Complexity of Propositional Proofs with Counting: Resolution over Linear Equations and Semi-Algebraic Proofs

Fedor Part

Royal Holloway, University of London

A thesis submitted for the degree of

Doctor of Philosophy

2018

Acknowledgements

I would like to thank Iddo Tzameret for his invaluable support and supervision.

Abstract

Propositional proof complexity studies the hardness of certifying that propositional statements are tautologies. One of its most important concerns is to answer the following question: is it possible to specify what counts as a proof of a tautology in such a way that all proofs are polynomial-time checkable and for every tautology there exists a proof of polynomial size? The negative answer to this question would imply the separation of complexity classes coNP and NP and, therefore, P and NP. Proof sizes have been extensively studied for various proof systems and for a number of weak systems the question above has been answered negatively. Many known examples of hard formulas, for which superpolynomial lower bounds on sizes of proofs in weak systems were obtained, are based on counting principles such as, for example, the Pigeonhole Principle or unsolvable linear systems. This is one of the reasons why exploration of the power of stronger proof systems that, loosely speaking, "can count" is one of the central topics in proof complexity.

In the first part of the thesis we develop new lower bounds techniques for resolution over linear equations and extend existing ones to work over different rings. We obtain a host of new lower bounds, separations and upper bounds, while calibrating the relative strength of different sub-systems. We first establish, over fields of characteristic zero, exponential-size *dag-like* lower bounds against resolution over linear equations refutations of instances with large coefficients. Specifically, we demonstrate that the subset sum principle $\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta$, for β not in the image of the linear form, requires refutations proportional to the size of the image. Moreover, for instances with small coefficients, we separate the tree and dag-like versions of $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$, when $\mathbb{F}$ is of characteristic zero, by employing the notion of essential covering of the hypercube from [48], among other techniques.

We then study resolution over linear equations over different *finite* fields, extending the work of Itsykson and Sokolov [40] who developed tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$ lower bounds techniques. We obtain new lower bounds and separations as follows: (**i**) exponential-size lower bounds for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_p})$ refutations of Tseitin mod q formulas, for every pair of distinct primes p, q. As a corollary we obtain an exponential-size separation between tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_p})$ and tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_q})$; (**ii**) exponential-size lower bounds for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_p})$ refutations of random k-CNF formulas, for every prime p and constant k; and (**iii**) exponential-size lower bounds for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutations of the pigeonhole principle, for *every* field $\mathbb{F}$.

Another important aspect of proof complexity is the study of weak "counting" proof systems, for which exponential lower bounds are already known, and specifically algebraic proof systems based on Hilbert's Nullstellensatz and semi-algebraic proof systems. Algorithms for certain central problems give rise to proof searching algorithms for these systems. For example, Groebner basis computation is dual to Polynomial Calculus ($\mathsf{PC}_{\mathcal{R}}$) proof search, while the branch-and-cut algorithm for integer programming is dual to Cutting Planes proof search and so on.

The so-called *Sum-of-Squares (SoS) (meta-)algorithm*, which is in duality with the SoS proof system, has become a cornerstone in the study of the complexity of optimization problems due to its conjectured optimality for a large class of problems. Analysis of the SoS algorithm often reduces to finding proofs of statements from, say, Boolean analysis in small degree SoS, which is a *propositional* proof system. Propositional proofs are complicated combinatorial objects and they can be hard to construct and understand.

In the second part of the thesis we investigate a formulation of first-order theories $\mathsf{TPC}_{\mathcal{R}}$ and TSoS that would correspond, via translation a la Paris-Wilkie, to constant degree $\mathsf{PC}_{\mathcal{R}}$ and SoS, respectively. We define these theories and show, that proofs in $\mathsf{TPC}_{\mathcal{R}}$ admit propositional translation to constant degree $\mathsf{PC}_{\mathcal{R}}$ if $\mathcal{R}$ is a field of positive characteristic. For all other rings $\mathcal{R}$ we define a translation to the constant degree $\mathsf{PC}^{\mathrm{rad}}_{\mathcal{R}}$, which extends $\mathsf{PC}_{\mathcal{R}}$ with the *radical rule*.

Contents

Chapter 1

Introduction

1.1 Background

A fundamental concept, arising in many areas of science and engineering, is that of an *algorithm*. Informally, an algorithm is a sequence of steps, required to perform certain task. A simple example – the task of multiplication of natural numbers. There is an elementary algorithm, learned by school children - a sequence of shifts and additions of digits (with carry), resulting in the product. If it starts with two n-digit numbers, the number of steps would be proportional to n^2, that is $O(n^2)$. But is there a better algorithm, performing less amount of shifts and additions of digits? It turns out, that there exists non-elementary algorithm, based on Fast Fourier Transform, performing $O(n \cdot \log n \cdot \log(\log n))$ steps [30]. There are number of algorithms, which slightly improve on this bound, but never do better than $O(n \log n)$. Thus, it is reasonable to ask whether we can prove that any other algorithm would perform at least $\Omega(n \cdot \log n)$ steps.

This kind of questions are addressed by *computational complexity* theory, which studies resources necessary for algorithms that solve a computational problem and classifies the problems accordingly. The mathematical study of computation is based on rigorous formulation of what an algorithm is. There is a number of mathematically precise definitions, that turn out to be equivalent to each other [37]. The most prominent of them is that of *Turing machine*, which closely matches the intuition: it consists of an *infinite tape*[1], where symbols from a finite alphabet can be read and written in line by the *head*, and a finite number of *states*, each of which contain specification either of whether the execution should halt or of what should be done by the head at the current step and which other state to jump to depending on what

[1] Or several tapes.

has been read at the current position of the head [7]. At the start, the tape contains an input of a computational problem and the output of the execution is defined to be the content of the tape once the execution has halted or is undefined in case the execution never halts. This simple mathematical abstraction is powerful enough to capture informal descriptions of algorithms as well as anything that can be written in programming languages as expressive as, for example, Kotlin or C.

Typically a computational problem is either a decision problem or can be efficiently reduced to one. A *decision problem* is a problem of the following form: given an input encoded as a finite bit string $x \in \{0,1\}^*$, decide whether x satisfies some property and output 1 or 0 if the answer is "yes" or "no", respectively. Clearly, the statement of a decision problem is just a subset $\mathcal{L} \subseteq \{0,1\}^*$, called *language*, and can be identified with it.

One of the main goals of complexity theory is to determine for a given decision problem $\mathcal{L}$ whether efficient algorithms exist for $\mathcal{L}$ according to some measure of complexity of algorithms. The most important such measure is *the worst-case time complexity* of an algorithm $\mathcal{A}$ – the function $f(n)$ such that the maximal number of steps the Turing machine, corresponding to $\mathcal{A}$, performs on inputs of length n is $f(n)$.

Consider the following problem. Suppose there are n cities and we are given distances between any two of them. The *Traveling Salesman Problem* asks to find a circular path of minimal length such that it passes through all the cities at most once [6]. It is easy to see, that this search problem can be reduced to the decision problem of checking whether there exists a path as above of length at most L by using binary search on L. The trivial brute force algorithm, just enumerating all paths and comparing their lengths, is of enormous time complexity $O(n!)$, clearly, it is highly infeasible. There are a bit smarter algorithms of complexity $O(2^n)$ [6], but they still can hardly be called feasible: already for 100 cities $2^{100} \cong 10^{30}$ steps seem too much even for modern supercomputers to be performed in a reasonable time.

It is a common agreement in complexity theory to classify an algorithm as feasible if its time complexity $f(n)$ grows polynomially, that is $f(n) = O(n^c)$ for some $c \in \mathbb{N}$. For that reason the class P of all decision problems, that can be solved by some polynomial-time Turing machine, plays an important role in complexity theory. Whether defined above TSP decision problem is solvable by a feasible algorithm is thus the question of whether TSP $\in \mathsf{P}$. The answer is not known and this question is of tremendous importance as we explain below.

Whether TSP $\in \mathsf{P}$ is not an isolated question, there is a vast number of important decision problems that admit polynomial-time reductions to and from TSP [32].

Although these problems are seemingly unrelated, they bear the following similarity: for each such problem $\mathcal{L}$ there exists a notion of effectively checkable "solution" such that $x \in \mathcal{L}$ iff there exists a "solution" $y \in \{0,1\}^*$ that can be checked in time polynomial in $|x|$. For example, for TSP a solution is any circular path of length at most L passing through all the cities at most once. Formally, the class NP of decision problems with a notion of effectively checkable "solution" or certificate can be defined as follows: $\mathcal{L} \in \mathsf{NP}$ iff there exists a polynomial-time Turing machine $M(x,y)$ and $c \in \mathbb{N}$ such that $x \in \mathcal{L}$ iff there exists $y \in \{0,1\}^*, |y| = O(|x|^c)$ such that $M(x,y) = 1$. It is easy to see that $\mathsf{P} \subseteq \mathsf{NP}$. Also, TSP $\in \mathsf{NP}$ and it possess the property, called NP-*hardness*, that every problem in NP is polynomial-time reducible to it. A decision problem $\mathcal{L}$ is called NP-*complete* iff $\mathcal{L} \in \mathsf{NP}$ and $\mathcal{L}$ is NP-hard. Thus, TSP is NP-complete and, as mentioned above, there are many other important problems that are NP-complete. If any of these problems is shown to be in P, then $\mathsf{P} = \mathsf{NP}$ and therefore, informally, whenever a solution to a problem can be checked effeciently it can be also found effeciently. Apart from the common sense, stating that it should be much more hard to find a solution than just to check it, there is a plenty of evidence in complexity theory that supports the conjecture $\mathsf{P} \neq \mathsf{NP}$. Proving $\mathsf{P} \neq \mathsf{NP}$ is one of the major open problems in complexity theory [7].

The canonical NP-complete problems are: CIRCUIT-SAT - satisfiability of a formula (or circuit) of propositional logic, and (k-)SAT - satisfiability of a (k-)CNF formula, $k \geq 3$. There are rather trivial reductions from CIRCUIT-SAT to (k-)SAT and vise versa. On the one hand SAT has rather simple statement, on the other hand it is universal, because it is often easy to reduce a specific NP problem to SAT. These two qualities of SAT have been motivating an active development of SAT-solving algorithms [51].

1.1.1 Proof Complexity

The connection of computational complexity to propositional proof complexity becomes apparent once we turn our attention to the class $\mathsf{coNP} := \{\overline{\mathcal{L}} \,|\, \mathcal{L} \in \mathsf{NP}\}$ - the class dual to NP. Consider the language of all unsatisfiable CNF formulas UNSAT $= \overline{\text{SAT}} \in \mathsf{coNP}$ and the language TAUT $\in \mathsf{coNP}$ of all tautological DNF formulas. Clearly, these two languages are coNP-complete and reducible to each other via $\phi \mapsto \neg\phi$.

Observe, that, by definition of NP, a language $\mathcal{L} \in \mathsf{NP}$ iff there is a way to certify $x \in \mathcal{L}$ by some polynomial-time checkable proof π of size bounded by a polynomial on $|x|$. In [28] Cook and Reckhow suggested to define a *proof system* for a language $\mathcal{L}$ as a polynomial-time Turing machine V such that $x \in \mathcal{L}$ iff there exists a V-proof π for x,

namely π such that $V(x, \pi) = 1$. For example, standard propositional proof systems like the sequent calculus PK or Hilbert-style systems are Cook-Reckhow systems for TAUT. Also, Cook-Reckhow proof system for UNSAT, which are called *refutation systems*, can be trivially interpreted as propositional proof systems via a trivial bijection between UNSAT and TAUT. The condition that $\mathcal{L} \in \mathsf{NP}$ is then equivalent to the existence of a proof system V for $\mathcal{L}$ such that any $x \in \mathcal{L}$ can be certified by a polynomial-size proof π in V. In this case V is called *polynomially bounded* proof system. Thus, in particular, TAUT $\in \mathsf{NP}$ iff there exists polynomially bounded propositional proof system. The negation of the latter condition implies TAUT $\notin \mathsf{NP}$, therefore $\mathsf{coNP} \neq \mathsf{NP}$ and $\mathsf{P} \neq \mathsf{NP}$. This means that by proving superpolynomial lower bounds on lengths of proofs in stronger and stronger propositional proof systems we get supposedly closer to proving $\mathsf{P} \neq \mathsf{NP}$.

1.1.1.1 Resolution and Its Extensions

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 general, unrestricted resolution refutations are refered to as *dag-like* refutations.

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. DPLL algorithms are simple recursive algorithms for solving SAT. 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.

Modern SAT-solvers are quite sophisticated and employ advanced techniques, which are beyond the scope of DPLL algorithms. For example, CDCL algorithms try to avoid deriving same clauses several times by using clause learning techniques. Such algorithms produce dag-like resolution refutations on unsatisfiable formulas and,

thus, dag-like resolution lower bounds imply lower bounds on the run-time of these algorithms (cf. [51]).

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_n^m, expressing that there is no (total) injective map from a set with cardinality m to a set with cardinality n if $m > n$ [36]. 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 [64].

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

One such extension of resolution was introduced by Raz and Tzameret [60] 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 $\mathbb{Z}$ and which contains Boolean axioms for variables $x_i = 0 \vee x_i = 1$, was studied in [60]. This work demonstrated the power of resolution with counting over the integers, and specifically provided polynomial upper bounds for the pigeonhole principle and the Tseitin formulas, as well as other basic counting formulas. It also established exponential lower bounds for a subsystem of R(lin), denoted $\mathrm{R}^0(\mathrm{lin})$. Subsequently, Itsykson and Sokolov [40] studied resolution over linear equations over $\mathbb{F}_2$, denoted $\mathrm{Res}(\oplus)$. They demonstrated the power of resolution with counting mod 2 as well as its limitations by means of several upper and tree-like lower bounds. Moreover, [40] introduces DPLL algorithms, which can "branch" on arbitrary linear forms over $\mathbb{F}_2$, as well as parity decision trees, and showed a correspondence between parity decision trees and tree-like $\mathrm{Res}(\oplus)$ refutations. In both [60] and [40] the dag-like lower bound question for resolution over linear equations remained open.

As it happens, resolution over linear equations, holds a special place in the theory of proof complexity: it can be viewed as a natural "minimal" subsystem of important propositional proof systems, as we now explain. Resolution operates

with clauses, which are De Morgan formulas ($\neg$, unbounded fan-in $\vee$ and $\wedge$) of a particular kind, namely, of depth 1. Thus, from the perspective of the theory of proof complexity, resolution is a fairly weak version of the propositional-calculus, where the latter operates with arbitrary De Morgan formulas. Under a natural and general definition, propositional-calculus systems go under the name *Frege systems*: they can be (axiomatic) Hilbert-style systems or sequent-calculus style systems. The task of proving lower bounds for general Frege systems is notoriously hard: no nontrivial lower bounds are known to date. Basically, the strongest fragment of Frege systems, for which lower bounds are known are $\mathbf{AC}^0$-**Frege** systems, which are Frege proofs operating with constant-depth formulas. For example, both PHP_n^m and TS_G do not admit sub-exponential proofs in $\mathbf{AC}^0$-**Frege** [1, 55, 47, 15]. However, if we extend the De Morgan language with counting connectives such as unbounded fan-in mod p ($\mathbf{AC}^0[p]$-Frege) or threshold gates ($\boldsymbol{TC}^0$-Frege), then we step again into the darkness: proving super-polynomial lower bounds for these systems is a long-standing open problem on what can be characterized as the "frontiers" of proof complexity. In this sense, resolution over linear equations over prime fields and over the integers is interesting as a first step towards $\mathbf{AC}^0[p]$-Frege lower and $\boldsymbol{TC}^0$-Frege lower bounds, respectively. Works by Krajíček [43], Garlik-Kołodziejczyk [33] and Krajíček-Oliveira [44] had suggested possible approaches to attack dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$ lower bounds.

1.1.1.2 Algebraic and Semi-algebraic Proof Systems

Algebraic proof systems arise as ways of certifying unsolvability of systems of polynomial equations over a ring or a field. One of such ways to certify unsolvability of a system $\mathcal{F} = \{f_1 = 0, \ldots, f_m = 0\}$ of polynomial equations over a field $\mathbb{F}$ is based on a weak version of Hilbert's Nullstellensatz Theorem [9]. It follows from this theorem that $\mathcal{F}$ has no solutions over algebraic closure of $\mathbb{F}$ iff there exist polynomials $g_1, \ldots, g_m$ over $\mathbb{F}$ such that $f_1 \cdot g_1 + \ldots + f_m \cdot g_m = 1$. These tuples of polynomials $(g_1, \ldots, g_m)$ are thus proofs of unsolvability of $\mathcal{F}$ and the corresponding proof system $\mathsf{NS}_{\mathbb{F}}$ is called *Nullstellensatz system* [13]. It is naturally a Cook-Reckhow proof system for the coNP-complete language of unsolvable systems of polynomial equations over $\mathbb{F}$.

$\mathsf{NS}_{\mathbb{F}}$ is also a Cook-Reckhow propositional proof system for (k-)UNSAT: if $\phi = \{C_1, \ldots, C_m\}$ is a set of clauses with variables $x_1, \ldots, x_n$, then ϕ is unsatisfiable iff the system of polynomial equations $a(C_1) = 0, \ldots, a(C_m) = 0, x_1^2 - x_1 = 0, \ldots, x_n^2 - x_n = 0$ is unsolvable over the algebraic closure of $\mathbb{F}$, where $a(\psi_1 \vee \psi_2) := a(\psi_1) \cdot a(\psi_2)$, $a(x_i) := x_i$ and $a(\neg x_i) = 1 - x_i$.

Another related algebraic proof system is the *Polynomial Calculus* ($\mathsf{PC}_{\mathcal{R}}$), where $\mathcal{R}$ is a ring. A refutation of $\mathcal{F}$ in $\mathsf{PC}_{\mathcal{R}}$ is a sequence of polynomials $(p_1, \ldots, p_s = 1)$, where every polynomial p_i is either from $\mathcal{F}$ or is obtained from previous polynomials as a linear combination of two of them or by multiplication by a variable. In contrast to the *static* form of $\mathsf{NS}_{\mathbb{F}}$ refutations, where all coefficients in a decomposition of 1 through $f_1, \ldots, f_m$ are written at once, $\mathsf{PC}_{\mathbb{F}}$ refutations *dynamically* derive consequences from $\mathcal{F}$ line-by-line. This makes $\mathsf{PC}_{\mathbb{F}}$ stronger than $\mathsf{NS}_{\mathbb{F}}$ because of the possibility to cancel the monomials [25].

The size $S(\pi)$ of a $\mathsf{NS}_{\mathbb{F}}$ or $\mathsf{PC}_{\mathcal{R}}$ refutation π is the total number of monomials in it and the degree $d(\pi)$ is the maximal degree of monomials in it. The size and degree of refutations $\pi : \mathcal{F} \vdash 1 = 0$ in these systems are related: $d(\pi) - d(\mathcal{F}) = O(\log S(\pi))$, where $d(\mathcal{F})$ is the maximal degree of polynomials in $\mathcal{F}$ ([25]). A number of linear lower bounds on degree and, thus, exponential lower bounds on size have been proven in [61, 38, 25, 23].

Consideration of unsolvability proofs for systems $\mathcal{H} = \{h_1 \geq 0, \ldots, h_k \geq 0\}$ of polynomial *inequalities* leads to much stronger propositional proof systems. One of the most prominent such systems has its roots in real algebraic geometry and is based on the Positivstellensatz Theorem [20]. It follows from this theorem that whenever a system $\mathcal{F}, \mathcal{H}$ of real polynomial equalities and inequalities is unsolvable there exist $a_1, \ldots, a_m \in \mathbb{R}[x_1, \ldots, x_n]$ and $\{u_\alpha\}_{\alpha \in \{0,1\}^k} \subset \Sigma^2_{\mathbb{R}}[x_1, \ldots, x_n]$, where $\Sigma^2_{\mathbb{R}}[x_1, \ldots, x_n]$ denotes the set of sums of squares of real polynomials, such that $\sum\limits_{\alpha \in \{0,1\}^k} u_\alpha \cdot h_1^{\alpha_1} \cdot \ldots \cdot h_k^{\alpha_m} + \sum\limits_{i=1}^{m} a_i \cdot f_i = -1$. The proof system PS for unsatisfiable systems $(\mathcal{F}, \mathcal{H})$, where proofs are tuples of polynomial coefficients $\{a_i\}_{i \in [m]}, \{u_\alpha\}_{\alpha \in \{0,1\}^k}$ as above, is called *Positivstellensatz* proof system. A resricted version of PS, where $u_\alpha = 0$ whenever α contains more than one 1, is called *Sum-of-Squares* proofs system (SoS) [34].

Like $\mathsf{NS}_{\mathbb{F}}$, systems PS and SoS are static. There is also a dynamic version of PS – the system $\mathsf{PC}_{>}$ where refutations of $(\mathcal{F}, \mathcal{H})$ are $\mathsf{PC}_{\mathbb{R}}$ derivations of $\sum\limits_{\alpha \in \{0,1\}^k} u_\alpha \cdot h_1^{\alpha_1} \cdot \ldots \cdot h_k^{\alpha_m} + 1$. Note that $\mathsf{PC}_{>}$ is only dynamic on equalities. The full dynamic system, which is dynamic on both equalities and inequalities, is very strong, the degree and size lower bounds for this system seem to be far beyond existing methods.

The system $\mathsf{PC}_{>}$ is also a dynamic version of SoS in case $\mathcal{H} = \emptyset$. A remarkable peculiarity of these semi-algebraic proof systems: although $\mathsf{PC}_{\mathbb{F}}$ is strictly stronger than $\mathsf{NS}_{\mathbb{F}}$, in semi-algebraic setting $\mathsf{PC}_{>}$ is *equivalent* to PS [16].

Linear lower bounds on the degree of proofs in these systems are known [34], however no non-trivial lower bounds on size are known to date. In contrast to

algebraic case, lower bounds on degrees of semi-algebraic proofs do not imply lower bounds on sizes.

Other, weak, semi-algebraic proof systems include: the Cutting-Planes (CP) proof system, operating with linear inequalities over integers, and the Lovasz-Schrijver (LS) proof system, which is degree 2 fragment of the full dynamic PS[35].

Semi-algebraic systems, including weak ones for which exponential lower bounds were proven, have been extensively studied due to their connection to integer programming, namely to LP and SDP hierarchies [29], [49]. And the connection between complexity of SoS proofs and approximability of NP combinatorial optimisation problems has placed SoS at the frontiers of current research in complexity theory [12].

1.1.1.3 Bounded Arithmetic

As explained above, there are close connections between propositional proof complexity and computational complexity. These two, in turn, can be studied via weak fragments of arithmetic. One of the key works in the origins of this approach is the work of Buss [21], where theories S_2^i and T_2^i of bounded arithmetic were defined. These theories are defined over the language of Peano Arithmetic (PA) plus function symbols $\lfloor x/2 \rfloor, |x|, x\#y$. The axioms of S_2^i, T_2^i are axioms for the new function symbols plus the axioms of PA but for induction, which is different from that of PA and is a cornerstone in the definition of these theories. In S_2^i and T_2^i induction is restricted to Σ_i^b-formulas with not more than i alternating bounded quantifiers of the form $\exists(y < t(x))$ and $\forall(y < t(x))$, where t is a term, and without unbounded quantifiers. The induction in T_2^i is just the normal induction axiom scheme for Σ_i^b-formulas and the induction axiom scheme for S_2^i is:

$$\phi(0) \wedge (\phi(\lfloor x/2 \rfloor) \supset \phi(x)) \supset \forall x \phi(x)$$

where ϕ is a Σ_i^b-formula.

