Electronic Colloquium on Computational Complexity, Report No. 79 (2025)

New Bounds for the Ideal Proof System in Positive Characteristic

Amik Raj Behera *

Nutan Limaye †

Varun Ramanathan ‡

Srikanth Srinivasan §

June 19, 2025

Abstract

In this work, we prove upper and lower bounds over fields of positive characteristics for several fragments of the Ideal Proof System (IPS), an algebraic proof system introduced by Grochow and Pitassi (J. ACM 2018). Our results extend the works of Forbes, Shpilka, Tzameret, and Wigderson (Theory of Computing 2021) and also of Govindasamy, Hakoniemi, and Tzameret (FOCS 2022). These works primarily focused on proof systems over fields of characteristic 0, and we are able to extend these results to positive characteristic.

The question of proving general IPS lower bounds over positive characteristic is motivated by the important question of proving $\mathsf{AC}^0[p]$-Frege lower bounds. This connection was observed by Grochow and Pitassi (J. ACM 2018). Additional motivation comes from recent developments in algebraic complexity theory due to Forbes (CCC 2024) who showed how to extend previous lower bounds over characteristic 0 to positive characteristic.

In our work, we adapt the functional lower bound method of Forbes et al. (Theory of Computing 2021) to prove exponential-size lower bounds for various subsystems of IPS. In order to establish these size lower bounds, we first prove a tight degree lower bound for a variant of *Subset Sum* over positive characteristic. This forms the core of all our lower bounds.

*Department of Computer Science, University of Copenhagen, Denmark, Email: ambe@di.ku.dk. Supported by Srikanth Srinivasan's start-up grant from the University of Copenhagen.

†IT University of Copenhagen, Denmark, Email: nuli@itu.dk. Supported by Independent Research Fund Denmark (grant agreement No. 10.46540/3103-00116B) and is also supported by the Basic Algorithms Research Copenhagen (BARC), funded by VILLUM Foundation Grant 54451.

‡School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India, Email: varun.ramanathan@tifr.res.in. Supported by the Department of Atomic Energy, Government of India, under project number RTI400112. A part of the work was done when the author was visiting the University of Copenhagen and was supported by the European Research Council (ERC) under grant agreement no. 101125652 (ALBA).

§Department of Computer Science, University of Copenhagen, Denmark, Email: srsr@di.ku.dk. Supported by the European Research Council (ERC) under grant agreement no. 101125652 (ALBA).

ISSN 1433-8092

Additionally, we derive upper bounds for the instances presented above. We show that they have efficient constant-depth IPS refutations. This demonstrates that constant-depth IPS refutations are stronger than the proof systems considered above even in positive characteristic. We also show that constant-depth IPS can efficiently refute a general class of instances, namely all symmetric instances, thereby further uncovering the strength of these algebraic proofs in positive characteristic.

Notably, our lower bounds hold for fields of arbitrary characteristic but require the field size to be $n^{\omega(1)}$. In a concurrent work, Elbaz, Govindasamy, Lu, and Tzameret have shown lower bounds against restricted classes of IPS over finite fields of any size by considering different hard instances.

Contents

1 Introduction

Propositional Proof Systems. A proof system consists of a set of axioms and inference rules. The goal is to start with the given set of axioms and apply the inference rules repeatedly to prove theorems (tautologies) within the proof system. A proof system is *sound* if it proves only true statements and it is *complete* if it proves all true statements. The area of *Propositional Proof Complexity* aims to understand the strength of different proof systems in the propositional setting. In a foundational work, Cook and Reckhow [CR79] showed that if we could prove that there exist tautologies such that they require exponential proof size (i.e., vaguely the number of times different inference rules are applied in the proof) in any proof system, then it would resolve the famous NP vs. coNP question in computational complexity theory.

Apart from the connection to this central question in complexity theory, understanding the power of different proof systems is also fundamental to mathematical reasoning. This has motivated a lot of research in the area for the last five decades. (See for instance these reference texts for more context [Kra95; CK02; Kra19].) There are many different kinds of propositional proof systems based on the set of axioms they start with and the kind of inference rules they are allowed to use. In this work, we will focus on algebraic proof systems. In algebraic proof systems, propositional tautologies are expressed as an unsatisfiable set of polynomial equations and the inference rules are algebraic, i.e. they involve reasoning based on polynomial arithmetic.

The study of algebraic proof systems originates from the work of Beame, Impagliazzo, Krajíček, Pitassi, and Pudlák [BIKPP96] who introduced the Nullstellensatz proof system (based on Hilbert's Nullstellensatz). Their work was followed by the work of Clegg, Edmonds, and Impagliazzo [CEI96] who introduced Polynomial Calculus as a *dynamic* variant of the Nullstellensatz proof system. Over the years, substantial work on these proof systems has helped us get a good understanding of their power in terms of complexity measures such as sparsity and degree [BIKPP96; BIKPRS97; Raz98; Gri98; IPS99; BGIP01; AR01].

However, as noted in [FSTW21], sparsity and degree only roughly capture the complexity of algebraic proofs. More recently, Grochow and Pitassi [GP18] proposed the Ideal Proof System (IPS) as a natural generalization of these well-studied algebraic proof systems such as Polynomial Calculus and Nullstellensatz proof systems. In the last decade, several papers studied this proof system. (See for instance [GP18; PT16; FSTW21; GHT22; HLT24].) This has allowed us to understand many other aspects of algebraic proofs, such as proof size and proof depth.

In this paper, we extend this line of work. Specifically, we revisit some of the known upper and lower bounds for Ideal Proof Systems over characteristic 0 and show similar bounds over fields of any characteristic[1].

1.1 Ideal Proof Systems

We start by describing the general setup for an algebraic (static[2]) proof system. Let $\mathbf{x}$ denote the set of variables $\{x_1, x_2, \ldots, x_n\}$. We are given a set of polynomial axioms $f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ and the goal is to show that there is no 0-1 assignment to the variables such that it simultaneously

[1] In all the results mentioned here, when we say that a result holds over characteristic 0, it in fact holds over large enough characteristic as well.

[2] In the literature, the following type of proof system is often referred to as a static proof system. There are other algebraic proof systems, where the proof is presented line-by-line and those are known as dynamic proof systems. Here, we will only discuss static proof systems.

satisfies $\{f_1(\mathbf{x}) = 0, f_2(\mathbf{x}) = 0, \ldots, f_m(\mathbf{x}) = 0\}$ over $\mathbb{F}$. To force a common Boolean solution, the set of axioms is appended with additional axioms, $\{x_i^2 - x_i = 0\}_{i\in[n]}$ for $i \in [n]$. These are called the *Boolean axioms*.

Based on Hilbert's Nullstellensatz, we know that if $\{f_1(\mathbf{x}) = 0, f_2(\mathbf{x}) = 0, \ldots, f_m(\mathbf{x}) = 0\} \cup \{x_i^2 - x_i = 0\}_{i\in[n]}$ are simultaneously not satisfiable, then such a refutation[3] can be given by polynomials $A_1(\mathbf{x}), A_2(\mathbf{x}), \ldots, A_m(\mathbf{x})$ and $B_1(\mathbf{x}), B_2(\mathbf{x}), \ldots, B_n(\mathbf{x})$ such that

$$\sum_{i\in[m]} A_i(\mathbf{x}) \cdot f_i(\mathbf{x}) + \sum_{i\in[n]} B_i(\mathbf{x}) \cdot (x_i^2 - x_i) = 1. \tag{1}$$

The complexity of such a proof can be defined using complexity parameters of the polynomials $\{A_i(\mathbf{x})\}$ and $\{B_i(\mathbf{x})\}$. In the case of the Ideal Proof System, Grochow, and Pitassi proposed that we assume that $A_i(\mathbf{x}), B_i(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ are computed by algebraic circuits. (See Section 1.3 for the formal definition.) Based on this, they defined complexity measures such as circuit size and circuit depth of IPS.

This proof system in its full generality is known to be quite strong. Specifically, it can polynomially simulate Extended Frege [GP18], which is one of the most powerful among well-studied propositional proof systems. Additionally, the same work also showed that proving lower bounds for this proof system would also imply strong algebraic circuit lower bounds, which is also a very challenging problem.

In light of this (and other reasons explained below), many restricted variants of the IPS have been studied. Let $\mathcal{C}$ be a class of polynomials. Then, a $\mathcal{C}$-IPS refutation is an IPS-refutation wherein $\{A_i(\mathbf{x})\}_{i\in[m]}$ and $\{B_i(\mathbf{x})\}_{i\in[n]}$ belong to the class $\mathcal{C}$. Forbes, Shpilka, Tzameret, and Wigderson [FSTW21], as well as Govindasamy, Hakoniemi, and Tzameret [GHT22], considered different classes of polynomials, for example, the class of polynomials computed by read-once oblivious algebraic branching programs (roABPs), by multilinear formulas, or by constant-depth algebraic formulas. They proved upper and lower bounds on the size of (some variants of) $\mathcal{C}$-IPS refutations over characteristic 0.

1.2 Motivation

We extend these works and prove similar bounds in arbitrary characteristic. Our work is motivated by the following important strands of research in proof complexity.

IPS-refutations and $\mathrm{AC}^0[p]$-Frege. A long-standing open question in proof complexity, open for almost three decades [Kra15], is to prove superpolynomial lower bounds against $\mathrm{AC}^0[p]$-Frege proof systems, i.e., a proof system in which the lines of the proof are constant-depth Boolean circuits that use modular gates. In the late 80s, Razborov [Raz87] and Smolensky [Smo87; Smo93] resolved the Boolean circuit lower bound question for $\mathrm{AC}^0[p]$, but the corresponding proof complexity question has proved to be elusive.

Over the years, several attempts have been made to resolve this question. The most relevant to our work is the result by Grochow and Pitassi [GP18, Theorem 3.5] which showed that constant-depth-IPS over characteristic p can efficiently simulate $\mathrm{AC}^0[p]$-Frege proofs. This means

[3]The words 'proofs' and 'refutations' are treated interchangeably in this paper. What we will be 'proving' is a statement that 'refutes' the existence of a common solution to a system of equations.

that proving superpolynomial lower bounds against constant-depth-IPS refutations will give superpolynomial lower bounds against $\mathrm{AC}^0[p]$-Frege. This gives a strong motivation to prove IPS lower bounds over small characteristics.

Functional lower bounds over any characteristic. Building on the work of [GP18], [FSTW21] further explored the power of IPS refutations. They proposed a concrete approach towards proving size lower bounds for IPS refutations via *functional lower bounds* (further explained in Section 1.4). Their method was inspired by the notion of functional lower bounds in Boolean circuit complexity [GR00; FKS16]. They demonstrated the promise of their method by proving several lower bounds for different fragments of IPS.

For example, the strong algebraic complexity lower bounds known for roABPs [Nis91] and multilinear formulas [Raz09] follow from understanding the *evaluation dimension* complexity measure in these models. Since this measure is essentially functional in nature, [FSTW21] used it to successfully prove lower bounds for $\mathcal{C}$-IPS when $\mathcal{C}$ is a class of read-once branching programs or multilinear formulas. Their bounds are over characteristic 0.

This approach of [FSTW21] was further adapted by Govindasamy, Hakoniemi, and Tzameret [GHT22] to prove superpolynomial lower bounds against (multilinear) constant-depth-IPS refutations. Their proof builds on some of the key components of the superpolynomial lower bound against constant-depth algebraic circuits by Limaye, Srinivasan, and Tavenas. The latter lower bound of [LST21] only worked over characteristic 0; for this and other reasons, the result of [GHT22] was also limited to characteristic 0. In a recent paper, however, Forbes [For24] improved the circuit lower bound result of [LST21] and proved the same[4] lower bound over any characteristic.

In light of these results, the next obvious step is to prove the lower bounds of [FSTW21; GHT22] over any characteristic. We achieve that in this work.[5]

1.3 Our Results

To describe our results, we start with the formal definitions of IPS refutations and its variants.

Definition 1.1 (IPS proof systems [GP18; FSTW21])**.** *Let $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ be a system of unsatisfiable polynomials over the Boolean cube $\{0,1\}^n$. In other words, there is no Boolean assignment $\mathbf{a} \in \{0,1\}^n$ to the variables $x_1, \ldots, x_n$ so that $f_i(\mathbf{a}) = 0$ for all $i \in [m]$.*

Given a class of algebraic circuits $\mathcal{C}$, a $\mathcal{C}$-IPS refutation *of the system of equations defined by $f_1, \ldots, f_m$ is an algebraic circuit $C \in \mathcal{C}$ in variables $x_1, \ldots, x_n, y_1, \ldots, y_m, z_1, \ldots, z_n$ such that*

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

The size of the refutation is the size of the circuit C.

Further, if the circuit C has individual degree at most 1 in the variables $\mathbf{y}$ and $\mathbf{z}$, then we say that C is a $\mathcal{C}$-$\mathrm{IPS}_{\mathrm{LIN}}$ refutation. *If the circuit C has individual degree at most 1 in the variables $\mathbf{y}$ (but not necessarily in $\mathbf{z}$), then C is said to be a* $\mathcal{C}$-$\mathrm{IPS}_{\mathrm{LIN}'}$ refutation.

[4] Some parameters in the lower bound by [LST21] were subsequently improved by [BDS24] and [For24] achieves those improved parameters.

[5] The subset-sum instances from [FSTW21; GHT22] are not always unsatisfiable over fields of positive characteristic; this requires that we tweak their instances to ensure unsatisfiability. Barring these changes, we qualitatively match their lower bounds over fields of positive characteristic.

Finally, we say that a circuit $C \in \mathcal{C}$ is a multilinear $\mathcal{C}$-IPS$_{\text{LIN}'}$ refutation *if additionally $C(\mathbf{x}, \mathbf{y}, \mathbf{0})$ is a multilinear polynomial in the variables $\mathbf{x} \cup \mathbf{y}$.*

Remark 1.2. *We mostly employ the above definition in the case that $m = 1$, i.e. the case when we have a single polynomial equation that is unsatisfiable over the Boolean cube. Further, while our upper bound results are proved in the more restrictive $\mathcal{C}$-IPS$_{\text{LIN}}$ proof system, our lower bounds results hold in the setting of the stronger $\mathcal{C}$-IPS$_{\text{LIN}'}$ proof systems.*

We also recall some standard notions about polynomials and algebraic models of computation, which will be useful below.

Multilinear and symmetric polynomials. A polynomial $f(\mathbf{x}) \in \mathbb{F}[x_1, \ldots, x_n]$ is a *multilinear* if the individual degree is at most 1. For a polynomial $f(\mathbf{x})$, the *multilinearization* operator, denoted by $\mathsf{ml}[\,\cdot\,]$, changes for each variable x_j and any k, every occurrence of x_j^k in $f(\mathbf{x})$ to x_j.

A polynomial $f(\mathbf{x}) \in \mathbb{F}[x_1, \ldots, x_n]$ is said to be a *symmetric polynomial* if the polynomial remains invariant under any permutation of the input variables. For a degree parameter $0 \leqslant d \leqslant n$, the d^{th} elementary symmetric polynomial $e_{n,d}(x_1, \ldots, x_n)$ is defined to be the following multilinear polynomial $e_{n,d}(x_1, \ldots, x_n) = \sum_{\substack{S \subseteq [n] \\ |S|=d}} \prod_{i \in S} x_i$. Whenever n is clear from the context, we will denote the d^{th} elementary symmetric polynomial by $e_d(\mathbf{x})$.

Algebraic models of computation. We recall the definitions of some of the standard models of computation relevant to our results.

Algebraic circuits and formulas. An *algebraic circuit* is a directed acyclic graph in which each node either computes a sum (or a linear combination) of its inputs, or a product of its inputs. The leaf nodes are either variables or constants. The size of an algebraic circuit is the number of edges in the circuit, and the depth of an algebraic circuit is the longest path from the output node (a sink) to a leaf node (a source). An *algebraic formula* is an algebraic circuit where the output of each node feeds into at most another node; in other words, the underlying graph of an algebraic formula is a tree. An algebraic formula is a *multilinear formula* if every gate of the formula computes a multilinear formula.

Sparse polynomials and constant-depth circuits. The class $\sum\prod$ consists of depth-2 formulas with an addition gate in the top layer and multiplication gates in the bottom (second) layer. All the gates have unbounded fan-in. $\sum\prod$ formulas essentially compute polynomials in the *sparse* representation i.e. as a sum of monomials. In general, a constant-depth algebraic circuit has $O(1)$ alternating layers of additional and multiplication gates.

Read-Once Oblivious Algebraic Branching Programs. A read-once oblivious algebraic branching program in the variable-order $\pi \in \mathcal{S}_n$[6] is a directed acyclic graph whose vertices are partitioned into n layers $V_0 = \{s\}, V_1, V_2, \ldots, V_n = \{t\}$. For each $i \in \{1, 2, \ldots, n\}$, there are edges directed from layer V_{i-1} to V_i that are labelled by univariate polynomials in the variable $x_{\pi(i)}$. For each s-to-t path p, the polynomial computed by p is defined to be product of the edge labels on p. The polynomial computed by the roABP is defined as the sum of polynomials computed by all s-to-t paths. The *width* of an roABP is $\max_{0 \leqslant i \leqslant n} |V_i|$ i.e. the size of the largest layer of vertices.

For more background on these models of computation, please refer to one of the standard surveys in algebraic complexity ([SY10],[Sap21]).

[6] $\mathcal{S}_n$ denotes the set of all permutation of $[n]$.

1.3.1 Lower Bounds Over Positive Characteristic

We start by stating our lower bound results.

Theorem 1.3 (Lower bounds for sparse-$\mathrm{IPS}_{\mathrm{LIN}'}$ in positive characteristic)**.** *The following holds for any large enough n. Let p be any prime number. Let $k \in \mathbb{N}$ such that $p^k > 2^{\Omega(n)}$. There exist $\alpha_i \in \mathbb{F}_{p^k}$ and $\beta \in \mathbb{F}_{p^{2k}} \setminus \mathbb{F}_{p^k}$ such that*

- *The polynomial $f = \sum_{i\in[n]} \alpha_i x_i - \beta$ has no Boolean satisfying assignment.*
- *Any sparse-$\mathrm{IPS}_{\mathrm{LIN}'}$ refutation[7] of f must have size at least $2^{\Omega(n)}$*

Note that the hard instance above is a sparse polynomial. We show that it has no small sparse refutation over positive characteristic.

Theorem 1.4 (Lower bounds for fixed-order roABP in positive characteristic)**.** *The following holds for any large enough n. Let p be any prime number. Let $k \in \mathbb{N}$ such that $p^k > 2^{\Omega(n)}$. There exist $\alpha_i \in \mathbb{F}_{p^k}$ and $\beta \in \mathbb{F}_{p^{2k}} \setminus \mathbb{F}_{p^k}$ such that*

- *The polynomial $f = \sum_{i\in[n]} \alpha_i x_i y_i - \beta$ has no Boolean satisfying assignment.*
- *Any* roABP-$\mathrm{IPS}_{\mathrm{LIN}'}$ *refutation of f in any order of variables where $\mathbf{x}$ variables come before $\mathbf{y}$ variables, must have width $2^{\Omega(n)}$.*

To obtain lower bounds against more powerful models such as roABP-$\mathrm{IPS}_{\mathrm{LIN}'}$ with respect to *any* order, or multilinear formulas, [FSTW21] used a slightly modified hard instance. We also use an instance the same as theirs up to the choice of coefficients.

Theorem 1.5 (Lower bounds for any order roABP-$\mathrm{IPS}_{\mathrm{LIN}'}$ and multilinear-formula-$\mathrm{IPS}_{\mathrm{LIN}'}$)**.** *The following holds for any large enough n. Let p be any prime number. Let $k \in \mathbb{N}$ such that $p^k > 2^{\Omega(n)}$. There exist $\alpha_{i,j} \in \mathbb{F}_{p^k}$ and $\beta \in \mathbb{F}_{p^{2k}} \setminus \mathbb{F}_{p^k}$ such that*

- *The polynomial $f = \sum_{1\leqslant i<j\leqslant n} \alpha_{i,j} z_{i,j} x_i x_j - \beta$ has no Boolean satisfying assignment.*
- *Any* roABP-$\mathrm{IPS}_{\mathrm{LIN}'}$ *refutation of f must have size at least $2^{\Omega(n)}$.*
- *Moreover, any multilinear-formula-$\mathrm{IPS}_{\mathrm{LIN}'}$ refutation of f must have size at least $n^{\Omega(\log n)}$ and for $\Delta = o(\log n/\log\log n)$, any product-depth[8]-$\Delta$ multilinear-formula-IPS refutation requires size $\geqslant n^{\Omega\left(\frac{1}{\Delta^2}(\frac{n}{\log n})^{1/\Delta}\right)}$.*

Again notice that, f is a sparse polynomial and hence has a polynomial size roABP. It is also efficiently computable by a multilinear formula.

In general, in Boolean proof complexity, it is typical that the hard-to-refute instances are themselves easy to compute. In algebraic proof complexity, there are some lower bound results that do not have this property. That is, the instances that are hard to refute are also hard to compute. For example, the set of results obtained by the approach of multiples in [FSTW21, Theorem 1.18, Theorem 1.19, Theorem 1.20] and in a paper by Andrews and Forbes [AF22]. Additionally, in a

[7] Note that sparse-$\mathrm{IPS}_{\mathrm{LIN}}$ (a weaker system than sparse-$\mathrm{IPS}_{\mathrm{LIN}'}$) is equivalent to the Nullstellensatz proof system of [BIKPP96].

[8] The product-depth of a circuit is the maximum number of product gates appearing in any leaf-to-root path.

recent work Hakoniemi, Limaye, and Tzameret [HLT24] presented instances that were hard to refute for roABP-IPS$_{\text{LIN}'}$ and for multilinear-formula-IPS$_{\text{LIN}'}$ over any characteristics, i.e., similar to what we prove here. However, unfortunately, their instances were hard to compute and specifically, they could not be computed by roABP or by multilinear formulas. Hence, our result here have the best of both the worlds; the lower bounds hold over any characteristic and the hard instances are easy to compute.

Theorem 1.6 (Lower bounds for multilinear constant-depth-IPS$_{\text{LIN}'}$ in positive characteristic)**.** *The following holds for any large enough n. Let p be any prime and let $k \in \mathbb{N}$ be large enough so that $p^k > 2^{\Omega((\log n)^2)}$. There exist $\alpha_{i,j,k,\ell} \in \mathbb{F}_{p^k}$ and $\beta \in \mathbb{F}_{p^{2k}} \setminus \mathbb{F}_{p^k}$ such that*

- *The polynomial $f = \sum_{1 \leqslant i<j<k<\ell \leqslant n} \alpha_{i,j,k,\ell} z_{i,j,k,\ell} x_i x_j x_k x_\ell - \beta$ has no Boolean satisfying assignment.*
- *Any multilinear constant-depth-IPS$_{\text{LIN}'}$ refutation of f must have size $n^{\omega(1)}$.*

The characteristic 0 (or large characteristic) version of the above theorem was presented in [GHT22]. Their lower bound is a step towards constant-depth-IPS lower bounds. Our result above can thus be thought of as another step forward in the right direction. Moreover, our input instance is the same as the input instance in Theorem 1 [GHT22] up to the choice of coefficients, and it is easy to compute (while being hard to refute). More specifically, it is computable by polynomial-sized constant-depth multilinear formulas.

Remark 1.7. *In all our results, the field characteristic is arbitrary, but the field size is quite large, i.e., p^k is either exponential or superpolynomial. This setting is non-trivial because the field elements have polynomial bit complexity. Other results in the area, such as the work of Alekseev, Grigoriev, Hirsch, and Tzameret [AGHT20] similarly use polynomial constraints with coefficients from exponentially large domains. Specifically [AGHT20] study a variant of the subset sum instance, called the Binary Value Principle, $\sum_{i \in [n]} 2^{i-1} x_i + 1 = 0$ in the context of IPS proof systems in fields of characteristic zero.*

It is an interesting open question to prove similar IPS *lower bounds over finite fields of small size. Unfortunately, as we show below, this forces the polynomial instances to become more complicated. See Section 1.5 for recent independent work that makes progress in this direction.*

1.3.2 Upper Bounds Over Positive Characteristic

A natural question for hard instances above is: what is the weakest proof system in which they are efficiently refutable? In personal communication, Tzameret observed that the above instances were refutable by constant-depth-IPS$_{\text{LIN}}$ hence showing that these proof systems can be exponentially more succinct than their multilinear counterpart. The theorem below shows that the above polynomials have efficient constant-depth-IPS$_{\text{LIN}}$ refutations, even in the setting of positive characteristic.

Theorem 1.8 (Upper bounds for (non-multilinear) constant-depth-IPS$_{\text{LIN}}$)**.** *Fix a prime number p. The following holds for any natural numbers n and k.*
Let $f \in \mathbb{F}_{p^k}[x_1, \ldots, x_n]$ be any polynomial with sparsity s and degree D with coefficients from the field $\mathbb{F}_{p^k}$ and let β be any element of $\mathbb{F} \setminus \mathbb{F}_{p^k}$ where $\mathbb{F}$ is a field extension of $\mathbb{F}_{p^k}$.
Then,

- *The polynomial $f(\mathbf{x}) - \beta$ has no satisfying assignment over the Boolean cube $\{0,1\}^n$*

- *There is a constant-depth-*IPS_{LIN} *refutation of degree* $O(k \cdot p \cdot D)$ *and size* $\text{poly}(s, p)$.

Note that since $\beta \notin \mathbb{F}_{p^k}$, the polynomial $f(\mathbf{x}) - \beta$ does not have a zero over $\{0,1\}^n$ (in fact it does not have a solution over $\mathbb{F}_{p^k}^n$). So the first item of above follows immediately. We also give non-trivial constant-depth-IPS_{LIN} refutations for degree-1 polynomials that are unsatisfiable over $\{0,1\}^n$ with all the coefficients in the same field.

Theorem 1.9 (Upper bound on degree of Nullstellensatz certificate)**.** *Fix a prime p. The following holds for any natural numbers n and k with $n > kp$.*
The following holds for every $\alpha_1, \ldots, \alpha_n, \beta \in \mathbb{F}_{p^k}$. Suppose the degree-1 polynomial $\sum_{i=1}^n \alpha_i x_i - \beta \in \mathbb{F}_{p^k}[x_1, \ldots, x_n]$ is unsatisfiable over the Boolean cube $\{0,1\}^n$ (i.e. there does not exist a Boolean point $\mathbf{a} \in \{0,1\}^n$ such that $\sum_{i=1}^n \alpha_i a_i - \beta = 0$).
Then, there is a constant-depth-IPS_{LIN} *refutation of degree $\mathcal{O}(k \cdot p)$ and size $\mathcal{O}(n/kp)^{\mathcal{O}(kp)}$.*

In particular, if $p = \mathcal{O}(1)$ and $k = o(n)$, then there is a constant-depth-IPS_{LIN} *refutation of degree $o(n)$ and size $2^{o(n)}$.*

Note that for degree-1 polynomials, the difference in Theorem 1.8 and Theorem 1.9 is in the constant-term β. If every $\alpha_i \in \mathbb{F}_{p^k}$ and $\beta \notin \mathbb{F}_{p^k}$, then the polynomial is always unsatisfiable over $\{0,1\}^n$ (no matter the choice of α_i's and β). In fact, it is unsatisfiable over $\mathbb{F}_p^n$. Our proof of Theorem 1.8 leverages this and yields an efficient refutation. However, if $\beta \in \mathbb{F}_{p^k}$, then our proof of Theorem 1.8 falls apart. We handle this separately in Theorem 1.9, but we do not match Theorem 1.8 qualitatively. More precisely, Theorem 1.8 yields a $\text{poly}(n, p)$-sized non-multilinear constant-depth refutations, but Theorem 1.9 yields a roughly $\binom{n}{k}$-sized non-multilinear constant-depth refutations.

