

New Bounds for the Ideal Proof System in Positive Characteristic

Amik Raj Behera ^{*}

Nutan Limaye [†]

Varun Ramanathan [‡]

Srikanth Srinivasan [§]

September 14, 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 $AC^0[p]$ -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).

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 [krabook95; CloteKSurvey02; krabook2019].) 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 [BIKPP94] 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 [BIKPP94; BIPRS97; Razborov98; Grigoriev98; IPS99; BGIP01; AR2001].

However, as noted in [FSTW21], sparsity and degree only roughly capture the complexity of algebraic proofs. More recently, Grochow and Pitassi [GP14] 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 [GP14; PitassiTzameretSurvey; FSTW21; GHT; 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 characteristic¹.

1.1 Ideal Proof Systems

We start by describing the general setup for an algebraic (static²) proof system. Let \mathbf{x} denote the set of variables $\{x_1, x_2, \dots, x_n\}$. We are given a set of polynomial axioms $f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$

¹In 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.

²In 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.

and the goal is to show that there is no 0-1 assignment to the variables such that it simultaneously satisfies $\{f_1(\mathbf{x}) = 0, f_2(\mathbf{x}) = 0, \dots, f_m(\mathbf{x}) = 0\}$ 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 $\{f_1(\mathbf{x}) = 0, f_2(\mathbf{x}) = 0, \dots, f_m(\mathbf{x}) = 0\} \cup \{x_i^2 - x_i = 0\}_{i \in [n]}$ are simultaneously not satisfiable, then such a refutation³ can be given by polynomials $A_1(\mathbf{x}), A_2(\mathbf{x}), \dots, A_m(\mathbf{x})$ and $B_1(\mathbf{x}), B_2(\mathbf{x}), \dots, B_n(\mathbf{x})$ such that

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

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 ?? 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 [GP14], 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 [GHT], 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 $\text{AC}^0[p]$ -Frege. A long-standing open question in proof complexity, open for almost three decades [kra2015], is to prove superpolynomial lower bounds against $\text{AC}^0[p]$ -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 [Razborov1987] and Smolensky [Smolensky1987; Smolensky1993] resolved the Boolean circuit lower bound question for $\text{AC}^0[p]$, 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 [GP14] which showed that constant-depth-IPS over characteristic p can efficiently simulate $\text{AC}^0[p]$ -Frege proofs. This means that proving

³The 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.

superpolynomial lower bounds against constant-depth-IPS refutations will give superpolynomial lower bounds against $AC^0[p]$ -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 [GP14], [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 ??). Their method was inspired by the notion of functional lower bounds in Boolean circuit complexity [Grigoriev-Razborov; 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 [Nisan] and multilinear formulas [Raz-2009] 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 [GHT] 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 [LST] only worked over characteristic 0; for this and other reasons, the result of [GHT] was also limited to characteristic 0. In a recent paper, however, Forbes [Forbes-LST-CCC] improved the circuit lower bound result of [LST] and proved the same⁴ lower bound over any characteristic.

In light of these results, the next obvious step is to prove the lower bounds of [FSTW21; GHT] 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 [GP14; FSTW21]). *Let $f_1, \dots, f_m \in \mathbb{F}[x_1, \dots, x_n]$ 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, \dots, 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, \dots, f_m is an algebraic circuit $C \in \mathcal{C}$ in variables $x_1, \dots, x_n, y_1, \dots, y_m, z_1, \dots, z_n$ such that

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

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} -IPS_{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} -IPS_{LIN'} refutation.

⁴Some parameters in the lower bound by [LST] were subsequently improved by [BDS24] and [Forbes-LST-CCC] achieves those improved parameters.

⁵The subset-sum instances from [FSTW21; GHT] 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 $\mathcal{C}\text{-IPS}_{\text{LIN}'}$ 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 $\mathcal{C}\text{-IPS}_{\text{LIN}}$ proof system, our lower bounds results hold in the setting of the stronger $\mathcal{C}\text{-IPS}_{\text{LIN}'}$ 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 $f(\mathbf{x}) \in \mathbb{F}[x_1, \dots, x_n]$ is a *multilinear* if the individual degree is at most 1. For a polynomial $f(\mathbf{x})$, the *multilinearization* operator, denoted by $\text{ml}[\cdot]$, changes for each variable x_j and any k , every occurrence of x_j^k in $f(\mathbf{x})$ to x_j .

A polynomial $f(\mathbf{x}) \in \mathbb{F}[x_1, \dots, x_n]$ is said to be a *symmetric polynomial* if the polynomial remains invariant under any permutation of the input variables. For a degree parameter $0 \leq d \leq n$, the d^{th} elementary symmetric polynomial $e_{n,d}(x_1, \dots, x_n)$ is defined to be the following multilinear polynomial $e_{n,d}(x_1, \dots, x_n) = \sum_{\substack{S \subseteq [n] \\ |S|=d}} \prod_{i \in S} x_i$. 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 $\Sigma\Pi$ 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. $\Sigma\Pi$ 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 addition 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 $V_0 = \{s\}, V_1, V_2, \dots, V_n = \{t\}$. For each $i \in \{1, 2, \dots, n\}$, 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 to be the sum of polynomials computed by all s -to- t paths. The *width* of an roABP is $\max_{0 \leq i \leq n} |V_i|$ 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],[sapharishisurvey]).