Theories S_2^i are intimately related to the polynomial hierarchy PH. For example, one of the main results in [21] states that a function f is strongly Σ_i^b-definable in S_2^i iff $f \in \mathrm{FP}^{\Sigma_{i-1}^p}$ (functional version of PH). In the work [46] it was proved that collapse of the hierarchy of the theories implies a collapse of PH. Subsequently, this result was strengthened independently in [22] and [65] by showing that $S_2 = \bigcup_i S_2^i$ is finitely axiomatizable iff PH collapses and this collapse is provable in S_2.

The connection of bounded arithmetic to propositional proof complexity is made by a *propositional translation* of first-order formulas with bounded quantifiers and proofs to propositional formulas and proofs of polynomial size respectively. There

are several such translations: for example, by Paris and Wilkie [54], by Cook [26] and by Krajicek and Pudlak [45]. These translations allow to apply thechniques from logic, in particular, from model theory to prove upper and lower bounds on sizes of propositional proofs. For example, one of the strongest results in propositional proof complexity – super-polynomial lower bounds on $\mathbf{AC}^0$-**Frege** system [1] – was achived by this method.

1.1.2 Complexity of optimization

A large class of problems, studied in theoretical computer science, is spanned by combinatorial optimization problems. These problems have the following form: given some disrcete structure and some set of objects, associated to it, the task is to find the optimal object according to some measure. Consider the following examples:

1. (Minimal spanning tree). Given a connected weighted graph $G = (V, E,$ $\omega : E \to \mathbb{N})$, among trees $T = (V, E' \subseteq E)$ such that T is connected (called *spanning trees* of G) find a tree of minimal weight $\omega(T) = \sum\limits_{e \in E'} \omega(e)$.

2. (Maximal independent set). Given a graph $G = (V, E)$, find a set $V' \subseteq V$ with maximal cardinality $|V'|$ such that $E(V', V') = \emptyset$, that is there is no edges between them.

3. (Sparsest cut). Given a d-regular graph $G = (V, E)$, find it's expansion $\phi_G = \min\limits_{S \subseteq V} \phi_G(S)$, where $\phi_G(S) := \frac{|E(S, V \setminus S)|}{d \cdot \min\{|S|, |V \setminus S|\}}$, and a set $S_* \subseteq V$ such that $\phi_G = \phi_G(S_*)$.

The complexity of the first two problems is fairly well-understood. The first one is easy: there are well-known classical polynomial-time algorithms for finding the minimal spanning tree, for example Prim's or Kruskal's algorithms [30]. And the second one is hard: unless the exponential time hypothesis fails, there is no algorithm, which given a graph G with n vertices and a number k checks that G has an independent set of size at least k in time $n^{o(k)}$, and, thus, it is impossible to beat the brute-force algorithm of time complexity $O(n^k)$ [24]. Once we know, that we probably cannot efficiently solve an optimization problem exactly, it is natural to ask whether at least an efficient approximation algorithm exists with a guarantee that the ratio between the value of a suggested solution and optimal value (approximation factor) is nicely bounded. However, unless $\mathsf{NP} \subseteq \mathsf{ZPTIME}(2^{(\log n)^{O(1)}})$ the best approximation factor for the maximal independent set a polynomial-time algorithm can achieve, namely $n^{1-o(n)}$, is close to the trivial one with approximation factor n [41].

The status of the third problem is less clear. Certainly, as the problem is NP-hard, it cannot be solved in polynomial time unless P=NP. And it also cannot be approximated with arbitrarily high precision, or, formally, it is not in class PTAS, unless $\mathsf{SAT} \in \mathsf{BPTIME}(2^{o(n)})$ [5]. But there exist polynomial-time computable non-trivial approximations to ϕ_G. For example, *Cheeger inequality* for ϕ_G implies the existence of a polynomial-time algorithm, which finds S such that $\phi_G(S) = O(\sqrt{\phi_G})$ [3]. Alternatively, the bound $\phi_G(S) = O(\sqrt{\log n} \cdot \phi_G)$ can be achieved by the algorithm of Arora, Rao, Vazirani [8]. However, the best known approximation factor is given by Cheeger inequality and it is not known, whether this bound is optimal. The precise characterisation of approximability of this problem is yet to be discovered.

Sparsest cut is not an isolated example. Usually, if there is a polynomial-time algorithm with non-trivial approximation factor for an optimization problem, which is not in PTAS under certain widely believed conjectures, then there is no proof of its optimality. In 2002 Khot formulated the *Unique Games Conjecture*, which is a conjecture on hardness of the Unique Games problem [42]. Since then, in a series of works, tightness of upper bounds have been derived from UGC for a number of problems. In particular, UGC implies optimality of the Cheeger inequality algorithm for the sparsest cut problem [59] and Raghavendra proved, assuming UGC, optimality results for all Constraint Satisfaction Problems [57].

A remarkable outcome of this research is that all problems in a certain class can be solved by a single "meta-algorithm", whose optimality follows from UGC. This meta-algorithm is based on *semidefinite programming* and, as soon as it phrased in this terms, it suggests a natural generalisation: the *SoS algorithm*. It applies to problems of the form:

$$\min_{\overline{x} \in K} g(\overline{x})$$

where $g \in \mathbb{R}[\overline{x}]$ and K is an algebraic set: $K = \{\overline{x} \in \mathbb{R}^n \mid f_1(\overline{x}) = \ldots = f_m(\overline{x}) = 0\}$ for some $f_1, \ldots, f_m \in \mathbb{R}[\overline{x}]$. The degree-$d$ Sum-of-Squares algorithm finds via binary search the largest $L^{(d)}$ such that the system of polynomial equations $g(\overline{x}) - L^{(d)} = 0, f_1(\overline{x}) = 0, \ldots, f_m(\overline{x}) = 0$ has a degree-d SoS refutation. This is a polynomial time procedure because of automatizability of constant degree SoS via semidefinite programming [12]. With the increase of d, the SoS algorithm gets closer and closer to the optimal value: $L^{(2)} \leq L^{(4)} \leq \ldots \leq \min_{\overline{x} \in K} g(\overline{x})$.

The best known algorithm for optimization problems, for which tightness of approximation follows from UGC, is the degree-2 SoS algorithm. Thus, UGC implies the optimality of degree-2 SoS algorithm on these problems and, in particular, means that degree-d SoS algorithm for any constant d cannot beat degree-2 SoS algorithm.

Challenging this conjecture as well as providing supporting evidence often relies on provability of certain inequalities, say, from Boolean analysis in low degree SoS. A good source of examples of this kind of SoS-ing results from Boolean analysis for approximability is the paper of O'Donnell and Zhou [52].

Consider *Small Set Expansion Problem* (SSEP). Like Sparsest Cut, it asks to find a set $S \subset V$ of minimal expansion $\phi_G(S)$, but under the condition that $|S| \leq \delta \cdot |V|$ for some δ. Recall, that in case of Sparsest Cut, Cheeger Inequality gives approximation of $\phi_G(S)$ via second largest eigenvalue $\lambda_2(G)$ of the adjacency matrix of G and thus reduces approximation to computing eigenvalues. Although it looks similar to Sparsest Cut, eigenvalues and similar methods do not work in this setting and, therefore, Cheeger Inequality cannot be applied. *Small Set Expansion Hypothesis* (SSEH) is equivalent to UGC and states that small set expansion is hard to approximate.

SSEP is a special case of finding "sparse" vectors in a linear space. Specifically, for $p > 1$ and $\delta \in (0, 1)$ call a vector $x \in \mathbb{R}^n$ (δ, p)-*sparse* if $(\|x\|_{2p})^{2p} \geq \delta^{1-p} \cdot (\|x\|_2)^{2p}$. Fix any $p \geq 2$ and $\phi \in (0, 1)$. Then if there exists $S \subseteq V$ with $|S| = o(|V|)$ and $\phi_G(S) \leq \phi$ then there exists an $(o(1), p)$-sparse vector $x \in W_{\phi+o(1)}$, where for every λ $W_{\leq\lambda}$ denotes the span of eigenvectors of the Laplacian matrix of G with eigenvalues at most λ. Conversely, if there exists a $(o(1), p)$-sparse vector $x \in W_\phi$, then there exists $S \subset V$ with $S = o(|V|)$ and $\phi_G(S) < \rho$ for some constant $\rho < 1$, depending on ϕ [10].

Thus we can say whether minimum of $\phi_G(S)$ is close to one or close to zero by estimating the maximum of the norm $\|x\|_{2p}$ over all unit vectors in some linear subspace. Therefore, potentially SSEH and UGC can be resolved by estimating the degree needed for SoS proofs to certify an inequality on $\|x\|_{2p}$ for unit vectors x in some linear $W \subseteq \mathbb{R}^n$. One such inequality is provided by the (2,4) hypercontractivity theorem [53], which states that for every d and every polynomial f with t variables and of degree at most d the subspace $W_d \subset \mathbb{R}^{2^t}$ of evaluations of f on $\{-1, +1\}^t$ does not contain $(o(1), 2)$-sparse vectors and satisfies for all $x \in W_d$:

$$(\|x\|_4)^4 \leq 9^t \cdot (\|x\|_2)^2$$

The existence of constant degree SoS proofs of this inequality was used in a number of works [58],[11] that showed that some hard instances are easy for SoS algorithm.

1.2 Contributions

This thesis contributes to the studies of the complexity of proofs, operating with algebraic expressions. Two topics within the subject are addressed: complexity of

proofs in resolution over linear equations and a formulation of first-order theories, capturing the strength of constant degree $\mathsf{PC}_{\mathcal{R}}$ and SoS. Although two stories, that underlie motivation and particular developments of these two topics, are different, there is a unifying objective behind the work: analysis of complexity of different forms of algebraic and semi-algebraic reasoning in propositional proof systems.

The work on resolution over linear equations focuses on elementary combinatorial approaches to the complexity of proofs. All lower bound techniques that have been developed for systems, operating with De Morgan formulas, like Resolution or $\mathbf{AC}^0$-Frege, fail to achieve strong lower bounds even for the case of minor extension of De Morgan formulas with algebraic expressions as simple as linear equations. This part of the thesis seeks for novel techniques, applicable in this context.

The second part of the thesis is devoted to a conceptual analysis of strength of constant degree SoS. The objective of this research is to formulate a first-order theory, corresponding to constant degree SoS under propositional translation, so that it is naturally capable of some amount of "ZFC" reasoning, used in, say, standard proofs of hypercontractive inequalities. Somewhat similar theory has been considered in [63] for the studies of complexity of linear algebra. Our propositional translation is defined in flavour of the one in [14].

1.2.1 Resolution over Linear Equations

In this part of the thesis we continue the study of the power of resolution over linear equations, while extending it to different rings $\mathcal{R}$, denoted $\mathrm{Res}(\mathrm{lin}_{\mathcal{R}})$, both finite and infinite. We prove a host of new lower bounds, separations and upper bounds for resolution over linear equations, including dag-like refutations. We focus mainly on finite fields $\mathbb{F}_q$, for different primes q, and fields of characteristic 0, most importantly the rational numbers $\mathbb{Q}$. Using our notation, R(lin) from [60] is simply $\mathrm{Res}(\mathrm{lin}_{\mathbb{Z}})$ and $\mathrm{Res}(\oplus)$ from [40] is $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$.

The refutation system $\mathrm{Res}(\mathrm{lin}_R)$ is defined as follows (see [60]). The proof-lines of $\mathrm{Res}(\mathrm{lin}_{\mathcal{R}})$ are ***linear clauses***, that is, disjunctions of linear equations. More formally, they are disjunctions of the form:

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

where k is some number (the *width* of the clause), and $a_{ji}, b_j \in \mathcal{R}$. 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 $\alpha, \beta \in \mathcal{R}$, and C, D some linear clauses. A Res(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 a some variable x_i, or was derived from previous proof-lines by the above resolution rule, or by the *weakening rule* that allows to extend clauses with arbitrary disjuncts, or a *simplification rule* allowing to discard false constant linear forms (e.g., $1 = 0$) from a linear clause. The last proof-line in a refutation is the empty clause (standing for the truth value false).

We are interested in the following questions:

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

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

(Q3) Can tree-like systems for different rings $\mathcal{R}$ be separated?

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

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

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

Let $k = char(\mathcal{R})$ be the characteristic of the ring $\mathcal{R}$. In case $k \notin \{1, 2, 3\}$, deciding whether an $\mathcal{R}$-linear clause D is a tautology (that is, holds for every 0-1 assignment to its variables) is at least as hard as deciding whether a 3-DNF is a tautology (because over characteristic $k \notin \{1, 2, 3\}$ linear equations can express conjunction of three conjuncts). For this reason $\text{Res}_{sw}(\text{lin}_\mathcal{R})$ proofs cannot be checked in polynomial time and thus $\text{Res}_{sw}(\text{lin}_\mathcal{R})$ 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 [28]).

The reason for studying $\text{Res}_{sw}(\text{lin}_\mathcal{R})$ is mainly the following: Let Γ be an arbitrary set of tautological $\mathcal{R}$-linear clauses. Then, lower bounds for tree-like $\text{Res}_{sw}(\text{lin}_\mathcal{R})$ imply lower bounds for tree-like Res($\text{lin}_\mathcal{R}$) 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 Res($\text{lin}_\mathbb{F}$) is

quite limited. For instance, we show that $2x_1 + \ldots + 2x_n = 1$ requires an exponential-size in n refutations. On the other hand, such contradictions *do* admit short tree-like $\text{Res}(\text{lin}_{\mathbb{F}})$ refutations in the presence of the following *generalized boolean axioms* (which is a tautological linear clause):

$$\mathsf{Im}(f) := \bigvee_{A \in im_2(f)} (f = A), \tag{1.1}$$

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

Lower bounds and separations in characteristic zero. First, we show that for $\mathbb{Q}$, whenever $\alpha_1 x_1 + \cdots + \alpha_n x_n + \beta = 0$ is unsatisfiable (over 0-1 assignments), it has polynomial dag-like $\text{Res}(\text{lin}_{\mathbb{Q}})$ refutations if the coefficients are polynomially bounded and it requires exponential dag-like $\text{Res}(\text{lin}_{\mathbb{Q}})$ refutations if coefficients are exponential. Note that $\alpha_1 x_1 + \cdots + \alpha_n x_n + \beta = 0$ expresses the *subset sum principle*: $\alpha_1 x_1 + \cdots + \alpha_n x_n = -\beta$ 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 23; Dag-like lower bound)**.** *All* $\text{Res}(\text{lin}_{\mathbb{Q}})$ *refutations of* $x_1 + 2x_2 + \cdots + 2^n x_n + 1 = 0$ *are of size* $2^{\Omega(n)}$.

The proof of this theorem introduces a new lower bound argument. Specifically, we show that every (dag- or tree-like) refutation π of a subset sum principle of the form $x_1 + 2x_2 + \cdots + 2^n x_n + 1 = 0$ can be transformed without much increase in size into a derivation of a clause C_π from Boolean axioms. We ensure that every disjunct $g = 0$ of C_π has at most 2^{cn} 0-1 satisfying assignments for some $c < 1$. Because C_π is derived from Boolean axioms, it must be a Boolean tautology and therefore it must contain at least $2^{(1-c)n}$ disjuncts. As our constructed derivation is not much larger than the original refutation, the size of the original refutation must be $2^{\Omega(n)}$.

This proof essentially relies on the fact that coefficients of the linear form are exponential: every contradiction of the form $f = 0$ can be shown to admit polynomial size dag-like $\text{Res}(\text{lin}_{\mathbb{Q}})$ refutations whenever coefficients of f are polynomially bounded. A natural question is whether in case of bounded coefficients $f = 0$ can be effeciently refuted already by tree-like $\text{Res}(\text{lin}_{\mathbb{Q}})$ refutations. The question turnes out to be non-trivial, we prove that the answer is negative:

Theorem (Theorem 35). *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 dag-like bound to reduce lower bound problem for refutations of $f = 0$ to lower bound problem for derivations of clauses of 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 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 C_π are large:

Theorem (Theorem 33). *Any tree-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ *derivation of any tautology of the form $\bigvee_{j\in[N]} g_j = 0$, 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 [56] for resolution, and developed further by Itsykson-Sokolov [40] for $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$, to general rings, including characteristic zero rings. We define a non-trivial strategy for Delayer in the corresponding game and prove that it gurantees $\sqrt{n}$ coins using a bound on size of essential coverings of the hypercube ([48]). The relation between Prover-Delayer games and tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ refutations allows to conclude that the size of tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ refutations must be $2^{\Omega(\sqrt{n})}$.

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

Corollary (Corollary 34). *Let f be any linear polynomial over $\mathbb{Q}$, which 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 [40] for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$. Together with the polynomial upper bound for PHP^m_n refutations in dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ for fields $\mathbb{F}$ of characteristic zero demonstrated in [60], our results establish a separation between dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ and tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ for characteristic zero fields.

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

Theorem (Theorem 19; Raz-Tzameret [60]; Short dag-like pigeonhole principle refutations)**.** *For every ring R of characteristic zero there exists a* $\mathrm{Res}(\mathrm{lin}_R)$ *refutation of* $\neg\mathrm{PHP}_n^m$ *of polynomial size.*

To prove Theorem 38 we need to prove that Delayer's strategy from [40] 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 $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ and tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$, as follows. We define the *image avoidance principle* to be:

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

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

Theorem (Theorem 15)**.** *For every ring $\mathcal{R}$ and every linear form f, there are polynomial-size* $\mathrm{Res}(\mathrm{lin}_R)$ *refutations of* $\mathsf{ImAv}\left(f\right)$.

Theorem (Theorem 37)**.** *Let $f=\epsilon_1x_1+\cdots+\epsilon_nx_n$, where $\epsilon_i\in\{-1,1\}\subset\mathbb{F}$, and let $\mathbb{F}$ be a field of characteristic zero. Then any tree-like* $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ *refutation of* $\mathsf{ImAv}\left(f\right)$ *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 explain briefly.

Let f be a linear form as in Theorem 37. We consider an instance of the Prover-Delayer game for $\mathsf{ImAv}\left(f\right)$. 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 $m\geq\frac{n}{4}$. 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 1.1 sums up our knowledge up to this point with respect to characteristic 0 fields.

	$\sum_{i=1}^{n} 2x_i = 1$	$\sum_{i=1}^{n} 2^i x_i = -1$	$\mathsf{ImAv}\left(\sum_{i=1}^{n} x_i\right)$	PHP_n^m	$\mathsf{Im}\left(\sum_{i=1}^{n} x_i\right)$
t-l $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$	$2^{\Omega(\sqrt{n})}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(\sqrt{n})}$
t-l $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$	poly	poly	$2^{\Omega(n)}$	$2^{\Omega(n)}$	poly
$\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$	poly	$2^{\Omega(n)}$	poly	poly [60]	poly

Table 1.1: Lower and upper bounds for fields of characteristic 0. The notation t-l $\mathrm{Res}(\mathrm{lin}_{\mathcal{R}})$ stands for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathcal{R}})$. The rightmost column describes bounds on *derivations*, in contract to refutations.

1.2.1.1 Lower Bounds and Separations in Finite Fields

We now turn to resolution over linear equations in *finite fields.* We obtain many new tree-like lower bounds over finite fields (Table 1.2).

We have already discussed above lower bounds for the pigeonhole principle which hold both for infinite and finite fields. 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 $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}_{p^k}})$ on random k-CNFs.[2]

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 $PC_{\mathbb{F}}$ [25], such that $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ proofs of width ω are translated to $PC_{\mathbb{F}}$ proofs of degree ω (the *width* ω of a clause is defined to be the total number of disjuncts in a clause). This establishes the lower bounds for the size of tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ proofs via lower bounds on $PC_{\mathbb{F}}$ degrees.

We show that

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

where ω_0 is what we call the *principal width*, which counts the number of linear equations in clauses when we treat as identical those defining parallel hyperplanes,

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

and $S_{\text{t-l Res}(\text{lin}_R)}(\phi \vdash \bot)$ denotes the minimal size of a tree-like $\text{Res}(\text{lin}_R)$ refutation of ϕ.

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

Theorem (Theorem 44; Size-width relation)**.** *Assume ϕ is an unsatisfiable CNF formula. The following relation between principal width and size holds for tree-like* $\text{Res}(\text{lin}_{\mathbb{F}})$ *and tree-like tree-like* $\text{Res}_{sw}(\text{lin}_{\mathbb{F}})$*: $S(\phi \vdash \bot) = 2^{\Omega(\omega_0(\phi \vdash \bot) - \omega_0(\phi))}$. 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 [33, Theorem 14] who showed a size-width relation for a system denoted tree-like $\text{PK}^{\text{id}}_{O(1)}(\oplus)$, which is a system extending tree-like $\text{Res}(\text{lin}_{\mathbb{F}_2})$ 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* $\text{Res}(\text{lin}_{\mathbb{F}})$ *refutation of an unsatisfiable CNF formula ϕ. Then, there exists a $PC_{\mathbb{F}}$ refutation π' of (the arithmetization of) ϕ of degree $\omega(\pi)$.*

Corollary (Corollary 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 would like to stress that the size-width relation of Theorem 44* ***cannot*** *be used for transfering $PC_{\mathbb{F}}$ 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 1.2 shows the results for $\text{Res}(\text{lin}_{\mathcal{R}})$ over finite fields.

	$A\overline{x}=\overline{b}$	$\mathrm{TS}^{(-)}_{G,\sigma}$	$\mathrm{TS}^{(q)}_{G,\sigma}$	random k-CNF	PHP^m_n
t-l $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_{p^k}})$	$2^{\Omega(dn)}$	poly	$2^{\Omega(dn)}$	$2^{\Omega\left(\frac{n}{\Delta^{2/(k-2)}\cdot\log\Delta}\right)}$	$2^{\Omega(n)}$
t-l $\mathrm{Res}(\oplus)$	poly [40]	poly [40]	$2^{\Omega(dn)}$	$2^{\Omega\left(\frac{n}{\Delta^{2/(k-2)}\cdot\log\Delta}\right)}$ [33]	$2^{\Omega(n)}$ [40]
t-l $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}_{p^k}})$	poly	poly	?	?	$2^{\Omega(n)}$

Table 1.2: Lower bounds over finite fields. Here G is d-regular graph and Δ is the clause density (number of clauses divided by the number of variables), $A\overline{x}=\overline{b}$ stands for a linear system over $\mathbb{F}_{p^k}$ that has no 0-1 solutions in the first and the third rows, and in the second row the linear system $A\overline{x}=\overline{b}$ is over $\mathbb{F}_2$. The notation $\mathrm{TS}^{(-)}_{G,\sigma}$ stands for $\mathrm{TS}^{(p)}_{G,\sigma}$ in the first and the third rows and for $\mathrm{TS}^{(2)}_{G,\sigma}$ in the second raw. t-l $\mathrm{Res}(\mathrm{lin}_{\mathcal{R}})$ stands for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathcal{R}})$, and $p\neq q$ are primes (in the second raw and third column we assume $q\neq 2$). Circled "?" denotes an open problem. The results marked with [40, 33] were proved in the corresponding papers. All other results are from the current work.

1.2.2 Complexity of Linear Systems

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

Proposition (Proposition 16; 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* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ *refutations of ϕ of size polynomial in the sum of sizes of encodings of all coefficients in A.*

The upper bound in Proposition 16 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 $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ 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 (see Sec 2.2.3). This means that unless P = NP there exist 0-1 unsatisfiable linear systems that require superpolynomial dag-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutations. Moreover, the reduction R from k-UNSAT is such that $\phi\in k$-UNSAT has $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutations of size S iff the system $R(\phi)$ has $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ refutations of size $O(S)$. Thus, in general proving lower bounds for linear systems can be as hard as proving lower bounds for CNFs: lower bounds for some linear systems imply lower bounds for CNFs.

