Towards a PIT for log-variate ROABPs

By

Subhayan Saha

Roll No: MCS201818

A Thesis submitted
in Partial Fulfillment of the Requirements
for the degree of
MASTER OF SCIENCE

CHENNAI MATHEMATICAL INSTITUTE

Department of Computer Science
Chennai Mathematical Institute
May 2020

To the Faculty of Chennai Mathematical Instiute:

The members of the Committee appointed to examine the thesis of SUBHAYAN SAHA find it satisfactory and recommend that it be accepted.

Prof. Nitin Saxena, Indian Institute of Technology, Kanpur, Chair

---

,

,

ACKNOWLEDGMENTS

I would start off by thanking Prof. Nitin Saxena for agreeing to be my advisor. I remember very clearly the first time I went as a summer intern under his guidance. I had just taken the complexity I course and as much as the mathematical nature of the problems excited me, I had absolutely no clue about Algebraic Complexity. It was there that I was introduced to polynomial identity testing problem, and honestly, it captured all of my attention. From patiently answering all of my stupid doubts, to all those countless hours of discussion - he is the most perfect guide one can have. I would rush to him whenever I had a doubt or an idea and he would always be welcoming, be it at 9 in the morning or very very late in the evening. He always had calculations, ideas ready whenever we went in for a discussion and I was mostly awestruck by his intuition about the problems. I am really fortunate to have an advisor like him and his dedication is something that has motivated me throughout.

At CMI, I was really fortunate to experience the teaching of Prof. Partha Mukhopadhyay, Prof. Sourav Chakraborty, Prof. KV Subrahmanyam. At IMSc, I had the wonderful opportunity to take courses under Prof. Meena Mahajan, Prof. V Arvind and Prof. Saket Saurabh. The two brilliant courses in Complexity theory by Prof. Partha Mukhopadhyay was the reason that I got interested in this beautiful subject. Also, the two brilliant courses taken by Prof. Meena Mahajan, Communication Complexity and Boolean Function Complexity, introduced me to the other really interesting fields in Complexity Theory.

I would like to thank Sumanta da, for guiding me throughout my stay at IITK. The work in the third chapter of the thesis is with him. His immense knowledge and impeccable clarity about the problems and his dedication towards his work, has been so inspiring for me. All of the work in this thesis is with Pranav and I cannot thank him enough for bearing with

my laziness, my eccentricities, my hand-waving and my stupid, stupid doubts. Even after I left IITK, when we thought we had solved the log-variate ROABP problem, he used to video-call me for hours and explain to me every bit of the discussion that he had with Prof. Saxena. I want to thank Amit da, Pranjal da, for all the discussions, for explaining things to me and clearing my doubts on problems which were not even related to what I was working on and also for the amazing company. I would also like to thank all the others in my lab and the PhD lab, Ashish, Prateek, Mahesh, Rajendra.

I would want to thank Abhiroop, who was my roommate throughout my Masters', my lab-mate, my partner-in-crime and my closest friend throughout. I would specifically thank Subhasis for being there, we had some of the most insane time together. Also, Abhibhav, for introducing me to random math, for the random interesting discussions and also, for introducing me to board games. I would also like to thank all the friends and juniors at CMI who made my last year absolutely enjoyable.

I would also take this opportunity to thank IITK for allowing me to stay there for two semesters as a non-degree student followed by 3 months as a Senior Student Research Associate. I would also like to thank IITK for sponsoring my visit to WACT, 2019 at ICTS, Bangalore. I would also thank CMI for allowing me to spend the first year of my masters' at IITK and for accepting the grades for the courses at IITK.

I would like to end by thanking my parents who made this all possible. I am indebted to them for their love, care and all the support. It would not have been possible to do any of it without them and I fondly dedicate my thesis to them.

ABSTRACT

The motivation of this thesis is to obtain a Polynomial Identity Testing algorithm for the class of log-variate ROABPs. Given a multivariate polynomial, in a certain fixed model of computation, the PIT problem asks whether the input polynomial is identically 0. We have a polynomial-time randomized algorithm for PIT. However, designing a deterministic polynomial-time algorithm for PIT is a long-standing open question in algebraic complexity theory. It has deep connections with both circuit lower bounds and many other algorithmic problems like perfect matching, multivariate polynomial factorization.

We consider the PIT problem in the black-box setting, where we are not allowed to see the internal structure of the circuit, but evaluations at points are allowed. For instance, the randomized algorithm for PIT is a black-box algorithm. Designing a deterministic black-box PIT algorithm for a circuit class is equivalent to finding a set of points such that for every nonzero circuit, the set contains a point where it evaluates to a nonzero value. Such a set is called *hitting set*. So by derandomizing PIT, we mean designing a poly(s)-time computable hitting set for s-variate size s degree s circuits.

Because of bootstrapping results in the PIT domain, there has been more focused research on discovering new techniques that can give efficient PIT algorithms for the "low"-variate models (Usually we use the notion of log-variate, by which, we mean the number of variables is logarithmic with respect to the circuit size)

We introduce the notion of cs $\leq k$ hypothesis. We get a structural characterization of the polynomial when it satisfies the cs $\leq k$ hypothesis with shift $t = (t, t, ..., t)$. The aim is to prove a result that if the cone-size hypothesis is satisfied by the polynomial, then it will have high width. This will give us a poly-time PIT for log-variate ROABPs. We prove this in the bivariate setting and get a $(2dk \ln k)$ hitting set. Proceeding along similar lines, we get ROABP width lower bound of $k^{\frac{1}{n-1}}$ on the polynomials satisfying the structural

characterisation and this gives us a new poly(d, w) hitting set for constant-variate ROABPs with individual degree $\leq d$ and width w.

We conjecture a stronger characterization lemma for polynomials that satisfy the cone-size hypothesis for which computing the width will be relatively easier. We prove this characterization lemma for the trivariate case and give the corresponding width lower bound for this. And we show examples of general families of polynomials which satisfy this new characterization lemma and have width $\geq \sqrt{k}$.

We now change the shift to the sparse PIT shift and then try to analyse the cone-size conjecture in terms of that shift. First we prove the structure lemma, where we give a characterization for the form of the polynomial, after it is shifted by the sparse-PIT map. We prove the width result for all polynomials that satisfy the structural result and have some high degree variable.

We consider polynomials over $\mathbb{F}^k$ and then show that, if we shift the polynomial by a basis isolating weight assignment (BIWA), the new polynomial has cs $\leq k$ concentration. This is already implied from [FGS18], who show that polynomials shifted by a BIWA have a cone-closed basis. We give a simpler proof of this fact and our proof is inspired from the [Gur+15] result who show $\log(k+1)$ support concentration. (Cone-size $\leq k$ concentration is strictly stronger than log-support concentration.)

Table of Contents

DEDICATION

This thesis is dedicated to my parents who have been putting up with all of my eccentricities over the years and have continued to love and support me unconditionally.

Chapter One

Introduction

Polynomials have probably been some of the most interesting and most studied functions in all of mathematics and computer science. Yet so little is understood about the complexity of evaluating them.We are often interested in calculating the number of computation steps (addition, multiplication) required to evaluate a polynomial at a certain point. This notion of complexity is best formalised by representing polynomials in the form of arithmetic circuits. (formal definition (2))

One of the most basic decision problems regarding polynomials is verifying in polynomial time (polynomial in input size) whether it is the 0 polynomial or not. This problem is known as the polynomial identity testing (PIT) problem. When the polynomial is given as a list of coefficients, this problem is trivial. But when a succinct representation of the polynomial is given as input, this problem suddenly becomes extremely non-trivial. There are two kinds of PIT algorithms, *white-box*, where one is allowed to make use of the circuit structure and other one in *black-box*, where one can only evaluate the polynomial at certain points. One way to check whether the polynomial computed by the input circuit is zero is to set its variables to constants. If the circuit evaluates to a non-zero value, we definitely know that the polynomial is non-zero. So one technique is to find a set of points in the underlying field (or maybe an extension), such that if the polynomial is non-zero it will evaluate to a non-zero value on at least one of these points. This is called a *hitting set* for the circuit and

it is equivalent to giving a black-box PIT algorithm.

The randomized PIT algorithm due to ([DL78], [Zip79],[Sch80]) is a black-box algorithm for PIT. If we try to derandomize that trivially, we get an exponential size hitting set which is also computable in same time complexity.

To derandomize the PIT algorithm, we need a polynomial time computable hitting set. [HS80] first showed that any random subset of size poly(s) is a hitting set for the class of size s degree s circuits. The best known construction is by [Mul12], who gave a PSPACE construction of a polynomial size hitting-sets.

A series of bootstrapping works [AGS18], [KST18], [Guo+19] finally show that even saving a single point over the trivial hitting set for arithmetic circuits (i.e giving an explicit hitting set of size $(s-1)^n - 1$) gives a polynomial size hitting set.

PIT has applications in designing various algorithm as well for proving various circuit lower bounds (both in algebraic and boolean setting). Kabanets and Impagliazzo in [[KI03]] proved that a polynomial-time white-box PIT algorithm implies a separation between two algebraic classes (VP $\neq$ VNP) or two boolean classes (NEXP $\notin$ P/poly). The famous polynomial-time primality testing by Agrawal, Kayal and Saxena can be seen as a special instance of PIT problem [AKS02]. Kopparty, Saraf and Shpilka showed the equivalence between polynomial identity testing and deterministic multivariate polynomial factorization [KSS14]. The perfect matching problem reduces to PIT question for a special class of polynomials [[Tut47], [Lov79],[FGT16]]

1.1 Polynomial Identity Testing

Definition 1. *Let $C_{s,d}$ be the set of algebraic circuits of size $\leq s$ computing polynomials in $\mathbb{F}[x_1, ..., x_n]$ of degree $\leq d$. Let $\phi \neq \mathcal{C} \subseteq C_{s,d}$. The* ***Polynomial Identity Testing (PIT)*** *problem is : Given a circuit $C \in \mathcal{C}$, computing a polynomial f_C, determine whether $f_C \equiv 0$. The size of the input circuit is given by 3 parameters s, n, d and the goal is to find an efficient*

algorithm that runs in poly(s, n, d) steps.

The algorithm for solving the PIT problem for a given circuit class $\mathcal{C}$, in the unit-cost model over $\mathbb{F}$ is called **black-box** if for all $C \in \mathcal{C}$, the algorithm only uses the given C to evaluate f_C on a set of points in $\mathbb{F}^n$. If it explores the entire structure of the given circuit, it is called **white-box**.

1.2 Models of Computation

We first fix an underlying field $\mathbb{F}$.

1.2.1 Arithmetic Circuits

Definition 2. *An arithmetic circuit is a directed acyclic graph with one sink (called the output gate). Each of the source vertices (input nodes) are either labeled by a variable x_i or an element from $\mathbb{F}$. Each of the internal nodes are labeled either by $+$ (addition gate) or $\times$ (multiplication gate). Sometimes edges may carry weights that are elements from $\mathbb{F}$.*

The computation is now conducted in a natural way. Every edge collects the polynomial computed at its tail node, scales it up the weight on the edge and sends it to the head node. Each addition gate then computes the sum of all the polynomials given by the incoming edges. Similarly, each multiplication gate then computes the product of all the polynomials given by the incoming edges. The polynomial computed at the output node is the polynomial computed by the circuit.
For each node, the *fan-in* is the in-degree of the node and the *fan-out* is the out-degree of the node.

Without loss of generality, the circuit or formula is assumed to be layered, with edges only between successive layers. The important parameters of an arithmetic circuit are the following:

- **size** - the number of edges which is equal to the number of additions and multiplications we perform to compute the polynomial computed by the circuit
- **depth** - The length of the longest path in the circuit from a leaf gate to an output gate
- **degree** - The syntactic degree of the polynomial computed at the output gate. The degree can be computed in a natural way. At an addition gate, we take the maximum over the degrees of all the children. At a multiplication gate, we take the sum of the degrees of all the children.

  Note, that the syntactic degree might not be the same as the actual degree of the polynomial (mostly owing to cancellations) but it is definitely an upper bound on the actual degree of the polynomial.

1.2.2 Algebraic Branching Programs

This section introduces the algebraic branching program model defined by Nisan [Nis91], and gives the definitions of the subclasses of this model that we will consider. In particular, we will define the restricted class of circuits called read-once oblivious algebraic branching program.

Definition 3. *An ABP over $\mathbb{F}$ is a directed acyclic layered graph with vertex set V and edge-set $E = E_1 \sqcup E_2 \sqcup ... \sqcup E_d$ where $E_i \subseteq V_{i-1} \times V_i$ with a set of labelling $\mathcal{L}_0, ..., \mathcal{L}_d$ such that each $\mathcal{L}_i : E_i \to \mathbb{F}[\boldsymbol{x}]$, where the labelling to every edge is a polynomial in $F[\boldsymbol{x}]$ of degree ≤ 1. We define the labelling function $\mathcal{L} : E \to \mathbb{F}[\boldsymbol{x}]$ such that $\mathcal{L}|_{E_i} = \mathcal{L}_i$*

- *The vertices are partitioned into $d+1$ layers i.e $V = V_0 = \{s\} \sqcup V_1 \sqcup ... \sqcup V_d = \{t\}$ such that s and t are the set of source and sink respectively.*
- *Each edge e goes from V_{i-1} to V_i for some $i \in [d]$, so $E \subseteq \sqcup_{i \in [d]} V_{i-1} \times V_i$*

- *An edge e from V_{i-1} to V_i is labelled with an element in $\mathcal{L}_i$*

The polynomial computed by the ABP is of the form

$$f = \sum_{p \in path(s,t)} \prod_{e \in p} \mathcal{L}(e)$$

The parameters of the ABP are

- width $(w) = \max_i |V_i|$
- The size of the ABP is the number of vertices which in this case is $w^2 d$.