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-IPS_{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-IPS_{LIN'} refutation⁷ 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-IPS_{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-IPS_{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-IPS_{LIN'} and multilinear-formula-IPS_{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 $f = \sum_{1 \leq i < j \leq n} \alpha_{i,j} z_{i,j} x_i x_j - \beta$ has no Boolean satisfying assignment.*
- *Any roABP-IPS_{LIN'} refutation of f must have size at least $2^{\Omega(n)}$.*
- *Moreover, any multilinear-formula-IPS_{LIN'} refutation of f must have size at least $n^{\Omega(\log n)}$ and for $\Delta = o(\log n / \log \log n)$, any product-depth⁸- Δ multilinear-formula-IPS refutation requires size $\geq n^{\Omega\left(\frac{1}{\Delta^2} \left(\frac{n}{\log n}\right)^{1/\Delta}\right)}$.*

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] and in a paper by Andrews and Forbes [AF22]. Additionally, in a recent work Hakoniemi, Limaye,

⁷Note that sparse-IPS_{LIN} (a weaker system than sparse-IPS_{LIN'}) is equivalent to the Nullstellensatz proof system of [BIKPP94].

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

and Tzameret [HLT24] presented instances that were hard to refute for $\text{roABP-IPS}_{\text{LIN}'}$ and for multilinear-formula- $\text{IPS}_{\text{LIN}'}$ 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- $\text{IPS}_{\text{LIN}'}$ 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 $f = \sum_{1 \leq i < j < k < \ell \leq n} \alpha_{i,j,k,\ell} z_{i,j,k,\ell} x_i x_j x_k x_\ell - \beta$ has no Boolean satisfying assignment.*
- *Any multilinear constant-depth- $\text{IPS}_{\text{LIN}'}$ refutation of f must have size $n^{\omega(1)}$.*

The characteristic 0 (or large characteristic) version of the above theorem was presented in [GHT]. 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 [GHT] 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 [AGHT] similarly use polynomial constraints with coefficients from exponentially large domains. Specifically [AGHT] 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 ?? 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_{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_{LIN} refutations, even in the setting of positive characteristic.

Theorem 1.8 (Upper bounds for (non-multilinear) constant-depth- IPS_{LIN}). *Fix a prime number p . The following holds for any natural numbers n and k .*

Let $f \in \mathbb{F}_{p^k}[x_1, \dots, x_n]$ 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 $\alpha_1, \dots, \alpha_n, \beta \in \mathbb{F}_{p^k}$. Suppose the degree-1 polynomial $\sum_{i=1}^n \alpha_i x_i - \beta \in \mathbb{F}_{p^k}[x_1, \dots, x_n]$ 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 $O(k \cdot p)$ and size $O(n/kp)^{O(kp)}$.

In particular, if $p = O(1)$ 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 ?? and ?? 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 ?? leverages this and yields an efficient refutation. However, if $\beta \in \mathbb{F}_{p^k}$, then our proof of ?? falls apart. We handle this separately in ??, but we do not match ?? qualitatively. More precisely, ?? yields a $\text{poly}(n, p)$ -sized non-multilinear constant-depth refutations, but ?? 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 .*

- ?? shows that the exponential field size in ??, ?? and ?? 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 $\text{roABP-IPS}_{\text{LIN}}$ refutations of size $2^{o(n)}$.¹⁰
- ?? also shows that the multilinearity assumption in ?? 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 $f(x_1, \dots, x_n)$ of the form

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

⁹Suppose 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 $d = o(n)$.

¹⁰When 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, \dots, 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 $f_1, \dots, f_m \in \mathbb{F}[x_1, \dots, x_n]$ 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 $A(\mathbf{x}), B_1(\mathbf{x}), \dots, B_n(\mathbf{x})$ 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 ?? 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 ??.

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

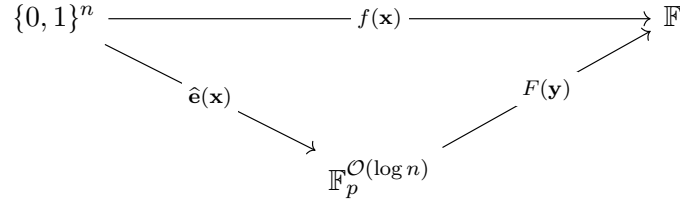
To prove that the polynomial has a refutation over constant-depth circuits, we first prove that for every j , $L_j(\mathbf{x}) = \alpha_1^{p^j} x_1 + \alpha_2^{p^j} x_2 + \dots + \alpha_n^{p^j} x_n - \beta^{p^j}$ can be expressed as a multiple of $L(\mathbf{x})$ modulo the ideal $\mathbf{x}^p - \mathbf{x}$, 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 ??, 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 ?? for complete details.

