Resolution with Counting: Dag-Like Lower Bounds and Different Moduli

Fedor Part

JetBrains Research, St. Petersburg, Russia
fedor.part@jetbrains.com

Iddo Tzameret

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

Abstract

Resolution over linear equations is a natural extension of the popular resolution refutation system, augmented with the ability to carry out basic counting. Denoted $\text{Res}(\text{lin}_R)$, this refutation system operates with disjunctions of linear equations with boolean variables over a ring R, to refute unsatisfiable sets of such disjunctions. Beginning in the work of [26], through the work of [17] which focused on tree-like lower bounds, this refutation system was shown to be fairly strong. Subsequent work (cf. [18, 17, 19, 13]) made it evident that establishing lower bounds against general $\text{Res}(\text{lin}_R)$ refutations is a challenging and interesting task since the system captures a "minimal" extension of resolution with counting gates for which no super-polynomial lower bounds are known to date.

We provide the first super-polynomial size lower bounds on general (dag-like) resolution over linear equations refutations in the large characteristic regime. In particular we prove that the subset-sum principle $1 + x_1 + \cdots + 2^n x_n = 0$ requires refutations of exponential-size over $\mathbb{Q}$. Our proof technique is nontrivial and novel: roughly speaking, we show that under certain conditions every refutation of a subset-sum instance $f = 0$, where f is a linear polynomial over $\mathbb{Q}$, must pass through a fat clause containing an equation $f = \alpha$ for each α in the image of f under boolean assignments. We develop a somewhat different approach to prove exponential lower bounds against tree-like refutations of any subset-sum instance that depends on n variables, hence also separating tree-like from dag-like refutations over the rationals.

We then turn to the finite fields regime, showing that the work of Itsykson and Sokolov [17] who obtained tree-like lower bounds over $\mathbb{F}_2$ can be carried over and extended to every finite field. We establish new lower bounds and separations as follows: (**i**) for every pair of distinct primes p, q, there exist CNF formulas with short tree-like refutations in $\text{Res}(\text{lin}_{\mathbb{F}_p})$ that require exponential-size tree-like $\text{Res}(\text{lin}_{\mathbb{F}_q})$ refutations; (**ii**) random k-CNF formulas require exponential-size tree-like $\text{Res}(\text{lin}_{\mathbb{F}_p})$ refutations, for every prime p and constant k; and (**iii**) exponential-size lower bounds for tree-like $\text{Res}(\text{lin}_{\mathbb{F}})$ refutations of the pigeonhole principle, for *every* field $\mathbb{F}$.

2012 ACM Subject Classification Theory of computation → Proof complexity

Keywords and phrases Proof complexity, concrete lower bounds, resolution, satisfiability, combinatorics

Digital Object Identifier 10.4230/LIPIcs.ITCS.2020.19

Related Version `https://eccc.weizmann.ac.il/report/2018/117/`

Acknowledgements 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 $\text{Res}(\text{lin}_{\mathbb{F}_2})$ that can be achieved using the results of Garlik and Kołodziejczyk. We thank Edward Hirsch for spotting a gap in the initial proof of the dag-like lower bound concerning the use of the contraction.

11th Innovations in Theoretical Computer Science Conference (ITCS 2020).
Editor: Thomas Vidick; Article No. 19; pp. 19:1–19:37
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

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 $D_1, \ldots, D_l = \emptyset$, where $\emptyset$ is the empty clause, such that each D_i is either a clause of the CNF or is derived from previous clauses $D_j, D_k, j \leq k < i$ 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. [23]). 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 (though it is important to clarify that contemporary SAT-solvers utilize more than the strength of tree-like resolution).

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 $m > n$ [15]. 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 [28].

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 many 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 [26] 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 [26]. 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 R^0(lin). Subsequently, Itsykson and Sokolov [17] studied resolution over linear equations over $\mathbb{F}_2$, denoted Res($\oplus$). They demonstrated the power of resolution with counting mod 2 as well as its limitations by means of several upper bounds and tree-like lower bounds. Moreover, [17] introduced 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 Res($\oplus$) refutations. In both [26] and [17] the dag-like lower bound question for resolution over linear equations remained open.

Apart from being a very natural refutation system, understanding the proof complexity of resolution over linear equations is important for the following reason: proving super-polynomial dag-like lower bounds against resolution over linear equations for prime fields and for the integers can be viewed as a first step towards the long-standing open problems of $\mathsf{AC}^0[p]$-Frege and TC^0-Frege lower bounds, respectively. We explain this in what follows.

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 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. A particular choice of the formalism is not important: a classical result by Reckhow [27] assures us that all Frege systems are polynomially equivalent. 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 [1, 24, 20, 7, 16]. 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. Recent works by Krajíček [18], Garlik-Kołodziejczyk [13] and Krajíček-Oliveira [19] had suggested possible approaches to attack dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$ lower bounds (though this problem remains open to date).

1.1 Our Results and Techniques

In this work we prove a host of new lower bounds, separations and upper bounds for resolution over linear equations. Our main novel technical contribution is a dag-like refutation lower bound over large characteristic fields. Conceptually, the proof idea exploits two main properties that recently have been found useful in proof complexity:

(i) Single axiom: the hard instance consists of a single unsatisfiable axiom (for boolean assignments)

$$1 + x_1 + \cdots + 2^n x_n = 0 \tag{1}$$

(unlike, for instance, a set of clauses).

(ii) Large coefficients: the hard instance uses coefficients of exponential magnitude.

Although employing different approaches, both of these properties played a recent role in proof complexity lower bounds. Forbes et al. [12] used subset-sum variants (that is, unsatisfiable linear equations with boolean variables) to establish lower bounds on subsystems of the ideal proof system (IPS) over large characteristic fields, where IPS is the strong proof system introduced by Grochow and Pitassi [14]. It is essential in both [12] and our work that the hard instance takes the form of a single unsatisfiable axiom. Subsequently, in a very recent work, Alekseev et al. [3] established conditional exponential-size lower bounds on full IPS refutations over the rationals of the same subset-sum instance (1), where the use of big coefficients is again essential to the lower bound. We explain our dag-like lower bound in Section 1.1.2.

The other novel contribution we make is a systematic development of new kinds of lower bound techniques against *tree-like* resolution over linear equations, both over the rationals and over finite fields. To this end we develop new and extend existing combinatorial techniques such as the Prover-Delayer game method as originated in Pudlak and Impagliazzo [25] for resolution, and developed further by Itsykson and Sokolov [17]. Moreover, we provide new applications in proof complexity of different combinatorial results; this include bounds on the size of essential coverings of the hypercube from Linial and Radhakrishnan [21], a result about the hyperplane coverings of the hypercube by Alon and Füredi [4] and the notion of immunity from Alekhnovich and Razborov [2]. We further non-trivially extend the well-established principle of size-width tradeoffs in resolution [8] to the setting of $\mathrm{Res}(\mathrm{lin}_R)$ (though it is important to note that most of our lower bounds do not follow from this tradeoff result).

1.1.1 Background

For a ring R, the refutation system $\mathrm{Res}(\mathrm{lin}_R)$ is defined as an extension of the resolution refutation system as follows (see Raz and Tzameret [26]). The *proof-lines* of $\mathrm{Res}(\mathrm{lin}_R)$ are called ***linear clauses*** (sometimes called simply *clauses*), which are defined as disjunctions of linear equations (with duplicate equations contracted). More formally, they are disjunctions of the form:

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

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 where C, D are linear clauses. A $\mathrm{Res}(\mathrm{lin}_R)$ *refutation* of an unsatisfiable over 0-1 set of linear clauses $C_1, \ldots, C_m$ 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 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).

The ***size*** of a $\mathrm{Res}(\mathrm{lin}_R)$ refutation 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) is the standard size of the binary representation of integers (nevertheless, when we talk about "big" or "exponential" coefficients and "polynomially bounded" coefficients, etc., we mean that the *magnitude* of the coefficients is big (exponential) or polynomially bounded).

We are generally interested in the following questions:

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

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

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

1.1.1.1 Tree-like $\mathrm{Res}(\mathrm{lin}_{\boldsymbol{R}})$ with semantic weakening

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.

The system $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ is obtained from $\mathrm{Res}(\mathrm{lin}_R)$ 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*):[1]

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

The reason for studying $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ is mainly the following: Let Γ be an arbitrary set of tautological R-linear clauses. Then, lower bounds for tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ imply lower bounds for tree-like $\mathrm{Res}(\mathrm{lin}_R)$ 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 $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ is quite limited. For instance, we show that $2x_1 + \cdots + 2x_n = 1$ requires refutations of exponential in n size (Theorem 35). On the other hand, such contradictions *do* admit short tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutations in the presence of the following *generalized boolean axioms* (which is a tautological linear clause):

$$\mathsf{Im}(f) := \bigvee\nolimits_{A \in im_2(f)} (f = A), \tag{2}$$

where $im_2(f)$ is the image of a linear polynomial 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 have. If a lower bound holds for tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ it also holds, in particular, for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ with the axioms $\mathsf{Im}(f)$, and this makes tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ a useful system, for which lower bounds against are sufficiently interesting.

1.1.2 Characteristic Zero Lower Bounds

For characteristic zero fields we will use mainly the rational number field $\mathbb{Q}$ (though many of the results hold over any characteristic zero rings). First, we show that over $\mathbb{Q}$, whenever $\alpha_1 x_1 + \cdots + \alpha_n x_n + \beta = 0$ is unsatisfiable (over 0-1 assignments), it has polynomial dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ refutations if the coefficients are polynomially bounded in magnitude, while it requires exponential dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ refutations for some subset-sum instances with exponential-magnitude coefficients. Note that $\alpha_1 x_1 + \cdots + \alpha_n x_n + \beta = 0$ expresses the *subset-sum principle*: $\alpha_1 x_1 + \cdots + \alpha_n x_n = -\beta$ is satisfiable iff there is a subset of the integral coefficients α_i whose sum is precisely $-\beta$. The lower bound is stated in the following theorem:

▶ **Theorem** (Theorem 21; Main dag-like lower bound). *Any* $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ *refutation of* $x_1 + 2x_2 + \cdots + 2^n x_n + 1 = 0$ *requires size* $2^{\Omega(n)}$.