An unconditional explicit bound for tree-like Res($\text{lin}_{\mathbb{F}}$) can be obtained via $PC_{\mathbb{F}}$ using size-width relation for finite fields (Theorem 44) and Proposition 7. In particular, hard instances of the form $A\,\overline{x} = \overline{b}$ can be constructed by applying the reduction in the proof of NP-completeness of 0-1 satisfiability of linear systems to, say, mod 2 Tseitin formulas. Our work implies an exponential lower bound for the size of tree-like Res($\text{lin}_{\mathbb{F}}$) refutations of these systems (for large enough, but constant, characteristic) and we conjecture that they are hard for dag-like Res($\text{lin}_{\mathbb{F}}$) as well.

We prove an upper bound for linear systems and suggest another, more direct, construction of a hard candidate, using error-correcting codes.

Theorem (Theorem 24; Upper bound on 0-1 unsatisfiable linear systems)**.** *Let $A_{f_1,\dots,f_m} : \mathbb{F}^n \to \mathbb{F}^m$ be an affine map $\overline{x} \mapsto (f_1(\overline{x}),\dots,f_m(\overline{x}))$, where $f_1,\dots,f_m$ are linear forms. If the system $f_1 = 0,\dots,f_m = 0$ is unsatisfiable over 0-1, that is, if $0 \notin im_2(A_{f_1,\dots,f_m}\,\overline{x})$, then there exists a* Res($\text{lin}_{\mathbb{F}}$) *refutation of this system of size $poly(n + |im_2(A_{f_1,\dots,f_m}\,\overline{x})|)$.*

The instance is constructed specifically to be provably hard for a simple and natural model of decision trees, which can be simulated both by tree-like Res($\text{lin}_{\mathbb{F}}$) and $PC_{\mathbb{F}}$ and reflects a natural strategy to refute 0-1 unsatisfiable linear systems. Such a strategy for refuting $A\,\overline{x} = \overline{b}$ can informally be described as follows: select variables and try to assign them 0-1 values until the system $(A\,\overline{x} = \overline{b})\restriction_\rho$ becomes unsatisfiable over $\mathbb{F}$, where ρ is the current assignment, and refute it by a polynomial-size refutation, guaranteed by Proposition 16 (above). Formally, a decision tree for $A\,\overline{x} = \overline{b}$ is a binary decision tree, where every leaf is marked with unsatisfiable over $\mathbb{F}$ system $(A\,\overline{x} = \overline{b})\restriction_\rho$, where ρ consists of variable assignments on the path from the root to the leaf.

The matrix A of the instance is constructed as a generator matrix of a linear error-correcting $(n,k,d)_q$ code, where n is the code length, k is the dimension of the code space, d is the minimal distance of the code and $q = |\mathbb{F}|$. The parameter k is chosen to be large enough to ensure that $q^k > 2^n$ and thus there exists some $\overline{b}$ such that $A\,\overline{x} = \overline{b}$ has no 0-1 solutions. On the other hand, $d = \Omega(\frac{n}{\log n})$ is chosen to be large enough to ensure that all the leaves of a decision tree for $A\,\overline{x} = \overline{b}$ are sufficiently deep in the tree: if ρ assigns at most $k < d$ variables, then the code generated by $A\restriction_\rho$ has a minimal distance at least $d - k$ and therefore $A\restriction_\rho$ has full rank. The existence of this code is guaranteed by the Gilbert bound.

Theorem (Theorem 26; Lower bound for decision trees on linear systems)**.** *For every $n \in \mathbb{N}$ there exists a 0-1 unsatisfiable linear system $A\,\overline{x} = \overline{b}$ over a finite field $\mathbb{F}_q$, $q > 2$, with n variables, such that any decision tree for this system is of size $2^{\Omega(\frac{n}{\log n})}$.*

1.2.2.1 Nondeterministic Linear Decision Trees

There is well-known size preserving (up to a constant factor) correspondence between tree-like resolution refutations for unsatisfiable formulas ϕ and decision trees, which solve the following problem: given an assignment ρ for the variables of ϕ, determine which clause $C \in \phi$ is falsified by querying values of the variables under the assignment ρ. In Itsykson-Sokolov [40] this correspondence was generalized to tree-like Res($\oplus$) refutations and parity decision trees. In the current work we initiate the study of linear decision trees and their properties over different characteristics, extending the correspondence to a correspondence between tree-like Res(lin_R) (and tree-like $\text{Res}_{sw}(\text{lin}_R)$) 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 $f \neq 0$.)

Theorem (Theorem 28)**.** *If ϕ is an unsatisfiable CNF formula, then every tree-like* $\text{Res}(\text{lin}_R)$ *or tree-like* $\text{Res}_{sw}(\text{lin}_R)$ *refutation can be transformed into an NLDT for ϕ of the same size up to a constant factor, and vise versa.*

This is joint work with Iddo Tzameret.

1.2.3 First-Order Theories for (Semi-)Algebraic Proof Systems

As we explained in Section 1.1.2, better understanding of what we can prove in constant degree SoS would contribute to our understanding of SoS algorithm and might, potentially, lead to a refutation of UGC. In certain cases, as, for example, shown in [52], whether we can obtain approximability results depends on whether some known theorems from, say, Boolean analysis can be formulated and proven in constant degree SoS. For the sake of adopting and adjusting known proofs or, perhaps, finding new proofs of such theorems for SoS_d, namely SoS of constant degree, it would be helpful to identify, what kind of reasoning patterns are feasible for SoS_d. For example, can we perform case analysis, can we reason by induction, can we reason

about fractional powers in SoS_d, and so on. We address this matter by defining a first-order theory TSoS such that proofs in this theory can be translated to a variant of SoS_d by a propositional translation. Our goal is to come up with a natural theory with a language as rich as possible and axioms and rules as strong as possible, provided refutations in the theory still can be translated to refutations in SoS_d.

1.2.3.1 Theory for $\mathsf{PC}_{\mathcal{R},d}$

We first define a theory $\mathsf{TPC}_{\mathcal{R}}$ for constant degree polynomial calculus, where $\mathcal{R}$ is a ring, and then obtain a theory TSoS for SoS_d as an extension of $\mathsf{TPC}_{\mathbb{R}}$. Theory $\mathsf{TPC}_{\mathcal{R}}$ is a two-sorted first-order theory over the language $\mathcal{L}^{\mathcal{R}}_{=}$ with ring sort for elements of a ring and index sort for natural numbers. Polynomials over $\mathcal{R}$ are represented as ring sort terms. For example, polynomial $(x_1 + \ldots + x_n) \cdot (x_k - a) + b$, where $a, b \in \mathcal{R}$, is represented as the term $\sum_i(X(i), n) \cdot (X(k) - a) + b$, where k, n are index-terms; i is index-variable; $\sum_i(r, n) \in \mathcal{L}^{\mathcal{R}}_{=}$ is the summation symbol for sums with varied number of summands; $+, -, \cdot \in \mathcal{L}^{\mathcal{R}}_{=}$ are symbols for standard ring operations; $a, b \in \mathcal{R} \subset \mathcal{L}^{\mathcal{R}}_{=}$ are ring constants and $X(i) \in \mathcal{L}^{\mathcal{R}}_{=}$ is a symbol for ring-sort valued oracle[3], which represents a sequence of variables. For the index sort $\mathcal{L}^{\mathcal{R}}_{=}$ contains a symbol for every function $f : \mathbb{N} \to \mathbb{N}$ such that $f = O(n^c)$ for some c. It follows that $\mathcal{L}^{\mathcal{R}}_{=}$ contains index-sort function symbols for all polynomially bounded k-ary functions.

Atomic formulas of $\mathsf{TPC}_{\mathcal{R}}$ are just equality predicates $=_{rng}, =_{ind} \in \mathcal{L}^{\mathcal{R}}_{=}$ for the ring sort and index sort respectively (we omit subscripts). Arbitrary index sort predicates are represented as the formula $f(n_1, \ldots, n_k) = 1$ for some $f : \mathbb{N}^k \to \{0, 1\}$. The logical language of $\mathsf{TPC}_{\mathcal{R}}$ apart from the usual elements of two-sorted first-order logic contains bounded index-sort universal quantifier $\forall(i < s)$, where i is an index-variable and s is any index-term such that i does not appear free in s.

The axioms of $\mathsf{TPC}_{\mathcal{R}}$ include, for instance, ring axioms for $+, -, \cdot$, integral domain axioms; axioms for $\sum_i$; axioms for all true sentences[4], not containing occurrences of the oracle X and free ring-variables. The theory $\mathsf{TPC}_{\mathcal{R}}$ has also induction rule for a class of $\mathcal{L}^{\mathcal{R}}_{=}$-formulas, which we denote $\Phi^{\mathcal{R}}_{=}$. Every formula $\phi(i, \overline{y}) \in \Phi^{\mathcal{R}}_{=}$, where i and $\overline{y}$ are free index- and ring-variables respectively, is such that for every $n \in \mathbb{N}$ $\phi(n, \overline{y})$ describes a property of the oracle X and ring-variables $\overline{y}$, which can also be defined by systems $\mathcal{P}_n$ of degree d polynomial equations with variables $X(0), X(1), \ldots$ and $\overline{y}$. Technically, these formulas are $\mathcal{L}^{\mathcal{R}}_{=}$-formulas with connectives $\vee, \wedge$, bounded index quantifier $\forall(i < s)$ and arbitrary subformulas as long as they do not contain

[3] Oracle is just a function symbol.

[4] True in the standard model.

occurrences of X or free ring-variables. This completes the sketch of the definition of $\mathsf{TPC}_{\mathcal{R}}$.

In order to relate first-order reasoning in $\mathsf{TPC}_{\mathcal{R}}$ to $\mathsf{PC}_{\mathcal{R},d}$ derivations we do the following. For all formulas $\phi(i)$ in $\Phi^{\mathcal{R}}_{=}$ we define a translation of ϕ to the family $\{\langle\phi\rangle_n\}_n$ of systems of polynomial equations of degree d with variables $x_0, \ldots, x_{s(n)}$ for some polynomially bounded $s(n)$. These translations $\langle\phi\rangle_n$ are natural phrasings of $\phi(n)$ in terms of systems of polynomial equations in the sense that atomic formulas $\langle t(i) = 0\rangle_n$ are translated to "equivalent" polynomial equations, where $X(0), X(1), \ldots$ are replaced with variables $x_0, x_1, \ldots$, and formula forming operations $\forall(i < s), \wedge, \vee$ are translated to semantically equivalent operations on systems. For example, we define $\langle\phi \vee \psi\rangle_n := \langle\phi\rangle_n \cdot \langle\psi\rangle_n$, where $\langle\phi\rangle_n \cdot \langle\psi\rangle_n := \{p \cdot q = 0 \,|\, p = 0 \in \langle\phi\rangle_n, q = 0 \in \langle\psi\rangle_n\}$.

Next, the goal is to define a translation of $\mathsf{TPC}_{\mathcal{R}}$ derivations $\phi(i) \vdash \psi(i)$, where $\phi, \psi \in \Phi^{\mathcal{R}}_{=}$, to families of $\mathsf{PC}_{\mathcal{R},d}$ derivations $\langle\phi\rangle_n \vdash \langle\psi\rangle_n$. By such a translation, once a family $\mathcal{P}_n$ of unsatisfiable systems of polynomial equations is phrased in a uniform way as a formula $\phi(i) \in \Phi^{\mathcal{R}}_{=}$ such that $\langle\phi\rangle_n = \mathcal{P}_n$, refuting $\mathcal{P}_n$ in $\mathsf{PC}_{\mathcal{R},d}$ can be reduced to refuting $\phi(i)$ in $\mathsf{TPC}_{\mathcal{R}}$.

In order to define this translation we represent $\mathsf{TPC}_{\mathcal{R}}$ derivations in the two-sorted version of sequent calculus LK. By the free-cut elimination theorem for the two-sorted first-order sequent calculus, every derivable sequent is derivable by free-cut free proofs. Free-cut free proofs possess the following subformula property, which is very useful for the translation: every formula in a free-cut free proof is a subformula of either a formula in the endsequent or a formula in an axiom. As we can represent every axiom of $\mathsf{TPC}_{\mathcal{R}}$ as a sequent, where all formulas are in $\Phi^{\mathcal{R}}_{=}$ and the induction rule is defined for $\Phi^{\mathcal{R}}_{=}$ formulas, the free-cut elimination theorem guarantees that if all formulas in a derivable sequent are in $\Phi^{\mathcal{R}}_{=}$, then there is a derivation of this sequent, where all formulas are in $\Phi^{\mathcal{R}}_{=}$.

It is, thus, enough to define the translation on free-cut free proofs, and this can be done inductively step by step. However, depending on the ring $\mathcal{R}$, some rules of LK can admit no translation to operations on $\mathsf{PC}_{\mathcal{R},d}$ derivations. For example, the contraction rule says that the sequent $t = 0 \longrightarrow r = 0 \vee r = 0$ derives the sequent $t = 0 \longrightarrow r = 0$. In order to translate this rule, once we have a $\mathsf{PC}_{\mathcal{R},d}$ derivation $\langle t = 0\rangle_n \vdash \langle r = 0 \vee r = 0\rangle_n$, where $\langle r = 0 \vee r = 0\rangle_n = \langle r = 0\rangle^2_n$, we should be able to construct a $\mathsf{PC}_{\mathcal{R},d}$ derivation $\langle t = 0\rangle_n \vdash \langle r = 0\rangle_n$. This can be done iff it is possible to derive $p = 0$ from $p^2 = 0$ in $\mathsf{PC}_{\mathcal{R},d}$ for all polynomials p.

In case $\mathbb{F}_q$ is a field of characteristic $q > 0$, there exist $\mathsf{PC}_{\mathbb{F}_q,d}$ derivations $p^2 = 0 \vdash p = 0$ (Proposition 49) and, consequently, we show that $\mathsf{TPC}_{\mathbb{F}_q}$ derivations admit

a translation to $\mathsf{PC}_{\mathbb{F}_q,d}$ derivations. On the other hand, it follows from the work in [31], that for fields $\mathbb{F}$ of characteristic 0, derivations $(x_1 + \ldots + x_n + 1)^2 = 0 \vdash (x_1 + \ldots + x_n + 1) = 0$ are of degree $\Omega(n)$ (Proposition 48). In this case, the translation requires $\mathsf{PC}_{\mathbb{F},d}$ to be extended with the *radical rule* $p^2 = 0 \vdash p = 0$. We denote this extension $\mathsf{PC}^{\mathrm{rad}}_{\mathcal{R},d}$. The translation is thus given by the following theorem:

Theorem (Theorem 51)**.** *Let Π be a $\mathit{TPC}_{\mathcal{R}}$ derivation of the sequent $\Gamma \longrightarrow \Delta$ such that all formulas in Γ and Δ are in $\Phi^{\mathcal{R}}_{=}$ and have free index-variable i. Then there exist $d \in \mathbb{N}$ such that for every $n \in \mathbb{N}$ there exists $\mathsf{PC}^{rad}_{\mathcal{R},d}$ ($\mathsf{PC}_{\mathcal{R},d}$ in case $\mathcal{R} = \mathbb{F}_q, q > 0$) derivation:*

$$\langle \bigwedge_{\phi \in \Gamma} \phi \rangle_n \vdash \langle \bigvee_{\phi \in \Delta} \phi \rangle_n$$

Unfortunately, this translation does not yet quite reach the original goal for $\mathbb{F}, \mathit{char}(\mathbb{F}) = 0$ as its destination $\mathsf{PC}^{\mathrm{rad}}_{\mathbb{F},d}$ is not $\mathsf{PC}_{\mathbb{F},d}$. However, although $\mathsf{PC}^{\mathrm{rad}}_{\mathbb{F},d}$ is strictly stronger than $\mathsf{PC}_{\mathbb{F},d}$ as a derivation system (Proposition 48), it still might happen that $\mathsf{PC}^{\mathrm{rad}}_{\mathbb{F},d}$ is not stronger than $\mathsf{PC}_{\mathbb{F},d}$ as a refutation system. This requires further investigation on the power of the radical rule.

1.2.3.2 Theory for SoS_d

We define two theories, which are built on top of a variant of $\mathsf{TPC}_{\mathbb{R}}$ and correspond to constant degree SoS: TSoS and $\mathsf{TSoS}_{\geq}$.

Theory TSoS is a "minimalistic" theory for SoS_d: it is $\mathsf{TPC}_{\mathbb{R}}$, extended with the following axiom for every term $t(i)$: $\sum_{i=0}^{n} t(i)^2 = 0 \supset \forall(i \leq n)\ t(i)^2 = 0$. Propositional translation of $\mathsf{TPC}_{\mathbb{R}}$ trivially extends to a translation of TSoS to an extension of $\mathsf{PC}^{\mathrm{rad}}_{\mathbb{R}}$, which we denote PC^{+}. The system PC^{+} adds to $\mathsf{PC}^{\mathrm{rad}}_{\mathbb{R}}$ the following rule: $f_1^2 + \cdots + f_m^2 = 0 \vdash f_1^2 = 0$. We extend the simulation of $\mathsf{PC}_{\mathbb{R}}$ by SoS, proven in [16], to a simulation of PC^{+} by SoS:

Theorem (Theorem 54)**.** *If there exists a PC^{+} refutation of degree d of a set of equalities $\mathcal{F}$, then there exists SoS refutation of $\mathcal{F}$ of degree $2d$.*

We also prove that TSoS has the right strength in the sense that it formalizes soundness of constant degree SoS:

Theorem (Theorem 57, Informal)**.** *TSoS proves soundness of constant degree SoS.*

Despite of all this, the theory TSoS is too poor and provides a little insight on strength of constant degree SoS. We introduce intuitionistic theory $\mathsf{TSoS}_{\geq}$, which

contains the theory $\mathsf{TPC}_{\mathbb{R}}$ as a subtheory, except for the integral domain axiom and induction axiom scheme. $\mathsf{TSoS}_{\geq}$ has marked inequality symbols $\{\geq_d\}_{d\in\mathbb{N}}$ and the square root $\sqrt{x}$ function symbol in the language. Expression $t \geq_d 0$ has informal meaning *"t is equal to a sum of squares of degree at most d"*. Additional axioms of $\mathsf{TSoS}_{\geq}$ are axioms of partially ordered ring for relations $\geq_d$, axioms for $\sqrt{x}$ and induction axiom scheme for formulas with connectives $\forall, \wedge$ and atomic formulas of the form $t = r$ and $t \geq_d r$ and all formulas without ring oracle or ring variables. We prove that this theory can be translated to $\mathsf{PC}^{+,\{2\}}$, which extends PC^{+} with auxiliary variables for square roots, and that $\mathsf{PC}^{+,\{2\}}$ is conservative over PC^{+}:

Theorem (Theorem 58)**.** *Let Π be a $\mathsf{TSoS}_{\geq}$ derivation of the sequent $\Gamma \longrightarrow \Delta$ such that all formulas in Γ and Δ are in Φ_{SDP} and have free index-variables $\bar{i}$. Then there exist $d \in \mathbb{N}$ such that for every assignment α for $\bar{i}$ and every witnessing function W_α for $\langle\Gamma\rangle^L_\alpha$ there exists a witnessing function W'_α for $\langle\Delta\rangle^R_\alpha$ and the following $\mathsf{PC}^{+,\{2\}}$ derivation of degree d:*

$$\langle\Gamma\rangle^L_\alpha(W_\alpha) \vdash \langle\Delta\rangle^R_\alpha(W'_\alpha)$$

Theorem (Theorem 55)**.** *Let $f_1, \ldots, f_m, g$ be real polynomials, not containing auxiliary variables of $\mathsf{PC}^{+,\{2\}}$. If there exist a $\mathsf{PC}^{+,\{2\}}$ derivation $\pi : f_1 = 0, \ldots, f_m = 0 \vdash g = 0$ of degree d and size S, then there exists PC^{+} derivation $\pi' : f_1 = 0, \ldots, f_m = 0 \vdash g = 0$ of degree $d^{2^{O(D)}}$ and size $2^{O(D)}S$, where D is the maximal level of nesting of square roots.*

This is joint work with Iddo Tzameret and Neil Thapen.

Chapter 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 $\mathcal{R}$, that is, a function of the form $\sum_{i=1}^{n} a_i x_i + a_0$ with $a_i \in \mathcal{R}$. We sometimes refer to a linear form as a *hyperplane*, since a linear form determines a hyperplane. We denote by $im_2(f)$ the image of f under 0-1 assignments to its variables; $\langle f \neq A \rangle := \bigvee_{A \neq B \in im_2(f)} (f = B)$, where $A \in \mathcal{R}$.

For ϕ a set of clauses or linear clauses (i.e., disjunctions of linear equations; see Section **??**), $vars(\phi)$ 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 $\mathcal{R}$ and A is a matrix over $\mathcal{R}$, 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 $\mathcal{R}$ and D is a linear clause over $\mathcal{R}$, denote by $\bigwedge_{C \in \phi} C \models D$ and $\bigwedge_{C \in \phi} C \models_{\mathcal{R}} D$ semantic entailment over 0-1 and $\mathcal{R}$-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 $\phi \restriction_\rho := (\ldots((\phi \restriction_{x_1 \leftarrow l_1}) \restriction_{x_2 \leftarrow l_2}) \ldots) \restriction_{x_n \leftarrow l_n}$.

2.2 Propositional Proof Systems

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

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

Definition 1 (Cook-Reckhow propositional proof system [28])**.** *A* propositional proof system Π *is a polynomial time computable onto function* $\Pi : \{0,1\}^* \to TAUT$.

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

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

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

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

A *resolution derivation* of a clause D from a set of clauses ϕ is a sequence of clauses $(D_1, \ldots, D_s \equiv D)$ such that for every $1 \leq i \leq s$ 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* [25] 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:

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

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

2.2.1 Hard Instances

2.2.1.1 Pigeonhole Principle

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

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

2.2.1.2 Mod p Tseitin Formulas

We use the version given in [2] (which is different from the one in [23, 60]). Let $G = (V, E)$ be a directed d-regular graph, that is a graph with incoming and outgoing degrees of every vertex are equal to d. 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 $u \in V$:

$$\sum_{(u,v)\in E} x_{(u,v)} - \sum_{(v,u)\in E} x_{(v,u)} \equiv \sigma(u) \mod p\,. \tag{2.1}$$

Note that we use the standard encoding of boolean functions as CNF formulas and the number of clauses, required to encode these equations is $O(2^d|V|)$. $\neg\mathrm{TS}^{(p)}_{G,\sigma}$ is unsatisfiable if and only if $\sum_{u\in V}\sigma(u) \not\equiv 0 \mod p$. To see this, note that if we sum (2.1) 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 total sum is 0, which is a contradiction.

In particular, $\neg\mathrm{TS}^{(2)}_{G,\sigma}$ are the classical Tseitin formulas [64] and $\mathrm{TS}^{(2)}_{G,1}$, where 1 is the constant function $v \mapsto 1$ (for all $v \in V$), expresses the fact that the sum of total degrees (incoming + outgoing) of the vertices is even.

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

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

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

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

Proposition 2 (Lubotzky *et. al* [50])**.** *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 $char(\mathbb{F}) \neq p$, 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 $\Omega(dn)$.

2.2.1.3 Random k-CNFs

A random k-CNF is a formula $\phi \sim \mathcal{F}_k^{n,\Delta}$ 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 $\phi \sim \mathcal{F}_k^{n,\Delta}, k \geq 3$ 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 $\mathbb{F}$.*

2.2.2 Error-Correcting Codes

Definition 2 ([?])**.** *Let $A : \mathbb{F}_q^k \hookrightarrow \mathbb{F}_q^n$ be a linear embedding. The image $C = im(A)$ of A is called $(n, k, d)_q$-code if for any $\overline{x}, \overline{y} \in C$ holds $d_H(\overline{x}, \overline{y}) \geq d$, where $d_H(\overline{x}, \overline{y}) = |\{i \,|\, x_i \neq y_i\}|$ is the Hamming distance. The matrix of A is called generator matrix for C.*

Theorem 5 (Gilbert bound [?])**.** *If q is a power of a prime and $n, k, d \in \mathbb{N}, n \geq k$ are such that inequality*

$$\sum_{i=1}^{d} \left(\begin{array}{c} n \\ i \end{array} \right) \cdot (q-1)^i < q^{n-k+1}$$

holds, then there exists $(n, k, d)_q$-code.