Remark 1.10. *Suppose the characteristic p is a fixed prime independent of the number of variables n.*

- *Theorem 1.8 shows that the exponential field size in Theorem 1.3, Theorem 1.4 and Theorem 1.5 is not an artifact of the proofs.*[9]*. For fields of subexponential size, the polynomials in these theorems have refutations of degree $o(n)$ and in particular have* roABP-IPS_{LIN} *refutations of size $2^{o(n)}$.* [10]

- *Theorem 1.8 also shows that the multilinearity assumption in Theorem 1.6 is not an artifact of the proof. Non-multilinear proofs, even over large fields, allow efficient constant-depth refutations for sparse instances.*

Our final result shows a constant-depth upper bound for multilinear and *symmetric* systems of polynomials, i.e. systems defined by polynomials $f(x_1, \ldots, x_n)$ of the form

$$\sum_{d=1}^{n} \alpha_d e_{n,d} + \alpha_0$$

[9] Suppose the field $\mathbb{F}_{p^k}$ is not large enough, say, $k = o(n)$. Then there is a refutation of degree $d = O(k \cdot p \cdot D)$, which is $o(n)$ when p and D are constants. In particular, the sparsity of the refutation is at most $\binom{n+d}{d}$, which is $2^{o(n)}$ when $d = o(n)$.

[10] When the characteristic p is a growing function of n, this argument breaks down. It might be possible to get rid of the exponential field size.

where $e_{n,d}$ denotes the elementary symmetric polynomial of degree d in variables $x_1, \ldots, x_n$. Such polynomial systems have been employed in [FSTW21] to prove lower bounds against restricted systems of constant-depth-IPS_{LIN}. Our results imply that general constant-depth circuit refutations can be exponentially more succinct than these restricted families, even for positive characteristic.

Theorem 1.11 (Upper bounds for multilinear symmetric systems)**.** *Fix a field $\mathbb{F}$. Let $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ be a family of multilinear and symmetric polynomials with no common Boolean solution i.e. there does not exist a $\mathbf{x} \in \{0,1\}^n$ such that each $f_i(\mathbf{x}) = 0$. This system has a constant-depth-IPS_{LIN} refutation of size $\mathcal{O}(m^2 n^5 \log n)$ and depth 8.*

1.4 Proof Techniques

Lower bounds. Our proof uses the functional lower bound method introduced by [FSTW21], which can be described as follows. We know that a $\mathcal{C}$-$\text{IPS}_{\text{LIN}'}$ refutation for $f(\mathbf{x})$ consists of $A(\mathbf{x})$, $B_i(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ such that

$$f(\mathbf{x}) \cdot A(\mathbf{x}) + \sum_{i \in [n]} (x_i^2 - x_i) \cdot B_i(\mathbf{x}) = 1,$$

where $A(\mathbf{x}), B_1(\mathbf{x}), \ldots, B_n(\mathbf{x})$ belong to $\mathcal{C}$. As $f(\mathbf{x})$ is unsatisfiable over the Boolean hypercube, this implies that over the Boolean hypercube, $A(\mathbf{x})$ is a well-defined reciprocal of $f(\mathbf{x})$. Hence, to show that $A(\mathbf{x})$ cannot belong to $\mathcal{C}$, it is enough to show that any polynomial that agrees with $1/f(\mathbf{x})$ cannot be computed by $\mathcal{C}$. That is, the problem of proving a lower bound on the size of $\mathcal{C}$-$\text{IPS}_{\text{LIN}'}$ is reduced to proving a functional lower bound for $1/f(\mathbf{x})$.

At the heart of such a functional lower bound lies a *degree lower bound*, i.e., a lower bound on the degree of $\tilde{f}(\mathbf{x})$, where $\tilde{f}(\mathbf{x})$ and $f(\mathbf{x})$ are related. In fact, $f(\mathbf{x})$ is a *lifted* version of $\tilde{f}(\mathbf{x})$. Once we have such a degree lower bound for $\tilde{f}(\mathbf{x})$, we can apply proof ideas from algebraic complexity theory such as the rank-based lower bound methods. These methods allow for the degree lower bounds for $\tilde{f}(\mathbf{x})$ to be lifted to size lower bounds for $f(\mathbf{x})$.

For their machinery to work over positive characteristic, we prove a *positive characteristic* version of the degree lower bound (see Lemma 2.2 for the formal statement). In the case of the lower bound argument in [FSTW21], it was important to obtain a tight degree lower bound of exactly n. They needed it for the next step, i.e., *lifting*, to work. In our case, we show that such a degree lower bound holds with high probability (over the choice of coefficients of the hard instance). Once we have the degree lower bound, the rest of the lower bound proof works similar to the proof by [FSTW21].

Upper bounds. We now describe the main ingredients in our upper bounds. We start by describing the main ideas in the proof of Theorem 1.8.

Constant-depth upper bounds. Here, we proceed in two steps. First, we observe that for any sparse polynomial of degree d, we can *flatten* it to a linear polynomial by renaming the monomials by fresh variables. Our hard instance is indeed sparse, hence the observation can be used to rewrite the polynomial as a linear polynomial over a fresh set of variables.

Now, consider a linear polynomial $L(\mathbf{x}) - \beta$ such that $L(\mathbf{x}) = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n$, where $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_{p^k}$ for some k and prime p and $\beta \in \mathbb{F} \backslash \mathbb{F}_{p^k}$ such that it is not satisfiable over 0-1 assignments.

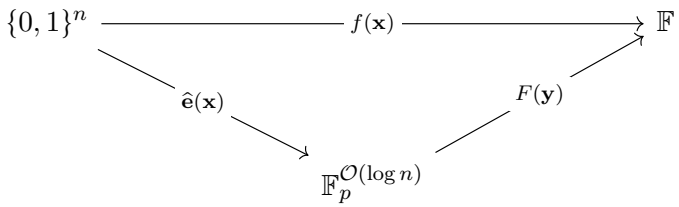
To prove that the polynomial has a refutation over constant-depth circuits, we first prove that for every j, $L_j(\mathbf{x}) = \alpha_1^{p^j} x_1 + \alpha_2^{p^j} x_2 + \ldots + \alpha_n^{p^j} x_n - \beta^{p^j}$ can be expressed as a multiple of $L(\mathbf{x})$ modulo the ideal $\mathbf{x}^p - \mathbf{x}$, which is a shorthand for the ideal generated by $\{x_i^p - x_i\}_{i \in [n]}$.

We then observe that for $j = k$, $L_k(\mathbf{x}) - L(\mathbf{x})$ is a non-zero constant and use this observation to construct small depth circuits for the refutation of $L(\mathbf{x}) - \beta$. Throughout, we use some standard but useful tricks available to positive characteristic fields.

For the proof of Theorem 1.9, we observe that the multilinear part of $(f(\mathbf{x}) - \beta)^{-1}$ has degree $\mathcal{O}(kp)$. This follows from Fermat's Little Theorem and using basic properties about multilinearization. See Section 3.2 for complete details.

Upper bounds for symmetric polynomials Now we discuss the proof outline for Theorem 1.11. For ease of exposition, we explain the ideas for the case of $m = 1$ in Theorem 1.11, i.e. there is one multilinear symmetric polynomial $f(\mathbf{x})$ that does not have a solution over the Boolean cube $\{0,1\}^n$. Suppose $\mathbb{F}$ has characteristic $p > 0$. Any symmetric polynomial is a polynomial of the n elementary symmetric polynomials[11] i.e. $e_1(\mathbf{x}), \ldots, e_n(\mathbf{x})$. However, if we restrict to the Boolean cube $\{0,1\}^n$, then any symmetric polynomial is a polynomial of just $\mathcal{O}(\log n)$ elementary symmetric polynomials. Let $\hat{\mathbf{e}}(\mathbf{x})$ denotes the tuple of those $\mathcal{O}(\log n)$ elementary symmetric polynomials (see Claim 4.2 for an explicit description of $\hat{\mathbf{e}}(\mathbf{x})$.)

Let $F(\mathbf{y})$ be the $\mathcal{O}(\log n)$ variate polynomial such that $F(\mathbf{y}) \circ \hat{\mathbf{e}}(\mathbf{x})$ agrees with $f(\mathbf{x})$ on the Boolean cube $\{0,1\}^n$. The Boolean cube $\{0,1\}^n$ is mapped to $\mathbb{F}_p^{\mathcal{O}(\log n)}$ under the map $\hat{\mathbf{e}}(\mathbf{x})$ because $\mathrm{char}(\mathbb{F}) = p$. The unsatisfiability of $f(\mathbf{x})$ over the Boolean cube $\{0,1\}^n$ implies the unsatisfiability of $F(\mathbf{y})$ over $\mathbb{F}_p^{\mathcal{O}(\log n)}$. Applying Hilbert's Nullstellensatz Theorem (see Theorem 1.19) on the unsatisfiability[12] of $F(\mathbf{y})$ over $\mathbb{F}_p^{\mathcal{O}(\log n)}$, we get a *low-variate* Nullstellensatz certificate (it is a Nullstellensatz certificate in just $\mathcal{O}(\log n)$ variables)[13]. The coefficients of this low-variate Nullstellensatz certificate can be computed via $\mathrm{poly}(n)$-sized constant-depth circuits. This follows from the fact that we are working over constant characteristic. Refer to the diagram below for a schematic representation of what we discussed so far.



Next we "lift" the Nullstellensatz back to the n variables $(x_1, \ldots, x_n)$. To do so, we plug-in $\hat{\mathbf{e}}(\mathbf{x})$ in place of $\mathbf{y}$. Observe that this substitution by $\hat{\mathbf{e}}(\mathbf{x})$ preserves the size and the depth of the coefficients of the low-variate Nullstellensatz certificate because of the Ben-Or's construction (see Theorem 1.15).

[11] This follows from the Fundamental Theorem of Symmetric Polynomials.

[12] To capture the restriction of $\mathbb{F}_p^n$, we add n univariate polynomials, each of which vanishes on one coordinate of $\mathbb{F}_p^n$.

[13] Loosely speaking, one can imagine this as a "dimension reduction" of our problem. The symmetric structure of $f(\mathbf{x})$ led us to convert a problem in n variables to a problem in just $\mathcal{O}(\log n)$ variables.

It remains to *prove* via constant-depth circuits that $F(\widehat{\mathbf{e}}(\mathbf{x}))$ agrees with $f(\mathbf{x})$ on the Boolean cube, i.e. $F(\widehat{\mathbf{e}}(\mathbf{x}))-f(\mathbf{x})$ lie in the ideal $(\mathbf{x}^2-\mathbf{x})$. Here "to prove in constant-depth circuits" refers to giving a certificate for the ideal membership whose coefficients can be computed by constant-depth circuits. More precisely, we want to prove that there exists polynomials $B_j(\mathbf{x})$'s which have $\mathrm{poly}(n)$-sized constant-depth circuits such that

$$F(\widehat{\mathbf{e}}(\mathbf{x})) \;=\; f(\mathbf{x}) + \sum_{j=1}^{n} B_j(\mathbf{x}) \cdot (x_j^2 - x_j).$$

This is the key step in our proof. To prove this, it suffices to prove the following special case, which we prove in Lemma 4.3.

Lemma 1.12. *Let $\ell = \mathcal{O}(\log n)$ and fix an arbitrary sequence $(\alpha_1, \ldots, \alpha_\ell)$ where each $\alpha_i \in [n]$. There exist polynomials $B_j(\mathbf{x})$'s such that*

$$\prod_{i=1}^{\ell} e_{\alpha_i}(\mathbf{x}) \;=\; \mathsf{ml}\left[\prod_{i=1}^{\ell} e_{\alpha_i}(\mathbf{x})\right] + \sum_{j=1}^{n} B_j(\mathbf{x}) \cdot (x_j^2 - x_j),$$

and each polynomial $B_j(\mathbf{x})$ can be computed by a $\mathrm{poly}(n)$-sized constant-depth circuit.

1.5 Related Work

In an independent work, Elbaz, Govindasamy, Lu, and Tzameret [EGLT25] consider related questions. Using the recent lower bound of Forbes [For24], which proves the positive characteristic version of the constant-depth formula lower bound of [LST21], they obtain lower bounds for fragments of the IPS over finite fields of *any* size.

1.6 Preliminaries

In this subsection, we present a few more definitions and standard facts on polynomials which will be used in our proofs later on.

For a polynomial $f(x_1, \ldots, x_n)$, the individual degree of f is an integer D such that for all $i \in [n]$, the degree of f when viewed as a univariate polynomial in the variable x_i is at most D.
We next mention some useful properties about multilinear polynomials.

Fact 1.13 (Standard facts on multilinear polynomials)**.** *Let $f(\mathbf{x}), g(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$.*

- *$f(\mathbf{x})$ and $\mathsf{ml}[f(\mathbf{x})]$ agree on the Boolean cube $\{0,1\}^n$.*
- *$f(\mathbf{x})$ and $g(\mathbf{x})$ agree on the Boolean cube $\{0,1\}^n$ if and only if $\mathsf{ml}[f(\mathbf{x})]$ is equal to the $\mathsf{ml}[g(\mathbf{x})]$.*
- $\mathsf{ml}[f(\mathbf{x})g(\mathbf{x})] = \mathsf{ml}[\mathsf{ml}[f(\mathbf{x})]\mathsf{ml}[g(\mathbf{x})]]$.

Theorem 1.14 (Fundamental Theorem of Symmetric Polynomials)**.** *Fix any arbitrary field $\mathbb{F}$. If $f \in \mathbb{F}[x_1, \ldots, x_n]$ is a symmetric polynomial of degree d, then there exists a unique polynomial $F \in \mathbb{F}[y_1, \ldots, y_d]$ such that $f(\mathbf{x}) = F(e_1(\mathbf{x}), \ldots, e_d(\mathbf{x}))$.*

A classical and beautiful construction of Ben-Or shows that every elementary symmetric polynomial can be computed by poly(n)-sized constant-depth circuits.

Theorem 1.15 (Ben-Or's construction for elementary symmetric polynomials)**.** *(See [SW01, Theorem 5.1]). Let $\mathbb{F}$ be a field with $|\mathbb{F}| > n$. Then for every $d \in [n]$, the d^{th} elementary symmetric polynomial $e_d(x_1, \ldots, x_n)$ has a circuit of size $\mathcal{O}(n^2)$ and depth 3 (a $\Sigma\Pi\Sigma$ circuit).*
More particularly, for any choice of $(n+1)$ distinct elements $\gamma_1, \ldots, \gamma_{n+1} \in \mathbb{F}$ and for every $k \in [n]$, there exists coefficients $c_{k,i}$'s such that

$$e_k(\mathbf{x}) \;=\; \sum_{i=1}^{n+1} c_{k,i} \prod_{j=1}^{n} (1 + \gamma_i x_j)$$

The following recursive definition of elementary symmetric polynomials will be used in the proofs.

$$e_d(x_1, \ldots, x_n) \;=\; x_1 \cdot e_{d-1}(x_2, \ldots, x_n) + e_d(x_2, \ldots, x_n), \qquad \text{for all } d \in [n] \tag{2}$$

Theorem 1.16 (Polynomial Identity Lemma)**.** *(See [GRS23, Lemma 9.2.2]). Let $\mathbb{F}$ be an arbitrary field. Let $f(\mathbf{x})$ be a nonzero polynomial of degree at most d and let $S \subseteq \mathbb{F}$. If we choose $\mathbf{a} \sim S^n$ uniformly at random, then:*

$$\Pr_{\mathbf{a} \sim S^n}[f(\mathbf{a}) = 0] \leqslant \frac{d}{|S|}$$

For a natural number k and variables $(z_1, \ldots, z_n)$, we will use $(\mathbf{z}^k - \mathbf{z})$ to denote the following ideal $(\mathbf{z}^k - \mathbf{z}) \;:=\; (z_1^k - z_1, \ldots, z_n^k - z_k) \;\subseteq\; \mathbb{F}[z_1, \ldots, z_n]$. We recall the following lemma which holds for fields with positive characteristic.

Lemma 1.17 (Freshman's Dream)**.** *Fix a prime number p and a field $\mathbb{F}$ of $\mathrm{char}(\mathbb{F}) = p$. Then for any $a, b \in \mathbb{F}$, we have, $(a+b)^p \;=\; a^p + b^p$. More generally, for any $a_1, \ldots, a_m \in \mathbb{F}$, we get, $(a_1 + \ldots + a_m)^p \;=\; a_1^p + \ldots + a_m^p$.*

Next we recall the definition of an ideal and a variety, and then we state Hilbert's Nullstellensatz.

Definition 1.18 (Ideal and Variety)**.** *Fix any field $\mathbb{F}$ and consider the commutative ring $\mathbb{F}[x_1, \ldots, x_n]$. For a set of polynomials $f_1, \ldots, f_m \in \mathbb{F}[\mathbf{x}]$, the ideal generated by f_i's, denoted by $(f_1, \ldots, f_m)$ is defined as:*

$$(f_1, \ldots, f_m) \;=\; \left\{ h \in \mathbb{F}[\mathbf{x}] \;\middle|\; \exists g_1, \ldots, g_m \in \mathbb{F} \text{ such that } h = \sum_{i=1}^{m} g_i f_i \right\}.$$

For a set of polynomials $f_1, \ldots, f_m \in \mathbb{F}$, their variety, denoted by $\mathbb{V}(f_1, \ldots, f_m)$ is a subset of the algebraic closure of $\overline{\mathbb{F}}^n$, defined as:

$$\mathbb{V}(f_1, \ldots, f_m) \;=\; \left\{ \mathbf{a} \in \bar{\mathbb{F}}^n \;\middle|\; f_1(\mathbf{a}) = \cdots = f_m(\mathbf{a}) = 0 \right\}.$$

Now we state Hilbert's Nullstellensatz which essentially says that if a set of polynomials do not have a common zero, then there exists "witness" for this, i.e. one can express 1 as a polynomial combination of f_i's.

Theorem 1.19 (Hilbert's Nullstellensatz)**.** *Fix any field $\mathbb{F}$. Let $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ be a set of multivariate polynomials such that they do not have any common zeros over the algebraic closure of $\mathbb{F}$. Then the constant 1 lies in the ideal $(f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$. In other words, there exists polynomials $A_1, \ldots, A_m \in \mathbb{F}[x_1, \ldots, x_n]$ such that*

$$A_1(\mathbf{x}) \cdot f_1(\mathbf{x}) + \cdots + A_m(\mathbf{x}) \cdot f_m(\mathbf{x}) \;=\; 1.$$

Strictly speaking, Hilbert's Nullstellensatz guarantees that the polynomials $A_i's$ are in $\overline{\mathbb{F}}[\mathbf{x}]$ ($\overline{\mathbb{F}}$ is the algebraic closure of $\mathbb{F}$). However, the above statement also follows easily by observing that we can solve for A_i's by solving a system of linear equations over $\mathbb{F}$. Throughout this article, we will refer to $(A_1(\mathbf{x}), \ldots, A_m(\mathbf{x}))$ as a *Nullstellensatz certificate*[14] for the system $\{f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\}$. We will also refer to A_i's as *coefficients* because if we take a polynomial combination of f_i's with A_i's being the coefficients, then we can generate 1.

Lemma 1.20 (Nullstellensatz certificate implies refutations)**.** *Fix any field $\mathbb{F}$. Let $P_1, \ldots, P_m \in \mathbb{F}[x_1, \ldots, x_n]$ be polynomials that have no common Boolean solution. Let the polynomials $A_i(\mathbf{x})$'s and $B_j(\mathbf{x})$'s be coefficients of the Nullstellensatz certificate, i.e.*

$$\sum_{i=1}^{m} A_i(\mathbf{x}) \cdot P_i(\mathbf{x}) + \sum_{j=1}^{n} B_j(\mathbf{x}) \cdot (x_j^2 - x_j) \;=\; 1.$$

Suppose for every $i \in [m]$ and for every $j \in [n]$, the polynomials $A_i(\mathbf{x})$ and $B_j(\mathbf{x})$ have a circuit of size s and depth Δ, then there exists a IPS *proof for the system $\{P_1, \ldots, P_r\}$ of size $\mathcal{O}(sm)$ and depth $\Delta + 2$.*

Proof of Lemma 1.20. Define the circuit $C(\mathbf{x}, \mathbf{y}, \mathbf{z})$ as follows:

$$C(\mathbf{x}, \mathbf{y}, \mathbf{z}) \;=\; \sum_{i=1}^{m} A_i(\mathbf{x}) \cdot y_i \;+\; \sum_{j=1}^{n} B_j(\mathbf{x}) \cdot z_j$$

Clearly $C(\mathbf{x}, \mathbf{0}, \mathbf{0}) = 0$ and $C(\mathbf{x}, f_1, \ldots, f_m, x_1^2 - x_1, \ldots, x_n^2 - x_n) = 1$. It is easy to verify the size and depth parameters of $C(\mathbf{x}, \mathbf{y}, \mathbf{z})$. ∎

Lemma 1.20 allows us to restrict our attention to finding an efficient (in terms of algebraic complexity) Nullstellensatz certificate, which yields a short IPS-proof.

2 Lower Bounds in Large Fields of Positive Characteristic

In this section, we will prove size lower bounds for several fragments of IPS over positive characteristic. As explained in Section 1.3.1, we start by proving a tight degree lower bound (Lemma 2.2) over positive characteristic. Using our positive characteristic variant of the degree lower bound, we then recover the lower bound results from [FSTW21] and [GHT22] over positive characteristic.

[14] There are infinitely many Nullstellensatz certificates for a system $\{f_1, \ldots, f_m\}$. To see this, suppose $m = 2$ and let (A_1, A_2) be a Nullstellensatz certificate. Then for any polynomial $g \in \mathbb{F}[\mathbf{x}]$, $(A_1 + gf_2, A_2 - gf_1)$ is also a Nullstellensatz certificate.

2.1 Degree Lower Bound for Arbitrary Characteristic

For any $\mathbf{a} \in \{0,1\}^n$, we use $|\mathbf{a}|$ to denote its Hamming weight. For any $\mathbf{a} = (a_1, \ldots, a_n) \in \{0,1\}^n$ and any subset of indices $S \subseteq [n]$, we use $\mathbf{a}_S$ to denote $\prod_{i \in S} a_i$. All the statements in this section work over fields of arbitrary characteristic.

First, we state a standard fact about multilinear polynomials, which will be useful in the main lemma.

Fact 2.1. *Let $f(\mathbf{x}) = \sum_{S \subseteq [n]} \lambda_S \mathbf{x}_S$ be a multilinear polynomial on n variables. Then,*

$$\lambda_{[n]} = \sum_{\mathbf{a} \in \{0,1\}^n} (-1)^{|\mathbf{a}|} f(\mathbf{a})$$

The next lemma is our main degree lower bound which shows that a multilinear polynomial for the inverse of a random linear form will have maximal degree. While similar statements have been observed in the literature (e.g. [Gri98, Proposition 2]), we give an explicit proof for the sake of completeness.

Lemma 2.2. *Let $\mathbb{F}$ and $\mathbb{F}'$ be fields such that $\mathbb{F}$ is a strict subfield of $\mathbb{F}'$. Let $n \in \mathbb{N}$ be a natural number and let $\mathbf{x}$ denote the tuple of variables $(x_1, \ldots, x_n)$. Fix any $\beta \in \mathbb{F}' \backslash \mathbb{F}$. For any $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$, let $f_{\boldsymbol{\alpha}}(\mathbf{x})$ be the unique multilinear polynomial that agrees with the function*

$$\frac{1}{\sum_{i=1}^n \alpha_i x_i - \beta}$$

on the Boolean cube $\{0,1\}^n$. Let $S \subseteq \mathbb{F}$ be any finite subset of the field. Then, for a uniformly random $\boldsymbol{\alpha} \sim S^n$:

$$\Pr_{\boldsymbol{\alpha} \sim S^n}[\deg f_{\boldsymbol{\alpha}}(\mathbf{x}) = n] \geqslant 1 - \frac{2^n - 1}{|S|}$$

Proof. By Fact 2.1, the coefficient of $\mathbf{x}_{[n]}$ in $f_{\boldsymbol{\alpha}}(\mathbf{x})$ is $\sum_{\mathbf{a} \in \{0,1\}^n} (-1)^{|\mathbf{a}|} f_{\boldsymbol{\alpha}}(\mathbf{a})$, or equivalently,

$$\sum_{V \subseteq [n]} (-1)^{|V|} \frac{1}{(\sum_{i \in V} \alpha_i) - \beta}$$

Based on the above expression, we define the rational function $\lambda_{[n]}(\mathbf{z})$ as follows.

$$\lambda_{[n]}(\mathbf{z}) := \sum_{V \subseteq [n]} (-1)^{|V|} \frac{1}{(\sum_{i \in V} z_i) - \beta}$$

We will use $N(\mathbf{z})$ and $D(\mathbf{z})$ to denote the numerator and denominator of $\lambda_{[n]}(\mathbf{z})$. For any $S \subseteq [n]$, we will use $L_S(\mathbf{z})$ to denote $\sum_{i \in S} z_i$. It follows that

$$N(\mathbf{z}) = \sum_{V \subseteq [n]} (-1)^{|V|} \prod_{T \subseteq [n] : T \neq V} (L_T(\mathbf{z}) - \beta)$$
$$D(\mathbf{z}) = \prod_{V \subseteq [n]} (L_V(\mathbf{z}) - \beta)$$

Since $\beta \in \mathbb{F}' \backslash \mathbb{F}$, $D(\boldsymbol{\alpha}) \neq 0$ for any $\boldsymbol{\alpha} \in \mathbb{F}$. If we prove that $N(\mathbf{z})$ is a non-zero polynomial, then by the Polynomial Identity Lemma (Theorem 1.16), for any finite subset $S \subseteq \mathbb{F}$, $\Pr_{\boldsymbol{\alpha} \sim S^n}[N(\boldsymbol{\alpha}) \neq 0] \geqslant$

$1 - \frac{2^n-1}{|S|}$, which implies that $\Pr_{\boldsymbol{\alpha}\sim S^n}[\lambda_{[n]}(\boldsymbol{\alpha}) \neq 0] \geqslant 1 - \frac{2^n-1}{|S|}$, and thus proves the theorem. Thus, it is enough to prove that some monomial in $N(\mathbf{z})$ has non-zero coefficient.

For $V \neq \varnothing$, $\prod_{T\subseteq[n]:T\neq V}(L_T(\mathbf{z}) - \beta)$ has degree at most $2^n - 2$ since $L_\varnothing(\mathbf{z}) - \beta$ will not increase the degree. The term $\prod_{T\neq\varnothing}(L_T(\mathbf{z}) - \beta)$ syntactically contributes monomials of degree $2^n - 1$ from $\prod_{T\neq\varnothing} L_T(\mathbf{z})$, but is possible that these coefficients vanish if the field $\mathbb{F}$ is of positive characteristic. We will show that there is a monomial of degree $2^n - 1$ with coefficient 1, and thus this monomial will survive over any field.

Claim 2.3. *The coefficient of the monomial*[15] $\prod_{i=1}^n z_i^{2^{i-1}}$ *in* $\prod_{T\neq\varnothing}(L_T(\mathbf{z}) - \beta)$ *is 1.*

Proof sketch. We would like to count the number of ways of collecting variables from each $L_T(\mathbf{z})$ to construct the required monomial. We first observe (via a simple counting argument) that for every $i \in [n]$, the number of subsets $T \subseteq [n]$ such that $\{j \in [n] : j > i\} \cap T = \varnothing$, and $i \in T$, is 2^{i-1}. Moreover, for each $i \in [n]$, if $\mathcal{T}_i$ is the collection of subsets with the above properties, then we observe that $\mathcal{T}_i \cap \mathcal{T}_j = \varnothing$ for all $i \neq j$, $i \in [n]$, $j \in [n]$.