The proof of this theorem introduces a new lower bound technique. We show that every (dag- or tree-like) refutation π of $x_1 + 2x_2 + \cdots + 2^n x_n + 1 = 0$ can be transformed without much increase in size into a derivation of a certain "fat" (exponential-size) clause C_π from boolean axioms only.[2] In order to prove that C_π is fat, we ensure that every disjunct $g = 0$

---

[1] 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 $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ proofs cannot be checked in polynomial time and thus $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ 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 [11]; see Section 2.2).

[2] The notion of showing that a refutation must go though a fat (i.e., wide) clause is well established in resolution lower bounds. However, we note that our lower bound is completely different from the known size-width based resolution lower bounds (as formulated in a generic way in the work of Ben-Sasson and Wigderson [8]).

in C_π has at most 2^{cn} satisfying boolean assignments, for some constant $c < 1$. Because C_π is derived from boolean axioms alone, it must be a boolean tautology, that is, it must have 2^n satisfying assignment. Since every disjunct in C_π is satisfied by at most 2^{cn} assignments, the number of disjuncts in the clause is at least $2^{(1-c)n}$. Since our constructed derivation is not much larger than the original refutation, the size of the original refutation must be $2^{\Omega(n)}$.

This proof relies in an essential way on the fact that the coefficients of the linear form have exponential magnitude. Indeed, every contradiction of the form $f = 0$ can be shown to admit polynomial-size dag-like $\mathrm{Res}(\mathrm{lin}_\mathbb{Q})$ refutations whenever the coefficients of f are polynomially bounded. A natural question is whether in the case of bounded coefficients, $f = 0$ can be efficiently refuted already by tree-like $\mathrm{Res}(\mathrm{lin}_\mathbb{Q})$ refutations. The question turns out to be non-trivial, and we provide a negative answer:

▶ **Theorem** (Theorem 35; Subset-sum tree-like lower bounds). *Let f be any linear polynomial over $\mathbb{Q}$, which depends on n variables. Then tree-like $\mathrm{Res}(\mathrm{lin}_\mathbb{Q})$ refutations of $f = 0$ are of size $2^{\Omega(\sqrt{n})}$.*

The proof is in two stages. First, we use a transformation analogous to the one used for the dag-like lower bound to reduce the lower bound problem for refutations of $f = 0$ to a lower bound problem for derivations of clauses of a certain kind. Namely, we transform any tree-like refutation π of $f = 0$ to a tree-like derivation of C_π from boolean axioms without much increase in size. The only difference is that this time we ensure that in every disjunct $g = 0$ of C_π, the linear polynomial g depends on at least $\frac{n}{2}$ variables.

Second, we prove that tree-like $\mathrm{Res}(\mathrm{lin}_\mathbb{Q})$ derivations of such a C_π are large:

▶ **Theorem** (Theorem 33). *Any tree-like $\mathrm{Res}(\mathrm{lin}_\mathbb{Q})$ derivation of any tautology of the form $\bigvee_{j\in[N]} g_j = 0$, for some positive N, where each g_j is linear over $\mathbb{Q}$ and depends on at least $\frac{n}{2}$ variables, is of size $2^{\Omega(\sqrt{n})}$.*

To prove this, as well as some other lower bounds, we extend the Prover-Delayer game technique as originated in Pudlak-Impagliazzo [25] for resolution, and developed further by Itsykson-Sokolov [17] for $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$, to general rings, including characteristic zero rings (see Sec. 5.2).[3]

We define a non-trivial strategy for Delayer in the corresponding game and prove that it guarantees $\sqrt{n}$ coins using a bound on the size of essential coverings of the hypercube from Linial and Radhakrishnan [21]. The relation between Prover-Delayer games and tree-like $\mathrm{Res}(\mathrm{lin}_\mathbb{Q})$ refutations allows us to conclude that the size of tree-like $\mathrm{Res}(\mathrm{lin}_\mathbb{Q})$ refutations must be $2^{\Omega(\sqrt{n})}$.

Moreover, as a corollary of Theorem 33 we obtain a lower bound on tree-like $\mathrm{Res}(\mathrm{lin}_\mathbb{Q})$ *derivations* (in contrast to refutations) of $\mathsf{Im}(f)$:

▶ **Corollary** (Corollary 34). *Let f be any linear polynomial over $\mathbb{Q}$ that depends on n variables. Then tree-like $\mathrm{Res}(\mathrm{lin}_\mathbb{Q})$ derivations of $\mathsf{Im}(f)$ are of size $2^{\Omega(\sqrt{n})}$.*

We also use Prover-Delayer games to 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 [17] for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$.

[3] We note here (see Remark 1 in the next sub-section) that the lower bounds that we prove using Prover-Delayer games techniques in case $char(\mathbb{F}) = 0$ *do not* follow from lower bounds for Polynomial Calculus using size-width relations.

▶ **Theorem** (Theorem 38; Pigeonhole principle lower bounds). *Let $\mathbb{F}$ be any (possibly finite) field. Then every tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ refutation of $\neg\mathrm{PHP}_n^m$ has size $2^{\Omega(\frac{n-1}{2})}$.*

Together with the polynomial upper bounds for PHP_n^m refutations in dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ for fields $\mathbb{F}$ of characteristic zero demonstrated by Raz and Tzameret [26], Theorem 38 establishes 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, for the language of unsatisfiable formulas in CNF:

▶ **Corollary.** *Over fields of characteristic zero $\mathbb{F}$, $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ has an exponential speed-up over tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ as refutation systems for unsatisfiable formulas in CNF.*

To prove Theorem 38 we need to prove that Delayer's strategy from [17] is successful over any field. This argument is new, and uses a result of Alon-Füredi [4] about the hyperplane coverings of the hypercube.

We prove another separation between dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ and tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{Q}})$, as follows. For any ring R 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 $0 \le i \le n$, $x_1 + \cdots + x_n$ equals some element in $\{0,\ldots,n\}\setminus i$. In more generality, let f be a linear form over $\mathbb{Q}$ and let $im_2(f)$ be the image of f under 0-1 assignments to its variables. Define $\langle f \neq A\rangle := \bigvee_{A\neq B\in im_2(f)}(f = B)$, where $A \in \mathbb{Q}$. We define

$$\mathsf{ImAv}\,(f) := \{\langle f \neq A\rangle : A \in im_2(f)\}\,. \tag{3}$$

▶ **Corollary** (Corollary 13). *For every ring R and every linear form f the contradiction $\mathsf{ImAv}\,(f)$ admits polynomial-size $\mathrm{Res}(\mathrm{lin}_R)$ refutations.*

▶ **Theorem** (Theorem 37). *We work over $\mathbb{Q}$. Let $f = \epsilon_1 x_1 + \cdots + \epsilon_n x_n$, where $\epsilon_i \in \{-1, 1\}$. Then any tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{Q}})$ refutation of $\mathsf{ImAv}\,(f)$ is of size at least $2^{\frac{n}{4}}$.*

The lower bound in Theorem 37 is one more novel application of the Prover-Delayer game argument, combined with the notion of immunity from Alekhnovich and Razborov [2], as we now briefly explain.

Let f be a linear form as in Theorem 37. We consider an instance of the Prover-Delayer game for $\mathsf{ImAv}\,(f)$. A position in the game 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. In the beginning Φ is empty. 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 $f \equiv A + 1 (\mathrm{mod}\ 2)$ ([2]) implies $m \ge \frac{n}{4}$. To conclude, by a standard argument 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 $\mathbb{Q}$ (and for some cases any characteristic 0 field):

1.1.3 Finite Fields Lower Bounds

We now turn to resolution over linear equations in *finite fields*. We obtain many new tree-like lower bounds (see Table 2).

Table 1 Lower and upper bounds for $\mathbb{Q}$. 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 contrast to refutations. All results except the upper bound on PHP are from the current work.

	$\sum_{i=1}^{n} 2x_i = 1$	$\sum_{i=1}^{n} 2^i x_i = -1$	$\mathsf{ImAv}\left(\sum_{i=1}^{n} x_i\right)$	PHP_n^m (CNF)	$\mathsf{Im}\left(\sum_{i=1}^{n} x_i\right)$
t-l $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$	$2^{\Omega(\sqrt{n})}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(\sqrt{n})}$
t-l $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{Q}})$	poly	poly	$2^{\Omega(n)}$	$2^{\Omega(n)}$	poly
$\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$	poly	$2^{\Omega(n)}$	poly	poly [26]	poly

We already discussed above lower bounds for the pigeonhole principle which hold both for positive and zero characteristic. We furthermore 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 $\mathrm{TS}^{(p)}_{G,\sigma}$ (written as CNFs), 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 $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}_{p^c}})$ on random k-CNFs.[4]

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 $PC_{\mathbb{F}}$ [10], 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 known lower bounds on $PC_{\mathbb{F}}$ degrees ([2]).

We show that

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

where ω_0 is what we call the *principal width*, which counts the number of linear equations in clauses when we treat as identical those defining parallel hyperplanes, and $S_{\text{t-l Res}(\mathrm{lin}_R)}(\phi \vdash \perp)$ 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 44; Size-width relation). *Let ϕ be an unsatisfiable set of linear clauses over a field $\mathbb{F}$. The following relation between principal width and size holds for both tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ and tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$: $S(\phi \vdash \perp) = 2^{\Omega(\omega_0(\phi \vdash \perp) - \omega_0(\phi))}$. If $\mathbb{F}$ is a finite field, then the same relation holds for the (standard) width of a clause ω.*

This extends to every field a result by Garlik-Kołodziejczyk [13, 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 45). *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 $\omega(\pi)$.*

---

[4] 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.

▶ **Corollary** (Corollary 46; 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* $\text{Res}(\text{lin}_{\mathbb{F}})$ *refutation of the Tseitin mod p formula $\neg\text{TS}^{(p)}_{G,\sigma}$ has size $2^{\Omega(dn)}$.*

▶ **Corollary** (Corollary 47; 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* $\text{Res}(\text{lin}_{\mathbb{F}})$ *refutation of ϕ has size $2^{\Omega\left(\frac{n}{\Delta^{2/(k-2)}\cdot\log\Delta}\right)}$ with probability $1 - o(1)$.*

▶ Remark 1. We stress that the size-width relation of Theorem 44 **cannot** be used for transferring $PC_{\mathbb{F}}$ degree lower bounds to tree-like $\text{Res}(\text{lin}_{\mathbb{F}})$ size lower bounds in case $char(\mathbb{F}) = 0$. This is due to the essential difference between principal width and width in this case. Thus, all the lower bounds that we prove using Prover-Delayer games techniques in case $char(\mathbb{F}) = 0$ **do not** follow from lower bounds for $PC_{\mathbb{F}}$.