Definition 4. *Let $\mathbb{F}$ be a field, $n \geq 1$ and let $\pi : [n] \to [n]$ be a permutation. A **read-once (oblivious) algebraic branching program** with variable order π is a depth $= n$ layered $\mathbb{F}[\boldsymbol{x}]$-ABP with labels*

$$\mathcal{L}_i \coloneqq \{f | f \in \mathbb{F}[x_{\pi(i)}]\}$$

The program computes in a known order if π is a fixed known permutation and computes in an unknown order if π is unknown.

An Algebraic Branching Program computes the polynomial by summing over all the paths, the weight of the paths, where the weight of the path is basically the product of the weights of all the edges along that path. An ABP is **read-once** if along each path each variable occurs in at most one label. A read-once ABP is **oblivious** if in each path the variables occur in the same order (some permutation π on $[n]$).

There is also another equivalent notion for an ROABP that is it can be written as a matrix product.

Lemma 1. *Let $f \in \mathbb{F}[\boldsymbol{x}_n]$ and let $\pi : [n] \to [n]$ be a permutation of $[n]$. Then the following statements are equivalent:*

- *The polynomial f is computed by a width -w and individual degree $\leq d$ ROABP in variable order π*

- *There exist matrices* $U, T \in \mathbb{F}^{w \times 1}$ *and* $M_i \in \mathbb{F}^{w \times w}[x_{\pi(i)}]$ *such that* $f = U^T \prod_{i=1}^{n} M_i T$, *writing* f *as a* 1×1 *matrix.*

An ROABP is called commutative when the corresponding matrix product is commutative. For example, when the matrices are diagonal matrices, which corresponds to the circuit model sum-of-products-of-univariate-polynomials. All the PIT results for ROABP (even with known order) also hold for commutative ROABP .

1.2.3 History of PIT for ROABPs

Raz and Shpilka [RS04] gave a poly(n, w, d)-time white-box algorithm for n-variate polynomials computed by a width-w ROABP with individual degree bound d. [Agr+15] give a $O(ndw)^{\log n}$ time hitting set for ROABPs even when the variable order is unknown. They also give a $(ndw)^{O(\log \log w)}$ hitting set for commutative ROABPs

1.3 Contributions of this Thesis

In this thesis, we look a possible approach towards finding a poly-time PIT algorithm for log-variate ROABPs. Because of bootstrapping results in the PIT domain, there has been more focused research on discovering new techniques that can give efficient PIT algorithms for the low-variate models (By log-variate, we mean the number of variables is logarithmic with respect to the circuit size). We already have a polynomial time PIT for the circuit model called the depth-3 diagonal circuits [FGS18] (Depth-3 diagonal circuits compute the sum of power of linear polynomials) . This model of log-variate ROABPs subsumes the depth-3 diagonal model. (2.6) No poly-time PIT is known for this model.

1.3.1 Cone-size Hypothesis

We first introduce the notion of the cone-size $\leq k$ hypothesis with shift $\mathbf{t} = (t, ..., t)$. A polynomial satisfies this hypothesis if after being shifted by $\mathbf{t}$, all of the monomials that have cone-size $\leq k$ have coefficient 0. We get a structural characterization of the polynomial when it satisfies the $cs \leq k$ hypothesis with shift $\mathbf{t}$. The aim is to prove a result that if the cone-size hypothesis is satisfied by the polynomial, then it will have high width (preferably $\sqrt{k}$). This will give us a poly-time PIT for log-variate ROABPs because given a polynomial that has ROABP width w, we need to just shift the polynomial by $\mathbf{t}$ and then check the coefficient of the monomials with cs $\leq w^2$ The number of such monomials in polynomial in the log-variate case. Then we use the algorithm from [FGS18] to extract the coefficient of the low-cone monomials in the black-box setting. This algorithm runs in polynomial time with respect to the cone-size of the monomial (which is again poly-time because of the log-variate regime). So if the given polynomial has an ROABP of width $< w$, then there must be a monomial in cone-size $\leq w^2$. Hence, just checking the coefficient of polynomially many number of monomials, we can output whether the polynomial is identically 0 or not.

We prove this in the bi-variate setting and get a $2dk \ln k + 1$ hitting set. Proceeding along similar lines, for trivariate we get a lower bound of $\sqrt{k}$ on the ROABP width of the polynomial that satisfies the $cs \leq k$ hypothesis. And we give an example to prove that this analysis is tight. For n-variate, we get a width lower bound of $k^{\frac{1}{n-1}}$ which is exponentially better than the trivial bound. This gives us a black-box hitting set of size poly(d, w) for ROABPs of any variable order with individual degree bounded by d and width w.

Our conjecture, though, is that if f satisfies cone-size $\leq k$ hypothesis, then it has width $\sqrt{k}$. So, we conjecture a stronger characterization lemma for which computing the width will be relatively easier. We consider some specific type of polynomials $P_{\mathbf{e}} = \prod_{i<j\in[n]}(x_i - x_j)^{e_{i,j}}$. We also define cs$(P_{\mathbf{e}}) > k$ for a fixed $e \in \mathbb{N}^n$ if for all monomial in the support of $P_{\mathbf{e}}$, cone-size of the monomial is $\geq k$. We conjecture that a polynomial satisfies the cone-size hypothesis

iff it is in the ideal (over the field $\mathbb{F}$) generated by polynomials which have $\text{cs}(P_{\mathbf{e}}) > k$. We prove this conjecture for the trivariate case. More formally, we prove that

Lemma 20: If $f \in \mathbb{F}[x_1, x_2, x_3]$ satisfies $cs \leq k$ hypothesis, then

$$f \in \left\langle P_{\bar{e}} = (x_1 - x_2)^{e_{12}}.(x_1 - x_3)^{e_{13}}.(x_2 - x_3)^{e_{23}} \mid cs(P_{\bar{e}}) > k \right\rangle_{\mathbb{F}[\bar{x}]}$$

And we show examples of families of polynomials which satisfy this new characterization lemma and have width $\sqrt{k}$. More formally we prove the following

Theorem 10:

$$f = \prod_{i<j\in[n]} (x_i - x_j)^l$$

Then any ROABP computing f must have width $> l^{\lfloor n/2 \rfloor}$

1.3.2 Stronger Cone-size Hypothesis

We now change the shift to a much stronger shift and then try to analyse the cone-size conjecture in terms of that shift. A lower bound on the width of an ROABP is given by the maximum possible rank of the partial derivative space with respect to a certain set of monomials. Now, the hardness of proving width lower bounds for the simple $(t, t, ..., t)$ shift is that a lot of monomials map to the same monomial under this shift and it is hard to keep track of the cancellations. A natural alternative which we try in this chapter is that we shift the polynomial by the sparse-PIT map and then we look at a similar version of the cone-size hypothesis.This ensures that after the shift, the cancellations can be controlled much better. First we prove the structure lemma where we describe the form of the polynomial after it is shifted by the sparse-PIT map. More formally, we show that

Lemma 22 If f satisfies $cs \leq k$ hypothesis, define $e' = (e_2, ..., e_n)$

$$f \in \left\langle P_{\bar{e}} = \prod_{i=2}^{n} \left(x_i - t_i(x_1)\right)^{e_i} \,\middle|\, cs(e') > k \right\rangle_{\mathbb{F}[x_1]}$$

Now, following a similar strategy like the previous section, we have to show corresponding width lower bounds. We conjecture that for polynomials. For formally, we conjecture that

Conjecture 3: If

$$f \in \left\langle P_{\bar{e}} = \prod_{i=2}^{n} \left(x_i - t_i(x_1)\right)^{e_i} \;\middle|\; cs(e') > k \right\rangle_{\mathbb{F}[x_1]}$$

then ROABP width $w(f) > k$, where $\bar{t}$ is the sparse PIT map for $\leq k$ sparse polynomials. The simplest case is that of a single summand for which just a simple rank calculation gives the lower bound.

Now let $e_i := \max_e\{\partial_{x_i^e}(f)|_{x_i=0} \neq 0\}$. Let $e := \max_i\{e_i\}$. Then, $w(f) \geq e$. Hence, this gives us a width result for all polynomials that satisfy the structural result and have some high degree variable. Let us assume without loss of generality that, $e = e_2$. More formally we prove that,

Lemma Let $f = a_1 \cdot (x_2 - t_2(x_1))^{e_2} \cdot g_1 + a_2 \cdot (x_2 - t_2)^{e_2'} \cdot g_2 + a_3 \cdot (x_2 - t_2)^{e_2''} \cdot g_3 + \cdots$, where $a_i \in \mathbb{F}$, $e_2 \geq e_2' \geq e_2'' \cdots$, f has arbitrary number of summands and each g_i contains the remaining product for that summand (For example $g_1 = (x_3 - t_3(x_1))^{e_3} \cdots (x_n - t_n(x_1))^{e_n}$). Then f has width, $w(f) > e_2$.

So the only case remains to be proved when all the variables have really low degree but they still combine to satisfy the cone-size hypothesis. An attempt would be to build on the proof of this lemma and proving the same for sets of variables.

1.3.3 Multinomial Matrix and Concentration results

For PIT of ROABPs, we can assume, without loss of generality, that the output gate is an addition gate and we can write the polynomial as

$$\begin{aligned} f &= \sum_i c_i f_i \\ &= \langle c, P \rangle \text{ where } P = (f_1, ..., f_k) \end{aligned}$$

So, we can now consider polynomial maps over a k-dimensional algebra $\mathbb{A}_k$ .We look at the vector-space V_P spanned by the coefficients of the input polynomial P. Usually the goal is to find a "small" set of monomials (S) such that the rank of the space spanned by their coefficients is the same as the rank of V_P. Then $f \neq 0$ iff the projection on the set S i.e. $f_S = \langle c, P_S \rangle \neq 0$. However it is not correct to expect the polynomial to have such a concentration on its own. So, the general idea is to apply some linear transformation to the variables such that the non-zeroness is preserved and low-rank concentration is achieved. Then to test non-zeroness of f, we have to verify if their exists a monomial from this set S in the transformed polynomial that has non-zero coefficient. The usual PIT literature [ASS13], [Agr+15], [FSS14], [Gur+15] looks at low-support monomials, but here following the theme of the thesis, we look at low-cone monomials. (i.e monomials with low cone-size)

A basis isolating weight assignment is an univariate map on the variables that can isolate a basis S for the coefficient space of the polynomial over $\mathbb{A}_k$. Now, let $f \neq 0$ and m^* be the minimum weight monomial which gives a non-zero inner product with c. We can show that substituting the variables by the weight assignment will also keep the coefficient of the monomial non-zero.

In this section, we show that if A is a polynomial over the k-dimensional algebra $\mathbb{A}_k$, shifting by a basis-isolating weight assignment gives us a $cs \leq k$ concentration. More formally, we show that,

Theorem 11 [Isolation to Concentration] Let $A(\mathbf{x}) \in \mathbb{A}_k[\mathbf{x}]$. Let $\mathbf{w}$ be a basis-isolating weight assignment for $A(\mathbf{x})$. Then $A(\mathbf{x} + t^{\mathbf{w}})$ is $cs \leq k$ concentrated.

[FGS18] show that if a polynomial $A \in \mathbb{A}_k[\mathbf{x}]$ is shifted by a basis isolating weight assignment then it has a cone-closed basis. This readily implies our theorem because if A has a cone-closed basis then it has cone-size $\leq k$ rank concentration. But this is a new proof and it is done by strengthening the combinatorial results in the [Gur+15] lemma who are able to show a $\log(k+1)$ concentration. For this, we prove that a certain multinomial matrix has full rank which we prove by proving an alternate equivalent statement.

Lemma 25 Let $f(\bar{x}) \in \mathbb{F}[\bar{x}]$ be a non-zero polynomial of sparsity at most k. Then $f'(\bar{x}) = f(\bar{x} + \bar{1})$ has a monomial of cone-size $\leq k$ with non-zero coefficient.

Chapter Two

Preliminaries

In this chapter, we give a guide to the basic definitions and notations used in this thesis. We also give an introduction to the basic results in polynomial identity testing, describe some algebraic models of computation and some structural results about them.

2.1 Notation and definitions

Let $\mathbf{a}, \mathbf{b} \in \mathbb{N}^n$. Then

$$\binom{\mathbf{a}}{\mathbf{b}} \coloneqq \prod_{i=1}^{n} \binom{a_i}{b_i}$$

By $\mathbf{x}$, we denote the set of variables $\{x_1, ..., x_n\}$. For any $\mathbf{e} \in \mathbb{N}^n$, $\mathbf{x}^\mathbf{e}$ denotes the monomial $\prod_{i=1}^{n} x_i^{e_i}$. The degree of a monomial $\mathbf{x}^\mathbf{e} = \prod_{i=1}^{n} e_i$ and is denoted by $\deg(\mathbf{x}^\mathbf{e})$.

