

Case 2: Assume π is a tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ derivation. We prove inductively that for every node u in T_π, $C_u \models \neg \Phi_u$ holds.
Base case: u is the root r and we have $\neg \Phi_r = \mathcal{C} = C_r$.
Induction step: u is a node which is not the root. If $C_p \models \neg \Phi_p$ holds for its parent p and $f \neq 0$ is the label on the edge from p to u, then $(C \vee D \vee \alpha f + \beta g = 0) \models C_p$, $C_u = (C \vee f = 0)$ for some $\alpha, \beta \in R$ a linear form g and some linear clauses C, D. Therefore, $C_u = (C \vee f = 0) \models (C_p \vee f = 0) \models (\neg \Phi_p \vee f = 0) = \neg \Phi_u$.

We now show that T_π satisfies the conditions of Definition 3 for $\mathrm{DT}_{sw}(R)$ trees.

- (Internal nodes) Let u be an internal node of T_π with outgoing edges labeled with $f \neq 0$ and $g \neq 0$. Then $(C \vee \alpha f + \beta g = 0) \models C_u$ for some $\alpha, \beta \in R$ and a linear clause C. Therefore $C_u \models \neg \Phi_u$ implies $\Phi_u \models \alpha f + \beta g \neq 0$.

- (Leaves) Let u be a leaf of T_π. Then C_u must be a weakening of some clause C in $\phi \cup \{0 = 0\}$, that is, $C_u = (C \vee D)$ for some clause D. Therefore $C_u \models \neg \Phi_u$ implies that C is falsified by Φ_u.

$\square$

An immediate corollary is this:

Proposition 18. *If $\phi \cup \{C\}$ is a set of linear clauses over a ring R such that $\phi \models C$, then there exists a tree-like* $\mathrm{Res}(\mathrm{lin}_R)$ *derivation of C from ϕ of size $O(2^n|C|)$, where $n = \left|vars(\phi \cup \{C\})\right|$.*

Proof: By Proposition 16 there exists a $\mathrm{DT}(R)$ tree for $(\phi \cup \{x = 0 \vee x = 1 \,|\, x \in vars(\phi \cup \{C\})\}, \neg C)$ of size $O(2^n|C|)$ and, thus, by Theorem 17 there exists a tree-like $\mathrm{Res}(\mathrm{lin}_R)$ derivation of C from ϕ of size $O(2^n|C|)$. $\square$

We construct an NLDT to prove the following upper bound:

Proposition 19. *Let R be a finite ring, $f = a_1x_1 + \cdots + a_nx_n$ a linear form over R, s_f the size of $\mathsf{Im}(f)$ (i.e., the size of its encoding) and $d_f = |im_2(f)|$. Then, there exists a tree-like* $\mathrm{Res}(\mathrm{lin}_R)$ *derivation of $\mathsf{Im}(f)$ of size $O(s_f n^{2d_f})$.*

Proof: We construct a decision tree of type $\mathrm{DT}(R)$ of size $O(s_f n^{2d_f})$ with the system $\Phi_r = \{f \neq A\}_{A \in im_2(f)}$ at its root r. By Theorem 17 this implies the existence of a tree-like $\mathrm{Res}(\mathrm{lin}_R)$ proof of $\mathsf{Im}(f)$ of the same size.

Let $f^{(1)} := a_1x_1 + \cdots + a_{\lfloor \frac{n}{2} \rfloor}x_{\lfloor \frac{n}{2} \rfloor}$ and $f^{(2)} := a_{\lfloor \frac{n}{2} \rfloor+1}x_{\lfloor \frac{n}{2} \rfloor+1} + \cdots + a_nx_n$. The decision tree for $\mathsf{Im}(f)$ is constructed recursively as a tree of height $2d_f$, where a subtree for $\mathsf{Im}\big(f^{(1)}\big)$ or for $\mathsf{Im}\big(f^{(2)}\big)$ is hanged from each leaf. At every node u of depth d the system of non-equalities is of the form: $\Phi_u = \Phi_r \cup \Phi_u^{(1)} \cup \Phi_u^{(2)}$, where $\Phi_u^{(i)} \subseteq \{f^{(i)} \neq A\}_{A \in im_2(f^{(i)})}$, $i \in \{1, 2\}$ and $|\Phi_u^{(1)}| + |\Phi_u^{(2)}| = d$. A node u is a leaf if and only if $\Phi_u^{(i)} = \{f^{(i)} \neq A\}_{A \in im_2(f^{(i)})}$ for some $i \in \{1, 2\}$. The branching at an internal node u is made by the non-equality $f^{(1)} - A_1 + f^{(2)} - A_2 \neq 0$, for some $A_i \in im_2(f^{(i)})$ where $f^{(i)} - A_i \notin \Phi_u^{(i)}, i \in \{1, 2\}$. The size s_n of this tree can be upper bounded as follows: $s_n \leq 2^{2d_f} s_{\lfloor \frac{n}{2} \rfloor + 1} + s_f 2^{2d_f} = O(s_f n^{2d_f})$. $\square$

5 Lower Bounds for $\mathrm{Res}(\mathrm{lin}_R)$

5.1 Prover-Delayer Games

The *Prover-Delayer game* is an approach to obtain lower bounds on resolution refutations introduced by Pudlák and Impagliazzo [PI00]. The idea is that the non-existence of small decision trees, and hence small tree-like resolution refutations, for an unsatisfiable formula, can be phrased in terms of the existence of a certain strategy for Delayer in a game against Prover, associated to the unsatisfiable formula. We define such games G^R and G^R_{sw} for decision trees $\mathrm{DT}(R)$ and $\mathrm{DT}_{sw}(R)$, respectively. Below we show (Lemma 20) that the existence of certain strategies for the Delayer in G^R and G^R_{sw} imply lower bounds on the size of $\mathrm{DT}(R)$ and $\mathrm{DT}_{sw}(R)$ trees, respectively.

The game. Let ϕ be a set of linear clauses and Φ_s be a set of linear non-equalities. Consider the following game between two parties called Prover and Delayer. The game goes in rounds, consisting of one move of Prover followed by one move of Delayer. The position in the game is determined by a system of linear non-equalities Φ, which is extended by one non-equality after every round. The starting position is Φ_s.

In each round, Prover presents to Delayer a possible branching $f \neq 0$ and $g \neq 0$ over a linear non-equality $f + g \neq 0$, such that $f + g \neq 0 \in \Phi \cup \{a \neq 0 \,|\, a \in R \setminus 0\}$ or $\Phi \models f + g \neq 0$ in G^R and G^R_{sw}, respectively. After that, Delayer chooses either $f \neq 0$ or $g \neq 0$ to be added to Φ, or leaves the choice to the Prover and thus earns a coin. The game G^R finishes, when $\neg C \subseteq \Phi$ for some $C \in \phi \cup \{0 = 0\}$, and G^R_{sw} finishes, when $\Phi \models \neg C$ for some clause $C \in \phi \cup \{0 = 0\}$.