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

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 $\text{TS}^{(-)}_{G,\sigma}$ stands for $\text{TS}^{(p)}_{G,\sigma}$ in the first and the third rows and for $\text{TS}^{(2)}_{G,\sigma}$ in the second row. t-l $\text{Res}(\text{lin}_R)$ stands for tree-like $\text{Res}(\text{lin}_R)$, and $p \neq q$ are primes (in the second row and third column we assume $q \neq 2$). Circled "?" denotes an open problem. The results marked with [17, 13] were proved in the respective papers. All other results are from the current work.

	$A\overline{x} = \overline{b}$	$\text{TS}^{(-)}_{G,\sigma}$	$\text{TS}^{(q)}_{G,\sigma}$	random k-CNF	PHP^m_n
t-l $\text{Res}(\text{lin}_{\mathbb{F}_{p^k}})$	$2^{\Omega(n)}$	poly	$2^{\Omega(dn)}$	$2^{\Omega\left(\frac{n}{\Delta^{2/(k-2)}\cdot\log\Delta}\right)}$	$2^{\Omega(n)}$
t-l $\text{Res}(\oplus)$	poly [17]	poly [17]	$2^{\Omega(dn)}$	$2^{\Omega\left(\frac{n}{\Delta^{2/(k-2)}\cdot\log\Delta}\right)}$ [13]	$2^{\Omega(n)}$ [17]
t-l $\text{Res}_{sw}(\text{lin}_{\mathbb{F}_{p^k}})$	poly	poly	?	?	$2^{\Omega(n)}$

1.1.4 Complexity of Linear Systems

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

▶ **Proposition** (Proposition 14; Upper bounds on unsatisfiable linear systems)**.** *Let $\mathbb{F}$ be a field and assume that the linear system $A\,\overline{x} = \overline{b}$, where A is a $k \times n$ matrix over $\mathbb{F}$, has no solutions (over $\mathbb{F}$). Let ϕ be a CNF formula encoding the linear system $A\,\overline{x} = \overline{b}$. Then, there exist tree-like* $\text{Res}(\text{lin}_{\mathbb{F}})$ *refutations of ϕ of size polynomial in the sum of sizes of encodings of all coefficients in A.*

The upper bound in Proposition 14 applies only to linear systems that are unsatisfiable over the *whole* field $\mathbb{F}$. But does any system $A\,\overline{x} = \overline{b}$ over $\mathbb{F}$ that has a satisfying assignment over $\mathbb{F}$, but *not* over 0-1 assignments, admit polynomial-size $\text{Res}(\text{lin}_{\mathbb{F}})$ refutations?

For fields $\mathbb{F}$ with $char(\mathbb{F}) \geq 5$ or $char(\mathbb{F}) = 0$ it is known that 0-1 satisfiability of $A\,\overline{x} = \overline{b}$ is NP-complete. This means that unless coNP = NP there exist 0-1 unsatisfiable linear systems that require superpolynomial dag-like Res($\mathrm{lin}_{\mathbb{F}}$) refutations.

If $char(\mathbb{F}) \geq k+1$ or $char(\mathbb{F}) = 0$, the canonical reduction R from the language k-UNSAT of unsatisfiable k-CNFs maps every $\phi(\overline{x}) \in k$-UNSAT to the system $R_\phi(\overline{x}, \overline{y})$ by encoding every clause in $\phi(\overline{x})$ as a linear equality with extra variables. This simple reduction allows to establish tight connections between proof complexity of CNF formulas and linear systems.

Firstly, lower bounds on $R_\phi(\overline{x}, \overline{y})$ imply lower bounds on $\phi(\overline{x})$: by implicational completeness there are polynomial-size derivations of $\phi(\overline{x})$ from $R_\phi(\overline{x}, \overline{y})$ in Res($\mathrm{lin}_{\mathbb{F}}$).

Secondly, if $\mathbb{F}$ is a finite field, tree-like Res($\mathrm{lin}_{\mathbb{F}}$) lower bounds on $\phi(\overline{x})$ imply tree-like Res($\mathrm{lin}_{\mathbb{F}}$) lower bounds on $R_\phi(\overline{x}, \overline{y})$. Each linear equation $l(\overline{x}, \overline{y}) = 0$ in $R_\phi(\overline{x}, \overline{y})$ is equivalent to a polynomial equation $l(\overline{x}, \overline{p}(\overline{x})) = 0$, where $\overline{p}$ are polynomials of constant degree. Therefore, there is a constant degree $PC_{\mathbb{F}}$ derivation π_ϕ of $R_\phi(\overline{x}, \overline{p}(\overline{x}))$ from $\phi(\overline{x})$ and vice versa. As any $PC_{\mathbb{F}}$ refutation of $R_\phi(\overline{x}, \overline{y})$ can be turned into a refutation of $R_\phi(\overline{x}, \overline{p}(\overline{x}))$ by substitution without much loss in degree, it is easy to see that $PC_{\mathbb{F}}$ refutes $R_\phi(\overline{x}, \overline{y})$ in degree d iff $PC_{\mathbb{F}}$ refutes $\phi(\overline{x})$ in degree $\Theta(d)$. By size-width relation for finite fields (Theorem 44), we obtain that for any formula $\phi(\overline{x})$ that is hard for $PC_{\mathbb{F}}$, $R_\phi(\overline{x}, \overline{y})$ is hard for tree-like Res($\mathrm{lin}_{\mathbb{F}}$).

1.1.5 Nondeterministic Linear Decision Trees

There is a 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 [17] this correspondence was generalized to tree-like Res($\oplus$) refutations and parity decision trees. In the paper by Beame et al. [5] an analogous correspondence was shown for tree-like R(CP) refutations [5] and decision trees that brunch on linear inequalities. In the current work we initiate the study of linear decision trees and their properties over different characteristics, extending the correspondence of [17] to a correspondence between tree-like Res(lin_R) (and tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$) derivations to what we call *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. (Note that in the same manner that in a (boolean) decision tree (which corresponds to a tree-like resolution refutation) we go along a path from the root to a leaf, choosing those edges that violate a literal x_i or $\neg x_i$, in an NLDT we branch along a path that violates equalities $f = 0$, or equivalently, certifies non-equalities of the form $f \neq 0$.)

2 Preliminaries

2.1 Notation

Denote by $[n]$ the set $\{1, \ldots, n\}$. We use $x_1, x_2, \ldots$ to denote variables, both propositional and algebraic. Let f be a linear polynomial (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

[5] R(CP) is a system operating with disjunctions of integer linear inequalities $f \geq 0$

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 $A \in R$.

A *linear clause* is a formula of the form $(\sum_{i=1}^{n} a_{1i}x_i + b_1 = 0) \vee \cdots \vee (\sum_{i=1}^{n} a_{ki}x_i + b_k = 0)$ with $x_1, \ldots, x_n$ variables, and a_{ij}, b_i's ring elements (when the ring is specified in advanced). We sometimes abuse notation by writing a linear equation as $\sum_{i=0}^{n} a_{1i}x_i = -b_1$ instead of $\sum_{i=0}^{n} a_{1i}x_i + b_1 = 0$. We assume that all the disjuncts in a linear clause are distinct.

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 polynomial 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, 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 polynomial 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 each linear polynomial l_i does not depend on x_i. For a clause or a set of clauses ϕ we define $\phi\restriction_\rho := (\ldots((\phi\restriction_{x_1\leftarrow l_1})\restriction_{x_2\leftarrow l_2})\ldots)\restriction_{x_n\leftarrow l_n}$.

Denote UNSAT $\subset \{0,1\}^*$ (resp. k-UNSAT $\subset \{0,1\}^*$) the language of unsatisfiable propositional CNF (resp. k-CNF) formulas. 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 $S_\Pi(\phi \vdash \perp)$ (we sometimes omit the subscript Π when it is clear from the context) the minimal size of a Π-refutation of ϕ.

2.2 Propositional Proof Systems

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 $1 \leq i \leq s$ 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* [10] 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 $1 \leq i \leq s$ 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:

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

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^m_n, 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 $\neg\mathrm{PHP}^m_n := \mathsf{Pigeons}^m_n \cup \mathsf{Holes}^m_n \in \mathrm{UNSAT}$, where $\mathsf{Pigeons}^m_n = \{\bigvee_{j \in [n]} x_{i,j}\}_{i \in [m]}$ are axioms for pigeons and $\mathsf{Holes}^m_n = \{\neg x_{i,j} \vee \neg x_{i',j}\}_{i \neq i' \in [m], j \in [n]}$ are axioms for holes.

2.3.2 Mod p Tseitin Formulas

We use the version given in [2] (which is different from the one in [9, 26]). 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* $\neg\mathrm{TS}^{(p)}_{G,\sigma}$ are the CNF encoding of the following equations for all $u \in V$:

$$\sum_{(u,v) \in E} x_{(u,v)} - \sum_{(v,u) \in E} x_{(v,u)} \equiv \sigma(u) \mod p\,. \tag{4}$$

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 $\sum_{u \in V} \sigma(u) \not\equiv 0 \mod p$. To see this, note that if we sum (4) 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 [28] 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 $e(U, U') := \{(u, u') \in E \,|\, u \in U, u' \in U'\}$. 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 2** (Lubotzky *et. al* [22])**.** *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 3** (Alekhnovich-Razborov [2])**.** *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 $\mathrm{char}(\mathbb{F}) \neq p$, 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 4** (Alekhnovich-Razborov [2])**.** *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 over Linear Equations for 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: $(\sum_{i=0}^n a_{1i}x_i + b_1 = 0) \vee \cdots \vee (\sum_{i=0}^n a_{ki}x_i + b_k = 0)$. Recall that disjunctions of this form are called *linear clauses*, and that we assume that all disjuncts are distinct, hence contract duplicate linear equations. We sometimes abuse notation by writing a linear equation as $(\sum_{i=0}^n a_{1i}x_i = -b_1)$ instead of $(\sum_{i=0}^n a_{1i}x_i + b_1 = 0)$.