For a field $\mathbb{F}$, $\mathbb{F}[\mathbf{x}]$ denotes the ring of polynomials over the variables $\{x_1, ..., x_n\}$ where the coefficients are coming from $\mathbb{F}$. For a monomial m and a polynomial P, we denote $\text{coeff}_m(P)$ to denote the coefficient of m in P. Again for a polynomial P, we define the **support** of P, denoted by $\text{supp}(f)$ to be the set of monomials m such that $\text{coeff}_m(P) \neq 0$. We define **sparsity** of a polynomial P denoted by $\text{sp}(P)$ to be the cardinality of $\text{supp}(P)$. We say that a polynomial P has degree d if for all monomials $m \in \text{supp}(P)$, $\deg(m) \leq d$. We say that a polynomial f has individual degree d, if for every monomial $m = \prod_{i=1}^{n} x_i^{e_i} \in \text{supp}(P)$, $e_i \leq d$.

A weight assignment on the set of variables is a function $\mathbf{w} : \mathbf{x} \to \mathbb{N}$. We usually attach the weight $\mathbf{w}(x_i) = w_i$ for all $i \in [n]$. For a monomial $m = \mathbf{x}^{\mathbf{e}}$, the weight of the monomial is $\mathbf{w}(m) = \sum_{i=1}^{n} e_i w_i$. Similarly for a set of monomials B, the weight of B is $\mathbf{w}(B) \coloneqq \sum_{m \in B} w(m)$

For a monomial $\mathbf{x}^{\mathbf{e}}$ and a polynomial $P \in \mathbb{F}[\mathbf{x}]$, we denote by $\partial_{\mathbf{x}^{\mathbf{e}}}(P)$ to be the partial derivative of P with respect to the monomial $\mathbf{x}^{\mathbf{e}}$. Also, by $\langle \partial^{\infty}(P) \rangle$, we denote the vector space over $\mathbb{F}$ spanned by $\partial_{\mathbf{x}^{\mathbf{e}}}(P)$ for all $\mathbf{e} \in \mathbb{N}^n$. This is the **partial derivative space** of f. Similarly, $\langle \partial^{=k}(P) \rangle$ denotes the the vector space over $\mathbb{F}$ spanned by $\partial_{\mathbf{x}^{\mathbf{e}}}(P)$ for all $\mathbf{e} \in \mathbb{N}^n$ such that $\deg(\mathbf{x}^{\mathbf{e}}) = k$.

Here, in the thesis, we look at polynomials over an algebra $\mathbb{A}$, An algebra is a vector space V over a certain field $\mathbb{F}$ equipped with a vector product, i.e., a binary operation from $V \times V$ to V . The product is associative (we will talk consider associative algebras) and distributive with the + operation of the vector space and compatible with the scalars from the underlying field. For two elements v_1, v_2 in algebra $\mathbb{A}$, $v_1 v_2$ denotes this vector product. The dimension of an algebra is the dimension of the underlying vector space. When this vector product is simply a coordinate wise product, then the resulting algebra is called the Hadamard algebra.

Let $\mathbb{A}_k$ denote the k-dimensional algebra over the field $\mathbb{F}$. For any 2 elements $\mathbf{a}, \mathbf{b} \in \mathbb{A}_k$, we denote $a \cdot b = \sum_{i=1}^{n} a_i b_i$.

We translate the same definitions from above here. For a polynomial P, we define the **support** of P, denoted by $\mathrm{supp}(f)$ to be the set of monomials m such that $\mathrm{coeff}_m(P) \neq \mathbf{0}$ where $\mathbf{0} = (0, ..., 0) \in \mathbb{N}^k$. The coefficient space of P is the subspace of $\mathbb{F}^k$ generated by the coefficients of P, and it is denoted by $\mathrm{sp}(P)$. For a set of monomials B, we say B is a basis of P, if the coefficients of all the monomials in B form a basis for $\mathrm{sp}(P)$ over $\mathbb{F}$

Definition 5 (Monomial ordering)**.** *A* ***monomial ordering*** $\prec$ *is a total order on the set of all monomial over* $\boldsymbol{x}$ *such that*

- *for all* $\boldsymbol{a} \in \mathbb{N}^n \setminus \boldsymbol{0} = (0, ..., 0)$, $1 \prec \boldsymbol{x^a}$
- *for all* $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{N}^n$, *if* $\boldsymbol{x^a} \prec \boldsymbol{x^b}$, *then* $\boldsymbol{x^{a+c}} \prec \boldsymbol{x^{b+c}}$

For a non-zero polynomial f, the leading monomial (with respect to a monomial ordering) is the largest monomial in the support of f.

2.2 Models of computation

2.2.1 Depth-3 Diagonal circuits

Depth-3 diagonal circuits compute polynomials of form

$$C(\mathbf{x}) = \sum_{i=1}^{s} l_i^{d_i}$$

where l_i's are the linear polynomials over the underlying field $\mathbb{F}$. We denote the class of depth-3 diagonal circuits by $\sum \bigwedge \sum$. For all $i \in [k]$, let f_i be the degree 1 part of l_i. Then the rank of a depth-3 diagonal circuit, denoted by $\mathrm{rk}(C)$, is the dimension of the subspace (over $\mathbb{F}$) generated by f_i's. Note that the rank of C can be equal or one less than the dimension of the subspace generated by l_i's. By log-variate depth-3 diagonal circuit we mean the class of depth-3 diagonal circuits where the number variables is at most logarithmic with respect to the circuit size.

2.3 Polynomial Identity Testing

Lemma 2. *[DL78] [Zip79] [Sch80] Let $f \in \mathbb{F}[x_1, ..., x_n]$ be a non-zero polynomial of degree $d \geq 0$. Let S be any finite subset of $\mathbb{F}$ and let $a_1, ..., a_n$ be sampled from S independently and uniformly at random. Then*

$$Pr_{a_1,...,a_n \in_r S}[f(a_1, ..., a_n) = 0] \leq \frac{d}{|S|}$$

Corollary 1. *For an n-variate polynomial $f \in \mathbb{F}[\boldsymbol{x}]$ of degree d, there exists a hitting set of size $(d+1)^n$ and computable in poly(d^n) time.*

Proof. Let $f \in \mathbb{F}[\mathbf{x}]$ be an n-variate polynomial of degree d. Let $S \subseteq \mathbb{F}$, such that $|S| = d+1$. If $|\mathbb{F}| < d+1$, we go to an extension $\mathbb{K}$ of $\mathbb{F}$ such that $|\mathbb{K}| \geq d+1$ and then, pick $S \subseteq \mathbb{K}$. Then according to lemma 2,

$$\mathbf{Pr}_{\mathbf{a} \in_r S^n}\Big[f(\mathbf{a}) = 0\Big] < 1$$

. Hence, there exists $\mathbf{a} \in S^n$, such that $f(\mathbf{a}) \neq 0$. □

Theorem 1 (Randomized black-box PIT algorithm)**.** *Let $\mathbb{F}$ be a field of size $\geq 2d$. Let $C_{s,d}$ be the set of algebraic circuits of size $\leq s$ computing polynomials in $\mathbb{F}[x_1, ..., x_n]$ of degree $\leq d$. There is a randomized black-box PIT algorithm for $\mathcal{C}$ running in poly(s, n, d) time, so that for every $C \in \mathcal{C}$, if $f_C = 0$, then the algorithm returns True else the algorithm returns False with probability $\geq \frac{1}{2}$*

Remark 1. *Suppose $f \in \mathbb{F}[\boldsymbol{x}]$ can be computed by a size $\leq s$ algebraic circuit C. Given C, for all $\alpha \in \mathbb{F}^n$, $f(\alpha)$ can be computed in poly(s, n) time in the unit-cost model over $\mathbb{F}$.*

Definition 6 (Hitting sets)**.** *Let $F \subseteq \mathbb{F}[x_1, ..., x_n]$ be a family of polynomials. $\mathcal{H} \subseteq \mathbb{F}^n$ is a **hitting-set** for all $f \in F$, $f \equiv 0$ if and only if $f(\alpha) = 0$ for all $\alpha \in \mathbb{F}^n$*

Now we show that the computational problem of finding the hitting set for a certain family of polynomials is equivalent to giving a deterministic black-box PIT algorithm. For a detailed proof, please refer to [For14a].

Lemma 3. *Let $\mathcal{C}$ be a subset of algebraic circuits of size s and computing polynomials of degree $\leq d$. If there exists a $t(s, n, d)$-explicit hitting set $\mathcal{H}$ for $\mathcal{C}$, then there is a deterministic poly$(|\mathcal{H}|, t(s, n, d), s, n, d)$-time black-box PIT algorithm for $\mathcal{C}$.*

Lemma 4. *Let $\mathcal{C} \subseteq \mathcal{C}_{s,d}$. If there is a $t(s, n, d)$-time deterministic black-box PIT algorithm $\mathcal{A}$ for $\mathcal{C}$, then there is a poly$(t(s, n, d), s, n, d)$-explicit hitting set $\mathcal{H}$ for $\mathcal{C}$ with $|\mathcal{H}| \leq t(s, n, d)$*

In the following lemma we describe a standard technique to design a hitting set for the set of polynomials having a "low-support" monomial in their support. It has been used in almost all PIT results regarding the construction of hitting sets for various restricted classes of circuits.

Lemma 5. *Let $\mathcal{P}$ be the set of n-variate degree d polynomials such that every non-zero polynomial in $\mathcal{P}$ has a l-support monomial with non-zero coefficient. Then there exists a hitting set for $\mathcal{P}$ that is computable in time $(nd)^{O(l)}$.*

Proof. Let $x = \{x_1, ..., x_n\}$ be the set of variables over which P is defined. For every l-size subset S of $[n]$, let ϕ_S be the projective of the a monomial $\mathbf{x}^{\mathbf{e}}$ to $\mathbf{x}_S^{\mathbf{e}_S}$. Now since, for every polynomial $P \in \mathcal{P}$, there exists an l-size subset $S \subseteq [n]$ such that $P(\phi_S(\mathbf{x})) \neq 0$. Hence it becomes a l-variate degree d monomial. So the trivial hitting set is of size $d^{O(l)}$. Since for given a P we do not know the l-size subset S of $[n]$ for which it is non-zero we have to try all possible subsets. This gives a $(nd)^{O(l)}$ hitting set. □

Lemma 6. *(Efficient Kronecker Map) Let M be the set of all monomials in $\boldsymbol{x} = \{x_1, ..., x_n\}$ such that for all $m \in M$, ideg$(m) \leq d$. For any value s, there is a polynomial-time constructible set of weight functions $\{w_i\}_{i\in[N]}$ such that $N := ns\log(d+1)$ where $w_i : x \to [2N \log N]$, such that for any $A \subseteq M \times M$, $|A| = s$, there exists i such that for all $(m, m') \in A$, $w_i(m) \neq w_i(m')$.*

We include a short proof of this lemma for completeness.

Proof. Let $M_{n,d}$ be the set of all monomials on n variables and degree $\leq d$. Now, we want to find $W : \mathbf{x} \to \mathbb{N}$ such that $w(m) \neq w(m')$ for $m \neq m' \in M_{n,d}$. One trivial approach towards this is to use the Kronecker map $W : x_i \to (d+1)^{i-1}$. This will give distinct weights to each monomial in $M_{n,d}$ but the weights given by W are exponentially high.
So, we take the weight function W modulo p for many small primes p. Each prime p leads to a different weight function. That is our set of candidate weight functions. We need to bound

the number N of primes which ensures that at least one of the weight functions separates all the monomial pairs in A. We choose the smallest N primes, say P is the set. By the effective version of the Prime Number Theorem, the highest value in the set P is less than $2N \log N$. □

2.4 Results for ROABPs

We include a few theorems which are necessary for better understanding of ROABPs.

2.4.1 Structural Results

Lemma 7. *(For proof refer to [For14b]) Let $f \in \mathbb{F}[\boldsymbol{x}]$ be a polynomial computed by a $t(n,w)$-explicit ROABP of width $\leq w$ in variable order $x_1 \leq ... \leq x_n$ so that $f = (\prod_{i=1}^{n} M_i(x_i))_{1,1}$, for matrices $M_i \in \mathbb{F}[x_i]^{w \times w}$. Then there are matrices $M_i' \in \mathbb{F}[x_i]^{w \times w}$ with $deg_{x_i}(M_i') \leq deg_{x_i}(f)$ with $f = (\prod_{i=1}^{n} M_i'(x_i))_{1,1}$ such that f can be computed by a poly$(t(n,w), n, w, ideg(f))$-explicit ROABP in variable order $x_1 < ... < x_n$ of individual degree $\leq$ $ideg(f)$ and width w.*

Corollary 2. *If polynomials $f, g \in \mathbb{F}[\boldsymbol{x}]$ are computed by an $ideg < d$ ROABPs in a variable order $\pi : [n] \to [n]$, where f is computed by a width $\leq w$ ROABP that is $t(n,w,d)$ explicit and g is computed by an ROABP of width $\leq s$ that is $r(n,s,d)$ explicit, the for all $a, b \in \mathbb{F}$, $af+bg$ can be computed by an $ideg < d$, width $(w+s)$, poly$(t(n,w,d), r(n,s,d), n, w, s, d)$-explicit ROABP in variable order π.*

The following lemma gives a trivial upper bound on the ROABP width with respect to the degree of the polynomial.