Upper bounds for symmetric polynomials Now we discuss the proof outline for ?. For ease of exposition, we explain the ideas for the case of $m = 1$ in ?, 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 polynomials¹¹ i.e. $e_1(\mathbf{x}), \dots, e_n(\mathbf{x})$. 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 ?? 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 $\text{char}(\mathbb{F}) = p$. 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 ??) on the unsatisfiability¹² 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 $\text{poly}(n)$ -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, \dots, 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 ??).

It remains to *prove* via constant-depth circuits that $F(\hat{\mathbf{e}}(\mathbf{x}))$ agrees with $f(\mathbf{x})$ on the Boolean cube,

¹¹This follows from the Fundamental Theorem of Symmetric Polynomials.

¹²To 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 .

¹³Loosely 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.

i.e. $F(\hat{\mathbf{e}}(\mathbf{x})) - f(\mathbf{x})$ 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 $\text{poly}(n)$ -sized constant-depth circuits such that

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

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

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

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

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

1.5 Related Work

In an independent work, Elbaz, Govindasamy, Lu, and Tzameret [EGLT-25] consider related questions. Using the recent lower bound of Forbes [Forbes-LST-CCC], which proves the positive characteristic version of the constant-depth formula lower bound of [LST], 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 $f(x_1, \dots, x_n)$, 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 $\text{ml}[f(\mathbf{x})]$ 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 $\text{ml}[f(\mathbf{x})]$ is equal to the $\text{ml}[g(\mathbf{x})]$.
- $\text{ml}[f(\mathbf{x})g(\mathbf{x})] = \text{ml}[\text{ml}[f(\mathbf{x})]\text{ml}[g(\mathbf{x})]]$.

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

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

Theorem 1.15 (Ben-Or's construction for elementary symmetric polynomials). *(See [Shpilka-Wigderson]).* Let \mathbb{F} be a field with $|\mathbb{F}| > n$. Then for every $d \in [n]$, the d^{th} elementary symmetric polynomial $e_d(x_1, \dots, x_n)$ 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 $\gamma_1, \dots, \gamma_{n+1} \in \mathbb{F}$ 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.

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

Theorem 1.16 (Polynomial Identity Lemma). *(See [GuruswamiRudraSudanCodingTheory]).* 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:

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

For a natural number k and variables (z_1, \dots, z_n) , we will use $(\mathbf{z}^k - \mathbf{z})$ to denote the following ideal $(\mathbf{z}^k - \mathbf{z}) := (z_1^k - z_1, \dots, z_n^k - z_n) \subseteq \mathbb{F}[z_1, \dots, z_n]$. 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 $\text{char}(\mathbb{F}) = p$. Then for any $a, b \in \mathbb{F}$, we have, $(a + b)^p = a^p + b^p$. More generally, for any $a_1, \dots, a_m \in \mathbb{F}$, we get, $(a_1 + \dots + a_m)^p = a_1^p + \dots + a_m^p$.*

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 $\mathbb{F}[x_1, \dots, x_n]$. For a set of polynomials $f_1, \dots, f_m \in \mathbb{F}[\mathbf{x}]$, the ideal generated by f_i 's, denoted by (f_1, \dots, f_m) is defined as:*

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

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

$$\mathbb{V}(f_1, \dots, f_m) = \{ \mathbf{a} \in \bar{\mathbb{F}}^n \mid f_1(\mathbf{a}) = \dots = f_m(\mathbf{a}) = 0 \}.$$

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 $f_1, \dots, f_m \in \mathbb{F}[x_1, \dots, x_n]$ 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 $(f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$. In other words, there exists polynomials $A_1, \dots, A_m \in \mathbb{F}[x_1, \dots, x_n]$ such that*

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

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 $(A_1(\mathbf{x}), \dots, A_m(\mathbf{x}))$ as a *Nullstellensatz certificate*¹⁴ for the system $\{f_1(\mathbf{x}), \dots, f_m(\mathbf{x})\}$. 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 $P_1, \dots, P_m \in \mathbb{F}[x_1, \dots, x_n]$ 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 $\{P_1, \dots, P_r\}$ of size $\mathcal{O}(sm)$ and depth $\Delta + 2$.

Proof. **TOPROVE 0** ■

?? 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 ??, we start by proving a tight degree lower bound (??) over positive characteristic. Using our positive characteristic variant of the degree lower bound, we then recover the lower bound results from [FSTW21] and [GHT] over positive characteristic.