Lemma 20. *If there exists a strategy with a starting position Φ_s for Delayer in the game G^R (respectively, G^R_{sw}) that guarantees at least c coins on a set of linear clauses ϕ, then the size of a DT(R) (respectively DT$_{sw}$(R)) tree for ϕ, with the system Φ_s in the root, must be at least 2^c.*

Proof: Assume that T is a tree of type DT(R) (respectively, DT$_{sw}$(R)) for ϕ. We define an embedding of the full binary tree B_c of height c to T inductively as follows. We simulate Prover in the game G^R (respectively, G^R_{sw}) by choosing branchings from T and following to a subtree chosen by the Delayer until Delayer decides to earn a coin and leaves the choice to the Prover or until the game finishes. In case we are at a position where Delayer earns a coin, and which corresponds to a vertex u in T, we map the root of B_c to u and proceed inductively by embedding two trees B_{c-1} to the left and right subtrees of u, corresponding to two choices of the Prover. $\square$

5.2 Dag-Like Lower Bounds for the Subset Sum Principle with Large Coefficients

One straightforward way to refute $\mathsf{SubSum}(f)$ in (either dag- or tree-like) Res(lin$_R$) is this: first use the Boolean axioms to derive $\mathsf{Im}(f)$, and then apply resolution with $f = 0$ (see, for example, Proposition 9). In this section we prove (Lemma 21) that if $\mathbb{F}$ is a field, then this is essentially *the only way to refute* $\mathsf{SubSum}(f)$. As a corollary, this establishes (Theorem 22) an exponential lower bound for *dag-like* Res(lin$_\mathbb{F}$) refutations of $\mathsf{SubSum}(f)$, for every f with exponentially large $|\mathsf{Im}(f)|$.

Note that for $|\mathsf{Im}(f)|$ to be exponentially large, the values of the coefficients in f must also be exponentially large. In the next section we will prove an exponential lower bound for tree-like Res(lin$_\mathbb{F}$) derivations of $\mathsf{Im}(f)$ for an f with small coefficients, which by Lemma 21 implies exponential lower bounds on *tree-like* Res(lin$_\mathbb{F}$) refutations of $\mathsf{SubSum}(f)$.

Lemma 21 (Normal form transformation)**.** *Let $f = 0$ be a single unsatisfiable linear equation over some field $\mathbb{F}$. Then, every* Res(lin$_\mathbb{F}$) *(resp. tree-like* Res(lin$_\mathbb{F}$)*) refutation of $f = 0$ can be transformed into the following derivation with the same size, up to a linear* *in the size of f factor: a* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ *(resp. tree-like* $\mathrm{Res}(\mathrm{lin}_R)$*) derivation of a weakening of* $\mathsf{Im}(f)$ *from the Boolean axioms followed by a sequence of applications of the resolution rule with $f = 0$ and the simplification rule.*

Proof: Let $\pi = (D_1, \dots, D_N)$ be a shortest $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutation of $f = 0$ and let n be the number of variables in f and A the size of the largest coefficient in f (where *size* here is the bit-size of the binary representation of A). We construct a derivation $\widehat{\pi}$ of $\mathsf{Im}(f)$ of size $O(A \cdot n \cdot S^2(\pi))$, followed by a sequence of applications of the resolution rule with $f = 0$ that eliminate all the disjuncts in $\mathsf{Im}(f)$ (so that $\widehat{\pi}$ combined with the eliminations of the disjuncts in $\mathsf{Im}(f)$ forms the final refutation). The derivation $\widehat{\pi}$ is achieved by eliminating all applications resolution with $f = 0$ from π.

Formally, we proceed by induction on k to prove the following:

Induction statement: If $\pi_{\le k} := (D_1, \dots, D_k)$ is the sequence of first k proof-lines in π, then there exists a $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ derivation $\widehat{\pi}_k = (\widehat{D}_1, \dots, \widehat{D}_l)$, for some $l \le k$, such that:

1. $\widehat{\pi}_k$ contains no application of the resolution rule with $f = 0$;

2. there is a (total) injective map $\tau : [l] \to [k]$ such that if $D_{\tau(i)}$ is $\bigvee_{t \in [m]} g_t = 0$, for $i \in [l]$, then

$$\widehat{D}_i = \left(\bigvee_{t \in [m]} g_t + a_t f = 0 \vee \bigvee_{t \in [s]} f + b_t = 0 \right),$$

   for some $a_1, \dots, a_m \in \mathbb{F}$ and $b_1, \dots, b_s \in \mathbb{F}^*$. In other words, $\widehat{D}_i$ can be viewed as a weakening of $D_{\tau(i)}$ with equations of the form $f - b_t = 0$, and with $a_t f$ added to all the linear equations in $D_{\tau(i)}$.

We assume without loss of generality that π does not contain applications of the weakening rule and whenever the simplification rule is used to derive D from $D \vee a = 0, a \in \mathbb{F}^*$, in π, everywhere further in π the clause $D \vee a = 0$ is never used as a premise, rather the simplified clause D is used instead.

Before proving the induction statement above, we now argue that this statement concludes the proof of the lemma. We need the following simple claim, which is evident by a simple inspection of the inductive construction of $\widehat{\pi}_{\le k}$ below:

Claim. *Every D_i in π has a corresponding clause according to τ in $\widehat{\pi}$, apart from those clauses in π whose all predecessors in the proof are (possibly a weakening of) $f = 0$.*[4]

Suppose that the number of lines in π is r. Since π is a refutation, the last linear clause D_r in π is the empty clause. Since, the empty clause is *not* semantically implied by $f = 0$ (over $\mathbb{F}$-elements)[5], it must be that the empty clause has a corresponding source according to τ. Hence, the last linear clause in $\widehat{\pi}_{\le r}$ is a disjunction of equations of the form $a_t f = 0$ or $f + b_t = 0$, for scalars a_t, b_t. Since $\widehat{\pi}_{\le r}$ is a legitimate $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ derivation,

[4] A weakening of $f = 0$ is $af = 0$, for $a \in \mathbb{F}$.

[5] Note that we must use both the Boolean axioms and $f = 0$ to semantically imply False.

where the only axioms used are the Boolean ones, by soundness of $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ it must hold that this last clause in $\widehat{\pi}_{\leq r}$ is a *tautology* (i.e., always holds over 0-1 assignments), which means that it must be a weakening of $\mathsf{SubSum}(f)$.

Base case: If $D_1 \neq (f=0)$, then let $\widehat{D}_1 = D_1$ and $\tau : [1] \to [1]$ be the identity. Otherwise let $\widehat{\pi}_1$ be empty.

Induction step: Assume $1 \leq k < N$, and assume that $\pi_{\leq k} = (D_1, \dots, D_k)$, $\widehat{\pi}_k = (\widehat{D}_1, \dots, \widehat{D}_l), l \leq k$ and $\tau : [l] \to [k]$ satisfy the conditions above. Consider the possible cases in which D_{k+1} is derived:

Case 1: Axiom $D_{k+1} = (f=0)$. Let $\widehat{\pi}_{k+1} := \widehat{\pi}_k$.

Case 2: Boolean axiom $D_{k+1} = (x = 0 \vee x = 1)$. Let $\widehat{\pi}_{k+1} := (\widehat{D}_1, \dots, \widehat{D}_l, D_{k+1})$ and $\tau(l+1) := k+1$.

Case 3: Resolution of $D_i = (\bigvee_{t\in[m]} g_t = 0 \vee h = 0)$, $i \leq k$, with $D_j = (\bigvee_{t\in[m']} g'_t = 0 \vee h' = 0)$, $j \leq k$, yielding:

$$D_{k+1} = \left(\bigvee_{t\in[m]} g_t = 0 \vee \bigvee_{t\in[m']} g'_t = 0 \vee \alpha h + \beta h' = 0 \right), \ \alpha, \beta \in \mathbb{F}\,.$$

If i, j are both not in the image of τ, then let $\widehat{\pi}_{k+1} := \widehat{\pi}_k$.