Lemma 8 (For proof refer to [For14b])**.** *Let $f \in \mathbb{F}[\boldsymbol{x}_n]$ be a polynomial with sparsity$(f) \leq s$. Then for any permutation $\pi : [n] \to [n]$, f can be computed by an ROABP of width s in variable order π. In particular, any f with $ideg(f) < d$, can be computed in width d^n.*

Lemma 9. *[Nis91] Let $A(x)$ be a polynomial of individual degree d, computed by an ROABP of width w with variable order $(x_1, ..., x_n)$. Let $k \leq n$ and $y = \{x_1, ..., x_k\}$. Then*

$$dim_{\mathbb{F}}\{A_{\boldsymbol{y},\boldsymbol{a}} | \boldsymbol{a} \in \{0, 1, ..., d\}^k\} \leq w$$

Proof. We know by the matrix product structure of ROABPs, $A(\mathbf{x}) = D_1(x_1)...D_n(x_n)$ such that $D_1 \in \mathbb{F}^{1\times w}[x_1]$, $D_n \in \mathbb{F}^{w\times 1}[x_n]$ and for all $2 \leq i \leq n-1$, $D_i \in \mathbb{F}^{w\times w}[x_i]$. Let $\mathbf{z} = \{x_{k+1}, ..., x_n\}$ be the remaining variables of $\mathbf{x}$. We define $P(\mathbf{y}) = D_1...D_k$ and $Q(\mathbf{z}) = D_{k+1}...D_n$. Then

$$P = [P_1(\mathbf{y}), ..., P_w(\mathbf{y})] \in \mathbb{F}^{1\times w}[\mathbf{y}]$$

and

$$Q = [Q_1(\mathbf{z}), ..., Q_w(\mathbf{z})]^T \mathbb{F}^{w\times 1}[\mathbf{z}]$$

and we have $A(\mathbf{x}) = P(\mathbf{y})Q(\mathbf{z})$.

Now, we know when $A(\mathbf{x}) = \prod_{i=1}^{n} D_i(x_i)$, then

$$\text{coeff}_{\mathbf{x}^\mathbf{a}}(A) = \prod_{i=1}^{n} \text{coeff}_{x_i^{a_i}}(D_i) \tag{2.1}$$

Then for $a \in \{0, 1, ..., d\}^k$, the coefficient $A_{(\mathbf{y},\mathbf{a})} \in \mathbb{F}[\mathbf{z}]$ of $\mathbf{y}^\mathbf{a}$ can be given as

$$A_{(\mathbf{y},\mathbf{a})} = \sum_{i=1}^{w} \text{coeff}_{\mathbf{y}^\mathbf{a}}(P_i)Q_i(\mathbf{z})$$

Hence, every $A_{(\mathbf{y},\mathbf{a})} \in \text{span}_{\mathbb{F}}\{Q_1, .., Q_w\}$. Hence the claim follows. □

Here, we have that y is a prefix of x. But note that this is not necessary for the construction to work. The variables in y can be arbitrarily distributed in x. We summarize this observation in the following corollary.

Corollary 3. *Let $A(x)$ be a polynomial of individual degree d, computed by an ROABP of width w , $k \leq n$ and $y = \{x_{i_1}, ..., x_{i_k}\}$. Then the polynomial $\{A_{\boldsymbol{y},\boldsymbol{a}}$ for all $\boldsymbol{a} \in \{0, 1, ..., d\}^k$ can be computed by an ROABP of width w in the same order inherited from the order of the ROABP.*

Lemma 10. *let* $f \in \mathbb{F}[\boldsymbol{x}, \boldsymbol{y}]$ *such that*

$$f(\boldsymbol{x}, \boldsymbol{y}) = \prod_i (x_i + y_i)$$

Then dim $(coeff_{\boldsymbol{x}^{\mathbf{b}}}) = dim\ (coeff_{\boldsymbol{y}^{\mathbf{b}}}) = 2^n$

Proof. For $\mathbf{b} \in \{0,1\}^n$, we look at the coefficients of $\mathbf{y^b}$ in f.

$$\begin{aligned} \text{coeff}_{\mathbf{y^b}}(f) &= \text{coeff}_{\mathbf{y^b}}(\prod_i (x_i + y_i)) \\ &= \partial_{\mathbf{y^b}}(f)\Big|_{\mathbf{y=0}} \\ &= \mathbf{x^{1-b}} \end{aligned}$$

Thus every $\text{coeff}_{\mathbf{y^b}}$ produces a distinct monomial in $\mathbb{F}[\mathbf{x}]$ and hence, these are linearly independent. So dim $(\text{coeff}_{\mathbf{y^b}}(f)) \geq 2^n$

The other direction follows from the fact that f is a multilinear polynomial and hence the coefficient space can have size at most 2^n.

Hence, we get that dim $(\text{coeff}_{\mathbf{y^b}}) = 2^n$. A similar argument works for dim $(\text{coeff}_{\mathbf{x^b}})$. □

Corollary 4. *let* $f \in \mathbb{F}[\boldsymbol{x}, \boldsymbol{y}]$ *such that*

$$f(\boldsymbol{x}, \boldsymbol{y}) = \prod_i (x_i + y_i)$$

- *For all permutations* π *of the variables* $\boldsymbol{x}, \boldsymbol{y}$, f *can be computed by an ROABP of width* $\leq 2^n$
- *There exist permutations* π *such that* $\pi(\boldsymbol{x}) < \pi(\boldsymbol{y})$, *so that any ROABP computing* f *in variable order* π *must have width* $\geq 2^n$
- *If we take the variable order* $x_1 < y_1 < ... < x_n < y_n$, *there is a poly*$(n)$ *explicit ROABP for computing* f *that has width* 2.

2.4.2 Standard results required for width calculation

Lemma 11. *(Monomial ordering)[For14b] Let $\prec$ be a monomial ordering on $\mathbb{F}[\boldsymbol{x}]$ and $S \subseteq \mathbb{F}[\boldsymbol{x}]$. Then dim$(S) = |LM(f) : f \in span(S)|$*

Proof. We first prove the $\geq$ direction.

Let $r \coloneqq |\mathrm{LM}(f) : f \in \mathrm{span}(S)|$. Let the polynomials in span(S) that have distinct leading monomials be $f_1, ..., f_r$ such that $\mathrm{LM}(f_1) \prec ... \prec \mathrm{LM}(f_r)$. We need to show that the f_i's are linearly independent. We take a non-trivial linear combination $\sum_i a_i f_i$. Let i_0 to be the largest $i \in [r]$ such that $a_i \neq 0$. We claim that $\mathrm{coeff}_{\mathrm{LM}(f_{i_0})}(\sum_i a_i f_i) \neq 0$.

Now $a_i = 0$ for all $i > i_0$. Also $\mathrm{LM}(f_i) \prec \mathrm{LM}(f_{i_0})$ for all $i < i_0$ so they do not contribute to $\mathrm{coeff}_{\mathrm{LM}(f_{i_0})}(\sum_i a_i f_i)$. It follows then that only the polynomial f_{i_0} contributes to this coefficient, so that

$$\mathrm{coeff}_{\mathrm{LM}(f_{i_0})}(\sum_i a_i f_i) = a_{i_0}\mathrm{coeff}_{\mathrm{LM}(f_{i_0})}(f_{i_0}) \neq 0$$

Hence, $\sum_i a_i f_i \neq 0$ and hence the f_is are linearly independent.

We now prove the $\leq$ direction. Let us consider f such that $f \notin \mathrm{span}\{f_i\}$ and then we consider the set $\{\mathrm{LM}(f - g) | g \in \mathrm{span}(f_i)\}$. Let $\mathbf{a}$ such that $\mathbf{x}^{\mathbf{a}}$ be a minimal monomial with respect to $\prec$ in this set. As $f \notin \mathrm{span}(f_i)$, this is a non-zero monomial. Now we claim $\mathbf{x}^{\mathbf{a}} \neq \mathrm{LM}(f_i)$ for any i. This is because if $\mathbf{x}^{\mathbf{a}} = \mathrm{LM}(f - g) = \mathrm{LM}(f_i)$, then we can write $f - g = c\mathbf{x}^{\mathbf{a}} + \sum_{\mathbf{b} \succ \mathbf{a}} c_{\mathbf{b}} \mathbf{x}^{\mathbf{b}}$ and $f_i = c'\mathbf{x}^{\mathbf{a}} + \sum_{\mathbf{b} \succ \mathbf{a}} c'_{\mathbf{b}} \mathbf{x}^{\mathbf{b}}$. From that we can get

$$f - \left(g + \frac{c}{c'} f_i\right) = \left(c - \frac{c}{c'}.c'\right)\mathbf{x}^{\mathbf{a}} + \sum_{\mathbf{b} \succ \mathbf{a}} \left(c_{\mathbf{b}} - \frac{c}{c'}.c'_{\mathbf{b}}\right)\mathbf{x}^{\mathbf{b}}$$

Now $g + \frac{c}{c'} f_i \in \mathrm{span}(f_i)$ and this has a minimal monomial $\succ \mathbf{x}^{\mathbf{a}}$, but this contradicts the minimality of $\mathbf{a}$.

Hence, there exists, $g \in \mathrm{span}(f_i) \subseteq \mathrm{span}(S)$, such that $\mathrm{LM}(f - g) \notin \{\mathrm{LM}(f_i)\}$. Thus $f - g \notin \mathrm{span}(S)$. As $f \in \mathrm{span}(S)$ implies $f - g \in \mathrm{span}(S)$, we have $f \notin \mathrm{span}(S)$ as desired. □

Corollary 5. *Let $f \in \mathbb{F}[\boldsymbol{x}]$ be a polynomial and let $\prec$ be any monomial ordering on $\mathbb{F}[\boldsymbol{x}]$. Then $dim(\partial^{<\infty}(f)) \geq dim(\partial^{<\infty}(LM(f)))$*

Lemma 12. *Let $\mathbb{F}$ be a field, $char(\mathbb{F}) = 0$ and $char(\mathbb{F}) > d$.Let $\mathcal{P}_{n,d,k} \subseteq \mathbb{F}[\boldsymbol{x}_n]$ such that for all $P \in \mathcal{P}$, $deg(P) \leq d$ and $dim(\partial^{<\infty}(f)) \leq k$. Then every $0 \neq P \in \mathcal{P}_{n,d,k}$ has a monomial $\boldsymbol{x}^{\boldsymbol{a}}$ such that $cs(\boldsymbol{x}^{\boldsymbol{a}}) \leq k$*

Proof. We first show that

$$|\{\mathrm{LM}(f)|f \in \mathrm{span}\{\partial^{<\infty}(P)\} \geq |\mathrm{cone}(\mathrm{LM}(P))|$$

Let $\mathbf{x}^{\mathbf{e}}$ be the leading monomial of P with respect to the monomial ordering $\prec$, Then P can be written as

$$P = c_{\mathbf{e}}\mathbf{x}^{\mathbf{e}} + \sum_{\mathbf{h} \in \mathrm{supp}(P)\setminus\mathbf{e}} c_{\mathbf{h}}\mathbf{x}^{\mathbf{h}}$$

Let $\mathbf{x}^{\mathbf{f}}$ belong to $\mathrm{cone}(\mathbf{x}^{\mathbf{e}})$. We define $\mathbf{g} \coloneqq \mathbf{e} - \mathbf{f}$. Then we get

$$\partial_{\mathbf{x}^{\mathbf{g}}}(P) = c'_{\mathbf{e}}\mathbf{x}^{\mathbf{e}-\mathbf{g}} + \sum_{\mathbf{h} \in \mathrm{supp}(P)\setminus\mathbf{e}} c'_{\mathbf{h}}\mathbf{x}^{\mathbf{h}-\mathbf{g}}$$

where

$$c'_{\mathbf{e}} = c_{\mathbf{e}} \prod_{i=1}^{n} \frac{e_i!}{(e_i - g_i)!}$$

Now due to the lexicographic monomial ordering, we get that $\mathbf{x}^{\mathbf{h}-\mathbf{g}} \prec \mathbf{x}^{\mathbf{e}-\mathbf{g}}$.

Hence, $\mathrm{LM}(\partial_{\mathbf{x}^{\mathbf{g}}}(P)) = \mathbf{x}^{\mathbf{f}}$.

Hence, we get that for every monomial $\mathbf{x}^{\mathbf{f}} \in \mathrm{cone}(\mathbf{x}^{\mathbf{e}})$, there exists polynomial $h \coloneqq \partial_{\mathbf{x}^{\mathbf{g}}}(P) \in \mathrm{span}\{\partial^{<\infty}(P)\}$ such that $\mathrm{LM}(h) = \mathbf{x}^{\mathbf{f}}$. This gives us that

$$\left|\left\{\mathrm{LM}(f)\middle| f \in \mathrm{span}\{\partial^{<\infty}(P)\right\}\right| \geq \left|\mathrm{cone}(\mathrm{LM}(P))\right|$$

We also know from the previous lemma that with respect to monomial ordering $\prec$,

$$\dim(\partial^{<\infty}(P)) \geq |\{LM(f)|f \in \langle\partial^{<\infty}(P)\rangle\}$$

This gives us the proof of our lemma. □

2.5 Low cone monomials are few

We prove the following lemma that we will use later.