2.2.3 Complexity of Linear Systems

It is a well-known fact that deciding 0-1 satisfiability of linear systems over $\mathbb{F}_p, p \geq 5$ or of linear systems over $\mathbb{Q}$ (even if coefficients are small) are NP-complete problems. Indeed, for example, the 3-clause $(x_1 \vee \neg x_2 \vee x_3)$ can be represented as the linear equation with additional Boolean variables y_1, y_2: $x_1 + (1 - x_2) + x_3 = 1 + y_1 + y_2$. In this way k-SAT reduces to 0-1 satisfiability of linear systems over a field of characteristic 0 or $p > k$.

Theorem 6. *The problem of deciding 0-1 satisfiability of linear systems over a field of characteristic 0 or $p \geq 5$ is NP-complete. In case of characteristic 0 this also holds if the size of coefficients is required to be bounded by a constant.*

The mapping R of k-CNFs to linear systems described above can be used to translate lower bounds on degree of $PC_{\mathbb{F}}$ refutations from k-CNFs to linear systems.

Proposition 7. *If $\phi \in k$-UNSAT and $\mathbb{F}$ is a field such that $char(\mathbb{F}) > k$ or $char(\mathbb{F}) = 0$, then ϕ admits $PC_{\mathbb{F}}$ refutations of degree d iff $R(\phi)$ admits $PC_{\mathbb{F}}$ refutations of degree $O(d)$.*

Proof: Denote σ the mapping from literals to linear polynomials such that: $\sigma(x) := x$ and $\sigma(\neg x) := 1 - x$. Let τ be the following mapping from clauses to linear polynomials: $\tau(l_1 \vee \cdots \vee l_s) := \sigma(l_1) + \cdots + \sigma(l_s) - 1 - y^{(1)}_{l_1 \vee \cdots \vee l_s} - \cdots - y^{(s-1)}_{l_1 \vee \cdots \vee l_s}$, where $y^{(i)}_{l_1 \vee \cdots \vee l_s}$ are auxiliary Boolean variables. Then R translates $\phi = \{C_i\}_{i \in [m]}$ to the 0-1 unsatisfiable linear system L: $\tau(C_1) = 0, \ldots, \tau(C_m) = 0$.

Assume L has $PC_{\mathbb{F}}$ refutation π of degree d. If $x_1, \ldots, x_n$ are variables of ϕ, then all the auxiliary variables $y^{(i)}_{C_j}$ can be substituted with polynomials $v^{(i)}_{C_j}(x_1, \ldots, x_n)$ of degree at most k such that $C_j \models (\tau(C_j) \restriction_{\rho_v}) = 0$, where ρ_v stands for the substitution and the entailment is over 0-1 assignments. It is easy to see that π can be extended to the proof $\pi \restriction_{\rho_v}$ of degree at most $k \cdot d$, where all the auxiliary variables are substituted with the corresponding polynomials. Due to implicational completeness of $PC_{\mathbb{F}}$, there are $PC_{\mathbb{F}}$ derivations $\pi_j : C_j \vdash (\tau(C_j) \restriction_{\rho_v}) = 0$ of degree at most k. Composition of $\{\pi_j\}_{j \in [m]}$ with $\pi \restriction_{\rho_v}$ gives a $PC_{\mathbb{F}}$ refutation of degree at most $k \cdot d$.

Conversely, if π is a $PC_{\mathbb{F}}$ refutation of ϕ of degree d, then the composition of derivations $\tau(C_j) = 0 \vdash C_j$ with π gives a refutation of L of degree at most $\max(k, d)$. □

2.2.4 Semi-Algebraic Proof Systems

Let $\mathcal{F} = \{f_i(x_1, \ldots, x_n) = 0\}_{i \in [m]}$ and $\mathcal{H} = \{h_j(x_1, \ldots, x_n) \geq 0\}_{j \in [k]}$ be sets of polynomial equalities and inequalities over $\mathcal{R}$. We call the pair $(\mathcal{F}, \mathcal{H})$ an *SDP pair*. The following defines semi-algebraic analogue of the notion of ideal:

Definition 3. *The cone $c(h_1, \ldots, h_k)$, generated by $h_1, \ldots, h_k \in \mathcal{R}[x_1, \ldots, x_n]$, is the set of polynomials in $\mathcal{R}[x_1, \ldots, x_n]$, derivable from $h_1, \ldots, h_k$ by a sequence of applications of the following rules:*

$$\frac{p \qquad q}{p + q} \qquad \frac{p \qquad q}{p \cdot q} \qquad \frac{}{p^2}$$

Sum-of-squares(SoS):

A SoS derivation of $q \geq 0$ from $(\mathcal{F}, \mathcal{H})$ is a tuple $(g_1, \ldots, g_m, u_0, \ldots, u_k)$ such that:

$$\sum_{i \in [m]} g_i \cdot f_i + \sum_{i \in [k]} u_i \cdot h_i + u_0 = q$$

and all u_i are sums of squares of polynomials. A SoS refutation of $(\mathcal{F}, \mathcal{H})$ is a SoS derivation of -1.

Positivstellensatz(PS):

A PS derivation of $q \geq 0$ from $(\mathcal{F}, \mathcal{H})$ is a SoS derivation of $q \geq 0$ from $(\mathcal{F}, \hat{\mathcal{H}})$, where $\hat{\mathcal{H}} = \{\prod_{i \in I} h_i\}_{I \subseteq [k]}$. An inequality $q \geq 0$ admits PS derivation from $(\mathcal{F}, \mathcal{H})$ iff $q \in (f_1, \ldots, f_m) + c(h_1, \ldots, h_k)$. A PS refutation of $(\mathcal{F}, \mathcal{H})$, which is a PS derivation of -1, exists iff $(\mathcal{F}, \mathcal{H})$ is unsatisfiable (Stengle's Positivstellensatz).

Positivstellensatz calculus($\mathsf{PC}_>$):

$\mathsf{PC}_>$ is a dynamic version of the static system PS defined above. If $q = f + h$ and $f \in (f_1, \ldots, f_m)$, $h \in c(h_1, \ldots, h_k)$, then $\mathsf{PC}_>$ derivation of $q \geq 0$ is a PC derivation of f from $f_1, \ldots, f_m$ together with a PS derivation of h from $h_1, \ldots, h_k$.

The work in [16] shows that the static and dynamic versions of Positivstellensatz system are equivalent:

Corollary 2.2 **in [16].** *If $(\mathcal{F}, \mathcal{H})$ has $\mathsf{PC}_>$ refutation of degree d and size S, then it has a PS refutation of degree $2d$ and size poly(S).*

2.3 Sequent Calculus LK

The logical symbols of LK are: $\wedge, \vee, \neg, \supset, \forall, \exists$. A line in LK proof is called a *sequent* and is of the form $\Gamma \longrightarrow \Delta$, where Γ and Δ are multisets of formulas. Γ and Δ are called *cedents*, Γ is called *antecedent*, Δ is called *succedent*. The intended meaning of sequent $\phi_1, \ldots, \phi_n \longrightarrow \psi_1, \ldots, \psi_m$ is:

$$\phi_1 \wedge \ldots \wedge \phi_n \supset \psi_1 \vee \ldots \vee \psi_m$$

Definition 4. *An LK proof is a tree of sequents, where leaves are sequents of the form $\phi \longrightarrow \phi$ (axioms), the root is what is proved and any sequent, that is not a leaf, is obtained from its children by one of the following rules:*

1. *Structural rules:*

$$\frac{\Gamma \longrightarrow \Delta}{\Gamma, \phi \longrightarrow \Delta} \quad \textit{(Left weakening)} \qquad \frac{\Gamma \longrightarrow \Delta}{\Gamma \longrightarrow \Delta, \phi} \quad \textit{(Right weakening)}$$

$$\frac{\Gamma \longrightarrow \Delta, \phi, \phi}{\Gamma \longrightarrow \Delta, \phi} \quad \textit{(Contraction)}$$

2. *Left and right $\wedge$-introduction:*

$$\frac{\Gamma \longrightarrow \Delta, \phi \qquad \Gamma \longrightarrow \Delta, \psi}{\Gamma \longrightarrow \Delta, \phi \wedge \psi} \quad \textit{(Right)} \qquad \frac{\phi, \Gamma \longrightarrow \Delta}{\phi \wedge \psi, \Gamma \longrightarrow \Delta} \quad \textit{(Left)}$$

3. *Left and right $\vee$-introduction:*

$$\frac{\phi, \Gamma \longrightarrow \Delta \qquad \psi, \Gamma \longrightarrow \Delta}{\phi \vee \psi, \Gamma \longrightarrow \Delta} \quad \textit{(Left)} \qquad \frac{\Gamma \longrightarrow \Delta, \phi}{\Gamma \longrightarrow \Delta, \phi \vee \psi} \quad \textit{(Right)}$$

4. *Left and right $\supset$-introduction:*

$$\frac{\Gamma \longrightarrow \Delta, \phi \qquad \psi, \Gamma \longrightarrow \Delta}{\phi \supset \psi, \Gamma \longrightarrow \Delta} \quad \textit{(Left)} \qquad \frac{\phi, \Gamma \longrightarrow \Delta, \psi}{\Gamma \longrightarrow \Delta, \phi \supset \psi} \quad \textit{(Right)}$$

5. *Left and right $\neg$-introduction:*

$$\frac{\Gamma \longrightarrow \Delta, \phi}{\neg\phi, \Gamma \longrightarrow \Delta} \quad \textit{(Left)} \qquad \frac{\Gamma, \phi \longrightarrow \Delta}{\Gamma \longrightarrow \Delta, \neg\phi} \quad \textit{(Right)}$$

6. *Left and right $\exists$-introduction:*

$$\frac{\phi(b), \Gamma \longrightarrow \Delta}{\exists x \phi(x), \Gamma \longrightarrow \Delta} \quad \textit{(Left)} \qquad \frac{\Gamma \longrightarrow \Delta, \phi(t)}{\Gamma \longrightarrow \Delta, \exists x \phi(x)} \quad \textit{(Right)}$$

7. *Left and right $\forall$-introduction:*

$$\frac{\phi(t), \Gamma \longrightarrow \Delta}{\forall x \phi(x), \Gamma \longrightarrow \Delta} \quad \textit{(Left)} \qquad \frac{\Gamma \longrightarrow \Delta, \phi(b)}{\Gamma \longrightarrow \Delta, \forall x \phi(x)} \quad \textit{(Right)}$$

***Case 5**: Cut rule:*

$$\frac{\Gamma \longrightarrow \Delta, \phi \qquad \phi \longrightarrow \Delta}{\Gamma \longrightarrow \Delta}$$

LK as defined above is sound and complete proof system for first-order logic. In Chapter 4 we use two sorted (index sort and ring sort) version of LK extended with index-sort bounded universal quantifier with rules:

$$\frac{\phi(t), \Gamma \longrightarrow \Delta}{t < s \forall (x < s)\phi(x), \Gamma \longrightarrow \Delta} \quad \text{(Left)} \qquad \frac{b < s, \Gamma \longrightarrow \Delta, \phi(b)}{\Gamma \longrightarrow \Delta, \forall (x < s)\phi(x)} \quad \text{(Right)}$$

2.4 Propositional Translations

Here we sketch propositional translations, which are precursors to our translations in Chapter 4.

One of the classical and most important propositional translations was given by Paris and Wilkie [54] for the theories of bounded arithmetic S^i_2 and T^i_2. Every Σ^b_i formula $\phi(x)$ with parameter x is translated to a family of propositional formulas $\langle \phi \rangle_n, n \in \mathbb{N}$ of Σ'-depth i [1]. The proofs in S^i_2 and T^i_2 are translated to propositional sequent calculus PK proofs of Σ'-depth i. The following theorem establishes formal relation between bounded arithmetic proofs in S^i_2 and T^i_2 and PK proofs:

Theorem. *Let $\phi(x)$ be a Σ^b_i formula.*

1. *Suppose $S^i_2 \vdash \phi(x)$. Then there exists a function $S(n) = 2^{n^{O(1)}}$ such that for all n $\langle \phi \rangle_n$ has a PK proof of Σ'-depth i and of size $S(n)$. This proof has height $O(\log \log S(n))$.*

2. *Suppose $T^i_2 \vdash \phi(x)$. Then there exists a function $S(n) = 2^{n^{O(1)}}$ such that for all n $\langle \phi \rangle_n$ has a PK proof of Σ'-depth i and of size $S(n)$. This proof has height $O(\log S(n))$.*

In [14] a theories $U_{d,k}$-IND were defined and for the theory $U_{2,1}$-IND a rather simple translation to resolution was established.

[1] Σ'-depth is a slightly adjusted version of depth of propositional formula, that is of maximal nesting depth of $\wedge$, $\vee$ blocks in it. It doesn't count depth of small formulas at the bottom.

Chapter 3

Resolution over Linear Equations

3.1 Resolution with Linear Equations over General Rings

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

The rules of Res(lin$_R$) are as follows (cf. [60]):

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

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

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

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

A Res(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 Res(lin$_R$) *refutation* of an unsatisfiable set of linear clauses ϕ is a Res(lin$_R$) derivation of the empty clause (which stands for false) from ϕ. The ***size*** of a Res(lin$_\mathcal{R}$) 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 $\mathcal{R}$-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 $A, B \in \mathcal{R}$. For $\mu \in \{\omega, \omega_0\}$, the measure μ associated with a $\mathrm{Res}(\mathrm{lin}_R)$ derivation $\pi = (D_1, \ldots, D_s)$ is $\mu(\pi) := \max_{1\leq i\leq s} \mu(D_i)$. For $\phi \in$ UNSAT, denote by $\mu(\phi \vdash \bot)$ the minimal value of $\mu(\pi)$ over all $\mathrm{Res}(\mathrm{lin}_R)$ refutations π.*

Proposition 8. *$\mathrm{Res}(\mathrm{lin}_R)$ is sound and complete. It is also implicationally complete, that is if ϕ is a set of linear clauses and C is a linear clause such that $\phi \models C$, then there exists a $\mathrm{Res}(\mathrm{lin}_R)$ derivation of C from ϕ.*

Proof: The soundness can be checked by inspecting that each rule of $\mathrm{Res}(\mathrm{lin}_R)$ is sound. Implicational completeness (and thus completeness) follows from Proposition 29. □

We now define two systems of resolution with linear equations over a ring, where some of the rules are semantic: $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ and Sem-$\mathrm{Res}(\mathrm{lin}_R)$. $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ is obtained from $\mathrm{Res}(\mathrm{lin}_R)$ 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 Sem-$\mathrm{Res}(\mathrm{lin}_R)$ 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 $\mathrm{Res}(\mathrm{lin}_R) \leq_p \mathrm{Res}_{sw}(\mathrm{lin}_R) \leq_p$ Sem-$\mathrm{Res}(\mathrm{lin}_R)$, where $P \leq_p Q$ denotes that Q polynomially simulates P.

In contrast to the case $\mathcal{R} = \mathbb{F}_2$ (see [40]), for rings $\mathcal{R}$ with $char(\mathcal{R}) \notin \{1, 2, 3\}$ both $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ and Sem-$\mathrm{Res}(\mathrm{lin}_R)$ are not Cook-Reckhow proof systems, unless $\mathsf{P} = \mathsf{NP}$:

Proposition 9. *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 $char(\mathcal{R}) = 3$ or in case $char(\mathcal{R}) = 2$ and $\mathcal{R} \neq \mathbb{F}_2$. In the case $\mathcal{R} = \mathbb{F}_2$ 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 [40], respectively[1]. As showed in [40], $\text{Res}(\text{lin}_{\mathbb{F}_2})$, $\text{Res}_{sw}(\text{lin}_{\mathbb{F}_2})$ and $\text{Sem-Res}(\text{lin}_{\mathbb{F}_2})$ are polynomially equivalent.

We now show that if $char(\mathcal{R}) \notin \{1,2,3\}$, 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 10. *If $char(\mathcal{R}) \notin \{1,2,3\}$, 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 $C_1, \ldots, C_m$.

[1]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*).

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 11. *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 tree-like $\text{Res}(\text{lin}_R)$ is also a proof in tree-like $\text{Res}(\text{lin}_{Frac(R)})$. To get the converse, just multiply every line by the least common multiple of all the coefficients in the tree-like $\text{Res}(\text{lin}_{Frac(R)})$ proof. □

3.1.1 Basic Counting in $\text{Res}(\text{lin}_R)$ and $\text{Res}_{sw}(\text{lin}_R)$

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

Our unsatisfiable instances are the following:

Linear systems: If $A = (B|b)$ is an $m \times (n+1)$ matrix over $\mathcal{R}$, where the B sub-matrix

consists of the first n columns, such that $B\overline{x} = b$ has no 0-1 solutions, then (B_i is the ith row in B):

$$\mathsf{LinSys}(A) := \{B_i \cdot \overline{x} = b_i\}_{i \in [m]}. \tag{3.1}$$

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

$$\mathsf{SubSum}(f) := \{f = 0\}\,. \tag{3.2}$$

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

$$\mathsf{ImAv}\,(f) := \{\langle f \neq A\rangle : A \in im_2(f)\}\,. \tag{3.3}$$

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

Image axiom: For f a linear form, define

$$\mathsf{Im}(f) := \bigvee_{A \in im_2(f)} f = A\,. \tag{3.4}$$

Dag-Like $\mathrm{Res}(\mathrm{lin}_R)$

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

Proposition 12. *Let $f = \sum_{i=1}^{n} a_i x_i + b$ be a linear form 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 $\mathsf{Im}\Big(\sum_{i=1}^{k} a_i x_i + b\Big)$, $0 \leq k \leq n$, 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 13. *For every linear form f such that $0 \notin im_2(f)$, the contradiction* $\mathsf{SubSum}(f)$ *admits* $\mathrm{Res}(\mathrm{lin}_R)$ *refutation of size polynomial in* $|\mathsf{Im}(f)|$.

Proof: First construct the shortest derivation of $\mathsf{Im}(f)$, and then by a sequence of $|im_2(f)|$ many application of the resolution rule with $f = 0$ derive the empty clause. By Proposition 12 the resulting refutation is of polynomial in $|\mathsf{Im}(f)|$ size. □

Proposition 14. *Let f be a linear form over R, $a \in im_2(f)$ and $\phi = \{\langle f \neq b\rangle\}_{b \in im_2(f), b \neq a}$. Then there exists* $\mathrm{Res}(\mathrm{lin}_R)$ *derivation π of $f = a$ from ϕ, such that $S(\pi) = poly(|\phi|)$ and $\omega_0(\pi) \leq 3$.*

Proof: Let $A_1, \ldots, A_N = a$ be an enumeration of all the elements in $im_2(f)$. By Proposition 12 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 15. *For every linear form f the contradiction* $\mathsf{ImAv}(f)$ *admits polynomial-size* $\mathrm{Res}(\mathrm{lin}_R)$ *refutations.*

Proof: Pick some $a \in im_2(f)$. By Proposition 14 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 $\langle f \neq a\rangle \in \mathsf{ImAv}(f)$. □

In Section 3.2.2.1 we prove an upper bound for $\mathsf{LinSys}(A)$ in terms of the size of the image of the affine map, corresponding to A (Theorem 24). All other $\mathrm{Res}(\mathrm{lin}_{\mathcal{R}})$ 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 $\mathrm{Res}(\mathrm{lin}_{\mathcal{R}})$ upper bounds further in this section.

Lower bounds. In Sec. 3.2.1 we prove an exponential lower bound for $\mathsf{SubSum}(f)$ in case f is a linear form with large coefficients (Theorem 23).

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

Upper bounds. In case R is a finite ring, in Sec. 3.3.1 we prove that the clauses in $\mathsf{Im}(f)$ admit derivations of polynomial size (Theorem 30). 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 16. *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. 3.2.1 we prove tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ 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 $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ (see below).

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

Upper bounds. Most of the instances above admit short derivations/refutations in tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$: $\mathsf{Im}(f)$ is semantic weakening of $0 = 0$ and thus derivable in one step; The empty clause is a semantic weakening of $\mathsf{SubSum}(f)$ and $\mathsf{LinSys}(A)$ and thus can be refuted via deriving $\bigvee_{i\in[m]}\langle A_i\overline{x} - b_i \neq 0\rangle$ as a semantic weakening of $0 = 0$ and resolving it with equalities in $\mathsf{LinSys}(A) = \{A_i\overline{x} - b_i = 0\}_{i\in[m]}$.
Lower bounds. In case $\mathbb{F}$ is a field of characteristic zero, $\mathsf{ImAv}(f)$ are hard even for tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$ whenever f is of the form $f = \epsilon_1 x_1 + \ldots + \epsilon_n x_n - A$ for some $\epsilon_i \in \{-1, 1\}, A \in \mathbb{F}$ (Theorem 37).

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

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

and tree-like $\text{Res}(\text{lin}_{\mathbb{F}'})$, 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 17. *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 16).

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 29 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 16 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 $O\left(\sum_{i\in[k]} \left|\phi_i\right| \cdot |A_i \cdot \overline{x} - b_i = 0|\right)$ 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. 2.2.1.2 for the definition):

Theorem 18. *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 $\neg\text{TS}^{(p)}_{G,\sigma}$ admit tree-like $\text{Res}(\text{lin}_{\mathbb{F}_p})$ refutations of polynomial size.*

Proof: $\neg\text{TS}^{(p)}_{G,\sigma}$ is an unsatisfiable system of linear equations over $\mathbb{F}_p$ (note that no assignment of $\mathbb{F}$-elements to the variables in $\neg\text{TS}^{(p)}_{G,\sigma}$ 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 17 there exists a tree-like $\text{Res}(\text{lin}_{\mathbb{F}_p})$ refutation of $\neg\text{TS}^{(p)}_{G,\sigma}$ of polynomial size. □

Theorem 19 ([60]). *Let R be a ring such that $char(R) = 0$. There exists a* $\text{Res}(\text{lin}_R)$ *refutation of $\neg PHP_n^m$ of polynomial size.*

Proof: This follows from the upper bound of [60] for $\text{Res}(\text{lin}_{\mathbb{Z}})$ and the fact that any $\text{Res}(\text{lin}_{\mathbb{Z}})$ proof can be interpreted as $\text{Res}(\text{lin}_R)$ if R is of characteristic 0. □

3.2 Dag-Like Lower Bounds

3.2.1 Dag-Like Lower Bounds for the Subset Sum Principle

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

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

1. π' is at most polynomially larger than π.
2. C_π must be 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 small number Z_g of 0-1 solutions, namely Z_g is at most 2^{cn} for some $c < 1$. This together with the oservation that C_π must be a Boolean tautology, because it is derivable from Boolean axioms, implies that C_π must be of exponential size. 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 $\text{Res}(\text{lin}_{\mathbb{Q}})$ refutations and, thus, there is no hope to prove strong bounds for dag-like refutations in this case. However, in Sec 3.3 we prove that any $f = 0$, as long as f depends on n variables, must have tree-like $\text{Res}(\text{lin}_{\mathbb{Q}})$ refutations of size at least $2^{\Omega(\sqrt{n})}$. The argument relies on the similar transformation from refutations π of $f = 0$ to derivations of some C_π and in this way reduces the problem to proving tree-like $\text{Res}(\text{lin}_{\mathbb{Q}})$ lower bound on the size derivations of C_π from Boolean axioms.

For that reason we formulate and prove generalised statement about the translation. For both dag-like and tree-like bounds we essentially need that for all the disjuncts $g = 0$ in C_π some specific 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 and in case of tree-like bound $\mathcal{P}(g) = 1$ iff g depends on at least $\frac{n}{2}$ variables. In Theorem 20 we prove that the translation can be done as long as $\mathcal{P}$ satisfies certain properties.