The rules of $\mathrm{Res}(\mathrm{lin}_R)$ are as follows (cf. [26]):

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

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

where $f(\overline{x}), g(\overline{x})$ are linear forms over R and C, D are linear clauses. Note that contraction of duplicates disjuncts is done automatically when applying the resolution rule. 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 $\text{Res}(\text{lin}_R)$ derivation of the empty clause (which stands for false) from ϕ. The ***size*** of a $\text{Res}(\text{lin}_R)$ 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 5.** *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* $\text{Res}(\text{lin}_R)$ *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 $\mu(\phi \vdash \bot)$ the minimal value of $\mu(\pi)$ over all* $\text{Res}(\text{lin}_R)$ *refutations π.*

▶ **Proposition 6.** $\text{Res}(\text{lin}_R)$ *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* $\text{Res}(\text{lin}_R)$ *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 28. ◀

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 [17]), 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 7.** *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). ◀

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 [17], respectively[6]. As showed in [17], $\text{Res}(\text{lin}_{\mathbb{F}_2})$, $\text{Res}_{sw}(\text{lin}_{\mathbb{F}_2})$ 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):

▶ **Proposition 8.** *If $char(R) \notin \{1, 2, 3\}$, then dag-like $\text{Res}_{sw}(\text{lin}_R)$ and tree-like $\text{Sem-Res}(\text{lin}_R)$ are polynomially bounded (not necessarily Cook-Reckhow) propositionally proof systems for 3-UNSAT.*

Proof. Let $\phi(x_1, \ldots, x_n) = \{C_i\}_{i \in [m]} \in$ 3-UNSAT. Given $C = (x_{j_1}^{\sigma_1} \vee \ldots \vee x_{j_k}^{\sigma_k})$ define $lin(\neg C) := ((2\sigma_1 - 1)x_{j_1} + \ldots + (2\sigma_k - 1)x_{j_k} - (\sigma_1 + \ldots + \sigma_k))$ 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 $\text{Res}_{sw}(\text{lin}_R)$ in a single step as a weakening of $0 = 0$ or resolving $0 = 0$ with $0 = 0$ in tree-like $\text{Sem-Res}(\text{lin}_R)$.

In tree-like $\text{Sem-Res}(\text{lin}_R)$ 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 $C_1, \ldots, C_m$.

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

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

▶ **Proposition 9.** *If R is an integral domain and $Frac(R)$ is its field of fractions, then $\text{Res}(\text{lin}_R)$ is equivalent to $\text{Res}(\text{lin}_{Frac(R)})$ and tree-like $\text{Res}(\text{lin}_R)$ is equivalent to tree-like $\text{Res}(\text{lin}_{Frac(R)})$.*

Proof. Every proof in $\text{Res}(\text{lin}_R)$ is also a proof in $\text{Res}(\text{lin}_{Frac(R)})$. To get the converse, just multiply every line by the least common multiple (lcm) of all the coefficients in the $\text{Res}(\text{lin}_{Frac(R)})$ proof. If $a_1, \ldots, a_N \in R$ is the list of denominators of all the coefficients in a $\text{Res}(\text{lin}_{Frac(R)})$ proof π, then under a reasonable encoding of R: $|lcm(a_1, \ldots, a_N)| \leq |a_1| + \cdots + |a_N| \leq |\pi|$. Therefore the corresponding $\text{Res}(\text{lin}_R)$ proof is of size at most $O(|\pi|^2)$. ◀

[6] 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*).

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

Here we introduce several unsatisfiable sets of linear clauses that express some counting principles, and serve to exemplify the ability of dag-like $\mathrm{Res}(\mathrm{lin}_R)$, tree-like $\mathrm{Res}(\mathrm{lin}_R)$ 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 instances 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):

$$\mathsf{LinSys}(A) := \{B_i \cdot \overline{x} = b_i\}_{i \in [m]}\,. \tag{5}$$

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

$$\mathsf{SubSum}(f) := \{f = 0\}\,. \tag{6}$$

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

$$\mathsf{ImAv}\,(f) := \{\langle f \neq A \rangle : A \in im_2(f)\}\,. \tag{7}$$

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

Image axiom: For f a linear form, define

$$\mathsf{Im}(f) := \bigvee_{A \in im_2(f)} f = A\,. \tag{8}$$

Dag-Like $\mathrm{Res}(\mathrm{lin}_{\boldsymbol{R}})$

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

▶ **Proposition 10.** *Let $f = \sum_{i=1}^n a_i x_i + b$ be a linear polynomial over R. There exists a $\mathrm{Res}(\mathrm{lin}_R)$ derivation of $\mathsf{Im}(f)$ of size polynomial in $|\mathsf{Im}(f)|$ and of principal width at most 3.*

Proof. We construct derivations of $\mathsf{Im}\Big(\sum_{i=1}^k a_i x_i + b\Big)$, $0 \le k \le n$, 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 $C_0 := \Big(\bigvee_{A \in im_2(f_k)} f_k + a_{k+1}x_{k+1} = A\Big) \vee x_{k+1} = 1$ 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 $C_1 := \Big(\bigvee_{A \in im_2(f_k)} f_k + a_{k+1}x_{k+1} = A + a_{k+1}\Big)$ $\vee\, x_{k+1} = 0$ 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. ◀

▶ **Proposition 11.** *For every linear polynomial 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 10 the resulting refutation is of polynomial in $|\mathsf{Im}(f)|$ size. ◀

▶ **Proposition 12.** *Let f be a linear polynomial 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 10 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 $(1 + \cdots + (N-2))|f| = \frac{(N-1)(N-2)}{2}|f|$. ◀

▶ **Corollary 13.** *For every ring R and every linear polynomial 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 12 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 $\langle f \neq a\rangle \in \mathsf{ImAv}(f)$. ◀

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

Lower bounds. In Sec. 4 we prove an exponential lower bound for $\mathsf{SubSum}(f)$ in case f is a linear polynomial with large coefficients (Theorem 21).

Tree-Like $\mathrm{Res}(\mathrm{lin}_{\boldsymbol{R}})$

Upper bounds. In case R is a finite ring, in Sec. 5.1 we prove that the clauses in $\mathsf{Im}(f)$ admit derivations of polynomial size (Theorem 29). 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 contrast, in case $R = \mathbb{Q}$ 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 14.** *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 $1 = 0$. ◀

Lower bounds. In Sec. 4 we prove tree-like $\text{Res}(\text{lin}_{\mathbb{Q}})$ exponential-size lower bounds for derivations of $\mathsf{Im}(f)$ and refutations of $\mathsf{SubSum}(f)$ for any f (Corollary 34 and Theorem 35). For $\mathsf{ImAv}\,(f)$ whenever f is of the form $f = \epsilon_1 x_1 + \ldots + \epsilon_n x_n - A$ for some $\epsilon_i \in \{-1, 1\}, A \in \mathbb{F}$ the lower bound holds even for the stronger system tree-like $\text{Res}_{sw}(\text{lin}_{\mathbb{F}})$ (see below).

Tree-Like $\text{Res}_{sw}(\text{lin}_{R})$

Upper bounds. Most of the instances above admit short derivations/refutations in tree-like $\text{Res}_{sw}(\text{lin}_R)$: $\mathsf{Im}(f)$ 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 $\mathsf{LinSys}(A) = \{A_i\overline{x} - b_i = 0\}_{i\in[m]}$.

Lower bounds. In case $\mathbb{F}$ is a field of characteristic zero, $\mathsf{ImAv}\,(f)$ are hard even for tree-like $\text{Res}_{sw}(\text{lin}_R)$ whenever f is of the form $f = \epsilon_1 x_1 + \ldots + \epsilon_n x_n - A$ for some $\epsilon_i \in \{-1, 1\}, A \in \mathbb{F}$ (Theorem 37).

3.2 CNF Upper Bounds for $\text{Res}(\text{lin}_{R})$

In this section we outline two basic polynomial upper bounds, which we use to establish our separations in subsequent sections: short tree-like $\text{Res}(\text{lin}_R)$ refutations for CNF encodings of linear systems over a ring R, and short $\text{Res}(\text{lin}_R)$ refutations for $\neg\text{PHP}^m_n$. Together with our lower bounds, these imply the separation between tree-like $\text{Res}(\text{lin}_{\mathbb{F}})$ and tree-like $\text{Res}(\text{lin}_{\mathbb{F}'})$, 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 $\text{Res}(\text{lin}_{\mathbb{F}})$ 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 15.** *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 $\text{Res}(\text{lin}_{\mathbb{F}})$ 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 14).

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 28 proved in the sequel there exists a tree-like $\text{Res}(\text{lin}_{\mathbb{F}})$ derivation of $A_i \cdot \overline{x} - b_i = 0$ from ϕ_i of size $O(2^{n_i}|A_i \cdot \overline{x} - b_i = 0|) = O(|\phi_i| \cdot \left|A_i \cdot \overline{x} - b_i = 0\right|)$.

By Proposition 14 there exists a tree-like $\text{Res}(\text{lin}_{\mathbb{F}})$ 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 16.** *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* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_p})$ *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 15 there exists a tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_p})$ refutation of $\neg \mathrm{TS}^{(p)}_{G,\sigma}$ of polynomial size. ◀

▶ **Theorem 17** (Raz and Tzameret [26]). *Let R be a ring such that $char(R) = 0$. There exists a* $\mathrm{Res}(\mathrm{lin}_R)$ *refutation of $\neg PHP^m_n$ of polynomial size.*

Proof. This follows from the upper bound of [26] for $\mathrm{Res}(\mathrm{lin}_{\mathbb{Z}})$ and the fact that any $\mathrm{Res}(\mathrm{lin}_{\mathbb{Z}})$ proof can be interpreted as $\mathrm{Res}(\mathrm{lin}_R)$ if R is of characteristic 0. ◀

4 Dag-Like Lower Bound

4.1 Lower Bound for Subset Sum with Large Coefficients

In this section we prove an exponential lower bound on the size of dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ refutations of $\mathsf{SubSum}(f)$, where $f = 1 + x_1 + \cdots + 2^n x_n$.

The lower bound is obtained by defining a mapping, that sends every refutation π of $f = 0$ to a derivation π' from the boolean axioms of some clause C_π, in such a way that π' satisfies two properties:

1. π' is at most polynomially larger than π;
2. C_π is exponentially large.

We ensure that the second property holds by defining the construction of π' in such a way that every disjunct $g = 0$ in C_π has a sufficiently small number Z_g of 0-1 solutions, namely Z_g is at most 2^{cn}, for some constant $c < 1$. This, together with the observation that C_π must be a boolean tautology, because it is derivable from the boolean axioms only, implies that C_π must be of exponential size (since C_π has 2^n satisfying assignments and each disjunct contributes at most 2^{cn} satisfying disjunctions). Therefore, by the first property, π must be of exponential size.