If exactly one of i, j is in the image of τ, then assume without loss of generality that i is in the image of τ. It must hold that $D_j = (h' = 0) = (f = 0)$, and we let $\widehat{D}_{l+1} = \widehat{D}_{\tau^{-1}(i)}$ and $\tau(l+1) := k+1$, where $\widehat{D}_{\tau^{-1}(i)} = (\bigvee_{t\in[m]} g_t + a_t f = 0 \vee \bigvee_{t\in[s]} f + b_t = 0 \vee (\alpha h + \beta h') - \beta f = 0)$.

If i, j both are in the image of τ, then we have $\widehat{D}_{\tau^{-1}(i)} = (\bigvee_{t\in[m]} g_t + a_t f = 0 \vee h + a_{m+1} f = 0 \vee \bigvee_{t\in[s]} f + b_t = 0)$ and $\widehat{D}_{\tau^{-1}(j)} = (\bigvee_{t\in[m']} g'_t + a'_t f = 0 \vee h' + a'_{m'+1} f = 0 \vee \bigvee_{t\in[s']} f + b'_t = 0)$. Let

$$\widehat{D}_{l+1} := \Big(\bigvee_{t\in[m]} g_t + a_t f = 0 \vee \bigvee_{t\in[m']} g'_t + a'_t f = 0 \vee \bigvee_{t\in[s]} f + b_t = 0 \\ \vee \bigvee_{t\in[s']} f + b'_t = 0 \vee \alpha h + \beta h' + (\alpha a_{m+1} + \beta a'_{m'+1}) f = 0 \Big)$$

and $\tau(l+1) := k+1$.

Case 4: The clause $D_{k+1} = (\bigvee_{t\in[m]} g_t = 0)$ is the result of a simplification of $D_i = (\bigvee_{t\in[m]} g_t = 0 \vee a = 0)$, for $a \in \mathbb{F}^*$ and $i \in [k]$. It must hold that i is in the image of τ and if $\widehat{D}_{\tau^{-1}(i)} = (\bigvee_{t\in[m]} g_t + a_t f = 0 \vee \bigvee_{t\in[s]} f + b_t = 0 \vee a + a_{m+1} f = 0)$ we define $\widehat{D}_{l+1} := \widehat{D}_{\tau^{-1}(i)}$ and $\tau(l+1) := k+1$. $\square$

As a corollary we obtain the following dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ lower bound:

Theorem 22. *Let $\mathbb{F}$ be a field. The size of the shortest (dag-like) $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutation of $\mathsf{SubSum}(f)$ is lower bounded by $|\mathsf{Im}(f)|$. In particular, if char$(\mathbb{F}) = 0$ the shortest $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutation of $\mathsf{SubSum}(x_1 + 2x_2 + \cdots + 2^n x_n + 1)$ is of size $2^{\Omega(n)}$.*

5.3 Tree-Like Lower Bounds for the Subset Sum with Small Coefficients

We now turn to tree-like lower bounds. Note that Theorem 22 only gives a lower bound for $\mathsf{SubSum}(f)$ if the coefficients of f are large enough. In what follows (Theorem 24) we prove that this lower bound holds for tree-like $\mathrm{Res}(\mathrm{lin}_R)$ even for small coefficients.

Lemma 23. *Let Φ be a satisfiable system of m non-equalities over a field $\mathbb{F}$ of characteristic 0. If $\Phi \models \epsilon_1 x_1 + \cdots + \epsilon_n x_n = A$ for some $\epsilon_i \in \{-1, 1\} \subset \mathbb{F}, A \in \mathbb{F}$, then $m \geq \frac{n}{4}$.*

Note that A must be an integer (inside $\mathbb{F}$), since the coefficients of variables are all $-1, 1$, and the variables themselves are Boolean (since $\models$ stands for semantic implication over 0-1 assignments only).

Proof: Let $\Phi = \{\overline{a}_1 \cdot \overline{x} + b_1 \neq 0, \dots, \overline{a}_m \cdot \overline{x} + b_m \neq 0\}$ and put $\sigma = A \mod 2$, $f = \epsilon_1 x_1 + \cdots + \epsilon_n x_n$. Then

$$\begin{aligned} f \equiv 1 - \sigma \pmod 2 &\models f \neq A \\ &\models (\overline{a}_1 \cdot \overline{x} + b_1) \cdot \ldots \cdot (\overline{a}_m \cdot \overline{x} + b_m) = 0. \end{aligned}$$

By Theorem 4.4 in Alekhnovich-Razborov [AR01], the function $f \equiv 1 - \sigma \pmod 2$ is $\frac{n}{4}$-*immune*, that is, the degree of any non-zero polynomial g such that $f \equiv 1 - \sigma \pmod 2 \models g = 0$ must be at least $\frac{n}{4}$. Therefore $m \geq \frac{n}{4}$. □

Theorem 24. *Let $f = \epsilon_1 x_1 + \cdots + \epsilon_n x_n$, where $\epsilon_i \in \{-1, 1\} \subset \mathbb{F}$, and $\mathbb{F}$ is field of char$(\mathbb{F}) = 0$. Then the following holds:*

1. *Any tree-like $\mathrm{Res}(\mathrm{lin}_\mathbb{F})$ derivation of the clause $\mathsf{Im}(f)$ is of size at least $2^{\frac{n}{4}}$.*

2. *Any tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_\mathbb{F})$ refutation of $\mathsf{ImAv}(f)$ is of size at least $2^{\frac{n}{4}}$.*

Proof: We use Prover-Delayer games to show the lower bounds. By the definition of the games (Sec. 5.1), in the former case the game $G^\mathbb{F}$ is on $\{x_i = 0 \vee x_i = 1\}_{i \in [n]}$ and starts with the position

$$\Phi_r = \{f - A \neq 0 \mid A \in im_2(f)\},$$

and in the latter case $G^\mathbb{F}_{sw}$ is on $\mathsf{ImAv}(f)$ and starts with the empty position $\Phi_r = \emptyset$.

The former game $G^\mathbb{F}$ finishes at a position Φ, where $\{x_i \neq 0, \; x_i \neq 1\} \subseteq \Phi$ for some $i \in [n]$ or $0 \neq 0 \in \Phi$. The latter game $G^\mathbb{F}_{sw}$ finishes at a position Φ, where $\Phi \models f = A, \; A \notin im_2(f)$.

We show that the following Delayer strategy guarantees $\frac{n}{4}$ coins for both games. This, together with Lemma 20, implies the lower bounds.

The strategy is as follows: let the position in the game be defined by a system Φ and let the branching chosen by the Prover be $g_1 \neq 0$ and $g_2 \neq 0$. Consider $\Phi' = \Phi \setminus \Phi_r$. Thus, Delayer does the following:

1. if $\Phi' \models g_2 = 0$, but $\Phi' \not\models g_1 = 0$, then choose $g_1 \neq 0$;

2. if $\Phi' \models g_1 = 0$, but $\Phi' \not\models g_2 = 0$ then choose $g_2 \neq 0$;

3. if none of the above holds, or both $\Phi' \models g_2 = 0$ and $\Phi' \models g_1 = 0$ hold, then leave the choice to the Prover and earn a coin.