With these observations, it inductively follows that for each $i \in [n]$, conditioned on the degree of variables $z_n, \ldots, z_{i+1}$ being correct (i.e. $z_j^{2^{j-1}}$), there is exactly one way of ensuring that the degree of z_i is 2^{i-1}: for each T that is one of the 2^{i-1} subsets satisfying the properties of the above observation, select the z_i's from $L_T(\mathbf{z})$. ■

■

Note that Lemma 2.2 is interesting only when the field size is large (at least 2^n), and that will be the case for subsequent lemmas as well. The next lemma proves a stronger version of the previous lemma: for a random linear form, the inverse of *every* restriction of the linear form (by setting some variables to 0) will have maximal degree.

Lemma 2.4. *Let $\mathbb{F}$ and $\mathbb{F}'$ be fields such that $\mathbb{F}$ is a strict subfield of $\mathbb{F}'$. Let $n \in \mathbb{N}$ be a natural number and let $\mathbf{x}$ denote the tuple of variables $(x_1, \ldots, x_n)$. Fix any $\beta \in \mathbb{F}'\backslash\mathbb{F}$. For any $\varnothing \neq U \subseteq [n]$, let $f_{\boldsymbol{\alpha},U}(\mathbf{x})$ be the unique multilinear polynomial that agrees with the function*

$$\frac{1}{\sum_{i\in U}\alpha_i x_i - \beta}$$

on the Boolean cube $\{0,1\}^n$. Let $S \subseteq \mathbb{F}$ be a finite subset of the field. Then, for an $\boldsymbol{\alpha} \sim S^n$ chosen uniformly at random:

$$\Pr_{\boldsymbol{\alpha}\sim S^n}[\exists \text{ a non-empty } U \subseteq [n] : \deg f_{\boldsymbol{\alpha},U}(\mathbf{x}) < |U|] \leqslant \sum_{\varnothing\neq U\subseteq[n]} \frac{2^{|U|}-1}{|S|} < \frac{2^{2n}}{|S|}$$

In particular, with probability at least $1 - (2^{2n}/|S|)$ over the choice of $\boldsymbol{\alpha} \sim S^n$, for every $U \subseteq [n]$, the leading monomial of $f_{\boldsymbol{\alpha},U}(\mathbf{x})$ is $c \cdot \prod_{i\in U} x_i$ for some $c \in \mathbb{F}\backslash\{0\}$.

Proof. This lemma is a simple application of union bound with the previous lemma. The previous lemma tells us that for a uniformly random $\boldsymbol{\alpha} \sim S^n$ and any $U \subseteq [n]$,

$$\Pr_{\boldsymbol{\alpha}\sim S^n}[\deg f_{\boldsymbol{\alpha},U}(\mathbf{x}) < |U|] \leqslant \frac{2^{|U|}-1}{|S|}$$

Union bound over all $U \subseteq [n]$ gives us the required statement. ■

[15] The same proof works for any monomial $\prod_{i=1}^n z_{\sigma(i)}^{2^{i-1}}$, where σ is an arbitrary permutation on $[n]$.

2.2 Sparse-$\mathrm{IPS}_{\mathrm{LIN}'}$ Lower Bound

The following claim from [FSTW21] proves a lower bound against sparse-$\mathrm{IPS}_{\mathrm{LIN}'}$ over fields of large characteristic.

Proposition 2.5 (Sparsity lower bound (Proposition 5.6 [FSTW21]))**.** *Let $n \geqslant 8$. Let $\mathbb{F}$ be a field of characteristic $> n$. Let $\beta \in \mathbb{F}\backslash\{0, \ldots, n\}$. Suppose $f(\mathbf{x})$ be a polynomial such that*

$$f(\mathbf{x}) \cdot \left(\sum_{i=1}^{n} x_i - \beta \right) \equiv 1 \left(\mathbf{x}^2 - \mathbf{x} \right)$$

where $\left(\mathbf{x}^2 - \mathbf{x}\right)$ denotes the ideal $(x_1^2 - x_1, \ldots, x_n^2 - x_n)$. Then, the sparsity of $f(\mathbf{x})$ is at least $2^{\frac{n}{4}-1}$.

The proof uses two observations.

1. ([FSTW21, Lemma 5.5]) If $f(\mathbf{x})$ has sparsity s, then a random restriction ρ will ensure that $\deg(\rho(f)) \leqslant \log(s) + 1$ with reasonable probability.
2. (Chernoff bound) A random restriction ρ will keep at least $n/4$ variables alive with reasonable probability.

By a union bound, we can find a random restriction ρ that ensures that the degree of $\rho(f)$ is at most $\log(s) + 1$ but at least $n/4$ variables survive ρ. In particular, $\rho(\sum_{i\in[n]} x_i - \beta) = \sum_{i\in S} x_i - \beta$ for some $S \subseteq [n]$ with $|S| \geqslant n/4$. But the degree lower bound in [FSTW21] tells us that the inverse of $\sum_{i\in S} x_i - \beta$ on the Boolean cube must have degree $\geqslant |S|$. Combining the above observations with the degree lower bound, we get that $n/4 \leqslant \log(s) + 1$ or $s \geqslant 2^{n/4-1}$.

The only part of the proof that requires $\operatorname{char} \mathbb{F} > n$ is the degree lower bound; the two observations work over all fields. Thus, we can replace their degree lower bound with Lemma 2.4 to recover the sparsity lower bound over large enough fields of arbitrary characteristic.

Theorem 2.6. *Let $n \geqslant 8$. Let $p \in \mathbb{N}$ be any prime. Let $\tilde{\mathbb{F}}$ be a field of characteristic p and size p^{2k}, where k is the smallest integer that satisfies $p^k > 2^{2n}$. Let β be an arbitrary element in $\tilde{\mathbb{F}}\backslash\mathbb{F}$, where $\mathbb{F}$ denotes the subfield of size p^k. For any $\boldsymbol{\alpha} \in \mathbb{F}^n$, let $f_{\boldsymbol{\alpha}}(\mathbf{x})$ be a polynomial which agrees with*

$$\frac{1}{\sum_{i\in[n]} \alpha_i x_i - \beta}$$

on the Boolean cube. Then there exists an $\boldsymbol{\alpha} \in \mathbb{F}^n$ such that $f_{\boldsymbol{\alpha}}$ has sparsity $\geqslant 2^{\Omega(n)}$

2.3 $\mathrm{roABP} - \mathrm{IPS}_{\mathrm{LIN}'}$ Lower Bound

Lemma 2.4 tells us that for a random choice of coefficients $\boldsymbol{\alpha}$ and any $U \subseteq [n]$, the inverse of $\sum_{i\in U}^{n} \alpha_i x_i - \beta$ has degree $|U|$ over the Boolean cube. The authors of [FSTW21] "lift" such maximal degree lower bounds to construct a polynomial $P(\mathbf{x})$ such that any roABP that computes (in *any* order of variables) the inverse of $P(\mathbf{x})$ over the Boolean cube requires exponential size. A high-level overview of their proof is as follows.

1. The optimal width of an roABP computing a polynomial g is captured exactly by the *coefficient dimension*[16] of g.

[16] These notions are defined with respect to a certain partition of the variables and any order of variables that is consistent with the specified partition.

2. The coefficient dimension of a polynomial g is at least as large as the *evaluation dimension* of g.

3. For $f(\mathbf{x},\mathbf{y}) := \sum_{i\in[n]} x_i y_i - \beta$, evaluations of f on $y \in \{0,1\}^n$ will be $f_S(\mathbf{x}) = \sum_{i\in S} x_i - \beta$ for various $S \subseteq [n]$.

4. By the degree lower bound in [FSTW21], any multilinear polynomial computing the inverse of f_S over the Boolean cube must have degree $|S|$. This eventually implies that the evaluation space of $g(\mathbf{x},\mathbf{y}) := \frac{1}{f(\mathbf{x},\mathbf{y})}$ over $y \in \{0,1\}^n$ will contain all the multilinear monomials on $\mathbf{x}$ variables. In particular, the evaluation dimension[17] of g is at least 2^n, and thus, any roABP computing g must have width $\geqslant 2^n$.

The only part of their proof that requires a restriction on the characteristic of the underlying field is the degree lower bound. The rest of their proof works with the degree lower bound in Lemma 2.4. In the rest of this section, we state the final theorems that follow using our degree lower bound in the proofs of [FSTW21]. For more details, we recommend the reader to refer to the appendix as well as [FSTW21].

Theorem 2.7 (Functional lower bound against roABP in a fixed order of variables)**.** *Let $n \in \mathbb{N}$. Let $p \in \mathbb{N}$ be any prime. Let $\tilde{\mathbb{F}}$ be a field of characteristic p and size p^{2k}, where k is the smallest integer that satisfies $p^k > 2^{2n}$. Let β be an arbitrary element in $\tilde{\mathbb{F}}\backslash\mathbb{F}$, where $\mathbb{F}$ denotes the subfield of size p^k. For any $\boldsymbol{\alpha} \in \mathbb{F}^n$, let $f_{\boldsymbol{\alpha}}(\mathbf{x},\mathbf{y})$ be a polynomial which agrees with*

$$\frac{1}{\sum_{i\in[n]} \alpha_i x_i y_i - \beta}$$

on the Boolean cube. Then there exists an $\boldsymbol{\alpha} \in \mathbb{F}^n$ such that any roABP *that computes $f_{\boldsymbol{\alpha}}$ in any order of variables where $\mathbf{x}$ precedes $\mathbf{y}$ requires width $\geqslant 2^n$.*

Theorem 2.8 (Functional lower bound against roABP in any order of variables)**.** *Let $n \in \mathbb{N}$. Let $p \in \mathbb{N}$ be any prime. Let $\tilde{\mathbb{F}}$ be a field of characteristic p and size p^{2k}, where k is the smallest integer that satisfies $p^k > \binom{2n}{n} 2^{2n}$. Let β be an arbitrary element in $\tilde{\mathbb{F}}\backslash\mathbb{F}$, where $\mathbb{F}$ denotes the subfield of size p^k. For any $\boldsymbol{\alpha} \in \mathbb{F}^{\binom{2n}{2}}$, let $f_{\boldsymbol{\alpha}}(\mathbf{x} = (x_i)_{i=1}^{2n}, \mathbf{z} = (z_{i,j})_{i,j\in[n]})$ be a polynomial which agrees with*

$$\frac{1}{\sum_{i<j} \alpha_{i,j} z_{i,j} x_i x_j - \beta}$$

on the Boolean cube. Then there exists an $\boldsymbol{\alpha} \in \mathbb{F}^{\binom{2n}{2}}$ such that any roABP *that computes $f_{\boldsymbol{\alpha}}$ in any order of variables requires size $\geqslant 2^n$.*

2.4 Multilinear-formula-IPS Lower Bound

Lower bounds against multilinear-formula-IPS follow from a coefficient dimension lower bound (see Lemma A.10) and the following theorem of Raz and Yehudayoff that connects multilinear formula size to coefficient dimension. Here, we present the version from [FSTW21, Theorem 3.13].

[17] Again, the order of variables will be important here, but one can also construct a polynomial which works against roABPs in *any* order of variables.

Theorem 2.9 (Raz-Yehudayoff [RY09][Raz09]). *Let $f \in \mathbb{F}[x_1, \ldots, x_{2n}, \mathbf{z}]$ be a multilinear polynomial and let $f_{\mathbf{z}}$ denote the polynomial f over the ring $\mathbb{F}[\mathbf{z}]$. Suppose for any balanced partition $(\mathbf{u}, \mathbf{v})$ of $\mathbf{x} = (x_1, \ldots, x_{2n})$:*

$$\dim_{\mathbb{F}(\mathbf{z})} \mathbf{Coeff}_{\mathbf{u}|\mathbf{v}}(f_{\mathbf{z}}) \geqslant 2^n$$

Then any multilinear formula for f requires size $\geqslant n^{\Omega(\log n)}$, and for $\Delta = o(\log n / \log\log n)$, any product-depth-$\Delta$ multilinear formula computing f will require size $\geqslant n^{\Omega\left(\frac{1}{\Delta^2}(\frac{n}{\log n})^{1/\Delta}\right)}$.

Theorem 2.10 (Functional lower bounds against multilinear formula). *Let $n \in \mathbb{N}$. Let $p \in \mathbb{N}$ be any prime. Let $\mathbb{F}_{p^{2k}}$ be a field of characteristic p and size p^{2k}, where k is the smallest integer that satisfies $p^k > \binom{2n}{n} 2^{2n}$. Let β be an arbitrary element in $\mathbb{F}_{p^{2k}} \backslash \mathbb{F}$, where $\mathbb{F}$ denotes the subfield of size p^k. For any $\boldsymbol{\alpha} \in \mathbb{F}^{\binom{2n}{n}}$, let $f_{\boldsymbol{\alpha}}(\mathbf{x} = (x_i)_{i=1}^{2n}, \mathbf{z} = (z_{i,j})_{i,j\in[n]})$ be a polynomial which agrees with*

$$\frac{1}{\sum_{i<j} \alpha_{i,j} z_{i,j} x_i x_j - \beta}$$

on the Boolean cube. There exists an $\boldsymbol{\alpha} \in \mathbb{F}^{\binom{2n}{n}}$ such that any multilinear-formula computing $f_{\boldsymbol{\alpha}}$ requires size $\geqslant n^{\Omega(\log n)}$ and for $\Delta = o(\log n / \log\log n)$, any product-depth-$\Delta$ multilinear-formula computing $f_{\boldsymbol{\alpha}}$ requires size $\geqslant n^{\Omega\left(\frac{1}{\Delta^2}(\frac{n}{\log n})^{1/\Delta}\right)}$.

While this immediately implies multilinear-formula-$\mathrm{IPS}_{\mathrm{LIN}'}$ lower bounds, one can observe (as noted in Lemma 5.2 of [FSTW21]) that any multilinear-formula-IPS refutation, by multilinearity, is a multilinear-formula-$\mathrm{IPS}_{\mathrm{LIN}'}$ refutation. Thus, the lower bounds work against multilinear-formula-IPS.

2.5 Constant-depth Multilinear $\mathrm{IPS}_{\mathrm{LIN}'}$ Lower Bound

In [GHT22], Govindasamy, Hakoniemi, and Tzameret prove super polynomial lower bounds against constant-depth multilinear $\mathrm{IPS}_{\mathrm{LIN}'}$ refutations of the subset sum variant

$$\sum_{i,j,k,l\in[n]} z_{i,j,k,l} x_i x_j x_k x_l - \beta$$

In particular, they prove the following theorem.

Theorem 2.11 (Constant-depth functional lower bounds [GHT22]). *Let $n, \Delta \in \mathbb{N}_+$ with $\Delta \leqslant \mathcal{O}(\log\log\log n)$ and assume that $\mathrm{char}(\mathbb{F}) = 0$. Let f be the multilinear polynomial such that*

$$f = \frac{1}{\sum_{i,j,k,l\in[n]} z_{i,j,k,l} x_i x_j x_k x_l - \beta}$$

over the Boolean cube. Then, any circuit of product-depth Δ computing f has size at least

$$n^{(\log n)^{\exp(-\mathcal{O}(\Delta))}}$$

We prove the same statement for large fields of arbitrary characteristic. Our proof exactly follows the structure of [GHT22]. Their proof requires the $\mathrm{char}\,\mathbb{F} = 0$ condition for two reasons:

1. They use the results of Limaye, Srinivasan, and Tavenas [LST21], which gave superpolynomial lower bounds against constant-depth circuits over any field $\mathbb{F}$ with $\mathrm{char}(\mathbb{F}) = 0$ or greater than the degree d of the hard polynomial. In particular, they use the result that over fields with $\mathrm{char}(\mathbb{F}) = 0$ or greater than d, any low-degree set-multilinear polynomial computed by a constant-depth circuit can also be computed by a set-multilinear constant-depth circuit.[18]

2. They use the degree lower bound for the multilinear representation of $1/(\sum_{i\in[n]} x_i - \beta)$, proved by Forbes, Shpilka, Tzameret, and Wigderson [FSTW21].

To deal with the first requirement, we use the recent beautiful result of Forbes [For24], which extends the results of [LST21] to arbitrary fields. In particular, we will use the following statement from [For24], which says that the set-multilinear projection of a constant-depth circuit can be efficiently computed by a constant-depth circuit over arbitrary fields.

Theorem 2.12. *[For24, Corollary 27]. Let $\mathbb{F}$ be an arbitrary field. Let $\mathbf{x} = \mathbf{x}_1 \sqcup \mathbf{x}_2 \sqcup \cdots \sqcup \mathbf{x}_d$ be a partition of the variables $\mathbf{x}$. Suppose f can be computed by a size s product-depth Δ arithmetic circuit. Then the set-multilinear projection of f (the restriction of f to monomials that are set-multilinear with respect to the specified partition) can be computed by a size $\mathrm{poly}(s, \Theta(\frac{d}{\log d})^d)$-size circuit of product-depth 2Δ.*

To deal with the second requirement, we use our degree lower bound from Lemma 2.4, which works for arbitrary fields of exponential size i.e. there is no restriction on the characteristic of the field.

Overview of [GHT22]

1. Using the *word polynomials* framework of [LST21], construct a *knapsack polynomial* $\mathrm{ks}_{\mathbf{w}}$ (for a partition given by a word $w \in \mathbb{Z}^d$) with the property that the set-multilinear projection of $\frac{1}{\mathrm{ks}_{\mathbf{w}}}$ over the Boolean cube requires superpolynomially large set-multilinear constant-depth circuits.

2. Consider a degree-4 subset-sum variant $f(\mathbf{z}, \mathbf{x}) := \sum_{i,j,k,l} z_{i,j,k,l} x_i x_j x_k x_l - \beta$ so that for the word $w \in \mathbb{Z}^d$ that will be used to instantiate the previous point, there exists an assignment of some of the variables in $\mathbf{z}$, $\mathbf{x}$ that maps $f(\mathbf{z}, \mathbf{x})$ to $\mathrm{ks}_{\mathbf{w}}$ (upto a renaming of variables).

3. If there is a multilinear polynomial computing $1/f(\mathbf{z}, \mathbf{x})$ over $\{0,1\}^n$ that has a small constant-depth circuit, then there is a multilinear polynomial computing $1/\mathrm{ks}_{\mathbf{w}}$ over $\{0,1\}^n$ that has a small constant-depth circuit. Moreover by the set-multilinearization of [LST21], there is a small set-multilinear constant-depth circuit computing the set-multilinear projection of $1/\mathrm{ks}_{\mathbf{w}}$.

4. Combining the first point with the contrapositive of the third point, conclude that any multilinear polynomial computing $1/f(\mathbf{z}, \mathbf{x})$ over $\{0,1\}^n$ requires superpolynomially large constant-depth circuits. The multilinear constant-depth $\mathrm{IPS}_{\mathrm{LIN}'}$ lower bound follows.

In [GHT22], the proof for the hardness of $\frac{1}{\mathrm{ks}_{\mathbf{w}}}$ requires the underlying field to be of large characteristic, essentially because it requires the degree lower bound from [FSTW21], which requires large

[18] They also use other ideas from [LST21] such as relative rank, word polynomial, etc., but those ideas do not require any restrictions on the characteristic of the underlying field.

characteristic. To make Theorem 2.11 work over fields of positive characteristic, we will employ our degree lower bound from Lemma 2.4 with a variant of the knapsack polynomial; the rest of the proof remains the same as that of Theorem 2.11. To provide the necessary details, we first describe the construction of the knapsack polynomial. Then, we state the particular claim from [GHT22] that uses the degree lower bound from [FSTW21]. Finally, we show how our degree lower bound Lemma 2.4 fits into the rest of the proof.

Constructing the knapsack polynomial We shall now recall the definitions required for defining the hard polynomial in [GHT22] via the word polynomials template of [LST21].

Let $\mathbf{w} \in \mathbb{Z}^d$ be an arbitrary word. For any $S \subseteq [d]$, let $w|_S$ denote the subword of w indexed by the set S. Consider the sequence $\overline{X}(w) = (X(w_1), \ldots, X(w_d))$ of sets of variables. Define the *positive indices* and *negative indices* of $\mathbf{w}$ as:

$$P_{\mathbf{w}} := \{i \in [d] : w_i \geqslant 0\}$$

$$N_{\mathbf{w}} := \{i \in [d] : w_i < 0\}$$

Let any $i \in P_{\mathbf{w}}$, the variables of $X(w_i)$ will be of the form $x_\sigma^{(i)}$, where σ is a binary string indexed by the set:

$$A_{\mathbf{w}}^{(i)} := \left[\sum_{\substack{i' \in P_{\mathbf{w}} \\ i' < i}} w_{i'} + 1, \sum_{\substack{i' \in P_{\mathbf{w}} \\ i' \leqslant i}} w_{i'} \right]$$

We will call these sets *positive indexing sets*. The size of each $A_{\mathbf{w}}^{(i)}$ is $|w_i|$. The number of strings in $A_{\mathbf{w}}^{(i)}$ is $2^{|w_i|}$.
For $i \in N_{\mathbf{w}}$, we similarly define the *negative indexing sets* $B_{\mathbf{w}}^{(i)}$ that will be used to index the variables of $X(w_i)$ for $i \in N_{\mathbf{w}}$.
A word $w \in \mathbb{Z}^d$ is *balanced* if:

- $\forall i \in P_{\mathbf{w}}\ \exists j \in N_{\mathbf{w}}$ such that $A_{\mathbf{w}}^{(i)} \cap B_{\mathbf{w}}^{(j)} \neq \varnothing$ (i.e. $j \in N_{\mathbf{w}}$ is a *witness* that $\mathbf{w}$ is balanced at $i \in P_{\mathbf{w}}$)
- $\forall j \in N_{\mathbf{w}}\ \exists i \in P_{\mathbf{w}}$ such that $A_{\mathbf{w}}^{(i)} \cap B_{\mathbf{w}}^{(j)} \neq \varnothing$ (i.e. $i \in P_{\mathbf{w}}$ is a *witness* that $\mathbf{w}$ is balanced at $j \in N_{\mathbf{w}}$)

For any $i \in P_{\mathbf{w}}, \sigma \in \{0,1\}^{A_{\mathbf{w}}^{(i)}}$, define:

$$f_\sigma^{(i)} := \prod_{\substack{j \in N_{\mathbf{w}} \\ A_{\mathbf{w}}^{(i)} \cap B_{\mathbf{w}}^{(j)} \neq \varnothing}} \sum_{\substack{\sigma_j \in \{0,1\}^{B_{\mathbf{w}}^{(j)}} \\ \sigma_j(k) = \sigma(k) \forall k \in A_{\mathbf{w}}^{(i)} \cap B_{\mathbf{w}}^{(j)}}} y_{\sigma_j}^{(j)} \tag{3}$$

The product ranges over each $j \in N_{\mathbf{w}}$ that witnesses the fact that $\mathbf{w}$ is balanced at i. The sum ranges over each σ_j that is consistent with σ on $A_{\mathbf{w}}^{(i)} \cap B_{\mathbf{w}}^{(j)}$. Now, we define the knapsack polynomial as

$$\mathrm{ks}_{\mathbf{w}} := \left(\sum_{i \in P_{\mathbf{w}}} \sum_{\sigma \in \{0,1\}^{A_{\mathbf{w}}^{(i)}}} x_\sigma^{(i)} f_\sigma^{(i)} \right) - \beta \tag{4}$$

where $\beta \in \mathbb{F}$ is any field element such that $\mathrm{ks}_{\mathbf{w}}$ has no Boolean roots.

To make the proof work over fields of positive characteristic, we define a variant of $\mathrm{ks}_{\mathbf{w}}$ as:

$$\mathrm{ks}_{\mathbf{w},\boldsymbol{\alpha}} := \left(\sum_{i \in P_{\mathbf{w}}} \alpha_i \sum_{\sigma \in \{0,1\}^{A^{(i)}_{\mathbf{w}}}} x^{(i)}_{\sigma} f^{(i)}_{\sigma} \right) - \beta \tag{5}$$

where $\boldsymbol{\alpha} = (\alpha_i)_{i \in P_{\mathbf{w}}} \in \mathbb{F}^{|P_{\mathbf{w}}|}$, and β will be chosen from an extension field $\tilde{\mathbb{F}} \supset \mathbb{F}$ so that $\mathrm{ks}_{\mathbf{w},\boldsymbol{\alpha}}$ has no Boolean roots.

For any word $\mathbf{w} \in \mathbb{Z}^d$, $M_{\mathbf{w}}(f)$ denotes the matrix with rows indexed by all monomials m that are set-multilinear over $\mathbf{w}|_{P_{\mathbf{w}}}$, and columns indexed by all monomials m' that are set-multilinear over $\mathbf{w}|_{N_{\mathbf{w}}}$. For each such pair of monomials (m, m'), the corresponding entry in $M_{\mathbf{w}}(f)$ carries the coefficient of mm' in f. To show that the set-multilinear projection of any multilinear polynomial f computing $1/\mathrm{ks}_{\mathbf{w}}$ over $\{0,1\}^n$ requires superpolynomially large set-multilinear constant-depth circuits, [GHT22] shows that $M_{\mathbf{w}}(f)$ is full-rank.

Lemma 2.13 (Rank lower bound lemma (Lemma 6 [GHT22]))**.** *Let $\mathbf{w} \in \mathbb{Z}^d$ be a balanced word, and let f be the multilinear polynomial such that*

$$f = \frac{1}{\mathrm{ks}_{\mathbf{w}}}$$

over $\{0,1\}^n$. Then, $M_{\mathbf{w}}(f)$ is full-rank.

With this lemma, the lower bound follows via the arguments from [LST21]. Importantly for us, this lemma uses the degree lower bound from [FSTW21]; we describe a sketch of the same.

The use of degree lower bound in [GHT22] Suppose $f = \sum_m g_m(\mathbf{x})m$, where the sum runs over all multilinear monomials m in the $\mathbf{y}$ variables, and $g_m(\mathbf{x})$ is some multilinear polynomial in the $\mathbf{x}$ variables. They show that for any m which is set-multilinear on $\mathbf{w}|_{N_{\mathbf{w}}}$, the leading monomial of $g_m(\mathbf{x})$ is the set-multilinear monomial m' on positive variables such that $\sigma(m')$ is consistent with $\sigma(m)$ ([GHT22] describes this formally). For each monomial m that is set-multilinear on $\mathbf{w}|_{N_{\mathbf{w}}}$, the leading monomial of $g_m(\mathbf{x})$ turns out to be a different set-multilinear monomial on the positive variables, and together, these leading monomials span the space of all set-multilinear monomials on the positive variables. This makes $M_{\mathbf{w}}(f)$ full-rank. To get a handle on $g_m(\mathbf{x})$ (for m being a monomial on $\mathbf{w}|_{N_{\mathbf{w}}}$, consisting only of $\mathbf{y}$ variables), [GHT22] sets all the variables in m to 1 and all the $\mathbf{y}$ variables outside m to 0. They call this transformation τ_m. For the proof of Lemma 2.13, an important requirement is that:

> For every $T \subseteq N_{\mathbf{w}}$ and for every set-multilinear monomial m on $\mathbf{w}|_T$, the leading monomial of $\tau_m(f)$ is $\prod_{i \in U_T} x^{(i)}_{\sigma_i}$, which is the product of all the variables that show up in the denominator of
>
> $$\frac{1}{\tau_m(\mathrm{ks}_{\mathbf{w}})} = \frac{1}{\sum_{i \in U_T} x^{(i)}_{\sigma(i)} - \beta}$$
>
> where $U_T = \{i \in P_{\mathbf{w}} : A^{(i)}_{\mathbf{w}} \subseteq B^T_{\mathbf{w}}\}$, and for each $i \in P_{\mathbf{w}}$, $\sigma(i)$ is the unique indexing string that agrees with $\sigma(m)$ on $A^{(i)}_{\mathbf{w}}$, the i^{th} positive indexing set.