The fact that f has exponentially large coefficients is essential in our proof that C_π is of exponential size. All contradictions of the form $f = 0$, where f has polynomially bounded coefficients, have polynomial dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ refutations and, thus, there is no hope to prove strong bounds for dag-like refutations in this case. However, in Sec 5 we prove that any $f = 0$, as long as f depends on n variables, must have tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ refutations of size at least $2^{\Omega(\sqrt{n})}$. The argument relies on a similar transformation from refutations π of $f = 0$ to derivations of some C_π and in this way reduces the problem to proving size lower bounds against tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ derivations of C_π from the boolean axioms.

In order to deal with both tree-like and dag-like lower bounds we formulate and prove a generalised statement about the translation. For both dag-like and tree-like lower bounds we need that for all the disjuncts $g = 0$ in C_π a certain predicate $\mathcal{P}$ holds for g. In case of the dag-like bound, $\mathcal{P}(g) = 1$ iff $g = 0$ has at most 2^{cn} 0-1 solutions, while in case of the tree-like bound $\mathcal{P}(g) = 1$ iff g depends on at least $\frac{n}{2}$ variables. In Theorem 18 we prove that the translation can be achieved as long as $\mathcal{P}$ satisfies certain properties (in what follows $\mathbb{F}[x_1, \ldots, x_n]_{\leq 1}$ denotes the linear polynomials in $\mathbb{F}[x_1, \ldots, x_n]$).

▶ **Theorem 18.** *Let f be a linear polynomial over a field $\mathbb{F}$ with n variables and let $\mathcal{P}$: $\mathbb{P}(\mathbb{F}[x_1, \ldots, x_n]_{\leq 1}) \to \{0, 1\}$ be a predicate on the projective space*[7] *of linear polynomials over $\mathbb{F}$ satisfying the following properties:*

1. *for all linear polynomials g and for all but at most one $a \in \mathbb{F}$: $\mathcal{P}(g + af) = 1$;*
2. *for all $b \in \mathbb{F}$: $\mathcal{P}(b + f) = 1$.*

If there exists $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ *(resp. tree-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$*) refutation of $f = 0$ of size S, then there exists* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ *(resp. tree-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$*) derivation of size $O(n \cdot S^3)$ of a linear clause $\bigvee_{j \in [N]} g_j = 0$ (for some positive N), where $\mathcal{P}(g_j) = 1$ for every $j \in [N]$.*

Proof. We now sketch the plan of the proof. Assume that π is a $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutation of $f = 0$. By taking out resolutions with $f = 0$ we transform π into a derivation π' of some clause C such that $\mathcal{P}(g) = 1$ for every disjunct $g = 0$ in C. We do this in such a way that π' is not much larger than π: $|\pi'| = O(n \cdot |\pi|^3)$.

Denote $\pi_{\leq k}$ the fragment of π, consisting of the first k lines of π. By induction on k we define the sequence π'_k of derivations of some clauses D_k from boolean axioms. The derivations π'_k are defined together with a surjective function τ_k from lines of $\pi_{\leq k}$ to lines of π'_k such that if $D = \left(\bigvee\limits_{t \in [m]} g_t = 0\right)$ is a line in $\pi_{\leq k}$, then

$$\tau_k(D) = \left(\bigvee_{t \in [m]} g_t + a_t f = 0\right) \vee \bigvee_{s \in [m']} h_s = 0$$

is a line in π'_k, where $a_t \in \mathbb{F}$ and each h_s is a linear polynomial. Moreover, $\tau_k(D)$ satisfies the following properties:

1. For each $h_s = 0$: $\mathcal{P}(h_s) = 1$.
2. The sets H_D of disjuncts $h_s = 0$ in $\tau_k(D)$ are not too large: $\left|\bigcup_{D \in \pi_{\leq k}} H_D\right| \leq 2|\pi_{\leq k}|$.
3. The numbers a_t and coefficients of h_s are not too large: their bit-size does not exceed the maximal bit-size of coefficients in π.

Before we proceed to the inductive definition of π'_k, we finish the proof assuming that π'_k described above exists. If l is the length of π, then $\pi' := \pi'_l$ contains a derivation of $\tau_l(\emptyset)$, where $\emptyset$ denotes the empty clause.

We now turn to the inductive definition of π'_k.

Base case: Define π'_0 to be the empty derivation.
Induction step: Assume π'_k and τ_k satisfy the properties above and k is smaller than the length of π. If D is the last line of $\pi_{\leq k+1}$, then τ_{k+1} extends τ_k to D and π'_{k+1} either extends π'_k with $\tau_{k+1}(D)$ or coincides with π'_k. Consider the possible cases in which the last line D of $\pi_{\leq k+1}$ is derived:

Case 1: Boolean axiom: $D = (x_i = 0 \vee x_i = 1)$. Then π'_{k+1} extends π'_k with D and $\tau_{k+1}(D) = D$.

Case 2: $D = (f = 0)$. Then π'_{k+1} extends π'_k with the axiom $0 = 0$ and $\tau_{k+1}(D) = (f - f = 0)$.

Case 3: D is derived by resolution: $D = (C_1 \vee C_2 \vee \alpha G_1 + \beta G_2 = 0)$ for some lines $(C_1 \vee G_1 = 0)$ and $(C_2 \vee G_2 = 0)$ in $\pi_{\leq k}$.

[7] Here, a *projective space* $\mathbb{P}(\mathbb{F}[x_1, \ldots, x_n]_{\leq 1})$ means the set of linear polynomials quotient by the relation $f \sim \alpha f$ for nonzero scalars α.

If $C_i = \bigvee_{t\in[m_i]} g_t^{(i)} = 0$, by induction hypothesis $\tau_k(C_i \vee G_i = 0)$ is of the form ($i = 1, 2$; $A_i \in \mathbb{F}$):

$$\tau_k(C_i \vee G_i = 0) = \left(G_i + A_i f = 0 \vee \bigvee_{t\in[m_i]} g_t^{(i)} + a_t^{(i)} f = 0 \right) \vee \bigvee_{s\in[m_i']} h_s^{(i)} = 0$$

Define $\tau_{k+1}(D)$ to be the following resolution of $\tau_k(C_1 \vee G_1 = 0) \in \pi_k'$ with $\tau_k(C_2 \vee G_2 = 0) \in \pi_k'$:

$$\tau_{k+1}(D) := \left(\alpha G_1 + \beta G_2 + (\alpha A_1 + \beta A_2) f = 0 \vee \bigvee_{i=1,2} \bigvee_{t\in[m_i]} g_t^{(i)} + a_t^{(i)} f = 0 \right) \vee \\ \vee \bigvee_{i=1,2} \bigvee_{s\in[m_i']} h_s^{(i)} = 0$$

The derivation π_{k+1}' extends π_k' with $\tau_{k+1}(D)$. It remains to be shown that $\tau_{k+1}(D)$ is of required form and that τ_{k+1} satisfies the required properties.
If we consider the clause $(\alpha G_1 + \beta G_2 = 0 \vee C_1 \vee C_2)$ as a *multiset* of disjuncts and C_1, C_2, as usual, as sets of disjuncts, there can be up to three identical copies of $g = 0$ (from C_1, from C_2 and from $\{\alpha G_1 + \beta G_2 = 0\}$), that are contracted to a single element in the set D. In $\tau_{k+1}(D)$ these copies can be different because of different $+af$ terms and, thus, can be non-contractible.
For every disjunct $g = 0$ in D, denote $\mathcal{F}_g$ the set of disjuncts in $\tau_{k+1}(D)$ that correspond to g, namely, $(g_j^{(i)} + a_j^{(i)} f = 0) \in \mathcal{F}_g$ iff $g_j^{(i)} = g$ and $(\alpha G_1 + \beta G_2 + (\alpha A_1 + \beta A_2) f = 0) \in \mathcal{F}_g$ iff $\alpha G_1 + \beta G_2 = g$. For every $g = 0 \in D$, pick one element $g + af = 0 \in \mathcal{F}_g$, which minimises $\mathcal{P}(g+af)$, and denote X the set of these elements. Denote $Y := \left(\bigcup_{g=0\in D} \mathcal{F}_g\right) \setminus X$. Write $\tau_{k+1}(D)$ as follows:

$$\tau_{k+1}(D) = \left(\bigvee_{g+af=0\in X} g + af = 0 \right) \vee \left(\bigvee_{i=1,2} \bigvee_{s\in[m_i']} h_s^{(i)} = 0 \vee \bigvee_{g+af=0\in Y} g + af = 0 \right)$$

We now show that τ_{k+1} satisfies all the desired properties:

1. For every $h_s^{(i)} = 0$, $\mathcal{P}(h_s^{(i)}) = 1$ holds by induction hypothesis. For every $g + af = 0 \in Y$, $\mathcal{P}(g + af) = 1$ holds by definition of Y.
2. Note that $|H_D \setminus \{h_s^{(i)} = 0\}_{i,s}| \leq 2|D|$. By induction hypothesis $|\bigcup_{\tilde{D}\in\pi_{\leq k}} H_{\tilde{D}}| \leq 2|\pi_{\leq k}|$. It follows that $|\bigcup_{\tilde{D}\in\pi_{\leq k}} H_{\tilde{D}} \cup H_D| = |\bigcup_{\tilde{D}\in\pi_{\leq k}} H_{\tilde{D}} \cup (H_D \setminus \{h_s^{(i)} = 0\}_{i,s})| \leq |\bigcup_{\tilde{D}\in\pi_{\leq k}} H_{\tilde{D}}| + |H_D \setminus \{h_s^{(i)} = 0\}_{i,s}| \leq 2|\pi_{\leq k}| + 2|D| \leq 2|\pi_{\leq k+1}|$.
3. The absolute values of coefficients in π_{k+1}' do not exceed the maximal absolute value of coefficients in π.