We now prove that this strategy guarantees the required number of coins in both games.

Consider the game $G^{\mathbb{F}}$. Suppose the game has finished at a position Φ. Let Φ_u be a position, where $\Phi'_u = \Phi_u \setminus \Phi_r$ *first* became unsatisfiable (over 0-1, that is). Such a position must exist, by the definition of a final state in the game. That is, if $\{x_i \neq 0,\ x_i \neq 1\} \subseteq \Phi$ for some $i \in [n]$ or $0 \neq 0 \in \Phi$, then $\{x_i \neq 0, x_i \neq 1\} \subseteq \Phi'$ or $0 \neq 0 \in \Phi'$, respectively.

Let $\Phi_{p(u)}$ be the position preceding immediately position Φ_u, and let Prover present the branching $g_1 \neq 0$ and $g_2 \neq 0$ to Delayer in position $\Phi_{p(u)}$, for some $g_1 + g_2 \neq 0 \in \Phi_{p(u)}$.

Claim. *Both* $\Phi'_{p(u)} \models g_1 = 0$ *and* $\Phi'_{p(u)} \models g_2 = 0$.

Proof of claim: $\Phi'_u = \Phi_u \setminus \Phi_r$ (and thus also Φ_u), is unsatisfiable by assumption. Suppose by a way of contradiction that $\Phi'_{p(u)} \not\models g_1 = 0$ or $\Phi'_{p(u)} \not\models g_2 = 0$.
Case 1: If $\Phi'_{p(u)} \not\models g_1 = 0$ and $\Phi'_{p(u)} \models g_2 = 0$, then by the strategy assumed above, Delayer chooses to branch on $g_1 \neq 0$. By definition of the game, $g_1 \neq 0$ is now added to $\Phi_{p(u)}$ and thus $\Phi_u = \Phi_{p(u)} \cup \{g_1 \neq 0\}$. But $\Phi'_u \subseteq \Phi'_{p(u)} \cup \{g_1 \neq 0\}$ is satisfiable in contrast to our assumption.
Case 2: If $\Phi'_{p(u)} \not\models g_2 = 0$ and $\Phi'_{p(u)} \models g_1 = 0$, then this is similar to Case 1.
Case 3: If *both* $\Phi'_{p(u)} \not\models g_1 = 0$ and $\Phi'_{p(u)} \not\models g_2 = 0$, then this is similar to the previous cases. $\blacksquare_{\text{Claim}}$

By the claim, $\Phi'_{p(u)} \models g_1 + g_2 = 0$. We know by assumption on position Φ_u that $\Phi'_{p(u)}$ is satisfiable and $\Phi'_{p(u)} \cup \{g_1 + g_2 \neq 0\}$ is unsatisfiable. Therefore, $g_1 + g_2 \neq 0 \notin \Phi'_{p(u)}$ is not a tautology over 0-1 assignments and this excludes the option that $g_1 + g_2$ is a non-zero constant. Recall that $g_1 + g_2 \neq 0$ is the non-equality picked by Prover to branch on when in state $\Phi_{p(u)}$. As $g_1 + g_2$ is non-constant, this means that $g_1 + g_2 \neq 0 \in \Phi_{p(u)} = \Phi'_{p(u)} \cup \Phi_r$. But since $g_1 + g_2 \neq 0 \notin \Phi'_{p(u)}$, we have $g_1 + g_2 \neq 0 \in \Phi_r$, which means that $g_1 + g_2 \equiv f - A,\ A \in im_2(f)$.

Let $\zeta_1, \ldots, \zeta_\ell$ be the set of non-equalities in $\Phi'_{p(u)}$, in the order they were added to $\Phi'_{p(u)}$. Let $\Psi'_{p(u)} \subseteq \Phi'_{p(u)}$ be the set of all ζ_i, $i \in [\ell]$, such that ζ_i is not implied by previous non-equalities ζ_j, for $j < i$. Note that at any position Φ if Case 1 or Case 2 of the Delayer's strategy hold, then the non-equality $g \neq 0$ chosen by Delayer satisfies $\Phi' \models g \neq 0$. Therefore the number of coins earned by Delayer at $\Phi_{p(u)}$ is at least $|\Psi'_{p(u)}|$ and $\Psi'_{p(u)} \models f = A$, by the previous paragraph. Lemma 23 implies that $|\Psi'_{p(u)}| \geq \frac{n}{4}$.

Consider the game $G^{\mathbb{F}}_{sw}$. This is similar to the argument for $G^{\mathbb{F}}$. Suppose that the game has finished at a position Φ. Thus, Φ must be satisfiable and contradict a clause $\langle f \neq A \rangle$ of $\mathsf{ImAv}(f)$. Therefore, $\Phi \models f = A$ for some $A \in im_2(f)$. Denote by $\Psi \subseteq \Phi$ the subsystem of non-equalities that are not implied by previous ones (similar to the argument for the game $G^{\mathbb{F}}$ above). Then, Delayer earns at least $|\Psi|$ coins, $\Psi \models f = A$, and by Lemma 23 we conclude that $|\Psi| \geq \frac{n}{4}$. □

Corollary 25. *Let f and $\mathbb{F}$ be as in Proposition 24. Then the shortest tree-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ *refutation of $f = n+1$ is of size at least $2^{\frac{n}{4}}$.*

Proof: Follows from Lemma 21 and Theorem 24. □

5.4 Lower Bounds for the Pigeonhole Principle

Here we prove that every tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ refutations of $\neg\mathrm{PHP}^m_n$ must have size at least $2^{\Omega(\frac{n-1}{2})}$ (see Sec. 2.3.1 for the definition of $\neg\mathrm{PHP}^m_n$). Together with the upper bound for dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ (see Sec. 3.2) this provides a separation between tree-like and dag-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ in the case $char(\mathbb{F}) = 0$. The lower bound argument is comprised of exhibiting a strategy for Delayer in the Prover-Delayer game. Delayer's strategy is similar to that in [IS14]. However, the proof that Delayer's strategy guarantees sufficiently many coins relies on Lemma 27, which is a generalization of Lemma 3.3 in [IS14] for arbitrary fields. Since the proof of Lemma 3.3 in [IS14] for the $\mathbb{F}_2$ case does not apply to arbitrary fields, our proof is different, and uses a result from Alon-Füredi [AF93] on the hyperplane coverings of the hypercube.

Theorem 26. *For every field $\mathbb{F}$, the shortest tree-like* $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ *refutation of $\neg PHP^m_n$ has size $2^{\Omega(\frac{n-1}{2})}$.*

Proof: We prove that there exists a strategy for Delayer in the $\neg\mathrm{PHP}^m_n$ game, which guarantees Delayer to earn $\frac{n-1}{2}$ coins. Following the terminology in [IS14], we call an assignment $x_{i,j} \mapsto \alpha_{ij}$, for $\alpha \in \{0,1\}^{mn}$, *proper* if it does not violate $\mathsf{Pigeons}^m_n$, namely, if it does not send two distinct pigeons to the same hole. We need to prove several lemmas before concluding the theorem.

