Proof Complexity Lower Bounds from Algebraic Circuit Complexity

Michael A. Forbes * Amir Shpilka † Iddo Tzameret ‡ Avi Wigderson §

Abstract

We give upper and lower bounds on the power of subsystems of the Ideal Proof System (IPS), the algebraic proof system recently proposed by Grochow and Pitassi [GP14], where the circuits comprising the proof come from various restricted algebraic circuit classes. This mimics an established research direction in the boolean setting for subsystems of Extended Frege proofs whose lines are circuits from restricted boolean circuit classes. Essentially all of the subsystems considered in this paper can simulate the well-studied Nullstellensatz proof system, and prior to this work there were no known lower bounds when measuring proof size by the algebraic complexity of the polynomials (except with respect to degree, or to sparsity).

Our main contributions are two general methods of converting certain algebraic lower bounds into proof complexity ones. Both require stronger arithmetic lower bounds than common, which should hold not for a specific polynomial but for a whole family defined by it. These may be likened to some of the methods by which Boolean circuit lower bounds are turned into related proof-complexity ones, especially the "feasible interpolation" technique. We establish algebraic lower bounds of these forms for several explicit polynomials, against a variety of classes, and infer the relevant proof complexity bounds. These yield separations between IPS subsystems, which we complement by simulations to create a partial structure theory for IPS systems.

Our first method is a *functional lower bound*, a notion of Grigoriev and Razborov [GR00], which is a function $\hat{f} : \{0,1\}^n \to \mathbb{F}$ such that any polynomial f agreeing with $\hat{f}$ on the boolean cube requires large algebraic circuit complexity. We develop functional lower bounds for a variety of circuit classes (sparse polynomials, depth-3 powering formulas, roABPs and multilinear formulas) where $\hat{f}(\overline{x})$ equals $1/p(\overline{x})$ for a constant-degree polynomial p depending on the relevant circuit class. We believe these lower bounds are of independent interest in algebraic complexity, and show that they also imply lower bounds for the size of the corresponding IPS refutations for proving that the relevant polynomial p is non-zero over the boolean cube. In particular, we show super-polynomial lower bounds for refuting variants of the subset-sum axioms in these IPS subsystems.

Our second method is to give *lower bounds for multiples*, that is, to give explicit polynomials whose all (non-zero) multiples require large algebraic circuit complexity. By extending known techniques, we give lower bounds for multiples for various restricted circuit classes such sparse polynomials, sums of powers of low-degree polynomials, and roABPs. These results are of independent interest, as we argue that lower bounds for multiples is the correct notion for instantiating the algebraic hardness versus randomness paradigm of Kabanets and Impagliazzo [KI04]. Further, we show how such lower bounds for multiples extend to lower bounds for refutations in the corresponding IPS subsystem.

*Email: miforbes@csail.mit.edu. Department of Computer Science, Princeton University. Supported by the Princeton Center for Theoretical Computer Science.

†Email: shpilka@post.tau.ac.il. Department of Computer Science, Tel Aviv University, Tel Aviv, Israel. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 257575.

‡Email: iddo.tzameret@rhul.ac.uk. Department of Computer Science, Royal Holloway, University of London, UK.

§Email: avi@math.ias.edu. Institute for Advanced Study, Princeton. This research was partially supported by NSF grant CCF-1412958.

Contents

1 Introduction

Propositional proof complexity aims to understand and analyze the computational resources required to prove propositional tautologies, in the same way that circuit complexity studies the resources required to compute boolean functions. A typical goal would be to establish, for a give proof system, super-polynomial lower bounds on the *size* of any proof of some propositional tautology. The seminal work of Cook and Reckhow [CR79] showed that this goal relates quite directly to fundamental hardness questions in computational complexity such as the NP vs. coNP question: establishing super-polynomial lower bounds for *every* propositional proof system would separate NP from coNP (and thus also P from NP). We refer the reader to Krajíček [Kra95] for more on this subject.

Propositional proof systems come in a large variety, as different ones capture different forms of reasoning, either reasoning used to actually prove theorems, or reasoning used by algorithmic techniques for different types of search problems (as failure of the algorithm to find the desired object constitutes a proof of its nonexistence). Much of the research in proof complexity deals with propositional proof systems originating from logic and from geometry. Logical proof systems include such systems as *resolution* (whose variants are related to popular algorithms for automated theory proving and SAT solving), as well as the *Frege* proof system (capturing the most common logic text-book systems) and its many subsystems. Geometric proof systems include *cutting-plane proofs*, capturing reasoning used in algorithms for integer programming, as well as proof systems arising from systematic strategies for rounding linear- or semidefinite-programming such as the lift-and-project or sum-of-squares hierarchies.

In this paper we focus on algebraic proof systems, in which propositional tautologies (or rather contradictions) are expressed as unsatisfiable systems of polynomial equations and algebraic tools are used to refute them. This study originates with the work of Beame, Impagliazzo, Krajíček, Pitassi and Pudlák [BIK+96a], who introduced the Nullstellensatz refutation system (based on Hilbert's Nullstellensatz), followed by the Polynomial Calculus system of Clegg-Edmonds-Impagliazzo [CEI96], which is a "dynamic version" of Nullstellensatz. In both systems the main measures of proof size that have been studied are the *degree* and *sparsity* of the polynomials appearing in the proof. Substantial work has lead to a very good understanding of the power of these systems with respect to these measures (see for example [BIK+96b,Raz98,Gri98,IPS99,BGIP01,AR01] and references therein).

However, the above measures of degree and sparsity are rather rough measures of a complexity of a proof. As such, Grochow and Pitassi [GP14] have recently advocated measuring the complexity of such proofs by their algebraic circuit size and shown that the resulting proof system can polynomially simulate strong proof systems such as the Frege system. This naturally leads to the question of establishing lower bounds for this stronger proof system, even for restricted classes of algebraic circuits.

In this work we establish such lower bounds for previously studied restricted classes of algebraic circuits, and show these lower bounds are interesting by providing non-trivial *upper* bounds in these proof systems for refutations of interesting sets of polynomial equations. This provides what are apparently the first examples of lower bounds on the algebraic circuit size of propositional proofs in the ideal proof system (IPS) framework of Grochow and Pitassi [GP14].

We note that obtaining proof complexity lower bounds from circuit complexity lower bounds is an established tradition, and takes many forms. Most prominent are the lower bounds for susbsystems of the Frege proof system defined by low-depth Boolean circuits, and lower bounds on Resolution and Cutting Planes system using the so-called feasible interpolation method [Pud97]. We refer the reader again to the monograph [Kra95] for more details. Our approach here for

algebraic systems shares features with both of these approaches.

The rest of this long introduction is arranged as follows. In Subsection 1.1 we give the necessary background in algebraic proof complexity, and explain IPS system. In subsection 1.2 we define the arithmetic complexity classes that will underlie the subsystems of IPS we will study. In subsection 1.3 we state our results and explain our techniques, for both the arithmetic and proof complexity worlds.

1.1 Algebraic Proof Systems

We now describe the algebraic proof systems that are the subject of this paper. If one has a set of polynomials (called *axioms*) $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ over some field $\mathbb{F}$, then (the weak version of) Hilbert's Nullstellensatz shows that the system $f_1(\overline{x}) = \cdots = f_m(\overline{x}) = 0$ is unsatisfiable (over the algebraic closure of $\mathbb{F}$) if and only if there are polynomials $g_1, \ldots, g_m \in \mathbb{F}[\overline{x}]$ such that $\sum_j g_j(\overline{x}) f_j(\overline{x}) = 1$ (as a formal identity), or equivalently, that 1 is in the ideal generated by the $\{f_j\}_j$.

Beame, Impagliazzo, Krajíček, Pitassi, and Pudlák [BIK+96a] suggested to treat these $\{g_j\}_j$ as a *proof* of the unsatisfiability of this system of equations, called a *Nullstellensatz refutation*. This is particular relevant for complexity theory as one can restrict attention to *boolean* solutions to this system by adding the *boolean axioms*, that is, adding the polynomials $\{x_i^2 - x_i\}_{i=1}^n$ to the system. As such, one can then naturally encode NP-complete problems such as the satisfiability of 3CNF formulas as the satisfiability of a system of constant-degree polynomials, and a Nullstellensatz refutation is then an equation of the form $\sum_{j=1}^m g_j(\overline{x}) f_j(\overline{x}) + \sum_{i=1}^n h_i(\overline{x})(x_i^2 - x_i) = 1$ for $g_j, h_i \in \mathbb{F}[\overline{x}]$. This proof system is sound (only refuting unsatisfiable systems over $\{0,1\}^n$) and complete (refuting any unsatisfiable system, by Hilbert's Nullstellensatz).

Given that the above proof system is sound and complete, it is then natural to ask what is its power to refute unsatisfiable systems of polynomial equations over $\{0,1\}^n$. To understand this question one must define the notion of the *size* of the above refutations. Two popular notions are that of the *degree*, and the *sparsity* (number of monomials). One can then show (see for example Pitassi [Pit97]) that for any unsatisfiable system which includes the boolean axioms, there exist a refutation where the g_j are multilinear and where the h_i have degree at most $O(n+d)$, where each f_j has degree at most d. In particular, this implies, when $d = O(n)$, that for any unsatisfiable system there is a refutation of degree $O(n)$ and involving at most $\exp(O(n))$ monomials. This intuitively agrees with the fact that coNP is a subset of non-deterministic exponential time.

Building on the suggestion of Pitassi [Pit97], Grochow and Pitassi [GP14] have recently considered more *succinct* descriptions of polynomials where one measures the size of a polynomial by the size of an algebraic circuit needed to compute it. This is potentially much more powerful as there are polynomials such as the determinant which are of high degree and involve exponentially many monomials and yet can be computed by small algebraic circuits. They named the resulting system the *Ideal Proof System (IPS)* which we now define.

Definition 1.1 (Ideal Proof System (IPS), Grochow-Pitassi [GP14])**.** *Let $f_1(\overline{x}), \ldots, f_m(\overline{x}) \in \mathbb{F}[x_1, \ldots, x_n]$ be a system of polynomials. An **IPS refutation** that the polynomials $\{f_j\}_j$ have no common solution in $\{0,1\}^n$ is an algebraic circuit $C(\overline{x}, \overline{y}, \overline{z}) \in \mathbb{F}[\overline{x}, y_1, \ldots, y_m, z_1, \ldots, z_n]$, such that*

1. *$C(\overline{x}, \overline{0}, \overline{0}) = 0$.*
2. *$C(\overline{x}, f_1(\overline{x}), \ldots, f_m(\overline{x}), x_1^2 - x_1, \ldots, x_n^2 - x_n) = 1$.*

*The **size** of the IPS refutation is the size of the circuit C. If C is of individual degree ≤ 1 in each y_j and z_i, then this is a **linear** IPS refutation (called* Hilbert *IPS by Grochow-Pitassi [GP14]), which we will abbreviate as IPS_{LIN} . If C is of individual degree ≤ 1 only in the y_j then we say this is a $IPS_{LIN'}$ refutation. If C comes from a restricted class of algebraic circuits $\mathcal{C}$, then this is a called a $\mathcal{C}$-IPS refutation, and further called a $\mathcal{C}$-IPS_{LIN} refutation if C is linear in $\overline{y}, \overline{z}$, and a $\mathcal{C}$-$IPS_{LIN'}$ refutation if C is linear in $\overline{y}$.*

Notice also that our definition here necessarily adds the equations $\{x_i^2 - x_i\}_i$ to the system $\{f_j\}_j$. For convenience we will often denote the equations $\{x_i^2 - x_i\}_i$ as $\overline{x}^2 - \overline{x}$. One need not add the equations $\overline{x}^2 - \overline{x}$ to the system in general, but this is the most interesting regime for proof complexity and thus we adopt it as part of our definition.

The above discussion shows that the above proof systems are all sound, and that for any complete class of algebraic circuits $\mathcal{C}$ (classes that can compute any polynomial, possibly requiring exponential complexity) that $\mathcal{C}$-IPS_{LIN} is a complete proof system. We note that we will also consider non-complete classes such as multilinear-formulas (which can only compute *multilinear* polynomials, but are complete for multilinear polynomials), where we will show that the multilinear-formula-IPS_{LIN} system is not complete (Subsection 4.3) but that the $\text{IPS}_{\text{LIN}'}$ version is complete (Theorem 4.6).

Grochow-Pitassi [GP14] proved the following theorem, showing that the IPS system has surprising power and that lower bounds on this system give rise to *computational* lower bounds.

Theorem 1.1 (Grochow-Pitassi [GP14])**.** *Let φ be a 3CNF. If there is a Frege proof that φ is unsatisfiable in size-s, then there is an IPS refutation of size $\mathsf{poly}(|\varphi|, s)$ computing a polynomial of degree $\mathsf{poly}(|\varphi|, s)$, and this refutation is checkable in randomized $\mathsf{poly}(|\varphi|, s)$ time. Conversely, if every IPS refutation requires size $\geq s$ then there is an explicit polynomial (that is, in VNP) that requires $\geq s$-size algebraic circuits.*

The fact that $\mathcal{C}$-IPS refutations are efficiently checkable (with randomness) follows from the fact that we need to verify the polynomial identities stipulated by the definition. That is, one needs to solve an instance of the *polynomial identity testing (PIT)* problem for the class $\mathcal{C}$: given a circuit from the class $\mathcal{C}$ decide whether it computes the identically zero polynomial. This problem is solvable in probabilistic polynomial time (BPP) for general algebraic circuits, and there are various restricted classes for which deterministic algorithms are known (see Subsection 3.1).

Motivated by the fact that PIT of non-commutative formulas can be solved deterministically ([RS05]), Li, Tzameret and Wang [LTW15] have shown that IPS over *non-commutative* polynomials can deterministically simulate Frege (see Li, Tzameret and Wang [LTW15] for a definition).

Theorem 1.2 (Li, Tzameret and Wang [LTW15])**.** *Let φ be a 3CNF. If Frege can prove that φ is unsatisfiable in size-s, then there is a non-commutative IPS refutation of formula size $\mathsf{poly}(|\varphi|, s)$ computing a polynomial of degree $\mathsf{poly}(|\varphi|, s)$, where the commutator axioms $x_i x_j - x_j x_i$ are also included in the polynomial system being refuted. Further, this refutation is checkable in deterministic $\mathsf{poly}(|\varphi|, s)$ time.*

The above results naturally motivate studying $\mathcal{C}$-IPS for various restricted classes of algebraic circuits, as lower bounds for such proofs then intuitively correspond to restricted lower bounds for the Extended Frege proof system. In particular, as exponential lower bounds are known for non-commutative formulas ([Nis91]), this possibly suggests that such methods could even attack the full Frege system itself.

1.2 Algebraic Circuit Classes

Having motivated $\mathcal{C}$-IPS for restricted circuit classes $\mathcal{C}$, we now give formal definitions of the algebraic circuit classes of interest to this paper, all of which were studied independently in algebraic complexity. Some of them define the state-of-art in our ability to prove lower bounds and provide efficient deterministic identity tests, so it is natural to attempt converting these to the proof complexity framework. We define each and briefly explain what we know about it. As the list is long, the reader may consider skipping to the results (Subsection 1.3), and refer to the definitions of these classes as they arise.

Arithmetic circuits and formula (over a fixed chosen field) compute polynomials via addition and multiplication gates, starting from the input variables and constants from the field. For background on arithmetic circuits in general and their complexity measures we refer the reader to the survey [SY10]. We next define the restricted circuit classes that we will be studying in this paper.

1.2.1 Low Depth Classes

We start by defining what are the simplest and most restricted classes of algebraic circuits. The first class simply represents polynomials as a sum of monomials. This is also called the *sparse representation* of the polynomial. Notationally we call this model $\Sigma\Pi$ formulas (to capture the fact that polynomials computed in the class are represented simply as sums of products), but we will more often call these polynomials "sparse".

Definition 1.2. *The class $\mathcal{C} = \Sigma\Pi$ compute polynomials in their sparse representation, i.e., as sum of monomials. The graph of computation has two layers with an addition gate at the top and multiplication gates at the bottom. The size of a $\Sigma\Pi$ circuit of a polynomial f is the number of monomials in f.*

This class of circuits is what is used in the Nullstellensatz proof system. In our terminology $\Sigma\Pi$-IPS$_{\text{LIN}}$ is exactly the Nullstellensatz proof system.

Another restricted class of algebraic circuits is that of *depth-3 powering formulas* (sometimes also called "diagonal depth-3 circuits"). We will sometimes abbreviate this name as a "$\sum\bigwedge\sum$ formula", where $\bigwedge$ denotes the powering operation. Specifically, polynomials that are efficiently computed by small formulas from this class can be represented as sum of powers of linear functions. This model appears implicitly in Shpilka [Shp02] and explicitly in the work of Saxena [Sax08].

Definition 1.3. *The class of depth-3 powering formulas, denoted $\sum\bigwedge\sum$, computes polynomials of the following form*

$$f(\overline{x}) = \sum_{i=1}^{s} \ell_i(\overline{x})^{d_i},$$

where $\ell_i(\overline{x})$ are linear functions. The degree of this $\sum\bigwedge\sum$ representation of f is $\max_i\{d_i\}$ and its size is $n \cdot \sum_{i=1}^{s}(d_i+1)$.

One reason for considering this class of circuits is that it is a simple, but non-trivial model that is somewhat well-understood. In particular, the partial derivative method of Nisan-Wigderson [NW96] implies lower bounds for this model and efficient PIT algorithms are known ([Sax08, ASS13, FS13a, FS13b, FSS14], as discussed further in Subsection 3.1).

We also consider a generalization of this model where we allow powering of low-degree polynomials.

Definition 1.4. *The class $\sum\bigwedge\sum\prod^t$ computes polynomials of the following form*

$$f(\overline{x}) = \sum_{i=1}^{s} f_i(\overline{x})^{d_i} ,$$

where the degree of the $f_i(\overline{x})$ is at most t. The size of this representation is $\binom{n+t}{t} \cdot \sum_{i=1}^{s}(d_i+1)$.

We remark that the reason for defining the size this way is that we think of the f_i as represented as sum of monomials (there are $\binom{n+t}{t}$ n-variate monomials of degree at most t) and the size captures the complexity of writing this as an algebraic formula. This model is the simplest that requires the method of *shifted partial derivatives* of Kayal [Kay12, GKKS14] to establish lower bounds, and this has recently been generalized to obtain PIT algorithms ([For14], as discussed further in Subsection 3.1).

1.2.2 Oblivious Algebraic Branching Programs

Algebraic branching programs (ABPs) form a model whose computational power lies between that of algebraic circuits and algebraic formulas, and certain *read-once* and *oblivious* ABPs are a natural setting for studying the *partial derivative matrix* lower bound technique of Nisan [Nis91].

Definition 1.5 (Nisan [Nis91])**.** *An **algebraic branching program (ABP) with unrestricted weights** of depth D and width $\leq r$, on the variables $x_1, \ldots, x_n$, is a directed acyclic graph such that:*

- *The vertices are partitioned in $D+1$ layers $V_0, \ldots, V_D$, so that $V_0 = \{s\}$ (s is the source node), and $V_D = \{t\}$ (t is the sink node). Further, each edge goes from V_{i-1} to V_i for some $0 < i \leq D$.*
- *$\max |V_i| \leq r$.*
- *Each edge e is weighted with a polynomial $f_e \in \mathbb{F}[\overline{x}]$.*

Each s-t path is said to compute the polynomial which is the product of the labels of its edges, and the algebraic branching program itself computes the sum over all s-t paths of such polynomials.

- *An algebraic branching program is said to be **oblivious** if for every layer ℓ, all the edge labels in that layer are univariate polynomials in a variable x_{i_ℓ}.*
- *An oblivious branching program is said to be a **read-once** oblivious ABP (roABP) if the x_{i_ℓ}'s are distinct variables, so that $D = n$. That is, each x_i appears in the edge labels in at exactly one layer. The layers thus define a **variable order**, which will be $x_1 < \cdots < x_n$ if not otherwise specified.*
- *An oblivious branching program is said to be a **read-k** oblivious ABP if each variable x_i appears in the edge labels of at most k layers, so that $D = kn$.*

Intuitively, roABPs are the algebraic analog of read-once boolean branching program, the non-uniform model of the class RL. Nisan [Nis91] proved lower bounds for non-commutative ABPs (and thus also for roABPs) and in a sequence of papers polynomial identity testing algorithms were devised for it ([RS05, FS13b, FSS14, AGKS14], see also Subsection 3.1). Recently Anderson, Forbes, Saptharishi, Shpilka, and Volk [AFS+15] obtained exponential lower bounds for read-k oblivious ABPs (when $k = o(\log n / \log\log n)$) as well as a slightly subexponential PIT algorithm.

We note that roABPs are known to simulate non-commutative formulas ([Nis91]). Thus, the result of Li, Tzameret and Wang [LTW15] (see Theorem 1.2) demonstrates the importance of studying IPS proofs over roABPs (see also Tzameret [Tza11]).

1.2.3 Multilinear Formulas

The last model that we consider is that of multilinear formulas.

Definition 1.6 (Multilinear formula)**.** *An algebraic formula is* a multilinear formula *(or equivalently,* multilinear algebraic formula*) if the polynomial computed by* each *gate of the formula is multilinear (as a formal polynomial, that is, as an element of* $\mathbb{F}[x_1, \ldots, x_n]$*).*

Raz [Raz09, Raz06] proved quasi-polynomial lower bounds for multilinear formulas and separated multilinear formulas from multilinear circuits. Raz and Yehudayoff proved exponential lower bounds for small depth multilinear formulas [RY09]. Only slightly sub-exponential polynomial identity testing algorithms are known for small-depth multilinear formulas ([OSV15]).

1.3 Our Results and Techniques

We now briefly summarize our results and techniques, stating some results in less than full generality to more clearly convey the result. We present the results in the order that we later prove them. We start by giving upper bounds for the IPS (subsubsection 1.3.1). We then describe our functional lower bounds and the IPS_{LIN} lower bounds they imply (subsubsection 1.3.2). Finally, we discuss lower bounds for multiples and state our lower bounds for IPS (subsubsection 1.3.3).

1.3.1 Upper Bounds for Proofs within Subclasses of IPS

Grochow and Pitassi [GP14] showed that the full IPS proof system can simulate powerful proof systems such as Frege. This left open the extent to which $\mathcal{C}$-IPS can refute interesting sets of polynomial equations for restricted classes $\mathcal{C}$. We establish here that even restricted classes of IPS are powerful, such as being able to refute interesting unsatisfiable systems of equations arising from particular instances of NP-complete problems.

Our first upper bound is to show that linear-IPS can simulate the full IPS proof system when the axioms are computationally simple.

Theorem (Theorem 4.1)**.** *For* $|\mathbb{F}| \geq \mathsf{poly}(d)$*, if* $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ *are degree-*d *polynomials computable by size-*s *algebraic circuits and they have a size-*t *IPS refutation, then they also have a size-*$\mathsf{poly}(d, s, t)$ *IPS_{LIN} refutation.*

This theorem is established by pushing the "non-linear" dependencies on the axioms into the IPS refutation itself, which is possible as the axioms are assumed to themselves be computable by small circuits. We note that Grochow and Pitassi [GP14] showed such a conversion, but only for IPS refutations computable by sparse polynomials.

We then turn our attention to IPS involving only restricted classes of algebraic circuits, and show that they are complete proof systems. This is clear for complete models of algebraic circuits such as sparse polynomials, depth-3 powering formulas [1] and roABPs. For multilinear formulas this is more subtle as not every polynomial is multilinear, however we can show a simulation of sparse-IPS_{LIN} by careful multilinearization.

Theorem (Subsection 4.3, Theorem 4.6)**.** *The proof systems of sparse-IPS_{LIN}, $\sum\bigwedge\sum$-IPS_{LIN} (in large characteristic fields), and roABP-IPS_{LIN} are complete proof systems. The multilinear-formula-IPS_{LIN} proof system is not complete, but the depth-2 multilinear-formula-$\text{IPS}_{LIN'}$ proof system is complete (for multilinear axioms) and can polynomially simulate sparse-IPS_{LIN} (for low-degree axioms).*

[1] Showing that depth-3 powering formulas are complete (in large characteristic) can be seen from the fact that any multilinear monomial can be computed in this model, see for example Fischer [Fis94].

We next consider the equation $\sum_{i=1}^n \alpha_i x_i - \beta$ along with the boolean axioms $\{x_i^2 - x_i\}_i$. Deciding whether this system of equations is satisfiable is the NP-complete *subset-sum* problem, and as such we do not expect small refutations in general (assuming NP $\neq$ coNP). Indeed, Impagliazzo, Pudlák, and Sgall [IPS99] have shown lower bounds for refutations in the *polynomial calculus* system (and thus also the Nullstellensatz system) even when $\overline{\alpha} = \overline{1}$. Specifically, they showed that such refutations require both $\Omega(n)$-degree and $\exp(\Omega(n))$-many monomials. In the language of this paper, they gave $\exp(\Omega(n))$-size lower bounds for refuting this system in $\sum\prod$-IPS$_{\text{LIN}}$. In contrast, we establish here poly(n)-size refutations for $\overline{\alpha} = \overline{1}$ in the restricted proof systems of roABP-IPS$_{\text{LIN}}$ and *depth-3* multilinear-formula-IPS$_{\text{LIN}}$ (despite the fact that multilinear-formula-IPS$_{\text{LIN}}$ is not complete).

IT: It should be Corollary 22, I think.

Theorem (Theorem 4.8, Theorem 4.9)**.** *Let $\mathbb{F}$ be a field of characteristic* $\text{char}(\mathbb{F}) > n$*. Then the system of polynomial equations $\sum_{i=1}^n x_i - \beta$, $\{x_i^2 - x_i\}_{i=1}^n$ is unsatisfiable for $\beta \in \mathbb{F} \setminus \{0, \ldots, n\}$, and there are explicit* poly(n)*-size refutations within roABP-IPS$_{LIN}$, as well as within depth-3 multilinear-formula-IPS$_{LIN}$.*

This claim is proven by noting that the polynomial $p(t) := \prod_{k=0}^n (t-k)$ vanishes on $\sum_i x_i$ modulo $\{x_i^2 - x_i\}_{i=1}^n$, but $p(\beta)$ is a non-zero constant. This implies that $\sum_i x_i - \beta$ divides $p(\sum_i x_i) - p(\beta)$. Denoting the quotient by $f(\overline{x})$, it follows that $\frac{1}{-p(\beta)} \cdot f(\overline{x}) \cdot (\sum_i x_i - \beta) \equiv 1 \mod \{x_i^2 - x_i\}_{i=1}^n$, which is nearly a linear-IPS refutation except for the complexity of establishing this relation over the boolean cube. We show that the quotient f is easily expressed as a depth-3 powering circuit. Unfortunately, proving the above equivalence to 1 modulo the boolean cube is not possible in the depth-3 powering circuit model. However, by moving to more powerful models (such as roABPs and multilinear formulas) we can give proofs of this multilinearization to 1 and thus give proper IPS refutations.

1.3.2 Linear-IPS Lower Bounds via Functional Lower Bounds

Having demonstrated the power of various restricted classes of IPS refutations by refuting the subset-sum axiom, we now turn to lower bounds. We give two paradigms for establishing lower bounds, the first of which we discus here, named a *functional circuit lower bound*. This term appeared in the work of Grigoriev and Razborov [GR00] as well as in the recent work of Forbes, Kumar and Saptharishi [FKS15]. We briefly motivate this type of lower bound as a topic of independent interest in algebraic circuit complexity, and then discuss the lower bounds we obtain and their implications to obtaining proof complexity lower bounds.

In algebraic complexity one computes polynomials *syntactically* as objects in the ring $\mathbb{F}[x_1, \ldots, x_n]$. Thus, even if one is only interested in evaluating the polynomial over the boolean cube, yielding a function $\hat{f} : \{0,1\}^n \to \mathbb{F}$, an algebraic computation of the polynomial necessarily gives a method for evaluating the polynomial over $\mathbb{F}$ as well as any extension of $\mathbb{F}$. However, some polynomials such as the permanent are known in boolean complexity to have complex behavior as functions even over boolean inputs, so one would expect that *any* polynomial f that agrees with the permanent on boolean inputs must require large algebraic circuits. We call such results functional circuit lower bounds. Prior work ([GK98, GR00, KS15]) has established functional lower bounds over fixed-size finite fields, and the recent work of Forbes, Kumar and Saptharishi [FKS15] has established some lower bounds for any field.

Goal 1.3 (Functional Circuit Lower Bound ([GR00, FKS15]))**.** *Obtain explicit functions $\hat{f} : \{0,1\}^n \to \mathbb{F}$ such that for any polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ such that $f(\overline{x}) = \hat{f}(\overline{x})$ for all $\overline{x} \in \{0,1\}^n$, it must be that f requires large algebraic circuits.*

While it is natural to hope that existing methods would yield such lower bounds, many lower bound techniques inherently use that algebraic computation is *syntactic*. In particular, techniques involving partial derivatives (which include the partial derivative method of Nisan-Wigderson [NW96] and the shifted partial derivative method of Kayal [Kay12, GKKS14]) cannot as is yield functional lower bounds as computing the partial derivative of a polynomial is *not local*, so that knowing a polynomial on $\{0,1\}^n$ is not enough information to conclude information about its partial derivatives.

We now explain how functional lower bounds imply lower bounds for linear-IPS refutations in certain cases. Suppose one considers refutations of the unsatisfiable polynomial system $f(\overline{x}), \{x_i^2 - x_i\}_{i=1}^n$. A linear-IPS refutation would yield an equation of the form $g(\overline{x}) \cdot f(\overline{x}) + \sum_i h_i(\overline{x}) \cdot (x_i^2 - x_i) = 1$ for some polynomials $g, h_i \in \mathbb{F}[\overline{x}]$. Viewing this equation modulo the boolean cube, we have that $g(\overline{x}) \cdot f(\overline{x}) \equiv 1 \mod \{x_i^2 - x_i\}_i$. Equivalently, since $f(\overline{x})$ is unsatisfiable over $\{0,1\}^n$, we see that $g(\overline{x}) = 1/f(\overline{x})$ for $\overline{x} \in \{0,1\}^n$, as $f(\overline{x})$ is never zero so this fraction is well-defined. It follows that *if* the function $\overline{x} \mapsto 1/f(\overline{x})$ induces a functional lower bound then $g(\overline{x})$ must require large complexity, yielding the desired linear-IPS lower bound.

Thus, it remains to instantiate this program. While we are successful, we should note that this approach as is seems to only yield proof complexity lower bounds for systems with one non-boolean axiom and thus cannot encode polynomial systems arising from 3CNFs.

Our starting point is to observe that the subset-sum axiom already induces a weak form of functional lower bound, where the complexity is measured by degree.

Theorem (Theorem 5.1)**.** *Let $\mathbb{F}$ be a field of a characteristic $\mathsf{poly}(n)$ and $\beta \notin \{0, \ldots, n\}$. Then $\sum_i x_i - \beta, \{x_i^2 - x_i\}_i$ is unsatisfiable and any polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ with $f(\overline{x}) = \frac{1}{\sum_i x_i - \beta}$ for $\overline{x} \in \{0,1\}^n$, satisfies $\deg f \geq n$.*

A lower bound of $\lceil \frac{n}{2} \rceil$ was previously established by Impagliazzo, Pudlák, and Sgall [IPS99], but the bound of 'n' (which is tight) will be crucial for our results.