Case 4: D is derived by simplification from a line $D \vee b = 0$ in $\pi_{\leq k}$. If $D = \left(\bigvee_{t\in[m]} g_t = 0 \right)$, then $\tau_k(D \vee b = 0)$ has the form: $\tau_k(D \vee b = 0) = \left(\bigvee_{t\in[m]} g_t + a_t f = 0 \right) \vee b + af = 0$.
If $a = 0$, we apply simplification to $\tau_k(D \vee b = 0)$ to derive $\tau_{k+1}(D) := \left(\bigvee_{t\in[m]} g_t + a_t f = 0 \right)$ and let π_{k+1}' extend π_k'.
Otherwise, if $a \neq 0$, we define $\tau_{k+1}(D)$ to be $\tau_{k+1}(D) := \tau_k(D \vee b = 0)$ and $\pi_{k+1}' := \pi_k'$.

Case 5: D is derived by weakening from a line C of $\pi_{\leq k}$: $D = (C \vee g = 0)$ for some g. Define $\tau_{k+1}(D) := (\tau_k(C) \vee g = 0)$ and let π'_{k+1} extend π'_k with $\tau_{k+1}(D)$. ◀

▶ **Lemma 19.** *Let $g : \mathbb{Z}^n \to \mathbb{Z}$ be a linear function. For the sets $I(g) := im_2(g)$ and $K(g) := g^{-1}(0) \cap \{0,1\}^n$ it holds that $|I(g)| \cdot |K(g)| \leq 3^n$.*

Proof. For every element $a \in I(g)$ choose some $v_a \in \{0,1\}^n$ such that $g(v_a) = a$. Consider the set $X := \{v_a + u\}_{a \in I(g), u \in K(g)} \subset \{0,1,2\}^n$.

It is easy to see that $|X| = |I(g)| \cdot |K(g)|$. Indeed, if $v_a + u = v_{a'} + u'$, then $g(v_a) + g(u) - g(0) = g(v_a + u) = g(v_{a'} + u') = g(v_{a'}) + g(u') - g(0)$ and therefore $a = a', v_a = v_{a'}, u = u'$.

On the other hand, $|X| \leq 3^n$. ◀

▶ **Lemma 20.** *Let $f = 1 + 2x_1 + \cdots + 2^n x_n$ and $g : \mathbb{Z}^n \to \mathbb{Z}$ be a linear function. For any $a \in \mathbb{Z} \backslash \{0\}$ one of the following holds:*
1. *$g = 0$ has at most $3^{\frac{n}{2}}$ 0-1 solutions.*
2. *$g + af = 0$ has at most $3^{\frac{n}{2}}$ 0-1 solutions.*

Proof. For every $b \in \mathbb{Z}$, there exists at most one boolean assignment that satisfies both $g = b$ and $b + af = 0$. Therefore the number of 0-1 solutions of $g + af = 0$ is at most the size of the boolean image $im_2(g)$ of g. By Lemma 19 either $|im_2(g)| \leq 3^{\frac{n}{2}}$ or $|g^{-1}(0) \cap \{0,1\}^n| \leq 3^{\frac{n}{2}}$. ◀

▶ **Theorem 21.** *Let $f = 1 + 2x_1 + \cdots + 2^n x_n$. Any $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ refutation of $f = 0$ is of size $2^{\Omega(n)}$.*

Proof. Define the predicate $\mathcal{P}(g)$ on linear polynomials over $\mathbb{Q}$ as follows: $\mathcal{P}(g) = 1$ iff $g = 0$ has at most $2^{(0.5 \cdot \log 3)n}$ 0-1 solutions. By Lemma 20, $\mathcal{P}$ satisfies the properties in Theorem 18. Therefore, by Theorem 18, if π is a refutation of $f = 0$, then there exists a derivation π' of some clause $C = \bigvee_{j \in [N]} g_j = 0$ from the boolean axioms, where each $g_j = 0$ has at most $2^{(0.5 \cdot \log 3)n}$ 0-1 solutions. Moreover $|\pi'| = O(n \cdot |\pi|^3)$. As C must be a boolean tautology, that satisfied by 2^n assignments, it must contain at least $2^{(1 - 0.5 \cdot \log 3)n}$ disjuncts (because every disjunct contributes at most $2^{(0.5 \cdot \log 3)n}$ satisfying assignments). Therefore $|\pi| = 2^{\Omega(n)}$. ◀

5 Tree-Like Lower Bounds

5.1 Nondeterministic Linear Decision Trees

In this section we extend the classical correspondence between tree-like resolution refutations and decision trees (cf. [6]) to tree-like $\mathrm{Res}(\mathrm{lin}_R)$ and tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$. We define *non-deterministic linear decision trees* (NLDT), which generalize parity decision trees, proposed in [17] 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 (though not for our dag-like 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 22** (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 $(\phi \cup \{x = 0 \vee x = 1 \mid x \in vars(\phi)\}, \Phi)$ 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 $\{x = 0 \vee x = 1 \mid x \in vars(\phi)\}$ and marking their roots with arbitrary clauses from ϕ.

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

▶ **Example 23.** Let ϕ be a set of clauses, representing unsatisfiable CNF. Then any standard decision tree on boolean variables is an NLDT for $\phi \cup \{x = 0 \vee x = 1 \mid x \in vars(\phi)\}$ of type $\mathrm{DT}(R)$, 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 24.** *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 $(\phi \cup \{x = 0 \vee x = 1 \mid x \in vars(\phi \cup \{\neg\Phi\})\}, \Phi)$ of size $O(2^n|\Phi|)$, where $n = |vars(\phi \cup \{\neg\Phi\})|$.*

Proof. Let $vars(\phi \cup \{\neg\Phi\}) = \{x_1, \ldots, x_n\}$ and fix an ordering on these variables. Construct a tree T_0 with 2^n nodes, that branches on $x_1, \ldots, x_n$, 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 $C = (f_1 = 0 \vee \cdots \vee f_m = 0)$ 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 $f_i = a_1x_1 + \ldots + a_nx_n + b$ then $a_1(1-\nu_1) + \ldots + a_n(1-\nu_n) + b \neq 0$ holds and the ith sequence is the following sequence of "substitutions": $(a_1x_1 + a_2(1-\nu_2) + \ldots + a_n(1-\nu_n) + b) + (a_1(1-\nu_1) - a_1x_1) \neq 0$ to $a_1x_1 + a_2(1-\nu_2) + \ldots + a_n(1-\nu_n) + b \neq 0$ and $a_1(1-\nu_1) - a_1x_1 \neq 0$, $\ldots$, $(a_1x_1 + \ldots + a_{n-1}x_{n-1} + a_n(1-\nu_n) + b) + (a_n(1-\nu_n) - a_nx_n) \neq 0$ 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 22. 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 $x_n \leftarrow 1 - \nu_n$, $\ldots$, $x_1 \leftarrow 1 - \nu_1$ to this non-equality.
In the leftmost node w in the end of the mth sequence, $\{f_1 \neq 0, \ldots, f_m \neq 0\} \subseteq \Phi_w$ holds and thus again C is violated at w in the sense of Definition 22 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 $g = a_1x_1 + \ldots + a_nx_n + b$. Attach to v a subtree that makes the following branches: $(a_1(1-\nu_1) + a_2x_2 + \ldots + a_nx_n + b) - (a_1(1-\nu_1) - a_1x_1) \neq 0$ to $(a_1(1-\nu_1) + a_2x_2 + \ldots + a_nx_n + b) \neq 0$ and $a_1(1-\nu_1) - a_1x_1 \neq 0$,$\ldots$, $(a_1(1-\nu_1) + \ldots + a_{n-1}(1-\nu_{n-1}) + a_n(1-\nu_n) + b) - (a_n(1-\nu_n) - a_nx_n) \neq 0$ 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 22.

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

▶ **Example 25.** Let ϕ be as in Example 23. *Parity decision trees*, as defined in [17], are NLDTs for ϕ of type $\mathrm{DT}_{sw}(\mathbb{F}_2)$: 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 26.** Let $\phi = \{f_1 = 0, \ldots, f_m = 0\}$, where $f_1, \ldots, f_m$ 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 $f_m \neq 0$, $\ldots$, $f_1 + f_2 \neq 0$ to $f_1 \neq 0$ and $f_2 \neq 0$.

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

▶ **Theorem 27.** *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* $\mathrm{Res}(\mathrm{lin}_R)$ *(resp. tree-like* $\mathrm{Res}_{sw}(\mathrm{lin}_R)$*) 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 $\mathrm{DT}_{sw}(R)$ for ϕ. We construct a tree-like $\mathrm{Res}(\mathrm{lin}_R)$ or tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ 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 22, $\neg\Phi_u$ must be a weakening of D (syntactic for $T_\phi \in \mathrm{DT}(R)$ and semantic for $T_\phi \in \mathrm{DT}_{sw}(R)$) 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 $T_\phi \in \mathrm{DT}(R)$, 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 $T_\phi \in \mathrm{DT}_{sw}(R)$, $\Phi_u \models \alpha f_u + \beta g_u \neq 0$, 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 $\mathrm{Res}(\mathrm{lin}_R)$ or a tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ 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 $\mathrm{Res}(\mathrm{lin}_R)$ or a tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ 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 $C_1, \ldots, C_k$ 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 22 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 22 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 $\phi \cup \{x = 0 \vee x = 1 \,|\, x \in \mathit{vars}(\phi)\} \cup \{0 = 0\}$, 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 22 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. ◀

An immediate corollary is the following:

▶ **Proposition 28.** *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 $n = \big|vars(\phi \cup \{C\})\big|$.*

Proof. By Proposition 24 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 27 there exists a tree-like $\mathrm{Res}(\mathrm{lin}_R)$ derivation of C from ϕ of size $O(2^n|C|)$. ◀

We construct an NLDT to prove the following upper bound:

▶ **Proposition 29.** *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 27 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: $s_n \leq 2^{2d_f} s_{\lfloor \frac{n}{2} \rfloor + 1} + s_f 2^{2d_f} = O(s_f n^{2d_f})$. ◀

5.2 Prover-Delayer Games

The *Prover-Delayer game* is an approach to obtain lower bounds on resolution refutations introduced by Pudlák and Impagliazzo [25]. 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 30) 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 30.** *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 $\mathrm{DT}(R)$ (respectively, $\mathrm{DT}_{sw}(R)$) 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. ◀

5.3 Lower Bounds for the Subset Sum with Small Coefficients

We now turn to tree-like lower bounds. In this section we prove tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ lower bound for $\mathsf{SubSum}(f)$ including instances, where coefficients of f are small, and tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ lower bound for $\mathsf{ImAv}\,(\pm x_1 \pm \cdots \pm x_n)$.