This requirement is satisfied due to the degree lower bound from [FSTW21], which requires the field to be of characteristic 0. The proof in [GHT22] includes helpful figures and the reader is encouraged to refer to the paper.
Let us recall our variant of $\mathrm{ks}_{\mathbf{w}}$:

$$\mathrm{ks}_{\mathbf{w},\boldsymbol{\alpha}} := \left(\sum_{i \in P_{\mathbf{w}}} \alpha_i \sum_{\sigma \in \{0,1\}^{A_{\mathbf{w}}^{(i)}}} x_{\sigma}^{(i)} f_{\sigma_i} \right) - \beta \tag{6}$$

where $\boldsymbol{\alpha} = (\alpha_i)_{i \in P_{\mathbf{w}}} \in \mathbb{F}^{|P_{\mathbf{w}}|}$. To prove Theorem 2.11 in positive characteristic, we use the following lemma that follows by a union bound over all $T \subseteq N_{\mathbf{w}}$ and all set-multilinear monomials on $\mathbf{w}|_T$, on top of Lemma 2.2.

Lemma 2.14. *Let $d \in \mathbb{N}$ be a natural number and $\mathbf{w} \in \mathbb{Z}^d$ be a balanced word. Let $m = |P_{\mathbf{w}}|$. For any $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m) \in \mathbb{F}^m$, $T \subseteq N_{\mathbf{w}}$ and any m_T that is a set-multilinear monomial on $\mathbf{w}|_T$, let $f_{\boldsymbol{\alpha},T,m_T}(\mathbf{x})$ be the unique multilinear polynomial that agrees with the function*

$$\tau_{m_T}\left(\frac{1}{\mathrm{ks}_{\mathbf{w},\boldsymbol{\alpha}}}\right) = \frac{1}{\sum_{i \in U_T} \alpha_i x_{\sigma(i)}^{(i)} - \beta}$$

on the Boolean cube, where $\beta \in \mathbb{F}$ is chosen so that $\mathrm{ks}_{\mathbf{w},\boldsymbol{\alpha}}$ has no Boolean roots, and $U_T = \{i \in P_{\mathbf{w}} : A_{\mathbf{w}}^{(i)} \subseteq B_{\mathbf{w}}^{T}\}$. Let $S \subseteq \mathbb{F}$ be a finite subset of the field. Let $\gamma := |N_{\mathbf{w}}| + \sum_{i \in N_{\mathbf{w}}} |w_i|$. Then, for an $\boldsymbol{\alpha} \in S^m$ chosen uniformly at random:

$$\Pr_{\boldsymbol{\alpha} \sim S^m}[\exists T \subseteq N_{\mathbf{w}}, m_T : \deg f_{\boldsymbol{\alpha},T,m_T}(\mathbf{x}) < |U_T|] < \frac{2^{\gamma+m}}{|S|}$$

In particular, with probability at least $1 - (2^{\gamma+m}/|S|)$ over the choice of $\boldsymbol{\alpha} \in S^m$, for every choice of $T \subseteq N_{\mathbf{w}}$ and set-multilinear monomial m_T over $\mathbf{w}|_T$, the leading monomial of $f_{\boldsymbol{\alpha},T,m_T}(\mathbf{x})$ is $c \cdot \prod_{i \in U_T} x_{\sigma_i}^{(i)}$ for some $c \in \mathbb{F} \setminus \{0\}$.

Proof. The number of $T \subseteq N_{\mathbf{w}}$ is $2^{|N_{\mathbf{w}}|}$. The number of set-multilinear monomials on $\mathbf{w}|_T$ for any $T \subseteq N_{\mathbf{w}}$ is $2^{\sum_{i \in T} |w_i|}$, which is at most $2^{\sum_{i \in N_{\mathbf{w}}} |w_i|}$. For any fixed $T \subseteq N_{\mathbf{w}}$ and m_T that is a set-multilinear monomial on $\mathbf{w}|_T$, Lemma 2.2 implies that for an $\boldsymbol{\alpha} \in S^m$ chosen uniformly at random:

$$\Pr_{\boldsymbol{\alpha} \in S^m}[\deg f_{\boldsymbol{\alpha},T,m_T}(\mathbf{x}) < |U_T|] < \frac{2^m}{|S|}$$

Applying a union bound over all $T \subseteq N_{\mathbf{w}}$ and m_T implies that for an $\boldsymbol{\alpha} \in S^m$ chosen uniformly at random:

$$\Pr_{\boldsymbol{\alpha} \in S^m}[\exists T \subseteq N_{\mathbf{w}}, m_T : \deg f_{\boldsymbol{\alpha},T,m_T}(\mathbf{x}) < |U_T|] < \sum_{T \subseteq N_{\mathbf{w}}, m_T} \frac{2^m}{|S|} \leqslant \frac{2^{\gamma+m}}{|S|}$$

■

With this lemma, the rest of the proof of [GHT22] works out verbatim. We state the final theorem, which is a version of Theorem 2.11 for finite fields of positive characteristic.

Theorem 2.15 ([GHT22] over positive characteristic)**.** *Let $n, \Delta \in \mathbb{N}_+$ with $\Delta \leqslant \mathcal{O}(\log\log\log n)$. Let $p \in \mathbb{N}$ be any prime. Let $\tilde{\mathbb{F}}$ be a field of characteristic p and size p^{2k}, where k is the smallest integer that satisfies $p^k > 2^{C(\log n)^2}$ for an absolute constant[19] $C \geqslant 1$. Let β be an arbitrary element in $\tilde{\mathbb{F}}\backslash\mathbb{F}$, where $\mathbb{F}$ denotes the subfield of size p^k. For any $\boldsymbol{\alpha} \in \mathbb{F}^{n^4}$, Let $f_{\boldsymbol{\alpha}}$ be the multilinear polynomial such that*

$$f = \frac{1}{\sum_{i,j,k,l\in[n]} \alpha_{i,j,k,l} z_{i,j,k,l} x_i x_j x_k x_l - \beta}$$

over the Boolean cube. Then, there exists an $\boldsymbol{\alpha} \in \mathbb{F}^{n^4}$ such that any circuit of product-depth Δ computing $f_{\boldsymbol{\alpha}}$ has size at least

$$n^{(\log n)^{\exp(-\mathcal{O}(\Delta))}}$$

The reason for $|\mathbb{F}| > 2^{\Omega((\log n)^2)}$ in Theorem 2.15 : When we instantiate Lemma 2.14 inside the proof of Theorem 2.15, the parameter d, which is the number of variable sets, will be $O(\log n)$, and the word $\mathbf{w} \in \mathbb{Z}^d$ will also be chosen so that for each $i \in [d]$, $|w_i| \leqslant O(\log n)$. Thus, $\sum_{i\in N_{\mathbf{w}}} |w_i| = O((\log n)^2)$, and fighting the union bound in Lemma 2.14 will require the field to be larger than $2^{O((\log n)^2)}$.

3 Non-multilinear Upper Bounds

3.1 Proof of Theorem 1.8

In this section, we prove Theorem 1.8. We start by proving it for a restricted setting when the polynomial $f(\mathbf{x})$ is a degree-1 polynomial. In particular, we prove Theorem 3.1, stated below.

Theorem 3.1 (Upper bounds for (non-multilinear) constant-depth-IPS_{LIN} in positive characteristic)**.** *Fix a prime number p. The following holds for any natural numbers n and k.*
Let $L \in \mathbb{F}[x_1, \ldots, x_n]$ be a degree-1 polynomial with coefficients from the $\mathbb{F}_{p^k}$ and let β be any element of $\mathbb{F}\backslash\mathbb{F}_{p^k}$ where $\mathbb{F}$ is a field extension of $\mathbb{F}_{p^k}$.
Then,

- *The polynomial $L(\mathbf{x}) - \beta$ has no satisfying assignment over the Boolean cube $\{0,1\}^n$*
- *There is a constant-depth-IPS_{LIN} refutation of degree $\mathcal{O}(k \cdot p)$ and size $\mathcal{O}(k \cdot np)$.*

Over fields of large enough characteristic, [FSTW21, Proposition 4.15] showed that $L(\mathbf{x}) - \beta$ has a constant-depth *multilinear*-IPS_{LIN} refutation of size that depends on the number of possible values $L(\mathbf{x})$ could take over $\{0,1\}^n$. Theorem 3.1 shows that if we allow non-multilinear IPS_{LIN} refutation, then the circuit size is small.

Proof of Theorem 3.1. Firstly, since the coefficients of the polynomial $L(\mathbf{x})$ are in the field $\mathbb{F}_{p^k}$, $L(\mathbf{x})$ cannot be equal to $\beta \notin \mathbb{F}_{p^k}$ for any $\mathbf{x} \in \{0,1\}^n$. In other words, $L(\mathbf{x}) - \beta$ has no satisfying assignment over the Boolean cube $\{0,1\}^n$.
To show the existence of a low-degree constant-depth-IPS_{LIN} refutation, we will use Lemma 1.20. In

[19] This C is a fixed constant that depends on the exact choice of parameters in the proof of [GHT22]

particular, Lemma 1.20 says that it is sufficient to prove that there exists polynomials $A(\mathbf{x}), B_1(\mathbf{x}), \ldots, B_n(\mathbf{x})$ such that

$$A(\mathbf{x}) \cdot (L(\mathbf{x}) - \beta) + \sum_{j=1}^{n} B_j(\mathbf{x}) \cdot (x_j^2 - x_j) \;=\; 1,$$

where $A(\mathbf{x}), B_1(\mathbf{x}), \ldots, B_n(\mathbf{x})$ are low-degree polynomials and have constant-depth circuits of size $\text{poly}(n)$.

Without loss of generality, we can assume that $L(\mathbf{x})$ is a homogeneous degree-1 polynomial[20] because of the following reason. If $L(\mathbf{x})$ has a non-zero constant term $\alpha_0 \in \mathbb{F}_{p^k}$, then we can work with $(\alpha_0 + \beta) \in \mathbb{F} \backslash \mathbb{F}_{p^k}$, instead of $\beta \in \mathbb{F} \backslash \mathbb{F}_{p^k}$.

Suppose $L(\mathbf{x}) \;=\; \alpha_1 x_1 + \cdots + \alpha_n x_n$, where for each $i \in [n]$, the coefficient $\alpha_i \in \mathbb{F}_{p^k}$. For any natural number $0 \leqslant j \leqslant k$, we define $L_j(\mathbf{x})$ to be the following degree-1 polynomial:

$$L_j(\mathbf{x}) \;:=\; \alpha_1^{p^j} x_1 + \cdots + \alpha_n^{p^j} x_n - \beta^{p^j}$$

In the above notation, $L_0(\mathbf{x}) = L(\mathbf{x}) - \beta$. The next claim shows that we can express $L_j(\mathbf{x})$ as a multiple of $L_0(\mathbf{x})$ modulo the ideal[21] $(\mathbf{x}^p - \mathbf{x})$.

Claim 3.2. *For every $j \in [k]$, there exists polynomials $A_j(\mathbf{x}), B_{j,1}(\mathbf{x}), \ldots, B_{j,n}(\mathbf{x})$ such that:*

$$L_j(\mathbf{x}) \;=\; A_j(\mathbf{x}) \cdot L_0(\mathbf{x}) + \sum_{i=1}^{n} B_{j,i}(\mathbf{x}) \cdot (x_i^p - x_i),$$

where each polynomial $A_j(\mathbf{x}), B_{j,1}(\mathbf{x}), \ldots, B_{j,n}(\mathbf{x})$

- *The polynomial $A_j(\mathbf{x})$ is a degree-$\mathcal{O}(j \cdot p)$ polynomial and has a circuit of size $\mathcal{O}(j \cdot (n+p))$ and depth 2.*
- *For each $j \in [n]$, the polynomial $B_j(\mathbf{x})$ is a degree-$\mathcal{O}(j \cdot p)$ polynomial and has a circuit of size $\mathcal{O}(j \cdot np + j^2)$ and depth 3.*

Proof of Claim 3.2. The proof is via induction on j.

Base case ($j = 1$): As we are working over a field $\mathbb{F}$ of characteristic p, we have:

$$L_0(\mathbf{x})^p \;=\; \left(\sum_{i=1}^{n} \alpha_i x_i - \beta\right)^p \;=\; \sum_{i=1}^{n} \alpha^p x_i^p - \beta^p \qquad \text{(Using Lemma 1.17)}$$

$$\Rightarrow L_0(\mathbf{x})^p \;=\; \underbrace{\left(\sum_{i=1}^{n} \alpha_i^p x_i - \beta^p\right)}_{=L_1(\mathbf{x})} + \sum_{i=1}^{n} \alpha_i^p \cdot (x_i^p - x_i) \qquad \text{(Adding and subtracting terms)}$$

[20] For the sake of less cumbersome notation

[21] Recall that $(\mathbf{x}^p - \mathbf{x}) = (x_1^p - x_1, \ldots, x_n^p - x_n)$

$$\Rightarrow L_1(\mathbf{x}) \;=\; \underbrace{L_0(\mathbf{x})^{p-1}}_{:=A_1(\mathbf{x})} \cdot L_0(\mathbf{x}) + \sum_{i=1}^{n} B_{1,i}(\mathbf{x}) \cdot (x_i^p - x_i),$$

where

- $A_1(\mathbf{x}) = L_0(\mathbf{x})^{p-1}$ has a circuit of size $\mathcal{O}(np)$ and depth 2 (a $\Pi\Sigma$ circuit). Additionally, $A_1(\mathbf{x})$ is a degree-$\mathcal{O}(p)$ polynomial.
- For each $i \in [n]$, $B_{1,i}(\mathbf{x}) = -\alpha_i^p$ has a circuit of size $\mathcal{O}(1)$ and depth 1. Additionally, $B_{1,i}(\mathbf{x})$ is a constant, so has degree-0.

Induction step: Now assume the induction hypothesis is true for some $1 \leqslant j < k$. Proceeding similarly to the base case, we have,

$$L_j(\mathbf{x})^p = \left(\sum_{i=1}^{n} \alpha_i^{p^j} x_i - \beta^{p^j}\right)^p \;=\; \sum_{i=1}^{n} \alpha^{p^{j+1}} x_i^p - \beta^{p^{j+1}} \qquad \text{(Using Lemma 1.17)}$$

$$\Rightarrow L_j(\mathbf{x})^p \;=\; \underbrace{\left(\sum_{i=1}^{n} \alpha_i^{p^{j+1}} x_i - \beta^{p^{j+1}}\right)}_{:=L_{j+1}(\mathbf{x})} + \sum_{i=1}^{n} \alpha_i^{p^{j+1}} \cdot (x_i^p - x_i) \qquad \text{(Adding and subtracting terms)}$$

$$\Rightarrow L_{j+1}(\mathbf{x}) \;=\; L_j(\mathbf{x}) \cdot L_j(\mathbf{x})^{p-1} + \sum_{i=1}^{n} (-\alpha_i^{p^{j+1}}) \cdot (x_i^p - x_i) \tag{7}$$

Using the induction hypothesis, we know there exists polynomials $A_j(\mathbf{x}), B_{j,1}(\mathbf{x}), \ldots, B_{j,n}(\mathbf{x})$ such that

$$L_j(\mathbf{x}) \;=\; A_j(\mathbf{x}) \cdot L_0(\mathbf{x}) + \sum_{i=1}^{n} B_{j,i}(\mathbf{x}) \cdot (x_i^p - x_i),$$

where the polynomials satisfy the size constraints as stated in Claim 3.2. Substituting this in Equation (7), we get,

$$L_{j+1}(\mathbf{x}) \;=\; \left(A_j(\mathbf{x}) \cdot L_0(\mathbf{x}) + \sum_{i=1}^{n} B_{j,i}(\mathbf{x}) \cdot (x_i^p - x_i)\right) \cdot L_j(\mathbf{x})^{p-1} + \sum_{i=1}^{n} (-\alpha_i^{p^{j+1}}) \cdot (x_i^p - x_i)$$

$$\Rightarrow L_{j+1}(\mathbf{x}) \;=\; \underbrace{(A_j(\mathbf{x}) \cdot L_j(\mathbf{x})^{p-1})}_{:=A_{j+1}(\mathbf{x})} \cdot L_0(\mathbf{x}) \;+\; \sum_{i=1}^{n} \underbrace{(B_{j,i}(\mathbf{x}) L_j(\mathbf{x})^{p-1} - \alpha_i^{p^{j+1}})}_{:=B_{j+1,i}(\mathbf{x})} \cdot (x_i^p - x_i)$$

Now,

- The polynomial $A_{j+1}(\mathbf{x})$ has a circuit of size $\mathcal{O}((j+1) \cdot np)$ and depth 2 (note that $A_j(\mathbf{x})$ is a product of powers of linear polynomials). Additionally, $A_{j+1}(\mathbf{x})$ is a degree-$((j+1) \cdot p)$ polynomial.

- For every $i \in [n]$, the polynomial $B_{j+1,i}(\mathbf{x})$ has a circuit of size $\mathcal{O}((j+1)\cdot np + (j+1)^2)$ and depth 3 (note that $B_j(\mathbf{x})$ is a $\Sigma\Pi\Sigma$ circuit). Additionally, $B_{j+1,i}(\mathbf{x})$ is a degree-$\mathcal{O}((j+1)\cdot p)$ polynomial.

This finishes the induction and also the proof of Claim 3.2. ■

So far in Claim 3.2, we have shown that the linear polynomial $L_k(\mathbf{x})$ is a multiple of the linear polynomial $L_0(\mathbf{x})$ modulo the ideal $(\mathbf{x}^p - \mathbf{x})$. Next we use the fact that $\beta \notin \mathbb{F}_{p^k}$ to show that $L_k(\mathbf{x})$ and $L_0(\mathbf{x})$ differ by a non-zero constant.

Observation 3.3. *The polynomial $L_k(\mathbf{x}) - L_0(\mathbf{x})$ is a* non-zero constant *polynomial. This is because $\alpha_i^{p^k} = \alpha_i$ (since $\alpha_i \in \mathbb{F}_{p^k}$) and on the other hand, $\beta^{p^k} \neq \beta$.*

Claim 3.2 gives us that there exists polynomials $A_k(\mathbf{x}), B_{k,1}(\mathbf{x}), \ldots, B_{k,n}(\mathbf{x})$ satisfying:

$$\begin{aligned} L_k(\mathbf{x}) &= A_k(\mathbf{x})\cdot L_0(\mathbf{x}) + \sum_{i=1}^{n} B_{k,i}(\mathbf{x})\cdot(x_i^p - x_i) \\ \Rightarrow L_k(\mathbf{x}) - L_0(\mathbf{x}) &= (A_k(\mathbf{x}) - 1)\cdot L_0(\mathbf{x}) + \sum_{i=1}^{n} B_{k,i}(\mathbf{x})\cdot(x_i^p - x_i) \\ \Rightarrow \frac{(A_k(\mathbf{x})-1)}{\beta^{p^k}-\beta}\cdot L_0(\mathbf{x}) &+ \sum_{i=1}^{n} \frac{B_{k,i}(\mathbf{x})}{\beta^{p^k}-\beta}\cdot(x_i^p - x_i) = 1, \end{aligned} \tag{8}$$

where in the final implication we used that $L_k(\mathbf{x}) - L_0(\mathbf{x}) = \beta^k - \beta$ is a non-zero constant. For each $i \in [n]$, the polynomial $(x_i^p - x_i)$ is a multiple of $(x_i^2 - x_i)$ because:

$$x_i^p - x_i = (x_i^{p-2} + \cdots + x_i + 1)\cdot(x_i^2 - x_i)$$

Substituting it back in Equation (8), we get,

$$\underbrace{\frac{(A_k(\mathbf{x})-1)}{\beta^{p^k}-\beta}}_{:=A(\mathbf{x})}\cdot L_0(\mathbf{x}) + \sum_{i=1}^{n} \underbrace{\left(\frac{B_{k,i}(\mathbf{x})}{\beta^{p^k}-\beta}\cdot(x_i^{p-2} + \cdots + x_i + 1)\right)}_{:=B_i(\mathbf{x})}\cdot(x_i^2 - x_i) = 1$$

Degree and Size Analysis We define $A(\mathbf{x})$ and $B_i(\mathbf{x})$ as follows:

- Claim 3.2 says that $A_k(\mathbf{x})$ is a degree-$\mathcal{O}(kp)$ polynomial and is computable by circuit of size $\mathcal{O}(k(n+p))$ and depth 2 (a $\Pi\Sigma$ circuit). Hence $A(\mathbf{x})$ is a degree-$\mathcal{O}(kp)$ polynomial and is computable by a circuit of size $\mathcal{O}(k(n+p))$ and depth 3 (a $\Sigma\Pi\Sigma$ circuit).
- Claim 3.2 says that $B_{k,i}(\mathbf{x})$ is a degree-$\mathcal{O}(kp)$ polynomial and is computable by a circuit of size $\mathcal{O}(knp + k^2)$ and depth 3 (a $\Sigma\Pi\Sigma$ circuit). Hence $B_i(\mathbf{x})$ is a degree-$\mathcal{O}(kp)$ polynomial and is computable by a circuit of size $\mathcal{O}(knp + k^2)$ and depth 3.

Thus we have shown that there is a low-degree constant-depth-IPS_{LIN} refutation of $L(\mathbf{x}) - \beta$ and this finishes the proof of Theorem 3.1. ■

Now we ready to prove Theorem 1.8 using Theorem 3.1. The idea is to replace each monomial in the sparse polynomial by a new variable, resulting in a linear polynomial in the new variables. A refutation of the resulting linear polynomial can be "lifted" to a refutation of the sparse polynomial in the original variables. We use the refutation of linear polynomials from Theorem 3.1, and to lift this refutation, we need to show that *monomial axioms* are in the ideal of the Boolean axioms. Before proceeding, we will prove the following claim on monomial axioms. It follows from a straightforward induction on the number of variables. We will omit the proof here, and it can be found in Appendix A.2.

Claim 3.4. *For any exponent vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ with $|\boldsymbol{\mu}| \leqslant D$, there exists polynomials $E_{\boldsymbol{\mu},1}(\mathbf{x}), \ldots, E_{\boldsymbol{\mu},n}(\mathbf{x})$ such that the following holds:*

$$((\mathbf{x}^{\boldsymbol{\mu}})^2 - \mathbf{x}^{\boldsymbol{\mu}}) \;=\; \sum_{\substack{j\in[n]\\ \mu_j>0}} E_{\boldsymbol{\mu},j}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

and for each $j \in [n]$ with $\mu_j > 0$, the polynomial $E_{\boldsymbol{\mu},j}(\mathbf{x})$ has a circuit of size $\mathcal{O}(nD^2)$ and depth 2 (a $\Pi\Sigma$ circuit).

Below we recall Theorem 1.8 and proceed to prove it.

Theorem 1.8 (Upper bounds for (non-multilinear) constant-depth-IPS$_{\text{LIN}}$)**.** *Fix a prime number p. The following holds for any natural numbers n and k.*
Let $f \in \mathbb{F}_{p^k}[x_1, \ldots, x_n]$ be any polynomial with sparsity s and degree D with coefficients from the field $\mathbb{F}_{p^k}$ and let β be any element of $\mathbb{F}\backslash\mathbb{F}_{p^k}$ where $\mathbb{F}$ is a field extension of $\mathbb{F}_{p^k}$.
Then,

- *The polynomial $f(\mathbf{x}) - \beta$ has no satisfying assignment over the Boolean cube $\{0,1\}^n$*
- *There is a constant-depth-IPS$_{\text{LIN}}$ refutation of degree $O(k \cdot p \cdot D)$ and size $\text{poly}(s, p)$.*

Proof of Theorem 1.8. From Lemma 1.20, it suffices to show that there exists coefficients $A(\mathbf{x})$ and $B_j(\mathbf{x})$'s in the ring $\mathbb{F}[x_1, \ldots, x_n]$ such that

$$A(\mathbf{x}) \cdot (f(\mathbf{x}) - \beta) \;+\; \sum_{j=1}^{n} B_j(\mathbf{x}) \cdot (x_j^2 - x_j) \;=\; 1,$$

where the $A(\mathbf{x})$ and $B_j(\mathbf{x})$'s have constant-depth circuits of poly(n)-size and degree $\mathcal{O}(kpD)$.

Let $f(\mathbf{x}) = \sum_{\boldsymbol{\mu}:|\boldsymbol{\mu}|\leqslant D} \alpha_{\boldsymbol{\mu}} \mathbf{x}^{\boldsymbol{\mu}}$, where $\boldsymbol{\mu}$ denotes an exponent vector. Define the support of $f(\mathbf{x})$:

$$\text{Supp}(f) \;=\; \{\boldsymbol{\mu} \subseteq [n] \mid \alpha_{\boldsymbol{\mu}} \neq 0\}$$

The cardinality of Supp(f) is equal to the sparsity of $f(\mathbf{x})$ which is s.

Reducing to linear polynomial For every $\boldsymbol{\mu} \in \text{Supp}(f)$, define a new variable $y_{\boldsymbol{\mu}}$, i.e. s new y variables. Let $F(\mathbf{y})$ denote the polynomial when we replace the monomials in $f(\mathbf{x})$ with the new y-variables, i.e.

$$F(\mathbf{y}) \;=\; \sum_{\boldsymbol{\mu}\in\text{Supp}(f)} \alpha_{\boldsymbol{\mu}} y_{\boldsymbol{\mu}}$$

Thus $F(\mathbf{y})$ is a degree-1 polynomial in s variables.

Observe that $F(\mathbf{y}) - \beta$ does not have a solution over the Boolean hypercube $\{0,1\}^s$ since $\beta \notin \mathbb{F}_{p^k}$. From the proof of Theorem 3.1 on the degree-1 polynomial $F(\mathbf{y}) - \beta$, we get that there exists polynomials $\tilde{A}(\mathbf{y})$ and $\tilde{B}_1(\mathbf{y}), \ldots, \tilde{B}_s(\mathbf{y})$ such that the following holds:

$$\tilde{A}(\mathbf{y}) \cdot (F(\mathbf{y}) - \beta) \; + \sum_{\boldsymbol{\mu} \in \mathrm{Supp}(f)} \tilde{B}_{\boldsymbol{\mu}}(\mathbf{y}) \cdot (y_{\boldsymbol{\mu}}^2 - y_{\boldsymbol{\mu}}) \; = \; 1, \tag{9}$$

where

- The polynomial $\tilde{A}(\mathbf{y})$ is a degree-$\mathcal{O}(kp)$ polynomial and is computable by a circuit of size $\mathcal{O}(k(s+p))$ and depth 3 (a $\Sigma\Pi\Sigma$ circuit).
- For each $\boldsymbol{\mu} \in \mathrm{Supp}(f)$, the polynomial $\tilde{B}_{\boldsymbol{\mu}}(\mathbf{y})$ is a degree-$\mathcal{O}(kp)$ polynomial and is computable by a circuit of size $\mathcal{O}(ksp + k^2)$ and depth 3 (a $\Sigma\Pi\Sigma$ circuit).

Lifting the Nullstellensatz certificate Plugging in $y_S = \mathbf{x}^S$ in the Equation (9), we get,

$$\underbrace{\tilde{A}(\mathbf{y}) \circ \mathbf{x}}_{:= A(\mathbf{x})} \cdot (f(\mathbf{x}) - \beta) \; + \sum_{\boldsymbol{\mu} \in \mathrm{Supp}(f)} \underbrace{\tilde{B}_{\boldsymbol{\mu}}(\mathbf{y}) \circ \mathbf{x}}_{B'_{\boldsymbol{\mu}}(\mathbf{x})} \cdot ((\mathbf{x}^{\boldsymbol{\mu}})^2 - \mathbf{x}^{\boldsymbol{\mu}}) \; = \; 1, \tag{10}$$

where

- The polynomial $A(\mathbf{x})$ is a degree-$\mathcal{O}(kp \cdot D)$ polynomial and is computable by a circuit of size $\mathcal{O}(k(s+p) + sD)$ and depth 4 (a $\Sigma\Pi\Sigma\Pi$ circuit)
- For each $\boldsymbol{\mu} \in \mathrm{Supp}(f)$, the polynomial $B'_{\boldsymbol{\mu}}(\mathbf{x})$ is a degree-$\mathcal{O}(kp \cdot D)$ polynomial and is computable by a circuit of size $\mathcal{O}(ksp + k^2 + sD)$ and depth 4 (a $\Sigma\Pi\Sigma\Pi$ circuit).