¹⁴There are infinitely many Nullstellensatz certificates for a system $\{f_1, \dots, f_m\}$. 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 $\mathbf{a} = (a_1, \dots, a_n) \in \{0, 1\}^n$ 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. [Grigoriev98]), 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, \dots, x_n) . Fix any $\beta \in \mathbb{F}' \setminus \mathbb{F}$. For any $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{F}^n$, let $f_\alpha(\mathbf{x})$ 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 $\alpha \sim S^n$:

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

Proof. **TOPROVE 1** ■

Note that ?? 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 2.3. *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, \dots, x_n) . Fix any $\beta \in \mathbb{F}' \setminus \mathbb{F}$. For any $\emptyset \neq U \subseteq [n]$, let $f_{\alpha, U}(\mathbf{x})$ 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 $\alpha \sim S^n$ chosen uniformly at random:

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

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

Proof. **TOPROVE 2** ■

2.2 Sparse-IPS_{LIN'} Lower Bound

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

Proposition 2.4 (Sparsity lower bound (Proposition 5.6 [FSTW21])). *Let $n \geq 8$. Let \mathbb{F} be a field of characteristic $> n$. Let $\beta \in \mathbb{F} \setminus \{0, \dots, n\}$. Suppose $f(\mathbf{x})$ be a polynomial such that*

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

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

The proof uses two observations.

1. ([FSTW21]) If $f(\mathbf{x})$ has sparsity s , then a random restriction ρ will ensure that $\deg(\rho(f)) \leq \log(s) + 1$ 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 $|S| \geq n/4$. 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 $\geq |S|$. Combining the above observations with the degree lower bound, we get that $n/4 \leq \log(s) + 1$ or $s \geq 2^{n/4-1}$.

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

Theorem 2.5. *Let $n \geq 8$. 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 $\tilde{\mathbb{F}} \setminus \mathbb{F}$, where \mathbb{F} denotes the subfield of size p^k . For any $\alpha \in \mathbb{F}^n$, let $f_\alpha(\mathbf{x})$ 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 $\alpha \in \mathbb{F}^n$ such that f_α has sparsity $\geq 2^{\Omega(n)}$

2.3 roABP – IPS_{LIN'} Lower Bound

?? tells us that for a random choice of coefficients α and any $U \subseteq [n]$, the inverse of $\sum_{i \in U} \alpha_i x_i - \beta$ 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*¹⁵ of g .

¹⁵These 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 dimension¹⁶ of g is at least 2^n , and thus, any roABP computing g must have width $\geq 2^n$.

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 ???. 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 2.6 (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^{2^k} , where k is the smallest integer that satisfies $p^k > 2^{2n}$. Let β be an arbitrary element in $\tilde{\mathbb{F}} \setminus \mathbb{F}$, where \mathbb{F} denotes the subfield of size p^k . For any $\alpha \in \mathbb{F}^n$, let $f_\alpha(\mathbf{x}, \mathbf{y})$ 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 $\alpha \in \mathbb{F}^n$ such that any roABP that computes f_α in any order of variables where \mathbf{x} precedes \mathbf{y} requires width $\geq 2^n$.

Theorem 2.7 (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^{2^k} , where k is the smallest integer that satisfies $p^k > \binom{2n}{n} 2^{2n}$. Let β be an arbitrary element in $\tilde{\mathbb{F}} \setminus \mathbb{F}$, where \mathbb{F} denotes the subfield of size p^k . For any $\alpha \in \mathbb{F}^{\binom{2n}{2}}$, let $f_\alpha(\mathbf{x} = (x_i)_{i=1}^{2n}, \mathbf{z} = (z_{i,j})_{i,j \in [n]})$ 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 $\alpha \in \mathbb{F}^{\binom{2n}{2}}$ such that any roABP that computes f_α in any order of variables requires size $\geq 2^n$.

2.4 Multilinear-formula-IPS Lower Bound

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

¹⁶Again, 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 2.8 (Raz-Yehudayoff [RY09][Raz-2009]). *Let $f \in \mathbb{F}[x_1, \dots, x_{2n}, \mathbf{z}]$ 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 $\mathbf{x} = (x_1, \dots, x_{2n})$:*

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

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

Theorem 2.9 (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 $\mathbb{F}_{p^{2k}} \setminus \mathbb{F}$, where \mathbb{F} denotes the subfield of size p^k . For any $\alpha \in \mathbb{F}^{\binom{2n}{n}}$, let $f_{\alpha}(\mathbf{x} = (x_i)_{i=1}^{2n}, \mathbf{z} = (z_{i,j})_{i,j \in [n]})$ 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 $\alpha \in \mathbb{F}^{\binom{2n}{n}}$ such that any multilinear-formula computing f_{α} requires size $\geq n^{\Omega(\log n)}$ and for $\Delta = o(\log n / \log \log n)$, any product-depth- Δ multilinear-formula computing f_{α} requires size $\geq n^{\Omega(\frac{1}{\Delta^2}(\frac{n}{\log n})^{1/\Delta})}$.

While this immediately implies multilinear-formula- $\text{IPS}_{\text{LIN}'}$ 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- $\text{IPS}_{\text{LIN}'}$ refutation. Thus, the lower bounds work against multilinear-formula-IPS.