We then lift this result to obtain lower bounds for stronger models of algebraic complexity. In particular, by replacing "x_i" by "$x_i y_i$" we show that the function $\frac{1}{\sum_i x_i y_i - \beta}$ has maximal *evaluation dimension* between $\overline{x}$ and $\overline{y}$, which is measure of correlation. This measure is essentially *functional*, so that one can lower bound this measure by understanding the functional behavior of the polynomial on finite sets such as the boolean cube. Our lower bound for evaluation dimension follows by examining the above degree bound. Using known relations between this complexity measure and algebraic circuit classes, we can obtain lower bounds for depth-3 powering linear-IPS.

Theorem (Theorem 5.8)**.** *Let $\mathbb{F}$ be a field of characteristic $\geq \mathsf{poly}(n)$ and $\beta \notin \{0, \ldots, n\}$. Then $\sum_{i=1}^n x_i y_i - \beta, \{x_i^2 - x_i\}_i, \{y_i^2 - y_i\}_i$ is unsatisfiable and any $\sum \bigwedge \sum$-IPS$_{LIN}$ refutation requires size $\geq \exp(\Omega(n))$.*

The above axiom only gets maximal correlation between a *fixed* partition of the variables. By introducing auxiliary variables we can create such correlation between *any* partition of (some) of the variables. By again invoking results showing such structure implies computational hardness we obtain more linear-IPS lower bounds.

Theorem (Theorem 5.12)**.** *Let $\mathbb{F}$ be a field of characteristic $\geq \mathsf{poly}(n)$ and $\beta \notin \{0, \ldots, \binom{2n}{2}\}$. Then $\sum_{i<j} z_{i,j} x_i x_j - \beta, \{x_i^2 - x_i\}_{i=1}^n, \{z_{i,j}^2 - z_{i,j}\}_{i<j}$ is unsatisfiable, and any roABP-IPS$_{LIN}$ refutation (in any variable order) requires $\exp(\Omega(n))$-size. Further, any multilinear-formula-IPS$_{LIN'}$ refutation requires $n^{\Omega(\log n)}$-size, and any depth-$(2d+1)$ multilinear-formula-IPS$_{LIN'}$ refutation requires $n^{\Omega((n/\log n)^{1/d}/d^2)}$-size.*

Thus, we show that even though roABP-IPS$_{\text{LIN}}$ and depth-3 multilinear formula-IPS$_{\text{LIN}'}$ can refute the subset-sum axiom in polynomial size, slight variants of this axiom do not have polynomial-size refutations.

1.3.3 Lower Bounds for Multiples

While the above paradigm can establish super-polynomial lower bounds for *linear*-IPS, it does not seem able to establish lower bounds for the general IPS proof system, even for restricted classes. This is because such systems would induce equations such as $h(\overline{x})f(\overline{x})^2 + g(\overline{x})f(\overline{x}) \equiv 1 \mod \{x_i^2 - x_i\}_{i=1}^n$, where we need to design a computationally simple axiom f so that this equation implies at least one of h or g is of large complexity. In the linear-IPS case we could assume h was zero, so that we can uniquely solve for $g(\overline{x})$ for $\overline{x} \in \{0,1\}^n$. However, in general knowing $f(\overline{x})$ does not uniquely determine $g(\overline{x})$ or $h(\overline{x})$, which makes this approach significantly more complicated. Further, even though we can efficiently simulate IPS by linear-IPS (Theorem 4.1) in general, this simulation increases the complexity of the proof so that even if one started with a $\mathcal{C}$-IPS proof for a restricted circuit class $\mathcal{C}$ the resulting IPS$_{\text{LIN}}$ proof may not be in $\mathcal{C}$-IPS$_{\text{LIN}}$.

As such, we introduce a second paradigm, called *lower bounds for multiples*, which can yield $\mathcal{C}$-IPS lower bounds for various restricted classes $\mathcal{C}$. We begin by defining this question.

Goal 1.4 (Lower Bounds for Multiples)**.** *Design an explicit polynomial $f(\overline{x})$ such that for any non-zero $g(\overline{x})$ we have that $g(\overline{x})f(\overline{x})$ is hard to compute.*

We now explain how such lower bounds yield IPS lower bounds. Consider the system $f, \{x_i^2 - x_i\}_i$ with a single non-boolean axiom. An IPS refutation is a circuit $C(\overline{x}, y, \overline{z})$ such that $C(\overline{x}, 0, \overline{0}) = 1$ and $C(\overline{x}, f, \overline{x}^2 - \overline{x}) = 1$, where (as mentioned) $\overline{x}^2 - \overline{x}$ denotes $\{x_i^2 - x_i\}_i$. Expressing this polynomial as a univariate in f, we obtain that $\sum_{i \geq 1} C_i(\overline{x}, \overline{x}^2 - \overline{x}) f^i = 1 - C(\overline{x}, 0, \overline{x}^2 - \overline{x})$ for some polynomials C_i. For many natural measures of circuit complexity $1 - C(\overline{x}, 0, \overline{x}^2 - \overline{x})$ has complexity roughly bounded by that of C itself. Though not strictly necessary for this method, it is worth noting that the complexity of each of the C_i is not much larger than that of C, as one can compute the C_i by homogenizing or interpolating C in the variable y (see for example the survey of Shpilka and Yehudayoff [SY10]). Thus, we see that a multiple of f has a small circuit, as $\left(\sum_{i \geq 1} C_i(\overline{x}, \overline{x}^2 - \overline{x}) f^{i-1}\right) \cdot f = 1 - C(\overline{x}, 0, \overline{x}^2 - \overline{x})$. Thus, if we can show that all multiples of f require large circuits then we rule out a small IPS refutation.

We now turn to methods for obtaining polynomials with hard multiples. Intuitively if a polynomial f is hard then so should small modifications such as $f^2 + x_1 f$, and this intuition is supported by the result of Kaltofen [Kal89] which shows that if a polynomial has a small algebraic circuit then so do all of its factors. As a consequence, if a polynomial requires super-polynomially large algebraic circuits than so do all of its multiples. However, Kaltofen's [Kal89] result is about *general* algebraic circuits, and there are very limited results about the complexity of factors of *restricted* algebraic circuits ([DSY09, Oli15b]) so that obtaining polynomials for hard multiples via factorization results seems difficult.

However, note that lower bound for multiples has a different order of quantifiers than the factoring question. That is, Kaltofen's [Kal89] result speaks about the factors of *any* small circuit, while the lower bound for multiples speaks about the multiples of a *single* polynomial. Thus, it seems plausible that existing methods could yield such explicit polynomials, and indeed we show this is the case.

We begin by noting that obtaining lower bounds for multiples is a natural instantiation of the algebraic *hardness versus randomness* paradigm. In particular, Heintz-Schnorr [HS80] and Agrawal [Agr05] showed that obtaining deterministic (black-box) PIT algorithms implies lower

bounds (see Subsection 3.1 for more on PIT), and we strengthen that connection here to lower bounds for multiples. We can actually instantiate this connection, and we use slight modifications of existing PIT algorithms to show that multiples of the determinant are hard in some models.

Theorem (Informal Version of Theorem 6.2, Theorem 6.7)**.** *Let $\mathcal{C}$ be a restricted class of n-variate algebraic circuits. Full derandomization of PIT algorithms for $\mathcal{C}$ yields an explicit polynomials all of whose multiples require $\exp(\Omega(n))$-size as $\mathcal{C}$-circuits.*

In particular, when $\mathcal{C}$ is the class of sparse polynomials, depth-3 powering formulas, $\sum\bigwedge\sum\prod^{\mathcal{O}(1)}$ formulas (in characteristic zero), or "every-order" roABPs, the $n \times n$ determinant has multiples which are all $\exp(\Omega(n))$-hard in these models.

The above statement shows that *derandomization* implies *hardness*. We also partly address the converse direction by arguing (Subsection 6.1) that hardness-to-randomness construction of Kabanets and Impagliazzo [KI04] only requires lower bounds for multiples to derandomize PIT. Unfortunately, this direction is harder to instantiate for restricted classes as it requires lower bounds for classes with suitable closure properties.[2]

Unfortunately the above result is slightly unsatisfying from a proof complexity standpoint as the (exponential-size) lower bounds for the subclasses of IPS one can derive from the above result would involve the determinant polynomial as an axiom. While the determinant is efficiently computable, it is not computable by the above restricted circuit classes (indeed, the above result proves that). As such, this would not fit the real goal of proof complexity which seeks to show that there are statements whose proofs must be *super-polynomial larger* than the length of the statement. Thus, if we measure the size of the IPS proof and the axioms with respect to the same circuit measure, the lower bounds for multiples approach *cannot* establish such super-polynomial lower bounds.

However, we believe that lower bounds for multiples could lead, with further ideas, to proof complexity lower bounds in the conventional sense. That is, it seems plausible that by adding *extension variables* we can convert complicated axioms to simple, local axioms by tracing through the computation of that axiom. That is, consider the axiom $xyzw$. This can be equivalently written as $\{a - xy, b - zw, c - ab, c\}$, where this conversion is done by considering a natural algebraic circuit for $xyzw$, replacing each gate with a new variable, and adding an axiom ensuring the new variables respect the computation of the circuit. While we are unable to understand the role of extension variables in this work, we aim to give as simple axioms as possible whose multiples are all hard as this may facilitate future work on extension variables.

We now discuss the lower bounds for multiples we obtain.[3]

Theorem (Theorem 6.8, Theorem 6.9, Theorem 6.10, Theorem 6.18,Theorem 6.20)**.** *We obtain the following lower bounds for multiples.*

- *All non-zero multiples of $x_1 \cdots x_n$ require $\exp(\Omega(n))$-size as a depth-3 powering formula (over any field), or as a $\sum\bigwedge\sum\prod^{\mathcal{O}(1)}$ formula (in characteristic zero).*
- *All non-zero multiples of $(x_1 + 1) \cdots (x_n + 1)$ require $\exp(\Omega(n))$-many monomials.*
- *All non-zero multiples of $\prod_i (x_i + y_i)$ require $\exp(\Omega(n))$-width as a roABPs in any variable order where $\overline{x}$ precedes $\overline{y}$.*

[2] Although, we note that one can instantiate this connection with depth-3 powering formulas (or even $\sum\bigwedge\sum\prod^{\mathcal{O}(1)}$ formulas) using the lower bounds for multiples developed in this paper, building on the work of Forbes [For15]. However, the resulting PIT algorithms are worse than those developed by Forbes [For15].

[3] While we discussed functional lower bounds for multilinear formulas, this class is not interesting for the lower bounds for multiples question. This is because a multiple of a multilinear polynomial may not be multilinear, and thus clearly cannot have a multilinear formula.

- *All non-zero multiples of* $\prod_{i,j=1}^{n}(z_{i,j}\cdot(x_i+x_j+x_ix_j)+(1-z_{i,j}))$ *require* $\exp(\Omega(n))$*-width as a roABP in any variable order, as well as* $\exp(\Omega(n))$*-width as a read-twice oblivious ABP.*

We now briefly explain our techniques for obtaining these lower bounds, focusing on the simplest case of depth-3 powering formulas. It follows from the partial derivative method of Nisan and Wigderson [NW94] (see Kayal [Kay08]) that such formulas require exponential size to compute the monomial $x_1 \dots x_n$ *exactly*. Forbes and Shpilka [FS13a], in giving a PIT algorithm for this class, showed that this lower bound can be *scaled down* and *made robust*. That is, if one has a size-s depth-3 powering formula, it follows that *if* it computes a monomial $x_{i_1}\cdots x_{i_\ell}$ for distinct i_j then $l \leq O(\log s)$ (so the lower bound is scaled down). One can then show that regardless of what this formula actually computes the *leading* monomial $x_{i_1}^{a_{i_1}}\cdots x_{i_\ell}^{a_{i_\ell}}$ (for distinct i_j and positive a_{i_j}) must have that $\ell \leq O(\log s)$. One then notes that leading monomials are *multiplicative*. Thus, for any non-zero g the leading monomial of $g\cdot x_1\dots x_n$ involves n variables so that if $g\cdot x_1\dots x_n$ is computed in size-s then $n \leq O(\log s)$, giving $s \geq \exp(\Omega(n))$ as desired. One can then obtain the other lower bounds using the same idea, though for roABPs one needs to define a leading *diagonal* (refining an argument of Forbes-Shpilka [FS12]).

We now conclude our IPS lower bounds.

Theorem (Theorem 7.1,Theorem 7.2)**.** *We obtain the following lower bound for subclasses of IPS.*

- *In characteristic zero, for* $m \neq n$*, the system of polynomials* $x_1\cdots x_n - 1, x_1+\cdots+x_n-m, \{x_i^2-x_i\}_{i=1}^n$ *is unsatisfiable, any* $\sum\bigwedge\sum$*-IPS refutation requires* $\exp(\Omega(n))$*-size.*
- *The system of polynomials,* $1+\prod_{i,j=1}^{n}(z_{i,j}(x_i+x_j-x_ix_j)+(1-z_{i,j})), \{x_i^2-x_i\}_i, \{z_{i,j}^2-z_{i,j}\}_{i,j}$ *is unsatisfiable, and any roABP-IPS refutation (in any variable order) must be of width* $\exp(\Omega(n))$*.*

Note that the first result is an encoding that $\mathrm{AND}(x_1,\dots,x_n)=1$ but $\mathrm{OR}(x_1,\dots,x_n)\neq 1$. The second is not as natural, but contains the simpler polynomial $\prod_i(u_i+v_i-u_iv_i)+1$ (up to renaming, and after appropriate substitution of the $z_{i,j}$ to values from $\{0,1\}$), which encodes that $\mathrm{AND}(\mathrm{OR}(u_1,v_1),\cdots,\mathrm{OR}(u_n,v_n))\notin\{0,1\}$.

1.4 Organization

The rest of the paper is organized as follows. Section 2 contains the basic notation for the paper. In Section 3 we give background from algebraic complexity, including several important complexity measures such as coefficient dimension and evaluation dimension (see Subsection 3.2 and Subsection 3.3). We present our upper bounds for IPS in Section 4. In Section 5 we give our functional lower bounds and from them obtain lower bounds for $\mathrm{IPS}_{\mathrm{LIN}}$. Section 6 contains our lower bounds for multiples of polynomials and in Section 7 we derive lower bounds for IPS using them. In Section 8 we study the relative strength of fragments of IPS, compared to other close algebraic proof systems that were studied in previous works. Finally, in Section 9 we list some problems which were left open by this work.

2 Notation

In this section we briefly describe notation used in this paper. We denote $[n] := \{1,\dots,n\}$. For a vector $\overline{a}\in\mathbb{N}^n$, we denote $\overline{x}^{\overline{a}} := x_1^{a_1}\cdots x_n^{a_n}$ so that in particular $\overline{x}^{\overline{1}} = \prod_{i=1}^n x_i$. The (total) degree of a monomial $\overline{x}^{\overline{a}}$, denoted $\deg\overline{x}^{\overline{a}}$, is equal to $|\overline{a}|_1$, and the individual degree, denoted $\mathrm{ideg}\,\overline{x}^{\overline{a}}$, is

equal to $|\overline{a}|_\infty$. Degree and individual degree can be defined for a polynomial f, denoted $\deg f$ and $\operatorname{ideg} f$ respectively, by taking the maximum over all monomials with non-zero coefficients in f. We will use $\prec$ to denote a monomial order on $\mathbb{F}[\overline{x}]$, see Subsection 3.6.

Polynomials will often be written out in their monomial expansion. At various points we will need to extract coefficients from polynomials. When "taking the coefficient of $\overline{y}^{\overline{b}}$ in $f \in \mathbb{F}[\overline{x}, \overline{y}]$" we mean that both $\overline{x}$ and $\overline{y}$ are treated as variables and thus the coefficient returned is a scalar in $\mathbb{F}$, and this will be denoted $\operatorname{Coeff}_{\overline{y}^{\overline{b}}}(f)$. However, when "taking the coefficient of $\overline{y}^{\overline{b}}$ in $f \in \mathbb{F}[\overline{x}][\overline{y}]$" we mean that $\overline{x}$ is now part of the ring of scalars, so the coefficient will be an element of $\mathbb{F}[\overline{x}]$, and this coefficient will be denoted $\operatorname{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(f)$.

For a vector $\overline{a} \in \mathbb{N}^n$ we denote $\overline{a}_{\leq i} \in \mathbb{N}^i$ to be the restriction of $\overline{a}$ to the first i coordinates. For a set $S \subseteq [n]$ we let $\overline{S}$ denote the complement set. We will denote the size-k subsets of $[n]$ by $\binom{[n]}{k}$. We will use $\operatorname{ml} : \mathbb{F}[\overline{x}] \to \mathbb{F}[\overline{x}]$ to denote the multilinearization operator, defined by Theorem 3.7. We will use $\overline{x}^2 - \overline{x}$ to denote the set of equations $\{x_i^2 - x_i\}_i$.

3 Algebraic Complexity Theory Background

In this section we state some known facts regarding the algebraic circuit classes that we will be studying. We also give some important definitions that will be used later in the paper.

3.1 Polynomial Identity Testing

In the *polynomial identity testing (PIT)* problem, we are given an algebraic circuit computing some polynomial f, and we have to determine whether "$f \equiv 0$". That is, we are asking whether f is the zero polynomial in $\mathbb{F}[x_1, \ldots, x_n]$. By the Schwartz-Zippel-DeMillo-Lipton Lemma [Zip79, Sch80, DL78], if $0 \neq f \in \mathbb{F}[\overline{x}]$ is a polynomial of degree $\leq d$ and $S \subseteq \mathbb{F}$, and $\overline{\alpha} \in S^n$ is chosen uniformly at random, then $f(\overline{\alpha}) = 0$ with probability at most[4] $d/|S|$. Thus, given the circuit, we can perform these evaluations efficiently,[5] giving an efficient randomized procedure for deciding whether "$f \equiv 0$?". It is an important open problem to find a derandomization of this algorithm, that is, to find a *deterministic* procedure for PIT that runs in polynomial time (in the size of circuit).

Note that in the randomized algorithm of Schwartz-Zippel-DeMillo-Lipton we only use the circuit to compute the evaluation $f(\overline{\alpha})$. Such algorithms are said to run in the *black-box* model. In contrast, an algorithm that can access the internal structure of the circuit runs in the *white-box* model. It is a folklore result that efficient deterministic black-box algorithms are equivalent to constructions of small *hitting sets*. That is, a hitting set is set of inputs so that any nonzero circuit from the relevant class evaluates to nonzero on at least one of the inputs in the set. For more on PIT we refer to the survey of Shpilka and Yehudayoff [SY10].

A related notion to that of a hitting set is that of a *generator*, which is essentially a low-dimensional curve whose image contains a hitting set. The equivalence between hitting sets and generators can be found in the above mentioned survey.

[4] Note that this is meaningful only if $d < |S| \leq |\mathbb{F}|$, which in particular implies that f is not the zero function.

[5] To present algorithms that are field independent, this paper works in a model of computation where field operations (such as addition, multiplication, inversion and zero-testing) over $\mathbb{F}$ can be computed at unit cost, see for example Forbes [For14, Appendix A]. We say that an algebraic circuit is **t-explicit** if it can be constructed in t steps in this unit-cost model.

Definition 3.1. *Let $\mathcal{C} \subseteq \mathbb{F}[x_1, \ldots, x_n]$ be a set of polynomials. A polynomial $\overline{\mathcal{G}} : \mathbb{F}^s \to \mathbb{F}^n$ is a* ***generator for $\mathcal{C}$ with seed-length*** *s if for all $f \in \mathcal{C}$, $f \equiv 0$ iff $f \circ \overline{\mathcal{G}} \equiv 0$. That is, $f(\overline{x}) = 0$ in $\mathbb{F}[\overline{x}]$ iff $f(\overline{\mathcal{G}}(\overline{y})) = 0$ in $\mathbb{F}[\overline{y}]$.*

In words, a generator for a circuit class $\mathcal{C}$ is a mapping $\overline{\mathcal{G}} : \mathbb{F}^t \to \mathbb{F}^n$, such that for any nonzero polynomial f, computed by a circuit from $\mathcal{C}$, it holds that the composition $f(\overline{\mathcal{G}})$ is nonzero as well. By considering the image of $\overline{\mathcal{G}}$ on S^t, where $S \subseteq \mathbb{F}$ is of polynomial size, we obtain a hitting set for $\mathcal{C}$.

In Subsection 6.2 we explain how one can use generators with small seed-length to obtain lower bounds for any nonzero multiple of a given polynomial f. Such generators are known for several of the models that we study in this paper.

Sparse Representation: There are many papers giving efficient black-box PIT algorithm for $\Sigma\Pi$ formulas. For example, Klivans and Spielman [KS01] gave a hitting set of polynomial size.

Depth-3 Powering Formulas: Saxena [Sax08] gave a polynomial time white-box PIT algorithm and Forbes, Shpilka, and Saptharishi [FSS14] gave a hitting set of size $s^{O(\log \log s)}$ for size-s depth-3 powering formulas.

$\sum \bigwedge \sum \prod^{\mathcal{O}(1)}$ Formulas: Forbes [For15] gave an $s^{\mathcal{O}(\lg s)}$-size hitting set for size-s computation in this model (in large characteristic).

Read-once ABPs: Raz and Shpilka [RS05] gave a polynomial time white-box PIT algorithm. A long sequence of papers calumniated in the work of Agrawal, Gurjar, Korwar, and Saxena [AGKS14], who gave a quasi-polynomial sized hitting set for roABPs.

Read-k Oblivious ABPs: In a very recent work, Anderson, Forbes, Saptharishi, Shpilka, Volk [AFS+15] obtained a white-box PIT algorithm for read-k oblivious ABPs that run in time $2^{\tilde{O}(n^{1-1/2^{k-1}})}$ and needs white box access only to know the order in which the variables appear in the ABP.

3.2 Coefficient Dimension and roABPs

This paper proves various lower bounds on roABPs using a complexity measures known as *coefficient dimension*. In this section, we define this measures and recall basic properties. Full proofs of these claims can be found for example in the thesis of Forbes [For14].

We first define the *coefficient matrix* of a polynomial, called the "partial derivative matrix" in the prior work of Nisan [Nis91] and Raz [Raz09]. This matrix is formed from a polynomial $f \in \mathbb{F}[\overline{x}, \overline{y}]$ by arranging its coefficients into a matrix. That is, the coefficient matrix has rows indexed by monomials $\overline{x}^{\overline{a}}$ in $\overline{x}$, columns indexed by monomials $\overline{y}^{\overline{b}}$ in $\overline{y}$, and the $(\overline{x}^{\overline{a}}, \overline{y}^{\overline{b}})$-entry is the coefficient of $\overline{x}^{\overline{a}}\overline{y}^{\overline{b}}$ in the polynomial f. We now define this matrix, recalling that $\mathrm{Coeff}_{\overline{x}^{\overline{a}}\overline{y}^{\overline{b}}}(f)$ is the coefficient of $\overline{x}^{\overline{a}}\overline{y}^{\overline{b}}$ in f.

Definition 3.2. *Consider $f \in \mathbb{F}[\overline{x}, \overline{y}]$. Define the* ***coefficient matrix of*** *f as the scalar matrix*

$$(C_f)_{\overline{a},\overline{b}} := \mathrm{Coeff}_{\overline{x}^{\overline{a}}\overline{y}^{\overline{b}}}(f) \ ,$$

where coefficients are taken in $\mathbb{F}[\overline{x}, \overline{y}]$, for $|\overline{a}|_1, |\overline{b}|_1 \leq \deg f$.

We now give the related definition of *coefficient dimension*, which looks at the dimension of the row- and column-spaces of the coefficient matrix. Recall that $\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(f)$ extracts the coefficient of $\overline{y}^{\overline{b}}$ in f as a polynomial in $\mathbb{F}[\overline{x}][\overline{y}]$.

Definition 3.3. *Let* $\mathbf{Coeff}_{\overline{x}|\overline{y}} : \mathbb{F}[\overline{x},\overline{y}] \to 2^{\mathbb{F}[\overline{x}]}$ *be the* ***space of*** $\mathbb{F}[\overline{x}][\overline{y}]$ ***coefficients****, defined by*

$$\mathbf{Coeff}_{\overline{x}|\overline{y}}(f) := \left\{ \mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(f) \right\}_{\overline{b}\in\mathbb{N}^n} ,$$

where coefficients of f *are taken in* $\mathbb{F}[\overline{x}][\overline{y}]$.

Similarly, define $\mathbf{Coeff}_{\overline{y}|\overline{x}} : \mathbb{F}[\overline{x},\overline{y}] \to 2^{\mathbb{F}[\overline{y}]}$ *by taking coefficients in* $\mathbb{F}[\overline{y}][\overline{x}]$.

The following basic lemma shows the rank of the coefficient matrix equals the coefficient dimension.

Lemma 3.1 (Nisan [Nis91])**.** *Consider* $f \in \mathbb{F}[\overline{x},\overline{y}]$. *Then the rank of the coefficient matrix* C_f *obeys*

$$\operatorname{rank} C_f = \dim \mathbf{Coeff}_{\overline{x}|\overline{y}}(f) = \dim \mathbf{Coeff}_{\overline{y}|\overline{x}}(f) .$$

Thus, the ordering of the partition ($(\overline{x},\overline{y})$ versus $(\overline{y},\overline{x})$) does not matter in terms of the resulting dimension. The above matrix-rank formulation of coefficient dimension can be rephrased in terms of low-rank decompositions.

Lemma 3.2. *Let* $f \in \mathbb{F}[\overline{x},\overline{y}]$. *Then* $\dim \mathbf{Coeff}_{\overline{x}|\overline{y}}(f)$ *equals the minimum* r *such that there are* $\overline{g} \in \mathbb{F}[\overline{x}]^r$ *and* $\overline{h} \in \mathbb{F}[\overline{y}]^r$ *such that* f *can be written as* $f(\overline{x},\overline{y}) = \sum_{i=1}^r g_i(\overline{x})h_i(\overline{y})$.

We now state a convenient normal form for roABPs (see for example Forbes [For14, Corollary 4.4.2]).

Lemma 3.3. *Consider a polynomial* $f \in \mathbb{F}[x_1,\ldots,x_n]$ *and let* $\pi : [n] \to [n]$ *be a permutation. The polynomial* f *is computed by width-*r *roABP in variable order* π *iff there exist matrices* $A_i(x_{\pi(i)}) \in \mathbb{F}[x_{\pi(i)}]^{r\times r}$ *of (individual) degree* $\leq \deg f$ *such that* $f = (\prod_{i=1}^n A_i(x_{\pi(i)}))_{1,1}$.

By splitting a roABP into such variable disjoint inner products one can obtain a lower bound for roABP width via coefficient dimension. In fact, this complexity measure *characterizes* roABP width.

Lemma 3.4. *Let* $f \in \mathbb{F}[x_1,\ldots,x_n]$ *be computed by a width-*r *roABP in variable order* $x_1 < \cdots < x_n$, *so that* f *can be computed as* $f(\overline{x}) = (\prod_{i=1}^n A_i(x_i))_{1,1}$ *for matrices* $A_i(x_i) \in \mathbb{F}[x_i]^{r\times r}$. *Then* $r \geq \max_i \dim \mathbf{Coeff}_{\overline{x}_{\leq i}|\overline{x}_{>i}}(f)$. *Further,* f *is computable by a roABP in variable order* $x_1 < \cdots < x_n$ *of width* $\max_i \dim \mathbf{Coeff}_{\overline{x}_{\leq i}|\overline{x}_{>i}}(f)$.

Using this complexity measure it is rather straightforward to prove the following closure properties of roABPs.

Fact 3.5. *If* $f, g \in \mathbb{F}[\overline{x}]$ *are computable by width-*r *and width-*s *roABPs respectively, then*

- $f + g$ *is computable by a width* $\max\{r, s\}$ *roABP.*
- $f \cdot g$ *is computable by a width-*(rs) *roABP.*

Further, if $f(\overline{x},\overline{y}) \in \mathbb{F}[\overline{x},\overline{y}]$ *is computable by a width-*r *roABP in some variable order then the partial substitution* $f(\overline{x},\overline{\alpha})$, *for* $\overline{\alpha} \in \mathbb{F}^{|\overline{y}|}$, *is computable by a width-*r *roABP in the induced order on* $\overline{x}$.

3.3 Evaluation Dimension

While coefficient dimension measures the size of a polynomial $f(\overline{x},\overline{y})$ by taking all coefficients in $\overline{y}$, *evaluation dimension* is a complexity measure due to Saptharishi [Sap12] that measures the size by taking all possible evaluations in $\overline{y}$ over the field. This measure will be important for our applications as one can restrict such evaluations to the boolean cube and obtain circuit lower bounds for computing $f(\overline{x},\overline{y})$ as a *polynomial* via its induced *function* on the boolean cube. We begin with the definition.

Definition 3.4 (Saptharishi [Sap12])**.** *Let $S \subseteq \mathbb{F}$. Let $\mathbf{Eval}_{\overline{x}|\overline{y},S} : \mathbb{F}[\overline{x},\overline{y}] \to 2^{\mathbb{F}[\overline{x}]}$ be the **space of $\mathbb{F}[\overline{x}][\overline{y}]$ evaluations over** S, defined by*

$$\mathbf{Eval}_{\overline{x}|\overline{y},S}(f(\overline{x},\overline{y})) := \left\{ f(\overline{x},\overline{\beta}) \right\}_{\overline{\beta}\in S^{|\overline{y}|}} .$$

Define $\mathbf{Eval}_{\overline{x}|\overline{y}} : \mathbb{F}[\overline{x},\overline{y}] \to 2^{\mathbb{F}[\overline{x}]}$ to be $\mathbf{Eval}_{\overline{x}|\overline{y},S}$ when $S = \mathbb{F}$.

Similarly, define $\mathbf{Eval}_{\overline{y}|\overline{x},S} : \mathbb{F}[\overline{x},\overline{y}] \to 2^{\mathbb{F}[\overline{y}]}$ by replacing $\overline{x}$ with all possible evaluations $\overline{\alpha} \in S^{|\overline{x}|}$, and likewise define $\mathbf{Eval}_{\overline{y}|\overline{x}} : \mathbb{F}[\overline{x},\overline{y}] \to 2^{\mathbb{F}[\overline{y}]}$.

The equivalence between evaluation dimension and coefficient dimension was shown by Forbes-Shpilka [FS13b] by appealing to interpolation.

Lemma 3.6 (Forbes-Shpilka [FS13b])**.** *Let $f \in \mathbb{F}[\overline{x},\overline{y}]$. For any $S \subseteq \mathbb{F}$ we have that $\mathbf{Eval}_{\overline{x}|\overline{y},S}(f) \subseteq \operatorname{span} \mathbf{Coeff}_{\overline{x}|\overline{y}}(f)$ so that $\dim \mathbf{Eval}_{\overline{x}|\overline{y},S}(f) \leq \dim \mathbf{Coeff}_{\overline{x}|\overline{y}}(f)$. In particular, if $|S| > \operatorname{ideg} f$ then $\dim \mathbf{Eval}_{\overline{x}|\overline{y},S}(f) = \dim \mathbf{Coeff}_{\overline{x}|\overline{y}}(f)$.*

3.4 Multilinear Polynomials and Multilinear Formulas

We state some well-known facts about multilinear polynomials.