Now applying Claim 3.4 for each subset $\boldsymbol{\mu} \in \mathrm{Supp}(f)$ in the "lifted" Nullstellensatz certificate Equation (10),

$$\begin{aligned} A(\mathbf{x}) \cdot (f(\mathbf{x}) - \beta) \; + \sum_{\boldsymbol{\mu} \in \mathrm{Supp}(f)} B'_{\boldsymbol{\mu}}(\mathbf{x}) \cdot \left(\sum_{j=1}^{n} E_{\boldsymbol{\mu},j}(\mathbf{x}) \right) \cdot (x_j^2 - x_j) \; &= \; 1 \\ \Rightarrow A(\mathbf{x}) \cdot (f(\mathbf{x}) - \beta) \; + \sum_{j=1}^{n} \underbrace{\left(\sum_{\boldsymbol{\mu} \in \mathrm{Supp}(f)} B'_{\boldsymbol{\mu}}(\mathbf{x}) \cdot E_{\boldsymbol{\mu},j}(\mathbf{x}) \right)}_{:= B_j(\mathbf{x})} \cdot (x_j^2 - x_j) \; &= \; 1, \end{aligned}$$

where for each $j \in [n]$, the polynomial $B_j(\mathbf{x})$ is a degree-$\mathcal{O}(kpD)$ polynomial and is computable by a circuit of size $\mathcal{O}(ksp + sD)$ and depth 5. This finishes the proof of Theorem 1.8. ■

3.2 Proof of Theorem 1.9

In this section, we are going to show Theorem 1.9, which we recall below.

Theorem 1.9 (Upper bound on degree of Nullstellensatz certificate)**.** *Fix a prime p. The following holds for any natural numbers n and k with $n > kp$.*
The following holds for every $\alpha_1, \ldots, \alpha_n, \beta \in \mathbb{F}_{p^k}$. Suppose the degree-1 polynomial $\sum_{i=1}^n \alpha_i x_i - \beta \in \mathbb{F}_{p^k}[x_1, \ldots, x_n]$ is unsatisfiable over the Boolean cube $\{0,1\}^n$ (i.e. there does not exist a Boolean point $\mathbf{a} \in \{0,1\}^n$ such that $\sum_{i=1}^n \alpha_i a_i - \beta = 0$).
Then, there is a constant-depth-$\mathrm{IPS}_{\mathrm{LIN}}$ refutation of degree $\mathcal{O}(k \cdot p)$ and size $\mathcal{O}(n/kp)^{\mathcal{O}(kp)}$.

In particular, if $p = \mathcal{O}(1)$ and $k = o(n)$, then there is a constant-depth-$\mathrm{IPS}_{\mathrm{LIN}}$ refutation of degree $o(n)$ and size $2^{o(n)}$.

Observe that the size bound is the "trivial" one, i.e. a n-variate multilinear polynomial with degree D has at most $\binom{n}{\leqslant D}$ monomials. Letting $D = \mathcal{O}(kp)$, we get the stated size bound in Theorem 1.9. So in our proof of Theorem 1.9, it will be enough to prove that there is a Nullstellensatz certificate of degree $\mathcal{O}(kp)$. As we will show, it will be sufficient to show that the multilinear polynomial equivalent to $1/(\sum \alpha_i x_i - \beta)$ on $\{0,1\}^n$ has degree $\mathcal{O}(kp)$. This will be our main technical lemma in the proof of Theorem 1.9, which we state and prove next.

Lemma 3.5 (Degree of the "inverse" polynomial)**.** *Fix a prime p, a parameter $k \in \mathbb{N}$ and finite field $\mathbb{F}_{p^k}$. The following holds for every $\alpha_1, \ldots, \alpha_n, \beta \in \mathbb{F}_{p^k}$ for which the equation $\sum_{i=1}^n \alpha_i x_i - \beta = 0$ is unsatisfiable over the Boolean cube $\{0,1\}^n$.*

If $f \in \mathbb{F}[x_1, \ldots, x_n]$ is a multilinear polynomial that agrees with $1/(\sum_{i=1}^n \alpha_i x_i - \beta)$ on $\{0,1\}^n$, i.e.

$$f \equiv \frac{1}{\sum_{i=1}^n \alpha_i x_i - \beta} \mod (\mathbf{x}^2 - \mathbf{x}),$$

then $\deg(f) \leqslant k \cdot (p-1)$.

Proof of Lemma 3.5. Let $L(\mathbf{x}) := \sum_{i=1}^n \alpha_i x_i - \beta$, $q = p^k$ and $(m_0, \ldots, m_{k-1})$ denote the p-ary representation of $(q-2)$ i.e.

$$q - 2 = \sum_{j=0}^{k-1} m_j p^j, \qquad \text{for all } i,\ 0 \leqslant m_i \leqslant (p-1).$$

The hypothesis says that for every $\mathbf{a} \in \{0,1\}^n$, $L(\mathbf{a}) \neq 0$. As we are working over the field $\mathbb{F}_q$, we get that for every $\mathbf{a} \in \{0,1\}^n$, $L(\mathbf{a}) \cdot (L(\mathbf{a}))^{q-2} = 1$. In other words,

$$\mathsf{ml}[(L(\mathbf{x}))^{q-2}] \equiv \frac{1}{L(\mathbf{x})} \mod (\mathbf{x}^2 - \mathbf{x}).$$

Since multilinear extension of a boolean function is unique, we get that $f = \mathsf{ml}[L^{q-2}]$, where $f \in \mathbb{F}_q[\mathbf{x}]$ is as defined in the statement of Lemma 3.5. So we will now show that $\deg(\mathsf{ml}[(L(\mathbf{x}))^{q-2}])$ is $k(p-1)$.

For every non-negative integer $j \geqslant 0$, repeated applications of Lemma 1.17 gives us:

$$(L(\mathbf{x}))^{p^j} = \sum_{i=1}^n \alpha_i^{p^j} x_i^{p^j} - \beta^{p^j} \quad \Rightarrow \quad \mathsf{ml}[(L(\mathbf{x}))^{p^j}] = \sum_{i=1}^n \alpha_i^{p^j} x_i - \beta^{p^j}.$$

For simplicity in notation, for every j, let $L_j(\mathbf{x}) := \mathsf{ml}[(L(\mathbf{x}))^{p^j}]$, and as we just showed, $\deg(L_j) = 1$. Using the p-ary expansion of $(q-2)$ and the third item of Fact 1.13, we have,

$$L(\mathbf{x})^{q-2} = \prod_{j=0}^{k-1} (L(\mathbf{x})^{p^j})^{m_j} \quad \Rightarrow \mathsf{ml}[(L(\mathbf{x}))^{q-2}] = \mathsf{ml}\left[\prod_{j=0}^{k-1} \mathsf{ml}[L_j(\mathbf{x})^{m_j}]\right]$$
$$\Rightarrow \deg(\mathsf{ml}[(L(\mathbf{x}))^{q-2}]) \leqslant \sum_{j=0}^{k-1} \deg(\mathsf{ml}[L_j(\mathbf{x})^{m_j}]) \leqslant \sum_{j=0}^{k-1} m_j \leqslant k \cdot (p-1).$$

Hence we have showed that the degree of f is $\leqslant k(p-1)$ and this finishes the proof of Lemma 3.5. ■

We now prove Theorem 1.9 using an almost straightforward application of Lemma 3.5.

Proof of Theorem 1.9. From Lemma 1.20, it suffices to show that there exists coefficients $A(\mathbf{x})$ and $B_j(\mathbf{x})$'s in the ring $\mathbb{F}_{p^k}[\mathbf{x}]$ such that

$$A(\mathbf{x}) \cdot \left(\sum_{i=1}^{n} \alpha_i x_i - \beta\right) + \sum_{j=1}^{n} B_j(\mathbf{x}) \cdot (x_j^2 - x_j) = 1,$$

where $A(\mathbf{x})$ and $B_j(\mathbf{x})$'s have constant-depth circuits of size $\mathcal{O}(n/kp)^{\mathcal{O}(kp)}$ and degree $\mathcal{O}(kp)$.

Let $L(\mathbf{x}) := \sum_{i=1}^{n} \alpha_i x_i - \beta$. Let $A \in \mathbb{F}[\mathbf{x}]$ be the multilinear polynomial such that for every $\mathbf{x} \in \{0,1\}^n$, $A(\mathbf{x})$ equals $1/L(\mathbf{x})$ (note that $L(\mathbf{x}) \neq 0$ for every $\mathbf{x} \in \{0,1\}^n$ because L is unsatisfiable over the Boolean cube). Applying Lemma 3.5, we get that $\deg(A) \leqslant k \cdot (p-1)$.
Since $A(\mathbf{x})$ is a n-variate multilinear polynomial of degree $\leqslant k(p-1)$, it has at most $\binom{n}{\leqslant D}$ monomials. Using Stirling's approximation, we get that the number of monomials is $\mathcal{O}(n/kp)^{\mathcal{O}(kp)}$, which implies a $\Sigma\Pi$ circuit for $A(\mathbf{x})$ of size $\mathcal{O}(n/kp)^{\mathcal{O}(kp)}$ and degree $\mathcal{O}(kp)$.

Now it remains to argue for $B_j(\mathbf{x})$'s. Let $B_1(\mathbf{x})$ be the quotient and $R_1(\mathbf{x})$ be the remainder when $A(\mathbf{x}) \cdot L(\mathbf{x})$ is divided by $(x_1^2 - x_1)$,

$$A(\mathbf{x}) \cdot L(\mathbf{x}) = B_1(\mathbf{x}) \cdot (x_1^2 - x_1) + R_1(\mathbf{x}).$$

Clearly $\deg(B_1), \deg(R_1) \leqslant \deg(A) + 1$. Next, let $B_2(\mathbf{x})$ denote the quotient and $R_2(\mathbf{x})$ denote the remainder when $R_1(\mathbf{x})$ is divided by $(x_2^2 - x_2)$, and so on. Since $A(\mathbf{x}) \cdot L(\mathbf{x}) - 1 \in (\mathbf{x}^2 - \mathbf{x})$, we know that $R_n(\mathbf{x}) = 1$. In other words,

$$A(\mathbf{x}) \cdot L(\mathbf{x}) = \sum_{j=1}^{n} B_j(\mathbf{x}) \cdot (x_j^2 - x_j) + 1$$
$$\Rightarrow A(\mathbf{x}) \cdot \left(\sum_{i=1}^{n} \alpha_i x_i - \beta\right) + \sum_{j=1}^{n} (-B_j(\mathbf{x})) \cdot (x_j^2 - x_j) = 1.$$

Here, for each $j \in [n]$, $\deg(B_j) \leqslant \deg(A) + 1$. Similar to $A(\mathbf{x})$, each $B_j(\mathbf{x})$ has a $\Sigma\Pi$ circuit of size $\mathcal{O}(n/kp)^{\mathcal{O}(kp)}$ and degree $\mathcal{O}(kp)$. This finishes the proof of Theorem 1.9. ■

4 Symmetric Refutations in Constant Depth

In this section, we will prove Theorem 1.11, which we recall below.

Theorem 1.11 (Upper bounds for multilinear symmetric systems)**.** *Fix a field $\mathbb{F}$. Let $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ be a family of multilinear and symmetric polynomials with no common Boolean solution i.e. there does not exist a $\mathbf{x} \in \{0,1\}^n$ such that each $f_i(\mathbf{x}) = 0$. This system has a constant-depth-$\mathrm{IPS}_{\mathrm{LIN}}$ refutation of size $\mathcal{O}(m^2 n^5 \log n)$ and depth 8.*

One of the steps in our proof of Theorem 1.11 is a *multilinearization* step, i.e. given a polynomial $f(\mathbf{x})$, we want to find a certificate in constant-depth circuits certifying that $f(\mathbf{x})$ and $\mathsf{ml}[f(\mathbf{x})]$ agree on the Boolean cube $\{0,1\}^n$. More formally, we are interested in finding polynomials $B_j(\mathbf{x})$'s such that

$$f(\mathbf{x}) \;=\; \mathsf{ml}[f(\mathbf{x})] + \sum_{j=1}^{n} B_j(\mathbf{x}) \cdot (x_j^2 - x_j),$$

and the polynomials $B_j(\mathbf{x})$ have a $\mathrm{poly}(n)$-sized constant-depth circuit.

We also need a few standard facts about elementary symmetric polynomials in fields of positive characteristic. A standard fact that is useful in our proof is that a symmetric function over the Boolean cube in constant positive characteristic only depends on $\mathcal{O}(\log n)$ elementary symmetric polynomials (instead of n elementary symmetric polynomials for symmetric polynomials over arbitrary domains). We now give a proof below for the sake of completeness.

Lemma 4.1 (Lucas's Theorem [Luc78])**.** *Fix a prime number p and any two natural numbers a and b. Denote a and b in their unique p-ary representations as:*

$$a = \sum_{i=0}^{\ell-1} a_i p^i, \qquad b = \sum_{i=0}^{\ell-1} b_i p^i, \qquad a_i, b_i \in \{0, 1, \ldots, p-1\}$$

Then,

$$\binom{a}{b} \;\equiv\; \prod_{i=0}^{\ell-1} \binom{a_i}{b_i} \mod p,$$

where we define $\binom{x}{y}$ to be 0 if $x < y$.

Next, we show that a symmetric function over the Boolean cube $\{0,1\}^n$ in characteristic p depends on $\mathcal{O}(\log n)$ elementary symmetric polynomials.

Claim 4.2 (Symmetric functions over $\{0,1\}^n$ in positive char)**.** *Fix a prime number p and a field $\mathbb{F}$ with $\mathrm{char}(\mathbb{F}) = p$. Fix a variable parameter $n \in \mathbb{N}$.*
Let $f(\mathbf{x}) \in \mathbb{F}[x_1, \ldots, x_n]$ be a multilinear and symmetric polynomial. Then $f(\mathbf{x})$ is a function of $\mathcal{O}(\log_p n)$ elementary symmetric polynomials on n variables.

Proof of Claim 4.2. Fix any natural number $0 \leqslant d \leqslant n$ and consider the d^{th} elementary symmetric polynomial, i.e., $e_d(x_1, \ldots, x_n)$. Fix an arbitrary point $\mathbf{a} \in \{0,1\}^n$ and let $k = |\mathbf{a}|$, where $|\mathbf{a}|$ denotes the Hamming weight of $\mathbf{a}$. We know that $e_d(\mathbf{a}) = \binom{k}{d}$. Denote k and d in their unique p-ary representation, i.e.

$$k = \sum_{i=0}^{\ell} k_i p^i, \quad \text{and} \quad d = \sum_{i=0}^{\ell} d_i p^i, \qquad k_i, d_i \in \{0, 1, \ldots, p-1\}$$

By Lucas's Theorem (Lemma 4.1), we have,

$$\binom{k}{d} = \binom{k_\ell}{d_\ell} \cdots \binom{k_0}{d_0} \mod p$$

Note that $e_{d_i p^i}(\mathbf{a}) = \binom{k}{d_i p^i}$. For every $0 \leqslant i \leqslant \ell$, using Lucas's Theorem (Lemma 4.1), we have,

$$\binom{k}{d_i p^i} = \binom{k_i}{d_i} \mod p \quad \Rightarrow \quad \binom{k}{d_i p^i} = \frac{1}{d_i!} \cdot \prod_{j=0}^{d-1} \left(\binom{k_i}{1} - j \right) \qquad (d_i! \neq 0 \mod p) \tag{11}$$

Using Lucas's Theorem (Lemma 4.1), we have

$$\binom{k_i}{1} = \binom{k}{p^i} \mod p \tag{12}$$

Define the polynomial $S_{d,i}(z) := \frac{1}{d_i!} \prod_{j=0}^{d_i - 1} (z - j)$. Note that $e_{p^i}(\mathbf{a}) = \binom{k}{p^i} \mod p$. Plugging in Equation (12) in Equation (11), we get

$$e_d(\mathbf{a}) = \binom{k}{d} = \prod_{i=1}^{\ell} S_{d,i}(e_{p^i}(\mathbf{a}))$$

We have shown that $e_d(\mathbf{a})$ is a polynomial of $e_{p^i}(\mathbf{a})$ for $i \in \{0, 1, \ldots, \ell\}$. Since $\mathbf{a}$ was an arbitrarily chosen point in $\{0,1\}^n$, we just argued that on the Boolean hypercube, $e_d(\mathbf{x})$ is a polynomial of $e_{p^0}(\mathbf{x}), \ldots, e_{p^\ell}(\mathbf{x})$. This holds for every $0 \leqslant d \leqslant n$. Hence, every symmetric function on $\{0,1\}^n$ in characteristic p is a polynomial of $e_{p^0}(\mathbf{x}), \ldots, e_{p^\ell}(\mathbf{x})$, i.e. of $\mathcal{O}(\log_p n)$ elementary symmetric polynomials. ■

A key lemma in our proof is the multilinearization lemma Lemma 4.3, which shows that multilinearization of a sparse polynomial in $\widehat{\mathbf{e}}(\mathbf{x})$ has a small constant-depth circuit.

Lemma 4.3 (Multilinearization of polynomial of elementary symmetric polynomials)**.** *Fix a prime number p and a field $\mathbb{F}$ with $\mathrm{char}(\mathbb{F}) = p$. Fix a variable parameter $r \in \mathbb{N}$.*
Let $F(\mathbf{y}) \in \mathbb{F}[y_1, \ldots, y_r]$ be a polynomial with individual degree strictly less than p. Then $\mathsf{ml}[F(e_1(\mathbf{y}), e_p(\mathbf{x}), \ldots, e_{p^{r-1}}(\mathbf{x}))]$ has a circuit of size $\mathcal{O}(n^5 \log n)$ and depth 5.

We will prove Lemma 4.3 later. For now, we show how it is useful in proving Theorem 1.11.

Proof of Theorem 1.11. We will first prove in the setting when the underlying field $\mathbb{F}$ has a small positive char, i.e. $\mathrm{char}(\mathbb{F}) = p$ for a constant prime p. The proof of characteristic 0 or $> n$ is similar and simpler too. We will come back to the setting $\mathrm{char}(\mathbb{F}) = 0$ or $> n$ towards the end of the proof.

From Lemma 1.20, we know that it suffices to prove there exists polynomials $A_i(\mathbf{x})$'s and $B_j(\mathbf{x})$'s in the ring $\mathbb{F}[x_1, \ldots, x_n]$ such that:

$$\sum_{i=1}^{m} A_i(\mathbf{x}) \cdot f_i(\mathbf{x}) + \sum_{j=1}^{n} B_j(\mathbf{x}) \cdot (x_j^2 - x_j) \;=\; 1,$$

where $A_i(\mathbf{x})$'s and $B_j(\mathbf{x})$'s have poly(n)-sized constant-depth circuits.

Reducing to few variables Let r denote the number of digits when n is expressed in p-ary representation. We have $r = \lfloor \log_p n \rfloor + 1 \leqslant 2 \log_p n$. Claim 4.2 tells us that there exists polynomials $F_1(\mathbf{y}), \ldots, F_m(\mathbf{y}) \in \mathbb{F}[y_1, \ldots, y_r]$ such that:

$$f_i(\mathbf{x}) \;=\; F_i(e_1(\mathbf{x}), e_p(\mathbf{x}), \ldots, e_{p^{r-1}}(\mathbf{x})) \mod \langle \mathbf{x}^2 - \mathbf{x} \rangle$$

We will denote the tuple of polynomials $(e_1(\mathbf{x}), e_p(\mathbf{x}), \ldots, e_{p^{r-1}}(\mathbf{x}))$ by $\widehat{\mathbf{e}}(\mathbf{x})$.
Since $f_i(\mathbf{x})$ and $F_i(\widehat{\mathbf{e}}(\mathbf{x}))$ agree on the Boolean cube $\{0,1\}^n$, their multilinear components are equal, i.e. $\mathsf{ml}[f_i(\mathbf{x})] = \mathsf{ml}[F_i(\widehat{\mathbf{e}}(\mathbf{x}))]$ (see Fact 1.13). Since $f_i(\mathbf{x})$ is a **multilinear** polynomial, we have $f_i(\mathbf{x}) = \mathsf{ml}[F_i(\widehat{\mathbf{e}}(\mathbf{x}))]$.

For every $1 \leqslant i \leqslant r$, the polynomial $e_{p^{i-1}}(\mathbf{x})$ take values in $\mathbb{F}_p$ over the Boolean cube. For every $j \in [r]$, let $p_j(t)$ be a univariate polynomial that vanishes on the set $\mathbb{F}_p$, i.e. $p_j(t) = \prod_{\alpha \in \mathbb{F}_p} (t - \alpha)$. For every $n < t < p^r - 1$, define the polynomial $Q_t \in \mathbb{F}[\mathbf{y}]$ as follows:

$$Q_t(\mathbf{y}) := \prod_{i=0}^{r-1} \frac{1}{t_i!} \prod_{j=0}^{t_i - 1} (y_i - j)\,, \qquad \text{where } t = \sum_{i=0}^{r-1} t_i p^{i-1}$$

Our first claim shows that if $f_i(\mathbf{x})$'s do not have a common Boolean solution, then $F_i(\mathbf{y})$'s along with some additional constraints do not have a common solution, even over the algebraic closure of $\mathbb{F}$.

Claim 4.4. *The system consisting of $F_i(\mathbf{y})$'s, $Q_t(\mathbf{y})$'s, and $p_j(y_j)$'s have no common solution in the closure $\overline{\mathbb{F}}^n$, i.e.*

$$\mathbb{V}(F_1(\mathbf{y}), \ldots, F_m(\mathbf{y}), Q_{n+1}(\mathbf{y}), \ldots, Q_{p^r-1}(\mathbf{y}), p_1(y_1), \ldots, p_r(y_r)) \;=\; \varnothing$$

Proof. We will prove this by contradiction. Assume for the sake of contradiction that there exists a common solution $\mathbf{b}$ to the above system of polynomials. Since for every $j \in [r]$, $p_j(b_j) = 0$, this implies that $\mathbf{b} \in \{0, \ldots, p-1\}^r$. We will now show the existence of a point $\mathbf{a} \in \{0,1\}^n$ such that

$$\widehat{\mathbf{e}}(\mathbf{a}) = \mathbf{b} \tag{13}$$

Observe that an $\mathbf{a} \in \{0,1\}^n$ which satisfies Equation (13) is a common Boolean solution to the system $\{f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\}$, which is a contradiction to our hypothesis of Theorem 1.11. So to finish the contradiction, all that remains is to show the existence of such a Boolean point $\mathbf{a}$.

We are only interested in showing the existence of a Boolean point satisfying Equation (13), i.e. we are only interested in the evaluation of $\mathbf{a}$ on symmetric polynomials. Thus we can focus on showing the existence of an appropriate Hamming weight $0 \leqslant k \leqslant n$ such that

$$\binom{k}{p^i} = \binom{k_i}{1} = b_i \mod p, \qquad \text{for all } 0 \leqslant i \leqslant r-1$$

where we used Lucas's Theorem (Lemma 4.1) for the first equality. Choose k to be $\sum_{i=0}^{\ell} b_i p^i$. If $k \leqslant n$, then k satisfies all the r constraints in Equation (13). In particular, we can set $\mathbf{a} = 1^k 0^{n-k}$ and it will satisfy Equation (13). Thus to complete the proof, we need to show that $k \leqslant n$.

By assumption, $Q_t(\mathbf{b}) = 0$ for all $n < t \leqslant p^r - 1$. Fix any $t = \sum_{i=0}^{r-1} t_i p^{i-1}$. By the definition of the polynomial $Q_t(\mathbf{y})$ (see the proof of Claim 4.2):

$$Q_t(\mathbf{b}) \;=\; \prod_{i=0}^{r-1} \binom{b_i}{t_i} \;=\; 0.$$

This means there exists an $i \in \{0, 1, \ldots, r-1\}$ such that $b_i < t_i$. Since this holds for every $n < t \leqslant p^r - 1$, this implies that for each coordinate i, $b_i \leqslant n_i$ where $n = \sum_{i=0}^{r-1} n_i p^i$. Thus $k \leqslant n$.

Hence we have found a Boolean point $\mathbf{a} \in \{0,1\}^n$ which satisfies Equation (13). Since for every $i \in [m]$, $f_i(\mathbf{x})$ and $F_i(\widehat{\mathbf{e}}(\mathbf{x}))$ agree on the Boolean cube $\{0,1\}^n$, $\mathbf{a}$ is a common Boolean solution to $f_i(\mathbf{x})$'s. This is a contradiction to our assumption in Theorem 1.11. ■

Low-variate Nullstellensatz We have shown that the unsatisfiability of the n-variate polynomials f_i's over the Boolean cube implies the unsatisfiability of $\mathcal{O}(\log n)$-variate polynomials F_i's (with some additional polynomials to reflect the Boolean cube restriction). Now we use Hilbert's Nullstellensatz to get a Nullstellensatz certificate for the $\mathcal{O}(\log n)$-variate polynomials and "lift" it to get a Nullstellensatz certificate for the original system of polynomial equations.

Claim 4.4 says that the system consisting of F_i's, Q_t's, and p_j's do not have a common zero over the algebraic closure $\bar{\mathbb{F}}$. Applying Hilbert's Nullstellensatz (Theorem 1.19) on this system, we know that there exist polynomials $\tilde{A}_i(\mathbf{y})$'s, $\tilde{S}_t(\mathbf{y})$'s, and $\tilde{B}_j(\mathbf{y})$'s such that:

$$\sum_{i=1}^{m} \tilde{A}_i(\mathbf{y}) \cdot F_i(\mathbf{y}) + \sum_{t=n+1}^{p^r-1} \tilde{S}_t(\mathbf{y}) \cdot Q_t(\mathbf{y}) \;+\; \sum_{j=1}^{r} \tilde{B}_j(\mathbf{y}) \cdot p_j(y_j) \;=\; 1 \tag{14}$$

Size analysis of low-variate certificate Next, we will show that the coefficients $\tilde{A}_i$'s, $\tilde{S}_t$'s, and $\tilde{B}_j$'s in the Nullstellensatz certificate (Equation (14)) have small constant-depth circuits. More precisely, we will show that $\tilde{A}_i$'s, $\tilde{S}_t$'s, and $\tilde{B}_j$'s are polynomials with sparsity $\mathrm{poly}(n)$, which in turn implies that they have $\mathrm{poly}(n)$-sized depth 2 circuits. Since these polynomials are $\mathcal{O}(\log n)$-variate polynomials, it will suffice to show that they have constant individual degrees. We will argue as follows:

- **Step (a)** The polynomials F_i's and Q_t's have constant individual degree.

- **Step (b)** Use the low-variate Nullstellensatz certificate Equation (14) to argue that the polynomials $\tilde{A}_i$'s and $\tilde{S}_t$'s can be assumed to have constant individual degree.
- **Step (c)** Use the previous two items to argue that the polynomials $\tilde{B}_j$'s have sparsity at most $\text{poly}(n)$.