Lemma 27. *Let $A\overline{x} \doteqdot \overline{b}$ be a system of k linear non-equalities over a field $\mathbb{F}$ with n variables and where $\overline{x} = 0$ is a solution, that is, $0 \doteqdot \overline{b}$. If $k < n$, then there exists a non-zero boolean solution to this system.*

Proof: Let $\overline{a}_1, \ldots, \overline{a}_k$ be the rows of the matrix A. The boolean solutions of the system $A\overline{x} \doteqdot \overline{b}$ are all the points of the n-dimensional boolean hypercube $B_n := \{0,1\}^n \subset \mathbb{F}^n$, that are not covered by the hyperplanes $H := \{\overline{a}_1\overline{x} - b_1 = 0, \ldots, \overline{a}_k\overline{x} - b_k = 0\}$. We need to show that if $k < n$ and $0 \in B_n$ is not covered by H, then some other point in B_n is not covered by H as well. This follows from a direct corollary from [AF93]:

Corollary from Alon-Füredi [AF93, Theorem 4]. *Let $Y(l) := \{(y_1, \ldots, y_n) \in \mathbb{F}^n \mid \forall i \in [n], 0 < y_i \leq 2,$ and $\sum_i^n y_i \geq l\}$. For any field $\mathbb{F}$, if k hyperplanes in $\mathbb{F}^n$ do not cover B_n completely, then they do not cover at least $M(2n-k)$ points, where*

$$M(l) := \min_{(y_1,\ldots,y_n)\in Y(l)} \prod_{1\leq i\leq n} y_i \,.$$

Thus, if $k < n$ hyperplanes do not cover B_n completely, then they do not cover at least $M(n+1)$ points. The set $Y(n+1)$ in the Corollary above consists of all tuples $(y_1, \ldots, y_n)$, where $y_i = 2$ for some $i \in [n]$ and $y_j = 1$ for $j \in [n], j \neq i$. Therefore $M(n+1) = 2$. $\square$

For two Boolean assignments $\alpha, \beta \in \{0,1\}^n$, denote by $\alpha \oplus \beta$ the bitwise XOR of the two assignments.

Lemma 28. *Let $A\overline{x} \doteqdot \overline{b}$ be a system of k linear non-equalities over a field $\mathbb{F}$ with $n > k$ variables and let $\alpha \in \{0,1\}^n$ be a solution to the system. Then, for every choice I of $k+1$ bits in α, there exists at least one $i \in I$ so that flipping the ith bit in α results in a new solution to $A\overline{x} \doteqdot \overline{b}$. In other words, if $I \subseteq [n]$ is such that $|I| = k+1$, then there exists a boolean assignment $\beta \neq 0$ such that $\{i \mid \beta_i = 1\} \subseteq I$ and $A(\alpha \oplus \beta) \doteqdot \overline{b}$.*

Proof: Let $I \subseteq \{0,1\}^n$. Denote by $A_I^\star$ the matrix with columns $\{(1-2\alpha_i)\overline{a}_i \mid i \in I\}$, where $\overline{a}_i$ is the ith *column* of A. That is, $A_I^\star$ is the matrix A restricted to columns i with $i \in I$ and where column i flips its sign iff α_i is 1.

Assume that $\beta \in \{0,1\}^n$ is nonzero and all its 1's must appear in the indices in I, that is, $\{i \mid \beta_i = 1\} \subseteq I$. Given a set of indices $J \subseteq [n]$, denote by β_J the restriction of β to the indices in J. Similarly, for a vector $v \in \mathbb{F}^n$, v_J denotes the restriction of v to the indices in J.

Claim. $A(\alpha \oplus \beta) \doteqdot \overline{b}$ *iff* $A_I^\star \beta_I \doteqdot \overline{b} - A\alpha$.

Proof of claim: We prove that $A(\alpha \oplus \beta) = A_I^\star \beta_I + A\alpha$. Consider any row $\mathbf{v}$ in A, and the corresponding row $\mathbf{v}_I^\star$ in $A_I^\star$. Notice that $\mathbf{v} \cdot (\alpha \oplus \beta)$ (for "$\cdot$" the dot product) equals the dot product of $\mathbf{v}$ and $\alpha \oplus \beta$, where both vectors are restricted only to those entries in which α and β differ. Considering entries outside I, by assumption we have $\beta_{[n]\setminus I} = 0$, which implies that

$$\mathbf{v}_{[n]\setminus I} \cdot (\alpha \oplus \beta)_{[n]\setminus I} = \mathbf{v}_{[n]\setminus I} \cdot \alpha_{[n]\setminus I} \,. \tag{7}$$

On the other hand, considering entries inside I, we have

$$\mathbf{v}_I \cdot (\alpha \oplus \beta)_I = \mathbf{v}_I \cdot \alpha_I + \mathbf{v}_I^\star \cdot \beta_I \,. \tag{8}$$

Equation (8) can be verified by inspecting all four cases for the ith bits in α, β, for $i \in I$, as follows: for those indices $i \in I$, such that $\alpha_i = 1$ and $\beta_i = 0$, only $\mathbf{v}_I \cdot \alpha$ contributes to the right hand side in (8). If $\alpha_i = 1$ and $\beta_i = 1$, then by the definition of $A_I^\star$, the two summands in the right hand side in (8) cancel out. The cases $\alpha_i = 0, \beta_i = 1$ and $\alpha_i = \beta_i = 0$, can also be inspected to contribute the same values to both sides of (8).

The two equations (7) and (8) concludes the claim. $\blacksquare_{\text{Claim}}$

We know that $A\alpha \doteqdot \overline{b}$, and we wish to show that for some nonzero $\beta \in \{0,1\}^n$ where $\{i \mid \beta_i = 1\} \subseteq I$, it holds that $A(\alpha \oplus \beta) \doteqdot \overline{b}$. By the claim above it remains to show the existence of such β where $A_I^\star \beta_I \doteqdot \overline{b} - A\alpha$. But notice that $\overline{b} - A\alpha \doteqdot 0$, since $A\alpha \doteqdot \overline{b}$, and that $A_I^\star \beta_I$ is a matrix of dimension $k \times (k+1)$. Therefore, by Lemma 20, the system

$A_I^\star \beta_I \doteqdot \overline{b} - A\alpha$ has a nonzero solution, that is, there exists a $\beta \neq 0$ for which all ones are in the I entries, such that $A_I^\star \beta_I \doteqdot \overline{b} - A\alpha$. $\square$

Lemma 29. *Assume that a system $A\overline{x} \doteqdot \overline{b}$ of $k \leq \frac{n-1}{2}$ non-equalities over $\mathbb{F}$ with variables $\{x_{i,j}\}_{(i,j)\in[m]\times[n]}$ has a proper solution. Then, for every $i \in [m]$ there exists a proper solution to the system, that satisfies the clause $\bigvee_{j\in[n]} x_{i,j}$. In other words, for every pigeon, there exists a proper solution that sends the pigeon to some hole.*

Proof: We first show that if there exists a proper solution of $A\overline{x} \doteqdot \overline{b}$, then there exists a proper solution of this system with at most k ones. Let α be a proper solution with at least $k+1$ ones. If I is a subset of $k+1$ ones in α, then Lemma 28 assures us that some other proper solution can be obtained from α by flipping some of these ones (note that flipping one to zero preserves the properness of assignments). Thus the number of ones can always be reduced until it is at most k.