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

In [GHT], Govindasamy, Hakoniemi, and Tzameret prove super polynomial lower bounds against constant-depth multilinear $\text{IPS}_{\text{LIN}'}$ 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 2.10 (Constant-depth functional lower bounds [GHT]). *Let $n, \Delta \in \mathbb{N}_+$ with $\Delta \leq \mathcal{O}(\log \log \log n)$ and assume that $\text{char}(\mathbb{F}) = 0$. 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 [GHT]. Their proof requires the $\text{char } \mathbb{F} = 0$ condition for two reasons:

1. They use the results of Limaye, Srinivasan, and Tavenas [LST], which gave superpolynomial lower bounds against constant-depth circuits over any field \mathbb{F} with $\text{char}(\mathbb{F}) = 0$ or greater than the degree d of the hard polynomial. In particular, they use the result that over fields with $\text{char}(\mathbb{F}) = 0$ 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.¹⁷
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 [Forbes-LST-CCC], which extends the results of [LST] to arbitrary fields. In particular, we will use the following statement from [Forbes-LST-CCC], 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 2.11. [Forbes-LST-CCC]. *Let \mathbb{F} be an arbitrary field. Let $\mathbf{x} = \mathbf{x}_1 \sqcup \mathbf{x}_2 \sqcup \dots \sqcup \mathbf{x}_d$ 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 $\text{poly}(s, \Theta(\frac{d}{\log d})^d)$ -size circuit of product-depth 2Δ .*

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

Overview of [GHT]

1. Using the *word polynomials* framework of [LST], construct a *knapsack polynomial* $\text{ks}_{\mathbf{w}}$ (for a partition given by a word $w \in \mathbb{Z}^d$) with the property that the set-multilinear projection of $\frac{1}{\text{ks}_{\mathbf{w}}}$ 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 $\text{ks}_{\mathbf{w}}$ (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 $1/\text{ks}_{\mathbf{w}}$ over $\{0, 1\}^n$ that has a small constant-depth circuit. Moreover by the set-multilinearization of [LST], there is a small set-multilinear constant-depth circuit computing the set-multilinear projection of $1/\text{ks}_{\mathbf{w}}$.
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 $\text{IPS}_{\text{LIN}'}$ lower bound follows.

In [GHT], the proof for the hardness of $\frac{1}{\text{ks}_{\mathbf{w}}}$ requires the underlying field to be of large characteristic, essentially because it requires the degree lower bound from [FSTW21], which requires large characteristic. To make ?? work over fields of positive characteristic, we will employ our degree lower bound from ?? with a variant of the knapsack polynomial; the rest of the proof remains

¹⁷They also use other ideas from [LST] such as relative rank, word polynomial, etc., but those ideas do not require any restrictions on the characteristic of the underlying field.

the same as that of ?? . To provide the necessary details, we first describe the construction of the knapsack polynomial. Then, we state the particular claim from [GHT] that uses the degree lower bound from [FSTW21]. Finally, we show how our degree lower bound ?? fits into the rest of the proof.

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

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 $\overline{X}(w) = (X(w_1), \dots, X(w_d))$ of sets of variables. Define the *positive indices* and *negative indices* of \mathbf{w} as:

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

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

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:

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

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 $A_{\mathbf{w}}^{(i)} \cap B_{\mathbf{w}}^{(j)} \neq \emptyset$ (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 $A_{\mathbf{w}}^{(i)} \cap B_{\mathbf{w}}^{(j)} \neq \emptyset$ (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:

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

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

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

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

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

$$\text{ks}_{\mathbf{w},\alpha} := \left(\sum_{i \in P_{\mathbf{w}}} \alpha_i \sum_{\sigma \in \{0,1\}^{A_{\mathbf{w}}^{(i)}}} x_{\sigma}^{(i)} f_{\sigma}^{(i)} \right) - \beta \quad (5)$$

where $\alpha = (\alpha_i)_{i \in P_{\mathbf{w}}} \in \mathbb{F}^{|P_{\mathbf{w}}|}$, and β will be chosen from an extension field $\tilde{\mathbb{F}} \supset \mathbb{F}$ so that $\text{ks}_{\mathbf{w},\alpha}$ 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 $1/\text{ks}_{\mathbf{w}}$ over $\{0,1\}^n$ requires superpolynomially large set-multilinear constant-depth circuits, [GHT] shows that $M_{\mathbf{w}}(f)$ is full-rank.