Fact 3.7. *For any two multilinear polynomials $f, g \in \mathbb{F}[x_1,\ldots,x_n]$, $f = g$ as polynomials iff they agree on the boolean cube $\{0,1\}^n$. That is, $f = g$ iff $f|_{\{0,1\}^n} = g|_{\{0,1\}^n}$.*

*Further, there is a **multilinearization** map $\operatorname{ml} : \mathbb{F}[\overline{x}] \to \mathbb{F}[\overline{x}]$ such that for any $f, g \in \mathbb{F}[\overline{x}]$,*

1. *$\operatorname{ml}(f)$ is multilinear.*
2. *f and $\operatorname{ml}(f)$ agree on the boolean cube, that is, $f|_{\{0,1\}^n} = \operatorname{ml}(f)|_{\{0,1\}^n}$.*
3. *$\deg \operatorname{ml}(f) \leq \deg f$.*
4. *$\operatorname{ml}(fg) = \operatorname{ml}(\operatorname{ml}(f)\operatorname{ml}(g))$.*
5. *ml is linear, so that for any $\alpha,\beta \in \mathbb{F}$, $\operatorname{ml}(\alpha f + \beta g) = \alpha \operatorname{ml}(f) + \beta \operatorname{ml}(g)$.*

Throughout the rest of this paper 'ml' will denote the multilinearization operator. Raz [Raz09, Raz06] gave lower bounds for multilinear formulas using the above notion of coefficient dimension, and Raz-Yehudayoff [RY08,RY09] gave simplifications and extensions to constant-depth multilinear formulas.

Theorem 3.8 (Raz [Raz09,RY09])**.** *Let $f \in \mathbb{F}[x_1,\ldots,x_{2n},\overline{z}]$ be a multilinear polynomial in the set of variables $\overline{x}$ and auxiliary variables $\overline{z}$. Let $f_{\overline{z}}$ denote the polynomial f in the ring $\mathbb{F}[\overline{z}][\overline{x}]$. Suppose that for any partition $[2n] = S \sqcup T$ with $|S| = |T| = n$ that*

$$\dim_{\mathbb{F}(\overline{x})} \mathbf{Coeff}_{\overline{x}|_S \left| \overline{x}|_T \right.} f_{\overline{z}} \geq 2^n ,$$

then f requires $\geq n^{\Omega(\log n)}$-size to be computed as a multilinear formula. For $d = o(\log n/\log\log n)$, f requires $n^{\Omega((n/\log n)^{1/d}/d^2)}$-size multilinear formulas of depth-$(2d+1)$.

3.5 Depth-3 Powering Formulas

In this section we review facts about depth-3 powering formulas. We begin with the *duality trick* of Saxena [Sax08], which shows that one can convert a power of a linear form to a sum of products of univariate polynomials.

Theorem 1 (Saxena's Duality Trick [SW01,Sax08,FSG13])**.** *Let $n \geq 1$, and $d \geq 0$. If $|\mathbb{F}| \geq nd+1$, then there are $\mathsf{poly}(n,d)$-explicit univariates $f_{i,j} \in \mathbb{F}[x_i]$ such that*

$$(x_1 + \cdots + x_n)^d = \sum_{i=1}^{s} f_{i,1}(x_1) \cdots f_{i,n}(x_n) \ ,$$

where $\deg f_{i,j} \leq d$ and $s = (nd+1)(d+1)$.

The original proof of Saxena [Sax08] only worked over field over large enough characteristic, and gave $s = nd + 1$. A similar version of this trick also appeared in Shpilka-Wigderson [SW01]. The parameters we use here are from the proof of Forbes, Gupta, and Shpilka [FSG13], which has the advantage of working over any large enough field.

Noting that the product $f_{i,1}(x_1) \cdots f_{i,n}(x_n)$ trivially has a width-1 roABP (in any variable order), it follows that $(x_1 + \cdots + x_n)^d$ has a $\mathsf{poly}(n,d)$-width roABP over a large enough field. Thus, size-s $\sum\bigwedge\sum$ formulas have $\mathsf{poly}(s)$-size roABPs over large enough fields by appealing to closure properties of roABPs (Theorem 3.5). As it turns out, this result also holds over any field as Forbes-Shpilka [FS13b] adapted Saxena's [Sax08] duality to work over any field. Their version works over any field, but loses the above clean form (sum of product of univariates).

Theorem 2 (Forbes-Shpilka [FS13b])**.** *Let $f \in \mathbb{F}[\overline{x}]$ be expressed as $f(\overline{x}) = \sum_{i=1}^{s}(\alpha_{i,0}+\alpha_{i,1}x_i+\cdots+\alpha_{i,n}x_n)^{d_i}$. Then f is computable by a width-r roABP in any variable order, where $r = \sum_i (d_i+1)$.*

One way to see this claim is to observe that for any variable partition, a linear function can be expressed as the sum of two variable disjoint linear functions $\ell(\overline{x}_1, \overline{x}_2) = \ell_1(\overline{x}_1) + \ell_2(\overline{x}_2)$. By the binomial theorem, the d-th power of this expression is a summation of $d+1$ variable disjoint products, which implies a coefficient dimension upper bound of $d+1$ (Theorem 3.2) and thus also a roABP-width upper bound (Theorem 3.4). One can then sum over the linear forms.

While this simulation suffices for obtaining roABP upper bounds, we will also want the clean form obtained via duality for application to multilinear-formula IPS proofs of the subset-sum axiom (Theorem 4.9).

3.6 Monomial Orders

We recall here the definition and properties of a *monomial order*, following Cox, Little and O'Shea [CLO07]. We first fix the definition of a *monomial* in our context.

Definition 3.5. *A **monomial** in $\mathbb{F}[x_1, \ldots, x_n]$ is a polynomial of the form $\overline{x}^{\overline{a}} = x_1^{a_1} \cdots x_n^{a_n}$ for $\overline{a} \in \mathbb{N}^n$.*

We will sometimes abuse notation and associate a monomial $\overline{x}^{\overline{a}}$ with its exponent vector $\overline{a}$, so that we can extend this order to the exponent vectors. Note that in this definition "1" is a monomial, and that scalar multiples of monomials such as $2x$ are not considered monomials. We now define a monomial order, which will be total order on monomials with certain natural properties.

Definition 3.6. *A **monomial ordering** is a total order $\prec$ on the monomials in $\mathbb{F}[\overline{x}]$ such that*

- *For all* $\overline{a} \in \mathbb{N}^n \setminus \{\overline{0}\}$, $1 \prec \overline{x}^{\overline{a}}$.
- *For all* $\overline{a}, \overline{b}, \overline{c} \in \mathbb{N}^n$, $\overline{x}^{\overline{a}} \prec \overline{x}^{\overline{b}}$ *implies* $\overline{x}^{\overline{a}+\overline{c}} \prec \overline{x}^{\overline{b}+\overline{c}}$.

For non-zero $f \in \mathbb{F}[\overline{x}]$, *the* ***leading monomial of*** f ***(with respect to a monomial order*** $\prec$***)***, *denoted* $\mathrm{LM}(f)$, *is the largest monomial in* $\mathrm{Supp}(f) := \{\overline{x}^{\overline{a}} : \mathrm{Coeff}_{\overline{x}^{\overline{a}}}(f) \neq 0\}$ *with respect to the monomial order* $\prec$. *The* ***trailing monomial of*** f, *denoted* $\mathrm{TM}(f)$, *is defined analogously to be the smallest monomial in* $\mathrm{Supp}(f)$. *The zero polynomial has neither leading nor trailing monomial.*

For non-zero $f \in \mathbb{F}[\overline{x}]$, *the* ***leading (resp. trailing) coefficient of*** f, *denoted* $\mathrm{LC}(f)$ *(resp.* $\mathrm{TC}(f)$*), is* $\mathrm{Coeff}_{\overline{x}^{\overline{a}}}(f)$ *where* $\overline{x}^{\overline{a}} = \mathrm{LM}(f)$ *(resp.* $\overline{x}^{\overline{a}} = \mathrm{TM}(f)$*).*

Henceforth in this paper we will assume $\mathbb{F}[\overline{x}]$ is equipped with some monomial order $\prec$. The results in this paper will hold for *any* monomial order. However, for concreteness, one can consider the lexicographic ordering on monomials, which is easily seen to be a monomial ordering (see also Cox, Little and O'Shea [CLO07]).

We begin with a simple lemma about how taking leading or trailing monomials (or coefficients) is homomorphic with respect to multiplication.

Lemma 3.9. *Let* $f, g \in \mathbb{F}[\overline{x}]$ *be non-zero so that* $fg \neq 0$. *Then the leading monomial and trailing monomials and coefficients are homomorphic with respect to multiplication, that is,* $\mathrm{LM}(fg) = \mathrm{LM}(f)\,\mathrm{LM}(g)$ *and* $\mathrm{TM}(fg) = \mathrm{TM}(f)\,\mathrm{TM}(g)$, *as well as* $\mathrm{LC}(fg) = \mathrm{LC}(f)\,\mathrm{LC}(g)$ *and* $\mathrm{TC}(fg) = \mathrm{TC}(f)\,\mathrm{TC}(g)$.

Proof: We do the proof for leading monomials and coefficients, the claim for trailing monomials and coefficients is symmetric.

Let $f(\overline{x}) = \sum_{\overline{a}} \alpha_{\overline{a}} \overline{x}^{\overline{a}}$ and $g(\overline{x}) = \sum_{\overline{b}} \beta_{\overline{b}} \overline{x}^{\overline{b}}$. Isolating the leading monomials,

$$f(\overline{x}) = \mathrm{LC}(f) \cdot \mathrm{LM}(f) + \sum_{\overline{x}^{\overline{a}} \succ \mathrm{LM}(f)} \alpha_{\overline{a}} \overline{x}^{\overline{a}}, \qquad g(\overline{x}) = \mathrm{LC}(g) \cdot \mathrm{LM}(g) + \sum_{\overline{x}^{\overline{b}} \succ \mathrm{LM}(g)} \beta_{\overline{b}} \overline{x}^{\overline{b}},$$

with $\mathrm{LC}(f) = \alpha_{\mathrm{LM}(f)}$ and $\mathrm{LC}(g) = \beta_{\mathrm{LM}(g)}$ being non-zero. Thus,

$$\begin{aligned} f(\overline{x})g(\overline{x}) = \mathrm{LC}(f)\,\mathrm{LC}(g) \cdot \mathrm{LM}(f)\,\mathrm{LM}(g) + \mathrm{LC}(f)\,\mathrm{LM}(f) \left(\sum_{\overline{x}^{\overline{b}} \succ \mathrm{LM}(g)} \beta_{\overline{b}} \overline{x}^{\overline{b}} \right) \\ + \mathrm{LC}(g)\,\mathrm{LM}(g) \left(\sum_{\overline{x}^{\overline{a}} \succ \mathrm{LM}(f)} \alpha_{\overline{a}} \overline{x}^{\overline{a}} \right) + \left(\sum_{\overline{x}^{\overline{a}} \succ \mathrm{LM}(f)} \alpha_{\overline{a}} \overline{x}^{\overline{a}} \right) \left(\sum_{\overline{x}^{\overline{b}} \succ \mathrm{LM}(g)} \beta_{\overline{b}} \overline{x}^{\overline{b}} \right). \end{aligned}$$

Using that $\overline{x}^{\overline{a}} \overline{x}^{\overline{b}} \succ \overline{x}^{\overline{a}}\,\mathrm{LM}(g)$, $\mathrm{LM}(f)\overline{x}^{\overline{b}} \succ \mathrm{LM}(f)\,\mathrm{LM}(g)$ shows that $\mathrm{LM}(f)\,\mathrm{LM}(g)$ is indeed the maximal monomial in the above expression with non-zero coefficient, and as its coefficient is $\mathrm{LC}(f)\,\mathrm{LC}(g)$. □

We now recall the well-known fact that for any set of polynomials the dimension of their span in $\mathbb{F}[\overline{x}]$ is equal to the number of distinct leading or trailing monomials in their span.

Lemma 3.10. *Let* $S \subseteq \mathbb{F}[\overline{x}]$ *be a set of polynomials. Then* $\dim \mathrm{span}\, S = |\mathrm{LM}(\mathrm{span}\, S)| = |\mathrm{TM}(\mathrm{span}\, S)|$. *In particular,* $\dim \mathrm{span}\, S \geq |\mathrm{LM}(S)|, |\mathrm{TM}(S)|$.

4 Upper Bounds for Linear-IPS

While the primary focus of this work is on *lower bounds* for restricted classes of the IPS proof system, we begin by discussing *upper bounds* to demonstrate that these restricted classes can prove the unsatisfiability of non-trivial systems of polynomials equations. This shows that obtaining lower bounds even for these restricted cases is non-trivial.

We begin by discussing the power of the linear-IPS proof system. While one of the most novel features of IPS proofs is their consideration of non-linear certificates, we show that in powerful enough models of algebraic computation, linear-IPS proofs can efficiently simulate IPS proofs. This result was obtained by Grochow and Pitassi [GP14] for the special case of sparse IPS. We then consider the *subset-sum* axioms, previously considered by Impagliazzo, Pudlák, and Sgall [IPS99], and show that they can be refuted in polynomial size by the $\mathcal{C}$-IPS$_{\text{LIN}}$ proof system where $\mathcal{C}$ is either the class of roABPs, or the class of multilinear formulas.

4.1 Simulating IPS Proofs with Linear-IPS

We show here that general IPS proofs can be efficiently simulated by linear-IPS, assuming that the axioms to be refuted are described by small algebraic circuits. Grochow and Pitassi [GP14] showed that whenever the IPS proof computes *sparse* polynomials, one can simulate it by linear-IPS using (possibly non-sparse) algebraic circuits. We give here a simulation of IPS when the proofs use general algebraic circuits.

We omit the boolean axioms in the below set of axioms as they do not play a role in this particular result.

Proposition 4.1. *Let $\mathbb{F}$ be a field with $|\mathbb{F}| \geq d+1$. Let $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ have an IPS proof $C \in \mathbb{F}[\overline{x}, y_1, \ldots, y_m]$, where $f_1, \ldots, f_m, C$ have degree-$(\leq d)$ and are computed by size-s algebraic circuits. Then $f_1, \ldots, f_m$ have a linear-IPS proof $C' \in \mathbb{F}[\overline{x}, \overline{y}]$ where $\deg C' \leq \mathsf{poly}(d)$ (and $\deg_{\overline{y}} C' \leq 1$), and C' is computable by a $\mathsf{poly}(d, s)$-size algebraic circuit.*

Proof: Express $C(\overline{x}, \overline{y})$ as a polynomial in $\mathbb{F}[\overline{x}][\overline{y}]$, so that $C(\overline{x}, \overline{y}) = \sum_{\overline{a} > \overline{0}} C_{\overline{a}}(\overline{x}) \overline{y}^{\overline{a}}$, where we use that $C(\overline{x}, \overline{0}) = 0$ to see that we can restrict $\overline{a} > \overline{0}$. Partitioning the $\overline{a} \in \mathbb{N}^n$ based on their first non-zero value, and denoting $\overline{a}_{<i}$ for the first $i-1$ coordinates of $\overline{a}$, we obtain

$$\begin{aligned} C(\overline{x}, \overline{y}) &= \sum_{\overline{a} > \overline{0}} C_{\overline{a}}(\overline{x}) \overline{y}^{\overline{a}} \\ &= \sum_{i=1}^{n} \sum_{\substack{\overline{a} : \overline{a}_{<i} = \overline{0}, \\ a_i > 0}} C_{\overline{a}}(\overline{x}) \overline{y}^{\overline{a}} \end{aligned}$$

Now define $C_i(\overline{x}, \overline{y}) := \sum_{\substack{\overline{a} : \overline{a}_{<i} = \overline{0}, \\ a_i > 0}} C_{\overline{a}}(\overline{x}) \overline{y}^{\overline{a} - \overline{e}_i}$, where $\overline{e}_i$ is the i-th standard basis vector. Note that this is a valid polynomial as in this summation we assume $a_i > 0$ so that $\overline{a} - \overline{e}_i \geq 0$,

$$= \sum_{i=1}^{n} C_i(\overline{x}, \overline{y}) y_i \ .$$

We now claim that $C'(\overline{x}, \overline{y}) := \sum_{i=1}^{n} C_i(\overline{x}, \overline{f}(\overline{x})) y_i$ is the desired linear-IPS refutation, where we have partially substituted in the f_i for the y_i. First, observe that it is a valid refutation, as $C'(\overline{x}, \overline{0}) = \sum_{i=1}^{n} C_i(\overline{x}, \overline{f}(\overline{x}))0 = 0$, and $C'(\overline{x}, \overline{f}(\overline{x})) = \sum_{i=1}^{n} C_i(\overline{x}, \overline{f}(\overline{x})) f_i(\overline{x}) = C(\overline{x}, \overline{f}(\overline{x})) = 1$ via the above definition.

We now argue that C' has a small circuit, for which it is enough to show the size claim for each C_i. First, note that

$$C_i(\overline{x},\overline{y})y_i = \sum_{\substack{\overline{a}:\overline{a}_{<i}=\overline{0},\\ a_i>0}} C_{\overline{a}}(\overline{x})\overline{y}^{\overline{a}} = C(\overline{x},\overline{0},y_i,\overline{y}_{>i}) - C(\overline{x},\overline{0},0,\overline{y}_{>i}) \ ,$$

where each "$\overline{0}$" here is a vector of $i-1$ zeros. Clearly each of $C(\overline{x},\overline{0},y_i,\overline{y}_{>i})$ and $C(\overline{x},\overline{0},0,\overline{y}_{>i})$ have size-s circuits, and it remains to show that $\frac{1}{y_i}(C(\overline{x},\overline{0},y_i,\overline{y}_{>i}) - C(\overline{x},\overline{0},0,\overline{y}_{>i}))$ (which is a polynomial) has a small circuit. A heavy-handed approach would be to use Strassen's [Str73] elimination of divisions. Another approach is to interpolation (or more generally, homogenization, see Shpilka and Yehudayoff [SY10]). That is, view $D(y_i) := C(\overline{x},\overline{0},y_i,\overline{y}_{>i}) - C(\overline{x},\overline{0},0,\overline{y}_{>i})$ as a univariate polynomial in the ring $\mathbb{F}[\overline{x},\overline{y}_{>i}][y_i]$, which thus has degree $\leq d$. By evaluating $D(y_i)$ at $d+1$ distinct points in $\mathbb{F}$, one can for any j take suitable linear combinations to obtain the coefficient of y_i^j in $D(y_i)$. By multiplying this coefficient by y_i^{j-1} and summing over j, one obtains $D(y_i)/y_i$. As this takes $\mathsf{poly}(d)$ evaluations of a size-s circuit one obtains that $C_i(\overline{x},\overline{y})$ has $\mathsf{poly}(s,d)$ size. As $\overline{f}$ also has such a size, it follows that $C_i(\overline{x},\overline{f}(\overline{x}))$ has $\mathsf{poly}(s,d)$ size as desired. □

We remark that this simulation also roughly preserves the *depth* of the IPS proof, assuming the axioms themselves also are computable by low-depth circuits. One can also remove the need for a large field by using homogenization, but this increases the depth complexity (see Shpilka and Yehudayoff [SY10]).

IT: I think it should be noted here that since the simulation depends on the degree d of C, then it doesn't follow from the theorem that linear-IPS polynomially simulates Extended Frege (while IPS does simulate Extended Frege).

4.2 Multilinearizing roABP-IPS$_{\text{LIN}}$

In this section we show that for roABP-IPS$_{\text{LIN}}$ and multilinear-formula-IPS$_{\text{LIN}}$ one can efficiently prove that a refutation equals its multilinearization modulo the boolean axioms. That is, for an axiom f and polynomial g we wish to prove that $g \cdot f \equiv \mathrm{ml}(g \cdot f) \mod \overline{x}^2 - \overline{x}$ by expressing $g \cdot f = \sum_i h_i \cdot (x_i^2 - x_i)$ where the h_i have small circuits. We will use this multilinearization in our construction of IPS refutations of the subset-sum axiom (Subsection 4.4).

We begin by noting that multilinearization for these two circuit classes is rather special, as these classes are both not too weak and not too strong. That is, some circuit classes are simply too weak to compute their multilinearizations. An example is the class of depth-3 powering formulas, where $(x_1 + \cdots + x_n)^n$ has a small $\sum\bigwedge\sum$ formula, but is multilinearization has leading term $n!x_1 \cdots x_n$ and thus requires exponential size as a $\sum\bigwedge\sum$ formula (by appealing to 3). On the other hand, some circuit classes are too strong to admit efficient multilinearization (under plausible complexity assumptions). That is, consider $f(X,\overline{y}) := (x_{1,1}y_1 + \cdots + x_{1,n}y_n) \cdots (x_{n,1}y_1 + \cdots + x_{n,n}y_n)$, which is clearly a simple depth-3 ($\prod\sum\prod$) circuit. Viewing this polynomial in $\mathbb{F}[X][\overline{y}]$ where X is an $n \times n$ matrix, one sees that $\mathrm{Coeff}_{X|y_1\cdots y_n} f = \mathrm{perm}(X)$, where $\mathrm{perm}(X)$ is the $n \times n$ permanent. Viewing $\mathrm{ml}(f)$, the multilinearization of f, in $\mathbb{F}[X][\overline{y}]$ one sees that $\mathrm{ml}(f)$ is of degree n and its degree n component is the coefficient of $y_1 \cdots y_n$ in $\mathrm{ml}(f)$, which is still $\mathrm{perm}(X)$. Hence, by interpolation, one can extract this degree n part and thus can compute a circuit for $\mathrm{perm}(X)$ given a circuit for $\mathrm{ml}(f)$. Since we believe $\mathrm{perm}(X)$ does not have small algebraic circuits it follows that the multilinearization of f does not have small circuits.

We begin by multilinearizing roABPs, where we multilinearize variable by variable via telescoping.

Proposition 4.2. *Let $f \in \mathbb{F}[x_1,\ldots,x_n]$ be computable by a width-r roABP in variable order $x_1 < \cdots < x_n$, so that $f(\overline{x}) = (\prod_{i=1}^n A_i(x_i))_{1,1}$ where $A_i \in \mathbb{F}[x_i]^{r\times r}$ have $\deg A_i \leq d$. Then* $\mathrm{ml}(f)$

has a width-r roABP in variable order $x_1 < \cdots < x_n$, and there are $\mathsf{poly}(r,n,d)$-explicit width-r roABPs $f_1,\ldots,f_n \in \mathbb{F}[\overline{x}]$ in variable order $x_1 < \cdots < x_n$ such that

$$f(\overline{x}) = \mathrm{ml}(f) + \sum_{i=1}^{n} f_i \cdot (x_i^2 - x_i) \ .$$

Further, the individual degree of the roABP for $\mathrm{ml}(f)$ is ≤ 1.

Proof: First observe that we can extend the multilinearization map $\mathrm{ml} : \mathbb{F}[\overline{x}] \to \mathbb{F}[\overline{x}]$ to matrices $\mathrm{ml} : \mathbb{F}[\overline{x}]^{r\times r} \to \mathbb{F}[\overline{x}]^{r\times r}$ by applying the map entry-wise. It follows then that $A_i(x_i) - \mathrm{ml}(A_i(x_i)) \equiv 0 \mod x_i^2 - x_i$, so that $A_i(x_i) - \mathrm{ml}(A_i(x_i)) = B(x_i)\cdot(x_i^2 - x_i)$ for some $B_i(x_i) \in \mathbb{F}[x_i]^{r\times r}$. Now define $\mathrm{ml}_{\leq i}$ be the map which multilinearizes the first i variables and leaves intact the others, so that $\mathrm{ml}_{\leq 0}$ is the identity map and $\mathrm{ml}_{\leq n} = \mathrm{ml}$. Telescoping,

$$\prod_{i=1}^{n} A_i(x_i) = \mathrm{ml}_{\leq n}\left(\prod_{i=1}^{n} A_i(x_i)\right) + \sum_{j=1}^{n}\left[\mathrm{ml}_{<j}\left(\prod_{i=1}^{n} A_i(x_i)\right) - \mathrm{ml}_{\leq j}\left(\prod_{i=1}^{n} A_i(x_i)\right)\right]$$

using that $\mathrm{ml}(gh) = \mathrm{ml}(\mathrm{ml}(g)\,\mathrm{ml}(h))$ (Theorem 3.7), even for these partial-multilinearization maps,

$$\begin{aligned} &= \mathrm{ml}_{\leq n}\left(\prod_{i=1}^{n} A_i(x_i)\right) + \sum_{j=1}^{n}\left[\mathrm{ml}_{<j}\left(\prod_{i<j}\mathrm{ml}_{<j}(A_i(x_i))\prod_{i\geq j} A_i(x_i)\right)\right. \\ &\qquad\qquad \left. - \mathrm{ml}_{\leq j}\left(\prod_{i\leq j}\mathrm{ml}_{\leq j}(A_i(x_i))\prod_{i>j} A_i(x_i)\right)\right] \end{aligned}$$

dropping the outside $\mathrm{ml}_{<j}$ and $\mathrm{ml}_{\leq j}$ as the inside polynomials are now multilinear in the appropriate variables, and replacing them with ml as appropriate,

$$\begin{aligned} &= \prod_{i=1}^{n}\mathrm{ml}(A_i(x_i)) + \sum_{j=1}^{n}\left[\prod_{i<j}\mathrm{ml}(A_i(x_i))\prod_{i\geq j} A_i(x_i)\right. \\ &\qquad\qquad \left. - \prod_{i\leq j}\mathrm{ml}(A_i(x_i))\prod_{i>j} A_i(x_i)\right] \\ &= \prod_{i=1}^{n}\mathrm{ml}(A_i(x_i)) + \sum_{j=1}^{n}\prod_{i<j}\mathrm{ml}(A_i(x_i))\Big(A_j(x_j) - \mathrm{ml}(A_j(x_j))\Big)\prod_{i>j} A_i(x_i) \\ &= \prod_{i=1}^{n}\mathrm{ml}(A_i(x_i)) + \sum_{j=1}^{n}\prod_{i<j}\mathrm{ml}(A_i(x_i))B_j(x_j)\prod_{i>j} A_i(x_i)\cdot(x_j^2 - x_j) \ . \end{aligned}$$

Taking the $(1,1)$-entry in the above yields that

$$\begin{aligned} f(\overline{x}) &= \left(\prod_{i=1}^{n} A_i(x_i)\right)_{1,1} \\ &= \left(\prod_{i=1}^{n}\mathrm{ml}(A_i(x_i))\right)_{1,1} + \sum_{j=1}^{n}\left(\prod_{i<j}\mathrm{ml}(A_i(x_i))B_j(x_j)\prod_{i>j} A_i(x_i)\right)_{1,1}\cdot(x_j^2 - x_j) \ . \end{aligned}$$

Thus, define $f_i := \left(\prod_{i<j}\mathrm{ml}(A_i(x_i))B_j(x_j)\prod_{i>j} A_i(x_i)\right)_{1,1}$ and define $f' := (\prod_{i=1}^{n}\mathrm{ml}(A_i(x_i)))_{1,1}$, which is an roABP of individual degree 1 as each $\mathrm{ml}(A_i(x_i))$ is linear. As the above yields that $f = f' + \sum_j f_j\cdot(x_j^2 - x_j)$ and f' is multilinear, it follows that $\mathrm{ml}(f) = f'$ and that this is the desired expression. □

4.3 Multilinear-Formula-IPS

We now turn to multilinear-formula-IPS, with the aim of showing that this efficiently simulates sparse-IPS$_{\text{LIN}}$. We begin by noting that *linear*-IPS over multilinear polynomials is not a complete proof system (for the language of all systems of polynomial equations with no 0-1 roots).

Example: Consider the unsatisfiable system of equations $xy+1, x^2-x, y^2-y$. A multilinear linear-IPS proof is a tuple of multilinear polynomials $(f,g,h) \in \mathbb{F}[x,y]$ such that $f\cdot(xy+1)+g\cdot(x^2-x)+h\cdot(y^2-y)=1$. In particular, $f(x,y)=\frac{1}{xy+1}$ for $x,y\in\{0,1\}$, which implies by interpolation over the boolean cube that $f(x,y)=1\cdot(1-x)(1-y)+\frac{1}{2}\cdot(1-x)y+\frac{1}{2}\cdot x(1-y)+\frac{1}{3}\cdot xy = 1-\frac{1}{2}\cdot x-\frac{1}{2}\cdot y+\frac{1}{3}\cdot xy$. Thus $f\cdot(xy+1)$ contains the monomial x^2y^2. However, as g,h are multilinear we see that x^2y^2 cannot appear in $g\cdot(x^2-x)+h\cdot(y^2-y)-1$, so that the equality $f\cdot(xy+1)+g\cdot(x^2-x)+h\cdot(y^2-y)=1$ cannot hold.

As such, to simulate sparse-IPS$_{\text{LIN}}$ (a complete proof system) we must resort to using *general* IPS. We first show how to multilinearize a monomial.

Lemma 4.3. *Let* $\overline{x}^{\overline{1}}=\prod_{i=1}^n x_i$. *Then,*

$$(\overline{x}^{\overline{1}})^2-\overline{x}^{\overline{1}}=\sum_{\overline{0}<\overline{a}\leq\overline{1}}\prod_{a_i=1}(x_i^2-x_i)\prod_{a_i=0}x_i\ .$$

Proof:

$$\begin{aligned}(\overline{x}^{\overline{1}})^2-\overline{x}^{\overline{1}}&=\prod_{i=1}^n((x_i^2-x_i)+x_i)-\prod_{i=1}^n x_i\\&=\sum_{\overline{0}\leq\overline{a}\leq\overline{1}}\prod_{a_i=1}(x_i^2-x_i)\prod_{a_i=0}x_i-\prod_{i=1}^n x_i\\&=\sum_{\overline{0}<\overline{a}\leq\overline{1}}\prod_{a_i=1}(x_i^2-x_i)\prod_{a_i=0}x_i\ .\end{aligned}$$

□

We now give an IPS proof for showing how a polynomial times a monomial equals its multilinearization.

Lemma 4.4. *Let* $f\in\mathbb{F}[\overline{x},y_1,\ldots,y_d]$ *be multilinear be expressed as* $f=\sum_{0\leq\overline{a}\leq\overline{1}}f_{\overline{a}}(\overline{x})\overline{y}^{\overline{a}}$ *in the ring* $\mathbb{F}[\overline{x}][\overline{y}]$. *Then*

$$f(\overline{x},\overline{y})\cdot\overline{y}^{\overline{1}}-\mathrm{ml}(f(\overline{x},\overline{y})\cdot\overline{y}^{\overline{1}})=C(\overline{x},\overline{y},\overline{y}^2-\overline{y})\ ,$$

where $C\in\mathbb{F}[\overline{x},\overline{y},z_1,\ldots,z_d]$ *is defined by* $C(\overline{x},\overline{y},\overline{z}):=\sum_{\overline{a}}f_{\overline{a}}(\overline{x})\sum_{\overline{0}<\overline{b}\leq\overline{a}}\prod_{b_i=1}z_i\prod_{b_i=0}y_i$, *so that* $C(\overline{x},\overline{y},\overline{0})=0$.