Step (a) We first argue that the individual degree of each $F_i(\mathbf{y})$ can be assumed to be $\leqslant (p-1)$. Recall that $f_i(\mathbf{x}) = F_i(\hat{\mathbf{e}}(\mathbf{x}))$ over $\{0,1\}^n$, i.e.: $f_i(\mathbf{x}) \equiv F_i(\hat{\mathbf{e}}(\mathbf{x})) \bmod (\mathbf{x}^2 - \mathbf{x})$.
As we are working over a field of characteristic p, for any $i \in [n]$, $e_i(\mathbf{x}) \in \{0, 1, \ldots, p-1\}$ for every $\mathbf{x} \in \{0,1\}^n$. This implies that $\hat{\mathbf{e}}(\mathbf{x})^p \equiv \hat{\mathbf{e}}(\mathbf{x}) \bmod (\mathbf{x}^2 - \mathbf{x})$ via Fermat's Little Theorem/Frobenius automorphism.
Let $F_i'(\mathbf{y}) := F_i(\mathbf{y})/(\mathbf{y}^p - \mathbf{y})$ be a "minimum individual-degree representative" of $F_i(\mathbf{y})$ modulo the ideal $(\mathbf{y}^p - \mathbf{y})$.

Thus, $F_i(\mathbf{y})$ is a polynomial of individual-degree $\leqslant p-1$ such that $F'(\mathbf{y}) \equiv F(\mathbf{y}) \bmod (\mathbf{y}^p - \mathbf{y})$. Combining these together, we get,

$$f_i(\mathbf{x}) \equiv F_i'(\hat{\mathbf{e}}(\mathbf{x})) \bmod (\mathbf{x}^2 - \mathbf{x}) \tag{15}$$

where F_i' has individual-degree $\leqslant p-1$. With a slight abuse of notation, we will now use "F_i" to denote F_i'.
By the definition of the polynomial $Q_t(\mathbf{y})$, for each t, the individual degree of $Q_t(\mathbf{y})$ is $\leqslant p$

Step (b) The polynomial $p_j(t)$, defined as $\prod_{\alpha \in \mathbb{F}_p}(t - \alpha)$, is equal to $(t^p - t)$ by Fermat's Little Theorem. From the Nullstellensatz certificate Equation (14),

$$\sum_{i=1}^{m} \tilde{A}_i(\mathbf{y}) \cdot F_i(\mathbf{y}) + \sum_{t=n+1}^{p^r-1} \tilde{S}_t(\mathbf{y}) \cdot Q_t(\mathbf{y}) = 1 \mod (\mathbf{y}^p - \mathbf{y})$$

We would now argue that the polynomials $\tilde{A}_i(\mathbf{y})$'s and $\tilde{S}_t(\mathbf{y})$'s has individual degree $\leqslant (p-1)$. Suppose there exists an $i \in [m]$ for which $\tilde{A}_i(\mathbf{y})$ has individual degree $> (p-1)$, then define $\tilde{A}_i'(\mathbf{y}) := A_i(\mathbf{y})/(\mathbf{y}^p - \mathbf{y})$ to be a "minimum individual-degree representative". Observe that replacing the polynomial $\tilde{A}_i$ with the polynomial $\tilde{A}_i'$, the low-variate Nullstellensatz certificate Equation (14) continues to holds.

Thus the sparsity of $\tilde{A}_i(\mathbf{y})$ is at most p^r. An analogous argument shows that for each $n < t \leqslant p^r - 1$, the polynomial $\tilde{S}_t(\mathbf{y})$ has individual degree $\leqslant (p-1)$ and thus has a $\mathcal{O}(n^2)$-sized circuit of depth 2 (a $\Sigma\Pi$ circuit).

Using $r \leqslant 2\log_p n$, we have that for every $i \in [m]$ and for every $n < t \leqslant p^r - 1$, the polynomials $\tilde{A}_i(\mathbf{y})$ and $\tilde{S}_t(\mathbf{y})$ have sparsity at most $\mathcal{O}(n^2)$. This also implies that the polynomials $\tilde{A}_i(\mathbf{y})$'s and $\tilde{S}_t(\mathbf{y})$'s have $\mathcal{O}(n^2)$-sized circuits of depth 2 (a $\Sigma\Pi$ circuit).

Step (c) Now it remains to show that the polynomials $\tilde{B}_j(\mathbf{y})$ have small constant-depth circuits. We will show that for each $j \in [n]$, the polynomial $\tilde{B}_j(\mathbf{y})$ has sparsity at most $\text{poly}(n)$. To show this, we will use the fact that $\sum_{i=1}^{m} \tilde{A}_i(\mathbf{y}) \cdot F_i(\mathbf{y}) + \sum_{t=n+1}^{p^r-1} \tilde{S}_t(\mathbf{y}) \cdot Q_t(\mathbf{y})$ is a polynomial of constant individual degree.

For a polynomial $H(\mathbf{y})$, the *individual-degree-p* operator, denoted by inddeg_p outputs the following polynomial: For each variable y_j, every occurrence of y_j^p in $H(\mathbf{y})$ is replaced by y_j until the individual degree of the polynomial is $< p$. For $p = 2$, inddeg_2 corresponds to multilinearization ml.
Our next claim shows that if there is a polynomial of low individual degree, then its individual-degree-p component can be extracted using polynomials of small constant-depth circuits. The proof is via a simple induction. We omit the proof here and it can be found in Appendix A.3.

Claim 4.5. *Let $H(\mathbf{y}) \in \mathbb{F}[y_1, \ldots, y_r]$ denote a polynomial whose individual degree is at most D. Then there exists polynomials $G_1(\mathbf{y}), \ldots, G_r(\mathbf{y})$ such that the following holds:*

$$H(\mathbf{y}) \;=\; \mathsf{inddeg}_p[H(\mathbf{y})] \;+\; \sum_{j=1}^{r} G_j(\mathbf{y}) \cdot (y_j^p - y_j),$$

and for each $j \in [r]$, the polynomial $G_j(\mathbf{y})$ has sparsity at most $D^{r+1}/(p-1)$.

Define the polynomial $H(\mathbf{y}) = \sum_{i=1}^{m} \tilde{A}_i(\mathbf{y}) \cdot F_i(\mathbf{y}) + \sum_{t=n+1}^{p^r-1} \tilde{S}_t(\mathbf{y}) \cdot Q_t(\mathbf{y})$. Observe that

$$H(\mathbf{y}) = 1 + \sum_{j=1}^{r} \tilde{B}_j(\mathbf{y}) \cdot (y_j^p - y_j)$$

So the polynomials $G_j(\mathbf{y})$ from Claim 4.5 correspond to the polynomials $\tilde{B}_j(\mathbf{y})$ from the Nullstellensatz certificate. Since both $\tilde{A}_i(\mathbf{y})$ and $F_i(\mathbf{y})$ have individual degree at most $(p-1)$, $H(\mathbf{y})$ has individual degree at most $2(p-1)$ in each of the r variables. Applying Claim 4.5 on $H(\mathbf{y})$, we get that the polynomials $\tilde{B}_j(\mathbf{y})$ has sparsity at most $(2(p-1))^r \cdot 2(p-1)/(p-1) = \mathrm{poly(n)}$ since $p \leqslant 2\log_p n$. This implies that the polynomials $\tilde{B}_j(\mathbf{y})$ has a circuit of size $\mathcal{O}(n^2)$ and depth 2 (a $\Sigma\Pi$ circuit).

Lifting the Nullstellensatz certificate By the definition of $F_i(\mathbf{y})$, we know that for every $i \in [m]$, $f_i(\mathbf{x}) = \mathsf{ml}[F_i(\hat{\mathbf{e}}(\mathbf{x}))]$. Applying the multilinearization Lemma 4.3 on $F_i(\hat{\mathbf{e}}(\mathbf{x}))$ for every $i \in [m]$, we know there exists polynomials $D_{ij}(\mathbf{x})$ for $i \in [m]$ and $j \in [n]$ such that:

$$F_i(\hat{\mathbf{e}}(\mathbf{x})) \;=\; f_i(\mathbf{x}) \;+\; \sum_{j=1}^{n} D_{ij}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

where each polynomial $D_{ij}(\mathbf{x})$ has a circuit of size $\mathcal{O}(n^5 \log n)$ and depth 5. This also implies that for each $i \in [m]$, the polynomial $f_i(\mathbf{x})$ has a circuit of size $\mathcal{O}(n^5 \log n)$ and depth 5.
Similarly, applying the multilinearization lemma Lemma 4.3 on $Q_t(\hat{\mathbf{e}}(\mathbf{x}))$ for every $n < t \leqslant p^r - 1$, we know there exists polynomials $R_t(tj)(\mathbf{x})$ for $n < t \leqslant p^r - 1$ and $j \in [n]$ such that:

$$Q_t(\hat{\mathbf{e}}(\mathbf{x})) \;=\; \mathsf{ml}[Q_t(\hat{\mathbf{e}}(\mathbf{x}))] \;+\; \sum_{j=1}^{n} R_{tj}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

where each polynomial $R_{tj}(\mathbf{x})$ has a circuit of size $\mathcal{O}(n^5 \log n)$ and depth 5.

Next, we multilinearize the coefficients $\tilde{A}_i(\hat{\mathbf{e}}(\mathbf{x}))$'s to get a multilinear constant-depth-IPS$_{\text{LIN}}$ proof. Applying the multilinearization Lemma 4.3 on $\tilde{A}_i(\hat{\mathbf{e}}(\mathbf{x}))$ for every $i \in [m]$, we know there exists polynomials $\tilde{D}_{ij}(\mathbf{x})$ for $i \in [m]$ and $j \in [n]$ such that:

$$\tilde{A}_i(\hat{\mathbf{e}}(\mathbf{x})) \;=\; \mathsf{ml}[\tilde{A}_i(\hat{\mathbf{e}}(\mathbf{x}))] + \sum_{j=1}^{n} \tilde{D}_{ij}(\mathbf{x}),$$

where each polynomial $\tilde{D}_{ij}(\mathbf{x})$ has a circuit of size $\mathcal{O}(n^5 \log n)$ and depth 5. This also implies that for each $i \in [m]$, the polynomial $\mathsf{ml}[\tilde{A}_i(\hat{\mathbf{e}}(\mathbf{x}))]$ has a circuit of size $\mathcal{O}(n^6)$ and depth 5.
Similarly, we also multilinearize the coefficients $\tilde{S}_t(\hat{\mathbf{e}}(\mathbf{x}))$. Applying the multilinearization lemma Lemma 4.3 on $\tilde{S}_t(\hat{\mathbf{e}}(\mathbf{x}))$ for every $n < t \leqslant p^r - 1$, we know there exists polynomials $\tilde{R}_{tj}(\mathbf{x})$ for $n < t \leqslant p^r - 1$ and $j \in [n]$ such that:

$$\tilde{S}_t(\hat{\mathbf{e}}(\mathbf{x})) \;=\; \mathsf{ml}[\tilde{S}_t(\hat{\mathbf{e}}(\mathbf{x}))] \;+\; \sum_{j=1}^{n} \tilde{R}_{tj}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

where each polynomial $\tilde{R}_{tj}(\mathbf{x})$ has a circuit of size $\mathcal{O}(n^5 \log n)$ and depth 5.
For every $i \in [r]$, applying the multilinearization lemma Lemma 4.3 on $p_i(e_{p^{i-1}}(\mathbf{x}))$, we know that there exists polynomials $E_{i1}(\mathbf{x}), \ldots, E_{in}(\mathbf{x})$ such that:

$$p_i(e_{p^{i-1}}(\mathbf{x})) \;=\; \sum_{j=1}^{n} E_{ij}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

where for each $j \in [n]$, the polynomial $E_{ij}(\mathbf{x})$ has a circuit of size $\mathcal{O}(n^5 \log n)$ and depth 5.

Substituting $\mathbf{y} = \hat{\mathbf{e}}(\mathbf{x})$ in the low-variate Nullstellensatz certificate Equation (14) and using the above polynomial relations, we get,

$$\sum_{i=1}^{m} \underbrace{\tilde{A}_i(\hat{\mathbf{e}}(\mathbf{x}))}_{:=A_i(\mathbf{x})} \cdot f_i(\mathbf{x}) \;+\; \sum_{j=1}^{n} B_j(\mathbf{x}) \cdot (x_j^2 - x_j) \;=\; 1,$$

where the polynomial $B_j(\mathbf{x})$ is:

$$\begin{aligned} B_j(\mathbf{x}) \;&=\; \sum_{i=1}^{m} (\mathsf{ml}[\tilde{A}_i(\hat{\mathbf{e}}(\mathbf{x}))] D_{ij}(\mathbf{x}) + f_i(\mathbf{x}) \tilde{D}_{ij}(\mathbf{x})) \\ &+ \sum_{t=n+1}^{p^r-1} (\mathsf{ml}[Q_t(\hat{\mathbf{e}}(\mathbf{x}))] \cdot \tilde{S}_t(\hat{\mathbf{e}}(\mathbf{x})) + Q_t(\hat{\mathbf{e}})(\mathbf{x}) \cdot \mathsf{ml}[\tilde{S}_t(\hat{\mathbf{e}}(\mathbf{x}))]) \\ &+ \sum_{i=1}^{m} \sum_{j' \leqslant j} \tilde{D}_{ij'}(\mathbf{x}) D_{ij}(\mathbf{x}) + \sum_{i=1}^{r} \tilde{B}_i(\hat{\mathbf{e}}(\mathbf{x})) E_{ij}(\mathbf{x}) \end{aligned}$$

We have,

- For each $i \in [m]$, using Ben-Or's construction Theorem 1.15, the polynomial $A_i(\mathbf{x})$ has a circuit of size $\mathcal{O}(n^2)$ and depth 5.

- For each $j \in [r]$, and again using Ben-Or's construction Theorem 1.15, the polynomial $B_j(\mathbf{x})$ has a circuit of size $\mathcal{O}(mn^5 \log n)$ and depth 7.

This finishes the proof of Theorem 1.11 in the setting when the underlying field has a positive characteristic p for a constant prime p.

We now discuss the proof of Theorem 1.11 in the setting when the underlying field has characteristic 0 or $> n$. The proof has the exact same steps as for positive characteristic. Instead of repeating the same steps again, for the sake of brevity, we only highlight the differences.

Over characteristic 0 or $> n$, every multilinear symmetric polynomial is a polynomial of $x_1 + \ldots + x_n$, i.e. of $e_1(\mathbf{x})$. Thus r in the above proof is just 1 and F_i's are univariate polynomials. Let $p(y) = \prod_{i=0}^{n}(y - i)$. Then it is easy to see that if $f_i(\mathbf{x})$ do not have a common Boolean solution, then the univariate polynomials $F_i(y)$'s and $p(y)$ do not have a common solution (following a similar strategy to the proof of Claim 4.4, if b is a common solution, then there exists a common Boolean solution of Hamming weight b, which is a contradiction).

To argue about the circuit size of the coefficients of the univariate Nullstellensatz certificate, it suffices to argue about their degrees since they are all univariate polynomials. The coefficients of the univariate certificate have degree at most $\mathcal{O}(n)$ because of the polynomial $p(y)$ (it is quite similar and simpler to the degree analysis of the coefficients of the low-variate Nullstellensatz certificate in the above proof).
In the end, we need to multilinearize $F_i(e_1(\mathbf{x}))$. This can again be done in a constant-depth circuit using Lemma 4.3. ■

4.1 Multilinearization

To show our multilinearization lemma (Lemma 4.3), it will be convenient to first define a notion of *partial multilinearization*, i.e., multilinearize with respect to a subset of variables. A key lemma used in our proofs of multilinearization statements is constant-depth multilinearization when $f(\mathbf{x})$ is a *product of univariate polynomials* (see Corollary 4.10). We now define the partial multilinearization and then use it to prove Corollary 4.10.

Definition 4.6 (Partial multilinearization)**.** *Fix any field $\mathbb{F}$ and let $f(\mathbf{x}) \in \mathbb{F}[x_1, \ldots, x_n]$. For any $j \in [n]$, let $f^{(\leqslant j)}(\mathbf{x}) \in \mathbb{F}[x_{j+1}, \ldots, x_n][x_1, \ldots, x_j]$ denote the polynomial $f(\mathbf{x})$ with variables $x_1, \ldots, x_j$ and coefficients in $\mathbb{F}[x_{j+1}, \ldots, x_n]$.*
The multilinearization of the polynomial $f(\mathbf{x})$ with respect to the variables $\{x_1, \ldots, x_j\}$, denoted by $\mathsf{ml}_{\leqslant j}[f(\mathbf{x})]$, is defined to be:

$$\mathsf{ml}_{\leqslant j}[f(\mathbf{x})] \;:=\; \mathsf{ml}[f^{(\leqslant j)}(\mathbf{x})]$$

Similarly, for any $k \in [n]$, let $f^{(k)}(\mathbf{x}) \in \mathbb{F}[x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n][x_k]$ denote the polynomial $f(\mathbf{x})$ with variable x_k only and coefficients in $\mathbb{F}[x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n]$. The multilinearization of the polynomial $f(\mathbf{x})$ with respect to the variable x_k only, denoted by $\mathsf{ml}_k[f(\mathbf{x})]$, is defined to be:

$$\mathsf{ml}_k[f(\mathbf{x})] \;=\; \mathsf{ml}[f^{(k)}(\mathbf{x})]$$

Sometimes we will denote $\mathsf{ml}_k[f(\mathbf{x})]$ by $\mathsf{ml}_{x_k}[f(\mathbf{x})]$ for sake of clarity.

Example: Let $f(\mathbf{x}) = x_1^2x_2^3 + x_2x_3^2$. Then,

$$\mathsf{ml}_{\leqslant 1}[f(\mathbf{x})] \;=\; x_1x_2^3 + x_2x_3^2, \quad \mathsf{ml}_{\leqslant 2}[f(\mathbf{x})] \;=\; x_1x_2 + x_2x_3^2, \quad \mathsf{ml}_2[f(\mathbf{x})] \;=\; x_1^2x_2 + x_2x_3^2$$

We make one observation on partial multilinearization, which will be helpful in the proofs.

Observation 4.7. *For every $j < n$, the following holds: For every polynomial $f(\mathbf{x})$,*

$$\mathsf{ml}_{\leqslant j+1}[f(\mathbf{x})] \;=\; \mathsf{ml}_{j+1}[\mathsf{ml}_{\leqslant j}[f(\mathbf{x})]]$$

In the rest of the section, we will use the notation $\mathbf{x}_{\leqslant j}$ to denote $(x_1, \ldots, x_j)$ and $\mathbf{x}_{>j}$ to denote $(x_{j+1}, \ldots, x_n)$.

Now we show that a product of univariate polynomials can be multilinearized using constant-depth poly(n)-sized circuits (see Corollary 4.10). We start by showing that we can do partial multilinearization with respect to a single variable.

Claim 4.8 (Multilinearize a single variable)**.** *Consider a univariate polynomial $h(z)$ of degree-D. Let $Q(\mathbf{y})$ be a polynomial with a circuit of size s and depth Δ. Let $\mathsf{ml}_z[h(z) \cdot Q(\mathbf{y})]$ denotes the partial multilinearization of the polynomial $h(z) \cdot Q(\mathbf{y})$ with respect to the z variable.*
Then,

$$h(z) \cdot Q(\mathbf{y}) \;=\; \mathsf{ml}_z[h(z) \cdot Q(\mathbf{y})] + B(z, \mathbf{y}) \cdot (z^2 - z),$$

- *The polynomial $\mathsf{ml}_z[h(z) \cdot Q(\mathbf{y})]$ is equal to $L(z) \cdot Q(\mathbf{y})$, where $L(z)$ is a degree-1 univariate polynomial in z.*
- *The polynomial $B(z, \mathbf{y})$ is equal to $\tilde{h}(z) \cdot Q(\mathbf{y})$ for a univariate polynomial $\tilde{h}(z)$.*

Proof of Claim 4.8. Let $h(z) = a_0 + a_1z + \cdots + a_Dz^D$. Then for every $2 \leqslant j \leqslant D$, rewriting a_jz^j as $a_j(z^j - z) + a_jz$, we get,

$$\begin{aligned} h(z) \cdot Q(\mathbf{y}) \;&=\; \left(a_0 + a_1z + \sum_{j=2}^{D} a_j(z^j - z + z)\right) \cdot Q(\mathbf{y}) \\ &= \underbrace{\left(a_0 + \left(\sum_{j=1}^{D} a_j\right) z\right)}_{:= L(z)} \cdot Q(\mathbf{y}) + \left(\sum_{j=2}^{D} a_j(z^j - z)\right) \cdot Q(\mathbf{y}) \end{aligned}$$

Observe that for any $j \geqslant 2$,

$$z^j - z \;=\; \left(z^{j-2} + \cdots + z + 1\right) \cdot (z^2 - z)$$

Using the above observation, we get,

$$h(z) \cdot Q(\mathbf{y}) \;=\; \underbrace{L(z) \cdot Q(\mathbf{y})}_{= \mathsf{ml}_z[h(z) \cdot Q(\mathbf{y})]} \;+\; \underbrace{Q(\mathbf{y}) \sum_{j=2}^{D} a_j \left(z^{j-1} + \cdots + z + 1\right) \cdot (z^2 - z)}_{= B(z, \mathbf{y})}$$

Let $\tilde{h}(z) = \sum_{j=2}^{D} a_j(z^{j-1} + \cdots + z + 1)$. Then the polynomial $B(z, \mathbf{y})$ is equal to $\tilde{h}(z) \cdot Q(\mathbf{y})$. The partial multilinearization $\mathsf{ml}_z[h(z) \cdot Q(\mathbf{y})]$ is of the form $L(z) \cdot Q(\mathbf{y})$ for a degree-1 polynomial $L(z)$. This finishes the proof of Claim 4.8. ■

The next claim shows that if a product of univariate polynomials, then we can do partial multilinearization with respect to a subset of variables. It follows with a simple induction using Claim 4.8. We omit the proof here and it can be found in Appendix A.4.

Claim 4.9 (Partial multilinearization of product of univariates)**.** *Let $h_1(z_1), \ldots, h_n(z_n)$ be univariate polynomials where each $h_i(z_i)$ has degree at most D.*
Then there exists degree-1 univariate polynomials $L_1(z_1), \ldots, L_n(z_n)$ and polynomials $B_1(\mathbf{z}), \ldots, B_n(\mathbf{z})$ satisfying the following: For every $k \in [n]$,

$$h_1(z_1) \cdots h_n(z_n) \;=\; \mathsf{ml}_{\leqslant k}\left[\prod_{i=1}^{n} h_i(z_i)\right] \;+\; \sum_{j=1}^{k} B_j(\mathbf{z}) \cdot (z_j^2 - z_j),$$

where

$$\mathsf{ml}_{\leqslant k}\left[\prod_{i=1}^{n} h_i(z_i)\right] \;=\; \prod_{i=1}^{k} L_i(z_i) \cdot \prod_{i=k+1}^{n} h_i(z_i),$$

and for each $j \in [n]$, the polynomial $B_j(\mathbf{z})$ has the following form:

$$B_j(\mathbf{x}) \;=\; \prod_{i=1}^{j-1} L_i(z_i) \cdot \tilde{h}_j(z_j) \cdot \prod_{i=j+1}^{n} h_i(z_i),$$

for some univariate polynomial $\tilde{h}_j(z_j)$.

Setting $k = n$ in Claim 4.9 immediately gives us the following corollary.

Corollary 4.10 (Multilinearization of product of univariates)**.** *Let $h_1(z_1), \ldots, h_n(z_n)$ be univariate polynomials where each $h_i(z_i)$ has degree at most D.*
Then there polynomials $B_1(\mathbf{z}), \ldots, B_n(\mathbf{z})$ such that,

$$h_1(z_1) \cdots h_n(z_n) \;=\; \mathsf{ml}\left[\prod_{i=1}^{n} h_i(z_i)\right] \;+\; \sum_{j=1}^{k} B_j(\mathbf{z}) \cdot (z_j^2 - z_j),$$

where for each $j \in [n]$, the polynomial $B_j(\mathbf{z})$ has a circuit of size $\mathcal{O}(nD^2)$ and depth 3 (a $\Pi\Sigma\Pi$) circuit.

In this section, we will prove the multilinearization lemma Lemma 4.3. The key step in our proof of Lemma 4.3 is Lemma 4.11 which is a special case of Lemma 4.3. In particular, Lemma 4.11 shows that the multilinearization of a product of two elementary symmetric polynomials has a small constant-depth circuit. Furthermore, it shows that the multlinearization of a product of two elementary symmetric polynomials has a nice structure which we use to prove Lemma 4.3.

Lemma 4.11 (Multilinearization of product of two elementary symmetric polynomials)**.** *Fix any two natural numbers α and β. Then*

- *There exists polynomials $R_{\alpha,\beta,j}(\mathbf{x})$'s such that*

$$\mathsf{ml}[e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x})] \;=\; e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x}) - \sum_{j=1}^{n} R_{\alpha,\beta,j}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

  where each polynomial $R_{\alpha,\beta,j}(\mathbf{x})$ has a circuit of size $\mathcal{O}(n^3)$ and depth 5 (a $\Sigma\Pi\Sigma\Pi\Sigma$ circuit).

- *There exists coefficients $c^{(i)}_{\alpha,\beta}$'s such that*

$$\mathsf{ml}[e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x})] \;=\; \sum_{i=1}^{n} c^{(i)}_{\alpha,\beta}\, e_i(\mathbf{x})$$

Proof of Lemma 4.11. Using Ben-Or's construction (Theorem 1.15) for $e_\alpha(\mathbf{x})$ and $e_\beta(\mathbf{x})$,

$$e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x}) \;=\; \sum_{i_1,i_2} c_{\alpha,i_1} c_{\beta,i_2} \prod_{j=1}^{n} (1+\gamma_{i_1} x_j)(1+\gamma_{i_2} x_j), \qquad \text{where all } \; c_{\alpha_{i_1}}, c_{\beta,i_2}, \gamma_{i_1}, \gamma_{i_2} \in \mathbb{F}$$

$$\Rightarrow e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x}) \;=\; \sum_{i=1}^{(n+1)^2} \prod_{j=1}^{n} h^{\alpha,\beta}_{i,j}(x_j), \tag{16}$$

where each polynomial $h^{\alpha,\beta}_{i,j}(x_j)$ is a degree-2 *univariate* polynomial. Fix any $i \in [(n+1)^2]$ and using Corollary 4.10 on $h^{\alpha,\beta}_{i,1}(x_1) \cdots h^{\alpha,\beta}_{i,n}(x_n)$, we know that there exists polynomials $B^{\alpha,\beta}_{i,j}(\mathbf{x})$'s such that:

$$h^{\alpha,\beta}_{i,1}(x_1) \cdots h^{\alpha,\beta}_{i,n}(x_n) \;=\; \mathsf{ml}[h^{\alpha,\beta}_{i,1}(x_1) \cdots h^{\alpha,\beta}_{i,n}(x_n)] + \sum_{j=1}^{n} B^{\alpha,\beta}_{i,j}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

where each polynomial $B_{i,j}(\mathbf{x})$ has a circuit of size $\mathcal{O}(n)$ and depth 3 (a $\Pi\Sigma\Pi$ circuit). Now summing it over all $i \in [(n+1)^2]$ (see Equation (16)), we get,

$$e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x}) \;=\; \mathsf{ml}[e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x})] + \sum_{j=1}^{n} R_{\alpha,\beta,j}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

where each polynomial $R_{\alpha,\beta,j}(\mathbf{x})$ has a circuit of size $\mathcal{O}(n^3)$ and depth 4 (a $\Sigma\Pi\Sigma\Pi$ circuit). This shows the *constant-depth circuit* item of Lemma 4.11.