Let α be a proper solution with at most k ones. The condition $k \leq \frac{n-1}{2}$ implies that there are $n-k \geq k+1$ free holes. Let J be a subset of size $k+1$ of the set of indices of free holes. Then for any $i \in [m]$ some of the bits in $I = \{(i,j) \mid j \in J\}$ can be flipped and still satisfy $A\overline{x} \doteqdot \overline{b}$, by Lemma 28. (As before, flipping from one to zero maintains the properness of the solution.) Hence, the resulting proper solution must satisfy the clause $\bigvee_{j\in[n]} x_{i,j}$. $\square$

We now describe the desired strategy for Delayer.

Delayer's Strategy: Let a position in the game be defined by the system of non-equalities Φ and assume that the branching chosen by Prover is $f_0 \neq 0$ or $f_1 \neq 0$, where $\Phi \models f_0 + f_1 \neq 0$. The only objective of Delayer is to ensure that the system Φ has proper solutions. Delayer uses the opportunity to earn a coin whenever both $\Phi \cup \{f_0 \neq 0\}$ and $\Phi \cup \{f_1 \neq 0\}$ have proper solutions by leaving the choice to Prover. Otherwise, in case $\Phi \wedge \mathsf{Pigeons}_n^m \models f_i = 0$, for some $i \in \{0,1\}$, Delayer chooses $f_{1-i} \neq 0$, which must satisfy $\Phi \wedge \mathsf{Pigeons}_n^m \models f_{1-i} \neq 0$, and so the sets of proper solutions of Φ and $\Phi \cup \{f_{1-i} \neq 0\}$ are identical.

This strategy ensures, that for every end-game position Φ, Φ has proper solutions and $\Phi \models \neg\mathsf{Holes}_n^m$. Note that Φ has the same proper solutions as Φ', obtained by throwing away from Φ all non-equalities that were added by Delayer when making a choice. Therefore, if $\Phi \models \neg\mathsf{Holes}_n^m$, then $\Phi' \wedge \mathsf{Pigeons}_n^m \models \neg\mathsf{Holes}_n^m$ and thus $|\Phi'| > \frac{n-1}{2}$ by Lemma 29.

Since $|\Phi'|$ is precisely the number of coins earned by Delayer, this gives the desired lower bound. $\square$

6 Size-Width Relation and Simulation by PC

In this section we prove a size-width relation for tree-like $\mathrm{Res}(\mathrm{lin}_R)$ (Theorem 32), which then implies an exponential lower bound on the size of tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ refutations in terms of the principal width of refutations (Definition 2). The connection between the principal width and the degree of PC refutations for finite fields $\mathbb{F}$, together with lower bounds on degree of PC refutations from [AR01] on Tseitin mod p formulas and random CNFs, imply exponential lower bounds for the size of tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ for these instances (Corollaries 34 and 35).

Proposition 30. *Let $\phi = \{C_i\}_{1\leq i\leq m}$ be a set of linear clauses and $x \in vars(\phi)$. Assume that l is a linear form in the variables $vars(\phi)\setminus\{x\}$. Then, there is a* $\mathrm{Res}(\mathrm{lin}_R)$ *derivation π of $\{C_i\restriction_{x\leftarrow l} \vee\langle x-l\neq 0\rangle\}_{1\leq i\leq m}$ from ϕ of size polynomial in $|\phi| + |\mathsf{Im}(l)|$ and such that $\omega_0(\pi) \leq \omega_0(\phi) + 2$.*

Proof: The clause $x-l = 0 \vee \langle x-l \neq 0\rangle$ is derivable in $\mathrm{Res}(\mathrm{lin}_R)$ in polynomial in $|\mathsf{Im}(l)|$ size by Proposition 8. Assume

$$C = \left(\bigvee_{j\in[k]} f_j + a_j x + b_j^{(1)} = 0 \vee \cdots \vee f_j + a_j x + b_j^{(N_j)} = 0\right),$$

where $x \notin vars(f_i)$ and we have grouped disjuncts so that $\omega_0(C) = k$. Then we resolve these groups one by one with $x-l = 0 \vee \langle x-l\neq 0\rangle$ and after $N_1 + \ldots + N_k$ steps yield $\left(\bigvee_{j\in[k]} f_j + a_j l + b_j^{(1)} = 0 \vee \cdots \vee f_j + a_j l + b_j^{(N_j)} = 0 \vee \langle x-l\neq 0\rangle\right)$. It is easy to see that the principal width never exceeds $k+2$ along the way. Therefore $\omega_0(\pi) \leq \omega_0(\phi)+2$. $\square$

Corollary 31. *Let $\phi = \{C_i\}_{1\leq i\leq m}$ be a set of linear clauses and $x \in vars(\phi)$. Suppose that l is a linear form with variables $vars(\phi)\setminus\{x\}$ and that π is a* $\mathrm{Res}(\mathrm{lin}_R)$ *refutation of $\phi\restriction_{x\leftarrow l} \cup\{l = 0 \vee l = 1\}$. Then, there exists a* $\mathrm{Res}(\mathrm{lin}_R)$ *derivation $\widehat{\pi}$ of $\langle x-l\neq 0\rangle$ from ϕ, such that $S(\widehat{\pi}) = O(S(\pi)+|\mathsf{Im}(l)|)$ and $\omega_0(\widehat{\pi}) \leq \max(\omega_0(\pi)+1, \omega_0(\phi)+2)$. Additionally, there is a refutation $\widehat{\pi}'$ of $\phi\cup\{x-l=0\}$ where $\omega_0(\widehat{\pi}') \leq \max(\omega_0(\pi), \omega_0(\phi)+2)$.*

Proof: By Proposition 30 there exists a derivation π_s of

$$\{C_i\restriction_{x\leftarrow l} \vee\langle x-l\neq 0\rangle\}_{1\leq i\leq m} \cup \{l = 0 \vee l = 1 \vee \langle x-l\neq 0\rangle\}$$

from ϕ of width at most $\omega_0(\phi)+2$. Composing π_s with $\pi\vee\langle x-l\neq 0\rangle$ yields the derivation $\widehat{\pi}$ of $\langle x-l\neq 0\rangle$ from ϕ.

Moreover, by taking the derivation π_s and adding to it the axiom $x - l = 0$, and then using a sequence of resolutions of π_s with $x-l = 0$, we obtain a derivation of $\phi\restriction_{x\leftarrow l} \cup\{l = 0\vee l = 1\}$ from $\phi\cup\{x-l=0\}$. The latter derivation composed with π yields the refutation $\widehat{\pi}'$ of $\phi\cup\{x-l=0\}$ of width at most $\max(\omega_0(\pi), \omega_0(\phi)+2)$. $\square$

Theorem 32. *Let ϕ be an unsatisfiable set of linear clauses over a field $\mathbb{F}$. The following size-width relation holds for both tree-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ *and tree-like* $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$*:*

$$S(\phi\vdash\perp) = 2^{\Omega(\omega_0(\phi\vdash\perp)-\omega_0(\phi))}.$$

Proof: We prove by induction on n, the number of variables in ϕ, the following:

$$\omega_0(\phi\vdash\perp) \leq \lceil\log_2 S(\phi\vdash\perp)\rceil + \omega_0(\phi) + 2\,.$$

Base case: $n = 0$. Thus ϕ must contain only linear clauses $a = 0$, for $a \in \mathbb{F}$, and the principal width for refuting ϕ is therefore 1.