Theorem 20. *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 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 clause $\bigvee_{j \in [N]} g_j = 0$, where $\mathcal{P}(g_j) = 1$ for every j.*

Proof: We now sketch the plan of the proof. Assume 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. 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 $D = \left(\bigvee_{t \in [m]} g_t = 0 \right)$ 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 function. Moreover, $\tau_k(D)$ satisfy 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: $|\bigcup_{D \in \pi_{\leq k}} H_D| \leq 2|\pi_{\leq k}|$.
3. The numbers a_t and coefficients of h_s are not too large: their size does not exceed the maximal size of coefficients in π.

Before we proceed to the inductive definition of π'_k, we finish the proof assuming π'_k described above exist. 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 lenght 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 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 $\pi_{\leq k}$.

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

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

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

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 22. *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 21 either $|im_2(g)| \leq 3^{\frac{n}{2}}$ or $|g^{-1}(0) \cap \{0,1\}^n| \leq 3^{\frac{n}{2}}$. □

Theorem 23. *Let $f = 1 + 2x_1 + \cdots + 2^n x_n$. Any* Res(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 22 $\mathcal{P}$ satisifes the properties in Theorem 20. Therefore, by Theorem 20, 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 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, it must contain at least $2^{(1-0.5 \cdot \log 3)n}$ disjuncts. Therefore $|\pi| = 2^{\Omega(n)}$. □

3.2.2 Linear Systems with Small Coefficients

In this section we study 0-1 unsatisfiable linear systems over finite fields.

Firslty, we prove an upper bound, which is polynomial in $|im_2(A\,\overline{x})|$, where $A = A_{f_1,\ldots,f_m} : \mathbb{F}^n \to \mathbb{F}^m$ is affine map $\overline{x} \mapsto (f_1(\overline{x}), \ldots, f_m(\overline{x}))$. In contrast to the case of a single equation $f = 0$, the size of the image $|im_2(A\,\overline{x})|$ does not fully characterise the size of the shortest Res(lin$_{\mathbb{F}}$) refutation of $f_1 = 0, \ldots, f_m = 0$: there is an example, where $|im_2(A\,\overline{x})|$ is large, but $S(f_1 = 0, \ldots, f_m = 0 \vdash \emptyset)$ is small.

Secondly, we prove a superpolynomial lower bound on a linear system for a restricted tree-like Res(lin$_{\mathbb{F}}$).

3.2.2.1 An Upper Bound

Denote $\langle A_{f_1,\ldots,f_m}\,\overline{x} \neq 0\rangle$ the linear clause $(\langle f_1 \neq 0\rangle \vee \cdots \vee \langle f_m \neq 0\rangle)$. The clause $\langle A_{f_1,\ldots,f_m}\,\overline{x} \neq 0\rangle$ is a tautology iff the system $f_1 = 0, \ldots, f_m = 0$ is 0-1 unsatisfiable. Therefore, any 0-1 unsatisfiable system $f_1 = 0, \ldots, f_m = 0$ can be refuted by first deriving $\langle A_{f_1,\ldots,f_m}\,\overline{x} \neq 0\rangle$ from Boolean axioms and then resolving it with $f_1 = 0, \ldots, f_m = 0$. We now prove an upper bound for derivations of $\langle A\,\overline{x} \neq 0\rangle$ in terms of $|im_2(A\,\overline{x})|$.

Theorem 24. *Let $f_1 = 0, \ldots, f_m = 0$ be a 0-1 unsatisfiable system with n variables. There exists a derivation of $\langle A_{f_1,\ldots,f_m}\,\overline{x} \neq 0\rangle$ of size $poly(n + |im_2(A_{f_1,\ldots,f_m}\,\overline{x})|)$.*

Proof: We arrange the derivation in n layers $L_0, \ldots, L_n$ in such a way that $L_0 := \{\langle A_{f_1,\ldots,f_m}\,\overline{x} \neq 0\rangle\}$ and

$$L_k := \{(\langle f_1 \restriction_{x_1\leftarrow\epsilon_1,\ldots,x_k\leftarrow\epsilon_k} \neq 0\rangle \vee \ldots \vee \langle f_m \restriction_{x_1\leftarrow\epsilon_1,\ldots,x_k\leftarrow\epsilon_k} \neq 0\rangle)\}_{\overline{\epsilon}\in\{0,1\}^k}$$

It is easy to see, that the following map is an embedding $L_k \hookrightarrow im_2(A_{f_1,\ldots,f_m}\,\overline{x})$:

$$(\langle f_1 \restriction_{x_1\leftarrow\epsilon_1,\ldots,x_k\leftarrow\epsilon_k} \neq 0\rangle \vee \ldots \vee \langle f_m \restriction_{x_1\leftarrow\epsilon_1,\ldots,x_k\leftarrow\epsilon_k} \neq 0\rangle) \mapsto$$
$$(f_1(\epsilon_1, \ldots, \epsilon_k, 0, \ldots, 0), \ldots, f_m(\epsilon_1, \ldots, \epsilon_k, 0, \ldots, 0))$$

Therefore $|L_k| \leq |im_2(A_{f_1,\ldots,f_m}\,\overline{x})|$.

It remains to note that every clause in L_k can be derived from clauses in L_{k+1} in $O(|im_2(A_{f_1,\ldots,f_m}\,\overline{x})|)$ steps. Indeed, if $C \in L_k$, then $C \restriction_{x_{k+1}\leftarrow 0} \in L_{k+1}$ and $C \restriction_{x_{k+1}\leftarrow 1} \in L_{k+1}$, and C can be derived from $C \restriction_{x_{k+1}\leftarrow 0}$ and $C \restriction_{x_{k+1}\leftarrow 1}$ and the axiom $(x_{k+1} = 0 \vee x_{k+1} = 1)$ in a standard way. □

Remark 25. *In contrast to the case of a single equation, dag-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$ *refutations of $f_1 = 0, \ldots, f_m = 0$ for $m \geq 2$ are not lower-bounded by $|im_2(A_{f_1,\ldots,f_m}\,\overline{x})|$ in general. For example, the system $x_1 - 2x_{n+1} = 0, x_n - 2x_{2n} = 0, x_{2n+1} + x_{n+1} + \ldots + x_{2n} - 2 = 0$ has refutation of size $O(n)$, but $|im_2(A_{f_1,\ldots,f_m}\,\overline{x})| = 2^{\Omega(n)}$.*

3.2.2.2 Lower Bound for Restricted Tree-Like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$

We define the following natural model of decision trees, certifying 0-1 unsatisfiability of linear systems over $\mathbb{F}$:

Definition 6. *Let $A\,\overline{x} = \overline{b}$ be a 0-1 unsatisfiable linear system over $\mathbb{F}$. A decision tree T for $A\,\overline{x} = \overline{b}$ is a binary tree, such that:*

- *Every internal node is labelled with a variable x_i and two branches correspond to assignments $x_i \leftarrow 0$ and $x_i \leftarrow 1$.*

- *If ρ_v is the variable assignment made along the path from the root to a leaf v, the system $(A\,\overline{x} = \overline{b}) \restriction_{\rho_v}$ is unsatisfiable over the whole field $\mathbb{F}$ (not just over 0-1).*

It is easy to see that this model of decision trees can be simulated by tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$. We argue that this model captures the strength of a natural fragment of tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$. If T is a decision tree for the system $f_1 = 0, \ldots, f_m = 0$ then a corresponding tree-like proof π for every leaf v in T derives the set of clauses

$$\left\{ \left(f_k \restriction_{\rho_v} = 0 \vee \bigvee_{i \in [n] \mid \rho_v(i) \neq *} x_i = 1 - \rho_v(i) \right) \right\}_{k \in [m]}$$

where $\rho_v : [n] \mapsto \{0, 1, *\}$ ($\rho_v(i) = *$ iff x_i is unassigned) is the assignment at v. By the leaf condition in Definition 6 the system $f_1 \restriction_{\rho_v} = 0, \ldots, f_m \restriction_{\rho_v} = 0$ is unsatisfiable over $\mathbb{F}$, therefore there exist $a_1, \ldots, a_m \in \mathbb{F}$ such that $a_1 f_1 \restriction_{\rho_v} + \cdots + a_m f_m \restriction_{\rho_v} = 1$ and the proof π uses this to derive further the clause $\bigvee_{i \in [n] \mid \rho_v(i) \neq *} x_i = 1 - \rho_v(i)$ from the clauses above for every leaf v. This is the *only place*, where counting is essentially used in π, the rest of the proof is just a standard resolution refutation obtained from T by the well-known correspondence between decision trees and tree-like resolution refutations. It is an interesting question whether this fragment is strictly weaker than full tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$.

We now prove a sub-exponential lower bound for this model and, consequently, for the corresponding fragment of tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}})$.

Theorem 26. *For every $n \in \mathbb{N}$ there exists a 0-1 unsatisfiable linear system $A\,\overline{x} = \overline{b}$ over a finite field $\mathbb{F}_q, q > 2$ with n variables such that any decision tree for this system is of size $2^{\Omega\left(\frac{n}{\log n}\right)}$.*

Proof: We construct the matrix A as a generator matrix of a linear $(n, k, d)_q = (n, \frac{n}{\log q} + 1, \Omega(\frac{n}{\log n}))_q$ error-correcting code (Definition 2).

The condition $k > \frac{n}{\log q}$, which this code satisfies, assures that $q^k > 2^n$ and therefore there exists $\overline{b} \in \mathbb{F}_q^k$ such that $A\,\overline{x} = \overline{b}$ is 0-1 unsatisfiable.

Note that depths of all leaves in any decision tree for $A\,\overline{x} = \overline{b}$ are at least d. Indeed, if $k < d$ variables are substituted at v by ρ_v, then the minimal distance of the code, generated by $A \restriction_{\rho_v}$, is at least $d - k$ and, in particular, $A \restriction_{\rho_v}$ has full rank, therefore v is not a leaf. Thus any decision tree for $A\,\overline{x} = \overline{b}$ has size at least $2^d = 2^{\Omega(\frac{n}{\log n})}$.

The existence of such a code is guaranteed by the Gilbert bound (Theorem 5). Recall that the Gilbert bound claims the existence of a linear $(n, k, d)_q$ code whenever

$$\sum_{i=1}^{d} \left(\begin{array}{c} n \\ i \end{array} \right) \cdot (q-1)^i < q^{n-k+1}$$

holds. In our case, if we assign $d = \frac{n}{10 \log n}$:

$$\sum_{i=1}^{d} \left(\begin{array}{c} n \\ i \end{array} \right) \cdot (q-1)^i < d \cdot q^{\frac{d \log n}{\log q}} \cdot q^d \leq \frac{n}{10 \log n} \cdot q^{n(\frac{1}{10 \log q} + \frac{1}{\log n})} < q^{n(1 - \frac{1}{\log q})+1}.$$

□

3.3 Tree-Like Lower Bounds

3.3.1 Nondeterministic Linear Decision Trees

In this section we extend the classical correspondence between tree-like resolution refutations and decision trees to tree-like $\mathrm{Res}(\mathrm{lin}_R)$ and tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_R)$. We define *nondeterministic linear decision trees* (NLDT), which generalize parity decision trees, proposed in [40] for $R = \mathbb{F}_2$, to arbitrary rings.

Both the Definition 7 of NLTDs and the proof of Theorem 28 are straightforward generalisations of standard decision trees and the proof of the correspondence between them and tree-like resolution respectively. We shall need these trees in the sequel to establish some of our novel upper and lower bounds.

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

DP1 Assume $\Phi \models \neg\phi$. Given a satisfying Boolean assignment ρ to Φ, determine which clause $C \in \phi$ is violated by ρ by making queries of the form: which of $f|_\rho \neq 0$ or $g|_\rho \neq 0$ hold for linear forms f, g in case $f|_\rho + g|_\rho \neq 0$.

DP2 Similar to DP1, only that we assume $\Phi \models_R \neg\phi$, and given R-valued assignment ρ, satisfying Φ, we ask to find a clause $C \in \phi$ falsified by ρ.

Below we define NLDTs of types $\mathrm{DT}_{sw}(R)$ and $\mathrm{DT}(R)$, which provide solutions to DP1 and DP2, respectively. The root of a tree is labeled with a system Φ, the edges in a tree are labeled with linear non-equalities of the form $f \neq 0$ and the leaves are labeled with clauses $C \in \phi$. Informally, at every node v there is a set Φ_v of all *learned*

non-equalities, which is the union of Φ and the set of non-equalities along the path from the root to the node. If v is an internal node, two outgoing edges $f \neq 0$ and $g \neq 0$ define a query to be made at v, where $f + g \neq 0$ is a consequence of Φ_v. If v is a leaf, then $\Phi_v \cup \Phi$ contradicts a clause $C \in \phi$.

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

Definition 7 (Nondeterministic linear decision tree NLDT; $\mathrm{DT}(R)$, $\mathrm{DT}_{sw}(R)$)**.** *Let ϕ be a set of linear clauses and Φ be a set of linear non-equalities over a ring R. A nondeterministic linear decision tree T of type $DT(R)$ and of type $DT_{sw}(R)$ for (ϕ, Φ) is a binary rooted tree, where every edge is labeled with some linear non-equality $f \neq 0$, in such a way that the conditions below hold. In what follows, for a node v, we denote by $\Phi_{r \rightsquigarrow v}$ the set of non-equalities along the path from the root r to v and by Φ_v the set $\Phi_{r \rightsquigarrow v} \cup \Phi$. We say that Φ_v is the set of* learned *non-equalities at v.*

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

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

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

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

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

Proposition 27. *If Φ is a set of linear non-equalities and ϕ is a set of linear clauses over $\mathcal{R}$ such that $\Phi \models \neg\phi$, then there exists a DT(R) tree for $(\phi \cup \{x = 0 \vee x = 1 \,|\, x \in vars(\phi \cup \{\neg\Phi\})\}, \Phi)$ of size $O(2^n|\Phi|)$, where $n = |vars(\phi \cup \{\neg\Phi\})|$.*

Proof: Let $vars(\phi \cup \{\neg\Phi\}) = \{x_1, \ldots, x_n\}$ 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 7. 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 7 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 7.

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

Example 2 Let ϕ be as in Example 1. *Parity decision trees*, as defined in [40], 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 3 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 $\mathrm{DT}(R)$ 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 28. *Let ϕ be a set of linear clauses over a ring $\mathcal{R}$ and Φ be a set of linear non-equalities over $\mathcal{R}$. 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 $\mathrm{DT}(R)$ 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 7, $\neg\Phi_u$ 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 7 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 $\alpha f + \beta g \neq 0 \in \{a \neq$

$0 \mid a \in R \setminus 0\}$, 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 7 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 $\phi \cup \{x = 0 \vee x = 1 \mid x \in vars(\phi)\} \cup \{0 = 0\}$, that is $C_u = (C \vee D)$ for some clause D. Therefore $\neg C_u \subseteq \Phi_u$ implies that C is falsified by Φ_u.

Case 2: Assume π is a tree-like $\text{Res}_{sw}(\text{lin}_R)$ 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 7 for $\text{DT}_{sw}(R)$ 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 this:

Proposition 29. *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* $\text{Res}(\text{lin}_R)$ *derivation of C from ϕ of size $O(2^n|C|)$, where $n = \left|vars(\phi \cup \{C\})\right|$.*

Proof: By Proposition 27 there exists a $\text{DT}(R)$ tree for $(\phi \cup \{x = 0 \vee x = 1 \mid x \in vars(\phi \cup \{C\})\}, \neg C)$ of size $O(2^n|C|)$ and, thus, by Theorem 28 there exists a tree-like $\text{Res}(\text{lin}_R)$ derivation of C from ϕ of size $O(2^n|C|)$. □

We construct an NLDT to prove the following upper bound:

Proposition 30. *Let R be a finite ring, $f = a_1x_1 + \cdots + a_nx_n$ a linear form over R, s_f the size of $\mathsf{lm}(f)$ (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 $\mathsf{lm}(f)$ 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 28 this implies the existence of a tree-like $\mathrm{Res}(\mathrm{lin}_R)$ proof of $\mathsf{lm}(f)$ 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 $\mathsf{lm}(f)$ is constructed recursively as a tree of height $2d_f$, where a subtree for $\mathsf{lm}\big(f^{(1)}\big)$ or for $\mathsf{lm}\big(f^{(2)}\big)$ 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: $s_n \leq 2^{2d_f} s_{\lfloor \frac{n}{2} \rfloor + 1} + s_f 2^{2d_f} = O(s_f n^{2d_f})$. □

3.3.2 Prover-Delayer Games

The *Prover-Delayer game* is an approach to obtain lower bounds on resolution refutations introduced by Pudlák and Impagliazzo [56]. 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 31) 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. Just as for NLDTs, our definition of Prover-Delayer games is not novel and is a straightforward generalisation of standard Prover-Delayer games as defined by Pudlák and Impagliazzo. They provide a handy language for tree-like $\mathrm{Res}(\mathrm{lin}_{\mathcal{R}})$ lower bound arguments. Very similar games for tree-like $\mathrm{Res}(\oplus)$ (that is tree-like $\mathrm{Res}(\mathrm{lin}_{\mathbb{F}_2})$ in our notation) were also studied in [40] and [39].

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 31. *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. □

Remark. The game, defined above, does not fully characterise the size of shortest NLDTs in the sense that lower bounds on size of NLDTs do not necessarily imply existence of a good strategy for Delayer. The characterisation gives tight bounds only for formulas, shortest NLDTs of which are symmetric, that is the size of the largest full binary tree, embedded in a shortest NLDT, is not much different from the size of shortest NLDT. In order to overcome this limitation, *asymmetric* Prover-Delayer games were introduced in [17], [18], [19] for the case of decision trees and tree-like resolution. At each round of such a game Prover and Delayer do the following:

1. Prover chooses an unassigned variable x.
2. Delayer assigns nonnegative weights p_0 and p_1, such that $p_0 + p_1 = 1$, to the two possible choices of the value for x.
3. Prover chooses value b, x is assigned to b and the score of Delayer is updated by $\log \frac{1}{p_b}$.

The standard game of Pudlák and Impagliazzo is a symmetric case of this game, where Delayer is only allowed to chooose weights $(1,0)$, $(0,1)$ or $(\frac{1}{2},\frac{1}{2})$. In [18] it was shown that this asymmetric game fully characterises decision trees. Namely, it was proved that:

1. If ϕ is unsatisfiable CNF, which has tree-like resolution refutation of size at most S, then there exists a strategy for Prover such that every strategy of Delayer scores at most $\log\lceil\frac{S}{2}\rceil$.

2. If ϕ is unsatisfiable CNF with shortest tree-like resolution refutation of size S. Then there is Delayer strategy, which scores at least $\log\lceil\frac{S}{2}\rceil$ against any strategy of the Prover.

Asymmetric games can similarly be defined for NLDTs and tree-like $\mathrm{Res}(\mathrm{lin}_{\mathcal{R}})$. An analogue of the statement (1.) trivially holds in this setting as well. It is interesting, though, whether an analogue of (2.) also holds. This is beyond the scope of our work, as the technique of symmetric games is enough for our needs.

3.3.3 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 3.2 we use Theorem 20 to transform refutations π of $f=0$ to derivations π' of a clause C_π from Boolean axioms. We ensure that π' is not much larger than π and C_π posesses 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 8. *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. *[[48]] 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 bounds below.

Theorem 33. *Any tree-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ *derivation of any tautology of the form* $\bigvee_{j\in[N]} g_j = 0$, *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 definitions in Sec. 3.3.2 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 31 obtain the lower bound.

If Φ is a position in the game, denote $\Phi_c \subset \Phi$ the subset of "coin" non-equalities, that is non-equalities that were chosen by Prover when Delayer decided to leave the choice to Prover and earned a coin. The number $|\Phi_c|$ is then precisely the number of coins earned by Delayer at Φ. Over 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 $|\Phi_I| = \Omega(\sqrt{|I|})$, for all $g \neq 0 \in (\Phi \restriction_{\rho_I}) \setminus (\Phi_c \restriction_{\rho_I})$ function g depends on at least $\frac{n}{2} - |I|$ variables, Φ_I contains variables only from I and $\Phi_c \restriction_{\rho_I}$ 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', which 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 hyperplanes $H_1, \ldots, H_k$ defined by equations $g_1 = 0, \ldots, g_k = 0$ form an essential cover of the cube $B_{|I'|}$. Therefore, by Theorem 32 $|\Psi''| = |\Psi'| - 1 \geq \sqrt{|I'|}$ and thus $|\Phi_{I\cup I'}| \geq \sqrt{|I|} + \sqrt{|I'|} \geq \sqrt{|I \cup I'|}$.

If necessary, Delayer repeats the above procedure constructing extensions $\rho_{I_1} \subset \ldots \subset \rho_{I_L}$ and $\Phi_{I_1} \subset \ldots \subset \Phi_{I_L}$, 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 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 $|\Phi_c| \geq |\Phi_I| \geq \sqrt{|I|} = \Omega(\sqrt{n})$. Thus the number of coins earned by Delayer is $\Omega(\sqrt{n})$.

□

Corollary 34. *If f is a linear polynomial over $\mathbb{Q}$, which depends on n variables, then tree-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ *derivations of* $\mathsf{lm}(f)$ *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 any tree-like* $\mathrm{Res}(\mathrm{lin}_{\mathbb{Q}})$ *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 conditions in Theorem 20 with respect to f. Therefore by Theorem 20 for every refutation π of $f = 0$ there exists a derivation π' of a clause C_π from 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

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

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. *Let f be a linear function over $\mathbb{F}$, which depends on n variables. Then tree-like $\mathrm{Res}_{sw}(\mathrm{lin}_{\mathbb{F}})$ refutation of $\mathsf{ImAv}(f)$ is of size $2^{\Omega(n)}$.*

Proof: According to definitions in Sec. 3.3.2 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 31 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 not inconsistent with Φ, then choose $g_1 \neq 0$;

2. if $g_1 \neq 0$ is inconsistent with Φ, but $g_2 \neq 0$ is not inconsistent 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 $|\Psi| \geq \frac{n}{4}$.

□

3.3.4 Lower Bounds for the Pigeonhole Principle

Here we prove that every tree-like $\text{Res}_{sw}(\text{lin}_{\mathbb{F}})$ refutations of $\neg\text{PHP}_n^m$ must have size at least $2^{\frac{n-1}{2}}$ (see Sec. 2.2.1.1 for the definition of $\neg\text{PHP}_n^m$). Together with the upper bound for dag-like $\text{Res}(\text{lin}_{\mathbb{F}})$ (see Sec. 3.1.2) this provides a separation between tree-like and dag-like $\text{Res}_{sw}(\text{lin}_{\mathbb{F}})$ in the case $char(\mathbb{F}) = 0$. The lower bound argument is comprised of exhibiting a strategy for Delayer in the Prover-Delayer game. Delayer's strategy is similar to that in [40]. However, the proof that Delayer's strategy guarantees sufficiently many coins relies on Lemma 39, which is a generalization of Lemma 3.3 in [40] for arbitrary fields. Since the proof of Lemma 3.3 in [40] 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 $\text{Res}_{sw}(\text{lin}_{\mathbb{F}})$ refutation of $\neg PHP_n^m$ has size at least $2^{\frac{n-1}{2}}$.*

Proof: We prove that there exists a strategy for Delayer in the $\neg\text{PHP}_n^m$ game, which guarantees Delayer to earn $\frac{n-1}{2}$ coins. Following the terminology in [40], we call an assignment $x_{i,j} \mapsto \alpha_{ij}$, for $\alpha \in \{0,1\}^{mn}$, *proper* if it does not violate Holes_n^m, 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 $A\overline{x} \doteq \overline{b}$ 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, $0 \doteq \overline{b}$. 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 $A\overline{x} \doteq \overline{b}$ 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*

$$Y(l) := \left\{ (y_1, \ldots, y_n) \in \mathbb{F}^n \mid \forall i \in [n], 0 < y_i \leq 2, \text{ and } \sum_{i=1}^{n} y_i \geq l \right\}.$$

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 $A\overline{x} \doteq \overline{b}$ 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 $A\overline{x} \doteq \overline{b}$. 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 $A(\alpha \oplus \beta) \doteq \overline{b}$.*

Proof: Let $I \subseteq \{0,1\}^n$. Denote by $A^\star_I$ 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, $A^\star_I$ 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. *$A(\alpha \oplus \beta) \doteq \overline{b}$ iff $A^\star_I \beta_I \doteq \overline{b} - A\alpha$.*

Proof of claim: We prove that $A(\alpha \oplus \beta) = A^\star_I \beta_I + A\alpha$. Consider any row $\mathbf{v}$ in A, and the corresponding row $\mathbf{v}^\star_I$ in $A^\star_I$. 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

$$\mathbf{v}_{[n]\setminus I} \cdot (\alpha \oplus \beta)_{[n]\setminus I} = \mathbf{v}_{[n]\setminus I} \cdot \alpha_{[n]\setminus I} \,. \tag{3.5}$$

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

$$\mathbf{v}_I \cdot (\alpha \oplus \beta)_I = \mathbf{v}_I \cdot \alpha_I + \mathbf{v}_I^\star \cdot \beta_I \,. \tag{3.6}$$

Equation (3.6) 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 (3.6). 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 (3.6) 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 (3.6).