Lemma 13. *For $0 < k \leq n$, $k \in \mathbb{Z}$,*

$$\sum_{i=0}^{k} \binom{n}{i} \leq \left(\frac{en}{k}\right)^k$$

Proof. For $0 < t \leq 1$

$$\sum_{i=0}^{k} \binom{n}{i} \leq \frac{1}{t^k} \sum_{i=0}^{k} \binom{n}{i} t^i \leq \frac{(1+t)^n}{t^k}$$

Since $(1+t) < e^t$ for all $t \neq 0$, from the above expression, we get

$$\sum_{i=0}^{k} \binom{n}{i} \leq \frac{e^{tn}}{t^k}$$

We now put $k = tn$ and this gives us the given inequality. □

Now, we show that the number of low-cone monomials are few.

Lemma 14. *[Gho19] The number of n-variate monomials with cone-size $\leq k$ is $O(rk^2)$ where $r = \left(\frac{3n}{\log k}\right)^{\log k}$*

Proof. First we prove that for any fixed support set S_k, the number of cone-size $\leq k$ monomials is $< k^2$. Then, we can multiply by the total number of support sets to get an upper bound.

Let $T(k, l)$ denote the number of cone-size monomials with support S of size l. The exponent of each $x_i \in S$ is at least 1 and $\leq k - 1$. This gives us the following recurrence

$$T(k, l) \leq \sum_{i=2}^{k} T(\frac{k}{i}, l-1)$$

We claim that $T(k, l) < k^2$. We prove this by induction on k. Using the previous equation, for all $t \leq k$, we get

$$T(k, l) < \sum_{i=2}^{k} \left(\frac{k}{i}\right)^2 < k^2 \sum_{i=2}^{k} \left(\frac{1}{i-1} - \frac{1}{i}\right) < k^2$$

From the definition of cone, a cone-size $\leq k$ must have support size $\leq l := \lfloor \log k \rfloor$. The number of possible support sets, is $\sum_{i=0}^{l} \binom{n}{i}$. Using lemma 13, $\sum_{i=0}^{l} \binom{n}{i} \leq \left(\frac{3n}{l}\right)^l$. □

Here we include a separate calculation for the bivariate case. We will use this later.

Lemma 15. *The number of bivariate cone-size $\leq k$ monomials is $\leq k \ln k$*

Proof. Let the variables be $\mathbf{x} = \{x_1, x_2\}$ We first count the number of monomials m with cs $\leq k$ such that $\deg_{x_1}(m) = i$. We denote that set by M_i. Then

$$|M_i| = \left\lfloor \frac{k}{i} \right\rfloor$$

Let M be the set of monomials with cs $\leq k$. Then

$$|M| = \sum_{i=1}^{k} |M_i| = \sum_{i=1}^{k} \left\lfloor \frac{k}{i} \right\rfloor \leq \sum_{i=1}^{k} \frac{k}{i} \leq k \sum_{i=1}^{k} \frac{1}{i} \leq k \ln k$$

□

2.6 Sum of log-variate ROABPs subsumes DD3

First we include small proofs to the results required for this proof.

In this next theorem, we will show that the polynomials in the depth-3 diagonal model has low dimensional partial derivative space.

Theorem 2. *[For14b] Let $f(\boldsymbol{x}) = \sum_{i=1}^{s} c_i (f_i(\boldsymbol{x}))^{d_i}$ where $f_i(\boldsymbol{x}) = a_0 + \sum_{j=1}^{n} a_{i,j} x_j$, $a_{i,j} \in \mathbb{F}$. Then $dim(\partial^{<\infty}(f))) \leq s(d+1)$ where $d = \max_i d_i$*

Proof. Let $\mathbf{x}^{\mathbf{e}}$ be such that $|\mathbf{e}|_1 = \sum_{i=1}^{n} e_i = b$. Then

$$\partial_{\mathbf{x}^{\mathbf{e}}}(f_i^{d_i}) = \binom{b}{e_1, ..., e_n} (a_0 + \sum_{j=1}^{n} a_{i,j} x_j)^{d_i - b} \prod_{j=1}^{n} a_{i,j}$$

Now for all $b \leq d_i$,

$$\dim(\partial^{=b}(f_i^{d_i})) \leq 1$$

and for $b > d_i$, it is zero. Hence, we get

$$\begin{aligned}\dim(\partial^{<\infty}(f)) &\leq \sum_{i=1}^{s} \dim(\partial^{<\infty}(f_i^{d_i})) \\ &\leq \sum_{i=1}^{k} (d_i + 1) \\ &\leq k(d+1)\end{aligned}$$

□

Theorem 3. *[Sax08] (Duality Trick) Let $m, d \in \mathbb{Z}_+$. Let $|\mathbb{F}| \geq d(m-1)$. Then for all distinct $a_0, ..., a_{d(m-1)} \in \mathbb{F}$, there exist $b_{i,j}$ such that*

$$(z_1 + ... + z_m)^d = \sum_{i=0}^{(m-1)d} \sum_{k=0}^{d} b_{i,j} \prod_{j=1}^{m} (z_j + a_i)^k$$

Proof. We define $p(t) \coloneqq \prod_{i=1}^{m}(z_i + t) - t^m$. Then

$$\partial_{t^{m-1}}(p)|_{t=0} = (z_1 + ... + z_m)$$

Let $g = p^d$. Then

$$\partial_{t^{(m-1)d}}(g)|_{t=0} = (z_1 + ... + z_m)^d$$

Now we consider $(d+1)$ distinct points $a_0, ..., a_d \in \mathbb{F}$ and look at the evaluations of g at these points. We know, for any t^i, the coefficient of t^i in $g(t)$ can be extracted by taking linear combinations of evaluations of g at $(i+1)$ many distinct points. Hence, there exists b_i''s such that

$$\begin{aligned}(z_1 + ... + z_m)^d &= \sum_{i=0}^{(m-1)d} b_i' g(a_i) \\ &= \sum_{i=0}^{(m-1)d} b_i' (\prod_{j=1}^{m} (z_j + a_i) - a_i^m)^d \\ &= \sum_{i=0}^{(m-1)d} \sum_{k=0}^{d} b_{i,k} (\prod_{j=1}^{m} (z_j + a_i))^k\end{aligned}$$

where $b_{i,k} = \binom{d}{k} b_i' (-a_i^m)^{d-k}$ □

Next, we show a variable reduction map from n to $O(\log n)$ variables preserving non-zeroness of polynomials.

Definition 7. *[Vai15] Define the map $\psi : \mathbb{F}[\boldsymbol{x}] \to \mathbb{F}[\boldsymbol{a}, \boldsymbol{b}, t]$ where $|\boldsymbol{a}| = |\boldsymbol{b}| = l$ as follows:*

$$x_i \to \sum_{j=1}^{l} a_j^i b_j^{i^2} t$$

Theorem 4. *[Vai15] Under the map $\psi : \mathbb{F}[\boldsymbol{x}] \to \mathbb{F}[\boldsymbol{a}, \boldsymbol{b}, t]$, each $\leq l$ support monomial in $P(\boldsymbol{x})$ is mapped to a unique monomial in $P(\psi(\boldsymbol{x}))$*

Proof. We fix an l-support monomial $\mathbf{x}^{\mathbf{r}} = x_{i_1}^{r_1} x_{i_2}^{r_2} ... x_{i_l}^{r_l}$. Now

$$\begin{aligned}\psi(\mathbf{x}^{\mathbf{r}}) &= \prod_{k=1}^{l} \Big(\sum_{j=1}^{l} a_j^{i_k} b_j^{i_k^2} t \Big)^{r_k} \\ &= a_0 \prod_{k=1}^{l} (a_k^{i_k} b_k^{i_k^2})_k^r t^{|r|_1} + \text{other terms}\end{aligned}$$

Now we claim that this monomial can be generated uniquely from the product and hence $a_0 = 1$.

Let us assume we picked the first term $a_1^{i_1} b_1^{i_1^2}$ from k_1 many brackets. And the corresponding exponents of a_1 are $(e_1, ..., e_{k_1})$. This gives us the following relations:

$$\sum_{j=1}^{k_1} e_j = r_1 i_1$$
$$\sum_{j=1}^{k_1} e_j^2 = r_1 i_1^2$$

Now using Cauchy-Schwarz, we get $r_1 \leq k_1$. Equality holds when $e_1 = e_2 = ... = e_{k_1} = i_1$ Similarly, for all j, we get $k_j \leq r_j$. Now $|k|_1 = |r|_1$. Hence, the monomial is generated uniquely. □

Now we include the proof here that the sum of log-variate ROABPs subsumes the diagonal depth-3.

Lemma 16. *[BS20] If we have poly-time black-box PIT for sum of width-1, log-variate (commutative) ROABPs, then we have poly-time black-box PIT for diagonal depth-3 circuits.*

Proof. We showed in Theorem (2) that diagonal depth-3 circuits have 'low' dimension partial derivative space, and that such polynomials have a nonzero log-support monomial. Under the promise of such a log-support monomial, we can apply variable-reduction map from Theorem (4) to get from n to O(log n) variables and we showed that this map preserves non-zeroness.

After applying the log-variate map, we will get to power-of-sums-of univariates form which we can convert to sum-of-products-of-univariates form using the duality-trick of that was shown in Theorem (3) . Moreover, each product-of-univariates has a width-1 ROABP; thus we have represented as sum of width-1 log-variate ROABPs (which are trivially commutative!). □

2.7 Results for Cone-closed Basis

Here we also state an interesting result regarding the structural properties of polynomials over $\mathbb{F}^k$ when they are shifted by a Basis-Isolating Weight Assignment (from [FGS18])

Definition 8 (Cone-closed Basis)**.** *A set of monomials B is called cone-closed set of monomials, if for every monomial in B, all its sub-monomials also belong to B. Let P be an n-variate polynomial over $\mathbb{F}^k$. We say that P has a cone-closed basis if there is a cone-closed set of monomials B whose coefficients in P form a basis for the coefficient space of P.*

Theorem 5. *Let $P(\boldsymbol{x})$ be a n-variate degree d polynomial over $\mathbb{F}^k[\boldsymbol{x}]$ and char$(\mathbb{F}) = 0$ or $> d$. Let $\boldsymbol{w} = (w_1, ..., w_n) \in \mathbb{N}^n$ be a basis isolating weight assignment for $P(\boldsymbol{x})$. Then $P(\boldsymbol{x} + t^{\boldsymbol{w}}) \coloneqq P(x_1 + t^{w_1}, ..., x_n + t^{w_n})$ has a cone-closed basis over $\mathbb{F}^t$*

For a detailed proof of this theorem, refer to Section (7) in [Gho19].

We show that the notion of cone-closed basis subsumes the other two notions of rank concentration, i.e., low-support rank concentration and low-cone rank concentration.

Lemma 17. *Let $P(\boldsymbol{x})$ be a polynomial in $\mathbb{F}^k[\boldsymbol{x}]$. Suppose that $P(\boldsymbol{x})$ has a cone-closed basis. Then, $P(\boldsymbol{x})$ has cone-size $\leq k$ rank concentration and $\log k$-support rank concentration.*

Proof. Let B be a cone-closed set of monomials such that it is a basis of P. Clearly $|B| \leq k$. Since B is cone-closed, for every monomial in B, all its sub-monomials are also in B. Thus, each $m \in B$ has cone-size $\leq k$

Moreover, each $m \in B$ has support size $\leq \log k$. Otherwise, there will be a monomial in B whose cone-size is greater than k. This is not possible, since B is cone-closed. So, P has $\log k$-support rank concentration. □

Chapter Three

Cone-Size Hypothesis

3.1 Introduction

Definitions: Let m be a monomial with coefficient c.

$$m = c \cdot {x_1}^{e_1} {x_2}^{e_2} \cdots {x_n}^{e_n}$$

We denote the exponent vector $(e_1, e_2, \cdots, e_n)$ as $\mathbf{e}$. A monomial $m_{\mathbf{a}}$ belongs in the cone of another monomial $m_{\mathbf{b}}$ if $m_{\mathbf{a}}$ divides $m_{\mathbf{b}}$, equivalently $\mathbf{a} \leq \mathbf{b}$. We define **cone** for a monomial or equivalently it's exponent vector $\mathbf{e}$ as:

$$cone(\mathbf{e}) = \{\, \mathbf{f} \in \mathbb{Z}^n \mid \mathbf{0} \leq \mathbf{f} \leq \mathbf{e} \,\}$$

For example $cone(x^d) = \{\, 1, x, x^2, \cdots, x^d \,\}$, and $cone(x^2y) = \{\, 1, x, x^2, y, xy, x^2y \,\}$. And now we define **cone size** of a monomial as simply the the number of monomials which divide it, that is the number of monomials in it's cone.

$$cs(\mathbf{e}) = |cone(\mathbf{e})| = \prod_{i \in [n]} (e_i + 1)$$

We define **cone of a polynomial** $P(\mathbf{x})$ (denoted $cs(P)$) as the minimal cone-size over all the monomials occurring in P.

$$cone(P) = min\{\, cone(\mathbf{e}) \quad | \quad \text{coeff}_{\mathbf{x^e}}(P) \neq 0\}$$

Note that a polynomial may have multiple minimal cone monomials with the same minimum cone size.