Lemma 2.12 (Rank lower bound lemma (Lemma 6 [GHT])). *Let $\mathbf{w} \in \mathbb{Z}^d$ be a balanced word, and let f be the multilinear polynomial such that*

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

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

With this lemma, the lower bound follows via the arguments from [LST]. 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 [GHT] 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 $\sigma(m)$ ([GHT] 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), [GHT] 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 ??, 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 $\prod_{i \in U_T} x_{\sigma_i}^{(i)}$, which is the product of all the variables that show up in the denominator of

$$\frac{1}{\tau_m(\text{ks}_{\mathbf{w}})} = \frac{1}{\sum_{i \in U_T} x_{\sigma(i)}^{(i)} - \beta}$$

where $U_T = \{i \in P_{\mathbf{w}} : A_{\mathbf{w}}^{(i)} \subseteq B_{\mathbf{w}}^T\}$, and for each $i \in P_{\mathbf{w}}$, $\sigma(i)$ is the unique indexing string that agrees with $\sigma(m)$ on $A_{\mathbf{w}}^{(i)}$, 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 [GHT] includes helpful figures and the reader is encouraged to refer to the paper.

Let us recall our variant of $\text{ks}_{\mathbf{w}}$:

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

where $\alpha = (\alpha_i)_{i \in P_{\mathbf{w}}} \in \mathbb{F}^{|P_{\mathbf{w}}|}$. To prove ?? 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 2.13. *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 $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{F}^m$, $T \subseteq N_{\mathbf{w}}$ and any m_T that is a set-multilinear monomial on $\mathbf{w}|_T$, let $f_{\alpha,T,m_T}(\mathbf{x})$ be the unique multilinear polynomial that agrees with the function*

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

on the Boolean cube, where $\beta \in \mathbb{F}$ is chosen so that $\text{ks}_{\mathbf{w},\alpha}$ 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 $\alpha \in S^m$ chosen uniformly at random:

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

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

Proof. **TOPROVE 3** ■

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

Theorem 2.14 ([GHT] over positive characteristic). *Let $n, \Delta \in \mathbb{N}_+$ with $\Delta \leq \mathcal{O}(\log \log \log n)$. Let $p \in \mathbb{N}$ be any prime. Let $\tilde{\mathbb{F}}$ be a field of characteristic p and size p^{2^k} , where k is the smallest integer that satisfies $p^k > 2^{C(\log n)^2}$ for an absolute constant¹⁸ $C \geq 1$. Let β be an arbitrary element in $\tilde{\mathbb{F}} \setminus \mathbb{F}$, where \mathbb{F} denotes the subfield of size p^k . For any $\alpha \in \mathbb{F}^{n^4}$, Let f_{α} 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 $\alpha \in \mathbb{F}^{n^4}$ such that any circuit of product-depth Δ computing f_{α} has size at least

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

¹⁸This C is a fixed constant that depends on the exact choice of parameters in the proof of [GHT]

The reason for $|\mathbb{F}| > 2^{\Omega((\log n)^2)}$ in ?? : When we instantiate ?? inside the proof of ??, 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 $i \in [d]$, $|w_i| \leq O(\log n)$. Thus, $\sum_{i \in N_{\mathbf{w}}} |w_i| = O((\log n)^2)$, and fighting the union bound in ?? will require the field to be larger than $2^{O((\log n)^2)}$.

3 Non-multilinear Upper Bounds

3.1 Proof of ??

In this section, we prove ?. We start by proving it for a restricted setting when the polynomial $f(\mathbf{x})$ is a degree-1 polynomial. In particular, we prove ??, 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 $L \in \mathbb{F}[x_1, \dots, x_n]$ be a degree-1 polynomial with coefficients from the \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 $L(\mathbf{x}) - \beta$ 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] 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 $\{0, 1\}^n$. ?? shows that if we allow non-multilinear IPS_{LIN} refutation, then the circuit size is small.

Proof. **TOPROVE 4** ■

Now we ready to prove ?? using ?. 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 ??, 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 ??.

Claim 3.2. *For any exponent vector $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)$ with $|\boldsymbol{\mu}| \leq D$, there exists polynomials $E_{\boldsymbol{\mu},1}(\mathbf{x}), \dots, E_{\boldsymbol{\mu},n}(\mathbf{x})$ 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 ?? and proceed to prove it.

Theorem ?? (Upper bounds for (non-multilinear) constant-depth- IPS_{LIN}). *Fix a prime number p . The following holds for any natural numbers n and k .*

Let $f \in \mathbb{F}_{p^k}[x_1, \dots, x_n]$ 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 $\text{poly}(s, p)$.*

Proof. **TOPROVE 5** ■

3.2 Proof of ??

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

Theorem ?? (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 $\alpha_1, \dots, \alpha_n, \beta \in \mathbb{F}_{p^k}$. Suppose the degree-1 polynomial $\sum_{i=1}^n \alpha_i x_i - \beta \in \mathbb{F}_{p^k}[x_1, \dots, x_n]$ 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 $2^{o(n)}$.

Observe that the size bound is the “trivial” one, i.e. a n -variate multilinear polynomial with degree D has at most $\binom{n}{\leq D}$ monomials. Letting $D = \mathcal{O}(kp)$, we get the stated size bound in ??. So in our proof of ??, 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 ??, which we state and prove next.