Proof: Simple calculation yields

$$\begin{aligned}f(\overline{x},\overline{y})\cdot\overline{y}^{\overline{1}}-\mathrm{ml}(f(\overline{x},\overline{y})\cdot\overline{y}^{\overline{1}})&=\sum_{\overline{a}}f_{\overline{a}}(\overline{x})\overline{y}^{\overline{a}}\cdot\overline{y}^{\overline{1}}-\mathrm{ml}\left(\sum_{\overline{a}}f_{\overline{a}}(\overline{x})\overline{y}^{\overline{a}}\cdot\overline{y}^{\overline{1}}\right)\\&=\sum_{\overline{a}}f_{\overline{a}}(\overline{x})\left(\overline{y}^{\overline{a}+\overline{1}}-\overline{y}^{\overline{1}}\right)\end{aligned}$$

appealing to Theorem 4.3,

$$=\sum_{\overline{a}}f_{\overline{a}}(\overline{x})\sum_{\overline{0}<\overline{b}\leq\overline{a}}\prod_{b_i=1}(y_i^2-y_i)\prod_{b_i=0}y_i$$

$$= C(\overline{x}, \overline{y}, \overline{y}^2 - \overline{y}) \ .$$

That $C(\overline{x}, \overline{y}, \overline{0}) = 0$ is clear. □

We now extend the above to a polynomial times a sparse polynomial, keeping track of the complexity of this IPS proof.

Corollary 4.5. *Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be computable by a multilinear formula of size-s, and let $g \in \mathbb{F}[\overline{x}]$ be a t-sparse multilinear polynomial with $\deg g \leq d$. Then*

$$g \cdot f - \mathrm{ml}(g \cdot f) = C(\overline{x}, \overline{x}^2 - \overline{x}) \ ,$$

for $C \in \mathbb{F}[\overline{x}, u_1, \ldots, u_n]$ where $C(\overline{x}, \overline{0}) = 0$ and C is computable by a $\mathsf{poly}(t, 2^d, n, s)$-size multilinear formula. Further, if f is actually depth-2 then C is computable in $\mathsf{poly}(t, 2^d, n, s)$-size and depth-2.

Proof: As ml is a linear map, it suffices to show the claim for $t = 1$ where $g = \prod_{i \in S}$ for some $S \subseteq [n]$ with $|S| \leq d$. Partition $\overline{x} = (\overline{y}, \overline{z})$ so $\overline{y} = \overline{x}|_{[n]\setminus S}$ and $\overline{z} = \overline{x}|_S$ so that $g = \overline{z}^{\overline{1}}$, and thus $\overline{z}$ has $\leq d$ variables. Express f in $\mathbb{F}[\overline{y}][\overline{z}]$ as $f = \sum_{\overline{0} \leq \overline{a} \leq \overline{1}} f_{\overline{a}}(\overline{y}) \overline{z}^{\overline{a}}$. By Theorem 4.4,

$$\begin{aligned} g \cdot f - \mathrm{ml}(g \cdot f) &= \overline{z}^{\overline{1}} \cdot f - \mathrm{ml}(\overline{z}^{\overline{1}} \cdot f) \\ &= \sum_{\overline{a}} f_{\overline{a}}(\overline{y}) \sum_{\overline{0} < \overline{b} \leq \overline{a}} \prod_{b_i = 1} (z_i^2 - z_i) \prod_{b_i = 0} z_i \\ &= C(\overline{y}, \overline{z}, \overline{z}^2 - \overline{z}) \ , \end{aligned}$$

where $C(\overline{y}, \overline{z}, \overline{w}) = \sum_{\overline{a}} f_{\overline{a}}(\overline{y}) \sum_{\overline{0} < \overline{b} \leq \overline{a}} \prod_{b_i = 1} w_i \prod_{b_i = 0} z_i$. We now claim that C has a $\mathsf{poly}(2^d, n, s)$-size multilinear formula. First notice that C is indeed multilinear as each $f_{\overline{a}}$ is multilinear Next notice that each $f_{\overline{a}}(\overline{y})$ has a $\mathsf{poly}(2^d, n, s)$-size formula as it can be computed by interpolating $f(\overline{y}, \overline{z})$ over $\overline{z} \in \{0,1\}^{|\overline{z}|}$, so that $f_{\overline{a}}(\overline{y}) = \sum_{\overline{\alpha} \in \{0,1\}^{|\overline{z}|}} \beta_{\overline{\alpha}} f(\overline{y}, \overline{\alpha})$ for some $\beta_{\overline{a}} \in \mathbb{F}$. Clearly $\sum_{\overline{0} < \overline{b} \leq \overline{a}} \prod_{b_i = 1} w_i \prod_{b_i = 0} z_i$ has the requisite size formula. Finally, summing over all $\overline{a}$ only multiplies the complexity by at most 2^d. The claim about sparsity (depth-2 computation) follows by inspection of the above process, noting that we can push all multiplication gates to the bottom of the above computation for C. □

We now conclude by showing that multilinear-formula-$\mathrm{IPS}_{\mathrm{LIN}'}$ can efficiently simulate sparse-$\mathrm{IPS}_{\mathrm{LIN}}$ when the axioms are low-degree. As this latter system is complete, this shows the former is as well. That is, we allow the refutation to depend non-linearly on the boolean axioms, but only linearly on the other axioms.

Corollary 4.6. *Let $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ be degree $\leq d$ multilinear polynomials unsatisfiable over the boolean cube. Suppose that there are s-sparse polynomials $g_1, \ldots, g_m, h_1, \ldots, h_n \in \mathbb{F}[\overline{x}]$ such that $\sum_{j=1}^m g_j f_j + \sum_{i=1}^n h_i \cdot (x_i^2 - x_i) = 1$. Then $\overline{f}, \overline{x}^2 - \overline{x}$ can be refuted by a depth-2 multilinear-formula-$\mathrm{IPS}_{\mathrm{LIN}'}$ proof of size $\mathsf{poly}(2^d, n, s, m)$.*

Proof: Clearly $\sum_{j=1}^m g_j f_j \equiv 1 \mod \overline{x}^2 - \overline{x}$, and thus by multilinearizing we also have the equation $\sum_{j=1}^m \mathrm{ml}(g_j) f_j \equiv 1 \mod \overline{x}^2 - \overline{x}$, where we used that the f_j are already multilinear. Now note that each $\mathrm{ml}(g_j)$ is an s-sparse multilinear polynomial of degree $\leq d$, and thus is computable by a $\mathsf{poly}(n, s)$-size multilinear formula. By Theorem 4.5, it follows for each j that $\mathrm{ml}(g_j) f_j - \mathrm{ml}(\mathrm{ml}(g_j) f_j) = C_j(\overline{x}, \overline{x}^2 - \overline{x})$, where $C_j(\overline{x}, \overline{0}) = 0$ and C_j is computable by a $\mathsf{poly}(2^d, n, s)$-size depth-2 multilinear formula, as each g_j is sparse. Summing over j, it follows that

$$\sum_j C_j(\overline{x}, \overline{x}^2 - \overline{x}) = \sum_{j=1}^m (\mathrm{ml}(g_j) f_j - \mathrm{ml}(\mathrm{ml}(g_j) f_j))$$

$$= \sum_{j=1}^{m} \mathrm{ml}(g_j) f_j - \mathrm{ml}\left(\sum_{j=1}^{m} \mathrm{ml}(g_j) f_j\right)$$
$$= \sum_{j=1}^{m} \mathrm{ml}(g_j) f_j - 1 \ ,$$

where in the last step we used that $\sum_{j=1}^{m} \mathrm{ml}(g_j) f_j \equiv 1 \mod \overline{x}^2 - \overline{x}$. Thus, defining the refutation $C(\overline{x}, y_1 \ldots, y_m, z_1 \ldots, z_n)$ by $\sum_j f_j(\overline{x}) \cdot y_j - \sum_j C_j(\overline{x}, \overline{z})$ we see that $C(\overline{x}, \overline{0}, \overline{0}) = 0$, $C(\overline{x}, \overline{f}, \overline{x}^2 - \overline{x}^2) = 1$, and C has a $\mathsf{poly}(2^d, n, s, m)$-size multilinear formula. This formula is even depth-2 as each C_j is depth-2, and $f_j(\overline{x}) \cdot y_j$ is depth-2 by pushing y_j to the bottom multiplication gate. □

4.4 Refutations of the Subset-Sum Axiom

We now give efficient IPS refutations of the subset-sum axiom, where these IPS refutations can be even placed in the restricted roABP-IPS$_{\mathrm{LIN}}$ or multilinear-formula-IPS$_{\mathrm{LIN}}$ subclasses. That is, we give such refutations for whenever the polynomial $\sum_i \alpha_i x_i - \beta$ is unsatisfiable over the boolean cube $\{0,1\}^n$, where the size of the refutation is polynomial in the size of the set $A := \{\sum_i \alpha_i x_i : \overline{x} \in \{0,1\}^n\}$. The most natural example is when $\overline{\alpha} = \overline{1}$ so that $A = \{0, \ldots, n\}$.

To construct these refutations, we first show that there is an efficiently computable polynomial f such that $f(\overline{x}) \cdot (\sum_i \alpha_i x_i - \beta) \equiv 1 \mod \overline{x}^2 - \overline{x}$. This will be done by considering the univariate polynomial $p_A(t) := \prod_{\alpha \in A}(t - \alpha)$. As for any univariate $p(x)$ we have that $x - y$ divides $p(x) - p(y)$, so that $p_A(\sum_i \alpha_i x_i) - p_A(\beta)$ is a multiple of $\sum_i \alpha_i x_i - \beta$. As $\sum_i \alpha_i x_i - \beta$ is unsatisfiable it must be that $\beta \notin A$. This implies that $p_A(\sum_i \alpha_i x_i) \equiv 0 \mod \overline{x}^2 - \overline{x}$ while $p_A(\beta) \neq 0$. Consequently, $p_A(\sum_i \alpha_i x_i) - p_A(\beta)$ is equivalent to a non-zero constant modulo $\overline{x}^2 - \overline{x}$, yielding the Nullstellensatz refutation

$$\frac{1}{-p_A(\beta)} \cdot \frac{p_A(\sum_i \alpha_i x_i) - p_A(\beta)}{\sum_i \alpha_i x_i - \beta} \cdot (\textstyle\sum_i \alpha_i x_i - \beta) \equiv 1 \mod \overline{x}^2 - \overline{x} \ .$$

By understanding the quotient $\frac{p_A(\sum_i \alpha_i x_i) - p_A(\beta)}{\sum_i \alpha_i x_i - \beta}$ we see that it can be efficiently computable as a small $\sum \bigwedge \sum$ formula and thus a roABP, and we can then multilinearize this roABP (Theorem 4.2). Over large fields, we can also convert the quotient to a sum of products of univariate linear forms using duality (1) and multilinearization.

Proposition 4.7. *Let $\overline{\alpha} \in \mathbb{F}^n$, $\beta \in \mathbb{F}$ and $A := \{\sum_{i=1}^n \alpha_i x_i : \overline{x} \in \{0,1\}^n\}$ be so that $\beta \notin A$. Then there is a multilinear polynomial $f \in \mathbb{F}[\overline{x}]$ such that*

$$f(\overline{x}) \cdot (\textstyle\sum_i \alpha_i x_i - \beta) \equiv 1 \mod \overline{x}^2 - \overline{x} \ .$$

For any $|\mathbb{F}|$, f is computable by a $\mathsf{poly}(|A|, n)$-explicit $\mathsf{poly}(|A|, n)$-width roABP of individual degree ≤ 1.

If $|\mathbb{F}| \geq \mathsf{poly}(|A|, n)$, then f is computable as

$$f(\overline{x}) = \sum_{i=1}^{s} f_{i,1}(x_1) \cdots f_{i,n}(x_n) \ ,$$

where each $f_{i,j} \in \mathbb{F}[x_i]$ has $\deg f_{i,j} \leq 1$, $s \leq \mathsf{poly}(|A|, n)$, and this expression is $\mathsf{poly}(|A|, n)$-explicit.

Proof: defining f: Define $p_A(t) \in \mathbb{F}[t]$ by $p_A := \prod_{\alpha \in A}(t - \alpha)$, so that $p_A(A) = 0$ and $p_A(\beta) \neq 0$. Express $p_A(t)$ in its monomial representation as $p_A(t) = \sum_{k=0}^{|A|} \gamma_k t^k$. Then observe that

$$p_A(t) - p_A(\beta) = \left(\sum_{k=0}^{|A|} \gamma_k \frac{t^k - \beta^k}{t - \beta}\right)(t - \beta)$$

$$= \left(\sum_{k=0}^{|A|} \gamma_k \sum_{j=0}^{k-1} t^j \beta^{(k-1)-j} \right) (t-\beta)$$
$$= \left(\sum_{j=0}^{|A|-1} \left(\sum_{k=j+1}^{|A|} \gamma_k \beta^{(k-1)-j} \right) t^j \right) (t-\beta) \, .$$

Thus, plugging in $t \leftarrow \sum_i \alpha_i x_i$, we can define the polynomial $g(\overline{x}) \in \mathbb{F}[\overline{x}]$ by

$$g(\overline{x}) := \frac{p_A(\sum_i \alpha_i x_i) - p_A(\beta)}{\sum_i \alpha_i x_i - \beta}$$
$$= \sum_{j=0}^{|A|-1} \left(\sum_{k=j+1}^{|A|} \gamma_k \beta^{(k-1)-j} \right) \left(\sum_i \alpha_i x_i \right)^j \, . \tag{1}$$

Hence,

$$g(\overline{x})(\textstyle\sum_i \alpha_i x_i - \beta) = p_A(\sum_i \alpha_i x_i) - p_A(\beta) \, . \tag{2}$$

For any $\overline{x} \in \{0,1\}^n$ we have that $\sum_i \alpha_i x_i \in A$. As $p_A(A) = 0$ it follows that $p_A(\sum_i \alpha_i x_i) = 0$ for all $\overline{x} \in \{0,1\}^n$. This implies that $p_A(\sum_i \alpha_i x_i) \equiv 0 \mod \overline{x}^2 - \overline{x}$, yielding

$$g(\overline{x})(\textstyle\sum_i \alpha_i x_i - \beta) \equiv -p_A(\beta) \mod \overline{x}^2 - \overline{x} \, .$$

As $-p_A(\beta) \in \mathbb{F} \setminus \{0\}$, we have that

$$\frac{1}{-p_A(\beta)} \cdot g(\overline{x}) \cdot (\textstyle\sum_i \alpha_i x_i - \beta) \equiv 1 \mod \overline{x}^2 - \overline{x} \, .$$

We now simply multilinearize, and thus define the multilinear polynomial $f(\overline{x}) := \mathrm{ml}(\frac{1}{-p_A(\beta)} \cdot g(\overline{x}))$. First, we see that this has the desired form, using the interaction of multilinearization and multiplication (Theorem 3.7).

$$1 = \mathrm{ml}\left(\frac{1}{-p_A(\beta)} g(\overline{x}) \cdot (\textstyle\sum_i \alpha_i x_i - \beta) \right)$$
$$= \mathrm{ml}\left(\mathrm{ml}\left(\frac{1}{-p_A(\beta)} \cdot g(\overline{x}) \right) \mathrm{ml}(\textstyle\sum_i \alpha_i x_i - \beta) \right)$$
$$= \mathrm{ml}\left(f \cdot \mathrm{ml}(\textstyle\sum_i \alpha_i x_i - \beta) \right)$$
$$= \mathrm{ml}\left(f \cdot (\textstyle\sum_i \alpha_i x_i - \beta) \right)$$

Thus, $f \cdot (\sum_i \alpha_i x_i - \beta) \equiv 1 \mod \overline{x}^2 - \overline{x}$ as desired.

computing f as a roABP: By Equation 1 we see that $g(\overline{x})$ is computable by a $\mathsf{poly}(|A|, n)$-size $\sum \bigwedge \sum$-formula, and by 2 $g(\overline{x})$ and thus $\frac{1}{-p_A(\beta)} \cdot g(\overline{x})$ are computable by $\mathsf{poly}(|A|, n)$-width roABPs. Noting that roABPs can be efficiently multilinearized (Theorem 4.2) we see that $f = \mathrm{ml}(\frac{1}{-p_A(\beta)} \cdot g(\overline{x}))$ can thus be computed by such a roABP also, where the individual degree of this roABP is at most 1.

computing f via duality: We apply duality (1) to see that over large enough fields there are univariates $g_{j,\ell,i}$ where

$$g(\overline{x}) = \sum_{j=0}^{|A|-1} \left(\sum_{k=j+1}^{|A|} \gamma_k \beta^{(k-1)-j} \right) \left(\sum_i \alpha_i x_i \right)^j$$

$$= \sum_{j=0}^{|A|-1} \left(\sum_{k=j+1}^{|A|} \gamma_k \beta^{(k-1)-j} \right) \sum_{\ell=1}^{(nj+1)(j+1)} g_{j,\ell,1}(x_1) \cdots g_{j,\ell,n}(x_n)$$

Absorbing the constant $\left(\sum_{k=j+1}^{|A|} \gamma_k \beta^{(k-1)-j}\right)$ into these univariates and re-indexing,

$$= \sum_{i=1}^{s} g_{i,1}(x_1) \cdots g_{i,n}(x_n)$$

for some univariates $g_{i,j}$, where $s \leq |A|(n|A|+1)(|A|+1) = \mathsf{poly}(|A|, n)$.

We now obtain f by multilinearizing the above expression, again appealing to multilinearization (Theorem 3.7).

$$\begin{aligned} f &= \mathrm{ml}\left(\frac{1}{-p_A(\beta)} g(\overline{x})\right) \\ &= \mathrm{ml}\left(\frac{1}{-p_A(\beta)} \sum_{i=1}^{s} g_{i,1}(x_1) \cdots g_{i,n}(x_n)\right) \end{aligned}$$

absorbing the constant $1/-p_A(\beta)$ and renaming,

$$\begin{aligned} &= \mathrm{ml}\left(\sum_{i=1}^{s} g'_{i,1}(x_1) \cdots g'_{i,n}(x_n)\right) \\ &= \mathrm{ml}\left(\sum_{i=1}^{s} \mathrm{ml}(g'_{i,1}(x_1)) \cdots \mathrm{ml}(g'_{i,n}(x_n))\right) \end{aligned}$$

defining $f_{i,j}(x_j) := \mathrm{ml}(g_{i,j}(x_j))$, so that $\deg f_{i,j} \leq 1$,

$$= \mathrm{ml}\left(\sum_{i=1}^{s} f_{i,1}(x_1) \cdots f_{i,n}(x_n)\right)$$

and we can drop the outside ml as this expression is now multilinear,

$$= \sum_{i=1}^{s} f_{i,1}(x_1) \cdots f_{i,n}(x_n) \ ,$$

showing that f is computable as desired. □

Note that computing f via duality also implies a roABP for f, but only in large enough fields $|\mathbb{F}| \geq \mathsf{poly}(|A|, n)$. Of course, $\mathbb{F}$ must have $|\mathbb{F}| \geq |A|$ at least, but by using the field-independent conversion of $\sum\bigwedge\sum$ to roABP (2) this shows that $\mathbb{F}$ need not be any larger than A for the refutation to be efficient.

The above shows that one can give an "IPS proof" $g(\overline{x})(\sum_i \alpha_i x_i - \beta) + \sum_i h_i(\overline{x})(x_i^2 - x_i) = 1$, where g is efficiently computable. However, this is not yet an IPS proof as it does not bound the complexity of the h_i. We now extend this to an actual IPS proof by using the above multilinearization results (Theorem 4.2). Note that while the above result gives a small $\sum\bigwedge\sum$ formula g such that $g \cdot (\sum_i \alpha_i x_i - \beta) \equiv -p_A(\beta) \mod \overline{x}^2 - \overline{x}$ for non-zero scalar $-p_A(\beta)$, this does not translate to a $\sum\bigwedge\sum$-IPS refutation as $\sum\bigwedge\sum$ formulas cannot be multilinearized efficiently (see the discussion in Subsection 4.2).

Corollary 4.8. *Let $\overline{\alpha} \in \mathbb{F}^n$, $\beta \in \mathbb{F}$ and $A := \{\sum_{i=1}^n \alpha_i x_i : \overline{x} \in \{0,1\}^n\}$ be so that $\beta \notin A$. Then $\sum_i \alpha_i x_i - \beta, \overline{x}^2 - \overline{x}$ has a $\mathsf{poly}(|A|, n)$-explicit $\mathsf{poly}(|A|, n)$-size roABP-IPS$_{LIN}$ refutation in any variable order.*

Proof: We prove the claim for the variable order $x_1 < \cdots < x_n$, the proof for other orders is symmetric. By Theorem 4.7 there is a multilinear polynomial $f \in \mathbb{F}[\overline{x}]$ such that $f(\overline{x}) \cdot (\sum_i \alpha_i x_i - \beta) \equiv 1 \mod \overline{x}^2 - \overline{x}$ and so that f has a $\mathsf{poly}(|A|, n)$-size roABP. Similarly, $\sum_i \alpha_i x_i - \beta$ is computable by a $\mathsf{poly}(n)$-size roABP. From Theorem 3.5 it follows that $f \cdot (\sum_i \alpha_i x_i - \beta)$ is computable by a $\mathsf{poly}(|A|, n)$-size roABP. Thus, by efficient multilinearization of roABPs (Theorem 4.2) $f \cdot (\sum_i \alpha_i x_i - \beta) = 1 + \sum_i f_i(\overline{x}) \cdot (x_i^2 - x_i)$ for $f_i \in \mathbb{F}[\overline{x}]$ computable by $\mathsf{poly}(|A|, n)$-size roABPs (all in the variable orders $x_1 < \cdots < x_n$). The desired roABP-IPS$_{\text{LIN}}$ refutation is thus $C(\overline{x}, y, z_1, \ldots, z_n) := f(\overline{x})y - f_1(\overline{x})z_1 - \cdots - f_n(\overline{x})z_n$, which has a $\mathsf{poly}(|A|, n)$-size roABP (in any variable order of the $\overline{x}, y, \overline{z}$ respecting $x_1 < \cdots < x_n$) by appealing to Theorem 3.5 again. □

We now turn to refuting the subset-sum axioms by multilinear-formula IPS$_{\text{LIN}}$ (which is not itself a complete proof system, but will suffice here). While one can use the multilinearization techniques for multilinear-formula-IPS of Subsection 4.3, we directly multilinearize the refutations we built above using that the subset-sum axiom is linear.

Proposition 4.9. *Let $\overline{\alpha} \in \mathbb{F}^n$, $\beta \in \mathbb{F}$ and $A := \{\sum_{i=1}^n \alpha_i x_i : \overline{x} \in \{0,1\}^n\}$ be so that $\beta \notin A$. If $|\mathbb{F}| \geq \mathsf{poly}(|A|, n)$, then $\sum_i \alpha_i x_i - \beta, \overline{x}^2 - \overline{x}$ has a $\mathsf{poly}(|A|, n)$-explicit $\mathsf{poly}(|A|, n)$-size depth-3 multilinear-formula-IPS$_{LIN}$ refutation.*

Proof: By Theorem 4.7, there is a multilinear polynomial $f \in \mathbb{F}[\overline{x}]$ such that $f(\overline{x}) \cdot (\sum_i \alpha_i x_i - \beta) \equiv 1 \mod \overline{x}^2 - \overline{x}$, and f is explicitly given as as

$$f(\overline{x}) = \sum_{i=1}^{s} f_{i,1}(x_1) \cdots f_{i,n}(x_n) \,,$$

where each $f_{i,j} \in \mathbb{F}[x_i]$ has $\deg f_{i,j} \leq 1$, $s \leq \mathsf{poly}(|A|, n)$.

We now efficiently prove that $f(\overline{x}) \cdot (\sum_{i=1}^n \alpha_i x_i - \beta)$ is equal to its multilinearization (which is 1) modulo the boolean cube. The key step is that for a linear function $p(x) = \gamma x + \delta$ we have that $(\gamma x + \delta)x = (\gamma + \delta)x + \gamma(x^2 - x) = p(1)x + (p(1) - p(0))(x^2 - x)$.

Thus,

$$\begin{aligned}
&f(\overline{x}) \cdot (\textstyle\sum_i \alpha_i x_i - \beta) \\
&\quad = \left(\sum_{i=1}^{s} f_{i,1}(x_1) \cdots f_{i,n}(x_n)\right) \cdot (\textstyle\sum_i \alpha_i x_i - \beta) \\
&\quad = \sum_{i=1}^{s} -\beta f_{i,1}(x_1) \cdots f_{i,n}(x_n) \\
&\qquad\qquad + \sum_{i=1}^{s} \sum_{j=1}^{n} \alpha_j \prod_{k \neq j} f_{i,k}(x_k) \cdot \Big(f_{i,j}(1)x_j + (f_{i,j}(1) - f_{i,j}(0))(x_j^2 - x_j)\Big) \\
&\quad = \sum_{i=1}^{s} -\beta f_{i,1}(x_1) \cdots f_{i,n}(x_n) + \sum_{i=1}^{s} \sum_{j=1}^{n} \alpha_j \prod_{k \neq j} f_{i,k}(x_k) \cdot f_{i,j}(1)x_j \\
&\qquad\qquad + \sum_{i=1}^{s} \sum_{j=1}^{n} \alpha_j \prod_{k \neq j} f_{i,k}(x_k) \cdot (f_{i,j}(1) - f_{i,j}(0)) \cdot (x_j^2 - x_j)
\end{aligned}$$

absorbing constants and renaming, using $j = 0$ to incorporate the above term involving β,

$$= \sum_{i=1}^{s} \sum_{j=0}^{n} \prod_{k=1}^{n} f_{i,j,k}(x_k) + \sum_{j=1}^{n} \left(\sum_{i=1}^{s} \prod_{k=1}^{n} h_{i,j,k}(x_k) \right) (x_j^2 - x_j)$$

where each $f_{i,j,k}$ and $h_{i,j,k}$ are linear univariates. As $f(\overline{x}) \cdot (\sum_{i=1}^{n} \alpha_i x_i - \beta) \equiv 1 \mod \overline{x}^2 - \overline{x}$ it follows that $\sum_i \sum_j \prod_k f_{i,j,k}(x_k) \equiv 1 \mod \overline{x}^2 - \overline{x}$, but as each $f_{i,j,k}$ is linear it must actually be that $\sum_i \sum_j \prod_k f_{i,j,k}(x_k) = 1$, so that,

$$= 1 + \sum_{j=1}^{n} \left(\sum_{i=1}^{s} \prod_{k=1}^{n} h_{i,j,k}(x_k) \right) (x_j^2 - x_j) \ .$$

Define $C(\overline{x}, y, \overline{z}) := f(\overline{x})y - \sum_{j=1}^{n} h_j(\overline{x}) z_j$, where $h_j(\overline{x}) := \sum_{i=1}^{s} \prod_{k=1}^{n} h_{i,j,k}(x_k)$. It follows that $C(\overline{x}, 0, \overline{0}) = 0$ and that $C(\overline{x}, \sum_i \alpha_i x_i - \beta, \overline{x}^2 - \overline{x}) = 1$, so that C is a linear-IPS refutation. Further, as each f, h_j is computable as a sum of products of linear univariates, these are depth-3 multilinear formulas. By distributing the multiplication of the variables $y, z_1, \ldots, z_n$ to the bottom multiplication gates, we see that C itself has a depth-3 multilinear formula of the desired complexity. □

5 Lower Bounds for Linear-IPS via Functional Lower Bounds

In this section we give *functional* circuit lower bounds for various measures of algebraic complexity, such as degree, sparsity, roABPs and multilinear formulas. That is, while algebraic complexity typically treats a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ as a *syntactic* object, one can also see that it defines a function on the boolean cube $\hat{f} : \{0,1\}^n \to \mathbb{F}$. If this function is particularly complicated then one would expect that this implies that the polynomial f requires large algebraic circuits. In this section we obtain such lower bounds, showing that for *any* polynomial f that agrees with a certain function on the boolean cube must in fact have large algebraic complexity. We stress that f need not be multilinear, though for the restricted classes we consider here one can assume that f is multilinear without increasing its size (see Subsection 4.2 and Subsection 4.3).

We then observe that by deriving such lower bounds for carefully crafted functions $\hat{f} : \{0,1\}^n \to \mathbb{F}$ one can easily obtain linear-IPS lower bounds for the above circuit classes. That is, suppose that the function $\hat{f}$ obeys the functional equation $\hat{f}(\overline{x}) = 1/p(\overline{x})$ for all $\overline{x} \in \{0,1\}^n$ for a polynomial $p(\overline{x}) \in \mathbb{F}[\overline{x}]$. Then consider the system of equations $p(\overline{x}), \overline{x}^2 - \overline{x}$, where $p(\overline{x})$ is chosen so this system is unsatisfiable. Any linear-IPS refutation yields an equation $g(\overline{x}) \cdot p(\overline{x}) + \sum_i h_i(\overline{x})(x_i^2 - x_i) = 1$, which implies $g(\overline{x}) = 1/p(\overline{x})$, for all $\overline{x} \in \{0,1\}^n$, so that the polynomial $g(\overline{x})$ agrees with the function $\hat{f}$ on the boolean cube. The functional circuit lower bound then implies that g must have large complexity.

5.1 Degree of a Polynomial

We begin with a particularly weak form of algebraic complexity, the degree of a polynomial. While it is trivial to obtain such bounds in some cases (as any polynomial that agrees with the AND function on the boolean cube $\{0,1\}^n$ must have degree $\geq n$), for our applications to proof complexity (Subsection 5.3) we will need such degree bounds for functions defined by $\hat{f}(\overline{x}) = 1/p(\overline{x})$ for simple polynomials $p(\overline{x})$.

We show that any multilinear polynomial agreeing with $1/p(\overline{x})$, where $p(\overline{x})$ is the subset-sum axiom $\sum_i x_i - \beta$, must have the maximal degree n. We note that a degree lower bound of $\lceil n/2 \rceil$

was established by Impagliazzo, Pudlák, and Sgall [IPS99]. They actually established this degree bound [6] when $p(\overline{x}) = \sum_i \alpha_i x_i - \beta$ for *any* $\overline{\alpha}$, while we only consider $\overline{\alpha} = \overline{1}$ here. However, we need the tight bound of n here as it will be used crucially in the proof of Theorem 5.6.

Proposition 5.1. *Let $n \geq 1$ and $\mathbb{F}$ be a field with $\mathrm{char}(\mathbb{F}) > n$. Suppose that $\beta \in \mathbb{F} \setminus \{0, \ldots, n\}$. Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a multilinear polynomial such that*

$$f(\overline{x}) \left(\sum_i x_i - \beta \right) = 1 \mod \overline{x}^2 - \overline{x} \, .$$

Then $\deg f = n$.

Proof: Clearly $\deg f \leq n$ as f is multilinear, so it remains to show the lower bound.

Begin by observing that as $\beta \notin \{0, \ldots, n\}$, this implies that $\sum_i x_i - \beta$ is never zero on the boolean cube, so that the above functional equation implies that for $\overline{x} \in \{0,1\}^n$ the expression

$$f(\overline{x}) = \frac{1}{\sum_i x_i - \beta} \, ,$$

is well defined.

Now observe that this implies that f is a symmetric polynomial. That is, define the multilinear polynomial g by symmetrizing f,

$$g(x_1, \ldots, x_n) := \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \, ,$$

where $\mathfrak{S}_n$ is the symmetric group on n symbols. Then we see that f and g agree on $\overline{x} \in \{0,1\}^n$, as

$$\begin{aligned} g(\overline{x}) &= \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \\ &= \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \frac{1}{\sum_i x_{\sigma(i)} - \beta} = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \frac{1}{\sum_i x_i - \beta} \\ &= \frac{1}{n!} \cdot n! \cdot \frac{1}{\sum_i x_i - \beta} = \frac{1}{\sum_i x_i - \beta} = f(\overline{x}) \, . \end{aligned}$$

It follows then that $g = f$ as polynomials, since they are multilinear and agree on the boolean cube (Theorem 3.7). As g is clearly symmetric, so is f. Thus f can be expressed as $f = \sum_{k=0}^{d} \gamma_k S_{n,k}(\overline{x})$, where $d := \deg f$, $S_{n,k} := \sum_{S \subseteq \binom{[n]}{k}} \prod_{i \in S} x_i$ is the k-th elementary symmetric polynomial, and $\gamma_k \in \mathbb{F}$ are scalars with $\gamma_d \neq 0$.