The two equations (3.5) and (3.6) concludes the claim. ■Claim

We know that $A\alpha \doteq \overline{b}$, 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 $A(\alpha \oplus \beta) \doteq \overline{b}$. By the claim above it remains to show the existence of such β where $A_I^\star \beta_I \doteq \overline{b} - A\alpha$. But notice that $\overline{b} - A\alpha \doteq 0$, since $A\alpha \doteq \overline{b}$, and that $A_I^\star \beta_I$ is a matrix of dimension $k \times (k+1)$. Therefore, by Lemma 39, the system $A_I^\star \beta_I \doteq \overline{b} - A\alpha$ has a nonzero solution, that is, there exists a $\beta \neq 0$ for which all ones are in the I entries, such that $A_I^\star \beta_I \doteq \overline{b} - A\alpha$. □

Lemma 41. *Assume that a system $A\overline{x} \doteq \overline{b}$ 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 $A\overline{x} \doteq \overline{b}$, 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 $A\overline{x} \doteq \overline{b}$, 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 $\Phi \wedge \mathsf{Holes}_n^m \models f_i = 0$, for some $i \in \{0, 1\}$, Delayer chooses $f_{1-i} \neq 0$, which must satisfy $\Phi \wedge \mathsf{Holes}_n^m \models f_{1-i} \neq 0$, 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 $\Phi \models \neg\mathsf{Pigeons}_n^m$. 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 $\Phi \models \neg\mathsf{Pigeons}_n^m$, then $\Phi' \wedge \mathsf{Holes}_n^m \models \neg\mathsf{Pigeons}_n^m$ 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. □

3.4 Size-Width Relation and Simulation by PC

In this section we prove a size-width relation for tree-like $\mathrm{Res}(\mathrm{lin}_R)$ (Theorem 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 12. Assume

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

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 $S(\widehat{\pi}) = O(S(\pi) + |\mathsf{lm}(l)|)$ 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 $\mathrm{Res}_{sw}(\mathrm{lin}_\mathbb{F})$:*

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

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

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

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

Induction step: Let π be a tree-like refutation of $\phi = \{C_1, \ldots, C_m\}$ such that $S(\pi) = S(\phi \vdash \bot)$ (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 Res($\mathrm{lin}_{\mathbb{F}}$) (resp. tree-like Res$_{sw}$($\mathrm{lin}_{\mathbb{F}}$)) as a simplification (resp. weakening) of an unsatisfiable $h = 0$ ($1 = 0$ in case of tree-like Res($\mathrm{lin}_{\mathbb{F}}$)) equation, which is derived by application of the resolution rule. Denote the left and right subtrees, corresponding to the premises of $h = 0$, by π_1 and π_2, respectively.

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

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

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

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

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

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

$$\phi \restriction_{\rho_A} \cup \left\{(A - l)a_n^{-1} = 0 \vee (A - l)a_n^{-1} = 1\right\} \tag{3.7}$$

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 (3.7) 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 14 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 Res($\mathrm{lin}_{\mathbb{F}}$) (resp. tree-like 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 $\mathrm{char}(\mathbb{F}) \neq p$, 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 $\neg\mathrm{TS}_{G,\sigma}^{(p)}$ is $\Omega(dn)$. Theorem 45 implies that the principal width of $\mathrm{Res}(\mathrm{lin}_\mathbb{F})$ refutations of $\neg\mathrm{TS}_{G,\sigma}^{(p)}$ 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 $1 - o(1)$. □

Chapter 4

First-Order Theories for Constant Degree $\mathsf{PC}_{\mathcal{R}}$ and SoS

In this chapter we present a formulation of first-order theories $\mathsf{TPC}_{\mathcal{R}}$ and TSoS, which are uniform versions of constant degree $\mathsf{PC}_{\mathcal{R}}$ and SoS, respectively. We start with the definition of $\mathsf{TPC}_{\mathcal{R}}$ and subsequently use $\mathsf{TPC}_{\mathbb{R}}$ as a basis for the definition of TSoS.

4.1 The Theory for Constant Degree $\mathsf{PC}_{\mathcal{R}}$

For a fixed ring $\mathcal{R}$, the theory $\mathsf{TPC}_{\mathcal{R}}$ is a two-sorted theory over the language $\mathcal{L}^{\mathcal{R}}_{=}$. Its two sorts are ring sort and index sort.

4.1.1 The Language $\mathcal{L}^{\mathcal{R}}_{=}$ of $\mathsf{TPC}_{\mathcal{R}}$

Index sort symbols:

- Index sort function symbols f for all functions $f : \mathbb{N} \to \mathbb{N}$. It follows that $\mathcal{L}^{\mathcal{R}}_{=}$, in particular, contains index sort constants for all $n \in \mathbb{N}$ and symbols $f(n_1, \ldots, n_k)$ for polynomially bounded functions of several arguments.
- Equality predicate symbol $=_{ind}$. We usually omit the subsrcipt.

Ring sort symbols:

- Constants for all $a \in \mathcal{R}$.
- Function symbols for all functions $f : \mathbb{N} \to \mathcal{R}$.
- Function symbols $+, -, \cdot$ for ring operations.

- Function symbol $\sum_{i,t}(n)$ for every index variable i and ring-term $t(i)$, where n is an index argument. Intended meaning: if $t(i)$ is ring term with free variable i and n is index term, then $\sum_{i,t}(n+1) := t(0) + \ldots + t(n), \sum_{i,t}(0) := 0$. We write sum-terms in the conventional form $\sum_{i=0}^{n} t(i)$.
- Oracle[1] $X(i)$, where i is index argument. Intended meaning: $X(i)$ states for i^{th} variable in the sequence of variables X.
- Equality predicate symbol $=_{rng}$. We usually omit the subscript.

4.1.2 The Axioms of $\mathsf{TPC}_{\mathcal{R}}$

Basic axioms:

- Every true sentence[2], not containing occurrences of the oracle X and ring-variables.
- Ring-sort and index-sort equalities axiom scheme:

$$\forall \overline{x}\ \forall \overline{y}\ \forall \overline{i}\ \forall \overline{j}\ \overline{x} = \overline{y}, \overline{i} = \overline{j}, t(\overline{x}, \overline{i}) = 0 \supset t(\overline{y}, \overline{j}) = 0$$

- Standard ring axioms for $0, 1 \in \mathcal{R}$ and $+, -, \cdot$.
- The big sum defining axiom schemes:

$$\sum_{i=0}^{j+1} t(i) = \sum_{i=0}^{j} t(i) + t(j) \qquad \sum_{i=0}^{0} t(i) = 0$$

Induction axiom:

For every formula $\phi(i)$ in the class $\Phi^{\mathcal{R}}_{=}$, which we define below, the axiom:

$$\phi(0) \wedge (\forall i\, \phi(i) \supset \phi(i+1)) \supset \forall n\, \phi(n)$$

Definition 9. *The class $\Phi^{\mathcal{R}}_{=}$ of $\mathcal{L}^{\mathcal{R}}_{=}$-formulas is defined by induction on complexity of formulas as follows:*

- *All atomic formulas are in $\Phi^{\mathcal{R}}_{=}$ and all formulas not containing occurrences of the oracle X or ring-variables are in $\Phi^{\mathcal{R}}_{=}$.*
- *If $\phi_1, \phi_2 \in \Phi^{\mathcal{R}}_{=}$, then $\phi_1 \vee \phi_2 \in \Phi^{\mathcal{R}}_{=}$ and $\phi_1 \wedge \phi_2 \in \Phi^{\mathcal{R}}_{=}$.*
- *If $\phi(i) \in \Phi^{\mathcal{R}}_{=}$, then $\forall (i < s)\phi(i) \in \Phi^{\mathcal{R}}_{=}$.*

[1] Technically, just a function symbol without any defining axioms.
[2] True in the standard model.

4.1.3 Propositional Translation for $\mathsf{TPC}_{\mathcal{R}}$

In this section we establish a connection between first-order $\mathsf{TPC}_{\mathcal{R}}$ derivations and propositional $\mathsf{PC}_{\mathcal{R},d}$ or $\mathsf{PC}^{\mathrm{rad}}_{\mathcal{R},d}$ derivations. Given an assignment α for index-variables $\bar{i}$, we define a translation of $\mathcal{L}^{\mathcal{R}}_{=}$-formulas $\phi(\bar{i}, y_1, \ldots, y_n) \in \Phi^{\mathcal{R}}_{=}$ with free index-variables $\bar{i}$ and free ring-variables $y_1, \ldots, y_n$ to sets of polynomial equations $\langle\phi\rangle_\alpha = \{f_1 = 0, \ldots, f_m = 0\}, f_i \in \mathcal{R}[x_0, \ldots, x_{s(\alpha)}, y_1, \ldots, y_n]$ such that

$$\{(a_1, \ldots, a_{s(\alpha)}, b_1, \ldots, b_n) \mid \phi(\alpha(\bar{i}), b_1, \ldots, b_n) \restriction_{X(j)\leftarrow a_j, j\in[s(\alpha)]} = \mathsf{True}\} = V(\langle\phi\rangle_\alpha)$$

where V denotes the set of solutions of a system of polynomial equations. We then show, that a first-order refutation of a set of formulas $\phi_1(\bar{i}), \ldots, \phi_k(\bar{i}) \in \Phi^{\mathcal{R}}_{=}$ can be translated to a family π_α of constant degree refutations of $\{\langle\phi_1\rangle_\alpha, \ldots, \langle\phi_k\rangle_\alpha\}_\alpha$ in $\mathsf{PC}_{\mathcal{R},d}$ or in its extension $\mathsf{PC}^{\mathrm{rad}}_{\mathcal{R},d}$.

4.1.3.1 Extension of $\mathsf{PC}_{\mathcal{R}}$ with The Radical Rule

The system $\mathsf{PC}^{\mathrm{rad}}_{\mathcal{R}}$ (respectively $\mathsf{PC}^{\mathrm{rad}}_{\mathcal{R},d}$) extends the system $\mathsf{PC}_{\mathcal{R}}$ (respectively $\mathsf{PC}_{\mathcal{R},d}$) with the following *radical rule*:

$$\frac{f^2 = 0}{f = 0}$$

This extension makes the system strictly stronger with respect to derivations (even with Boolean axioms) in case of fields of characteristic 0 as shown in the following proposition:

Proposition 48. *If $\mathbb{F}$ is a field of characteristic 0, then $\mathsf{PC}_{\mathbb{F}}$ derivations*

$$\{x_i^2 - x_i = 0\}, (x_1 + \ldots + x_n + 1)^2 = 0 \vdash x_1 + \ldots + x_n + 1 = 0$$

are of degree $\Omega(n)$.

Proof: Write such a derivation in static form as $(x_1 + \ldots + x_n + 1) = a \cdot (x_1 + \ldots + x_n + 1)^2 + \sum_i h_i \cdot (x_i^2 - x_i)$, where a, h_i are some polynomials. Clearly, a on 0-1 assignments is the function $a(\overline{x}) \equiv \frac{1}{x_1+\ldots+x_n+1}$. From Corollary 5.4 in [31] it follows that the degree of a is $\Omega(n)$. □

However, whether in this case the system $\mathsf{PC}^{\mathrm{rad}}_{\mathbb{F}}$ is strictly stronger than $\mathsf{PC}_{\mathbb{F}}$ as a refutation system is still an open question.

The following proposition shows that in case of fields of positive characteristic the situation is different:

Proposition 49. *If $\mathbb{F}$ is a field of positive characteristic, then there exist $\mathsf{PC}_{\mathbb{F}}$ derivations $f^2 = 0 \vdash f = 0$ of degree $O(deg(f))$.*

Proof: The derivation is $f = f^{p-2} \cdot f^2$, where p is characteristic of $\mathbb{F}$. □

4.1.3.2 Translation of Terms and Formulas

Translation of terms. Let α be an assignment for index variables $\bar{i}$. By induction on terms, we define the translation $\langle t(\bar{i}, \overline{y})\rangle_\alpha$ to polynomials over $\mathcal{R}$ of terms with free index-variables $\bar{i}$ and free ring-variables $\overline{y}$ as follows:

- All the constants $a \in \mathcal{R}$ are translated to the corresponding element in the ring $\mathcal{R}$. For $i \in \mathbb{N}$, $X(i)$ is translated to the variable x_i. A ring-variable y_i is translated to the variable y_i.
- Operations $+, -, \cdot$ are translated to the corresponding operations on polynomials.
- $\langle \sum_i (t(i), n)\rangle_\alpha := \langle t(1)\rangle_\alpha + \ldots + \langle t(n)\rangle_\alpha$.

Translation of formulas. If ϕ is a formula in $\Phi^{\mathcal{R}}_{=}$ we define propositional translation ϕ to a set of polynomial equations $\langle\phi\rangle_\alpha$ as follows:

- If ϕ is atomic formula $t = r$, then $\langle\phi\rangle_\alpha := \{\langle t\rangle_\alpha - \langle r\rangle_\alpha = 0\}$.
- If ϕ is a formula, not containing occurrences of the oracle X or ring-variables, then $\langle\phi\rangle_\alpha := \{0 = 0\}$ if $\phi \restriction_\alpha = \mathsf{True}$ and $\langle\phi\rangle_\alpha := \{1 = 0\}$ otherwise.
- If $\phi = \psi \wedge \psi'$, then $\langle\phi\rangle_\alpha := \langle\psi\rangle_\alpha \cup \langle\psi'\rangle_\alpha$.
- If $\phi = \psi \vee \psi'$, then $\langle\phi\rangle_\alpha := \langle\psi\rangle_\alpha \cdot \langle\psi'\rangle_\alpha$, where the product of two sets of polynomials is defined to be $\mathcal{P} \cdot \mathcal{Q} := \{p \cdot q = 0 \,|\, p = 0 \in \mathcal{P}, q = 0 \in \mathcal{Q}\}$.
- If $\phi = \forall(i < s)\psi(i)$, then $\langle\phi\rangle_\alpha := \{\langle\psi(v)\rangle_\alpha\}_{v<s\restriction_\alpha}$

4.1.3.3 Propositional Translation of $\mathsf{TPC}_{\mathcal{R}}$ Proofs

We work with $\mathsf{TPC}_{\mathcal{R}}$ proofs as the sequent calculus LK derivations. If Γ is an antecedent, denote $\langle\Gamma\rangle^L_\alpha := \langle\bigwedge_{\phi\in\Gamma}\phi\rangle_\alpha$ and if Δ is a succedent denote $\langle\Delta\rangle^R_\alpha := \langle\bigvee_{\phi\in\Delta}\phi\rangle_\alpha$. The following Lemma is a routine verification:

Lemma 50. *Every basic axiom of $\mathsf{TPC}_{\mathcal{R}}$ can be written as a sequent, where all formulas are in $\Phi^{\mathcal{R}}_{=}$. The induction axiom can be defined by the rule:*

$$\frac{\phi(i) \longrightarrow \phi(i+1)}{\phi(0) \longrightarrow \phi(i)}$$

The following theorem relates TPC derivations to $\mathsf{PC}^{\mathrm{rad}}_d$ derivations:

Theorem 51. *Let Π be a $TPC_{\mathcal{R}}$ derivation of the sequent $\Gamma \longrightarrow \Delta$ such that all formulas in Γ and Δ are in $\Phi^{\mathcal{R}}_{=}$ and have free index-variables $\bar{i}$. Then there exist $d \in \mathbb{N}$ such that for every assignment α for $\bar{i}$ there exists $\mathsf{PC}^{rad}_{\mathcal{R},d}$ ($\mathsf{PC}_{\mathcal{R},d}$ in case $\mathcal{R}$ is a field of positive characteristic) derivation:*

$$\langle \Gamma \rangle^L_\alpha \vdash \langle \Delta \rangle^R_\alpha$$

Proof: By the free-cut elimination theorem for two-sorted LK, there exists free-cut free proof Π' of $\Gamma \longrightarrow \Delta$. By Lemma 50, all axioms of $\mathsf{TPC}_{\mathcal{R}}$ can be written as sequents, where all formulas are in $\Phi^{\mathcal{R}}_{=}$. Additionally, all formulas appearing in the induction rule are from $\Phi^{\mathcal{R}}_{=}$. Therefore, by the subformula property of free-cut free proofs, all formulas in Π' must be in $\Phi^{\mathcal{R}}_{=}$.

The proof is by induction on the number of steps in Π'.
Base case: It is easy to see, that if t is a term, then $\langle t \rangle_\alpha$ is a family of polynomials with degree bounded by a constant. All axioms but for equality axioms are translated to trivial statements. Recall the equality axiom:

$$\overline{x} = \overline{y}, \bar{i} = \bar{j}, t(\overline{x}, \bar{i}) = 0 \longrightarrow t(\overline{y}, \bar{j}) = 0$$

Denote $\Gamma_=$ antecedent above. In case $\alpha(\bar{i}) \neq \alpha(\bar{j})$, 1 is in $\langle \Gamma_= \rangle^L_\alpha$. Otherwise $\langle \Gamma_= \rangle^L_\alpha = \{\overline{x} - \overline{y}, p(\overline{x})\}$, where $p(\overline{x}) := \langle t(\overline{x}, \alpha(\bar{i})) \rangle_\alpha = \langle t(\overline{x}, \alpha(\bar{j})) \rangle_\alpha$, and there is an obvious derivation $\{\overline{x} - \overline{y}, p(\overline{x})\} \vdash p(\overline{y})$ in $\mathsf{PC}^{\mathrm{rad}}_{\mathcal{R},d}$.

Induction step: The cases of structural rules (weakening and contraction) are trivial, except for the contraction rule:

$$\frac{\Gamma \longrightarrow \Delta, \phi, \phi}{\Gamma \longrightarrow \Delta, \phi} \quad \text{(Contraction)}$$

By induction hypothesis there exists $\mathsf{PC}^{\mathrm{rad}}_{\mathcal{R},d}$ derivation $\langle \Gamma \rangle^L_\alpha \vdash \langle \Delta \rangle^R_\alpha \cdot \langle \phi \rangle^2_\alpha$. Applying the radical rule we obtain $\langle \Gamma \rangle^L_\alpha \vdash \langle \Delta \rangle^R_\alpha \cdot \langle \phi \rangle_\alpha$.

Other rules are handled as follows:
Case 1: Left and right $\wedge$-introduction:

$$\frac{\Gamma \longrightarrow \Delta, \phi \qquad \Gamma \longrightarrow \Delta, \psi}{\Gamma \longrightarrow \Delta, \phi \wedge \psi} \quad \text{(Right)} \qquad \frac{\phi, \Gamma \longrightarrow \Delta}{\phi \wedge \psi, \Gamma \longrightarrow \Delta} \quad \text{(Left)}$$

(Left) We just use the proof of $\langle \phi, \Gamma \rangle^L_\alpha \vdash \langle \Delta \rangle^R_\alpha$ as the proof of $\langle \phi \wedge \psi, \Gamma \rangle^L_\alpha \vdash \langle \Delta \rangle^R_\alpha$.
(Right) The $\mathsf{PC}^{\mathrm{rad}}_{\mathcal{R},d}$ derivation of $\langle \Gamma \rangle^L_\alpha \vdash \langle \Delta, \phi \wedge \psi \rangle^R_\alpha = \langle \Delta, \phi \rangle^R_\alpha \cup \langle \Delta, \psi \rangle^R_\alpha$ is just the union of $\mathsf{PC}^{\mathrm{rad}}_{\mathcal{R},d}$ derivations of $\langle \Gamma \rangle^R_\alpha \vdash \langle \Delta, \phi \rangle^R_\alpha$ and of $\langle \Gamma \rangle^L_\alpha \vdash \langle \Delta, \psi \rangle^R_\alpha$.
Case 2: Left and right $\vee$-introduction:

$$\frac{\phi, \Gamma \longrightarrow \Delta \qquad \psi, \Gamma \longrightarrow \Delta}{\phi \vee \psi, \Gamma \longrightarrow \Delta} \quad \text{(Left)} \qquad \frac{\Gamma \longrightarrow \Delta, \phi}{\Gamma \longrightarrow \Delta, \phi \vee \psi} \quad \text{(Right)}$$

(Right) By induction hypothesis there is $\mathsf{PC}^{\mathrm{rad}}_{\mathcal{R},d}$ derivation $\pi : \langle \Gamma \rangle^L_\alpha \vdash \langle \Delta \rangle^R_\alpha \cdot \langle \phi \rangle_\alpha$. It is trivially extended to a derivation of $\langle \Delta \rangle^R_\alpha \cdot \langle \phi \rangle_\alpha \cdot \langle \psi \rangle_\alpha$.
(Left) By induction hypothesis there are derivations $\pi_1 : \langle \phi \rangle_\alpha, \langle \Gamma \rangle^L_\alpha \vdash \langle \Delta \rangle^R_\alpha$ and $\pi_2 : \langle \psi \rangle_\alpha, \langle \Gamma \rangle^L_\alpha \vdash \langle \Delta \rangle^R_\alpha$. Obvious composition of π_1 and π_2 yields $\langle \phi \rangle_\alpha \cdot \langle \psi \rangle_\alpha, \langle \Gamma \rangle^L_\alpha \vdash (\langle \Delta \rangle^R_\alpha)^2 \vdash \langle \Delta \rangle^R_\alpha$.
Case 3: Left and right bounded index $\forall$-introduction:

$$\frac{\phi(l), \Gamma \longrightarrow \Delta}{l < s, \forall_{i<s} \phi(i), \Gamma \longrightarrow \Delta} \quad \text{(Left)} \qquad \frac{i < s, \Gamma \longrightarrow \Delta, \phi(i)}{\Gamma \longrightarrow \Delta, \forall_{j<s} \phi(j)} \quad \text{(Right)}$$

where variable i does not occur in Γ or Δ in the (Right) rule.
(Left) If $l \restriction_\alpha \geq s \restriction_\alpha$, then $1 \in \langle l < s, \forall_{i<s} \phi(i), \Gamma \rangle^L_\alpha$. Otherwise, $\langle \phi(l) \rangle_\alpha$ is subset of $\langle \forall_{i<s} \phi(i), \Gamma \rangle_\alpha$ and the statement trivially follows.
(Right) By induction hypothesis there exists derivation $\langle i < s, \Gamma \rangle^L_{\alpha[i \leftarrow v]} \vdash \langle \Delta, \phi(i) \rangle^R_{\alpha[i \leftarrow v]}$ for all $v \in \mathbb{N}$ and all assignments α. The family of derivations $\{\langle \Gamma \rangle^L_\alpha \vdash \langle \Delta \rangle^R_\alpha \cdot \langle \phi(v) \rangle_\alpha\}_{v < s \restriction_\alpha}$ constitute the derivation $\langle \Gamma \rangle^L_\alpha \vdash \langle \Delta \rangle^R_\alpha \cdot \{\langle \phi(v) \rangle_\alpha\}_{v < s \restriction_\alpha} = \langle \Delta \rangle^R_\alpha \cdot \langle \forall_{j<s} \phi(j) \rangle_\alpha$.
Case 4: Induction rule:

$$\frac{\Gamma, \phi(i) \longrightarrow \phi(i+1), \Delta}{\Gamma, \phi(0) \longrightarrow \phi(i), \Delta}$$

where variable i does not occur in Γ or Δ.