The proof of tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ lower bound for $\mathsf{SubSum}(f)$ goes in two stages. Assume f depends on n variables. First, as in the proof of dag-like lower bound in Sec. 4 we use Theorem 18 to transform refutations π of $f = 0$ to derivations π' of a clause C_π from only the boolean axioms. We ensure that π' is not much larger than π and C_π possesses the following property, which makes it hard to derive: for every disjunct $g = 0$ in C_π the linear polynomial g depends on at least $\frac{n}{2}$ variables. Second, we use Prover-Delayer games to prove the lower bound for derivations of any clause with this property. The proof that Delayer's strategy succeeds to earn sufficiently many coins is guaranteed by a bound on size of essential coverings of hypercubes.

▶ **Definition 31.** *Let $\mathcal{H}$ be a set of hyperplanes in $\mathbb{Q}^n$. We say that $\mathcal{F}$ forms **essential cover** of the cube $B_n = \{0,1\}^n$ if:*

- *Every point of B_n is covered by some hyperplane in $\mathcal{H}$.*
- *No proper subset $\mathcal{H}' \subsetneq \mathcal{H}$ covers B_n.*
- *No axis in $\mathbb{Q}^n$ is parallel to all hyperplanes in $\mathcal{H}$. In other words, if $\mathcal{H} = \{H_1, \ldots, H_m\}$ and $f_i = 0$ is the linear equation defining H_i, $i \in [m]$, then every variable x_j, $j \in [n]$, occurs with nonzero coefficient in some f_i.*

▶ **Theorem 32** ([21]). *Any essential cover of the cube B_n in $\mathbb{Q}^n$ must contain at least $\frac{1}{2}(\sqrt{4n+1}+1)$ hyperplanes.*

We use Prover-Delayer games to prove the lower bound below.

▶ **Theorem 33.** *Any tree-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ *derivation of any tautology of the form* $\bigvee_{j\in[N]} g_j = 0$*, for some positive* N*, where each* g_j *is linear over* $\mathbb{Q}$ *and depends on at least* $\frac{n}{2}$ *variables, is of size* $2^{\Omega(\sqrt{n})}$.

Proof. According to the definitions in Sec. 5.2 the corresponding Prover-Delayer game is on $0 = 0$ and starts with the position

$$\Phi_r = \{g_j \neq 0 \mid j \in [N]\}\,.$$

The game finishes at a position Φ, where $\{x_i \neq 0,\ x_i \neq 1\} \subseteq \Phi$ for some $i \in [n]$ or $0 \neq 0 \in \Phi$.

We now define a Delayer's strategy that guarantees $\Omega(\sqrt{n})$ coins and by Lemma 30 obtain the lower bound.

If Φ is a position in the game, denote by $\Phi_c \subset \Phi$ the subset of so-called "coin" non-equalities, that is, non-equalities that were chosen by Prover when Delayer decided to leave the choice to Prover and earn a coin. The number $|\Phi_c|$ is then precisely the number of coins earned by Delayer at Φ. Throughout the game Delayer constructs a partial assignment ρ_I for variables in $I \subseteq [n]$ and a set of non-equalities $\Phi_I \subseteq \Phi_c$, such that:

1. $|\Phi_I| = \Omega(\sqrt{|I|})$;
2. for all $g \neq 0 \in (\Phi \restriction_{\rho_I}) \setminus (\Phi_c \restriction_{\rho_I})$, the function g depends on at least $\frac{n}{2} - |I|$ variables;
3. Φ_I contains variables only from I; and
4. $\Phi_c \restriction_{\rho_I}$ is 0-1 satisfiable.

In the beginning both ρ_I and Φ_I are empty.

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 $g_2 \neq 0$, where $g_1 + g_2 \neq 0 \in \Phi$. Delayer does the following. Before making any decision Delayer checks if there exists some nonconstant linear g with variables in $[n] \setminus I$ such that $(\Phi_c \restriction_{\rho_I}) \cup \{g \neq 0\}$ is unsatisfiable over 0-1.

In case it holds, $\Psi := (\Phi_c \setminus \Phi_I) \restriction_{\rho_I} \cup \{g \neq 0\}$ must be 0-1 unsatisfiable. Consider a minimal subset $\Psi' \subseteq \Psi$ such that Ψ' is 0-1 unsatisfiable and denote $I' \subseteq [n]$ the set of variables that occur in Ψ'. As $\Psi'' := \Psi' \setminus \{g \neq 0\}$ is 0-1 satisfiable, there exists an assignment $\rho_{I'}$ for variables in I', that satisfies Ψ''. Delayer extends the assignment ρ_I with $\rho_{I'}$ to $\rho_{I\cup I'}$ and defines $\Phi_{I\cup I'} := \Phi_I \cup \Psi''$.

If $\Psi' = \{g_1 \neq 0, \ldots, g_k \neq 0\}$, then the hyperplanes $H_1, \ldots, H_k$ defined by the equations $g_1 = 0, \ldots, g_k = 0$ form an essential cover of the cube $B_{|I'|}$. Therefore, by Theorem 32, $|\Psi''| = |\Psi'| - 1 \geq \frac{1}{2} \cdot \sqrt{|I'|}$ and thus $|\Phi_{I\cup I'}| \geq \frac{1}{2} \cdot \sqrt{|I|} + \frac{1}{2} \cdot \sqrt{|I'|} \geq \frac{1}{2} \cdot \sqrt{|I \cup I'|}$.

If necessary, Delayer repeats the above procedure constructing extensions $\rho_{I_1} \subset \cdots \subset \rho_{I_L}$ and $\Phi_{I_1} \subset \cdots \subset \Phi_{I_L}$, where $I_1 = I \subset \ldots \subset I_L$, until there is no $g \neq 0$ inconsistent with $\Phi_c \restriction_{\rho_{I_L}}$ as described above. The new value of I is set to I_L. After that Delayer does the following:

1. if $g_1 \restriction_{\rho_I} = 0$, then choose $g_2 \neq 0$;
2. otherwise, if $g_2 \restriction_{\rho_I} = 0$, then choose $g_1 \neq 0$;
3. if none of the above cases hold, leave the choice to Prover and earn a coin.

Denote by Φ' and $\Phi'_c \subseteq \Phi'$ the new position and the subset of "coin" non-equalities, respectively, after the choice is made. It is easy to see that the property that any $g \neq 0 \in (\Phi' \restriction_{\rho_I}) \setminus (\Phi'_c \restriction_{\rho_I})$ depends on at least $\frac{n}{2} - |I|$ variables still holds.

It follows from the definition of Delayer's strategy that Φ_c is always 0-1 satisfiable. Therefore if Φ is the endgame position, that is if $0 \neq 0 \in \Phi$ or $\{x_i \neq 0, x_i \neq 1\} \subset \Phi$ for some $i \in [n]$, then $0 \neq 0 \in (\Phi\restriction_{\rho_I}) \setminus (\Phi_c\restriction_{\rho_I})$ or $\{x_i \neq 0, x_i \neq 1\} \subset (\Phi\restriction_{\rho_I}) \setminus (\Phi_c\restriction_{\rho_I})$ respectively. This implies that $|I| \geq \frac{n}{2} - 1$ and therefore $|\Phi_c| \geq |\Phi_I| \geq \frac{1}{2} \cdot \sqrt{|I|} = \Omega(\sqrt{n})$. Thus the number of coins earned by Delayer is $\Omega(\sqrt{n})$. ◀

▶ **Corollary 34.** *Let f be any linear polynomial over $\mathbb{Q}$ that depends on n variables. Then tree-like* $\text{Res}(\text{lin}_{\mathbb{Q}})$ *derivations of* $\mathsf{Im}(f)$ *are of size* $2^{\Omega(\sqrt{n})}$.

▶ **Theorem 35.** *If f is a linear polynomial over $\mathbb{Q}$, which depends on n variables and $0 \notin im_2(f)$, then every tree-like* $\text{Res}(\text{lin}_{\mathbb{Q}})$ *refutation of $f = 0$ is of size* $2^{\Omega(\sqrt{n})}$.

Proof. Consider the following predicate $\mathcal{P}$ on linear polynomials: $\mathcal{P}(g) = 1$ iff g depends on at least $\frac{n}{2}$ variables. It is easy to see that $\mathcal{P}$ satisfies the conditions in Theorem 18 with respect to f. Therefore by Theorem 18 for every refutation π of $f = 0$ there exists a derivation π' of a clause C_π from the boolean axioms such that $|\pi'| = O(n \cdot |\pi|^3)$ and $\mathcal{P}(g)$ for every $g = 0$ in C_π. Thus, by Theorem 33 $|\pi'| = 2^{\Omega(\sqrt{n})}$ and $|\pi| = 2^{\Omega(\sqrt{n})}$. ◀

▶ **Lemma 36.** *Let Φ be a satisfiable system of m non-equalities over $\mathbb{F}$. 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 $\Phi = \{\overline{a}_1 \cdot \overline{x} + b_1 \neq 0, \ldots, \overline{a}_m \cdot \overline{x} + b_m \neq 0\}$ 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 [2], 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 37.** *We work over $\mathbb{Q}$. Let $f = \epsilon_1 x_1 + \cdots + \epsilon_n x_n$, where $\epsilon_i \in \{-1, 1\}$. Then any tree-like* $\text{Res}_{sw}(\text{lin}_{\mathbb{Q}})$ *refutation of* $\mathsf{ImAv}(f)$ *is of size at least* $2^{\frac{n}{4}}$.

Proof. According to the definitions in Sec. 5.2 the corresponding Prover-Delayer game is on $\mathsf{ImAv}(f)$ and starts with the empty position. The game finishes at a position Φ, where $\Phi \models f - A = 0$ for some $A \in im_2(f)$.

We now define a Delayer's strategy that guarantees $\frac{n}{4}$ coins and by Lemma 30 obtain the lower bound.

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 $g_2 \neq 0$, where $\Phi \models g_1 + g_2 \neq 0$. Delayer does the following:

1. if $g_2 \neq 0$ is inconsistent with Φ, but $g_1 \neq 0$ is consistent with Φ, then choose $g_1 \neq 0$;
2. if $g_1 \neq 0$ is inconsistent with Φ, but $g_2 \neq 0$ is consistent with Φ, then choose $g_2 \neq 0$;
3. if none of the above holds, then leave the choice to the Prover and earn a coin.