Now observe that for $k < n$, we can clearly understand the action of multiplying $S_{n,k}$ by $\sum_i x_i - \beta$.

$$\begin{aligned} (\textstyle\sum_i x_i - \beta) S_{n,k}(\overline{x}) &= \sum_{S \in \binom{[n]}{k}} (\textstyle\sum_i x_i - \beta) \prod_{i \in S} x_i \\ &= \sum_{S \in \binom{[n]}{k}} \left(\textstyle\sum_{j \in S} x_j \prod_{i \in S} x_i + \sum_{j \notin S} x_j \prod_{i \in S} x_i - \beta \prod_{i \in S} x_i \right) \end{aligned}$$

[6] The degree lower bound of Impagliazzo, Pudlák, and Sgall [IPS99] actually holds for the (dynamic) polynomial calculus proof system, while we only consider the (static) Nullstellensatz proof system here. Note that for polynomial calculus there is also a matching upper bound of $\lceil n/2 \rceil$ for $\overline{\alpha} = \overline{1}$.

$$= \sum_{S \in \binom{[n]}{k}} \left(\sum_{\substack{|T|=k+1 \\ T \supseteq S}} \prod_{i \in T} x_i + (k-\beta) \prod_{i \in S} x_i \right) \mod \overline{x}^2 - \overline{x}$$
$$= (k+1) S_{n,k+1} + (k-\beta) S_{n,k} \mod \overline{x}^2 - \overline{x} .$$

Note that we used that each subset of $[n]$ of size k is contained in $n-k$ subsets of size $k+1$, and every subset of size $k+1$ contains $k+1$ subsets of size k.

Putting the above together, suppose for contradiction that $d < n$. Then

$$\begin{aligned}
1 &= f(\overline{x}) \left(\sum_i x_i - \beta \right) \mod \overline{x}^2 - \overline{x} \\
&= \left(\sum_{k=0}^{d} \gamma_k S_{n,k} \right) \left(\sum_i x_i - \beta \right) \mod \overline{x}^2 - \overline{x} \\
&= \left(\sum_{k=0}^{d} \gamma_k \Big((k+1) S_{n,k+1} + (k-\beta) S_{n,k} \Big) \right) \mod \overline{x}^2 - \overline{x} \\
&= \gamma_d (d+1) S_{n,d+1} + (\text{degree} \leq d) \mod \overline{x}^2 - \overline{x}
\end{aligned}$$

However, as $\gamma_d \neq 0$, $d+1 \leq n$ (so that $d+1 \neq 0$ in $\mathbb{F}$ and $S_{n,d+1}$ is multilinear) this is a contradiction to the uniqueness of representation of multilinear polynomials modulo $\overline{x}^2 - \overline{x}$. Thus, we must have $d = n$. □

To paraphrase the above argument, it shows that for multilinear f of $\deg f < n$ with $\mathrm{ml}(f(\overline{x}) \cdot (\sum_i x_i - \beta)) = 1$ it holds that $\deg \mathrm{ml}(f(\overline{x}) \cdot (\sum_i x_i - \beta)) = \deg f + 1$. This contradicts the fact that $\deg 1 = 0$, so that $\deg f = n$. It is tempting to attempt to argue this claim without using that $\mathrm{ml}(f(\overline{x}) \cdot (\sum_i x_i - \beta)) = 1$ in some way. That is, one could hope to argue that $\deg(\mathrm{ml}(f(\overline{x}) \cdot (\sum_i x_i - \beta))) = \deg f + 1$ directly. Unfortunately this is false, as seen by the example $\mathrm{ml}((x+y)(x-y)) = \mathrm{ml}(x^2 - y^2) = x - y$. However, one can make this approach work to obtain a degree lower bound of $\lceil n/2 \rceil$, as shown by Impagliazzo, Pudlák, and Sgall [IPS99].

Putting the above together we obtain that any polynomial agreeing with $\frac{1}{\sum_i x_i - \beta}$ on the boolean cube must be of degree $\geq n$.

Corollary 5.2. *Let $n \geq 1$ and $\mathbb{F}$ be a field with $\mathrm{char}(\mathbb{F}) > n$. Suppose that $\beta \in \mathbb{F} \setminus \{0, \ldots, n\}$. Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a multilinear polynomial such that*

$$f(\overline{x}) \left(\sum_i x_i - \beta \right) = 1 \mod \overline{x}^2 - \overline{x} .$$

Then $\deg f \geq n$.

In Appendix A we give another, more concrete, proof of the above fact, which exactly computes the unique multilinear polynomial agreeing with $\frac{1}{\sum_i x_i - \beta}$ (see Theorem A.1).

5.2 Sparse polynomials

We now use the above functional lower bounds for degree, along with random restrictions, to obtain functional lower bounds for sparsity. We then apply this to obtain exponential lower bounds for

sparse-IPS$_{\mathrm{LIN}}$ refutations of the subset-sum axioms. Recall that sparse-IPS$_{\mathrm{LIN}}$ is equivalent to the Nullstellensatz proof system when we measure the size of the proof in terms of the number of monomials. While we provide the proof here for completeness, we note that this result has already been obtained by Impagliazzo-Pudlák-Sgall [IPS99], who also gave such a lower bound for the stronger polynomial calculus proof system.

We first recall the random restrictions lemma. This lemma shows that by randomly setting half of the variables to zero, sparse polynomials become sums of monomials involving few variables, which after multilinearization is a low-degree polynomial.

Lemma 5.3. *Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be an s-sparse polynomial. Let $\rho : \mathbb{F}[\overline{x}] \to \mathbb{F}[\overline{x}]$ be the homomorphism induced by randomly and independently setting each variable x_i to 0 with probability $1/2$ and leaving x_i intact with probability $1/2$. Then with probability $\geq 1/2$, each monomial in $\rho(f(\overline{x}))$ involves $\leq \lg s + 1$ variables. Thus, with probability $\geq 1/2$, $\deg \mathrm{ml}(\rho(f)) \leq \lg s + 1$.*

Proof: Consider a monomial $\overline{x}^{\overline{a}}$ involving $\geq t$ variables. Then the probability that $\rho(\overline{x}^{\overline{a}})$ is non-zero is at most 2^{-t}. Now consider $f(\overline{x}) = \sum_{j=1}^{s} \alpha_j \overline{x}^{\overline{a}_j}$. By a union bound, the probability that any monomial $\overline{x}^{\overline{a}_j}$ involving at least t variables survives the random restriction is at most $s2^{-t}$. For $t = \lg s + 1$ this is at most $\frac{1}{2}$. The claim about the multilinearization of $\rho(f(\overline{x}))$ follows by observing that for a monomial $\overline{x}^{\overline{a}}$ involving $\leq \lg s + 1$ variables it must be that $\deg \rho(\overline{x}^{\overline{a}}) \leq \lg s + 1$. □

We now give our functional lower bound. This follows from taking any refutation of the subset-sum axiom and applying a random restriction. The subset-sum axiom will be relatively unchanged, but any sparse polynomial will become (after multilinearization) low-degree, to which our degree lower bounds (Subsection 5.1) can then be applied.

Proposition 5.4. *Let $n \geq 8$ and $\mathbb{F}$ be a field with $\mathrm{char}(\mathbb{F}) > n$. Suppose that $\beta \in \mathbb{F} \setminus \{0, \ldots, n\}$. Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a polynomial such that*

$$f(\overline{x}) = \frac{1}{\sum_i x_i - \beta} ,$$

for $\overline{x} \in \{0,1\}^n$. Then f requires $\geq 2^{n/4-1}$ monomials.

Proof: Suppose that f is s-sparse so that $f(\overline{x}) = \sum_{j=1}^{s} \alpha_j \overline{x}^{\overline{a}_j}$. Take a random restriction ρ as in Theorem 5.3, so that with probability at least $1/2$ we have that $\deg \mathrm{ml}(\rho(f)) \leq \lg s + 1$. By the Chernoff bound,[7] we see that ρ keeps alive at least $n/4$ variables with probability at least $1 - \mathrm{e}^{-2 \cdot 1/4^2 \cdot n}$, which is $\geq 1 - \mathrm{e}^{-1}$ for $n \geq 8$. Thus, by a union bound the probability that ρ fails to have either that $\deg \mathrm{ml}(\rho(f)) \leq \lg s + 1$ or that it keeps at least $n/4$ variables alive is at most $1/2 + \mathrm{e}^{-1} < 1$. Thus a ρ exists obeying both properties.

Thus, the functional equation for f implies that

$$f(\overline{x}) \left(\sum_i x_i - \beta \right) = 1 + \sum_i h_i(\overline{x})(x_i^2 - x_i) ,$$

for some $h_i \in \mathbb{F}[\overline{x}]$. Applying the random restriction and multilinearization to both sizes of this equation, we obtain that

$$\mathrm{ml}(\rho(f)) \cdot \left(\sum_{\rho(x_i) \neq 0} x_i - \beta \right) \equiv 1 \mod \{x_i^2 - x_i\}_{\rho(x_i) \neq 0} .$$

[7] For independent $[0,1]$-valued random variables $\mathsf{X}_1, \ldots, \mathsf{X}_n$, $\Pr\left[\sum_i \mathsf{X}_i - \sum_i \mathbb{E}[\mathsf{X}_i] \leq -\epsilon n\right] \leq \mathrm{e}^{-2\epsilon^2 n}$.

Thus, by appealing to the degree lower bound for this functional equation (Theorem 5.1) we obtain that $\lg s + 1 \geq \deg \mathrm{ml}(\rho(f))$ is at least the number of variables which is $\geq n/4$, so that $s \geq 2^{n/4-1}$ as desired. □

As sparse-IPS$_{\mathrm{LIN}}$ refutations of $\sum_i x_i - \beta, \overline{x}^2 - \overline{x}$ give rise to functional equations of the above form we obtain the lower bound for such refutations.

Corollary 5.5. *Let $n \geq 1$ and $\mathbb{F}$ be a field with $\mathrm{char}(\mathbb{F}) > n$. Suppose that $\beta \in \mathbb{F} \setminus \{0, \ldots, n\}$. Then $\sum_i x_i - \beta, \overline{x}^2 - \overline{x}$ is unsatisfiable and any sparse-IPS$_{LIN}$ refutation requires size $\exp(\Omega(n))$.*

5.3 Coefficient Dimension in a Fixed Partition

We now seek to prove functional circuit lower bounds for more powerful models of computation such as roABPs and multilinear formulas. As recalled in Section 3, the coefficient dimension complexity measure can give lower bounds for such models. However, by definition it is a *syntactic* measure as it speaks about the coefficients of a polynomial. Unfortunately, knowing that a polynomial $f \in \mathbb{F}[\overline{x}]$ agrees with a function $\hat{f} : \{0,1\}^n \to \mathbb{F}$ on the boolean cube $\{0,1\}^n$ does not in general give enough information to determine its coefficients. In contrast, the *evaluation* dimension measure is concerned with evaluations of a polynomial (which is functional). Obtaining lower bounds for evaluation dimension, and leveraging the fact that the evaluation dimension lower bounds coefficient dimension (Theorem 3.6) we can obtain the desired lower bounds for this complexity measure.

We now proceed to the lower bound. It will follow from the degree lower bound for the subset-sum axiom (Theorem 5.1). That is, this degree bound shows that if $f(\overline{z}) \cdot (\sum_i z_i - \beta) \equiv 1 \mod \overline{z}^2 - \overline{z}$ then f must have degree $\geq n$. We then replace $\overline{z} \leftarrow \overline{x} \circ \overline{y}$ where '$\circ$' is the Hadamard (entry-wise) product. We then leverage the degree bound to show that the evaluation dimension, which can be thought as measure of the "correlation" between $\overline{x}$ and $\overline{y}$ is maximal.

Proposition 5.6. *Let $n \geq 1$ and $\mathbb{F}$ be a field with $\mathrm{char}(\mathbb{F}) > n$. Suppose that $\beta \in \mathbb{F}$ has that $\beta \in \mathbb{F} \setminus \{0, \ldots, n\}$. Let $f \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ be a polynomial such that*

$$f(\overline{x}, \overline{y}) = \frac{1}{\sum_i x_i y_i - \beta} ,$$

for $\overline{x}, \overline{y} \in \{0,1\}^n$. Then $\dim \mathbf{Coeff}_{\overline{x}|\overline{y}} f \geq 2^n$.

Proof: By lower bounding coefficient dimension by the evaluation dimension over the boolean cube (Theorem 3.6),

$$\begin{aligned} \dim \mathbf{Coeff}_{\overline{x}|\overline{y}} f &\geq \dim \mathbf{Eval}_{\overline{x}|\overline{y},\{0,1\}} f \\ &= \dim\{f(\overline{x}, \mathbb{1}_S) : S \subseteq [n]\} \\ &\geq \dim\{\mathrm{ml}(f(\overline{x}, \mathbb{1}_S)) : S \subseteq [n]\} , \end{aligned}$$

where $\mathbb{1}_S \in \{0,1\}^n$ is the indicator vector for a set S, and ml is the multilinearization operator. Note that we used that ml is linear (Theorem 3.7) and that dimension is non-increasing under linear maps. Now note that for $\overline{x} \in \{0,1\}^n$,

$$f(\overline{x}, \mathbb{1}_S) = \frac{1}{\sum_{i \in S} x_i - \beta} ,$$

It follows then $\mathrm{ml}(f(\overline{x}, \mathbb{1}_S))$ is a multilinear polynomial only depending on $\overline{x}|_S$, and by its functional behavior it follows from Theorem 5.1 that $\deg \mathrm{ml}(f(\overline{x}, \mathbb{1}_S)) = |S|$. As $\mathrm{ml}(f(\overline{x}, \mathbb{1}_S))$ is multilinear

it thus follows that the leading monomial of $\mathrm{ml}(f(\overline{x}, \mathbb{1}_S))$ is $\prod_{i \in S} x_i$, which is distinct for each distinct S. This is also readily seen from the explicit description of $\mathrm{ml}(f(\overline{x}, \mathbb{1}_S))$ given by Theorem A.1. Thus, we can lower bound the dimension of this space by the number of leading monomials (Theorem 3.10),

$$\begin{aligned}
\dim \mathbf{Coeff}_{\overline{x}|\overline{y}} f &\geq \dim\{\mathrm{ml}(f(\overline{x}, \mathbb{1}_S)) : S \subseteq [n]\} \\
&\geq \left|\mathrm{LM}\left(\{\mathrm{ml}(f(\overline{x}, \mathbb{1}_S)) : S \subseteq [n]\}\right)\right| \\
&= \left|\left\{\prod_{i \in S} x_i : S \subseteq [n]\right\}\right| \\
&= 2^n \,.
\end{aligned}$$

□

Note that in the above proof we crucially leveraged that the degree bound of Theorem 5.1 is *exactly* n, not just $\Omega(n)$. This exact bound allows us to determine the leading monomials of these polynomials, which seems not to follow from $\Omega(n)$ degree lower bounds.

As coefficient dimension lower bounds roABP width (Theorem 3.4) and depth-3 powering formulas can be computed by roABPs in any variable order (2), we obtain as a corollary our functional lower bound for these models.

Corollary 5.7. *Let $n \geq 1$ and $\mathbb{F}$ be a field with $\mathrm{char}(\mathbb{F}) > n$. Suppose that $\beta \in \mathbb{F}$ has that $\beta \in \mathbb{F} \setminus \{0, \ldots, n\}$. Let $f \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ be a polynomial such that*

$$f(\overline{x}, \overline{y}) = \frac{1}{\sum_i x_i y_i - \beta} \,,$$

for $\overline{x}, \overline{y} \in \{0,1\}^n$. Then f requires width $\geq 2^n$ to be computed as a roABP in any variable order where $\overline{x}$ precedes $\overline{y}$. In particular, f requires $\exp(\Omega(n))$-size as a depth-3 powering formula.

We now conclude with a lower bound for linear-IPS over roABPs in certain variable orders, and thus also for depth-3 powering formulas.

Corollary 5.8. *Let $n \geq 1$ and $\mathbb{F}$ be a field with $\mathrm{char}(\mathbb{F}) > n$. Suppose that $\beta \in \mathbb{F} \setminus \{0, \ldots, n\}$. Then $\sum_i x_i y_i - \beta, \overline{x}^2 - \overline{x}, \overline{y}^2 - \overline{y}$ is unsatisfiable and any roABP-IPS$_{LIN}$ refutation, where the roABP reads $\overline{x}$ before $\overline{y}$, requires width $\geq \exp(\Omega(n))$. In particular, any $\sum \bigwedge \sum$-IPS$_{LIN}$ refutation requires size $\geq \exp(\Omega(n))$.*

Proof: That this system is unsatisfiable is clear from construction. Now consider a width-r roABP $C(\overline{x}, \overline{y}, z, \overline{u}, \overline{v})$ which reads $\overline{x}$ before $\overline{y}$, and which is a roABP-IPS$_{\mathrm{LIN}}$ refutation. That is, $C(\overline{x}, \overline{y}, z, \overline{u}, \overline{v}) = f(\overline{x}, \overline{y}) \cdot z + \sum_{i=1}^n g_i(\overline{x}, \overline{y}) u_i + \sum_{i=1}^n h_i(\overline{x}, \overline{y}) v_i$ and $C(\overline{x}, \overline{y}, \sum_i x_i y_i - \beta, \overline{x}^2 - \overline{x}, \overline{y}^2 - \overline{y}) = 1$. In particular, $f(\overline{x}, \overline{y}) = C(\overline{x}, \overline{y}, 1, 0, 0)$ is computable by a width-r roABP reading $\overline{x}$ before $\overline{y}$ (Theorem 3.5). Thus, $f(\overline{x}, \overline{y}) \cdot (\sum_i x_i y_i - \beta) \equiv 1 \mod \overline{x}^2 - \overline{x}, \overline{y}^2 - \overline{y}$ so that our functional lower bound (Theorem 5.7) implies that the width of f (and thus C) must be $\geq 2^n$. One obtains the conclusion about depth-3 powering formulas similarly, as these are also closed under partial substitution. □

Note that while Theorem 5.7 gives a lower bound for roABP refutations in any variable order where $\overline{x}$ precedes $\overline{y}$, if we allow the roABP to read the variables in the interleaved order $x_1, y_1, x_2, \ldots, x_n, y_n$ then a simple modification of Theorem 4.8 would show that there is a simple $\mathsf{poly}(|A|, n)$-size roABP-IPS$_{\mathrm{LIN}}$ refutation of $\sum_i x_i y_i - \beta, \overline{x}^2 - \overline{x}, \overline{y}^2 - \overline{y}$. We thus obtain the following separation between roABP-IPS$_{\mathrm{LIN}}$ and IPS$_{\mathrm{LIN}}$ over depth-3 powering formulas.

Corollary 5.9. *Let $\overline{\alpha} \in \mathbb{F}^n$, $\beta \in \mathbb{F}$ and $A := \{\sum_{i=1}^n \alpha_i x_i y_i : \overline{x}, \overline{y} \in \{0,1\}^n\}$ be so that $\beta \notin A$. Then $\sum_i \alpha_i x_i y_i - \beta, \overline{x}^2 - \overline{x}, \overline{y}^2 - \overline{y}$ has a $\mathsf{poly}(|A|, n)$-explicit $\mathsf{poly}(|A|, n)$-size roABP-IPS$_{LIN}$ refutation in the variable order $x_1 < y_1 < \ldots < x_n < y_n$. On the other hand, any $\sum \bigwedge \sum$-IPS$_{LIN}$ refutation requires size $\geq \exp(\Omega(n))$.*

5.4 Coefficient Dimension in any Variable Partition

The previous section gave functional lower bounds for coefficient dimension, and thus roABP width, in the $\overline{x}|\overline{y}$ variable partition. However, this lower bound fails for other variable orderings. In this section we extend the lower bound to *any* variable ordering by using suitable auxiliary variables to plant the previous lower bound into any partition we desire by suitably evaluating the auxiliary variables.

We begin by developing some preliminaries for how coefficient dimension works in the presence of auxiliary indicator variables. That is, consider a polynomial $f(\overline{x}, \overline{y}, \overline{z})$ where we wish to study the coefficient dimension of f in the $\overline{x}|\overline{y}$ partition. We can view this polynomial as lying in $\mathbb{F}[\overline{z}][\overline{x}, \overline{y}]$ so that its coefficients are polynomials in $\overline{z}$ and one studies the dimension of the coefficient space in the field of rational functions $\mathbb{F}(\overline{z})$. Alternatively one can evaluate $\overline{z}$ at some point $\overline{z} \leftarrow \overline{\alpha}$ so that $f(\overline{x}, \overline{y}, \overline{\alpha}) \in \mathbb{F}[\overline{x}, \overline{y}]$ and study its coefficient dimension over $\mathbb{F}$. The following straightforward lemma shows the first dimension over $\mathbb{F}(\overline{z})$ is lower-bounded by the second dimension over $\mathbb{F}$.

Lemma 5.10. *Let $f \in \mathbb{F}[\overline{x}, \overline{y}, \overline{z}]$. Let $f_{\overline{z}}$ denote f as a polynomial in $\mathbb{F}[\overline{z}][\overline{x}, \overline{y}]$, so that for any $\overline{\alpha} \in \mathbb{F}^{|\overline{z}|}$ we have that $f_{\overline{\alpha}} \in \mathbb{F}[\overline{x}, \overline{y}]$. Then for any such $\overline{\alpha}$,*

$$\dim_{\mathbb{F}(\overline{z})} \mathbf{Coeff}_{\overline{x}|\overline{y}} f_{\overline{z}}(\overline{x}, \overline{y}) \geq \dim_{\mathbb{F}} \mathbf{Coeff}_{\overline{x}|\overline{y}} f_{\overline{\alpha}}(\overline{x}, \overline{y}) \ .$$

Proof: By Theorem 3.1 we see that $\dim_{\mathbb{F}(\overline{z})} \mathbf{Coeff}_{\overline{x}|\overline{y}} f_{\overline{z}}(\overline{x}, \overline{y})$ is equal to the rank (over $\mathbb{F}(\overline{z})$) of the coefficient matrix $C_{f_{\overline{z}}}$, which has entries in $\mathbb{F}[\overline{z}]$. Similarly, $\dim_{\mathbb{F}} \mathbf{Coeff}_{\overline{x}|\overline{y}} f_{\overline{\alpha}}(\overline{x}, \overline{y})$ is equal to the rank (over $\mathbb{F}$) of the coefficient matrix $C_{f_{\overline{\alpha}}}$, which has entries in $\mathbb{F}$. It follows that $C_{f_{\overline{z}}}|_{\overline{z} \leftarrow \overline{\alpha}} = C_{f_{\overline{\alpha}}}$. The claim then follows by noting that for a matrix $M(\overline{w}) \in \mathbb{F}[\overline{w}]^{r \times r}$ it holds that $\mathrm{rank}_{\mathbb{F}(\overline{w})} M(\overline{w}) \geq \mathrm{rank}_{\mathbb{F}} M(\overline{\beta})$ for any $\overline{\beta} \in \mathbb{F}^{|\overline{w}|}$. This follows as the rank of $M(\overline{w})$ is equal to the maximum size of a minor with a non-vanishing determinant. As such determinants are polynomials in $\overline{w}$, they can only further vanish when $\overline{w} \leftarrow \overline{\beta}$. □

Proposition 5.11. *Let $n \geq 1$ and $\mathbb{F}$ be a field with $\mathrm{char}(\mathbb{F}) > \binom{2n}{2}$. Suppose that $\beta \in \mathbb{F}$ has that $\beta \in \mathbb{F} \setminus \{0, \ldots, \binom{2n}{2}\}$. Let $f \in \mathbb{F}[x_1, \ldots, x_{2n}, z_1, \ldots, z_{\binom{2n}{2}}]$ be a polynomial such that*

$$f(\overline{x}, \overline{z}) = \frac{1}{\sum_{i<j} z_{i,j} x_i x_j - \beta} \ ,$$

for $\overline{x} \in \{0,1\}^{2n}$, $\overline{z} \in \{0,1\}^{\binom{2n}{2}}$. Let $f_{\overline{z}}$ denote f as a polynomial in $\mathbb{F}[\overline{z}][\overline{x}]$. Then for any partition $[2n] = S \sqcup T$ with $|S| = |T| = n$,

$$\dim_{\mathbb{F}(\overline{z})} \mathbf{Coeff}_{\overline{x}|_S \overline{x}|_T} f_{\overline{z}} \geq 2^n \ .$$

Proof: As S and T are of the same size, define an arbitrary bijection $\sigma : S \to T$. Then define the $\overline{z}$-evaluation $\overline{\alpha} \in \{0,1\}^{\binom{n}{2}}$ to restrict f to sum over $x_i x_j$ in the matching, so that

$$\alpha_{i.j} = \begin{cases} 1 & i = \sigma(j) \\ 0 & \text{else} \end{cases} \ .$$

It follows then that $f(\overline{x},\overline{\alpha}) = \frac{1}{\sum_{i\in S} x_i x_{\sigma(i)} - \beta}$ for $\overline{x} \in \{0,1\}^{2n}$, which is, up to renaming, the polynomial studied in the previous section. Thus, by appealing to our lower bound for a fixed partition (Theorem 5.6) and the relation between the coefficient dimension in $f_{\overline{z}}$ versus $f_{\overline{\alpha}}$ (Theorem 5.10),

$$\begin{aligned}\dim_{\mathbb{F}(\overline{z})} \mathbf{Coeff}_{\overline{x}|_S|\overline{x}|_T} f_{\overline{z}}(\overline{x}|_S,\overline{x}|_T) &\geq \dim_{\mathbb{F}} \mathbf{Coeff}_{\overline{x}|_S|\overline{x}|_T} f_{\overline{\alpha}}(\overline{x}|_S,\overline{x}|_T)\\ &\geq 2^n\ .\end{aligned}$$

□

We remark that this lower bound is only $\exp(\Omega(\sqrt{m}))$ where $m = 2n + \binom{2n}{2}$ is the number of total variables, while one could hope for an $\exp(\Omega(m))$ lower bound. One can achieve such a lower bound by replacing the above auxiliary variable scheme (which corresponds to a complete graph) with one derived from constant-degree expander graphs. That is because in such graphs any large partition of the vertices induces a large matching across that partition, where one can then embed the fixed-partition lower bounds of the previous section (Subsection 5.3). While this would strengthen our result for functional lower bounds of roABPs, it would not suffice for the below application to multilinear formulas (Theorem 5.12) as that application requires a full rank polynomial. Indeed, the lower bound here of 2^n is also the trivial *upper* bound for the coefficient dimension of any multilinear polynomial, so that the above result is particularly sharp.

We now obtain our desired functional lower bounds for roABPs and multilinear formulas.

IT: Note that here it's corollary but in the intro it is referred to as Theorems.

Corollary 5.12. *Let $n \geq 1$ and $\mathbb{F}$ be a field with $\mathrm{char}(\mathbb{F}) > \binom{2n}{2}$. Suppose that $\beta \in \mathbb{F}$ has that $\beta \in \mathbb{F} \setminus \{0,\ldots,\binom{2n}{2}\}$. Let $f \in \mathbb{F}[x_1,\ldots,x_{2n},z_1,\ldots,z_{\binom{2n}{2}}]$ be a polynomial such that*

$$f(\overline{x},\overline{z}) = \frac{1}{\sum_{i<j} z_{i,j}x_ix_j - \beta}\ ,$$

for $\overline{x} \in \{0,1\}^{2n}$, $\overline{z} \in \{0,1\}^{\binom{2n}{2}}$. Then f requires width $\geq 2^n$ to be computed by a roABP in any variable order. Also, f requires $n^{\Omega(\log n)}$-size to be computed as a multilinear formula. For $d = o(\log n/\log\log n)$, f requires $n^{\Omega((n/\log n)^{1/d}/d^2)}$-size multilinear formulas of depth-$(2d+1)$.

Proof: roABPs: Suppose that $f(\overline{x},\overline{z})$ is computable by a width-r roABP in some variable order. By pushing the $\overline{z}$ variables into the fraction field, it follows that $f_{\overline{z}}$ (f as a polynomial in $\mathbb{F}[\overline{z}][\overline{x}]$) is also computable by a width-r roABP over $\mathbb{F}(\overline{z})$ in the induced variable order on $\overline{x}$ (Theorem 3.7). By splitting $\overline{x}$ in half along its variable order one obtains the lower bound by combining the coefficient dimension lower bound of Theorem 5.11 with its relation to roABPs (Theorem 3.4).

multilinear formulas: This follows immediately from our coefficient dimension lower bound (Theorem 5.11) and the Raz [Raz09] and Raz-Yehudayoff [RY09] results (Theorem 3.8). □

As before, this immediately yields the desired linear-IPS lower bounds.

Corollary 5.13. *Let $n \geq 1$ and $\mathbb{F}$ be a field with $\mathrm{char}(\mathbb{F}) > \binom{2n}{2}$. Suppose that $\beta \in \mathbb{F}$ has that $\beta \in \mathbb{F} \setminus \{0,\ldots,\binom{2n}{2}\}$. Then $\sum_{i<j} z_{i,j}x_ix_j - \beta, \overline{x}^2 - \overline{x}, \overline{z}^2 - \overline{z}$ is unsatisfiable, and any roABP-IPS$_{LIN}$ refutation (in any variable order) requires $\exp(\Omega(n))$-size. Further, any multilinear-formula-IPS$_{LIN'}$ refutation requires $n^{\Omega(\log n)}$-size, and any depth-$(2d+1)$ multilinear-formula-IPS$_{LIN'}$ refutation requires $n^{\Omega((n/\log n)^{1/d}/d^2)}$-size.*

Proof: Clearly the system is unsatisfiable. The roABP-IPS$_{\mathrm{LIN}}$ lower bound follows as in Theorem 5.8. For the multilinear-formula result, note that a IPS$_{\mathrm{LIN}'}$-refutation must be linear in those axioms aside from $\overline{x}^2 - \overline{x}, \overline{z}^2 - \overline{z}$. Thus, taking the refutation modulo these equations we obtain $f(\overline{x},\overline{z}) \cdot (\sum_{i<j} z_{i,j}x_ix_j - \beta) \equiv 1 \mod \overline{x}^2 - \overline{x}, \overline{z}^2 - \overline{z}$ which by the above functional lower bound implies the desired lower bound for the complexity of f. As the size of f is at most the size of the refutation (as argued in Theorem 5.8) this gives the claim. □

6 Lower Bounds for Multiples of Polynomials

In this section we consider the problem of finding explicit polynomials whose non-zero multiples are all hard. Such polynomials are natural to search for, as intuitively if f is hard to compute than so should small modifications such as $x_1 f^2 + 4f^3$. This intuition is buttressed by Kaltofen's [Kal89] result that if a polynomial has a small algebraic circuit then so do all its factors. The problem is that this result is only known to hold for general algebraic circuits, or for small depth circuits (with some additional restrictions ([DSY09, Oli15b])) and not when the polynomial is computed by a restricted circuit such as roABPs or depth-3 powering formulas.