Let α be assignments and let $n := \alpha(i)$. By induction hypothesis there are derivations $\pi_v : \langle \Gamma \rangle^L_\alpha \cup \langle \phi(i) \rangle_{\alpha[i \leftarrow v]} \vdash \langle \phi(i+1) \rangle_{\alpha[i \leftarrow v]} \cdot \langle \Delta \rangle^R_\alpha$. By multiplying π_v by $\langle \Delta \rangle^R_\alpha$ and applying radical rule we obtain derivation $\pi'_v : \langle \Gamma \rangle^L_\alpha \cdot \langle \Delta \rangle^R_\alpha \cup \langle \phi(i) \rangle_{\alpha[i \leftarrow v]} \cdot \langle \Delta \rangle^R_\alpha \vdash \langle \phi(i+1) \rangle_{\alpha[i \leftarrow v]} \cdot (\langle \Delta \rangle^R_\alpha)^2 \vdash \langle \phi(i+1) \rangle_{\alpha[i \leftarrow v]} \cdot \langle \Delta \rangle^R_\alpha$. Multiplication by $\langle \Delta \rangle^R_\alpha$ and concatenation with $\pi'_0, \ldots, \pi'_{n-1}$ results in the derivation $\langle \Gamma \rangle^L_\alpha \cup \langle \phi(0) \rangle_\alpha \vdash \langle \Gamma \rangle^L_\alpha \cdot \langle \Delta \rangle^R_\alpha \cup \langle \phi(0) \rangle_\alpha \cdot \langle \Delta \rangle^R_\alpha \vdash \ldots \vdash \langle \phi(n) \rangle_\alpha \cdot \langle \Delta \rangle^R_\alpha$.
Case 5: Cut rule:

$$\frac{\Gamma \longrightarrow \Delta, \phi \qquad \phi, \Gamma \longrightarrow \Delta}{\Gamma \longrightarrow \Delta}$$

By induction hypothesis there are derivations $\pi_1 : \langle\Gamma\rangle^L_\alpha \vdash \langle\Delta\rangle^R_\alpha \cdot \langle\phi\rangle_\alpha$ and $\pi_2 : \langle\phi\rangle_\alpha \cup \langle\Gamma\rangle^L_\alpha \vdash \langle\Delta\rangle^R_\alpha$. Construct the desired derivation as follows: $\langle\Gamma\rangle^L_\alpha \vdash \langle\Delta\rangle^R_\alpha \cdot \langle\phi\rangle_\alpha \cup \langle\Gamma\rangle^L_\alpha \cdot \langle\Delta\rangle^R_\alpha \vdash \langle\Delta\rangle^R_\alpha$. □

4.2 Theories for Constant Degree SoS

We define TSoS as a minimalistic extension of $\mathsf{TPC}_\mathbb{R}$, which reflects the strength of SoS_d. The theory TSoS extends $\mathsf{TPC}_\mathbb{R}$ with just one axiom, expressing that if a sum of squares is zero, then every square in the sum is zero. We ensure that TSoS is strong enough by showing that TSoS proves soundness of SoS_d.

Subsequently, we discuss possible formulations of theories with inequality symbol in the language.

Theory $\mathsf{TSoS}_\geq$ is the strongest theory with inequality, for which we can prove the existence of a translation to SoS. The language of $\mathsf{TSoS}_\geq$ contains marked inequalities $\{\geq_d\}_{d\in\mathbb{N}}$ and square root function symbol $\sqrt{x}$. Expression $t \geq_d r$ informally means *"$t - r$ is a sum-of-squares of degree at most d"*. $\mathsf{TSoS}_\geq$ is intuitionistic and is built on top of restricted version of $\mathsf{TPC}_\mathbb{R}$: $\mathsf{TPC}_\mathbb{R}$ without integral domain axiom and induction axiom scheme restricted to formulas with $\forall, \wedge$ connectives (that is without $\vee$). It also contains axioms of *partially* ordered ring for marked inequality relations $\geq_d$ and natural axioms for square root $\sqrt{x}$, excluding monotonicity axiom: $x \geq_d y \supset \sqrt{x} \geq_d \sqrt{y}$. Induction axiom scheme is defined for formulas of the form $\forall\overline{x}\ \bigwedge_i t_i(\overline{x}) \geq_d 0 \wedge \phi$, where ϕ is a $(\forall, \wedge)$-formula, where all atomic subformulas are equalities. The proofs in $\mathsf{TSoS}^1_\geq$ are naturally translated to constant degree derivations in an extension of $\mathsf{PC}^{\mathrm{rad}}_{\mathbb{R},d}$. We extend one of the results in [16] to show that these extensions are simulated by SoS.

One might consider the following modifications to the theory $\mathsf{TSoS}_\geq$:

1. Replace marked inequalities $\{\geq_d\}_{d\in\mathbb{N}}$ with a single unmarked inequality $\geq$.
2. Remove universal quantifier for inequalities in induction.
3. Add integral domain axiom for equalities and $\vee$ connective in formulas in induction.
4. Add totality axiom for inequality.
5. Allow classical reasoning.
6. Add monotonicity axiom for $\sqrt{x}$.

7. Add other fractional powers $1/k, k \geq 3$.

All modification, except for (2), are strengthenings. Currently, we only know that if we apply one of (1), (3) or (4) and do not apply (2), then the resulting theory T will be too strong, because T proves soundness of resolution. All other comibations are not yet ruled out by a proof that corresponding theory is stronger than constant degree SoS.

4.2.1 Extensions of $\mathsf{PC}^{\mathbf{rad}}_{\mathbb{R}}$

4.2.1.1 The system PC^{+}

This system extends $\mathsf{PC}^{\mathrm{rad}}_{\mathbb{R}}$ with the following rule:

$$\frac{f_1^2 + \ldots + f_m^2 = 0}{f_1^2 = 0}$$

We show that PC^{+} can be simulated by SoS. Our argument is an extension of simulation of $\mathsf{PC}_{\mathbb{R}}$ in SoS described in [16]. The following lemma demonstrates that SoS "almost simulates" the radical rule.

Lemma 52. *Let f be a polynomial of degree $d/2$. Then for every $\epsilon > 0$ there exist degree d SoS_d derivations of $f^2 = 0 \vdash f \geq -\epsilon$ and $f^2 = 0 \vdash f \leq \epsilon$. Moreover the coefficient of f^2 in the derivation is a negative real number.*

Proof: Let $\epsilon > 0$. The following is SoS_d derivation of $\epsilon \pm f \geq 0$:

$$\epsilon \pm f = \frac{1}{4\epsilon} \cdot (-f^2 + (2\epsilon \pm f)^2)$$

□

Using this lemma we prove that SoS "almost simulates" PC^{+} with respect to derivations.

Proposition 53. *Let $r_1 = 0, \ldots, r_L = 0$ be PC^{+} derivation of degree d from a set of equalities $\mathcal{F} = \{f_1, \ldots, f_m\}$. Then for every $\epsilon > 0$ there exists degree $2d$ SoS derivation of $-r_L^2 + \epsilon \geq 0$ from $\mathcal{F}$.*

Proof: We prove by induction on L that $-r_L^2 + \epsilon \geq 0$ has SoS proof of degree $2d$. Consider all possible ways of how $r_L = 0$ is derived. In case $r_L = 0$ is an axiom from $\mathcal{F}$ or a Boolean axiom, the SoS derivation is trivial.

Let $r_L = 0$ be derived by variable rule from $r_k = 0$: $r_L = x_j r_k$ for some variable x_j. By induction hypothesis there exists SoS derivation π of $-r_k^2 + \epsilon \geq 0$ of degree $2d$. We derive $-r_L^2$ by adding to π the following expression: $(r_k - x_j r_k)^2 + (-2r_k^2)(x_j^2 - x_j)$.

Let $r_L = 0$ be derived from $r_k = 0$ and $r_{k'} = 0$ by sum rule: $r_L = ar_k + br_{k'}$, where $a, b \in \mathbb{R}$. By induction hypothesis there exist SoS derivations π of $-r_k^2 + \frac{\epsilon}{2a^2} \geq 0$ and π' of $-r_{k'}^2 + \frac{\epsilon}{2b^2} \geq 0$ both of degree $2d$. The following is a derivation of $-r_L^2 + \epsilon \geq 0$: $2a^2\pi + 2b^2\pi' + (ar_k - br_{k'})^2$.

Let $r_L = 0$ be derived from $r_k = 0$ by the radical rule: $r_L^2 = r_k$. By induction hypothesis there exists an SoS derivation of $-r_L^4 + \delta \geq 0$ for every $\delta > 0$. From the proof of Lemma 52 we conclude that there exists SoS derivation of $-r_L^2 + 2\epsilon' \geq 0$ from $-r_L^4 + 4\epsilon'^2 \geq 0$ for every $\epsilon' > 0$ and, in particular, for $\epsilon' = \frac{\sqrt{\delta}}{2}$. Thus there exists derivation of $-r_L^2 + \sqrt{\delta} \geq 0$, where we choose $\delta = \epsilon^2$.

Let $r_L = 0$ be derived from $r_k = 0$, where $r_L = f_1^2$ and $r_k = f_1^2 + \cdots + f_m^2$, by sum-of-squares rule. By induction hypothesis there exists an SoS derivation of $-(f_1^2 + \cdots + f_m^2)^2 + \delta \geq 0$ for every $\delta > 0$. It follows that there exists an SoS derivation of $-f_1^4 + \delta \geq 0$ for every $\delta \geq 0$. □

As a corollary we obtain that SoS simulates PC^+.

Theorem 54. *If there exists a PC^+ refutation of degree d of a set of equalities $\mathcal{F}$, then there exists SoS refutation of $\mathcal{F}$ of degree $2d$.*

4.2.1.2 The system $\mathsf{PC}^{+,\mathcal{P}}$

Let $V = \{x_1, \ldots, x_n, \ldots\}$ be the set of variables used in systems of polynomials being refuted. The system $\mathsf{PC}^{+,\mathcal{P}}$ extends PC^+ with auxiliary variables $X_{k,Q}$ for $Q^{1/k}$ for every $k \in \mathcal{P} \subseteq \mathbb{N}\backslash\{0,1\}$ and every polynomial Q, possibly containing auxiliary variables. Circularity is avoided by arranging variables of $\mathsf{PC}^{+,\mathcal{P}}$ in the family of sets U_i: $U_0 = V$ and U_{i+1} consists of all variables $X_{k,Q}$, where Q is a real polynomial with variables in $\bigcup_{j \leq i} U_j$. For all $k \in \mathcal{P}$, all polynomials P, Q and all s.o.s polynomials A, not containing $X_{k,Q}$, $\mathsf{PC}^{+,\mathcal{P}}$ has the rules:

$$\frac{Q - A = 0}{(X_{k,Q})^k - Q = 0} \qquad \frac{Q - A = 0}{X_{k,Q^k} - Q = 0}$$

if k is even and

$$\frac{}{(X_{k,Q})^k - Q = 0} \qquad \frac{}{X_{k,Q^k} - Q = 0}$$

if k is odd.

We now prove that $\mathsf{PC}^{+,\{2\}}$ is a conservative extension of PC^{+}.

Theorem 55. *Let $f_1, \ldots, f_m, g$ be real polynomials, not containing auxiliary variables of $\mathsf{PC}^{+,\{2\}}$. If there exist a $\mathsf{PC}^{+,\{2\}}$ derivation $\pi : f_1 = 0, \ldots, f_m = 0 \vdash g = 0$ of degree d and size S, then there exists PC^{+} derivation $\pi' : f_1 = 0, \ldots, f_m = 0 \vdash g = 0$ of degree $d^{2^{O(D)}}$ and size $2^{O(D)}S$, where D is the maximal level of nesting of square roots.*

Proof: Let $x_1, \ldots, x_n$ be variables of $f_1, \ldots, f_m, g$. Denote $z_1, \ldots, z_M$ the auxiliary variables in π of the maximal nesting level D and $y_1, \ldots, y_N$ all other auxiliary variables. We prove, that π can be converted to a proof π' without variables $\overline{z}$ of degree $O(d^2)$ and size $O(S)$. The claim will follow by induction on D.

Denote $Q_i(\overline{x}, \overline{y})$ the square of z_i and π_i the derivation of $Q_i - \sum_j u_{i,j}^2 = 0$, that is subderivation of π. Denote $\mathcal{I}$ the ideal, generated by $\{z_i^2 - Q_i\}$. By induction on the steps in π we show that if $p = 0$ is a line in π and $p \equiv \sum\limits_{I \subseteq [M]} P_I(\overline{x}, \overline{y}) z_I \ mod\ \mathcal{I}$, where $z_I = \prod\limits_{i \in I} z_i$, then for each $I \subseteq [M]$ there exist a proof π_I of $P_I \cdot Q_I$, where $Q_I = \prod\limits_{i \in I} Q_i$. Moreover, each π_I is of degree at most $O(d^2)$ and size $O(S)$.

The base case is obvious. Consider rules of $\mathsf{PC}^{+,\{2\}}$ one by one:

Case 1: Sum rule. This case is obvious.

Case 2: Variable rule. For x_i or y_i variables the case is obvious. Assuming the statement holds for $p = 0$, we prove that it also holds for $p \cdot z_i = 0$. We multily separatly each monomial. For all monomials z_I such that $i \notin I$, $P_I \cdot Q_I \cdot Q_i$ is derived by multiplication of $P_I \cdot Q_I$ by Q_i. If $i \in I$, then $P_I \cdot z_I \cdot z_i \equiv P_I \cdot Q_i \cdot z_{I \setminus \{i\}} \mod \mathcal{I}$ and $P_I \cdot Q_I$ derivable by hypothesis.

Case 3: Radical rule. Assuming the statement holds for $p^2 = 0$, we prove it for $p = 0$. Note that $p^2 \equiv \sum_I \left(\sum\limits_{J,J' : J \Delta J' = I} P_J \cdot P_{J'} \cdot Q_{J \cap J'} \right) z_I \ mod\ \mathcal{I}$. By induction hypothesis we know that $\sum\limits_{J,J' : J \Delta J' = I} P_J \cdot P_{J'} \cdot Q_{J \cap J'} \cdot Q_I = \sum\limits_{J,J' : J \Delta J' = I} P_J \cdot P_{J'} \cdot Q_{J \cup J'} = 0$ are derivable for every I. In particular, for $I = \emptyset$, $\sum\limits_J P_J^2 \cdot Q_J = 0$ is derivable. Because $\sum\limits_J P_J^2 \cdot Q_J$ is a sum of squares (each Q_J is sum of squares as a product of sums of squares $Q_i, i \in J$ by proofs $\pi_i, i \in J$), by sum-of-squares rule $P_J \cdot Q_J = 0$.

Case 4: Sum-of-squares rule. Analogous to the previous case.

Case 5: Auxiliary variables rules. This case is trivial. □

4.2.2 Theory TSoS

The languages of TSoS and $\mathsf{TPC}_{\mathbb{R}}$ coincide, inequality $t \geq 0$ is phrased in TSoS by means of saying that t equals to a sum of squares.
The axioms of TSoS are axioms of $\mathsf{TPC}_{\mathbb{R}}$ plus the following axiom for all terms $t(i)$:

$$\sum_{i}^{n} t(i)^2 = 0 \wedge j < n \supset t(j) = 0$$

The translation of $\mathsf{TPC}_{\mathbb{R}}$ trivially extends to a translation from TSoS to PC^+.

4.2.2.1 Soundness of SoS_d in TSoS

We show that whenever there exists an SoS_d refutation π of a system of equations $\mathcal{F}$, the TSoS encoding of the statement that $\mathcal{F}$ has satisfying assignment A can be refuted in TSoS using the TSoS encoding of π.

We first describe TSoS phrasing of the statement. Let n be an index-variable and let a polynomial[3] $P \in \mathbb{R}[x_1, \ldots, x_n]_d$ of degree at most d be represented in TSoS as a ring-sort function symbol $a_P(i, n)$, where $a_P(i, n)$ represents the coefficient of i^{th} monomial in P and monomials are ordered according to the deglex ordering. In particular, linear combination $\alpha P + \beta Q$ and product $P \cdot Q$ are represented by functions $a_{\alpha P+\beta Q}(i, n)$ and $a_{P \cdot Q}(i, n)$ respectively, and the following equalities are axioms of TSoS:

$$a_{\alpha P+\beta Q}(i, n) = \alpha a_P(i, n) + \beta a_Q(i, n) \qquad a_{P \cdot Q}(i, n) = \sum_{j,k}^{M_d(n)} \delta_{\odot(j,k,i,n)} \cdot a_P(j, n) \cdot a_Q(k, n)$$

where $M_d(n)$ is the function symbol for the number of monomials with n variables and of degree at most d, and $\odot(j, k, i, n)$ is the predicate symbol, expressing that i^{th} monomial is the product of j^{th} monomial and k^{th} monomial.

An assignment for variables is represented by the ring-oracle A: x_i is assigned to $A(i)$. Let $\nu_1(i, n), \ldots, \nu_d(i, n)$ be functions such that $\nu_1(i, n) \leq \ldots \leq \nu_d(i, n)$ and m_i is the monomial $\prod_{j|\nu_j(i,n)>0} x_{\nu_j(i,n)}$. The evaluation $m_i[A]_d$ of i^{th} monomial m_i on A depends on the chosen upper bound d for the degree and is defined as follows:

$$m_i[A]_d := \prod_{1 \leq j \leq d} \left(\delta_{\nu_j(i,n)>0} \cdot (A(\nu_j(i, n)) - 1) + 1\right)$$

This extends to evaluation $P[A]_d := \sum_i a_P(i, n) m_i[A]_d$ of a polynomial P on A.

[3] Actually, a family of polynomials parameterized by n. We refer to it as a polynomial for brevity.

Lemma 56. *Let P, Q be some polynomials of degree at most $d \in \mathbb{N}$. TSoS proves $(P+Q)[A]_d = P[A]_d + Q[A]_d$ and $(P \cdot Q)[A]_{2d} = P[A]_{2d} \cdot Q[A]_{2d}$.*

Proof: $(P+Q)[A]_d = P[A]_d + Q[A]_d$ follows from $\sum_{i=0}^{N}(t(i)+r(i)) = \sum_{i=0}^{N} t(i) + \sum_{i=0}^{N} r(i)$ for any terms t and r, which is in turn provable by induction in TSoS.

The destributivity $\left(\sum_{i=0}^{N} t(i)\right) \cdot r = \sum_{i=0}^{N} t(i) \cdot r$ for sum-terms is provable by induction in TSoS. It follows that TSoS proves:

$$P[A]_{2d} \cdot Q[A]_{2d} = \left(\sum_{i=0}^{M_{2d}(n)} a_P(i,n) m_i[A]_{2d}\right) \cdot \left(\sum_{i=0}^{M_{2d}(n)} a_Q(i,n) m_i[A]_{2d}\right) =$$

$$= \sum_{i=0}^{M_{2d}(n)} \left(\sum_{j=0}^{M_{2d}(n)} a_P(i,n) \cdot a_Q(j,n) \cdot m_i[A]_{2d} \cdot m_j[A]_{2d}\right)$$

Degree bounds $deg(P) \leq d$ and $deg(Q) \leq d$ imply that the following statement is an axiom of TSoS: $\left(\bigwedge_{s \leq d} \nu_s(i,n) = 0 \wedge \bigwedge_{s \leq d} \nu_s(j,n) = 0\right) \vee (a_P(i,n) \cdot a_Q(j,n) = 0)$. Both disjuncts imply $a_P(i,n) \cdot a_Q(j,n) \cdot m_i[A]_{2d} \cdot m_j[A]_{2d} = a_P(i,n) \cdot a_Q(j,n) \cdot m_{k(i,j)}[A]_{2d}$, where $m_{k(i,j)}$ is the product of m_i and m_j. Therefore:

$$P[A]_{2d} \cdot Q[A]_{2d} = \sum_{i=0}^{M_{2d}(n)} \left(\sum_{j=0}^{M_{2d}(n)} a_P(i,n) \cdot a_Q(j,n) \cdot m_{k(i,j)}[A]_{2d}\right) =$$

$$= \sum_{i=0}^{M_{2d}(n)} \left(\sum_{j=0}^{M_{2d}(n)} \sum_{k=0}^{M_{2d}(n)} a_P(i,n) \cdot a_Q(j,n) \cdot \delta_{k=k(i,j)} \cdot m_k[A]_{2d}\right) =$$

$$= \sum_{k=0}^{M_{2d}(n)} m_k[A]_{2d} \cdot \left(\sum_{i=0}^{M_{2d}(n)} \sum_{j=0}^{M_{2d}(n)} a_P(i,n) \cdot a_Q(j,n) \cdot \delta_{\odot(i,j,k)}\right) = (P \cdot Q)[A]_{2d}$$

□

We now prove the soundness:

Theorem 57. *If there exists a family $\pi_n = (\{g_{i,n}\}, \{u_{k,n}\}_{k \in [M_n]})$ of SoS_d refutations of $\mathcal{F}_n = \{f_{i,n} = 0\}_{i \in [N_n]}$, that is $\sum_i g_{i,n} f_{i,n} + \sum_k u_{k,n}^2 = -1$, then TSoS proves:*

$$\forall n \exists (i < N_n)\ \neg(f_{i,n}[A]_{2d} = 0)$$

.

Proof: Assume, for contradiction, $\forall(i < N_n)\ f_{i,n}[A]_{2d} = 0$ and derive $\left(\sum_{i=1}^{N_n} f_{i,n}g_{i,n}\right)[A]_{2d} = 0$ using Lemma 56 and the induction of TSoS. As a consequence we derive $\left(\sum_{k=1}^{M_n} u_{k,n}^2 + 1\right)[A]_{2d} = 0$, which, in turn, by Lemma 56 and TSoS induction implies $\sum_{k=1}^{M_n}(u_{k,n}[A]_{2d})^2 + 1 = 0$ and by the sum-of-squares rule this implies $1 = 0$. □

4.2.2.2 Theory $\mathsf{TSoS}_{\geq}$

The language of $\mathsf{TSoS}_{\geq}$ extends the language of $\mathsf{TPC}_{\mathbb{R}}$ with predicate symbols $\{\geq_d\}_{d\in\mathbb{N}}$ and function symbol $\sqrt{x}$. The underlying logic of $\mathsf{TSoS}_{\geq}$ is intuitionistic, that is excluded middle $\phi \vee \neg\phi$ is not an axiom.

Axioms of $\mathsf{TSoS}_{\geq}$

- All axioms of $\mathsf{TPC}_{\mathbb{R}}$, except for integral domain axioms and induction, are also axioms of $\mathsf{TSoS}_{\geq}$.
- Axioms of partial order for $\geq_d$.
- Axioms of interaction of $\geq_d$ with ring operations:

$$\forall x \forall y \forall z\ x \geq_d y \supset x + z \geq_d y + z$$

$$\forall x \forall y\ x \geq_d 0 \wedge y \geq_{d'} 0 \supset x \cdot y \geq_{d+d'} 0$$

- Axioms $\forall x \forall y\ x \geq_d y \supset x \geq_{d+d'} y$.
- Squares are nonnegative: if t is a term of degree d (i.e. $\sup_\alpha deg(\langle t\rangle_\alpha) = d$), then $t^2 \geq_{2d} 0$.
- Axioms for square roots $\sqrt{x}$:

$$\forall x\ x \geq_d 0 \supset \sqrt{x} \geq_d 0$$

$$\forall x\ x \geq_d 0 \supset (\sqrt{x})^2 = x$$

$$\forall x\ x \geq_d 0 \supset \sqrt{x^2} = x$$

- Induction axiom scheme for formulas with connectives $\forall$ and $\wedge$. We denote Φ_{SDP} the resulting set of formulas for $\mathsf{TSoS}_{\geq_d}$ induction.