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

Suppose that the game has finished at a position Φ. The strategy of Delayer guarantees that Φ is satisfiable and Φ contradicts a clause $\langle f \neq A \rangle$ of $\mathsf{ImAv}(f)$, that is $\Phi \models f - A = 0$ for some $A \in im_2(f)$. Let $\zeta_1, \ldots, \zeta_\ell$ be the set of non-equalities in Φ, in the order they were added to Φ. Let $\Psi \subseteq \Phi$ be the set of all ζ_i, $i \in [\ell]$, such that ζ_i is not implied by previous non-equalities ζ_j, for $j < i$. Then, Delayer earns at least $|\Psi|$ coins, $\Psi \models f = A$, and by Lemma 36 we conclude that $|\Psi| \geq \frac{n}{4}$. ◀

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 $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ (Theorem 17) this provides a separation between tree-like and dag-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ in the case $char(\mathbb{F}) = 0$, for formulas in CNF. 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 [17]. However, the proof that Delayer's strategy guarantees sufficiently many coins relies on Lemma 39, which is a generalization of Lemma 3.3 in [17] for arbitrary fields. Since the proof of Lemma 3.3 in [17] for the $\mathbb{F}_2$ case does not apply to arbitrary fields, our proof is different, and uses a result from Alon-Füredi [4] on the hyperplane coverings of the hypercube.

▶ **Theorem 38.** *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 [17], 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 39.** *Let $A\overline{x} \dot{=} \overline{b}$ 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, $0 \dot{=} \overline{b}$. 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 to the system $A\overline{x} \dot{=} \overline{b}$ 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 [4]:

▶ **Corollary from Alon-Füredi [4, 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 $\sum_{i=1}^n y_i \geq l\}$. 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 from B_n, 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 $M(n+1) = 2$. ◀

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

▶ **Lemma 40.** *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. 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

$$\mathbf{v}_{[n]\setminus I} \cdot (\alpha \oplus \beta)_{[n]\setminus I} = \mathbf{v}_{[n]\setminus I} \cdot \alpha_{[n]\setminus I} \,. \tag{9}$$

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

$$\mathbf{v}_I \cdot (\alpha \oplus \beta)_I = \mathbf{v}_I \cdot \alpha_I + \mathbf{v}_I^\star \cdot \beta_I \,. \tag{10}$$

Equation (10) 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 (10). 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 (10) 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 (10).

The two equations (9) and (10) concludes the 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 39, 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 $A_I^\star\beta_I \doteqdot \overline{b} - A\alpha$. ◀

▶ **Lemma 41.** *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 40 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 40. (As before, flipping from one to zero maintains the properness of the solution.) Hence, the resulting proper solution must satisfy the clause $\bigvee_{j\in[n]} x_{i,j}$. ◀

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 41.

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

6 Size-Width Relation and Simulation by Polynomial Calculus

In this section we prove a size-width relation for tree-like $\mathrm{Res}(\mathrm{lin}_R)$ (Theorem 44), 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 5). 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 [2] 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 46 and 47).

▶ **Proposition 42.** *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 10. 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 $\omega_0(\pi) \leq \omega_0(\phi) + 2$. ◀

▶ **Corollary 43.** *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 42 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 $\max(\omega_0(\pi), \omega_0(\phi) + 2)$. ◀

▶ **Theorem 44.** *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 $\phi = \{C_1, \dots, C_m\}$ 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,

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

for some $l(x_1, \dots, x_{n-1}) = \sum_{i=1}^{n-1} a_i x_i + B$, 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

$$\phi \restriction_{\rho_A} \cup \left\{(A - l)a_n^{-1} = 0 \vee (A - l)a_n^{-1} = 1\right\} \tag{11}$$

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 (11) 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 43 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 12 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 43 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 45.** *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 44 and 45, and the relation $\omega_0 \geq \frac{1}{|\mathbb{F}|}\omega$ as well as the results from [2], we have the following:

▶ **Corollary 46.** *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 $\neg TS_{G,\sigma}^{(p)} \in UNSAT$, every tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutation of $\neg TS_{G,\sigma}^{(p)}$ has size $2^{\Omega(dn)}$.*

Proof. Corollary 4.5 from [2] states that the degree of $PC_{\mathbb{F}}$ refutations of $\neg\mathrm{TS}^{(p)}_{G,\sigma}$ is $\Omega(dn)$. Theorem 45 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 44 the size is $2^{\Omega(dn)}$. ◀

▶ **Corollary 47.** *Let $\phi \sim \mathcal{F}_k^{n,\Delta}, k \geq 3$ 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 [2] 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 45 implies that the principal width of $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutations of $\phi \sim \mathcal{F}_k^{n,\Delta}$ is $\Omega(\frac{1}{|\mathbb{F}|}dn) = \Omega(dn)$ and thus by Theorem 44 the size of the refutations is $2^{\Omega(dn)}$ with probability $1 - o(1)$. ◀

7 Conclusion

By the discussion in Sec. 1.1.4, for finite fileds we can take any CNF $\phi(\overline{x})$ known to be hard for $PC_{\mathbb{F}}$ (e.g. Tseitin formulas, random CNFs etc) and turn it into the linear system $R_\phi(\overline{x}, \overline{y})$, which we can prove is hard for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$. It is reasonable to conjecture that these linear systems are also hard for dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ and to try to prove a lower bound for them. However, this would require dealing with particular systems, arising from these CNFs and, therefore, having a specific structure. Alternatively, we may turn our attention to fields $char(\mathbb{F}) = 0$: the hard instance in this case can be chosen freely among systems $L(\overline{x})$, where all coefficients are bounded by a constant: every equation in $L(\overline{x})$ can be coded as a short CNF formula, which admits short $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ derivations from $L(\overline{x})$. $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ lower bound for such a linear system would imply $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ CNF lower bound if $char(\mathbb{F}) = 0$ and by generalization of the proof of simulation of $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$ by $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ in [17] to arbitrary finite fields, this would imply CNF $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}'})$ bounds for any finite field $\mathbb{F}'$.

Thus, for any field $\mathbb{F}$, the general dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ lower bound problem for CNF can be reduced to the following problem: find a 0-1 unsatisfiable linear system $L(\overline{x})$ over $\mathbb{Z}$ with coefficients bounded by a constant such that any $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ refutation of $L(\overline{x})$ is of superpolynomial size.

Note that $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ is pretty strong proof system: classical tautologies such as Pigeonhole Principle, Clique-Coclique Principle or (mod p)-Tseitin Tautologies are all easy for $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$[26]. Therefore, even indentifying explicit hard candidate for $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ is a non-trivial problem. Linear systems in many respects are more handy to work with while indentifying hardness conditions as well as analysing structure of $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ proofs.

References

1 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.

2 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.

3 Yaroslav Alekseev, Dima Grigoriev, Edward A. Hirsch, and Iddo Tzameret. Semi-Algebraic Proofs, IPS Lower Bounds and the τ-Conjecture: Can a Natural Number be Negative? *Electronic Colloquium on Computational Complexity TR19-142*, 2019.

4 Noga Alon and Zoltán Füredi. Covering the Cube by Affine Hyperplanes. *Eur. J. Comb.*, 14(2):79–83, March 1993. doi:10.1006/eujc.1993.1011.

5 Paul Beame, Noah Fleming, Russell Impagliazzo, Antonina Kolokolova, Denis Pankratov, Toniann Pitassi, and Robert Robere. Stabbing Planes. In Anna R. Karlin, editor, *9th Innovations in Theoretical Computer Science Conference (ITCS 2018)*, volume 94 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10:1–10:20, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.ITCS.2018.10.

6 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. doi:10.1613/jair.1410.

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

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

9 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).

10 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.

11 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 Reckhow [27]. doi:10.2307/2273702.

12 Michael A. Forbes, Amir Shpilka, Iddo Tzameret, and Avi Wigderson. Proof Complexity Lower Bounds from Algebraic Circuit Complexity. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 32:1–32:17, 2016. doi:10.4230/LIPIcs.CCC.2016.32.

13 Michal Garlik and Lezsek Kołodziejczyk. Some Subsystems of Constant-Depth Frege with Parity. *ACM Transactions on Computational Logic*, 19(4), 2018. URL: https://www.mimuw.edu.pl/~lak/jansparity.pdf.

14 Joshua A. Grochow and Toniann Pitassi. Circuit Complexity, Proof Complexity, and Polynomial Identity Testing: The Ideal Proof System. *J. ACM*, 65(6):37:1–37:59, 2018. doi:10.1145/3230742.

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

16 J. Hastad. On Small-Depth Frege Proofs for Tseitin for Grids. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 97–108, Los Alamitos, CA, USA, October 2017. IEEE Computer Society. doi:10.1109/FOCS.2017.18.

17 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. doi:10.1007/978-3-662-44465-8_32.

18 Jan Krajícek. A feasible interpolation for random resolution. *Logical Methods in Computer Science*, 13(1), 2017. doi:10.23638/LMCS-13(1:5)2017.

19 Jan Krajícek and Igor Carboni Oliveira. On monotone circuits with local oracles and clique lower bounds. *Chicago J. Theor. Comput. Sci.*, 2018, 2018. URL: http://cjtcs.cs.uchicago.edu/articles/2018/1/contents.html.

20 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.

21 Nathan Linial and Jaikumar Radhakrishnan. Essential covers of the cube by hyperplanes. *Journal of Combinatorial Theory, Series A*, 109:331–338, 2005.

22 A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, September 1988. doi:10.1007/BF02126799.

23 Jakob Nordström. On the Interplay Between Proof Complexity and SAT Solving. *ACM SIGLOG News*, 2(3):19–44, August 2015. `doi:10.1145/2815493.2815497`.

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

25 Pavel Pudlák and Russell Impagliazzo. A lower bound for DLL algorithms for k-SAT (preliminary version). In *Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms, January 9-11, 2000, San Francisco, CA, USA.*, pages 128–136, 2000. URL: `http://dl.acm.org/citation.cfm?id=338219.338244`.

26 Ran Raz and Iddo Tzameret. Resolution over linear equations and multilinear proofs. *Ann. Pure Appl. Logic*, 155(3):194–224, 2008. `doi:10.1016/j.apal.2008.04.001`.

27 Robert Reckhow. *On the lengths of proofs in the propositional calculus*. PhD thesis, University of Toronto, 1976. Technical Report No . 87.

28 Grigori Tseitin. *On the complexity of derivations in propositional calculus*, pages 466–483. Studies in constructive mathematics and mathematical logic Part II. Consultants Bureau, New-York-London, 1968.