We will begin by discussing the applications of this problem to the hardness versus randomness paradigm in algebraic complexity. We then use existing derandomization results to show that multiples of the determinant are hard for certain restricted classes. However, this method is very rigidly tied to the determinant. Thus, we also directly study existing lower bound techniques for restricted models of computation (depth-3 powering formulas, sparse polynomials, and roABPs) and extend these results to also apply to multiples. We will show the applications of such polynomials to proof complexity in section Section 7.

6.1 Connections to Hardness versus Randomness and Factoring Circuits

To motivate the problem of finding polynomials with hard multiples, we begin by discussing the hardness versus randomness approach to derandomizing polynomial identity testing. That is, Kabanets and Impagliazzo [KI04] extended the hardness versus randomness paradigm of Nisan and Wigderson [NW94] to the algebraic setting, showing that sufficiently good algebraic circuit lower bounds for an explicit polynomial would qualitatively derandomize PIT. While much of the construction is similar (using combinatorial designs, hybrid arguments, etc.) to the approach of Nisan and Wigderson [NW94] for boolean derandomization, there is a key difference. In the boolean setting, obtaining a hardness versus randomness connection requires converting *worst-case* hardness (no small computation can compute the function everywhere) to *average-case* hardness (no small computation can compute the function on most inputs). Such a reduction (obtained by Impagliazzo and Wigderson [IW97]) can in fact be obtained using certain error-correcting codes based in multivariate polynomials (as shown by Sudan, Trevisan and Vadhan [STV01]).

Such a worst-case to average-case reduction is also needed in the algebraic setting, but as multivariate polynomials are one source of this reduction in the boolean regime, it is natural to expect it to be easier in the algebraic setting. Specifically, the notion of average-case hardness for a polynomial $f(\overline{x})$ used in Kabanets-Impagliazzo [KI04] is that for any $g(\overline{x}, y)$ satisfying $g(\overline{x}, f(\overline{x})) = 0$, it must be that g then requires large algebraic circuits (by taking $g(\overline{x}, y) := y - f(\overline{x})$ this implies f itself requires large circuits). This can be interpreted as average-case hardness because if such a g existed with a small circuit, then for any value $\overline{\alpha}$ we have that $g(\overline{\alpha}, y)$ is a univariate polynomial that vanishes on $f(\overline{\alpha})$. By factoring this univariate (which can be done efficiently), we see that such g give a small list (of size at most $\deg g$) of possible values for $f(\overline{\alpha})$. By picking a random element from this list, one can correctly compute $f(\overline{x})$ with noticeable probability, which by an averaging argument one can convert to a (non-uniform) deterministic procedure to compute $f(\overline{x})$ on most inputs (over any fixed finite set). While this procedure (involving univariate factorization) is not an algebraic circuit, the above argument shows that the Kabanets-Impagliazzo [KI04] notion is a natural form of average case hardness.

To obtain this form of average-case hardness from worst-case hardness, Kabanets and Impagliazzo [KI04] used a result of Kaltofen [Kal89], who showed that (up to pathologies in low-characteristic fields), factors of small (general) circuits have small circuits. As $g(\overline{x}, f(\overline{x})) = 0$ iff

$y - f(\overline{x})$ divides $g(\overline{x}, y)$, it follows that if $g(\overline{x}, y)$ has a small circuit then so does $y - f(\overline{x})$, and thus so does $f(\overline{x})$. Taking the contrapositive, if f requires large circuits (worst-case hardness) then any such $g(\overline{x}, y)$ with $g(\overline{x}, f(\overline{x})) = 0$ also requires large circuits (average-case hardness). Note that this says that *any* worst-case hard polynomial is *also* average-case hard. In contrast, this is provably false for boolean functions, where such worst-case to average-case reductions thus necessarily modify the original function.

Unfortunately, Kaltofen's [Kal89] factoring algorithm does not preserve structural restrictions (such as multilinearity, homogeneousness, low-depth, read-once-ness, etc.) of the original circuit, so that obtaining average-case hardness for restricted classes of circuits requires worst-case hardness for much stronger classes. While follow-up work has reduced the complexity of the circuits resulting from Kaltofen's [Kal89] algorithm (Dvir-Shpilka-Yehudayoff [DSY09] and Oliveira [Oli15b] extended Kaltofen's [Kal89] to roughly preserve the depth of the original computation) these works are limited to factoring polynomials of small individual degree and do not seem applicable to other types of computations such as roABPs. Indeed, it even remains an open question to show any non-trivial upper bounds on the complexity of the factors of sparse polynomials. In fact, we actually have non-trivial *lower* bounds. Specifically, von zur Gathen and Kaltofen [vzGK85] gave an explicit s-sparse polynomial (over any field) which has a factor with $s^{\Omega(\log s)}$ monomials, and Volkovich [Vol15b] gave, for a prime p, an explicit n-variate n-sparse polynomial which in characteristic p has a factor with $\binom{n+p-2}{n}$ monomials (an exponential separation for $p \geq \mathsf{poly}(n)$). .

While showing the equivalence of worst-case and average-case hardness for restricted circuit classes seems difficult, to derandomize PIT via Kabanets-Impagliazzo [KI04] only requires a *single* polynomial which is average case hard. To facilitate obtaining such hard polynomials, we now give a lemma showing that polynomials with only hard multiples are average case hard.

Lemma 6.1. *Let $f(\overline{x}) \in \mathbb{F}[\overline{x}]$ and $g(\overline{x}, y) \in \mathbb{F}[\overline{x}, y]$ both be non-zero. If $g(\overline{x}, f(\overline{x})) = 0$ then $f(\overline{x}) | g(\overline{x}, 0)$.*

Proof: Let $g(\overline{x}, y) = \sum_i g_i(\overline{x}) y^i$, so that $g(\overline{x}, 0) = g_0(\overline{x})$. That $g(\overline{x}, f(\overline{x})) = 0$ implies that

$$0 = g(\overline{x}, f(\overline{x})) = \sum_i g_i(\overline{x})(f(\overline{x}))^i = g_0(\overline{x}) + \sum_{i \geq 1} g_i(\overline{x})(f(\overline{x}))^i$$

so that $g_0(\overline{x}) = f(\overline{x}) \cdot \left(- \sum_{i \geq 1} g_i(\overline{x})(f(\overline{x}))^{i-1} \right)$ as desired. □

That is, as the size of computing $g(\overline{x}, 0)$ is bounded by that of $g(\overline{x}, y)$ (for most natural measures of circuit size), we have that if $f(\overline{x})$ has only hard multiples then it is also average-case hard in the sense needed for Kabanets-Impagliazzo [KI04]. However, note that the converse of this lemma is false, as seen by considering $g(x, y) := y - x(x+1)$, so that $x | g(x, 0)$ but $g(x, x) \neq 0$.

6.2 Lower Bounds for Multiples via PIT

This above discussion shows that obtaining lower bounds for multiples is a weaker condition sufficient for instantiating the hardness versus randomness paradigm. While at first this may seem no easier than the more general problem of bounding the complexity of factors, we now observe that one can obtain such polynomials with hard multiples via derandomizing (black-box) PIT, or equivalently, producing generators with small seed-length. That is, Heintz-Schnorr [HS80] and Agrawal [Agr05] showed that one can use explicit hitting sets for small circuits to obtain explicit hard polynomials, and we observe here that the resulting polynomials also have only hard multiples.

Lemma 6.2. *Let $\mathcal{C} \subseteq \mathbb{F}[\overline{x}]$ be a class of polynomials and let $\overline{\mathcal{G}} : \mathbb{F}^{\ell} \to \mathbb{F}^{\overline{x}}$ be a generator for $\mathcal{C}$. Suppose $0 \neq h \in \mathbb{F}[\overline{x}]$ has $h \circ \overline{\mathcal{G}} = 0$. Then for any non-zero $g \in \mathbb{F}[\overline{x}]$ we have that $g \cdot h \notin \mathcal{C}$.*

Proof: By definition of $\overline{\mathcal{G}}$, for any $f \in \mathcal{C}$, $f = 0$ iff $f \circ \overline{\mathcal{G}} = 0$. Then for any such g, $g \cdot h \neq 0$ and $(g \cdot h) \circ \overline{G} = (g \circ \overline{\mathcal{G}}) \cdot (h \cdot \overline{\mathcal{G}}) = (g \circ \overline{\mathcal{G}}) \cdot 0 = 0$. Thus, we must have that $g \cdot h \notin \mathcal{C}$. □

While one can analogously prove the hitting-set version of this claim, it is a weaker claim. That is, it is possible to consider classes $\mathcal{C}$ of unbounded degree and still have generators with small seed-length (see Theorem 6.5 below), but for such classes one must have hitting sets with infinite size (as hitting univariate polynomials of unbounded degree requires an infinite number of points).

Thus the above claim shows that obtaining black-box PIT yields the existence of a polynomial with hard multiples, which yields average-case hardness, which (for general enough classes) will allow the Kabanets-Impagliazzo [KI04] reduction to again yield black-box PIT. Thus, we see that obtaining such polynomials with hard multiples is essentially what is needed for this hardness versus randomness approach.

While there are now a variety of restricted circuit classes with non-trivial (black-box) PIT results, it seems challenging to find for any given generator $\mathcal{G}$ an *explicit* non-zero polynomial f with $f \circ \mathcal{G} = 0$. Indeed, to the best of our knowledge no such examples have ever been furnished for interesting generators. Aside from the quest for polynomials with hard multiples, this question is independently interesting as it demonstrates the limits of the generator in question, especially for generators that are commonly used. There is not even a consensus as to whether the generators currently constructed could suffice to derandomize PIT for general circuits. Agrawal [Agr05] has even conjectured that a certain generator for depth-2 circuits (sparse polynomials) would actually suffice for PIT of constant-depth circuits.

We consider here the generator of Shpilka-Volkovich [SV09]. This generator has a parameter ℓ, and intuitively can seen as an algebraic analogue of the boolean pseudorandomness notion of a (randomness efficient) ℓ-wise independent hash function. Just as ℓ-wise independent hash functions are ubiquitous in boolean pseudorandomness, the Shpilka-Volkovich [SV09] generator has likewise been used in a number of papers on black-box PIT (for example [SV09, AvMV11, FS13a, FSS14, Vol15b, For15] is a partial list). As such, we believe it is important to understand the limits of this generator.

However, ℓ-wise independence is a *property* of a hash function and likewise the Shpilka-Volkovich [SV09] generator is really a family of generators that all share a certain property. Specifically, the map $\overline{\mathcal{G}}^{\mathrm{SV}}_{\ell,n} : \mathbb{F}^r \to \mathbb{F}^n$ has the property [8] that the image $\overline{\mathcal{G}}^{\mathrm{SV}}_{\ell,n}(\mathbb{F}^r)$ contains all ℓ-sparse vectors in $\mathbb{F}^n$. The most straightforward construction of a randomness efficient generator with this property has that $r = 2\ell$, though even in this construction there is freedom to choose the finite set of points where Lagrange interpolation will be performed. To understand the power of the above *property* we are free to construct another generator $\overline{\mathcal{G}}^{\mathrm{SV}'}_{\ell,n}$ with this property for which we can find an explicit f where $f \circ \overline{\mathcal{G}}^{\mathrm{SV}'}_{\ell,n} = 0$ for small ℓ. We choose a variant of the original construction so that we can take f as the determinant.

In the original Shpilka-Volkovich [SV09] generator, one first obtains the $\ell = 1$ construction by using two variables, the control variable y and another variable z. By using Lagrange polynomials to simulate indicator functions, the value of y can be set to choose between the outputs $z\overline{e}_1, \ldots, z\overline{e}_n \in \mathbb{F}[z]^n$, where $\overline{e}_i \in \mathbb{F}^n$ is the i-th standard basis vector. By varying z one obtains all 1-sparse vectors in $\mathbb{F}^n$. To obtain $\overline{\mathcal{G}}^{\mathrm{SV}}_{\ell,n}$ one can sum ℓ independent copies of $\overline{\mathcal{G}}^{\mathrm{SV}}_{1,n}$. In contrast, our construction will

[8] The most obvious algebraic analogue of a ℓ-wise independent hash function would require that for a generator $\overline{G} : \mathbb{F}^r \to \mathbb{F}^n$ that any subset $S \subseteq [n]$ with $|S| \leq \ell$ the output of $\overline{G}$ restricted to S is all of $\mathbb{F}^S$. This property is implied by the Shpilka-Volkovich [SV09] property, but is strictly weaker, and is in fact too weak to be useful for PIT. That is, consider the map $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_n, x_1 + \cdots + x_n)$. This map has this naive "algebraic n-wise independence" property, but does not even fool linear polynomials (which the Shpilka-Volkovich [SV09] generator would).

simply use a different control mechanism, where instead of using univariate polynomials we use bivariates.

Construction 6.3. *Let $n, \ell \geq 1$. Let $\mathbb{F}$ be a field of size $> n$. Let $\Omega := \{\omega_1, \ldots, \omega_n\}$ be distinct elements in $\mathbb{F}$. Define $\overline{\mathcal{G}}^{\mathrm{SV}'}_{1,n} : \mathbb{F}^3 \to \mathbb{F}^{n \times n}$ by*

$$\left(\overline{\mathcal{G}}^{\mathrm{SV}'}_{1,n}(x,y,z)\right)_{i,j} = z \cdot \mathbb{1}_{\omega_i,\Omega}(x) \cdot \mathbb{1}_{\omega_j,\Omega}(y) \ .$$

where $\mathbb{1}_{\omega_i,\Omega}(x) \in \mathbb{F}[x]$ is the unique univariate polynomial of degree $< n$ such that

$$\mathbb{1}_{\omega_i,\Omega}(\omega_j) = \begin{cases} 1 & i = j \\ 0 & else \end{cases} \ .$$

Define $\overline{\mathcal{G}}^{\mathrm{SV}'}_{\ell,n} : \mathbb{F}^{3\ell} \to \mathbb{F}^{n \times n}$ by the polynomial map

$$\overline{\mathcal{G}}^{\mathrm{SV}'}_{\ell,n}(x_1, y_1, z_1, \ldots, x_\ell, y_\ell, z_\ell) := \overline{\mathcal{G}}^{\mathrm{SV}'}_{1,n}(x_1, y_1, z_1) + \cdots + \overline{\mathcal{G}}^{\mathrm{SV}'}_{1,n}(x_\ell, y_\ell, z_\ell) \ ,$$

working in the ring $\mathbb{F}[\overline{x}, \overline{y}, \overline{z}]$.

Note that this map has n^2 outputs. Now observe that it is straightforward to see that this map has the desired property.

Lemma 6.4. *Assume the setup of Theorem 6.3. Then the image of the generator, $\overline{\mathcal{G}}^{\mathrm{SV}'}_{\ell,n}(\mathbb{F}^{3\ell})$, contains all ℓ-sparse vectors in $\mathbb{F}^{n \times n}$.*

As this property is all that is used [9] about the Shpilka-Volkovich [SV09] generator, we replace it with our construction in known black-box PIT results (such as [SV09, ASS13, FS13a, FSS14, GKST15, For15]), some of which we now state.

Corollary 6.5. *Assume the setup of Theorem 6.3. Then $\overline{\mathcal{G}}^{\mathrm{SV}'}_{O(\log s),n}$ is a generator for the following classes of n-variate polynomials, of arbitrary degree.*

- *Polynomials of sparsity s ([SV09, GKST15, For15]).*
- *Polynomials computable as a depth-3 powering formula of size s ([ASS13, FS13a]).*
- *Polynomials computable as a $\sum \bigwedge \sum \prod^{\mathcal{O}(1)}$ formula of size s ([For15]), in characteristic zero.*
- *Polynomials computable by width-s roABPs in every variable order, also known as commutative roABPs ([FSS14]).*

The above result shows the power of the $\overline{\mathcal{G}}^{\mathrm{SV}'}_{\ell,n}$ generator to hit restricted classes of circuits. We now observe that it is also limited by its inability to hit the determinant.

Proposition 6.6. *Assume the setup of Theorem 6.3. The output of $\overline{\mathcal{G}}^{\mathrm{SV}'}_{\ell,n}$ is an $n \times n$ matrix of rank $\leq \ell$, when viewed as a matrix over the ring $\mathbb{F}(\overline{x}, \overline{y}, \overline{z})$. Thus, taking $\det_n \in \mathbb{F}[X]$ to be the $n \times n$ determinant, we have that $\det_n \circ \overline{\mathcal{G}}^{\mathrm{SV}'}_{\ell} = 0$ for $\ell < n$.*

[9] Note that for black-box PIT it is important that we use a *generator* that contains all sparse vectors, instead of just the *set* of sparse vectors. As an example, the monomial $x_1 \ldots x_n$ is zero on all k-sparse vectors for $k < n$, but is non-zero when evaluated on the Shpilka-Volkovich [SV09] generator for any $\ell \geq 1$.

Proof: $\underline{\ell = 1}$: For a field $\mathbb{K}$, a (non-zero) matrix $M \in \mathbb{K}^{n \times n}$ is rank-1 if it can be expressed as an outer-product, so that $(M)_{i,j} = \alpha_i \beta_j$ for $\overline{\alpha}, \overline{\beta} \in \mathbb{K}^n$. Taking $\overline{\alpha}, \overline{\beta} \in \mathbb{F}(\overline{x}, \overline{y}, \overline{z})^n$ defined by $\alpha_i := z \mathbb{1}_{\omega_i, \Omega}(x)$ and $\beta_j := \mathbb{1}_{\omega_j, \Omega}(y)$ we immediately see that $\overline{\mathcal{G}}^{\mathrm{SV}'}_{1,n}(x, y, z)$ is rank-1.

$\underline{\ell > 1}$: This follows as rank is subadditive, and $\overline{\mathcal{G}}^{\mathrm{SV}'}_{\ell}$ is the sum of ℓ copies of $\overline{\mathcal{G}}^{\mathrm{SV}'}_{1}$.

$\underline{\det_n \text{ vanishes:}}$ This follows as the $n \times n$ determinant vanishes on matrices of rank $< n$. □

Note that in the above we could hope to find an f such that $f \circ \overline{\mathcal{G}}^{\mathrm{SV}'}_{\ell} = 0$ for all $\ell < n^2$, but we are only able to handle $\ell < n$. Given the above result, along with Theorem 6.2, we obtain that the multiples of the determinant are hard.

Corollary 6.7. *Let* $\det_n \in \mathbb{F}[X]$ *denote the* $n \times n$ *determinant. Then any non-zero multiple* $f \cdot \det_n$ *of* $\det_n$*, for* $0 \neq f \in \mathbb{F}[X]$*, has the following lower bounds.*

- $f \cdot \det_n$ *involves* $\exp(\Omega(n))$ *monomials.*
- $f \cdot \det_n$ *requires size* $\exp(\Omega(n))$ *to be expressed as a depth-3 powering formula.*
- $f \cdot \det_n$ *requires size* $\exp(\Omega(n))$ *to be expressed as a* $\sum \bigwedge \sum \prod^{\mathcal{O}(1)}$ *formula, in characteristic zero.*
- $f \cdot \det_n$ *requires width-*$\exp(\Omega(n))$ *to be computed as a roABP in some variable order.*

Proof: By Theorem 6.5, $\overline{\mathcal{G}}^{\mathrm{SV}'}_{O(\log s),n}$ is a generator for the above size-s computations in the above classes. However, following Theorem 6.2, $(f \cdot \det_n) \circ \left(\overline{\mathcal{G}}^{\mathrm{SV}'}_{\ell,n}\right) = 0$ for $\ell < n$. Thus, if $f \cdot \det_n$ (which is non-zero) is computable in size-s it must be that $O(\log s) \geq n$, so that $s \geq \exp(\Omega(n))$. □

Note that the above results do not directly apply to other polynomials, as we crucially leveraged that the determinant vanishes on low-rank matrices. However, one can extend some of these results to (say) the permanent by using VNP-completeness of the permanent, but we do not do so here as it does not qualitatively change the result.

6.3 Lower Bounds for Multiples via Leading/Trailing Monomials

We now use the theory of leading (and trailing) monomials to obtain explicit polynomials with hard multiples. We aim at finding as simple polynomials as possible so they will give rise to simple "axioms" with no small refutations. These results will essentially be immediate corollaries of previous work.

6.3.1 Depth-3 Powering Formulas

Kayal [Kay08] observed that using the partial derivative method of Nisan and Wigderson [NW96] one can show that these formulas require $\exp(\Omega(n))$ size to compute the monomial $x_1 \cdots x_n$. Forbes and Shpilka [FS13a] later observed that this result can be made *robust* by modifying the *hardness of representation* technique of Shpilka and Volkovich [SV09], in that similar lower bounds apply when the leading monomial involves many variables, as we now quote.

Theorem 3 (Forbes-Shpilka [FS13a])**.** *Let* $f(\overline{x}) \in \mathbb{F}[\overline{x}]$ *be computed a* $\sum \bigwedge \sum$ *formula of size* $\leq s$*. Then the leading monomial* $\overline{x}^{\overline{a}} = \mathrm{LM}(f)$ *involves* $|\overline{a}|_0 \leq \lg s$ *variables.*

We now note that as the leading monomial is multiplicative (Theorem 3.9) this lower bound automatically extends to multiples of the monomial.

Corollary 6.8. *All non-zero multiples of $x_1 \cdots x_n$ require size $\geq 2^n$ to be computed as $\sum\bigwedge\sum$ formula.*

Proof: Consider any $0 \neq g(\overline{x}) \in \mathbb{F}[x_1, \ldots, x_n]$. Then as the leading monomial is multiplicative (Theorem 3.9) we have that $\mathrm{LM}(g \cdot x_1 \cdots x_n) = \mathrm{LM}(g) \cdot x_1 \cdots x_n$, so that $\mathrm{LM}(g \cdot x_1 \cdots x_n)$ involves n variables. By the robust lower bound (3) this implies $g(\overline{x}) \cdot x_1 \cdots x_n$ requires size $\geq 2^n$ as a $\sum\bigwedge\sum$ formula. □

6.3.2 $\sum\bigwedge\sum\prod^{\mathcal{O}(1)}$ Formulas

Kayal [Kay12] introduced the method of shifted partial derivatives, and Gupta-Kamath-Kayal-Saptharishi [GKKS14] refined it to give exponential lower bounds for various sub-models of depth-4 formulas. In particular, it was shown that the monomial $x_1 \cdots x_n$ requires $\exp(\Omega(n))$-size to be computed as a $\sum\bigwedge\sum\prod^{\mathcal{O}(1)}$ formula. Applying the hardness of representation approach of Shpilka and Volkovich [SV09], Mahajan-Rao-Sreenivasaiah [MRS14] showed how to develop a deterministic black-box PIT algorithm for multilinear polynomials computed by $\sum\bigwedge\sum\prod^{\mathcal{O}(1)}$ formulas. Independently, Forbes [For15] (following Forbes-Shpilka [FS13a]) showed that this lower bound can again be made to apply to leading monomials [10] (which implies a deterministic black-box PIT algorithm for all $\sum\bigwedge\sum\prod^{\mathcal{O}(1)}$ formulas, with the same complexity as Mahajan-Rao-Sreenivasaiah [MRS14]). This leading monomial lower bound, which we now state, is important for its applications to polynomials with hard multiples.

Theorem 4 (Forbes [For15])**.** *Let $f(\overline{x}) \in \mathbb{F}[\overline{x}]$ be computed as a $\sum\bigwedge\sum\prod^{t}$ formula of size $\leq s$. If $\mathrm{char}(\mathbb{F}) \geq \mathrm{ideg}(\overline{x}^{\overline{a}})$, then the leading monomial $\overline{x}^{\overline{a}} = \mathrm{LM}(f)$ involves $|\overline{a}|_0 \leq O(t \lg s)$ variables.*

As for depth-3 powering formulas (Theorem 6.8), this immediately yields that all multiples (of degree below the characteristic) of the monomial are hard.

Corollary 6.9. *All non-zero multiples of $x_1 \cdots x_n$ of degree $< \mathrm{char}(\mathbb{F})$ require size $\geq \exp(n/t)$ to be computed as $\sum\bigwedge\sum\prod^{t}$ formula.*

6.3.3 Sparse Polynomials

While the above approaches for $\sum\bigwedge\sum$ and $\sum\bigwedge\sum\prod^{\mathcal{O}(1)}$ formulas focus on leading monomials, one cannot show that the leading monomials of sparse polynomials involve few variables as sparse polynomials can easily compute the monomial $x_1 \cdots x_n$. However, following the *translation* idea of Agrawal-Saha-Saxena [ASS13], Gurjar-Korwar-Saxena-Thierauf [GKST15] showed that sparse polynomials under full-support translations have *some* monomial involving few variables, and Forbes [For15] (using different techniques) showed that in fact the *trailing* monomial involving few variables (translations do not affect the leading monomial, so the switch to trailing monomials is necessary here).

Theorem 5 (Forbes [For15])**.** *Let $f(\overline{x}) \in \mathbb{F}[\overline{x}]$ be $(\leq s)$-sparse, and let $\overline{\alpha} \in (\mathbb{F} \setminus \{0\})^n$ so that $\overline{\alpha}$ has full-support. Then the trailing monomial $\overline{x}^{\overline{a}} = \mathrm{TM}(f(\overline{x} + \overline{\alpha}))$ involves $|\overline{a}|_0 \leq \lg s$ variables.*

This result thus allows one to construct polynomials whose multiples are all non-sparse.

Corollary 6.10. *All non-zero multiples of $(x_1 + 1) \cdots (x_n + 1) \in \mathbb{F}[\overline{x}]$ require sparsity $\geq 2^n$. Similarly, all non-zero multiples of $(x_1 + y_1) \cdots (x_n + y_n) \in \mathbb{F}[\overline{x}, \overline{y}]$ require sparsity $\geq 2^n$.*

[10] The result there is stated for trailing monomials, but the argument equally applies to leading monomials

Proof: Define $f(\overline{x}) = \prod_{i=1}^{n}(x_i + 1)$. For any $0 \neq g(\overline{x}) \in \mathbb{F}[\overline{x}]$ the multiple $g(\overline{x})f(\overline{x})$ under the translation $\overline{x} \mapsto \overline{x} - \overline{1}$ is equal to $g(\overline{x} - \overline{1}) \prod_i x_i$. Then all monomials (in particular the trailing monomial) involves n variables (as $g(\overline{x}) \neq 0$ implies $g(\overline{x} - \overline{1}) \neq 0$). Thus, by 5 it must be that $g(\overline{x})f(\overline{x})$ requires $\geq 2^n$ monomials.

The second part of the claim follows from the first, noting that setting $\overline{y} \leftarrow \overline{1}$ does not increase sparsity in a multiple $g(\overline{x}, \overline{y}) \cdot \prod_i (x_i + y_i)$. □

6.4 Lower Bounds for Multiples of Sparse Multilinear Polynomials

While the previous section established that all multiples of $(x_1 + 1) \cdots (x_n + 1)$ are non-sparse, the argument was somewhat specific to that polynomial and fails to obtain an analogous result for $(x_1 + 1) \cdots (x_n + 1) + 1$. Further, while that argument can show for example that all multiples of the $n \times n$ determinant or permanent require sparsity $\geq \exp(\Omega(n))$, this is the best sparsity lower bound obtainable for these polynomials with this method.[11] In particular, one cannot establish a sparsity lower bound of "$n!$" for the determinant or permanent (which would be tight) via this method.

We now give a different argument, due to Oliveira [Oli15a] that establishes a much more general result showing that multiples of *any* multilinear polynomial have at least the sparsity of the original polynomial. While Oliveira [Oli15a] gave a proof using Newton polytopes, we give a more compact proof here using induction on variables (loosely inspired by a similar result of Volkovich [Vol15a] on the sparsity of factors of multi-quadratic polynomials).

Proposition 6.11 (Oliveira [Oli15a])**.** *Let $f(\overline{x}) \in \mathbb{F}[x_1, \ldots, x_n]$ be a non-zero multilinear polynomial with sparsity exactly s. Then any non-zero multiple of f has sparsity $\geq s$.*

Proof: By induction on variables.

$\underline{n = 0}$: Then f is a constant, so that $s = 1$ as $f \neq 0$. All non-zero multiples are non-zero polynomials so have sparsity ≥ 1.

$\underline{n \geq 1}$: Partition the variables $\overline{x} = (\overline{y}, z)$, so that $f(\overline{y}, z) = f_1(\overline{y})z + f_0(\overline{y})$, where $f_i(\overline{y})$ has sparsity s_i and $s = s_1 + s_0$. Consider any non-zero $g(\overline{y}, z) = \sum_{i=d_0}^{d_1} g_i(\overline{y})z^i$ with $g_{d_0}(\overline{y}), g_{d_1}(\overline{y}) \neq 0$ (possibly $d_0 = d_1$). Then

$$\begin{aligned} g(\overline{y}, z)f(\overline{y}, z) &= \Big(f_1(\overline{y})z + f_0(\overline{y})\Big) \cdot \left(\sum_{i=d_0}^{d_1} g_i(\overline{y})z^i\right) \\ &= f_1(\overline{y})g_{d_1}(\overline{y})z^{d_1+1} + \left[\sum_{d_0 < i \leq d_1} \Big(f_1(\overline{y})g_{i-1}(\overline{y}) + f_0(\overline{y})g_i(\overline{y})\Big)z^i\right] + f_0(\overline{y})g_{d_0}(\overline{y})z^{d_0} \end{aligned}$$

By partitioning this sum by powers of z and focusing on the extreme points,

$$|\operatorname{Supp}\Big(g(\overline{y}, z)f(\overline{y}, z)\Big)| \geq |\operatorname{Supp}\Big(f_1(\overline{y})g_{d_1}(\overline{y})\Big)| + |\operatorname{Supp}\Big(f_0(\overline{y})g_{d_0}(\overline{y})\Big)|$$

so that appealing to the induction hypothesis, as f_0 and f_1 are multilinear polynomials of sparsity s_0 and s_1 respectively,

$$\geq s_1 + s_0 = s\ .$$ □

[11] Specifically, as the determinant and permanent are degree n multilinear polynomials, and thus so are their translations, their monomials always involve $\leq n$ variables so no sparsity bound better than 2^n can be obtained by using 5.

We note that multilinearity is essential in the above lemma, even for univariates. This is seen by noting that the 2-sparse polynomial $x^n - 1$ is a multiple of $x^{n-1} + \cdots + x + 1$.

Thus, the above not only gives a different proof of the non-sparsity of multiples of $\prod_i (x_i + 1)$ (Theorem 6.10), but also establishes that non-zero multiples of $\prod_i (x_i + 1) - 1$ are $\geq 2^n - 1$ sparse, and non-zero multiples of the determinant or permanent are $n!$ sparse, which is tight. Note further that this lower bound proof is "monotone" in that it applies to any polynomial with the same support, whereas the proof of Theorem 6.10 is seemingly not monotone as seen by contrasting $\prod_i (x_i + 1)$ and $\prod_i (x_i + 1) - 1$.

6.5 Lower Bounds for Multiples by Leading/Trailing Diagonals

In the previous sections we obtained polynomials with hard multiples for various circuit classes by appealing to the fact that lower bounds for these classes can be reduced to studying the number of variables in leading or trailing monomials. Unfortunately this approach is restricted to circuit classes where monomials (or translations of monomials) are hard to compute, which in particular rules out this approach for roABPs. Thus, to develop polynomials with hard multiples for roABPs we need to develop a different notion of a "leading part" of a polynomial. In this section, we define such a notion called a *leading diagonal*, establish its basic properties, and obtain the desired polynomials with hard multiples. The ideas of this section are a cleaner version the techniques used in the PIT algorithm of Forbes and Shpilka [FS12] for commutative roABPs.