Coefficients of a shifted polynomial: Let us shift a general polynomial randomly, and try to find relation among its coefficients. Let $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a non-zero polynomial with individual degree d. Let $f'(\mathbf{x}+\mathbf{t})$ be the shifted polynomial where $x_i \to x_i + t_i$. Assume $ch(\mathbb{F}) = 0$.

$$\text{Let } f(\mathbf{x}) = \sum_{\mathbf{0} \le \mathbf{e} \le \mathbf{d}} z_{\mathbf{e}} \cdot \mathbf{x}^{\mathbf{e}}, \text{ where } z_{\mathbf{e}} \text{ are coefficients } \in \mathbb{F}$$

$$f'(\mathbf{x}+\mathbf{t}) = \sum_{0 \le \mathbf{e} \le \mathbf{d}} z'_{\mathbf{e}} \cdot \mathbf{x}^{\mathbf{e}}, \text{ where } z'_{\mathbf{e}} \in \mathbb{F}[\mathbf{t}]$$

Now by using the binomial theorem and collecting the coefficients, we can show that

$$z'_{\mathbf{e}} = \sum_{\mathbf{e} \le \mathbf{f} \le \mathbf{d}} \binom{\mathbf{f}!}{\mathbf{e}!} \cdot z_{\mathbf{f}} \cdot \mathbf{t}^{\mathbf{f}-\mathbf{e}} \tag{3.1}$$

Equation (3.1) also has an alternate viewpoint using the generalized Taylor series expansion of $f(\mathbf{x})$ around the point $\mathbf{t}$.

$$f(\mathbf{x}) = \sum_{\mathbf{e}} (1/\mathbf{e}!) \cdot \partial_{\mathbf{x}^{\mathbf{e}}} f(\mathbf{t}) \cdot (\mathbf{x}-\mathbf{t})^{\mathbf{e}}$$

$$f(\mathbf{x}+\mathbf{t}) = \sum_{\mathbf{e}} (1/\mathbf{e}!) \cdot \partial_{\mathbf{x}^{\mathbf{e}}} f(\mathbf{t}) \cdot \mathbf{x}^{\mathbf{e}}$$

This immediately gives the desired expression as:

$$\begin{aligned} z'_{\mathbf{e}} &= \frac{1}{\mathbf{e}!} \cdot \partial_{\mathbf{x}^{\mathbf{e}}} f(\mathbf{t}) \\ &= \sum_{\mathbf{e} \le \mathbf{f} \le \mathbf{d}} \binom{\mathbf{f}!}{\mathbf{e}!} \cdot z_{\mathbf{f}} \cdot \mathbf{t}^{\mathbf{f}-\mathbf{e}} \end{aligned}$$

3.2 Cone size Hypothesis

Let $f(\mathbf{x}) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ be shifted by a single variable t, $x_i \to x_i + t$ to get $f' \in \mathbb{F}[t][\mathbf{x}]$.

Cone size$\le k$ Hypothesis: All cone size $\le k$ monomials in $f' = f(\mathbf{x}+t)$ have coefficients

$= 0$.

Note that the coefficients of f' are univariate polynomials in t, and the hypothesis demands that $\forall \mathbf{e}$ such that $cs(\mathbf{e}) \leq k$, $\text{coeff}_{\mathbf{x}^\mathbf{e}}(f') = 0$. As, we saw in Equation (3.1), the cone size hypothesis can be re-framed as :

$$\partial_{\mathbf{x}^\mathbf{e}} f(t, \ldots, t) = 0 \text{ for all } \mathbf{e}.\ \text{cs}(\mathbf{e}) \leq k$$

What can we say about a polynomial f which satisfies $cs \leq k$ hypothesis? We wish to exactly characterize the structure of f. We start with the simple case of bivariate polynomials, where we derive a very nice structure for f. And to motivate we will show how that structure gives us a faster hitting set for bivariate polynomials computed by an ROABP.

3.2.1 Bivariate Case

We prove the following structure in the case when number of variables $n = 2$. Let $\mathcal{I} = \langle x - y \rangle_{\mathbb{F}[x,y]}$ be a principal ideal of $\mathbb{F}[x, y]$. Note that $\mathcal{I}^k = \left\langle (x-y)^k \right\rangle_{\mathbb{F}[x,y]}$. Then:

Lemma 18. *If $f \in \mathbb{F}[x, y]$ satisfies $cs \leq k$ hypothesis, then $f \in \mathcal{I}^k$.*

Proof. The proof is by induction on k.

Base Case: $k = 1$. Since the zeroth partial derivative of f is f itself, f satisfies $cs \leq 0$ hypothesis simply means $f(t, t) = 0$, which by Factor Theorem implies that $(x-y)|f \Rightarrow f \in \mathcal{I}$.

Inductive case: Suppose $f \in \mathcal{I}^k$, and f satisfies $cs \leq k + 1$ hypothesis. This means

$f = g(x,y) \cdot (x-y)^k$, and $\partial_{x^k} f(t,t) = 0$, respectively.

$$\begin{aligned}
\partial_{x^k} f(t,t) &= 0 \\
\Big(k! \cdot g + (x-y)^k \cdot \partial_{x^k} g\Big)(t,t) &= 0 \\
g(t,t) &= 0 \\
g \in \mathcal{I} \Rightarrow g &= h \cdot (x-y), \text{ where h } \in \mathbb{F}[x,y] \\
f = g \cdot (x-y)^k &= h \cdot (x-y)^{k+1} \\
f &\in \mathcal{I}^{k+1}
\end{aligned}$$

□

Remark 2. *For $n \geq 3$, the ideal will look like $\mathcal{I} = \langle (x_1 - x_2), (x_1 - x_3), \ldots, (x_1 - x_n) \rangle$. In that case, if f satisfies cs $\leq k$ hypothesis, then $f \notin \mathcal{I}^k$, since the above proof will break down because $\mathcal{I}$ is not a principal ideal. We will however still find the exact structure of f, as we will see later.*

Now we will show an ROABP width lower bound when $f \in \mathcal{I}^k$.

Theorem 6. *If $f = (x-y)^k \cdot g(x,y)$, then ROABP width $w(f) > k$.*

Proof. Note that $(x-y)^k$ has ROABP width exactly $k+1$ by Nisan's characterization (coeff(y^i) for $i \in [k+1]$ are linearly independent as the coefficients are different degree x polynomials). The idea is that width will only increase when multiplied by g.

Suppose g is a degree d polynomial. Let $g = g_d + g_{d-1} + \ldots + g_0$, where g_i is degree i homogeneous part of g. It suffices to prove that $(x-y)^k g_d$ has width $> k$. This is because $(x-y)^k$ is a homogeneous polynomial of degree k, hence $(x-y)^k g_d$ contributes a $\deg_x = k + d - i$ term in $\text{coeff}_f(y^i)$, while $g_{d-1}, \ldots, g_0$ contribute $\deg_x < k + d - i$ terms in $\text{coeff}_f(y^i)$. Hence they will not affect $\mathbb{F}$ linear independence of coefficients as the degrees are different.

Let $h = (x-y)^k g_d(x,y)$. Let $sp(p)$ denote the sparsity of a univariate polynomial p. To prove $w(h) > k$, it suffices to prove that

$$sp\Big(h(x,1)\Big) = sp\Big((x-1)^k g_d(x,1)\Big) > k$$

This is because $h(x,y)$ is a homogeneous polynomial of degree $d+k$. If $sp(h(x,1)) = l > k$, then

$$h(x,1) = c_1 \cdot x^{j_1} + \ldots + c_l \cdot x^{j_l}$$

with each $c_1, \ldots, c_l \neq 0$. Note that $\text{coeff}_h(x^{j_i})$ has $\deg_y = k + d - j_i$ in $h(x,y)$. Therefore $\text{coeff}_h(x^{j_1}), \ldots,$ $\text{coeff}_h(x^{j_l})$ are linearly independent as they are of different degrees.
Now, we focus on the final step of proving the following claim

Claim 1. $sp\Big((x-1)^k \cdot g_d(x,1)\Big) > k$

Note that $(x-1)^k \cdot g_d(x,1)$ is a non-zero polynomial, since partial evaluation of a homogeneous polynomial is always non-zero. For the sake of contradiction, suppose $sp <= k$. That means:

$$(x-1)^k \cdot g_d(x,1) = \sum_{i=0}^{k-1} c_i \cdot x^{j_i}$$
$$x^k \cdot g_d(x+1,1) = \sum_{i=0}^{k-1} c_i \cdot (x+1)^{j_i}$$

On LHS, all degree $< k$ terms are 0. This sets up a homogeneous system of linear equations.

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ j_0 & j_1 & \cdots & j_{k-1} \\ \binom{j_0}{2} & \binom{j_1}{2} & \cdots & \binom{j_{k-1}}{2} \\ & & \ddots & \\ \binom{j_0}{k-1} & \binom{j_1}{k-1} & \cdots & \binom{j_{k-1}}{k-1} \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{k-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

This binomial coefficient matrix is full rank, a simple proof of which we include later for completeness. Hence this system only has a trivial solution which contradicts the non-

zeroness of $(x-1)^k \cdot g_d(x,1)$. Hence, $sp > k$, which suffices to prove width of $f = (x-y)^k \cdot g(x,y) > k$. □

Lemma 19. *[Gur+15] Let C be a $w \times w$ matrix with $C(a,b) = \binom{j_a}{b-1}$ for all $a, b \in [w]$ where $\{j_a\}_a$ are all distinct numbers. Then C has full rank*

Proof. We will show that for any $\alpha \neq \mathbf{0} \in \mathbb{F}^{w\times 1}$, $C\alpha \neq 0$. We consider the polynomial $h(y) = \sum_{b=1}^{w} \alpha_b \frac{y(y-1)\ldots(y-b+2)}{(b-1)!}$. As, $h(y)$ is a polynomial with degree $\leq w-1$, we have at most $w-1$ roots. Then there exists $a \in [w]$, such that $h(j_a) = \sum_{b=1}^{w} \alpha_b \binom{j_a}{b-1} \neq 0$. □

Hitting sets for bivariate ROABPs

Note that Theorem 6 and Lemma 18 together give a hitting set for bivariate ROABPs. Let f be the input bivariate polynomial computed by an ROABP of width k given as black-box. We simply need to shift f with t and test zeroness of the coefficients (univariates in t) of $cs \leq k$ monomials in the shifted polynomial. This is a valid hitting set because if f is a non-zero polynomial, then one of the $cs \leq k$ coefficients must be non-zero. If not, then by Lemma 18 and Theorem 6, f has ROABP width $> k$, which is a contradiction.

Now we compute the size of the hitting set for the ROABPs. The trivial hitting set is $(d+1)^2$ where $d \coloneqq \max\{\text{ideg}(x_1), \text{ideg}(x_2)\}$. Now, for this hitting set we need to check the coefficient of all bivariate monomials that have cs $\leq k$. The number of bivariate monomials with cs $\leq k$ is $k \ln k$. (Lemma 15) Now the coefficient of each such monomial is an univariate polynomial of *degree* $\leq 2d$. So the size of the required hitting set is $2dk \ln k + 1$. This is strictly better than the trivial hitting set for $d >> k$.

3.2.2 Extending to general ROABPs

Theorem 7 (Structural Result)**.** *Let*

$$\mathcal{I} = \Big\langle \prod_{j\geq 2} (x_j - x_1)^{e_{i,j}} | cs(\overline{e}) \leq k \Big\rangle_{\mathbb{F}[x_1]} \tag{3.2}$$

If f satisfies cone-size $\leq k$ hypothesis, then $f \in \mathcal{I}$

Proof. The proof is same as the proof of Lemma (22). Replace the $t_i = t$ and this follows. □

We now want to prove a similar ROABP width lower bound for polynomials satisfying this structural property.

Theorem 8 (Width Result)**.** *If $f \in \mathcal{I}$, then any ROABP computing f has width $\geq k^{\frac{1}{n-1}}$*

We first write the proof for $n = 3$ and then generalize it to higher variables.

Proof. We take $f \in \mathcal{I}$. Now f can be split into degree -i homogeneous parts. Each of these contribute separately to the width because there are no cancellations and taking the partial derivative with respect to them will always give distinct leading monomials. Hence,without loss of generality, we can, assume that f is a degree d homogeneous polynomial.

Hence,

$$f = \sum_{\mathbf{e}:\mathrm{cs}(\mathbf{e})\geq k} a_{\mathbf{e}} x_1^{d-|e|_1} (x_2 - x_1)_2^e (x_3 - x_1)_3^e$$

We now use the following substitution $x_2 = t_2 x_1$ and $x_3 = t_3 x_1$. Hence the new polynomial becomes

$$f = x_1^d \sum_{\mathbf{e}:\mathrm{cs}(\mathbf{e})\geq k} a_{\mathbf{e}} (t_2 - 1)^{e_2} (t_3 - 1)^{e_3}$$

So using this structural characterization, we make the following claim regarding the sparsity of such polynomials

Claim 2. *Let*

$$f' = \sum_{\boldsymbol{e}:cs(\boldsymbol{e})\geq k} a_{\boldsymbol{e}} (t_2 - 1)^{e_2} (t_3 - 1)^{e_3}$$

Then $sp(f) \geq \sqrt{k}$

Now since, $\text{cs}(\mathbf{e}) \geq k$, then either e_2 or e_3 must be $\geq \sqrt{k} - 1$. If not, then $\text{cs}(\mathbf{e}) = (e_2 + 1)(e_3 + 1) < k$.