Propositional translation for $\mathsf{TSoS}_{\geq}$

The propositional translation for formulas of $\mathsf{TSoS}_{\geq}$ extends inductive definition of the translation of formulas of TSoS to atomic formulas of the form $t \geq_d 0$. The translation $\langle t \geq_d 0 \rangle_\alpha(u)$ is parameterized by a sum of squares u and is defined to be $\langle t \rangle_\alpha - u = 0$. Consequently, if ϕ is a formula, then its translation $\langle \phi \rangle_\alpha(W_\alpha)$ is parameterized by a *witnessing function* $W_\alpha : t \geq_d 0 \mapsto (u_t)$, where the domain of W_α includes all inequalities in the corresponding formula and the degree bound is respected: $deg(u_t) \leq d$. This gives rise to a translation of cedents with inequalities.

Theorem 58. *Let Π be a $\mathsf{TSoS}_{\geq}$ derivation of the sequent $\Gamma \longrightarrow \Delta$ such that all formulas in Γ and Δ are in Φ_{SDP} and have free index-variables $\bar{i}$. Then there exist $d \in \mathbb{N}$ such that for every assignment α for $\bar{i}$ and every witnessing function W_α for $\langle \Gamma \rangle^L_\alpha$ there exists a witnessing function W'_α for $\langle \Delta \rangle^R_\alpha$ and the following $\mathsf{PC}^{+,\{2\}}$ derivation of degree d:*

$$\langle \Gamma \rangle^L_\alpha(W_\alpha) \vdash \langle \Delta \rangle^R_\alpha(W'_\alpha)$$

Proof: The cut-elimination theorem for LK implies that there exists a free-cut free derivation Π' of $\Gamma \longrightarrow \Delta$. By the subformula property of free-cut free proofs all formulas in Π' are from Φ_{SDP}.

The proof is by induction on the number of steps in Π'.

Base case: The proof for all axioms of $\mathsf{TPC}_{\mathbb{R}}$ is the same as in Theorem 51. The proofs for most of axioms for $\geq_d$ are trivial. We give the proof for the case of the following axioms:

Case 1: Antisymmetry for $\geq_d$: $t \geq_d r, r \geq_d t \longrightarrow r = t$. We need to prove that for all sums-of-squares A, B of degree at most d holds $\langle t \rangle_\alpha - \langle r \rangle_\alpha - A = 0, \langle r \rangle_\alpha - \langle t \rangle_\alpha - B = 0 \vdash \langle t \rangle_\alpha - \langle r \rangle_\alpha = 0$. Indeed, the sum of the premises is $A + B = 0$, which by the sum-of-squares rule imples $A = 0$ and $B = 0$.

Case 2: Axiom $t \geq_d 0, r \geq_{d'} 0 \longrightarrow t \cdot r \geq_{d+d'} 0$. For any SoS A of degree $\leq d$ and any SoS B of degree $\leq d'$ we derive $\langle t \rangle_\alpha \cdot \langle r \rangle_\alpha - A \cdot B = 0$ as follows: $\langle t \rangle_\alpha \cdot \langle r \rangle_\alpha - A \cdot B = (\langle t \rangle_\alpha - A) \cdot (\langle r \rangle_\alpha - B) + A \cdot (\langle r \rangle_\alpha - B) + B \cdot (\langle t \rangle_\alpha - A)$.

Case 3: Axioms for suqare root are interpreted using corresponding rules for auxiliary variables in $\mathsf{PC}^{+,\{2\}}$.

Induction step: The cases of structural rules are trivial. Other rules are handled as follows:

Case 1: Left and right $\wedge$-introduction. This case is trivial, the argument is as in Theorem 51.

Case 2: Left and right bounded index $\forall$-introduction. Analogous to the argument in Theorem 51.

Case 3: Induction rule:

$$\frac{\Gamma, \phi(i) \longrightarrow \phi(i+1)}{\Gamma, \phi(0) \longrightarrow \phi(j)}$$

where variable j does not occur in Γ.

Let α be assignments and let $n := \alpha(j)$. By induction hypothesis, there exists $d \in \mathbb{N}$ such that for every $v \in \mathbb{N}$ and every witnessing function $W_{\alpha[i \leftarrow v]}$ of degree $\leq d$ there exists $W'_{\alpha[i \leftarrow v]}$ of degree $\leq d$ and a derivation $\pi_v(W_{\alpha[i \leftarrow v]}) : \langle \Gamma \rangle^L_\alpha(W_{\alpha[i \leftarrow v]}) \cup \langle \phi(i) \rangle_{\alpha[i \leftarrow v]}(W_{\alpha[i \leftarrow v]}) \vdash \langle \phi(i+1) \rangle_{\alpha[i \leftarrow v]}(W'_{\alpha[i \leftarrow v]})$. Concatenation of $\pi_0(W_{\alpha[i \leftarrow 0]}), \pi_1(W'_{\alpha[i \leftarrow 1]}), \ldots$ gives a derivation $\langle \Gamma \rangle^L_\alpha \cup \langle \phi(0) \rangle_\alpha(W_{\alpha[i \leftarrow 0]}) \vdash \langle \phi(n) \rangle_\alpha(W^{(n)}_{\alpha[i \leftarrow n]})$.

Case 4: Cut rule:

$$\frac{\Gamma \longrightarrow \phi \qquad \phi, \Gamma \longrightarrow \psi}{\Gamma \longrightarrow \psi}$$

By induction hypothesis there are derivations $\pi_1(W_\alpha) : \langle \Gamma \rangle^L_\alpha(W_\alpha) \vdash \langle \phi \rangle_\alpha(W'_\alpha)$ and $\pi_2(W_\alpha) : \langle \phi \rangle_\alpha(W_\alpha) \cup \langle \Gamma \rangle^L_\alpha(W_\alpha) \vdash \langle \psi \rangle^R_\alpha(W'_\alpha)$. Construct the desired derivation as follows: $\langle \Gamma \rangle^L_\alpha(W_\alpha) \vdash \langle \phi \rangle_\alpha(W'_\alpha) \cup \langle \Gamma \rangle^L_\alpha(W'_\alpha) \vdash \langle \psi \rangle^R_\alpha(W''_\alpha)$, where $\langle \Gamma \rangle^L_\alpha(W'_\alpha) = \langle \Gamma \rangle^L_\alpha(W_\alpha)$. $\square$

Chapter 5

Conclusion and Open Problems

This thesis contributes to better understanding of counting in propositional proof complexity, but also broadens perspectives of what is yet to be discovered.

We demonstrated limitations of dag- and tree-like Res($\text{lin}_{\mathcal{R}}$) via a number of upper and lower bounds on some basic contradictions. Of course, the picture is far from complete and there are many open questions, but one question seems to be standing out. In this work we proved an exponential lower bound for dag-like Res($\text{lin}_{\mathbb{F}}$) refutation for a single 0-1 unsatisfiable equation with large coefficients, where $\mathbb{F}$ of characteristic 0. It is interesting to find out, whether this result can be extended to 0-1 unsatisfiable systems of equations over fields of different characteristics, where coefficients are polynomially bounded. It would be particularly interesting to prove dag-like lower bound for a system, which is in the image of the reduction from 3-SAT to 0-1 unsatisfiable linear systems over $\mathbb{F}_p, p \geq 5$, because this lower bound would imply CNF lower bound.

Another interesting question is whether separations between tree-like Res($\text{lin}_{\mathbb{F}_p}$) and tree-like Res($\text{lin}_{\mathbb{F}_q}$) for $p \neq q$ can be proven directly, not via $\mathsf{PC}_{\mathbb{F}}$ degree lower bounds. Perhaps the technique, employed by Razborov and Alekhnovich for $\mathsf{PC}_{\mathbb{F}}$, can inspire some direct argument, proving lower bound on Res($\text{lin}_{\mathbb{F}_p}$) width. And, of course, separation between tree-like Res($\text{lin}_{\mathbb{F}_p}$) and tree-like Res($\text{lin}_{\mathbb{F}}$), where $char(\mathbb{F}) = 0$, is still missing.

Finally, it would be interesting to develop further Prover-Delayer games technique and to extend it to, for example, Tseitin formulas.

The work on first-order theories completely resolves the question of what should be a first-order theory for constant degree $\mathsf{PC}_{\mathcal{R}}$ if $\mathcal{R}$ is a field of positive characteristic. For field $\mathbb{F}$ of characteristic 0 we only know that we can translate the theory $\mathsf{TPC}_{\mathbb{F}}$ to an extension $\mathsf{PC}^{\text{rad}}_{\mathbb{F},d}$ of $\mathsf{PC}_{\mathbb{F},d}$ with the radical rule. Although we know, that there is a

lower bound on derivations $f^2 = 0 \vdash f = 0$, we do not know whether $\mathsf{PC}_{\mathbb{F}}^{\mathrm{rad}}$ is strictly stronger than $\mathsf{PC}_{\mathbb{F},d}$ as a refutation system.

In case of constant degree SoS it is much less clear how the right theory for it should look like. The gap between the theory $\mathsf{TSoS}_{\geq}$, which we can translate to SoS_d, and theories, which we know are too string, is quite substantial.

Currently we show that a theory T is too strong by showing that it proves soundness of resolution. Alternatively, one can show that T proves some statement, for propositional translation of which there is $\omega(1)$ SoS degree lower bound. For example, linear degree SoS lower bounds are known for the Subset Sum principle and Tseitin tautologies, therefore T should not be able to prove them.

Attempts to prove that T is too strong raise new SoS degree lower bound questions in this way. For example, the modification $\mathsf{TSoS}^1_{\geq}$ of $\mathsf{TSoS}_{\geq}$, where marked inequalities $\geq_d$ are replaced with unmarked one $\geq$, is too strong, because it simulates resolution. If we restrict induction of $\mathsf{TSoS}^1_{\geq}$ to formulas without universally quantified inequalities $\forall \overline{x}\ t(\overline{x}) \geq 0$ (but possibly with inequalities $t \geq 0$), it does not prove soundness of resolution, but it still proves a statement, which is supposedly beyond SoS_d. Namely, it proves the following statement a la telescopic system: $X(1) \leq \epsilon \wedge \forall (0 \leq i < n) X(i+1) = X(i)^2 \supset X(n) \leq \epsilon^{2^n}$.

From the other side, one may try to show that a theory T is not stronger than SoS_d, even though proofs in T cannot be directly translated to SoS_d, by proving that T conservatively extends TSoS with inequality. A possible way to approach this is via extending any model of TSoS with an ordering, satisfying axioms in T.

Index

Bibliography

[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. 1.1.1.1, 1.1.1.3

[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. 1.2.1, 2.2.1.2, 3, 4, 3.3.3, 3.4, 3.4, 3.4, 3.4

[3] Noga Alon. Decomposition of the complete$r$-graph into complete$r$-partite$r$-graphs. *Graphs and Combinatorics*, 2(1):95–100, 1986. 1.1.2

[4] Noga Alon and Zoltán Füredi. Covering the cube by affine hyperplanes. *Eur. J. Comb.*, 14(2):79–83, March 1993. 1.2.1, 3.3.4, 3.3.4

[5] Christoph Ambuhl, Monaldo Mastrolilli, and Ola Svensson. Inapproximability results for sparsest cut, optimal linear arrangement, and precedence constrained scheduling. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '07, pages 329–337, Washington, DC, USA, 2007. IEEE Computer Society. 1.1.2

[6] David L. Applegate, Robert E. Bixby, Vasek Chvatal, and William J. Cook. *The Traveling Salesman Problem: A Computational Study (Princeton Series in Applied Mathematics)*. Princeton University Press, Princeton, NJ, USA, 2007. 1.1

[7] Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, New York, NY, USA, 1st edition, 2009. 1.1

[8] Sanjeev Arora, Satish Rao, and Umesh Vazirani. Expander flows, geometric embeddings and graph partitioning. *J. ACM*, 56(2):5:1–5:37, April 2009. 1.1.2

[9] Michael Francis Atiyah and I. G. MacDonald. *Introduction to commutative algebra.* Addison-Wesley-Longman, 1969. 1.1.1.2

[10] Boaz Barak, Fernando G. S. L. Brandão, Aram Wettroth Harrow, Jonathan A. Kelner, David Steurer, and Yuan Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *STOC*, pages 307–326. ACM, 2012. 1.1.2

[11] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. Making the long code shorter. *SIAM J. Comput.*, 44(5):1287–1324, 2015. 1.1.2

[12] Boaz Barak and David Steurer. Sum-of-squares proofs and the quest toward optimal algorithms. *CoRR*, abs/1404.5236, 2014. 1.1.1.2, 1.1.2

[13] Paul Beame, Russell Impagliazzo, Jan Krajíček, Toniann Pitassi, and Pavel Pudlák. Lower bounds on Hilbert's Nullstellensatz and propositional proofs. *Proc. London Math. Soc. (3)*, 73(1):1–26, 1996. 1.1.1.2

[14] Arnold Beckmann, Pavel Pudlák, and Neil Thapen. Parity games and propositional proofs. *ACM Transactions on Computational Logic.* 1.2, 2.4

[15] Eli Ben-Sasson. Hard examples for the bounded depth Frege proof system. *Comput. Complexity*, 11(3-4):109–136, 2002. 1.1.1.1

[16] Christoph Berkholz. The Relation between Polynomial Calculus, Sherali-Adams, and Sum-of-Squares Proofs. In Rolf Niedermeier and Brigitte Vallée, editors, *35th Symposium on Theoretical Aspects of Computer Science (STACS 2018)*, volume 96 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 11:1–11:14, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. 1.1.1.2, 1.2.3.2, 2.2.4, 4.2, 4.2.1.1

[17] Olaf Beyersdorff, Nicola Galesi, and Massimo Lauria. A lower bound for the pigeonhole principle in tree-like resolution by asymmetric prover-delayer games. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:81, 2010. 3.3.2

[18] Olaf Beyersdorff, Nicola Galesi, and Massimo Lauria. A characterization of tree-like resolution size. *Inf. Process. Lett.*, 113(18):666–671, 2013. 3.3.2, 3.3.2

[19] Olaf Beyersdorff, Nicola Galesi, and Massimo Lauria. Parameterized complexity of DPLL search procedures. *ACM Trans. Comput. Log.*, 14(3):20:1–20:21, 2013. 3.3.2

[20] J. Bochnak, M. Coste, and M-F. Roy. *Real Algebraic Geometry*. Springer, 1998. 1.1.1.2

[21] Samuel R. Buss. *Bounded Arithmetic*, volume 3 of *Studies in Proof Theory*. Bibliopolis, 1986. 1.1.1.3

[22] Samuel R. Buss. Relating the bounded arithmetic and polynomial time hierarchies. *Ann. Pure Appl. Logic*, 75(1-2):67–77, 1995. Proof theory, provability logic, and computation (Berne, 1994). 1.1.1.3

[23] 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). 1.1.1.2, 2.2.1.2

[24] Jianer Chen, Xiuzhen Huang, Iyad A. Kanj, and Ge Xia. Strong computational lower bounds via parameterized complexity. *J. Comput. Syst. Sci.*, 72(8):1346–1367, December 2006. 1.1.2

[25] Matthew Clegg, Jeffery Edmonds, and Russell Impagliazzo. Using the Groebner basis algorithm to find proofs of unsatisfiability. In *Proceedings of the 28th Annual ACM Symposium on the Theory of Computing (Philadelphia, PA, 1996)*, pages 174–183, New York, 1996. ACM. 1.1.1.2, 1.2.1.1, 2.2

[26] Stephen A. Cook. Feasibly constructive proofs and the propositional calculus (preliminary version). In *STOC*, pages 83–97, 1975. 1.1.1.3

[27] Stephen A. Cook and Robert A. Reckhow. On the lengths of proofs in the propositional calculus (preliminary version). In *1974*, pages 135–148, 1974. For corrections see Cook-Reckhow [**?**]. 28

[28] Stephen A. Cook and Robert A. Reckhow. The relative efficiency of propositional proof systems. *J. Symb. Log.*, 44(1):36–50, 1979. This is a journal-version of Cook-Reckhow [27] and Reckhow [62]. 1.1.1, 1.2.1, 1

[29] W. Cook, C. R. Coullard, and G. Turan. On the complexity of cutting plane proofs. *Discrete Applied Mathematics*, 18:25–38, 1987. 1.1.1.2

[30] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms, Third Edition*. The MIT Press, 3rd edition, 2009. 1.1, 1.1.2

[31] Michael A. Forbes, Amir Shpilka, Iddo Tzameret, and Avi Wigderson. Proof complexity lower bounds from algebraic circuit complexity. In *Conference on Computational Complexity*, volume 50 of *LIPIcs*, pages 32:1–32:17. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2016. 1.2.3.1, 4.1.3.1

[32] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., New York, NY, USA, 1979. 1.1

[33] M. Garlik and L. A. Kołodziejczyk. Some subsystems of constant-depth Frege with parity. 2017. Unpublished manuscript. 1.1.1.1, 1.2.1.1, 1.2.1.1, 1.2

[34] Dima Grigo riev and Nicolai Vorobjov. Complexity of Null- and Positivstellensatz proofs. *Ann. Pure Appl. Logic*, 113(1-3):153–160, 2002. First St. Petersburg Conference on Days of Logic and Computability (1999). 1.1.1.2

[35] Dima Grigoriev, Edward Hirsch, and Dmitrii Pasechnik. Complexity of semi-algebraic proofs. 09 2002. 1.1.1.2

[36] Armin Haken. The intractability of resolution. *Theoret. Comput. Sci.*, 39(2-3):297–308, 1985. 1.1.1.1

[37] R. Herken, editor. *A Half-century Survey on The Universal Turing Machine*, New York, NY, USA, 1988. Oxford University Press, Inc. 1.1

[38] Russell Impagliazzo, Pavel Pudlák, and Jiří Sgall. Lower bounds for the polynomial calculus and the gröbner basis algorithm. *ComputationalComplexity*, 8(2):127–144, 1999. 1.1.1.2

[39] Dmitry Itsykson and Alexander Knop. Hard satisfiable formulas for splittings by linear combinations. In Serge Gaspers and Toby Walsh, editors, *Theory and Applications of Satisfiability Testing – SAT 2017*, pages 53–61. Springer International Publishing, 2017. 3.3.2

[40] Dmitry Itsykson and Dmitry Sokolov. Lower bounds for splittings by linear combinations. In *Mathematical Foundations of Computer Science 2014 - 39th International Symposium, MFCS 2014, Budapest, Hungary, August 25-29, 2014.*

Proceedings, Part II, pages 372–383, 2014. (document), 1.1.1.1, 1.2.1, 1.2.1, 1.2.1.1, 1.2, 1.2.2.1, 3.1, 3.1, 3.3.1, 3.3.1, 3.3.2, 3.3.4, 3.3.4

[41] S. Khot. Improved inapproximability results for MaxClique, chromatic number and approximate graph coloring. In *Proceedings of the 42Nd IEEE Symposium on Foundations of Computer Science*, FOCS '01, pages 600–, Washington, DC, USA, 2001. IEEE Computer Society. 1.1.2

[42] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the Thiry-fourth Annual ACM Symposium on Theory of Computing*, STOC '02, pages 767–775, New York, NY, USA, 2002. ACM. 1.1.2

[43] Jan Krajícek. Randomized feasible interpolation and monotone circuits with a local oracle. *CoRR*, abs/1611.08680, 2016. 1.1.1.1

[44] Jan Krajícek and Igor Carboni Oliveira. On monotone circuits with local oracles and clique lower bounds. *Chicago J. Theor. Comput. Sci.*, 2018, 2018. 1.1.1.1

[45] Jan Krajícek and Pavel Pudlák. Quantified propositional calculi and fragments of bounded arithmetic. *Math. Log. Q.*, 36(1):29–46, 1990. 1.1.1.3

[46] Jan Krajíček, Pavel Pudlák, and Gaisi Takeuti. Bounded arithmetic and the polynomial hierarchy. *Ann. Pure Appl. Logic*, 52(1-2):143–153, 1991. International Symposium on Mathematical Logic and its Applications (Nagoya, 1988). 1.1.1.3

[47] Jan Krajíček, Pavel Pudlák, and Alan Woods. An exponential lower bound to the size of bounded depth Frege proofs of the pigeonhole principle. *Random Structures Algorithms*, 7(1):15–39, 1995. 1.1.1.1

[48] Nathan Linial and Jaikumar Radhakrishnan. Essential covers of the cube by hyperplanes. *Journal of Combinatorial Theory, Series A*, 109. (document), 1.2.1, 32

[49] László Lovász and Alexander Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, 1991. 1.1.1.2

[50] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, Sep 1988. 2

[51] Jakob Nordström. On the interplay between proof complexity and sat solving. *ACM SIGLOG News*, 2(3):19–44, August 2015. 1.1, 1.1.1.1

[52] R. O'Donnell and Y. Zhou. Approximability and proof complexity. In *Proceedings of SODA*, 2013. 1.1.2, 1.2.3

[53] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, New York, NY, USA, 2014. 1.1.2

[54] J. Paris and A. Wilkie. Counting problems in bounded arithmetic. In *Methods in mathematical logic (Caracas, 1983)*, volume 1130 of *Lecture Notes in Math.*, pages 317–340. Springer, Berlin, 1985. 1.1.1.3, 2.4

[55] Toniann Pitassi, Paul Beame, and Russell Impagliazzo. Exponential lower bounds for the pigeonhole principle. *Comput. Complexity*, 3(2):97–140, 1993. 1.1.1.1

[56] 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. 1.2.1, 3.3.2

[57] Prasad Raghavendra. Optimal algorithms and inapproximability results for every csp? In *STOC*, pages 245–254. ACM, 2008. 1.1.2

[58] Prasad Raghavendra and David Steurer. Integrality gaps for strong SDP relaxations of UNIQUE GAMES. In *50th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2009, October 25-27, 2009, Atlanta, Georgia, USA*, pages 575–585, 2009. 1.1.2

[59] Prasad Raghavendra, David Steurer, and Madhur Tulsiani. Reductions between expansion problems. In *IEEE Conference on Computational Complexity*, pages 64–73. IEEE Computer Society, 2012. 1.1.2

[60] Ran Raz and Iddo Tzameret. Resolution over linear equations and multilinear proofs. *Ann. Pure Appl. Logic*, 155(3):194–224, 2008. 1.1.1.1, 1.2.1, 1.2.1, 2.2.1.2, 3.1, 19, 3.1.2

[61] Alexander A. Razborov. Lower bounds for the polynomial calculus. *Comput. Complex.*, 7(4):291–324, December 1998. 1.1.1.2

[62] Robert A. Reckhow. *On the lengths of proofs in the propositional calculus*. PhD thesis, University of Toronto, 1976. 28

[63] Michael Soltys and Stephen Cook. The proof complexity of linear algebra. *Ann. Pure Appl. Logic*, 130(1-3):277–323, 2004. 1.2

[64] Grigori Tseitin. *On the complexity of derivations in propositional calculus*. Studies in constructive mathematics and mathematical logic Part II. Consultants Bureau, New-York-London, 1968. 1.1.1.1, 2.2.1.2

[65] Domenico Zambella. Notes on polynomially bounded arithmetic. *The Journal of Symbolic Logic*, 61(3):942–966, 1996. 1.1.1.3