We begin with the definition.

Definition 6.1. *Let $f \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ be non-zero. The **leading diagonal of** f, denoted $\mathrm{LD}(f)$, is the leading coefficient of $f(\overline{x} \circ \overline{z}, \overline{y} \circ \overline{z})$ when this polynomial is considered in the ring $\mathbb{F}[\overline{x}, \overline{y}][z_1, \ldots, z_n]$, and where $\overline{x} \circ \overline{z}$ denotes the Hadamard product $(x_1 z_1, \ldots, x_n, z_n)$. The **trailing diagonal of** f is defined analogously. The zero polynomial has no leading or trailing diagonal.*

As this notion has not explicitly appeared prior in the literature, we now establish several straightforward properties. The first is that extremal diagonals are homomorphic with respect to multiplication.

Lemma 6.12. *Let $f, g \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ be non-zero. Then $\mathrm{LD}(fg) = \mathrm{LD}(f)\,\mathrm{LD}(g)$ and $\mathrm{TD}(fg) = \mathrm{TD}(f)\,\mathrm{TD}(g)$.*

Proof: As $\mathrm{LD}(f) = \mathrm{LC}_{\overline{x},\overline{y}|\overline{z}}(f(\overline{x} \circ \overline{z}, \overline{y} \circ \overline{z}))$, where this leading coefficient is taken in the ring $\mathbb{F}[\overline{x}, \overline{y}][\overline{z}]$, this automatically follows from the fact that leading coefficients are homomorphic with respect to multiplication (Theorem 3.9). The result for trailing diagonals is symmetric. □

We now show how to relate the leading monomials of the coefficient space of f to the respective monomials associated to the leading diagonal of f.

Proposition 6.13. *Let $f \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots, y_n]$. For any $\overline{b}$, if $\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(\mathrm{LD}(f)) \neq 0$, then*

$$\mathrm{LM}\left(\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(\mathrm{LD}(f))\right) = \mathrm{LM}\left(\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(f)\right) .$$

The respective trailing statement also holds.

Proof: We prove the leading statement, the trailing version is symmetric. Let $f = \sum_{\overline{a},\overline{b}} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}} \overline{y}^{\overline{b}}$. We can then expand $f(\overline{x} \circ \overline{z}, \overline{y} \circ \overline{z})$ as follows.

$$f(\overline{x} \circ \overline{z}, \overline{y} \circ \overline{z}) = \sum_{\overline{c}} \left(\sum_{\overline{a}+\overline{b}=\overline{c}} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}} \overline{y}^{\overline{b}} \right) \overline{z}^{\overline{c}}$$

choose $\overline{c}_0$ so that $\mathrm{LC}_{\overline{x},\overline{y}|\overline{z}}(f) = \mathrm{Coeff}_{\overline{x},\overline{y}|\overline{z}^{\overline{c}_0}}(f)$, we get that

$$= \left(\sum_{\overline{a}+\overline{b}=\overline{c}_0} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}} \overline{y}^{\overline{b}}\right) \overline{z}^{\overline{c}_0} + \sum_{\overline{c} \prec \overline{c}_0} \left(\sum_{\overline{a}+\overline{b}=\overline{c}} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}} \overline{y}^{\overline{b}}\right) \overline{z}^{\overline{c}} \,,$$

where $\mathrm{LD}(f) = \sum_{\overline{a}+\overline{b}=\overline{c}_0} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}} \overline{y}^{\overline{b}}$ and $\sum_{\overline{a}+\overline{b}=\overline{c}} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}} \overline{y}^{\overline{b}} = 0$ for $\overline{c} \succ \overline{c}_0$. In particular, this means that for any $\overline{b}$ we have that $\alpha_{\overline{a},\overline{b}} = 0$ for $\overline{a} \succ \overline{c}_0 - \overline{b}$.

Thus we have that

$$\begin{aligned}
\mathrm{LM}\left(\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(\mathrm{LD}(f))\right) &= \mathrm{LM}\left(\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}\left(\sum_{\overline{a}+\overline{b}=\overline{c}_0} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}} \overline{y}^{\overline{b}}\right)\right) \\
&= \mathrm{LM}\left(\alpha_{\overline{c}_0-\overline{b},\overline{b}} \overline{x}^{\overline{c}_0-\overline{b}}\right) \\
&= \overline{x}^{\overline{c}_0-\overline{b}} \,,
\end{aligned}$$

as we assume this leading monomial exists, which is equivalent here to $\alpha_{\overline{c}_0-\overline{b},\overline{b}} \neq 0$.

In comparison,

$$\begin{aligned}
\mathrm{LM}\left(\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(f)\right) &= \mathrm{LM}\left(\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}\left(\sum_{\overline{a},\overline{b}} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}} \overline{y}^{\overline{b}}\right)\right) \\
&= \mathrm{LM}\left(\sum_{\overline{a}} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}}\right) \\
&= \mathrm{LM}\left(\sum_{\overline{a} \succ \overline{c}_0-\overline{b}} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}} + \alpha_{\overline{c}_0-\overline{b},\overline{b}} \overline{x}^{\overline{c}_0-\overline{b}} + \sum_{\overline{a} \prec \overline{c}_0-\overline{b}} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}}\right)
\end{aligned}$$

as $\alpha_{\overline{a},\overline{b}} = 0$ for $\overline{a} \succ \overline{c}_0 - \overline{b}$,

$$\begin{aligned}
&= \mathrm{LM}\left(\alpha_{\overline{c}_0-\overline{b},\overline{b}} \overline{x}^{\overline{c}_0-\overline{b}} + \sum_{\overline{a} \prec \overline{c}_0-\overline{b}} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}}\right) \\
&= \overline{x}^{\overline{c}_0-\overline{b}} \,,
\end{aligned}$$

where in the last step we used that $\alpha_{\overline{c}_0-\overline{b},\overline{b}} \neq 0$. This establishes the desired equality. □

We now relate the extremal monomials of the coefficient space of f to the monomials of the coefficient space of the extremal diagonals of f.

Corollary 6.14. *Let $f \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots, y_n]$. Then*

$$\mathrm{LM}(\mathbf{Coeff}_{\overline{x}|\overline{y}}(f)) \supseteq \mathrm{LM}(\mathbf{Coeff}_{\overline{x}|\overline{y}}(\mathrm{LD}(f))) \,,$$

and likewise for trailing monomials.

Proof: This follows as $\mathrm{LM}(\mathbf{Coeff}_{\overline{x}|\overline{y}}(\mathrm{LD}(f)))$ is equal to $\{\mathrm{LM}(\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(\mathrm{LD}(f)))|\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(\mathrm{LD}(f)) \neq 0\}$, but by Theorem 6.13 this set equals $\{\mathrm{LM}(\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(f))|\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(\mathrm{LD}(f)) \neq 0\}$, which is clearly contained in $\mathrm{LM}(\mathbf{Coeff}_{\overline{x}|\overline{y}}(f))$. □

We now observe that the number of leading monomials of the coefficient space of a leading diagonal is equal to its sparsity.

Lemma 6.15. *Let $f \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots, y_n]$. For a polynomial g, let $|g|_0$ denotes its sparsity. Then*

$$|\,\mathrm{LM}(\mathbf{Coeff}_{\overline{x}|\overline{y}}(\mathrm{LD}(f)))| = |\,\mathrm{LD}(f)|_0 \,,$$

where the respective statement also holds for trailing monomials.

Proof: We prove the claim for the leading diagonal, the trailing statement is symmetric. As above, let $f = \sum_{\overline{a},\overline{b}} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}} \overline{y}^{\overline{b}}$ so that $\mathrm{LD}(f) = \sum_{\overline{a}+\overline{b}=\overline{c}_0} \alpha_{\overline{a},\overline{b}} \overline{x}^{\overline{a}} \overline{y}^{\overline{b}} = \sum_{\overline{b}} \alpha_{\overline{c}_0-\overline{b},\overline{b}} \overline{x}^{\overline{c}_0-\overline{b}} \overline{y}^{\overline{b}}$ for some $\overline{c}_0 \in \mathbb{N}^n$. Then $\mathrm{Coeff}_{\overline{x}|\overline{y}^{\overline{b}}}(\mathrm{LD}(f)) = \alpha_{\overline{c}_0-\overline{b},\overline{b}} \overline{x}^{\overline{c}_0-\overline{b}}$. As the monomials $\overline{x}^{\overline{c}_0-\overline{b}}$ are linearly independent for distinct $\overline{b}$, it follows that $\dim \mathbf{Coeff}_{\overline{x}|\overline{y}}(\mathrm{LD}(f)) = |\{\overline{b} | \alpha_{\overline{c}_0-\overline{b},\overline{b}} \neq 0\}|$, which is equal the sparsity $|\mathrm{LD}(f)|_0$. □

Finally, we now lower bound the coefficient dimension of a polynomial by the sparsity of its extremal diagonals.

Corollary 6.16. *Let $f \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots, y_n]$. Then*

$$\dim \mathbf{Coeff}_{\overline{x}|\overline{y}}(f) \geq |\mathrm{LD}(f)|_0, |\mathrm{TD}(f)|_0 \ ,$$

where for a polynomial g, $|g|_0$ denotes its sparsity.

Proof: Combining Theorem 6.14 with Theorem 6.15 we get

$$\dim \mathbf{Coeff}_{\overline{x}|\overline{y}}(f) \geq \dim \mathbf{Coeff}_{\overline{x}|\overline{y}}(\mathrm{LD}(f)) \geq |\mathrm{LD}(f)|_0, |\mathrm{TD}(f)|_0 \ .$$

□

6.6 Lower Bounds for Multiples for Read-Once and Read-Twice ABPs

Having developed the theory of leading diagonals in the previous section, we now turn to using this theory to obtain explicit polynomials whose non-zero multiples all require large roABPs and read-twice ABPs. As $\sum \bigwedge \sum$ formulas and sparse polynomials have small roABPs, these polynomials will also have multiples requiring large complexity in these models as well and thus qualitatively reproving some of the above results in this section. However, we included the previous sections as the hard polynomials there are simpler (being monomials or translations of monomials), and more importantly we will need those results for the proofs below.

The proofs will use the characterization of roABPs by their coefficient dimension (Theorem 3.2), the lower bound for coefficient dimension in terms of the sparsity of the extremal diagonals (Theorem 6.16), and polynomials whose multiples are all non-sparse (Theorem 6.10).

Proposition 6.17. *Let $f(\overline{x}, \overline{y}) := \prod_{i=1}^{n} (x_i + y_i + \alpha_i x_i y_i) \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots, y_n]$, for any $\alpha_i \in \mathbb{F}$. Then for any $0 \neq g \in \mathbb{F}[\overline{x}, \overline{y}]$,*

$$\dim \mathbf{Coeff}_{\overline{x}|\overline{y}}(g \cdot f) \geq 2^n \ .$$

In particular, all non-zero multiples of f require width at least 2^n to be computed by a roABP in any variable order where $\overline{x} \prec \overline{y}$.

Proof: Observe that trailing diagonal of f is insensitive to the α_i. That is, $\mathrm{TD}(x_i+y_i+\alpha_i x_i y_i) = x_i + y_i$, so by multiplicativity of the trailing diagonal (Theorem 6.12) we have that $\mathrm{TD}(f) = \prod_i (x_i + y_i)$. Thus, appealing to Theorem 6.16 and Theorem 6.10,

$$\begin{aligned} \dim \mathbf{Coeff}_{\overline{x}|\overline{y}}(g \cdot f) &\geq |\mathrm{TD}(g \cdot f)|_0 \\ &= |\mathrm{TD}(g) \cdot \mathrm{TD}(f)|_0 \\ &= |\mathrm{TD}(g) \cdot \prod_i (x_i + y_i)|_0 \\ &\geq 2^n \ . \end{aligned}$$

The claim about roABP width follows from Theorem 3.2. □

Note that this lower bound works in the "monotone" setting, as we only used the zero/non-zero pattern of the coefficients.

Corollary 6.18. *Let $f(\overline{x},\overline{y}) := \prod_{i=1}^{n}(x_i + y_i + x_i y_i) \in \mathbb{F}[x_1,\ldots,x_n,y_1,\ldots,y_n]$. Then, any roABP for $g \cdot f$, for any $0 \neq g \in \mathbb{F}[\overline{x},\overline{y}]$, has width $\geq 2^n$.*

Proof: The proof follows from combining Theorem 6.17 with Theorem 3.2. □

To obtain a lower bound for read-twice ABPs we need the following theorem of Anderson et al. [AFS+15].

Theorem 6.19 (Theorem 4.3 of [AFS+15])**.** *Let $f \in \mathbb{F}[x_1,\ldots,x_n]$ be a polynomial computed by width-w read-k oblivious ABP. Then, there exist three disjoint subsets $U \sqcup V \sqcup W = [n]$, such that*

1. *$|U|, |V| \geq n/k^{O(k)}$,*
2. *$|W| \leq n/10$, and*
3. *$\dim \mathbf{Eval}_{\overline{x}_U|\overline{x}_V,\mathbb{F}}(f) \leq w^{2k}$, where we compute the dimension over the field $\mathbb{F}(\overline{x}_W)$, and by $\overline{x}_U$ (similarly, $\overline{x}_V, \overline{x}_W$) we mean the variables in $\overline{x}$ whose indices belong to U.*

Corollary 6.20. *Let $f(\overline{x},\overline{z}) := \prod_{i,j=1}^{n}(z_{i,j}\cdot(x_i+x_j+x_ix_j)+(1-z_{i,j})) \in \mathbb{F}[x_1,\ldots,x_n,z_{1,1},\ldots,z_{n,n}]$. Then, any read-2 oblivious ABP for $g \cdot f$, for any $0 \neq g \in \mathbb{F}[\overline{x},\overline{z}]$, has width $\geq \exp(n)$.*

Proof: Consider a read-2 oblivious ABP for $g \cdot f$ of width w. Let us think of the variable $\overline{z}$ as being part of the field. By Theorem 6.19 (for $k = 2$) we can partition to variable $\overline{x}$ to three sets $\overline{x}_U, \overline{x}_V$ and $\overline{x}_W$ such that when we add $\overline{x}_W$ to the field, $\dim \mathbf{Eval}_{\overline{x}_U|\overline{x}_V,\mathbb{F}}(f) \leq w^{2k}$. To simplify notation let us assume that $|U| \leq |V|$ and that $U = \{1,3,5,\ldots,2|U|-1\}$ and $V = \{2,4,6,\ldots,2|V|\}$. Then we can express f as $f = \prod_{i=1}^{|U|}(z_{2i-1,2i}(x_{2i-1} + x_{2i} + x_{2i-1}\cdot x_{2i}) + (1 - z_{2i-1,2i}))\cdot f'$. In particular,

$$f \cdot g = \prod_{i=1}^{|U|}(z_{2i-1,2i}(x_{2i-1} + x_{2i} + x_{2i-1}\cdot x_{2i}) + (1 - z_{2i-1,2i}))\cdot g' \,,$$

for some nonzero $g'(\overline{x},\overline{z})$. From Theorem 6.17 (or, in fact, a small variation of it) we get that $\dim \mathbf{Coeff}_{\overline{x}_U|\overline{x}_V}(g \cdot f) \geq 2^{|U|}$. Combining with Theorem 6.18 we get $w^4 \geq 2^{|U|}$, which implies $w = 2^{\Omega(|U|)} = 2^{\Omega(n)}$. □

7 IPS Lower Bounds via Lower Bounds for Multiples

In this section we use the lower bounds for multiples of Section 6 to derive lower bounds for $\mathcal{C}$-IPS proofs for various restricted algebraic circuit classes $\mathcal{C}$. While we consider most of the same classes for which we proved linear-IPS lower bounds in Section 5, our simulation of IPS by linear-IPS (Theorem 4.1) does not work for these restricted classes and thus we need further ideas to obtain IPS lower bounds. Unfortunately, as discussed in the introduction (subsubsection 1.3.3), this approach will necessarily give lower bounds for $\mathcal{C}$-IPS where the axioms themselves are hard to compute within $\mathcal{C}$ (but may be computable by general algebraic circuits).

7.1 IPS Lower Bounds for Depth-3 Powering Formulas

We begin by proving lower bounds for IPS proofs written as depth-3 powering formulas.

Proposition 7.1. *Define $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ by $f := x_1 \cdots x_n - 1$ and $g = x_1 + \cdots + x_n - m$, for $m \in \mathbb{F}$. Then for $m \neq n$, the equations $f, g, \overline{x}^2 - \overline{x}$ are unsatisfiable, and any $\sum\bigwedge\sum$-IPS refutation must be of size $\exp(\Omega(n))$.*

Proof: To see that the equations $f, g, \overline{x}^2 - \overline{x}$ are unsatisfiable note that in order to satisfy f, all variables must be set to 1. This assignment however does not satisfy g.

We now move to proving the lower bound. By definition, an IPS refutation gives rise to a computation of the form

$$1 = C(f, g, B_1, \ldots, B_n, \overline{x})$$

where C is a $\sum\bigwedge\sum$ formula. In particular, C computes a polynomial of the form

$$C(f, g, B_1, \ldots, B_n, \overline{x}) = \sum_{i=1}^{s} \left(\alpha_{i,1} \cdot f + \alpha_{i,2} \cdot g + \sum_{j=1}^{n} \beta_{i,j} B_j + \sum_{t=1}^{n} \gamma_t \cdot x_t \right)^{d_i},$$

where the α, β, γ are constants in the field. As g is a linear polynomial and and the B_j are quadratic we can rewrite C as

$$C(f, g, B_1, \ldots, B_n, \overline{x}) = \sum_{i=1}^{s} (\alpha_i \cdot \prod_{i=1}^{n} x_i + Q_i)^{d_i},$$

where Q_i are quadratic polynomials. Thus, $1 = \sum_{i=1}^{s} (\alpha_i \cdot \prod_{i=1}^{n} x_i + Q_i)^{d_i}$. We now observe that this implies that there exists some nonzero polynomial h such that

$$h(\overline{x}) \cdot \prod_{i=1}^{n} x_i = 1 - \sum_{i=1}^{s} Q_i^{d_i}. \tag{3}$$

Indeed, to prove that h is nonzero it is enough to notice that if h was zero then the equality $1 = C(f, g, B_1, \ldots, B_n, \overline{x})$ would not depend on f and thus replacing it with, say 0, would make the set of initial polynomials satisfiable, thus contradicting the fact that we have a refutation.

4 implies that the leading monomial of the RHS in (3) contains at most $O(\log s)$ many variables. This implies $s = \exp(n)$, which is what we wanted to prove. □

7.2 IPS Lower Bounds for roABPs

We now prove lower bounds for roABP-IPS. The argument is similar to the previous section and is based on Theorem 6.20. Let $f = 1 + \prod_{i,j}(z_{i,j}(x_i + x_j - x_i x_j) + (1 - z_{i,j}))$. It is easy to see that over the boolean cube f is never zero, this is it never satisfiable.

Proposition 7.2. *Let $f = 1 + \prod_{i,j}(z_{i,j}(x_i + x_j - x_i x_j) + (1 - z_{i,j}))$. Then, the polynomial set $\{f, \overline{x}^2 - \overline{x}\}$ is not satisfiable, and any roABP-IPS refutation must be of width $\exp(\Omega(n))$.*

Proof: Let $C(\overline{x}, \overline{z}, y, \overline{v})$ be a roABP of width w so that $C(\overline{x}, 0, \overline{0}) = 0$ and $C(\overline{x}, \overline{z}, f, x_1^2 - x_1, \ldots, z_{n,n}^2 - z_{n,n}) = 1$. Let $C'(\overline{x}, \overline{z}, \overline{v}) \triangleq C(\overline{x}, \overline{z}, 0, \overline{v})$. Clearly C' is a roABP, and its width is at most w. It is also immediate that

$$C(\overline{x}, \overline{z}, y, \overline{v}) = C'(\overline{x}, \overline{z}, \overline{v}) + y \cdot C''(\overline{x}, \overline{z}, \overline{v}) \ ,$$

for some polynomial $g(\overline{x}, \overline{z}, \overline{v})$. Substituting the boolean axioms to the variables $\overline{v}$ and f to y we get

$$1 = C(\overline{x}, \overline{z}, f, x_1^2 - x_1, \ldots, z_{n,n}^2 - z_{n,n}) = C'(\overline{x}, \overline{z}, x_1^2 - x_1, \ldots, z_{n,n}^2 - z_{n,n}) + f \cdot g' \ ,$$

for some polynomial $g'(\overline{x},\overline{z})$. Note that $g' \neq 0$ as otherwise we will get a contradiction by setting all variables to zero using $C(\overline{x},0,\overline{0}) = 0$. By rearranging we get that $f \cdot g' = 1 - C'(\overline{x},\overline{z},x_1^2 - x_1,\ldots,z_{n,n}^2 - z_{n,n})$. We now observe that as C' is read-once in the variables $\overline{x},\overline{z}$, what we get after substituting the boolean axioms to $\overline{v}$ is a read-2 oblivious ABP in the variables $\overline{x},\overline{z}$. Corollary 6.20 implies that $w = \exp(n)$ as claimed. □

8 The Relative Strength of IPS Fragments

Here we further study the relative strength of the IPS fragments considered in previous sections, comparing them to several related propositional proof systems studied in proof complexity literature. We have already seen that our fragments of IPS (excluding depth-3 powering formulas IPS) sit strictly above the Nullstellensatz refutation system (measured by sparsity). Here we show that our fragments are polynomially simulated by semantic *tree-like* versions of the Polynomial Calculus in which proof-lines are written with either roABP or multilinear formulas, respectively, as considered in [RT08, Tza11]. We further observe that both roABP-IPS and multilinear-formulas-IPS have exponential speedups over PC, from which we derive new separations between tree-like PC over multilinear-formulas and roABPs, and PC (measured by sparsity). We discuss the relations between fragments of IPS to Frege systems, as well as to the non-commutative-IPS proof system that efficiently simulates Frege.

We start by a brief formal summary of basic notions from propositional proof complexity.

8.1 Basic Concepts in Propositional Proof Complexity

A *propositional proof system*, sometimes called a *Cook-Reckhow proof system* following the definition given in [CR79], is simply a polynomial-time function f from a set of finite strings over some given alphabet ***onto*** the set of propositional tautologies. Thus, $f(x) = y$ means that the string x is the proof of the tautology y. Note that since f is onto, all tautologies and only tautologies have proofs (and thus the proof system is complete and sound). The idea behind this definition is that a purported proof x may be much longer then the tautology y it proves, but given the proof it should be possible to efficiently check (efficient with respect to the proof length) that the proof is indeed a correct proof of the tautology.

This definition of a Cook-Reckhow proof system encompasses most standard proof systems for propositional tautologies, such as resolution and Frege proofs. We can also relax the verification of propositional proofs to be efficient in probabilistic polynomial-time, namely the proof system f above is now assumed only to be in BPP. This relaxation is relevant to fragments of IPS, and it was already considered before in several works in proof complexity (cf. [Pit97, RT08, GP14]).

Considering a BPP verifiable propositional proof system is frequent when dealing with algebraic proof systems, in which verification of proofs naturally depends on a polynomial identity testing (PIT) algorithm for arithmetic circuits. For general arithmetic circuits only a BPP (in fact, coRP) efficient PIT is currently known [Zip79, Sch80]). But for certain restricted circuit classes, such as roABPs [RS05] (and trivially, ΣΠ formulas) there is a deterministic polynomial time PIT algorithm.

8.1.1 Simulations

For the purpose of comparing the relative complexity of different proof systems we have the concept of a ***simulation***. Concrete examples of simulation results where given in this work above. Formally, we say that a propositional proof system P *polynomially simulates* another propositional proof system Q if there is a polynomial-time computable function f that maps Q-proofs to P-proofs of

the same tautologies (if P and Q use different representations for tautologies, we fix a (polynomial) translation from one representation to the other). In case f is computable in time $t(n)$ (for n the input-size), we say that P $t(n)$-*simulates* Q. We say that P and Q are *polynomially equivalent* in case P polynomially simulates Q and Q polynomially simulates P. If P polynomially simulates Q but Q does not polynomially simulates P we say that P *is strictly stronger than* Q (equivalently, that Q *is strictly weaker than* P), and we also say that P *is separated from* Q.

8.2 Relations with Polynomial Calculus

Most propositional proof systems considered in proof complexity are so-called *sequential proof systems*, sometimes also termed *dynamic* proof system to contrast them from "static" ones. In sequential proof systems, a proof is a sequence of formulas, each derived by a set of finite derivation rules from previous formulas. Each formula in a proof-sequence is referred to as a *proof-line*.

The Polynomial Calculus (introduced by Clegg et al. [CEI96]) and its variants, are sequential proof systems for finite collections of polynomial equations having no 0-1 solutions over a given field.[12]

Definition 8.1 (**Polynomial Calculus (PC)**). *Let $\mathbb{F}$ be a field and let $F = \{f_1, \ldots, f_m\}$ be a collection of multivariate polynomials from $\mathbb{F}[x_1, \ldots, x_n]$. A* PC proof from Q of a polynomial g *is a finite sequence $\pi = (p_1, ..., p_\ell)$ of multivariate polynomials from $\mathbb{F}[x_1, \ldots, x_n]$, where $p_\ell = g$ and for every $1 \leq i \leq \ell$, either $p_i = f_j$ for some $j \in [m]$, or p_i is a Boolean axiom $x_i \cdot (1 - x_i)$ for some $i \in [n]$, or p_i was deduced from p_j, p_k, for $j, k < i$, by one of the following inference rules:*

(i) product rule*: from p, derive $g \cdot p$, for g any polynomial in $\mathbb{F}[x_1, \ldots, x_n]$;*

(ii) addition rule*: from p, derive $ap + bq$, for $a, b \in \mathbb{F}$.*

A **PC refutation** of F *is a proof of* 1 *(which is interpreted as $1 = 0$, that is the unsatisfiable equation standing for* false*) from F. The* degree *of a PC-proof is the maximal degree of a polynomial in the proof. The* ***size*** *of a* PC *proof π is the total number of monomials (with nonzero coefficients) in all the proof-lines, denoted $|\pi|$.*

We say that a PC proof is a ***tree-like*** proof if the underlying proof-structure is a tree, or in other words, if every proof-line can be used *at most once as the premise of an inference rule*.

Note: It is important to notice that the product rule for PC we use in Definition 8.1 is slightly different from the standard product rule [CEI96] which assumes that g *is a variable*. For PC proofs in which polynomials are written as sums of monomials or as algebraic *formulas*, both definitions are polynomially equivalent to each other (cf. [RT08]); this is because we can polynomially simulate a product by g, by a straightforward induction on the formula size of g. However, if we consider *tree-like* PC proofs, we *cannot* in general polynomially simulate the rule "from f derive $f \cdot g$" using the rule "from f derive $x_i \cdot f$", since in a tree-like PC proof the latter rule would amount to multiplying f by *all* the monomials in g, *one by one*. Because we are going to consider tree-like PC proofs, we therefore define PC with the more general rule "from f derive $f \cdot g$".

Notice also that the size of PC proofs in Definition 8.1 can be defined equivalently (up to a factor of n) as the total formula size of all proof-lines, where polynomials are written as sums of monomials and equivalently as $\Sigma\Pi$ formulas (Definition 1.2). We will consider in what follows PC proofs in which proof-lines are taken from different algebraic circuit classes.

Clegg et al. [CEI96] proved that every PC proof of the subset sum principle $\sum_{i=1}^{n} x_i = m$ (for $m > n$) over the reals cannot have subexponential size (namely, it must contain exponentially many monomials). Thus, by Theorem 4.8 and Theorem 4.9:

[12] Formally, each different field yields a different algebraic proof system.

Corollary 8.1. *PC (over the reals) does not polynomially simulate neither roABP-IPS$_{LIN}$ nor multilinear-formulas-IPS$_{LIN}$.*

We leave it open whether either roABP-IPS$_{\text{LIN}}$ or multilinear-formulas-IPS$_{\text{LIN}}$ polynomially simulate PC.

8.3 Relations with PC over roABPs

Tzameret [Tza11] considered the strength of polynomial calculus where polynomials are written as roABPs. This work focused on *ordered formulas* and defined the system OFPC, standing for *ordered formulas PC*. Ordered formulas is the formula class corresponding to roABPs, similar to the way that non-commutative formulas correspond to non-commutative ABPs. More precisely, an ordered formula is a non-commutative formula in which every node computes a polynomial such that the multiplication order of variables in each monomial respects a fixed total order on the variables (e.g., ascending order $x_1 < x_2 < \ldots < x_n$). Nevertheless, as discussed in [Tza11] *all results in that paper also hold for roABPs.*

Definition 8.2 (PC over roABP)**.** PC over roABPs *is the semantic version of PC where proof-size is defined to be the total roABPs size of the polynomials in the proof (for some fixed total order on variables).*

Note that we say that PC over roABP is *semantic* to mean that a polynomial p in a proof can be written as any roABP that computes that polynomial. Though the proof system is semantic in this sense, it still polynomially verifiable (i.e., it is a Cook-Reckhow proof system), because roABPs have an efficient PIT algorithm (due to [RS05]; see [Tza11] for details).

The following is an easy observation:

Corollary 8.2. *Tree-like PC over roABP polynomially simulates roABP-IPS$_{LIN}$.*

Proof: Let $P = \sum_{i=1}^{m} p_i(\mathbf{x}) \cdot f_i(\mathbf{x}) = q(\mathbf{x})$ be an roABP-IPS$_{\text{LIN}}$ proof of the polynomial q from the initial polynomials $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$. Our desired tree-like PC over roABPs proof of q simply adds one by one the products of the initial polynomials f_i by p_i, using the product and addition rules of PC. □

From Theorem 8.1 we can conclude the following:

Corollary 8.3. *Tree-like PC over roABPs has an exponential speed-up over PC (over the reals).*

Previously, only a speed-up between (dag-like) PC over roABPs and PC was known [Tza11].

8.4 Relations with Non-Commutative-IPS and Frege

Li et al. [LTW15] studied $\mathcal{C}$-IPS for $\mathcal{C}$ the set of *non-commutative formulas*. They showed that $\mathcal{C}$-IPS, using non-commutative polynomials over finite fields or the rational numbers, and adding the commutator $x_ix_j - x_jx_i$ for any pair of variables as an axiom (for the sake of completeness), polynomially simulates Frege propositional proofs. They also showed that Frege can quasipolynomially simulate $\mathcal{C}$-IPS over $GF(2)$. Since the class of non-commutative ABPs (Definition 1.5) is at least as strong as non-commutative formulas, non-commutative-ABP-IPS (Definition 8.3) polynomially simulates Frege proofs as well.