Next we argue about the *structure* item of Lemma 4.11. Since $e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x})$ is a symmetric polynomial, its multilinearization $\mathsf{ml}[e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x})]$ is also a symmetric polynomial. The Fundamental Theorem of Symmetric Polynomials (see Theorem 1.14) implies that $\mathsf{ml}[e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x})]$ is a polynomial of $e_k(\mathbf{x})$'s. Note that any multilinear symmetric polynomial is a linear combination of $e_k(\mathbf{x})$'s. Thus $\mathsf{ml}[e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x})]$ is a linear combination of $e_k(\mathbf{x})$'s. This finishes the *structure* item of Lemma 4.11. This finishes the proof of Lemma 4.11. ■

Now we are ready to prove Lemma 4.3. The idea for the proof is as follows:

- We use the fact that $F(\mathbf{y})$ has at most poly(n) sparsity. So for each monomial $\mathbf{y}^{\boldsymbol{\mu}}$, we multilinearize $\mathbf{y}^{\boldsymbol{\mu}} \circ \widehat{\mathbf{e}}(\mathbf{x})$ individually.
- For any fixed monomial $\mathbf{y}^{\boldsymbol{\mu}} \circ \widehat{\mathbf{e}}(\mathbf{x})$, we note that it is a product of elementary symmetric polynomials. Lemma 4.11 shows how to multilinearize a product of two elementary symmetric polynomials. We repeatedly apply this on $\mathbf{y}^{\boldsymbol{\mu}} \circ \widehat{\mathbf{e}}(\mathbf{x})$.

We recall the statement of Lemma 4.3 below and then proceed to prove it.

Lemma 4.3 (Multilinearization of polynomial of elementary symmetric polynomials)**.** *Fix a prime number p and a field $\mathbb{F}$ with $\mathrm{char}(\mathbb{F}) = p$. Fix a variable parameter $r \in \mathbb{N}$.*
Let $F(\mathbf{y}) \in \mathbb{F}[y_1, \ldots, y_r]$ be a polynomial with individual degree strictly less than p. Then $\mathsf{ml}[F(e_1(\mathbf{y}), e_p(\mathbf{x}), \ldots, e_{p^{r-1}}(\mathbf{x}))]$ has a circuit of size $\mathcal{O}(n^5 \log n)$ and depth 5.

Proof of Lemma 4.3. Suppose the polynomial $F(\mathbf{y}) \in \mathbb{F}[y_1, \ldots, y_r]$ is:

$$F(\mathbf{y}) \;=\; \sum_{\boldsymbol{\mu}} \lambda_{\boldsymbol{\mu}} \mathbf{y}^{\boldsymbol{\mu}},$$

where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_r)$ denotes the exponent vector of a monomial. Recall that the individual degree of $F(\mathbf{y})$ is at most $\leqslant (p-1)$. Consider a monomial $\mathbf{y}^{\boldsymbol{\mu}}$ with a non-zero coefficient $\lambda_{\boldsymbol{\mu}}$ in $F(\mathbf{y})$. We will multilinearize $\mathfrak{m} = \prod_{i=1}^{r} e_{p^{i-1}}(\mathbf{x})^{\mu_i}$. We do it by multilinearizing two products at a time using Lemma 4.11. Defining $\mu_0 := 1$, we have,

$$\mathfrak{m} \;=\; \underbrace{e_1(\mathbf{x}) \cdots e_1(\mathbf{x})}_{\mu_1 \text{ times}} \;\cdots\; \underbrace{e_{p^{r-1}}(\mathbf{x}) \cdots e_{p^{r-1}}(\mathbf{x})}_{\mu_r \text{ times}} \;=\; \prod_{\ell=1}^{\boldsymbol{\mu}} e_{\alpha_\ell}(\mathbf{x}),$$

where $\alpha_\ell = p^{i-1}$ if $\ell \in \left[\sum_{\ell=1}^{i-1} \mu_\ell + 1,\ \sum_{\ell=1}^{i} \mu_i\right]$.

Claim 4.12. *Let $\boldsymbol{\mu} \in \{0, 1, \ldots, p-1\}^r$ denote an exponent vector as described above. Then for any $k \in [|\boldsymbol{\mu}|]$, the following holds:*

- *(Constant-depth circuit). There exists polynomials $R_{\leqslant \alpha_k, j}(\mathbf{x})$'s such that*

$$\mathsf{ml}\left[\prod_{\ell=1}^{k} e_{\alpha_\ell}(\mathbf{x})\right] \;=\; \prod_{\ell=1}^{k} e_{\alpha_\ell}(\mathbf{x}) - \sum_{j=1}^{n} R_{\leqslant \alpha_k, j}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

  and the polynomials $R_{\leqslant \alpha_k, j}(\mathbf{x})$'s have circuits of size $\mathcal{O}(n^3 \cdot k)$ and depth 5.

- *(Structure). There exists coefficients $c^{(1)}_{\leqslant \alpha_k}, \ldots, c^{(n)}_{\leqslant \alpha_k}$ such that*

$$\mathsf{ml}\left[\prod_{\ell=1}^{k} e_{\alpha_\ell}(\mathbf{x})\right] \;=\; \sum_{i=1}^{n} c^{(i)}_{\leqslant \alpha_k}\, e_i(\mathbf{x})$$

Proof of Claim 4.12. We will prove this using induction on k.

Base case: For $k = 1$, we have $\mathsf{ml}[e_{\alpha_1}] = e_{\alpha_1}(\mathbf{x})$ and $R_{\leqslant \alpha_1, j}(\mathbf{x}) = 0$. The claim holds for the base case.

Induction step: Now we assume the induction is true for k and prove it for $(k+1)$. We will first prove the *constant-depth circuit* item for $(k+1)$ and then prove the *structure* item for $(k+1)$.

Using the **structure** item of the induction hypothesis, we have,

$$\mathsf{ml}\left[\prod_{\ell=1}^{k} e_{\alpha_\ell}(\mathbf{x})\right] \;=\; \sum_{i=1}^{n} c^{(i)}_{\leqslant \alpha_k}\; e_i(\mathbf{x}) \tag{17}$$

Using the third item of Fact 1.13,

$$\begin{aligned}
\mathsf{ml}\left[\prod_{\ell=1}^{k+1} e_{\alpha_\ell}(\mathbf{x})\right] \;&=\; \mathsf{ml}\left[\mathsf{ml}\left[\prod_{\ell=1}^{k} e_{\alpha_\ell}(\mathbf{x})\right] \cdot e_{\alpha_{k+1}}(\mathbf{x})\right] \\
&= \mathsf{ml}\left[\sum_{i=1}^{n} c^{(i)}_{\leqslant \alpha_k}\; e_i(\mathbf{x}) \cdot e_{\alpha_{k+1}}(\mathbf{x})\right] \qquad \text{(Using Equation (17))} \\
&= \sum_{i=1}^{n} c^{(i)}_{\leqslant \alpha_k}\; \mathsf{ml}[e_i(\mathbf{x}) \cdot e_{\alpha_{k+1}}(\mathbf{x})] \qquad (18)
\end{aligned}$$

For each $i \in [n]$, we apply the **constant-depth circuit** item from Lemma 4.11 on $e_i(\mathbf{x}) \cdot e_{\alpha_{k+1}}(\mathbf{x})$ to get:

$$\mathsf{ml}[e_i(\mathbf{x}) \cdot e_{\alpha_{k+1}}(\mathbf{x})] \;=\; e_i(\mathbf{x}) \cdot e_{\alpha_{k+1}}(\mathbf{x}) - \sum_{j=1}^{n} D_{i,\alpha_{k+1}}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

where the polynomials $D_{i,\alpha_{k+1}}(\mathbf{x})$ have a circuit of size $\mathcal{O}(n^3)$ and depth 5. Substituting it in Equation (18),

$$\begin{aligned}
\mathsf{ml}\left[\prod_{\ell=1}^{k+1} e_{\alpha_\ell}(\mathbf{x})\right] \;&=\; \sum_{i=1}^{n} c^{(i)}_{\leqslant \alpha_k}\; \left(e_i(\mathbf{x}) \cdot e_{\alpha_{k+1}}(\mathbf{x}) - \sum_{j=1}^{n} D_{i,\alpha_{k+1}}(\mathbf{x}) \cdot (x_j^2 - x_j)\right) \\
&= \left(\sum_{i=1}^{n} c^{(i)}_{\leqslant \alpha_k} e_i(\mathbf{x})\right) e_{\alpha_{k+1}}(\mathbf{x}) - \sum_{j=1}^{n} \left(\sum_{i=1}^{n} c^{(i)}_{\leqslant \alpha_k} D_{i,\alpha_{k+1}}(\mathbf{x})\right) \cdot (x_j^2 - x_j)
\end{aligned}$$

Using Equation (17),

$$\sum_{i=1}^{n} c^{(i)}_{\leqslant \alpha_k} e_i(\mathbf{x}) \;=\; \mathsf{ml}\left[\prod_{\ell=1}^{k} e_{\alpha_\ell}(\mathbf{x})\right] \;=\; \prod_{\ell=1}^{k} e_{\alpha_\ell}(\mathbf{x}) - \sum_{j=1}^{n} R_{\leqslant \alpha_k, j}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

where we use the **constant-depth circuit** item from the induction hypothesis for the last equality. The polynomials $R_{\leqslant\alpha_k,j}(\mathbf{x})$'s have circuits of size $\mathcal{O}(n^3 \cdot k)$ and depth 5 (a $\Sigma\Pi\Sigma\Pi\Sigma$ circuit). Using this in the previous expression, we have

$$\begin{aligned}
&\mathsf{ml}\left[\prod_{\ell=1}^{k+1} e_{\alpha_\ell}(\mathbf{x})\right] \\
&= \left(\prod_{\ell=1}^{k} e_{\alpha_\ell}(\mathbf{x}) - \sum_{j=1}^{n} R_{\leqslant\alpha_k,j}(\mathbf{x}) \cdot (x_j^2 - x_j)\right) e_{\alpha_{k+1}}(\mathbf{x}) - \sum_{j=1}^{n} \left(\sum_{i=1}^{n} c^{(i)}_{\leqslant\alpha_k} D_{i,\alpha_{k+1}}(\mathbf{x})\right) \cdot (x_j^2 - x_j) \\
&= \prod_{\ell=1}^{k+1} e_{\alpha_\ell}(\mathbf{x}) - \sum_{j=1}^{n} \underbrace{\left(\sum_{i=1}^{n} c^{(i)}_{\leqslant\alpha_k} D_{i,\alpha_{k+1}}(\mathbf{x}) + R_{\leqslant\alpha_k,j}(\mathbf{x})\right)}_{:=R_{\leqslant\alpha_{k+1},j}(\mathbf{x})} \cdot (x_j^2 - x_j)
\end{aligned}$$

The polynomial $R_{\leqslant\alpha_{k+1},j}(\mathbf{x})$ has a circuit of size $\mathcal{O}(n^3 \cdot (k+1))$ and depth 5 (a $\Sigma\Pi\Sigma\Pi\Sigma$ circuit). This shows the *constant-depth circuit* item of the induction.

By applying the **structure item** of Lemma 4.11 on $e_i(\mathbf{x}) \cdot e_{\alpha_{k+1}}(\mathbf{x})$, we get,

$$\mathsf{ml}\left[\prod_{\ell=1}^{k} e_{\alpha_\ell}(\mathbf{x})\right] = \sum_{i=1}^{n} c^{(i)}_{\leqslant\alpha_k}\, e_i(\mathbf{x})$$

Substituting it in Equation (18),

$$\mathsf{ml}\left[\prod_{\ell=1}^{k+1} e_{\alpha_\ell}(\mathbf{x})\right] = \sum_{i=1}^{n} c^{(i)}_{\leqslant\alpha_k} \sum_{i'=1}^{n} d_i^{(i')} e_{i'}(\mathbf{x}) = \sum_{i=1}^{n} c^{(i)}_{\leqslant\alpha_{k+1}}\, e_i(\mathbf{x}),$$

where $c^{(i)}_{\leqslant\alpha_{k+1}} = \sum_{j=1}^{n} c^{(j)}_{\leqslant\alpha_k} d_j^{(i)}$. This completes the *structure* item of the induction.
This finishes the induction and thus we have finished the proof of Claim 4.12. ■

Now we employ Claim 4.12 on each monomial $\boldsymbol{\mu}$ with non-zero coefficient and then sum them together. It is easy to verify that there exists polynomials $R_j(\mathbf{x})$'s such that

$$F(\hat{\mathbf{e}}(\mathbf{x})) = \mathsf{ml}[F(\hat{\mathbf{e}}(\mathbf{x}))] + \sum_{j=1}^{n} R_j(\mathbf{x}) \cdot (x_j^2 - x_j),$$

where each polynomial $R_j(\mathbf{x})$ has a circuit of size $\mathcal{O}(p^r r n^3)$ and depth 5 (a $\Sigma\Pi\Sigma\Pi\Sigma$ circuit). Using $r \leqslant 2\log_p n$, we get that each polynomial $R_j(\mathbf{x})$ has a circuit of size $\mathcal{O}(n^5 \log n)$ and depth 5. This finishes the proof of Lemma 4.3. ■

References

[AR01] M. Alekhnovich and A.A. Razborov. "Lower bounds for polynomial calculus: non-binomial case". In: *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*. 2001, pp. 190–199. DOI: 10.1109/SFCS.2001.959893 (cit. on p. 4).

[AGHT20] Yaroslav Alekseev, Dima Grigoriev, Edward A. Hirsch, and Iddo Tzameret. "Semi-algebraic proofs, IPS lower bounds, and the τ-conjecture: can a natural number be negative?" In: *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2020. Chicago, IL, USA: Association for Computing Machinery, 2020, pp. 54–67. ISBN: 9781450369794. DOI: 10.1145/3357713.3384245. URL: https://doi.org/10.1145/3357713.3384245 (cit. on p. 9).

[AF22] Robert Andrews and Michael A. Forbes. "Ideals, determinants, and straightening: proving and using lower bounds for polynomial ideals". In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2022. Rome, Italy: Association for Computing Machinery, 2022, pp. 389–402. ISBN: 9781450392648. DOI: 10.1145/3519935.3520025. URL: https://doi.org/10.1145/3519935.3520025 (cit. on p. 8).

[BIKPP96] Paul Beame, Russell Impagliazzo, Jan Krajíček, Toniann Pitassi, and Pavel Pudlák. "Lower Bounds on Hilbert's Nullstellensatz and Propositional Proofs". In: *Proceedings of the London Mathematical Society* s3-73.1 (1996), pp. 1–26. DOI: https://doi.org/10.1112/plms/s3-73.1.1. URL: https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/plms/s3-73.1.1 (cit. on pp. 4, 8).

[BDS24] C.S. Bhargav, Sagnik Dutta, and Nitin Saxena. "Improved Lower Bound, and Proof Barrier, for Constant Depth Algebraic Circuits". In: 16.4 (Nov. 2024). ISSN: 1942-3454. DOI: 10.1145/3689957. URL: https://doi.org/10.1145/3689957 (cit. on p. 6).

[BGIP01] Sam Buss, Dima Grigoriev, Russell Impagliazzo, and Toniann Pitassi. "Linear Gaps between Degrees for the Polynomial Calculus Modulo Distinct Primes". In: *Journal of Computer and System Sciences* 62.2 (2001), pp. 267–289. ISSN: 0022-0000. DOI: https://doi.org/10.1006/jcss.2000.1726. URL: https://www.sciencedirect.com/science/article/pii/S0022000000917264 (cit. on p. 4).

[BIKPRS97] Samuel R. Buss, Russell Impagliazzo, Jan Krajícek, Pavel Pudlák, Alexander A. Razborov, and Jirí Sgall. "Proof Complexity in Algebraic Systems and Bounded Depth Frege Systems with Modular Counting". In: *Comput. Complex.* 6.3 (1997), pp. 256–298 (cit. on p. 4).

[CEI96] Matthew Clegg, Jeffery Edmonds, and Russell Impagliazzo. "Using the Groebner basis algorithm to find proofs of unsatisfiability". In: *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*. STOC '96. Philadelphia, Pennsylvania, USA: Association for Computing Machinery, 1996, pp. 174–183. ISBN: 0897917855. DOI: 10.1145/237814.237860. URL: https://doi.org/10.1145/237814.237860 (cit. on p. 4).

[CK02] Peter Clote and Evangelos Kranakis. *Boolean Functions and Computation Models*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2002. DOI: https://doi.org/10.1007/978-3-662-04943-3 (cit. on p. 4).

[CR79] Stephen A. Cook and Robert A. Reckhow. "The relative efficiency of propositional proof systems". In: *Journal of Symbolic Logic* 44.1 (1979), pp. 36–50. DOI: 10.2307/2273702 (cit. on p. 4).

[EGLT25] Tal Elbaz, Nashlen Govindasamy, Jiaqi Lu, and Iddo Tzameret. *Lower bounds against the Ideal proof system in finite fields*. 2025 (cit. on p. 13).

[For14] Michael A. Forbes. "Polynomial identity testing of read-once oblivious algebraic branching programs". PhD thesis. Massachusetts Institute of Technology, Cambridge, MA, USA, 2014. URL: https://hdl.handle.net/1721.1/89843 (cit. on pp. 50, 51).

[For24] Michael A. Forbes. "Low-Depth Algebraic Circuit Lower Bounds over Any Field". In: *39th Computational Complexity Conference (CCC 2024)*. Ed. by Rahul Santhanam. Vol. 300. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 31:1–31:16. ISBN: 978-3-95977-331-7. DOI: 10.4230/LIPIcs.CCC.2024.31. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.CCC.2024.31 (cit. on pp. 6, 13, 21).

[FKS16] Michael A. Forbes, Mrinal Kumar, and Ramprasad Saptharishi. "Functional Lower Bounds for Arithmetic Circuits and Connections to Boolean Circuit Complexity". In: *31st Conference on Computational Complexity (CCC 2016)*. Ed. by Ran Raz. Vol. 50. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2016, 33:1–33:19. ISBN: 978-3-95977-008-8. DOI: 10.4230/LIPIcs.CCC.2016.33. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.CCC.2016.33 (cit. on p. 6).

[FSTW21] Michael A. Forbes, Amir Shpilka, Iddo Tzameret, and Avi Wigderson. "Proof Complexity Lower Bounds from Algebraic Circuit Complexity". In: *Theory Comput.* 17 (2021), pp. 1–88. URL: https://theoryofcomputing.org/articles/v017a010/ (cit. on pp. 4–6, 8, 11, 15, 18–25, 50–52).

[GHT22] Nashlen Govindasamy, Tuomas Hakoniemi, and Iddo Tzameret. "Simple Hard Instances for Low-Depth Algebraic Proofs". In: *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*. 2022, pp. 188–199. DOI: 10.1109/FOCS54457.2022.00025 (cit. on pp. 4–6, 9, 15, 20–25).

[Gri98] D. Grigoriev. "Tseitin's tautologies and lower bounds for Nullstellensatz proofs". In: *Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No.98CB36280)*. 1998, pp. 648–652. DOI: 10.1109/SFCS.1998.743515 (cit. on pp. 4, 16).

[GR00] Dima Grigoriev and Alexander A. Razborov. "Exponential Lower Bounds for Depth 3 Arithmetic Circuits in Algebras of Functions over Finite Fields". In: *Applicable Algebra in Engineering, Communication and Computing* 10 (2000), pp. 465–487. URL: https://api.semanticscholar.org/CorpusID:26189857 (cit. on p. 6).

[GP18] Joshua A. Grochow and Toniann Pitassi. "Circuit Complexity, Proof Complexity, and Polynomial Identity Testing: The Ideal Proof System". In: *J. ACM* 65.6 (Nov. 2018). ISSN: 0004-5411. DOI: 10.1145/3230742. URL: https://doi.org/10.1145/3230742 (cit. on pp. 4–6).

[GRS23] Venkatesan Guruswami, Atri Rudra, and Madhu Sudan. *Essential Coding Theory*. Available online: https://cse.buffalo.edu/faculty/atri/courses/coding-theory/book/web-coding-book.pdf. 2023. URL: https://cse.buffalo.edu/faculty/atri/courses/coding-theory/book/web-coding-book.pdf (cit. on p. 14).

[HLT24] Tuomas Hakoniemi, Nutan Limaye, and Iddo Tzameret. "Functional Lower Bounds in Algebraic Proofs: Symmetry, Lifting, and Barriers". In: *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*. STOC 2024. Vancouver, BC, Canada: Association for Computing Machinery, 2024, pp. 1396–1404. ISBN: 9798400703836. DOI: 10.1145/3618260.3649616. URL: https://doi.org/10.1145/3618260.3649616 (cit. on pp. 4, 9).

[IPS99] Russell Impagliazzo, Pavel Pudlák, and Jirí Sgall. "Lower Bounds for the Polynomial Calculus and the Gröbner Basis Algorithm". In: *Comput. Complex.* 8.2 (1999), pp. 127–144. DOI: https://doi.org/10.1007/s000370050024 (cit. on p. 4).

[Kra95] Jan Krajícek. *Bounded arithmetic, propositional logic, and complexity theory*. Vol. 60. Encyclopedia of mathematics and its applications. Cambridge University Press, 1995. DOI: https://doi.org/10.1017/CBO9780511529948 (cit. on p. 4).

[Kra19] J. Krajíček. *Proof Complexity*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2019. ISBN: 9781108416849. URL: https://books.google.dk/books?id=uOyKuQEACAAJ (cit. on p. 4).

[Kra15] Jan Krajíček. "A reduction of proof complexity to computational complexity for $\mathsf{AC}^0[p]$ Frege systems". In: *Proceedings of the American Mathematical Society* 143.11 (2015), pp. 4951–4965. ISSN: 00029939, 10886826. URL: http://www.jstor.org/stable/24507779 (cit. on p. 5).

[LST21] Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. "Superpolynomial Lower Bounds Against Low-Depth Algebraic Circuits". In: *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*. 2021, pp. 804–814. DOI: 10.1109/FOCS52979.2021.00083 (cit. on pp. 6, 13, 21–23).

[Luc78] Edouard Lucas. "Théorie des Fonctions Numériques Simplement Périodiques". In: *American Journal of Mathematics* 1.2 (1878), pp. 184–196. ISSN: 00029327, 10806377. URL: http://www.jstor.org/stable/2369308 (visited on 11/04/2024) (cit. on p. 33).

[Nis91] Noam Nisan. "Lower bounds for non-commutative computation". In: *Proceedings of the Twenty-Third Annual ACM Symposium on Theory of Computing*. STOC '91. New Orleans, Louisiana, USA: Association for Computing Machinery, 1991, pp. 410–418. ISBN: 0897913973. DOI: 10.1145/103418.103462. URL: https://doi.org/10.1145/103418.103462 (cit. on pp. 6, 51).

[PT16] Toniann Pitassi and Iddo Tzameret. "Algebraic proof complexity: progress, frontiers and challenges". In: *ACM SIGLOG News* 3.3 (Aug. 2016), pp. 21–43. DOI: 10.1145/2984450.2984455. URL: https://doi.org/10.1145/2984450.2984455 (cit. on p. 4).

[Raz09] Ran Raz. "Multi-linear formulas for permanent and determinant are of super-polynomial size". In: *J. ACM* 56.2 (Apr. 2009). ISSN: 0004-5411. DOI: 10.1145/1502793.1502797. URL: https://doi.org/10.1145/1502793.1502797 (cit. on pp. 6, 20).

[RY09] Ran Raz and Amir Yehudayoff. "Lower Bounds and separations for constant depth multilinear circuits". English. In: *Computational Complexity* 18.2 (June 2009), pp. 171–207. ISSN: 1016-3328. DOI: 10.1007/s00037-009-0270-8 (cit. on p. 20).

[Raz87] Alexander A. Razborov. "Lower bounds on the size of bounded depth circuits over a complete basis with logical addition". In: *Mathematical notes of the Academy of Sciences of the USSR* 41 (1987), pp. 333–338. URL: https://api.semanticscholar.org/CorpusID:121744639 (cit. on p. 5).

[Raz98] Alexander A. Razborov. "Lower Bounds for the Polynomial Calculus". In: *Comput. Complex.* 7.4 (1998), pp. 291–324. DOI: https://doi.org/10.1007/s000370050013 (cit. on p. 4).

[Sap21] Ramprasad Saptharishi. "A survey of lower bounds in arithmetic circuit complexity". In: *Github survey* (2021). URL: https://github.com/dasarpmar/lowerbounds-survey/releases/tag/v9.0.3 (cit. on p. 7).

[SW01] Amir Shpilka and Avi Wigderson. "Depth-3 arithmetic circuits over fields of characteristic zero". In: *Comput. Complex.* 10.1 (2001), pp. 1–27. DOI: 10.1007/PL00001609. URL: https://doi.org/10.1007/PL00001609 (cit. on p. 14).

[SY10] Amir Shpilka and Amir Yehudayoff. "Arithmetic Circuits: A Survey of Recent Results and Open Questions". In: *Foundations and Trends® in Theoretical Computer Science* 5.3–4 (2010), pp. 207–388. ISSN: 1551-305X. DOI: 10.1561/0400000039. URL: http://dx.doi.org/10.1561/0400000039 (cit. on p. 7).

[Smo87] Roman Smolensky. "Algebraic methods in the theory of lower bounds for Boolean circuit complexity". In: *Proceedings of the nineteenth annual ACM symposium on Theory of computing* (1987). URL: https://api.semanticscholar.org/CorpusID:2214101 (cit. on p. 5).

[Smo93] Roman Smolensky. "On representations by low-degree polynomials". In: *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science* (1993), pp. 130–138. URL: https://api.semanticscholar.org/CorpusID:206559792 (cit. on p. 5).

A Appendix

A.1 Details of roABP-IPS$_{\text{LIN}'}$ Lower Bound

We recall some standard definitions and lemmas that are useful for understanding the complexity of roABPs. For more details, please refer to [FSTW21; For14].

Definition A.1 (Coefficient matrix)**.** *Consider $f \in \mathbb{F}[\mathbf{x}, \mathbf{y}]$. The coefficient matrix of C_f is defined with the following entries from $\mathbb{F}$:*

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

where $\mathrm{Coeff}_{\mathbf{x}^\mathbf{a},\mathbf{y}^\mathbf{b}}(f)$ denotes the coefficient of the monomial $\mathbf{x}^\mathbf{a}\mathbf{y}^\mathbf{b}$ in f.

Definition A.2 (Coefficient space)**.** *Consider $f \in \mathbb{F}[\mathbf{x}, \mathbf{y}]$. The* space of $\mathbb{F}[\mathbf{x}][\mathbf{y}]$ coefficients of f *is defined as:*

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

where $\mathrm{Coeff}_{\mathbf{x}|\mathbf{y}^\mathbf{b}}$ denotes the coefficient of $\mathbf{y}^\mathbf{b}$ when f is viewed as a polynomial in the $\mathbf{y}$-variables, with coefficients from the ring $\mathbb{F}[\mathbf{x}]$. The space of $\mathbb{F}[\mathbf{y}][\mathbf{x}]$ coefficients of f *is defined similarly.*

For any subset S of polynomials over a field $\mathbb{F}$, we will use $\dim(S)$ to denote the dimension of the $\mathbb{F}$-linear span of polynomials in S.

Lemma A.3 (Coefficient dimension equals rank of C_f [Nis91])**.** *For any $f \in \mathbb{F}[\mathbf{x}, \mathbf{y}]$:*

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

Lemma A.4 (Coefficient dimension captures roABP width [Nis91][For14])**.** *For any $f(x_1, \ldots, x_n)$, if f is computable by a width-r* roABP*, then $r \geqslant \max_{i\in n} \dim(\mathbf{Coeff}_{\mathbf{x}_{\leqslant i}|\mathbf{x}_{>i}}(f))$. Further, there is a width-r* roABP *for f, where $r = \max_{i\in n} \dim(\mathbf{Coeff}_{\mathbf{x}_{\leqslant i}|\mathbf{x}_{>i}}(f))$.*

Definition A.5 (Evaluation space)**.** *For $f \in \mathbb{F}$, the* space of $\mathbb{F}[\mathbf{x}][\mathbf{y}]$ evaluations of f over a set $S \subseteq \mathbb{F}$ *is defined as:*

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

Omitting the S in the notation will denote that $S = \mathbb{F}$. The space of $\mathbb{F}[\mathbf{y}][\mathbf{x}]$ evaluations of f over a set S *is defined similarly.*

Lemma A.6 (Evaluation dimension $\leqslant$ coefficient dimension)**.** *For $f \in \mathbb{F}[\mathbf{x}][\mathbf{y}]$ and $S \subseteq \mathbb{F}$,*

$$\mathbf{Eval}_{\mathbf{x}|\mathbf{y},S}(f) \subseteq \mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f)$$

which implies that $\dim(\mathbf{Eval}_{\mathbf{x}|\mathbf{y},S}(f)) \leqslant \dim(\mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f))$. If $|S|$ is greater than the individual degree of each variable in f, then $\mathbf{Eval}_{\mathbf{x}|\mathbf{y},S}(f) = \mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f)$.