Then we can write the polynomial into two parts. One where $e_2 \geq \sqrt{k} - 1$ and $e_3 \geq \sqrt{k} - 1$

$$f' = (t_2 - 1)^{\sqrt{k}} \sum_{\mathbf{e}} (t_2 - 1)^{e'_2} (t_3 - 1)^{e_3} + (t_3 - 1)^{\sqrt{k}} \sum_{\mathbf{e}} (t_2 - 1)^{e_2} (t_3 - 1)^{e'_3}$$

There are two cases.

Case 1: If $(t_2 - 1)^{\sqrt{k}} g(x_2, x_3)$ is a summand. (This works even if both of them exist.) We define e' to be the highest power such that $(t_3 - 1)^e | f'$ but $(t_3 - 1)^{e+1} \nmid f'$

$$f'' = \frac{f'}{(t_3 - 1)^e}$$

Now we want to prove that $\text{sp}(f'') \geq \sqrt{k}$

Now, $f''(t_2, 1) \neq 0$. Hence,

$$f''(t_2, 1) = (t_2 - 1)^{\sqrt{k}} \sum_{e} a'_e (t_2 - 1)^3$$

Now, $f''(t_2, 1) = (t_2 - 1)^{\sqrt{k}} g(t_2)$.We use Claim (1) here to prove that $\text{sp}(f''(t_2, 1)) \geq \sqrt{k}$. Now, replacing variables by constants can only decrease the sparsity by adding more cancellations. Hence, $\text{sp}(f'') \geq \sqrt{k}$. Hence, multiplying back by $(t_3 - 1)^e$ will keep sparsity unchanged. Hence $\text{sp}(f') \geq \sqrt{k}$. . Hence, we can write

$$f' = \sum_{i=1}^{w \geq \sqrt{k}} a_i(t_3) t_2^{e_i}$$

Now we can write back

$$f = \sum_{i=1}^{w \geq \sqrt{k}} a_i(t_3) x_1^{d - e_i} x_2^{e_i}$$

Hence, if we take the set of partial derivatives

$$\Delta = \{\partial_{x_1^{d-e_i}}(f)|_{x_1 = 0} \text{ where } i \in [w]\}$$

These partial derivatives are all linearly independent due to different powers of x_2 in them. Hence, we use Lemma (9), to get that $w(f) \geq \sqrt{k}$.

Case 2: If $(t_2 - 1)^{\sqrt{k}} g(x_2, x_3)$ is not a summand. Then we can write

$$f'' = (t_3 - 1)^{\sqrt{k}} \sum_{\mathbf{e}} a_{\mathbf{e}} (t_2 - 1)^{e_2} (t_3 - 1)^{e'_3}$$

We can use the same proof strategy as **Case 1** but now by switching t_2 and t_3. This will again give us that $\text{sp}(f') \geq \sqrt{k}$. And then we can write

$$f = \sum_{i=1}^{w \geq \sqrt{k}} a_i(t_2) x_1^{d-e_i} x_3^{e_i}$$

We take the same set of partial derivatives Δ and these are all linearly independent due to different powers of x_3 in them. Hence, we use Lemma (9), to get that $w(f) \geq \sqrt{k}$. □

Now this proof can be extended to higher variables by a similar observation that there exists $i \in \{2, ..., n\}$ such that $e_i \geq k^{\frac{1}{n-1}}$. The rest of the proof follows in an exactly similar fashion. We get that $w(f) \geq k^{\frac{1}{n-1}}$

3.3 A structural Conjecture for Cone-Size hypothesis

As mentioned in Remark 2, the picture is not so simple and clear for general number of variables n. But, inspired from bivariate and many calculations, we give the following generalized conjecture for the structure of f.

Conjecture 1. *If f satisfies $cs \leq k$ hypothesis, then*

$$f \in \left\langle P_{\boldsymbol{e}} = \prod_{i \neq j} (x_i - x_j)^{e_{ij}} \mid cs(P_{\boldsymbol{e}}) \geq k + 1 \right\rangle_{\mathbb{F}[\boldsymbol{x}]}$$

One can easily verify that Lemma 18 is a special case of above conjecture. The benefit of having this form of f will be justified by the next conjecture:

Conjecture 2. *If $f \in \left\langle P_{\boldsymbol{e}} = \prod_{i \neq j} (x_i - x_j)^{e_{ij}} \mid cs(P_{\boldsymbol{e}}) \geq k + 1 \right\rangle_{\mathbb{F}[\boldsymbol{x}]}$, then ROABP width $w(f) > \sqrt{k}$.*

The idea is to first show that $P_{\mathbf{e}} = \prod_{i \neq j}(x_i - x_j)^{e_{ij}}$, where $cs(P_{\mathbf{e}}) \geq k+1$ has $w(P_{\mathbf{e}}) > k$, then show the same lower bound for $g \cdot P_{\mathbf{e}}$, where g is any n variate polynomial. The final nail in the coffin is to then show the width lower bound for $\sum_i g_i \cdot P_i$, which is expected to be the most difficult part.

Remark 3. *Here one is forced to consider* ***commutative ROABP width****, as there exists the following example:*

$$f = (x_1 - y_1)^{e_1}(x_2 - y_2)^{e_2} \dots (x_n - y_n)^{e_n}$$

has ROABP width $= 1 + max_i\{e_i\}$ *in the variable order* $x_1 < y_1 < \ldots < x_n < y_n$*. However it still has width* $= cs(\boldsymbol{e})$ *in all the variable orders where* $\sigma(\boldsymbol{x}) < \sigma(\boldsymbol{y})$*, that is in the variable order, where all* $\boldsymbol{x}$ *variables must occur before any* $\boldsymbol{y}$ *variables although allowing any permutations within* $\boldsymbol{x}$ *or* $\boldsymbol{y}$*.*

Hitting set overview: Suppose we are given an input polynomial f as a black-box ROABP of size k (width $\leq k$), and number of variables $n = O(\log k)$. We know that number of $cs \leq k$ monomials when $n = O(\log k)$ is $poly(k)$. Hence, in polynomial time we will test zeroness of the coefficients of $cs \leq k$ monomials in $f(x_1+t, \ldots, x_n+t)$, which can be done in polynomial time as they are polynomially many and each coefficient is a univariate polynomial in t. If the input polynomial f was non-zero, then one of these coefficients must be non-zero. Because otherwise, f satisfies $cs \leq k$ hypothesis, which by Conjecture 1 and 2 will contradict the width of input polynomial.

3.3.1 A proof for the trivariate case

Definition 9 (Cone set)**.** *The* ***cone set*** S_m *of a monomial* m *is defined as:*

$$S_m = \{P_{\boldsymbol{e}} \mid cone(P_{\boldsymbol{e}}) = m\}$$

Note that different cone sets may share some polynomials since a polynomial can have multiple least cone monomials. The reason for defining S_m is driven by partial derivatives.

In $\langle P_{=k} \rangle$, we wish to take partial derivative by a monomial $m = \mathbf{x}^{\mathbf{e}}$ of cone-size $= k$. Suppose $\langle S_m \rangle \subseteq \langle P_{=k} \rangle = \sum_{i:P_i \in S_m} a_i P_i$, then:

$$\frac{1}{\mathbf{e!}} \cdot \partial_m \langle P_{=k} \rangle(\mathbf{t}) = \frac{1}{\mathbf{e!}} \cdot \partial_m \langle S_m \rangle(\mathbf{t}) = \sum_{i:P_i \in S_m} a_i(\mathbf{t})$$
$$\frac{1}{\mathbf{e!}} \cdot \partial_m \langle P_{=k} \rangle(\mathbf{t}) = \sum_{i:P_i \in S_m} a_i(\mathbf{t}) = 0$$

The last step follows from $cs \leq k$ hypothesis, where $\mathbf{t}$ simply means evaluating at $x_1 = x_2 = x_3 = t$ point. The first step is true since only $P_i \in S_m$ will contribute in the partial derivatives evaluated at $\mathbf{t}$. In this section, we will show that the cone size conjecture is true when the input polynomial is trivariate. This will require much more non-trivial ideas than the bivariate case proved earlier. The hope is to generalize the techniques here to prove $cs \leq k$ conjecture for general n. Let us formally state this as a lemma:

Lemma 20. *If $f \in \mathbb{F}[x_1, x_2, x_3]$ satisfies $cs \leq k$ hypothesis, then*

$$f \in \left\langle P_{\boldsymbol{e}} = (x_1 - x_2)^{e_{12}}.(x_1 - x_3)^{e_{13}}.(x_2 - x_3)^{e_{23}} \,|\, cs(P_{\boldsymbol{e}}) > k \right\rangle_{\mathbb{F}[\boldsymbol{x}]}$$

Proof Sketch: The proof is again by induction on k.

Base case: $k = 1$ which means $f(t,t,t) = 0$. This implies $f \in \mathcal{I}$, where $\mathcal{I} = \langle (x_1 - x_2), (x_1 - x_3) \rangle$. Hence, f satisfies the induction hypothesis since $cs(x_1 - x_2) = cs(x_1 - x_3) > 1$.

Inductive case: If f satisfies $cs \leq k$ hypothesis, then we wish to prove that $f \in \left\langle P_{\mathbf{e}} = \prod_{i \neq j} (x_i - x_j)^{e_{i,j}} \,|\, cs(P_{\mathbf{e}}) > k \right\rangle_{\mathbb{F}[\mathbf{x}]}$. Note that f satisfies $cs \leq (k-1)$ hypothesis, which by induction hypothesis implies

$$f \in \left\langle P_{\mathbf{e}} = \prod_{i \neq j} (x_i - x_j)^{e_{i,j}} \,|\, cs(P_{\mathbf{e}}) = k \right\rangle + \left\langle P_{\mathbf{e}} = \prod_{i \neq j} (x_i - x_j)^{e_{i,j}} \,|\, cs(P_{\mathbf{e}}) > k \right\rangle \tag{3.3}$$

By taking partial derivatives of f by monomials of cone size exactly k (the $\mathbf{e}$ in the first ideal), we will show that f is actually only in the second ideal above. Note that $\partial_{\mathbf{x}^{\mathbf{e}}} f(t,t,t)$ has 0 contribution from the second ideal. For simplicity, let us write Equation 3.3 in short as $f = \langle P_{=k} \rangle + \langle P_{>k} \rangle$. Incrementally, we will push all the terms in $\langle P_{=k} \rangle$ to $\langle P_{>k} \rangle$, by taking

partial derivatives by monomials of $cs = k$. Once, a term has been moved to $\langle P_{>k} \rangle$, we need not worry about its contribution in the partial derivatives since it will be zero due to the form of $\langle P_{>k} \rangle$.

Telescopic Differences: Suppose $\langle S_m \rangle \subseteq \langle P_{=k} \rangle = a_1 P_1 + a_2 P_2 + \ldots + a_r P_r$. This can be rewritten in a telescopic form as follows:

$$\langle S_m \rangle = a_1 P_1 - a_1 P_2 + (a_1 + a_2) P_2 - (a_1 + a_2) P_3 + \ldots + (a_1 + a_2 + \ldots + a_r) P_r$$

The reason behind such decomposition is well motivated. These differences are *"well-behaved"*. Also, the last lone summand $(a_1 + a_2 + \ldots + a_r)P_r$ is special as it will no longer be in $\langle P_{=k} \rangle$. This follows from the cone set argument above

$$1/\mathbf{e}! \cdot \partial_m \langle P_{=k} \rangle (\mathbf{t}) = a_1(\mathbf{t}) + a_2(\mathbf{t}) + \ldots + a_r(\mathbf{t}) = 0$$

This implies $\sum_{i=1}^{r} a_i \in \mathcal{I}$. Since, $P_r \in \langle P_{=k} \rangle$, with $cs(P_r) = k$,

$$\mathrm{cs}((\sum_{i=1}^{r} a_i) P_r) > k$$

That is, $(\sum_{i=1}^{r}) P_r \in \langle P_{>k} \rangle$, thus taking us closer to proving our induction step. Now, we are left with handling the telescopic differences.

Initial settings: Remember we were in the induction step: $f = \langle P_{=k} \rangle + \langle P_{>k} \rangle$. We focus on

$$\langle P_{=k} \rangle = \left\langle P_{\mathbf{e}} = (x_1 - x_2)^{e_{12}} . (x_1 - x_3)^{e_{13}} . (x_2 - x_3)^{e_{23}} \mid cs(P_{\mathbf{e}}) = k \right\rangle$$

. If $m_{12} = x_1^{e_{12}+e_{13}} \cdot x_2^{e_{23}}$ is a least cone monomial of $P_{\mathbf{e}}$, then $m_{13} = x_1^{e_{13}+e_{12}} \cdot x_3^{e_{23}}$ is also a least cone monomial of $P_{\mathbf{e}}$. Similarly $m_{21} \sim m_{23}$ and $m_{31} \sim m_{32}$ are the only other possible least cone monomials due to the form of $P_{\mathbf{e}}$.