Lemma 3.3 (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 $\alpha_1, \dots, \alpha_n, \beta \in \mathbb{F}_{p^k}$ for which the equation $\sum_{i=1}^n \alpha_i x_i - \beta = 0$ is unsatisfiable over the Boolean cube $\{0, 1\}^n$.*

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

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

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

Proof. **TOPROVE 6** ■

We now prove ?? using an almost straightforward application of ??.

Proof. **TOPROVE 7** ■

4 Symmetric Refutations in Constant Depth

In this section, we will prove ??, which we recall below.

Theorem ?? (Upper bounds for multilinear symmetric systems). *Fix a field \mathbb{F} . Let $f_1, \dots, f_m \in \mathbb{F}[x_1, \dots, x_n]$ 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-IPSLIN refutation of size $\mathcal{O}(m^2 n^5 \log n)$ and depth 8.*

One of the steps in our proof of ?? 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 $\text{ml}[f(\mathbf{x})]$ agree on the Boolean cube $\{0, 1\}^n$. More formally, we are interested in finding polynomials $B_j(\mathbf{x})$'s such that

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

and the polynomials $B_j(\mathbf{x})$ have a $\text{poly}(n)$ -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 [Lucas]). *Fix a prime number p and any two natural numbers a and b . Denote a and b in their unique p -ary representations as:*

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

Then,

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

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 $\text{char}(\mathbb{F}) = p$. Fix a variable parameter $n \in \mathbb{N}$.*

Let $f(\mathbf{x}) \in \mathbb{F}[x_1, \dots, x_n]$ 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. TOPROVE 8 ■

A key lemma in our proof is the multilinearization lemma ??, which shows that multilinearization of a sparse polynomial in $\hat{\mathbf{e}}(\mathbf{x})$ 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 $\text{char}(\mathbb{F}) = p$. Fix a variable parameter $r \in \mathbb{N}$.*

Let $F(\mathbf{y}) \in \mathbb{F}[y_1, \dots, y_r]$ be a polynomial with individual degree strictly less than p . Then $\text{ml}[F(e_1(\mathbf{y}), e_p(\mathbf{x}), \dots, e_{p^{r-1}}(\mathbf{x}))]$ has a circuit of size $\mathcal{O}(n^5 \log n)$ and depth 5.

We will prove ?? later. For now, we show how it is useful in proving ??.

Proof. TOPROVE 9 ■

4.1 Multilinearization

To show our multilinearization lemma (??), 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 ??). We now define the partial multilinearization and then use it to prove ??.

Definition 4.4 (Partial multilinearization). *Fix any field \mathbb{F} and let $f(\mathbf{x}) \in \mathbb{F}[x_1, \dots, x_n]$. For any $j \in [n]$, let $f^{(\leq j)}(\mathbf{x}) \in \mathbb{F}[x_{j+1}, \dots, x_n][x_1, \dots, x_j]$ denote the polynomial $f(\mathbf{x})$ with variables x_1, \dots, x_j and coefficients in $\mathbb{F}[x_{j+1}, \dots, x_n]$.*

The multilinearization of the polynomial $f(\mathbf{x})$ with respect to the variables $\{x_1, \dots, x_j\}$, denoted by $\text{ml}_{\leq j}[f(\mathbf{x})]$, is defined to be:

$$\text{ml}_{\leq j}[f(\mathbf{x})] := \text{ml}[f^{(\leq j)}(\mathbf{x})]$$

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

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

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

Example: Let $f(\mathbf{x}) = x_1^2 x_2^3 + x_2 x_3^2$. Then,

$$\text{ml}_{\leq 1}[f(\mathbf{x})] = x_1 x_2^3 + x_2 x_3^2, \quad \text{ml}_{\leq 2}[f(\mathbf{x})] = x_1 x_2 + x_2 x_3^2, \quad \text{ml}_2[f(\mathbf{x})] = x_1^2 x_2 + x_2 x_3^2$$

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

Observation 4.5. *For every $j < n$, the following holds: For every polynomial $f(\mathbf{x})$,*

$$\text{ml}_{\leq j+1}[f(\mathbf{x})] = \text{ml}_{j+1}[\text{ml}_{\leq j}[f(\mathbf{x})]]$$

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

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

Claim 4.6 (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 $\text{ml}_z[h(z) \cdot Q(\mathbf{y})]$ denotes the partial multilinearization of the polynomial $h(z) \cdot Q(\mathbf{y})$ with respect to the z variable.*

Then,

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

- The polynomial $\text{ml}_z[h(z) \cdot Q(\mathbf{y})]$ 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. **TOPROVE 10** ■

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 ??. We omit the proof here and it can be found in ??.

Claim 4.7 (Partial multilinearization of product of univariates). *Let $h_1(z_1), \dots, h_n(z_n)$ be univariate polynomials where each $h_i(z_i)$ has degree at most D .*

Then there exists degree-1 univariate polynomials $L_1(z_1), \dots, L_n(z_n)$ and polynomials $B_1(\mathbf{z}), \dots, B_n(\mathbf{z})$ satisfying the following: For every $k \in [n]$,

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

where

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

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 ?? immediately gives us the following corollary.