Induction step: Let π be a tree-like refutation of $\phi = \{C_1, \ldots, C_m\}$ such that $S(\pi) = S(\phi \vdash \perp)$ (i.e., π is of minimal size). Without loss of generality, we assume that the resolution rule in π is only applied to simplified clauses, that is clauses not containing disjuncts $1 = 0$ in case of tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ and not containing unsatisfiable $f = 0$, $0 \notin im_2(f)$ in case of tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$. The former can be eliminated by the simplification rule and the latter by the semantic weakening rule. By this assumption, the empty clause at the root of π is derived in tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ (resp. tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$) as a simplification (resp. weakening) of an unsatisfiable $h = 0$ ($1 = 0$ in case of tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$) equation, which is derived by application of the resolution rule. Denote the left and right subtrees, corresponding to the premises of $h = 0$, by π_1 and π_2, respectively.

The roots of π_1 and π_2 must be of the form $f_1 = 0$ and $f_2 = 0$, respectively, where $f_1 - f_2 = h$. Therefore,

$$f_1 = l(x_1, \ldots, x_{n-1}) + a_n x_n \quad \text{and} \quad f_2 = l(x_1, \ldots, x_{n-1}) + a_n x_n - h\,,$$

for some $l(x_1, \ldots, x_{n-1}) = \sum_{i=1}^{n-1} a_i x_i + B$, where $a_i, B \in \mathbb{F}$.

Assume without loss of generality that $a_n \neq 0$ and $S(\pi_1) \leq S(\pi_2)$. We now use the induction hypothesis to construct a narrow derivation $\pi_1^\bullet$ of $f_1 = 0$ such that

$$\begin{aligned} \omega_0(\pi_1^\bullet) &\leq \lceil \log_2 S(\pi_1) \rceil + 1 + \omega_0(\phi) + 2 \\ &\leq \lceil \log_2 S(\pi) \rceil + \omega_0(\phi) + 2\,. \end{aligned}$$

For every nonzero $A \in im_2(f_1)$ define the partial linear substitution ρ_A as $x_n \leftarrow (A - l(x_1, \ldots, x_{n-1}))a_n^{-1}$. Thus, $f_1 \restriction \rho_A = A$. The set of linear clauses

$$\phi \restriction_{\rho_A} \cup \left\{ (A - l)a_n^{-1} = 0 \vee (A - l)a_n^{-1} = 1 \right\} \tag{9}$$

is unsatisfiable and has $n - 1$ variables, and is refuted by $\pi_1 \restriction_{\rho_A}$.

By induction hypothesis there exists a (narrow) refutation π_1^A of (9) with

$$\begin{aligned} \omega_0(\pi_1^A) &\leq \lceil \log_2 S(\pi_1 \restriction_{\rho_A}) \rceil + \omega_0(\phi) + 2 \\ &\leq \lceil \log_2 S(\pi_1) \rceil + \omega_0(\phi) + 2\,. \end{aligned}$$

By Corollary 31 there exists a derivation $\widehat{\pi}_1^A$ of $\langle l + a_n x_n \neq A \rangle$ from ϕ such that $\omega_0(\widehat{\pi}_1^A) \leq \max(\omega_0(\pi_1^A) + 1, \omega_0(\phi) + 2) \leq \lceil \log_2 S(\pi_1) \rceil + \omega_0(\phi) + 3$. By Proposition 10 there exists a derivation $\pi_1^\bullet$ of $f_1 = 0$ such that $\omega_0(\pi_1^\bullet) \leq \lceil \log_2 S(\pi_1) \rceil + \omega_0(\phi) + 3 \leq \lceil \log_2 S(\pi) \rceil + \omega_0(\phi) + 2$.

Consider the following substitution ρ: $x_n \leftarrow -l \cdot a_n^{-1}$. Then, $\pi_2|_\rho$ is a derivation of $h = 0$ from $\phi|_\rho \cup \{-l \cdot a_n^{-1} = 0 \vee -l \cdot a_n^{-1} = 1\}$, which we augment to refutation π_2'

by taking composition with simplification (resp. weakening) in case of tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ (resp. tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$). By induction hypothesis there exists a refutation $\pi_2^{\bullet}$ of width

$$\begin{aligned}\omega_0(\pi_2^{\bullet}) &\leq \lceil \log_2(S(\pi_2') + 1)\rceil + \omega_0(\phi) + 2 \\ &\leq \lceil \log_2 S(\pi)\rceil + \omega_0(\phi) + 2\,,\end{aligned}$$

and thus by Corollary 31 there exists a refutation $\widehat{\pi}_2^{\bullet}$ of $\phi \cup \{f_1 = 0\}$ of width $\omega_0(\widehat{\pi}_2^{\bullet}) \leq \lceil \log_2 S(\pi)\rceil + \omega_0(\phi) + 2$. The combination of $\widehat{\pi}_2^{\bullet}$ and $\pi_1^{\bullet}$ gives a refutation of ϕ of the desired width. □

Theorem 33. *Let $\mathbb{F}$ be a field and π be a $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutation of an unsatisfiable set of linear clauses ϕ. Then, there exists a $PC_{\mathbb{F}}$ refutation π' of (the arithmetization of) ϕ of degree $\omega(\pi)$.*

Proof: The idea is to replace every clause $C = (f_1 = 0 \vee \ldots \vee f_m = 0)$ in π by its arithmetization $a(C) := f_1 \cdot \ldots \cdot f_m$, and then augment this sequence to a valid $PC_{\mathbb{F}}$ derivation by simulating all the rule applications in π by several $PC_{\mathbb{F}}$ rule applications.
Case 1: If $D = (C \vee g_1 = 0 \vee \ldots \vee g_m = 0)$ is a weakening of C, then apply the product and the addition rules to derive $a(D) = a(C) \cdot g_1 \cdot \ldots \cdot g_m$ from $a(C)$.
Case 2: If D is a simplification of $D \vee 1 = 0$, then $a(D) = a(D \vee 1 = 0)$.
Case 3: If $D = (x = 0 \vee x = 1)$ is a a Boolean axiom, then $a(D) = x^2 - x$ is an axiom of $PC_{\mathbb{F}}$.
Case 4: If $D = (C \vee C' \vee E \vee \alpha f + \beta g = 0)$ is a result of resolution of $(C \vee E \vee f = 0)$ and $(C' \vee E \vee g = 0)$, where C and C' do not contain the same disjuncts, then by the product and addition rules of PC we derive $a(C) \cdot a(C') \cdot a(E) \cdot f$ from $a(C \vee E \vee f = 0) = a(C) \cdot a(E) \cdot f$, and also derive $a(C) \cdot a(C') \cdot a(E) \cdot g$ from $a(C' \vee E \vee f = 0) = a(C') \cdot a(E) \cdot f$, and then apply the addition rule to derive $a(C) \cdot a(C') \cdot a(E) \cdot (\alpha f + \beta g) = a(D)$.