Definition 8.3 (Non-commutative IPS)**.** *Let $\mathbb{F}$ be a field. Assume that $f_1(\overline{x}) = f_2(\overline{x}) = \cdots = f_m(\overline{x}) = 0$ is a system of non-commutative polynomial equations from the ring of non-commutative polynomials denoted $\mathbb{F}\langle x_1, \ldots, x_n\rangle$, and suppose that the following set of equations (axioms) are included in the $f_i(\overline{x})$'s: (i) $x_i \cdot (1 - x_i)$, for all $1 \leq i \leq n$; (ii) $x_i \cdot x_j - x_j \cdot x_i$, for all $1 \leq i < j \leq n$. Suppose that the $f_i(\overline{x})$'s have no common 0-1 solutions.*[13] *A* non-commutative-IPS refutation *that the system of $f_i(\overline{x})$'s is unsatisfiable is a non-commutative polynomial $\mathfrak{F}(\overline{x}, \overline{y})$ in the variables $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ (i.e. $\mathfrak{F} \in \mathbb{F}\langle \overline{x}, \overline{y}\rangle$), such that:*

1. $\mathfrak{F}(x_1, \ldots, x_n, \overline{0}) = 0$
2. $\mathfrak{F}(x_1, \ldots, x_n, F_1(\overline{x}), \ldots, F_m(\overline{x})) = 1$.

We can assume that non-commutative-IPS refutations are written as a *non-commutative-ABPs* (namely, an ABP as in Definition 1.5, in which the order of multiplication along a path in the ABP corresponds to non-commutative multiplication; note that a non-commutative ABP may not be a read once ABP, since the multiplication of variables can be done in every possible order). In this case, the *size* of a non-commutative IPS refutation is the minimal size of a non-commutative-ABP computing the non-commutative-IPS refutation.

We have the following:

Corollary 8.4. *Non-commutative-ABP-IPS polynomially simulates roABP-IPS (and hence, also roABP-IPS$_{LIN}$).*

Proof: This is almost immediate from the definitions. We observe that there is no need to use the commutator axioms $x_i x_j - x_j x_i$ in roABP-IPS.

Indeed, notice that the polynomial computed by an roABP, considered as a *non-commutative*-ABP, is a non-commutative polynomial in which the product in every monomial respects the *same* total order on variables. Thus, every roABP-IPS refutation is also a non-commutative ABP-IPS (since there is no need to use the commutator axioms, as all products are ordered according to the same fixed total order on variables). □

This shows that getting rid of the 'read-once' restriction in the roABP-IPS$_{\text{LIN}}$ lower bound (Theorem 7.2) is very close to proving super-polynomial Frege lower bounds.

8.5 Relations with PC over Multilinear Formulas

We recall the concept of *multilinear proofs* introduced by Raz and Tzameret in [RT08], denoted fMC (formula multilinear calculus). Essentially, fMC is a semantic variant of PC in which proof-lines are written as multilinear formulas (instead of by sparse representation). However, to guarantee that every PC proof-line is indeed a *multilinear* polynomial one needs *to add new formal variables* $\overline{x}_i$, that will stand for the "negation of x_i". Formally, to make $\overline{x}_i$ equal to the negation of x_i we add the axiom $x_i + \overline{x}_i - 1$, for each ("original") variable x_i (and so x_i is 0 iff $\overline{x}_i$ is 1).

Definition 8.4 (PC over multilinear formulas (fMC) [RT08])**.** *Fix a field $\mathbb{F}$ and let $F := \{f_1, \ldots, f_m\}$ be a collection of multilinear polynomials from $\mathbb{F}[x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n]$. Call the set of polynomials consisting of $x_i + \overline{x}_i - 1$ and $x_i \cdot \overline{x}_i$ for $1 \leq i \leq n$, the* Boolean axioms of fMC*. An* fMC proof from F of a polynomial g *is a sequence $\pi = (p_1, ..., p_\ell)$ of multilinear polynomials from $\mathbb{F}[x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n]$, such that $p_\ell = g$, and for each $i \in [\ell]$, either $p_i = f_j$ for some $j \in [m]$, or p_i is a Boolean axiom of* fMC*, or p_i was deduced by one of the following inference rules using p_j, p_k for $j, k < i$:*

[13] One can check that the $f_i(\overline{x})$'s have no common 0-1 solutions in $\mathbb{F}$ iff they do not have a common 0-1 solution in every $\mathbb{F}$-algebra.

(i) Product rule: *from p deduce $q \cdot p$, for some polynomial $q \in \mathbb{F}[x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n]$* such that $p \cdot q$ is multilinear*;*

(ii) Addition rule*: from p, q deduce $\alpha \cdot p + \beta \cdot q$, for some $\alpha, \beta \in \mathbb{F}$.*

All the polynomials in an fMC proof are ***written as multilinear formulas****. An* fMC refutation *of F is a proof of 1 from F. The* size *of an* fMC *proof π is defined as the total sum of all the formula sizes in π and is denoted by $|\pi|$.*

Note, once more, that the fMC system is a ***semantic*** proof system, namely in every proof-line computing the multilinear polynomial f we can choose to write any multilinear formula that computes f.

[RT08] showed that fMC is strictly stronger than PC (as well as PC with resolution, denoted PCR), and can refute efficiently the pigeonhole principle and the graph Tseitin's formulas, among other principles. Here we show that *tree-like* fMC polynomially simulates multilinear-formula-IPS_{LIN} (for the language of unsatisfiable systems of low degree polynomials). Since a proof in multilinear-formulas-IPS_{LIN} may be a sum of polynomials that are not necessarily multilinear we need to simulate the proof using the negative variable $\overline{\mathbf{x}}$ in fMC.

Theorem 8.5. *Let $\mathcal{C}$ be the class of multilinear formulas and let F be an unsatisfiable system of multilinear polynomials of degree at most d. If there is a $\mathcal{C}$-IPS_{LIN} refutation of F with size s, then there exists a $\mathsf{poly}(2^d, s)$ tree-like fMC refutation of F. In particular, when $d = O(\log s)$ the tree-like fMC refutation is polynomial in s.*

Note that standard polynomial translations of unsatisfiable 3CNFs, as well as the subset-sum principle (and their extensions) considered in previous sections, are all multilinear and of constant degree.

Proof: Let $P = \sum_{i=1}^{m} p_i(\mathbf{x}) \cdot f_i(\mathbf{x}) + \sum_{i=1}^{n} h_i(\mathbf{x}) \cdot (x_i^2 - x_i) = 1$ be a multilinear-formula-IPS_{LIN} refutation from the initial polynomials $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ (for the sake of convenience we assume the Boolean axioms are not included in the f_i's. We also assume for simplicity that all $p_i's$ are nonzero). We will use the following claim:

Claim 8.6. *For every $i \in [m]$ and every set $I \subseteq [n]$ with $|I| = d$, $p_i = \sum_{J \subseteq I}(p_J \cdot \prod_{j \in J} x_j)$, where each summand in this sum is multilinear and can be computed by a multilinear formula of size at most $\mathsf{poly}(2^d, s)$.*

Proof of claim. This is proved by "factoring out" from p_i all the monomials determined by the indices in I. This is done by using the (interpolation) formula $g(x_j) = x_j \cdot (g(1) - g(0)) - g(0)$, inductively on the size of I. □

The tree-like fMC refutation is constructed as follows. Recall that in fMC we have the "negative" variables $\overline{x}_i$ that equal $1 - x_i$ using the Boolean axioms $x_i + \overline{x}_i - 1$, as well as the other Boolean axioms $x_i\overline{x}_i$. We denote the set of negative variables by $\overline{\mathbf{x}}$. Recall also the fMC is a semantic proof system, and so any multilinear polynomial can be written in any desired way.

Let $p_i'(\overline{\mathbf{x}})$ be $p_i(\mathbf{x})$ in which every occurrence of the variables $x_1, \ldots, x_n$ is substituted by $(1 - \overline{x}_1), \ldots, (1 - \overline{x}_n)$, respectively (we shall sometimes suppress the explicit mention of the variables $\overline{\mathbf{x}}$ and $\mathbf{x}$ in what follows). We start the proof with $p_i' \cdot f_i$, for all $i \in [m]$ (note that indeed these products are multilinear).

Since the degree of the multilinear polynomial f_i is at most d, we can write f_i as a sum of $2^{O(d)}$ monomials. Thus, $p_i' \cdot f_i = \sum_k (p_i' \cdot M_k)$, where M_k are the monomials in f_i.

For a monomial M, denote by $I_M \subseteq [n]$ the indices of the variables in M. By the claim above (when we substitute each variable x_i by $(1-\overline{x}_i)$), for any $I \subseteq [n]$ there exists $P'_J s$ such that $p'_i = \sum_{J \subseteq I} \prod_{j \in J}(1-\overline{x}_j) \cdot p'_J$, where each summand is multilinear over the variables $\overline{\mathbf{x}}$, and can be computed by a multilinear formula of size at most $\mathsf{poly}(2^d, s)$. We thus get

$$p'_i(\overline{\mathbf{x}}) \cdot f_i(\mathbf{x}) = \sum_k (p'_i \cdot M_k) = \sum_k M_k \cdot \sum_{J \subseteq I_{M_k}} \prod_{j \in J}(1-\overline{x}_j) \cdot p'_J \tag{4}$$

(the p'_J's depend of course on M_k).

Since fMC is a semantic proof system, we can actually write $p'_i \cdot f_i$ as the right hand side of (4). Note that in the right hand side of (4) there may be variables x_j in M_k that appear as $(1-\overline{x}_j)$ in $\prod_{j \in J}(1-\overline{x}_j)$ (but not in p'_J). Using the axioms $x_i \cdot \overline{x}_i$ we can derive (for each k that appears in (4))

$$\mathrm{ml}(M_k \cdot \prod_{j \in J} x_j) - M_k \cdot \prod_{j \in J}(1-\overline{x}_j)\,,$$

with a tree-like fMC proof of size $\mathsf{poly}(\deg(M_k), d)$. Thus, using substitutions in $\sum_k M_k \cdot \sum_{J \subseteq I} \prod_{j \in J}(1-\overline{x}_j) \cdot p'_J$ in (4) we can derive, with a linear size tree-like fMC proof (linear in the size of the term in which we substitute),

$$\sum_k \sum_{J \subseteq I} \mathrm{ml}\left(M_k \cdot \prod\nolimits_{j \in J} x_j\right) \cdot p'_J\,.$$

By using the axiom $x_i + \overline{x}_i - 1$ we can substitute in the last term every occurrence of $(1-\overline{x}_i)$ by x_i in p'_J. This also is doable with a tree-like fMC proof of polynomial-size in the formula in which we substitute. Notice that by construction the variables in $M_k \cdot \prod_{j \in J} x_j$ are disjoint from those in p_J. We thus derive

$$\begin{aligned}\sum_k \sum_{J \subseteq I} \mathrm{ml}\left(M_k \cdot \prod\nolimits_{j \in J} x_j\right) \cdot p_J(\mathbf{x}) &= \sum_k \sum_{J \subseteq I} \mathrm{ml}\left(M_k \cdot \prod\nolimits_{j \in J} x_j \cdot p_J(\mathbf{x})\right) \\ &= \mathrm{ml}\left(p_i(\mathbf{x}) \cdot f_i(\mathbf{x})\right)\,.\end{aligned}$$

Since the multilinearization of the Boolean axioms $x_i^2 - x_i$ is 0, we get

$$\begin{aligned}\mathrm{ml}(P) &= \mathrm{ml}\left(\sum_{i=1}^m p_i(\mathbf{x}) \cdot f_i(\mathbf{x})\right) + \mathrm{ml}\left(\sum_{i=1}^n h_i(\mathbf{x}) \cdot (x_i^2 - x_i)\right) \\ &= \mathrm{ml}\left(\sum_{i=1}^m p_i(\mathbf{x}) \cdot f_i(\mathbf{x})\right) = 1\,.\end{aligned}$$

Now, the tree-like fMC refutation of $f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})$ will simply add one by one the proof-lines $\mathrm{ml}\left(p_i(\mathbf{x}) \cdot f_i(\mathbf{x})\right)$ that we have derived above, until we reach the polynomial 1, concluding the refutation.

Note that the total size of the tree-like fMC refutation is $\mathsf{poly}(2^d, s)$, as each proof-line was of this size-order, and each of the derivations used only $\mathsf{poly}(2^d, s)$ many proof-lines. □

From Theorem 8.1 we conclude:

Corollary 8.7. *Tree-like fMC has an exponential speed-up over PC (over the reals).*

Previously, an exponential speed-up over PC was known only for (dag-like) fMC.

9 Open Problems

Open Problem 9.1. *Can the lower bounds on roABP-IPS$_{LIN}$ and multilinear-formulas-IPS$_{LIN}$ from Theorem 5.13 be extended to (tree-like or dag-like) PC over roABPs or PC over multilinear formulas fMC, respectively (Section 8)?*

Acknowledgments

We would like to thank Rafael Oliviera for telling us of Theorem 6.11, as well as Mrinal Kumar and Ramprasad Saptharishi for conversations [FKS15] clarifying the roles of functional lower bounds in this work. We would also like to thank Joshua Grochow for helpful discussions regarding this work.

References

[AFS+15] Matthew Anderson, Michael A. Forbes, Ramprasad Saptharishi, Amir Shpilka, and Ben Lee Volk. Identity testing and lower bounds for read-k oblivious algebraic branching programs. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:184, 2015.

[AGKS14] Manindra Agrawal, Rohit Gurjar, Arpita Korwar, and Nitin Saxena. Hitting-sets for ROABP and sum of set-multilinear circuits. *arXiv*, 1406.7535, 2014.

[Agr05] Manindra Agrawal. Proving lower bounds via pseudo-random generators. In *Proceedings of the 25th International Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2005)*, pages 92–105, 2005.

[AR01] Michael Alekhnovich and Alexander A. Razborov. Lower bounds for polynomial calculus: Non-binomial case. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2001)*, pages 190–199, 2001.

[ASS13] Manindra Agrawal, Chandan Saha, and Nitin Saxena. Quasi-polynomial hitting-set for set-depth-Δ formulas. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC 2013)*, pages 321–330, 2013. Full version at `arXiv:1209.2333`.

[AvMV11] Matthew Anderson, Dieter van Melkebeek, and Ilya Volkovich. Derandomizing polynomial identity testing for multilinear constant-read formulae. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity (CCC 2011)*, pages 273–282, 2011. Full version in the Electronic Colloquium on Computational Complexity (ECCC), Technical Report TR10-188.

[BGIP01] Samuel R. Buss, Dima Grigoriev, Russell Impagliazzo, and Toniann Pitassi. Linear gaps between degrees for the polynomial calculus modulo distinct primes. *JCSS*, 62(2):267–289, 2001. Preliminary version in the *14th Annual IEEE Conference on Computational Complexity (CCC 1999)*.

[BIK+96a] 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. Preliminary version in the *35th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1994)*.

[BIK+96b] Samuel R. Buss, Russell Impagliazzo, Jan Krajíček, Pavel Pudlák, Alexander A. Razborov, and Jiří Sgall. Proof complexity in algebraic systems and bounded depth Frege systems with modular counting. *Computational Complexity*, 6(3):256–298, 1996.

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

[CLO07] David Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer, New York, third edition, 2007. An introduction to computational algebraic geometry and commutative algebra.

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

[CR74b] Stephen A. Cook and Robert A. Reckhow. Corrections for "On the lengths of proofs in the propositional calculus (preliminary version)". *SIGACT News*, 6(3):15–22, July 1974.

[CR79] Stephen A. Cook and Robert A. Reckhow. The relative efficiency of propositional proof systems. *J. Symb. Log.*, 44(1):36–50, 1979. This is a journal-version of Cook-Reckhow [CR74a] and Reckhow [Rec76].

[DL78] Richard A. DeMillo and Richard J. Lipton. A probabilistic remark on algebraic program testing. *Inf. Process. Lett.*, 7(4):193–195, 1978.

[DSY09] Zeev Dvir, Amir Shpilka, and Amir Yehudayoff. Hardness-randomness tradeoffs for bounded depth arithmetic circuits. *SICOMP*, 39(4):1279–1293, 2009. Preliminary version in the *40th Annual ACM Symposium on Theory of Computing (STOC 2008)*.

[Fis94] Ismor Fischer. Sums of like powers of multivariate linear forms. *Mathematics Magazine*, 67(1):59–61, 1994.

[FKS15] Michael A. Forbes, Mrinal Kumar, and Ramprasad Saptharishi. Functional lower bounds for arithmetic circuits and boolean circuit complexity. Manuscript, 2015.

[For14] Michael A. Forbes. *Polynomial Identity Testing of Read-Once Oblivious Algebraic Branching Programs*. PhD thesis, Massachusetts Institute of Technology, June 2014.

[For15] Michael A. Forbes. Deterministic divisibility testing via shifted partial derivatives. In *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2015)*, 2015.

[FS12] Michael A. Forbes and Amir Shpilka. On identity testing of tensors, low-rank recovery and compressed sensing. In *Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC 2012)*, pages 163–172, 2012. Full version at `arXiv:1111.0663`.

[FS13a] Michael A. Forbes and Amir Shpilka. Explicit Noether Normalization for simultaneous conjugation via polynomial identity testing. In *Proceedings of the 17th International Workshop on Randomization and Computation (RANDOM 2013)*, pages 527–542, 2013. Full version at `arXiv:1303.0084`.

[FS13b] Michael A. Forbes and Amir Shpilka. Quasipolynomial-time identity testing of non-commutative and read-once oblivious algebraic branching programs. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2013)*, pages 243–252, 2013. Full version at `arXiv:1209.2408`.

[FSG13] Michal A. Forbes, Amir Shpilka, and Ankit Gupta. Personal Communication to Gupta, Kamath, Kayal, Saptharishi [GKKS13], 2013.

[FSS14] Michael A. Forbes, Ramprasad Saptharishi, and Amir Shpilka. Hitting sets for multilinear read-once algebraic branching programs, in any order. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC 2014)*, pages 867–875, 2014. Full version at `arXiv:1309.5668`.

[GK98] Dima Grigoriev and Marek Karpinski. An exponential lower bound for depth 3 arithmetic circuits. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC 1998)*, pages 577–582, 1998.

[GKKS13] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic circuits: A chasm at depth three. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2013)*, pages 578–587, 2013. Full version in the Electronic Colloquium on Computational Complexity (ECCC), Technical Report TR13-026.

[GKKS14] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the chasm at depth four. *JACM*, 61(6):33:1–33:16, December 2014. Preliminary version in the *28th Annual IEEE Conference on Computational Complexity (CCC 2013)*.

[GKST15] Rohit Gurjar, Arpita Korwar, Nitin Saxena, and Thomas Thierauf. Deterministic identity testing for sum of read once ABPs. In *Proceedings of the 30th Annual Computational Complexity Conference (CCC 2015)*, 2015. Full version at `arXiv:1411.7341`.

[GP14] Joshua A. Grochow and Toniann Pitassi. Circuit complexity, proof complexity, and polynomial identity testing. In *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2014)*, pages 110–119, 2014. Full version at `arXiv:abs/1404.3820`.

[GR00] Dima Grigoriev and Alexander A. Razborov. Exponential lower bounds for depth 3 arithmetic circuits in algebras of functions over finite fields. *Appl. Algebra Eng. Commun. Comput.*, 10(6):465–487, 2000. Preliminary version in the *39th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1998)*.

[Gri98] Dima Grigoriev. Tseitin's tautologies and lower bounds for Nullstellensatz proofs. In *Proceedings of the 39th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1998)*, pages 648–652, 1998.

[HS80] Joos Heintz and Claus-Peter Schnorr. Testing polynomials which are easy to compute (extended abstract). In *Proceedings of the 12th Annual ACM Symposium on Theory of Computing (STOC 1980)*, pages 262–272, 1980.

[IPS99] Russell Impagliazzo, Pavel Pudlák, and Jiří Sgall. Lower bounds for the polynomial calculus and the gröbner basis algorithm. *Computational Complexity*, 8(2):127–144, 1999.

[IW97] Russell Impagliazzo and Avi Wigderson. P=BPP if E requires exponential circuits: Derandomizing the XOR lemma. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing (STOC 1997)*, pages 220–229, 1997.

[Kal89] Erich L. Kaltofen. Factorization of polynomials given by straight-line programs. In Silvio Micali, editor, *Randomness and Computation*, volume 5 of *Advances in Computing Research*, pages 375–412. JAI Press, Inc., Greenwich, CT, USA, 1989.

[Kay08] Neeraj Kayal. Personal Communication to Saxena [Sax08], 2008.

[Kay12] Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 19(81), 2012.

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *Computational Complexity*, 13(1-2):1–46, 2004. Preliminary version in the *35th Annual ACM Symposium on Theory of Computing (STOC 2003)*.

[Kra95] Jan Krajíček. *Bounded arithmetic, propositional logic, and complexity theory*, volume 60 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995.

[KS01] Adam Klivans and Daniel A. Spielman. Randomness efficient identity testing of multivariate polynomials. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC 2001)*, pages 216–223, 2001.

[KS15] Mrinal Kumar and Ramprasad Saptharishi. An exponential lower bound for homogeneous depth-5 circuits over finite fields. *arXiv*, 1507.00177, 2015.

[LTW15] Fu Li, Iddo Tzameret, and Zhengyu Wang. Non-commutative formulas and Frege lower bounds: a new characterization of propositional proofs. In *Proceedings of the 30th Computational Complexity Conference (CCC), June 17-19, 2015*, 2015.

[MRS14] Meena Mahajan, B.V. Raghavendra Rao, and Karteek Sreenivasaiah. Building above read-once polynomials: Identity testing and hardness of representation. In *cocoon2014*, volume 8591 of *LNCS*, pages 1–12, 2014.

[Nis91] Noam Nisan. Lower bounds for non-commutative computation. In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing (STOC 1991)*, pages 410–418, 1991.

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.*, 49(2):149–167, 1994. Preliminary version in the *29th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1988)*.

[NW96] Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6(3):217–234, 1996. Preliminary version in the *36th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1995)*.

[Oli15a] Rafael Oliveira. Personal Communication, 2015.

[Oli15b] Rafael Oliveira. Factors of low individual degree polynomials. In *Proceedings of the 30th Annual Computational Complexity Conference (CCC 2015)*, volume 33 of *LIPIcs*, pages 198–216, 2015.

[OSV15] Rafael Oliveira, Amir Shpilka, and Ben Lee Volk. Subexponential size hitting sets for bounded depth multilinear formulas. In *Proceedings of the 30th Annual Computational Complexity Conference (CCC 2015)*, volume 33 of *LIPIcs*, pages 304–322, 2015. Full version at `arXiv:1411.7492`.

[Pit97] Toniann Pitassi. Algebraic propositional proof systems. In *Descriptive complexity and finite models (Princeton, NJ, 1996)*, volume 31 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 215–244. Amer. Math. Soc., Providence, RI, 1997.

[Pud97] Pavel Pudlák. Lower bounds for resolution and cutting plane proofs and monotone computations. *The Journal of Symbolic Logic*, 62(3):981–998, Sept. 1997.

[Raz98] Alexander A. Razborov. Lower bounds for the polynomial calculus. *Computational Complexity*, 7(4):291–324, 1998.

[Raz06] Ran Raz. Separation of multilinear circuit and formula size. *Theory of Computing*, 2(6):121–135, 2006. Preliminary version in the *45th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2004)*.

[Raz09] Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *J. ACM*, 56(2), 2009. Preliminary version in the *36th Annual ACM Symposium on Theory of Computing (STOC 2004)*.

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

[RS05] Ran Raz and Amir Shpilka. Deterministic polynomial identity testing in non-commutative models. *Comput. Complex.*, 14(1):1–19, April 2005. Preliminary version in the *19th Annual IEEE Conference on Computational Complexity (CCC 2004)*.

[RT08] Ran Raz and Iddo Tzameret. The strength of multilinear proofs. *Computational Complexity*, 17(3):407–457, 2008.

[RY08] Ran Raz and Amir Yehudayoff. Balancing syntactically multilinear arithmetic circuits. *Computational Complexity*, 17(4):515–535, 2008.

[RY09] Ran Raz and Amir Yehudayoff. Lower bounds and separations for constant depth multilinear circuits. *Computational Complexity*, 18(2):171–207, 2009. Preliminary version in the *23rd Annual IEEE Conference on Computational Complexity (CCC 2008)*.

[Sap12] Ramprasad Saptharishi, 2012. Personal communication to Forbes-Shpilka [FS13b].

[Sax08] Nitin Saxena. Diagonal circuit identity testing and lower bounds. In *Proceedings of the 35th International Colloquium on Automata, Languages and Programming (ICALP 2008)*, pages 60–71, 2008. Preliminary version in the Electronic Colloquium on Computational Complexity (ECCC), Technical Report TR07-124.

[Sch80] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, October 1980. Preliminary version in the *International Symposium on Symbolic and Algebraic Computation (EUROSAM 1979)*.

[Shp02] Amir Shpilka. Affine projections of symmetric polynomials. *JCSS*, 65(4):639–659, 2002. Preliminary version in the *16th Annual IEEE Conference on Computational Complexity (CCC 2001)*.

[Str73] Volker Strassen. Vermeidung von divisionen. *J. Reine Angew. Math.*, 264:184–202, 1973.

[STV01] Madhu Sudan, Luca Trevisan, and Salil P. Vadhan. Pseudorandom generators without the XOR lemma. *J. Comput. Syst. Sci.*, 62(2):236–266, 2001. Preliminary version in the *31st Annual ACM Symposium on Theory of Computing (STOC 1999)*.

[SV09] Amir Shpilka and Ilya Volkovich. Improved polynomial identity testing for read-once formulas. In *Proceedings of the 13th International Workshop on Randomization and Computation (RANDOM 2009)*, volume 5687 of *LNCS*, pages 700–713, 2009. Full version in the Electronic Colloquium on Computational Complexity (ECCC), Technical Report TR10-011.

[SW01] Amir Shpilka and Avi Wigderson. Depth-3 arithmetic circuits over fields of characteristic zero. *Computational Complexity*, 10(1):1–27, 2001. Preliminary version in the *14th Annual IEEE Conference on Computational Complexity (CCC 1999)*.

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, 2010.

[Tza11] Iddo Tzameret. Algebraic proofs over noncommutative formulas. *Information and Computation*, 209(10):1269–1292, 2011.

[Vol15a] Ilya Volkovich. Computations beyond exponentiation gates and applications. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:42, 2015.

[Vol15b] Ilya Volkovich. Deterministically factoring sparse polynomials into multilinear factors and sums of univariate polynomials. In *Proceedings of the 19th International Workshop on Randomization and Computation (RANDOM 2015)*, volume 40 of *LIPIcs*, pages 943–958, 2015. Preliminary version in the Electronic Colloquium on Computational Complexity (ECCC), Technical Report TR14-168.

[vzGK85] Joachim von zur Gathen and Erich L. Kaltofen. Factoring sparse multivariate polynomials. *JCSS*, 31(2):265–287, 1985. Preliminary version in the *24th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1983)*.

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation (EUROSAM 1979)*, pages 216–226. Springer-Verlag, 1979.

A Explicit Multilinear Polynomial Satisfying a Functional Equation

In Subsection 5.1 we showed that any polynomial that agrees with function $\overline{x} \mapsto 1/(\sum_i x_i - \beta)$ on the boolean cube must have degree $\geq n$. However, as there is a unique multilinear polynomial obeying this functional equation it is natural to ask for an explicit description of this polynomial, which we now give.

Proposition A.1. *Let $n \geq 1$ and $\mathbb{F}$ be a field with $\mathrm{char}(\mathbb{F}) > n$. Suppose that $\beta \in \mathbb{F} \setminus \{0, \ldots, n\}$. Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be the unique multilinear polynomial such that*

$$f(\overline{x}) = \frac{1}{\sum_i x_i - \beta} ,$$

for $\overline{x} \in \{0,1\}^n$. Then

$$f(\overline{x}) = - \sum_{S \subseteq [n]} \frac{(|S|)!}{\prod_{k=0}^{|S|}(\beta - k)} \prod_{i \in S} x_i .$$

Proof: It follows from the uniqueness of the evaluations of multilinear polynomials over the boolean cube that

$$f(\overline{x}) = \sum_{T \subseteq [n]} f(\mathbb{1}_T) \prod_{i \in T} x_i \prod_{i \notin T} (1 - x_i)$$

where $\mathbb{1}_T \in \{0,1\}^n$ is the indicator vector of the set T, so that

$$= \sum_{T \subseteq [n]} \frac{1}{|T| - \beta} \prod_{i \in T} x_i \prod_{i \notin T} (1 - x_i) .$$

Using this, let us determine the coefficient of $\prod_{i \in S} x_i$ in $f(\overline{x})$, for $S \subseteq [n]$ with $|S| = m$. First observe that setting $x_i = 0$ for $i \notin S$ preserves this coefficient, so that

$$\begin{aligned}
\mathrm{Coeff}_{\prod_{i \in S} x_i}\Big(f(\overline{x})\Big) &= \mathrm{Coeff}_{\prod_{i \in S} x_i}\Big(f(\overline{x}|_S, \overline{0})\Big) \\
&= \mathrm{Coeff}_{\prod_{i \in S} x_i} \left(\sum_{T \subseteq [n]} \frac{1}{|T| - \beta} \prod_{i \in T} x_i \prod_{i \notin T} (1 - x_i) \right) \Bigg|_{x_i \leftarrow 0, i \in S}
\end{aligned}$$

and thus those sets T with $T \not\subseteq S$ are zeroed out,

$$\begin{aligned}
&= \mathrm{Coeff}_{\prod_{i \in S} x_i} \left(\sum_{T \subseteq S} \frac{1}{|T| - \beta} \prod_{i \in T} x_i \prod_{i \in S \setminus T} (1 - x_i) \right) \\
&= \sum_{T \subseteq S} \frac{1}{|T| - \beta} \mathrm{Coeff}_{\prod_{i \in S} x_i} \left(\prod_{i \in T} x_i \prod_{i \in S \setminus T} (1 - x_i) \right) \\
&= \sum_{T \subseteq S} \frac{1}{|T| - \beta} (-1)^{m - |T|} \\
&= \sum_{k=0}^{m} \binom{m}{k} \frac{1}{k - \beta} (-1)^{m-k}
\end{aligned}$$

$$= -\frac{m!}{\prod_{k=0}^{m}(\beta-k)} = -\frac{(|S|)!}{\prod_{k=0}^{|S|}(\beta-k)} ,$$

where the last step uses the below subclaim.

Sub claim A.2.

$$\sum_{k=0}^{m}\binom{m}{k}\frac{1}{k-\beta}(-1)^{m-k} = -\frac{m!}{\prod_{k=0}^{m}(\beta-k)} .$$

Sub-Proof: Clearing denominators,

$$\prod_{j=0}^{m}(j-\beta)\cdot\sum_{k=0}^{m}\binom{m}{k}\frac{1}{k-\beta}(-1)^{m-k} = \sum_{k=0}^{m}\binom{m}{k}(-1)^{m-k}\prod_{j\neq k}(j-\beta)$$

Note that the right hand side is a degree $< m$ polynomial in β, so it is determined by its value on $\ell \in \{0,\dots,m\}$. Note that on these values,

$$\begin{aligned}\sum_{k=0}^{m}\binom{m}{k}(-1)^{m-k}\prod_{j\neq k}(j-\ell) &= \binom{m}{\ell}(-1)^{m-\ell}\prod_{0\le j<\ell}(j-\ell)\cdot\prod_{\ell<j\le m}(j-\ell)\\ &= \binom{m}{\ell}(-1)^{m-\ell}\cdot(-1)^{\ell}\ell!\cdot(m-\ell)!\\ &= (-1)^m m! .\end{aligned}$$

Thus $\sum_{k=0}^{m}\binom{m}{k}(-1)^{m-k}\prod_{j\neq k}(j-\beta) = (-1)^m m!$ for all β, and thus dividing by $\prod_{k=0}^{m}(k-\beta)$ and clearing -1's yields the claim. □

This then gives the claim. □