Corollary 4.8 (Multilinearization of product of univariates). *Let $h_1(z_1), \dots, h_n(z_n)$ be univariate polynomials where each $h_i(z_i)$ has degree at most D .*

Then there polynomials $B_1(\mathbf{z}), \dots, B_n(\mathbf{z})$ such that,

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

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

Lemma 4.9 (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*

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

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 $c_{\alpha,\beta}^{(i)}$'s such that*

$$\text{ml}[e_\alpha(\mathbf{x}) \cdot e_\beta(\mathbf{x})] = \sum_{i=1}^n c_{\alpha,\beta}^{(i)} e_i(\mathbf{x})$$

Proof. **TOPROVE 11** ■

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

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

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

Lemma ??? (Multilinearization of polynomial of elementary symmetric polynomials). *Fix a prime number p and a field \mathbb{F} with $\text{char}(\mathbb{F}) = p$. Fix a variable parameter $r \in \mathbb{N}$.*

Let $F(\mathbf{y}) \in \mathbb{F}[y_1, \dots, y_r]$ be a polynomial with individual degree strictly less than p . Then $\text{ml}[F(e_1(\mathbf{y}), e_p(\mathbf{x}), \dots, e_{p^{r-1}}(\mathbf{x}))]$ has a circuit of size $\mathcal{O}(n^5 \log n)$ and depth 5.

Proof. **TOPROVE 12** ■

A Appendix

A.1 Details of $\text{roABP-IPS}_{\text{LIN}}$ 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; Forbes-thesis].

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} :*

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

where $\text{Coeff}_{\mathbf{x}^{\mathbf{a}}, \mathbf{y}^{\mathbf{b}}}(f)$ 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:*

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

where $\text{Coeff}_{\mathbf{x}|\mathbf{y}^{\mathbf{b}}}(f)$ 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 [Nisan]). *For any $f \in \mathbb{F}[\mathbf{x}, \mathbf{y}]$:*

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

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

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:*

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

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 \leq 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 $\dim(\mathbf{Eval}_{\mathbf{x}|\mathbf{y}, S}(f)) \leq \dim(\mathbf{Coeff}_{\mathbf{x}|\mathbf{y}}(f))$. 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 [Forbes-thesis]). *Let $S = \{f_1(\mathbf{x}), \dots, f_m(\mathbf{x})\} \subseteq \mathbb{F}[\mathbf{x}]$. For each f_i , let $\text{LM}(f_i)$ denote the leading monomial of f_i based on some monomial ordering. Then, $\dim \text{span } S = \dim \text{span}\{\text{LM}(f_i) : f_i \in S\}$.*

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

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

Proof. **TOPROVE 13** ■

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 $\mathbb{F}(\mathbf{z}) \geq$ 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 $\mathbf{b} \in \mathbb{F}^n$:*

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

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 (??) 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 $\alpha \in \mathbb{F}^{\binom{2n}{2}}$ and $\beta \in B_\alpha$, let $f_{\alpha,\beta}(\mathbf{x} = (x_i)_{i \in [2n]}, \mathbf{z} = (z_{i,j})_{i < j \leq 2n})$ be a polynomial which computes*

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

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

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

where $f_{\alpha,\beta}$ 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 $\alpha \in S^{\binom{2n}{2}}$ is good with probability $\geq 1 - \frac{\binom{2n}{n} 2^{2n}}{|S|}$.

Proof. **TOPROVE 14** ■

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^{2^k} , where k is the smallest integer that satisfies $p^k > \binom{2n}{n} 2^{2n}$. Let β be an arbitrary element in $\tilde{\mathbb{F}} \setminus \mathbb{F}$, where \mathbb{F} denotes the subfield of size p^k . For any $\alpha \in \mathbb{F}^{\binom{2n}{2}}$, let $f_{\alpha}(\mathbf{x} = (x_i)_{i=1}^{2n}, \mathbf{z} = (z_{i,j})_{i,j \in [n]})$ 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 $\alpha \in \mathbb{F}^{\binom{2n}{2}}$ such that any roABP that computes f_α in any order of variables requires size $\geq 2^n$.

Proof. TOPROVE 15 ■

A.2 Proof of ??

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

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

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

Proof. TOPROVE 16 ■

A.3 Proof of ??

*

Proof. TOPROVE 17 ■

A.4 Proof of ??

Claim ?? (Partial multilinearization of product of univariates). Let $h_1(z_1), \dots, h_n(z_n)$ be univariate polynomials where each $h_i(z_i)$ has degree at most D .

Then there exists degree-1 univariate polynomials $L_1(z_1), \dots, L_n(z_n)$ and polynomials $B_1(\mathbf{z}), \dots, B_n(\mathbf{z})$ satisfying the following: For every $k \in [n]$,

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

where

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

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. TOPROVE 18 ■