It is easy to see that the degree of the resulting $PC_{\mathbb{F}}$ refutation is at most $\omega(\pi)$. □

As a consequence of Theorems 32 and 33, and the relation $\omega_0 \geq \frac{1}{|\mathbb{F}|}\omega$ as well as the results from [AR01], we have the following:

Corollary 34. *For every prime p there exists a constant $d_0 = d_0(p)$ such that the following holds. If $d \geq d_0$, G is a d-regular Ramanujan graph on n vertices (augmented with arbitrary orientation to its edges) and $\mathbb{F}$ is a finite field with $char(\mathbb{F}) \neq p$, then for every function σ such that $\neg TS^{(p)}_{G,\sigma} \in UNSAT$, every tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutation of $\neg TS^{(p)}_{G,\sigma}$ has size $2^{\Omega(dn)}$.*

Proof: Corollary 4.5 from [AR01] states that the degree of $PC_{\mathbb{F}}$ refutations of $\neg \mathrm{TS}^{(p)}_{G,\sigma}$ is $\Omega(dn)$. Theorem 33 implies that the principal width of $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutations of $\neg \mathrm{TS}^{(p)}_{G,\sigma}$ is $\Omega(\frac{1}{|\mathbb{F}|}dn) = \Omega(dn)$ and thus by Theorem 32 the size is $2^{\Omega(dn)}$. □

Corollary 35. *Let $\phi \sim \mathcal{F}^{n,\Delta}_k, k \geq 3$ and $\Delta = \Delta(n)$ be such that $\Delta = o(n^{\frac{k-2}{2}})$ and let $\mathbb{F}$ be any finite field. Then every tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutation of ϕ has size $2^{\Omega\left(\frac{n}{\Delta^{2/(k-2)} \cdot \log \Delta}\right)}$ with probability $1 - o(1)$.*

Proof: Corollary 4.7 from [AR01] states that the degree of $PC_{\mathbb{F}}$ refutations of $\phi \sim \mathcal{F}_k^{n,\Delta}$, where $k \geq 3$, is $\Omega(dn)$ with probability $1 - o(1)$. Theorem 33 implies that the principal width of Res($\text{lin}_{\mathbb{F}}$) refutations of $\phi \sim \mathcal{F}_k^{n,\Delta}$ is $\Omega(\frac{1}{|\mathbb{F}|}dn) = \Omega(dn)$ and thus by Theorem 32 the size of the refutations is $2^{\Omega(dn)}$ with probability $1 - o(1)$. □

Acknowledgments

We wish to thank Dima Itsykson and Dima Sokolov for very helpful comments concerning this work, and telling us about the lower bound on random k-CNF formulas for tree-like Res($\text{lin}_{\mathbb{F}_2}$) that can be achieved using the results of Garlik and Kołodziejczyk.

References

[AF93] Noga Alon and Zoltán Füredi. Covering the cube by affine hyperplanes. *Eur. J. Comb.*, 14(2):79–83, March 1993. 1.2, 5.4, 5.4

[Ajt88] Miklós Ajtai. The complexity of the pigeonhole principle. In *Proceedings of the IEEE 29th Annual Symposium on Foundations of Computer Science*, pages 346–355, 1988. 1.1

[AR01] Michael Alekhnovich and Alexander A. Razborov. Lower bounds for polynomial calculus: non-binomial case. In *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science (Las Vegas, NV, 2001)*, pages 190–199. IEEE Computer Soc., Los Alamitos, CA, 2001. (document), 1.2, 2.3.2, 2, 3, 5.3, 6, 6, 6, 6

[BGIP01] Samuel R. Buss, Dima Grigoriev, Russell Impagliazzo, and Toniann Pitassi. Linear gaps between degrees for the polynomial calculus modulo distinct primes. *J. Comput. System Sci.*, 62(2):267–289, 2001. Special issue on the 14th Annual IEEE Conference on Computational Complexity (Atlanta, GA, 1999). 2.3.2

[BKS04] Paul Beame, Henry A. Kautz, and Ashish Sabharwal. Towards understanding and harnessing the potential of clause learning. *J. Artif. Intell. Res.*, 22:319–351, 2004. 4

[BS02] Eli Ben-Sasson. Hard examples for the bounded depth Frege proof system. *Comput. Complexity*, 11(3-4):109–136, 2002. 1.1

[CEI96] Matthew Clegg, Jeffery Edmonds, and Russell Impagliazzo. Using the Groebner basis algorithm to find proofs of unsatisfiability. In *Proceedings of the 28th Annual ACM Symposium on the Theory of Computing (Philadelphia, PA, 1996)*, pages 174–183, New York, 1996. ACM. 1.2, 2.2

[CR74a] Stephen A. Cook and Robert A. Reckhow. Corrections for "On the lengths of proofs in the propositional calculus (preliminary version)". *SIGACT News*, 6(3):15–22, July 1974. 6

[CR74b] Stephen A. Cook and Robert A. Reckhow. On the lengths of proofs in the propositional calculus (preliminary version). In *Proceedings of the 6th Annual ACM Symposium on Theory of Computing (STOC 1974)*, pages 135–148, 1974. For corrections see Cook-Reckhow [CR74a]. 6

[CR79] Stephen A. Cook and Robert A. Reckhow. The relative efficiency of propositional proof systems. *J. Symb. Log.*, 44(1):36–50, 1979. This is a journal-version of Cook-Reckhow [CR74b] and Reckhow [Rec76]. 1.2, 1

[GK17] Michal Garlik and Lezsek Kołodziejczyk. Some subsystems of constant-depth Frege with parity. 2017. Unpublished manuscript. (document), 1.1, 1.2

[Hak85] Armin Haken. The intractability of resolution. *Theoret. Comput. Sci.*, 39(2-3):297–308, 1985. 1.1

[IS14] Dmitry Itsykson and Dmitry Sokolov. Lower bounds for splittings by linear combinations. In *Mathematical Foundations of Computer Science 2014 - 39th International Symposium, MFCS 2014, Budapest, Hungary, August 25-29, 2014. Proceedings, Part II*, pages 372–383, 2014. (document), 1.1, 1.2, 1.2, 1.2, 1.2, 2, 3, 3, 4, 4, 5.4, 5.4

[KO18] Jan Krajícek and Igor Carboni Oliveira. On monotone circuits with local oracles and clique lower bounds. *Chicago J. Theor. Comput. Sci.*, 2018, 2018. (document), 1.1

[KPW95] 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. 1.1

[Kra17] Jan Krajícek. A feasible interpolation for random resolution. *Logical Methods in Computer Science*, 13(1), 2017. (document), 1.1

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, Sep 1988. 1

[Nor15] Jakob Nordström. On the interplay between proof complexity and sat solving. *ACM SIGLOG News*, 2(3):19–44, August 2015. 1.1

[PBI93] Toniann Pitassi, Paul Beame, and Russell Impagliazzo. Exponential lower bounds for the pigeonhole principle. *Comput. Complexity*, 3(2):97–140, 1993. 1.1

[PI00] Pavel Pudlák and Russell Impagliazzo. A lower bound for DLL algorithms for k-sat. In *Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms, January 9-11, 2000, San Francisco, CA, USA.*, pages 128–136, 2000. 1.2, 5.1

[Rec76] Robert A. Reckhow. *On the lengths of proofs in the propositional calculus*. PhD thesis, University of Toronto, 1976. 6

[RT08] Ran Raz and Iddo Tzameret. Resolution over linear equations and multilinear proofs. *Ann. Pure Appl. Logic*, 155(3):194–224, 2008. (document), 1.1, 1.2, 1.2, 1.2, 2.3.2, 3, 15, 3.2

[Tse68] Grigori Tseitin. *On the complexity of derivations in propositional calculus*. Studies in constructive mathematics and mathematical logic Part II. Consultants Bureau, New-York-London, 1968. 1.1, 2.3.2

— Page left blank for ECCC stamp —

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il