Fact A.7 (Dimension of polynomials = dimension of leading monomials [For14])**.** *Let $S = \{f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\} \subseteq \mathbb{F}[\mathbf{x}]$. For each f_i, let $\mathrm{LM}(f_i)$ denote the leading monomial of f_i based on some monomial ordering. Then, $\dim \mathrm{span}\, S = \dim \mathrm{span}\{\mathrm{LM}(f_i) : f_i \in S\}$.*

The following lemma proves an analog of the coefficient dimension lower bound from [FSTW21] for the positive characteristic case using the degree lower bound in Lemma 2.4.

Lemma A.8 (Coefficient dimension lower bound from degree lower bound for fixed partition(Proposition 5.8 [FSTW21]))**.** *Let $n \in \mathbb{N}$. For any $\boldsymbol{\alpha} \in \mathbb{F}^n$ and $\beta \in B_{\boldsymbol{\alpha}}$, let $f_{\boldsymbol{\alpha},\beta}(\mathbf{x}, \mathbf{y})$ be a polynomial that computes*

$$\frac{1}{\sum_{i=1}^n \alpha_i x_i y_i - \beta}$$

on $\{0, 1\}^n$. Let S be a finite subset of $\mathbb{F}$. Then, for a uniformly randomly chosen $\boldsymbol{\alpha} \sim S^n$:

$$\Pr_{\boldsymbol{\alpha}\sim S^n}[\dim(\mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f_{\boldsymbol{\alpha},\beta})) \geqslant 2^n] \geqslant 1 - \frac{2^{2n}}{|S|}$$

Proof. Lemma A.6 implies that

$$\dim(\mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f_{\boldsymbol{\alpha},\beta})) \geqslant \dim\{f_{\boldsymbol{\alpha},\beta}(\mathbf{x},\mathbf{b})\}_{\mathbf{b}\in\{0,1\}^n}$$

For any $\mathbf{b} = (b_1,\ldots,b_n) \in \{0,1\}^n$, if $U_{\mathbf{b}} := \{i \in [n] : b_i = 1\}$ then :

$$f_{\boldsymbol{\alpha},\beta}(\mathbf{x},\mathbf{b}) = \frac{1}{\sum_{i\in U_{\mathbf{b}}} \alpha_i x_i - \beta}$$

Lemma 2.4 tells us that for a randomly chosen $\boldsymbol{\alpha} \sim S^n$:

$$\Pr_{\boldsymbol{\alpha}\sim S^n}[\forall \mathbf{b} \in \{0,1\}^n : \deg f_{\boldsymbol{\alpha},\beta,U_{\mathbf{b}}}(\mathbf{x}) = |\mathbf{b}|] \geqslant 1 - \frac{2^{2n}}{|S|}$$

In particular, for a uniformly random $\boldsymbol{\alpha} \sim S^n$, for any $\mathbf{b} \in \{0,1\}^n$, the leading monomial of $f_{\boldsymbol{\alpha},\beta,U_{\mathbf{b}}}(\mathbf{x})$ is $c_{\mathbf{b}} \cdot \prod_{i:b_i=1} x_i$ for some $c_{\mathbf{b}} \in \mathbb{F}\backslash\{0\}$. Combining this with Fact A.7, we get that with probability at least $1 - (2^{2n}/|S|)$:

$$\dim(\mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f_{\boldsymbol{\alpha},\beta})) \geqslant \dim\{f_{\boldsymbol{\alpha},\beta}(\mathbf{x},\mathbf{b})\}_{\mathbf{b}\in\{0,1\}^n} \geqslant \dim\{\mathrm{ml}(f_{\boldsymbol{\alpha},\beta}(\mathbf{x},\mathbf{b}))\}_{\mathbf{b}\in\{0,1\}^n} \geqslant 2^n$$

since each multilinear restriction $\mathrm{ml}(f_{\boldsymbol{\alpha},\beta}(\mathbf{x},\mathbf{b}))$ generates a different multilinear monomial as its leading monomial, and thus the space contains all 2^n multilinear monomials on $\mathbf{x}$. Here, we also used the fact the multilinearization operator is a linear map and does not increase the dimension. ■

The following fact relates the coefficient dimension of a polynomial $f \in \mathbb{F}[\mathbf{x},\mathbf{y},\mathbf{z}]$ over $\mathbb{F}(\mathbf{z})$ to the coefficient dimension of $f(\mathbf{x},\mathbf{y},\mathbf{b})$ over $\mathbb{F}$ for any $\mathbf{b} \in \mathbb{F}^n$.

Fact A.9 (Coefficient dimension over $\mathbb{F}(\mathbf{z})$ $\geqslant$ coefficient dimension over $\mathbb{F}$ (Lemma 5.12 [FSTW21]))**.** *Let $f \in \mathbb{F}[\mathbf{x},\mathbf{y},\mathbf{z}]$. Let $f_{\mathbf{z}}$ denote f as a polynomial in $\mathbb{F}[\mathbf{z}][\mathbf{x},\mathbf{y}]$ so that for any $\mathbf{b} \in \mathbb{F}^n$, $f_{\mathbf{b}}(\mathbf{x},\mathbf{y}) = f(\mathbf{x},\mathbf{y},\mathbf{b}) \in \mathbb{F}[\mathbf{x},\mathbf{y}]$. Then for any $\mathbf{b} \in \mathbb{F}^n$:*

$$\dim_{\mathbb{F}(\mathbf{z})} \mathbf{Coeff}_{\mathbf{x}|\mathbf{y}} f_{\mathbf{z}}(\mathbf{x},\mathbf{y}) \geqslant \dim_{\mathbb{F}} \mathbf{Coeff}_{\mathbf{x}|\mathbf{y}} f_{\mathbf{b}}(\mathbf{x},\mathbf{y})$$

Using this fact, [FSTW21] proves a coefficient dimension lower bound over $\mathbb{F}(\mathbf{z})$ for any partition of variables, using the coefficient dimension lower bound over $\mathbb{F}$ for a fixed partition of variables. We observe that their proofs work even when we replace their coefficient dimension lower bound by a suitable version over fields of positive characteristic (Lemma A.8) using the degree lower bound over positive characteristic.

Lemma A.10 (Coefficient dimension lower bound for any partition of variables (Proposition 5.13 [FSTW21]))**.** *Let $n \in \mathbb{N}$. For any $\boldsymbol{\alpha} \in \mathbb{F}^{\binom{2n}{2}}$ and $\beta \in B_{\boldsymbol{\alpha}}$, let $f_{\boldsymbol{\alpha},\beta}(\mathbf{x} = (x_i)_{i\in[2n]}, \mathbf{z} = (z_{i,j})_{i<j\leqslant 2n})$ be a polynomial which computes*

$$\frac{1}{\sum_{i<j\leqslant n} \alpha_{i,j} z_{i,j} x_i x_j - \beta}$$

on the Boolean cube. Let $S \subseteq \mathbb{F}$. Call an $\boldsymbol{\alpha} \in S^{\binom{2n}{2}}$ good *if for any partition $\mathbf{x} = (\mathbf{u},\mathbf{v})$ with $|\mathbf{u}| = |\mathbf{v}| = n$:*

$$\dim_{\mathbb{F}(\mathbf{z})}(\mathbf{Coeff}_{\mathbf{u}|\mathbf{v}}(f_{\boldsymbol{\alpha},\beta})) \geqslant 2^n$$

where $f_{\boldsymbol{\alpha},\beta}$ is viewed as a polynomial in $\mathbb{F}[\mathbf{z}][\mathbf{x},\mathbf{y}]$ with coefficients in $\mathbb{F}[\mathbf{z}]$.
Then, a uniformly randomly chosen $\boldsymbol{\alpha} \in S^{\binom{2n}{2}}$ is good with probability $\geqslant 1 - \frac{\binom{2n}{n}2^{2n}}{|S|}$.

Proof. For any balanced partition $\mathbf{x} = (\mathbf{u}, \mathbf{v})$ where $|\mathbf{u}| = |\mathbf{v}| = n$, we can embed $\sum_{i\in[n]} u_i v_i - \beta$ in $\sum_{i<j\leqslant n} \alpha_{i,j} z_{i,j} x_i x_j - \beta$ by a natural restriction $\mathbf{z} = \mathbf{b}_{\mathbf{u},\mathbf{v}} \in \{0,1\}^{\binom{2n}{2}}$ that sets $z_{i,j}$ to 1 if $x_i = u_k$, $x_j = v_k$, and 0 otherwise. So, for every such restriction $\mathbf{b}_{\mathbf{u},\mathbf{v}}$ that corresponds to a balanced partition:

$$f(\mathbf{u}, \mathbf{v}, \mathbf{b}_{\mathbf{u},\mathbf{v}}) = \frac{1}{\sum_{i\in[n]} u_i v_i - \beta}$$

For any fixed choice of balanced partition $\mathbf{b}_{\mathbf{u},\mathbf{v}} \in \{0,1\}^{\binom{2n}{2}}$, Lemma A.8 tells us that for a uniformly randomly chosen $\boldsymbol{\alpha} \in S^{\binom{2n}{n}}$:

$$\Pr_{\boldsymbol{\alpha}\in S^{\binom{2n}{n}}}[\dim(\mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f_{\boldsymbol{\alpha},\beta}(\mathbf{u}, \mathbf{v}, \mathbf{b}))) \geqslant 2^n] \geqslant 1 - \frac{2^{2n}}{|S|}$$

Applying a union bound over all $\binom{2n}{n}$ choices of balanced partitions $\mathbf{x} = (\mathbf{u}, \mathbf{v})$ implies that for a uniformly randomly chosen $\boldsymbol{\alpha} \in S^{\binom{2n}{2}}$:

$$\Pr_{\boldsymbol{\alpha}\in S^{\binom{2n}{n}}}[\forall \mathbf{x} = (\mathbf{u}, \mathbf{v}) : \dim_{\mathbb{F}}(\mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f_{\boldsymbol{\alpha},\beta}(\mathbf{u}, \mathbf{v}, \mathbf{b}_{\mathbf{u},\mathbf{v}}))) \geqslant 2^n] \geqslant 1 - \frac{\binom{2n}{n}2^{2n}}{|S|}$$

Finally, applying Fact A.9 implies that for a uniformly randomly chosen $\boldsymbol{\alpha} \in S^{\binom{2n}{2}}$:

$$\Pr_{\boldsymbol{\alpha}\in S^{\binom{2n}{n}}}[\forall \mathbf{x} = (\mathbf{u}, \mathbf{v}) : \dim_{\mathbb{F}(\mathbf{z})}(\mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f_{\boldsymbol{\alpha},\beta}(\mathbf{u}, \mathbf{v}, \mathbf{b}_{\mathbf{u},\mathbf{v}}))) \geqslant 2^n] \geqslant 1 - \frac{\binom{2n}{n}2^{2n}}{|S|}$$

■

Theorem A.11 (Functional lower bound against roABP in any order of variables)**.** *Let $n \in \mathbb{N}$. Let $p \in \mathbb{N}$ be any prime. Let $\tilde{\mathbb{F}}$ be a field of characteristic p and size p^{2k}, where k is the smallest integer that satisfies $p^k > \binom{2n}{n}2^{2n}$. Let β be an arbitrary element in $\tilde{\mathbb{F}}\backslash\mathbb{F}$, where $\mathbb{F}$ denotes the subfield of size p^k. For any $\boldsymbol{\alpha} \in \mathbb{F}^{\binom{2n}{2}}$, let $f_{\boldsymbol{\alpha}}(\mathbf{x} = (x_i)_{i=1}^{2n}, \mathbf{z} = (z_{i,j})_{i,j\in[n]})$ be a polynomial which agrees with*

$$\frac{1}{\sum_{i<j} \alpha_{i,j} z_{i,j} x_i x_j - \beta}$$

on the Boolean cube. Then there exists an $\boldsymbol{\alpha} \in \mathbb{F}^{\binom{2n}{2}}$ such that any roABP *that computes $f_{\boldsymbol{\alpha}}$ in any order of variables requires size $\geqslant 2^n$.*

Proof of Theorem A.11. We will instantiate Lemma A.10 for the field $\tilde{\mathbb{F}}$ and the set $S = \mathbb{F}$. Thus, choosing $\beta \in \tilde{\mathbb{F}}\backslash\mathbb{F}$ ensures that for any choice of $\boldsymbol{\alpha} = (\alpha_{i,j})_{1\leqslant i<j\leqslant 2n} \in S^{\binom{2n}{2}}$, β will be in $B_{\boldsymbol{\alpha}}$ (which we recall to be the complement of all possible subset sums of $\boldsymbol{\alpha}$). With the above choices, it follows from Lemma A.10 that for a uniformly randomly chosen $\boldsymbol{\alpha} \in S^{\binom{2n}{2}}$,

$$\Pr_{\boldsymbol{\alpha}\in S^m}[\forall \mathbf{x} = (\mathbf{u}, \mathbf{v}) : \dim_{\mathbb{F}(\mathbf{z})}(\mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f_{\boldsymbol{\alpha},\beta}(\mathbf{u}, \mathbf{v}, \mathbf{b}_{\mathbf{u},\mathbf{v}}))) \geqslant 2^n] \geqslant 1 - \frac{\binom{2n}{n}2^{2n}}{|S|} > 0$$

for $|S| = |\mathbb{F}| > \binom{2n}{n}2^{2n}$, where $\mathbf{x} = (\mathbf{u}, \mathbf{v})$ denotes any balanced partition of $\mathbf{x}$. In particular, this implies that there exists an $\boldsymbol{\alpha} \in S^{\binom{2n}{n}}$ such that for any balanced partition $\mathbf{x} = (\mathbf{u}, \mathbf{v})$,

$$\dim_{\mathbb{F}(\mathbf{z})}(\mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f_{\boldsymbol{\alpha},\beta}(\mathbf{u}, \mathbf{v}, \mathbf{b}_{\mathbf{u},\mathbf{v}}))) \geqslant 2^n \tag{19}$$

Now, suppose $f(\mathbf{x}, \mathbf{z})$ is computable by a width-r roABP in some order of variables. Using $f_{\mathbf{z}}$ to denote f as a polynomial in $\mathbb{F}[z][\mathbf{x}]$, it follows that $f_{\mathbf{z}}$ is also computable by a width-r roABP over the fraction field $\mathbb{F}(\mathbf{z})$ in the induced order of variables on $\mathbf{x}$. By splitting the $\mathbf{x}$ variables in half along the induced order, using Equation (19) along with Nisan's characterization of width of roABPs (Lemma A.4), we obtain the required lower bound. ■

A.2 Proof of Claim 3.4

Claim 3.4. *For any exponent vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ with $|\boldsymbol{\mu}| \leqslant D$, there exists polynomials $E_{\boldsymbol{\mu},1}(\mathbf{x}), \ldots, E_{\boldsymbol{\mu},n}(\mathbf{x})$ such that the following holds:*

$$((\mathbf{x}^{\boldsymbol{\mu}})^2 - \mathbf{x}^{\boldsymbol{\mu}}) \;=\; \sum_{\substack{j \in [n] \\ \mu_j > 0}} E_{\boldsymbol{\mu},j}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

and for each $j \in [n]$ with $\mu_j > 0$, the polynomial $E_{\boldsymbol{\mu},j}(\mathbf{x})$ has a circuit of size $\mathcal{O}(nD^2)$ and depth 2 (a $\Pi\Sigma$ circuit).

Proof of Claim 3.4. We will prove it by induction on the cardinality of $\mathrm{Supp}(\boldsymbol{\mu})$, which is defined as follows:

$$\mathrm{Supp}(\boldsymbol{\mu}) \;=\; \{j \in [n] \mid \mu_j > 0\}\,.$$

Base case: Suppose $|\mathrm{Supp}(\boldsymbol{\mu})| = 1$ and $\mu_1 > 0$. If $\mu_1 = 1$, then we can set $E_{\boldsymbol{\mu},1}(\mathbf{x}) = 1$. Otherwise, if $\mu_1 > 1$, then

$$x_1^{2\mu_1} - x_1^{\mu_1} \;=\; (x_1^{2\mu_1} - x_1) - (x_1^{\mu_1} - x_1)$$

We have the following identity for any $j \geqslant 2$:

$$z^j - z \;=\; (z^{j-2} + \ldots + z + 1) \cdot (z^2 - z)$$

Using this we get,

$$\begin{aligned} x_1^{2\mu_1} - x_1^{\mu_1} \;&=\; (x_1^{2\mu_1 - 2} + \ldots + x_1 + 1) \cdot (x_1^2 - x_1) - (x_1^{\mu_1 - 2} + \ldots + x_1 + 1) \cdot (x_1^2 - x_1) \\ &= \underbrace{(x_1^{2\mu_1 - 2} + \ldots + x_{\mu_1 - 1})}_{:= E_{\boldsymbol{\mu},1}(\mathbf{x})} \cdot (x_1^2 - x_1) \end{aligned}$$

The polynomial $E_{\boldsymbol{\mu},1}(\mathbf{x})$ has a circuit of size $\mathcal{O}(D^2)$ and depth 2 (a $\Sigma\Pi$ circuit).

Induction step: Assume this is true for all $\boldsymbol{\mu}$ with $|\boldsymbol{\mu}| \leqslant D$ and $|\mathrm{Supp}(\boldsymbol{\mu})| = k$. Consider any arbitrary exponent vector $\boldsymbol{\mu}$ with $|\boldsymbol{\mu}| \leqslant D$ and $|\mathrm{Supp}(\boldsymbol{\mu})| = (k+1)$. Let t be the largest element in $\mathrm{Supp}(\boldsymbol{\mu})$ and let $\boldsymbol{\nu}$ be the exponent vector with $\nu_t = 0$ and $\nu_i = \mu_i$ for all $i \neq t$. We have,

$$((\mathbf{x}^{\boldsymbol{\mu}})^2 - \mathbf{x}^{\boldsymbol{\mu}}) \;=\; (x_t^{2\mu_t} - x_t + x_t) \cdot (\mathbf{x}^{\boldsymbol{\nu}})^2 - (x_t^{\mu_t} - x_t + x_t) \cdot \mathbf{x}^{\boldsymbol{\nu}}$$

$$\underbrace{((x_t^{2\mu_t - 2} + \ldots + 1) \cdot (\mathbf{x}^{\boldsymbol{\nu}})^2 - (x_t^{\mu_t - 2} + \ldots + 1) \cdot \mathbf{x}^{\boldsymbol{\nu}})}_{:= E_{\boldsymbol{\mu},t}(\mathbf{x})} \cdot (x_t^2 - x_t) \;+\; x_t \cdot ((\mathbf{x}^{\boldsymbol{\nu}})^2 - \mathbf{x}^{\boldsymbol{\nu}}), \tag{20}$$

where we used the identity $(z^j - z) = (z^{j-2} + \ldots + z + 1) \cdot (z^2 - z)$. Since the exponent vector $\boldsymbol{\nu}$ satisfies $|\boldsymbol{\nu}| \leqslant D$ and $|\mathrm{Supp}(\boldsymbol{\nu})| = k$, we can apply the induction hypothesis on $(\mathbf{x}^{\boldsymbol{\nu}})^2 - \mathbf{x}^{\boldsymbol{\nu}}$. From induction, we know there exists polynomials $E_{\boldsymbol{\nu},j}(\mathbf{x})$ for all $j \in \mathrm{Supp}(\boldsymbol{\nu})$ such that:

$$((\mathbf{x}^{\boldsymbol{\nu}})^2 - \mathbf{x}^{\boldsymbol{\nu}}) \;=\; \sum_{j \in \mathrm{Supp}(\boldsymbol{\nu})} E_{\boldsymbol{\nu},j}(\mathbf{x}) \cdot (x_j^2 - x_j),$$

and the polynomials $E_{\boldsymbol{\nu},j}(\mathbf{x})$ have circuits of size $\mathcal{O}(nD^2)$ and depth 2 (a $\Pi\Sigma$ circuit). Substituting it in Equation (20), we get,

$$((\mathbf{x}^{\boldsymbol{\mu}})^2 - \mathbf{x}^{\boldsymbol{\mu}}) \;=\; E_{\boldsymbol{\mu},t}(\mathbf{x}) \cdot (x_t^2 - x_t) \;+\; \sum_{j \in \mathrm{Supp}(\boldsymbol{\nu})} \underbrace{x_t \cdot E_{\boldsymbol{\nu},j}(\mathbf{x})}_{:= E_{S,j}} \cdot (x_j^2 - x_j),$$

where the polynomials $E_{\boldsymbol{\mu},j}(\mathbf{x})$ have a circuit of size $\mathcal{O}(nD^2)$ and depth 2 (a $\Pi\Sigma$ circuit). Moreover, the polynomials $E_{\boldsymbol{\mu},j}(\mathbf{x})$ are of degree-$2D$ polynomials. This finishes the proof of Claim 3.4. ■

A.3 Proof of Claim 4.5

Claim 4.5. *Let $H(\mathbf{y}) \in \mathbb{F}[y_1, \ldots, y_r]$ denote a polynomial whose individual degree is at most D. Then there exists polynomials $G_1(\mathbf{y}), \ldots, G_r(\mathbf{y})$ such that the following holds:*

$$H(\mathbf{y}) \;=\; \mathsf{inddeg}_p[H(\mathbf{y})] \;+\; \sum_{j=1}^{r} G_j(\mathbf{y}) \cdot (y_j^p - y_j),$$

and for each $j \in [r]$, the polynomial $G_j(\mathbf{y})$ has sparsity at most $D^{r+1}/(p-1)$.

Proof of Claim 4.5. Fix an arbitrary monomial $\mathfrak{m}$ with a non-zero coefficient in the polynomial $H(\mathbf{y})$. Say $\mathfrak{m} = y_1^{\mu_1} \cdots y_r^{\mu_r}$ where for every $j \in [r]$, $0 \leqslant e_i \leqslant D$. Let $S_{\mathfrak{m}} \subseteq [r]$ denote the set of variables whose exponent in $\mathfrak{m}$ is at least p, i.e.

$$S_{\mathfrak{m}} \;=\; \{j \in [r] \mid p \leqslant \mu_j \leqslant D\}$$

Let $\ell \in S_{\mathfrak{m}}$, and let $\mathfrak{m}_{-\ell} := \mathfrak{m}/y_\ell^{\mu_\ell}$. In other words, $\mathfrak{m} = y_\ell^{\mu_\ell} \cdot \mathfrak{m}_{-\ell}$. Then,

$$\begin{aligned} \mathfrak{m} \;&=\; (y_\ell^p - y_\ell + y_\ell)\, y_\ell^{\mu_\ell - p} \cdot \mathfrak{m}_{-\ell} \\ &= y_\ell^{\mu_\ell - p + 1} \cdot \mathfrak{m}_{-j} \;+\; y_\ell^{\mu_\ell - p} \cdot \mathfrak{m}_{-j} \cdot (y_\ell^p - y_\ell) \end{aligned}$$

The monomial $y_\ell^{\mu_\ell - p} \cdot \mathfrak{m}_{-j}$ is a monomial in the polynomial $G_\ell(\mathbf{y})$. We repeat the above step on the monomial $y_\ell^{\mu_\ell - p+1} \cdot \mathfrak{m}_{-j}$. In each step with respect to the variable y_ℓ (for the monomial $\mathfrak{m}$), the

degree of y_ℓ is reducing by $(p-1)$. Thus this step can be repeated $\leqslant D/(p-1)$ times because the individual degree of $H(\mathbf{y})$ is $\leqslant D$. In each step, we get one monomial for the polynomial $G_\ell(\mathbf{y})$, and thus we get $D/(p-1)$ monomials in the polynomial $G_\ell(\mathbf{y})$ from the monomial $\mathfrak{m}$. We do this for every variable in the set $S_{\mathfrak{m}}$.

Finally, we iterate the above steps for every monomial with non-zero coefficient in the polynomial $H(\mathbf{y})$. The sparsity of the polynomial $H(\mathbf{y})$ is at most D^r, since the individual degree of $H(\mathbf{y})$ is $\leqslant D$. For each monomial in the support of $H(\mathbf{y})$, each polynomial $G_\ell(\mathbf{y})$ gets at most $D/(p-1)$ monomials, and hence each polynomial $G_\ell(\mathbf{y})$,

$$\text{sparsity}(G_\ell(\mathbf{y})) \;=\; \text{sparsity}(H(\mathbf{y})) \cdot D/(p-1) \;\leqslant\; D^r \cdot D/(p-1)$$

This finishes the proof of Claim 4.5. ■

A.4 Proof of Claim 4.9

Claim 4.9 (Partial multilinearization of product of univariates)**.** *Let $h_1(z_1), \ldots, h_n(z_n)$ be univariate polynomials where each $h_i(z_i)$ has degree at most D.*
Then there exists degree-1 univariate polynomials $L_1(z_1), \ldots, L_n(z_n)$ and polynomials $B_1(\mathbf{z}), \ldots, B_n(\mathbf{z})$ satisfying the following: For every $k \in [n]$,

$$h_1(z_1) \cdots h_n(z_n) \;=\; \mathsf{ml}_{\leqslant k}\left[\prod_{i=1}^{n} h_i(z_i)\right] \;+\; \sum_{j=1}^{k} B_j(\mathbf{z}) \cdot (z_j^2 - z_j),$$

where

$$\mathsf{ml}_{\leqslant k}\left[\prod_{i=1}^{n} h_i(z_i)\right] \;=\; \prod_{i=1}^{k} L_i(z_i) \cdot \prod_{i=k+1}^{n} h_i(z_i),$$

and for each $j \in [n]$, the polynomial $B_j(\mathbf{z})$ has the following form:

$$B_j(\mathbf{x}) \;=\; \prod_{i=1}^{j-1} L_i(z_i) \cdot \tilde{h}_j(z_j) \cdot \prod_{i=j+1}^{n} h_i(z_i),$$

for some univariate polynomial $\tilde{h}_j(z_j)$.

Proof of Claim 4.9. We will prove this via induction on k.

Base case: For $k = 1$, this is exactly Claim 4.8 where $z = z_1$ and $\mathbf{y} = (z_2, \ldots, z_n)$.

Induction case: Assume the claim is true up to k. Let $Q(\mathbf{z}) \;=\; h_{k+2}(z_{k+2}) \cdots h_n(z_n)$. By induction hypothesis, we have,

$$\mathsf{ml}_{\leqslant k}\left[\prod_{i=1}^{n} h_i(z_i)\right] \;=\; h_{k+1}(z_{k+1}) \;\cdot\; \underbrace{\prod_{i=1}^{k} L_i(z_i) \cdot \prod_{i=k+2}^{n} h_i(z_i)}_{=Q(\mathbf{y})},$$

where $\mathbf{y} = (z_1, \ldots, z_k, z_{k+2}, \ldots, z_n)$. From Observation 4.7,

$$\mathsf{ml}_{\leqslant k+1}\left[\prod_{i=1}^{n} h_i(z_i)\right] \;=\; \mathsf{ml}_{z_{k+1}}\left[\mathsf{ml}_{\leqslant k}\left[\prod_{i=1}^{n} h_i(z_i)\right]\right],$$

Now applying Claim 4.8 on $h_{k+1}(z_{k+1}) \cdot Q(\mathbf{y})$ with respect to the variable z_{k+1}, we get the claim for $k+1$. This finishes the proof of Claim 4.9. ■

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il