Observe that if $P_{\mathbf{e}} \in S_{m_{12}}$, then $(P_{\mathbf{e}})_{x_2 \leftrightarrow x_3} \in S_{m_{12}}$, where in the latter we have swapped the variables x_2 and x_3. This is a nice structural property owing to the form of $P_{\mathbf{e}}$ which we will

exploit later in the next concept of atomic operations. WLOG, henceforth we will assume m_{12} of cone-size $= k$ is a least cone monomial of $P_{\mathbf{e}}$, unless stated otherwise. This puts us in the settings: $e_{12} \geq e_{23}$ and $e_{13} \geq e_{23}$, which can be easily verified. Also wlog, let $e_{12} \geq e_{13}$, because if not we can always swap $x_2 \leftrightarrow x_3$. This gives us:

$$e_{12} \geq e_{13} \geq e_{23}$$

Atomic Operations: Let us use the shorthand (e_{12}, e_{13}, e_{23}) for $P_{\mathbf{e}}$. We assumed that $(e_{12}, e_{13}, e_{23}) \in S_{m_{12}}$. We define atomic operation by decrementing/incrementing e_{12}, e_{13} by -1, +1 respectively (Total degree should not change). The question is whether $(e_{12} - 1, e_{13} + 1, e_{23}) \in S_{m_{12}}$? This question is non-trivial because although monomial m_{12} still has the same cone-size k, this operation might give rise to some other least cone monomial with cone size $< k$, for example m_{31} which has cone-size $(e_{13} + e_{23} + 1) \cdot (e_{12} - 1)$ (actually both the products also have a $+1$ but we will ignore these in cone sizes since it will not affect the calculations). But, we will rule out such possibilities because we knew that $cs(m_{31}) \geq k$ in both $P_{\mathbf{e}}$ and $(P_{\mathbf{e}})_{x_2 \leftrightarrow x_3}$, that is $(e_{13} + e_{23}).(e_{12}) \geq k$ and $(e_{12} + e_{23}).(e_{13}) \geq k$. We formalize this, in the following claim:

Lemma 21. $(e_{12} - \delta, e_{13} + \delta, e_{23}) \in S_{m_{12}}$ *for* $\delta = \{0, 1, \ldots, e_{12} - e_{13}\}$

Proof. We first show that $m_{31} \sim m_{32}$ cannot get a cone-size $< k$ in $(e_{12} - \delta, e_{13} + \delta, e_{23})$. The claim for end points $\delta = 0, e_{12} - e_{13}$ is given, that is $(e_{13} + e_{23}).(e_{12}) \geq k$ and $(e_{12} + e_{23}).(e_{13}) \geq k$. We need to use this to prove it for the rest of the values of δ. Note that this can be formulated as a function $h(x) = x \cdot (c - x)$, where $c = e_{12} + e_{13} + e_{23}$ is a constant. Since the second derivative $h''(x) < 0$, this function is concave, hence all intermediate values on the graph of $h(x)$ between the two end points for $\delta = 0$ and $\delta = e_{12} - e_{13}$ are $\geq k$. Note that start point $\delta = 0$ polynomial has cone-size $\geq$ end point $\delta = e_{12} - e_{13}$. Therefore, if we take difference of two consecutive polynomials, the resulting polynomial can have cone-size $= k$ only in the last difference.

Similar argument also shows that $m_{21} \sim m_{23}$ cannot get cone-size $< k$ in $(e_{12} - \delta, e_{13} + \delta, e_{23})$. Here, if we take consecutive differences, the resulting polynomial can have cone-size $= k$ only in the first difference. □

Atomic Differences: Now we shall combine telescopic differences and atomic operations to show the aforementioned *"well-behaved"*ness of differences.

If $S_{m_{12}}$ is singleton, then we simply take partial derivative by m_{12}, evaluate at $\mathbf{t}$ to show that it's coefficient $\in \mathcal{I}$, and hence it is in $\langle P_{>k} \rangle$, and we are done. If we have more than one polynomial in $S_{m_{12}}$, pick the one with bigger e_{12}. Then we subtract and add telescopically with $(e_{12} - \delta, e_{13} + \delta, e_{23})$ till we reach the next $\mathbf{e}'$ with smaller e_{12} and so on. Let us analyze any single atomic difference step:

$$(e_{12}, e_{13}, e_{23}) - (e_{12} - 1, e_{13} + 1, e_{23}) = (e_{12} - 1, e_{13}, e_{23} + 1)$$

This nice factorization of two differences is useful for us in making progress in our induction step.. By Claim 21, we know that in any intermediate step, the polynomials being subtracted, both of them $\in S_{m_{12}}$. Thus the difference of the two will give a resulting polynomial also with minimal cone size $\geq k$. We handle it in two cases:

Case 1: The resulting polynomial has cone-size $> k$. This puts it in $\langle P_{>k} \rangle$ which is a good case. Note that this will happen when both the polynomials being subtracted had only $m_{12} \sim m_{13}$ as the only least cone monomial, which got subtracted out.

Case 2: The resulting polynomial has cone-size $= k$. Then, the resulting polynomial either belongs in $S_{m_{21} - m_{12}}$ or $S_{m_{31} - m_{12}}$. The proof of Claim 21 reveals that it will move to $S_{m_{21}}$ only in the first difference, if at all and $S_{m_{31}}$ only in the last difference, if at all. Rest of the differences will be Case 1.

Now we will repeat this process for polynomials $\in S_{m_{21}}$. The good thing is that on taking derivative with m_{21} at $\mathbf{t}$, common polynomials from $S_{m_{12}}$ cannot interfere as they have already been removed. Here also, either the atomic differences will have cone-size$> k$,

which is good or they will have cone-size = k. In the latter case, either it will give m_{31} uniquely, if at all, and it may give m_{12} uniquely. The m_{31} case will be handled in next step, and the m_{12} case can be handled simply by taking derivative with m_{12} at $\mathbf{t}$. Since, m_{12} was produced uniquely, no other polynomial can interfere, and hence its coefficient $\in \mathcal{I}$ pushing the polynomial in $\langle P_{>k} \rangle$.

Again, we repeat the process for the remaining polynomials of $\langle P_{=k} \rangle$ which are in $S_{m_{31}}$ but do not have m_{12} or m_{21} as their least cone monomials. Here again the difficult case is when the atomic differences might produce m_{12} or m_{21} but since they produce it just once, they can be moved to $\langle P_{>k} \rangle$ by taking partial derivatives respectively.

Finally, we have showed through the method of telescopic atomic differences, that f is generated by polynomials of cone-size $> k$, that is

$$f \in \left\langle P_{\mathbf{e}} = \prod_{i \neq j} (x_i - x_j)^{e_{i,j}} \mid cs(P_{\mathbf{e}}) > k \right\rangle$$

which completes the induction step and proof of Lemma 20.

3.3.2 Some related width Results

We first prove a width result for the $n = 3$ case.

Theorem 9 (Width result for the trivariate case)**.** *Let*

$$\mathcal{I} \coloneqq \left\langle P_{\boldsymbol{e}} = (x_1 - x_2)^{e_{12}}.(x_1 - x_3)^{e_{13}}.(x_2 - x_3)^{e_{23}} \mid cs(P_{\boldsymbol{e}}) > k \right\rangle_{\mathbb{F}[\boldsymbol{x}]}$$

If $f \in \mathcal{I}$, then it has width $\geq O(\sqrt{k})$

Proof. We will use techniques similar to the proof of Theorem 8.
Now, $f \in \mathcal{I}$ can be split into degree-d homogeneous parts. Each of these contribute separately to the width because there are no cancellations and taking the partial derivative with respect to them, will give distinct leading monomials. Let f_d be the degree-d homogeneous part of

f. Then $w(f) \geq w(f_d)$.

Hence, we already can write

$$f = \sum_{cs(P_{\mathbf{e}})>k} a_{\mathbf{e}}(x_2 - x_1)^{e_{2,1}}(x_3 - x_1)^{e_{3,1}}(x_2 - x_3)^{e_{2,3}}$$

where $a_{\mathbf{e}} \in \mathbb{F}$ and $\mathbf{e} = (e_{2,1}, e_{3,1}, e_{2,3}) \in \mathbb{N}^n$

Now, because of the assumption, we can write

$$f_d = \sum_{cs(P_{\mathbf{e}})>k} a_{\mathbf{e}} x_1^d (t_2 - 1)^{e_{2,1}}(t_3 - 1)^{e_{3,1}}(t_2 - t_3)^{e_{2,3}}$$

where $x_2 = t_2 x_1$ and $x_3 = t_3 x_1$.

Now $\text{cs}(P_{\mathbf{e}}) > k$ means that one of the following things will happen

$$(e_{2,1} + e_{3,1} + 1)(e_{2,3} + 1) > k \text{ or}$$

$$(e_{2,1} + e_{2,3} + 1)(e_{3,1} + 1) > k \text{ or}$$

$$(e_{2,3} + e_{3,1} + 1)(e_{2,1} + 1) > k$$

By Pigeonhole principle, we get that for each of these situations, there exists some i, j such that $e_{i,j} \geq \frac{\sqrt{k}}{2} = l$.

We can then write the polynomial f_d as

$$f_d = (t_2 - 1)^{\frac{l}{2}} g_2(t_2, t_3) + (t_3 - 1)^{\frac{l}{2}} g_3(t_2, t_3) + (t_2 - t_3)^l g_{2,3}(t_2, t_3)$$

Now we can split $(t_2 - t_3)^l$ as $((t_2 - t_1) - (t_3 - t_1))^l = \sum_i \binom{l}{i} (t_2 - 1)^i (t_3)^{l-i}$. Now, either $i \geq \frac{l}{2}$ or $l - i \geq \frac{l}{2}$. So, we can split this part and collect the terms in the first two summands and write the polynomial as

$$f_d = (t_2 - 1)^{\frac{l}{2}} g_2'(t_2, t_3) + (t_3 - 1)^{\frac{l}{2}} g_3'(t_2, t_3)$$

We can use the same proof as 2 to show that $\text{sp}(f_d) \geq \frac{\sqrt{k}}{4}$.

This gives us a proof that

$$w(f) \geq O(\sqrt{k})$$

□

We prove a result for a general class of polynomials that satisfies the structural conjecture. These polynomials can serve to be a toy case and the techniques used here could potentially be generalised to the class of polynomials satisfying the cone-size hypothesis.

Theorem 10.

$$f = \prod_{i<j\in[n]} (x_i - x_j)^l$$

Then any ROABP computing f must have width $> l^{\lfloor n/2 \rfloor}$

We first give a proof in the $n = 4$ case to give an idea of the proof strategy.

Claim 3. *Let*

$$f = \prod_{i<j\in[4]} (x_i - x_j)^l$$

Then any ROABP computing f must have width $> l^2$

Proof. We look at the dimension of the following coefficient space

$$\text{coeff}_{x_1^i x_3^{2l+j}}(f) \text{ where } i, j \in [0, l]$$

We also apply a monomial ordering on the set of monomials that is guided by the following ordering on the set of variables $x_2 \succ x_4$. We look at the leading monomial with respect to this monomial ordering. We claim that

$$\text{LM}\Big(\text{coeff}_{x_1^i x_3^{2l+j}}(f)\Big) \text{ where } i, j \in [0, l]$$

We claim that $x_2^{2l-i} x_4^{l-j}((x_2 - x_4)^l)$ is the leading monomial and has non-zero coefficient.

Now $(x_2 - x_4)^l | \text{LM}(\text{coeff}_{x_1^i x_3^j}(f))$ for all i, j. This is because

$$\text{coeff}_{x_1^i x_3^j}(f) = \partial_{x_1^i x_3^j}(f)\Big|_{x_1, x_3 = 0}$$

Using the product rule, the partial derivative operator $\partial_{x_1^i x_3^j}$ leaves $(x_2 - x_4)^l$ factor unchanged.

Now for the leading monomial to have non-zero coefficient $(x_1 - x_3) \nmid \text{LM}(\partial_{x_1^i x_3^{2l+j}}(f))$. Now, we want to show that, $m = x_2^{2l-i}x_4^{l-j}$ can be computed in only one way. Let us assume on the contrary, that m can also be computed if x_1 takes $i' < i$ from the $(x_1 - x_4)^l$ bracket and takes the remaining $i - i'$ from the $(x_1 - x_2)^l$ bracket. More formally,

$$\begin{aligned}
&LM(\partial_{x_3^{l+j}}(\partial_{x_1^i}((x_1 - x_2)^l(x_1 - x_4)^l)(x_3 - x_4)^l(x_3 - x_2)^l)) \\
=&LM(\partial_{x_3^{l+j}}((x_1 - x_2)^{l-i+i'}(x_1 - x_4)^{l-i'}(x_3 - x_4)^l(x_3 - x_2)^l)) \\
=&x_2^{2l-i+i'-j}x_4^{l-i'}
\end{aligned}$$

But $i' < i$, hence, $x_2^{2l-i+i'-j}x_4^{l-i'} \prec x_2^{2l-i}x_4^{l-j}$ due to the monomial ordering.

Similarly, we can also assume on the contrary, m can also be computed if x_3 takes $j' > j$ from the $(x_3 - x_2)^l$ bracket and takes the remaining $l - j + j'$ from the $(x_3 - x_4)^l$ bracket. But this gives a leading monomial $x_2^{2l-j'}x_3^{l-i+j'-j}$ which is $\prec x_2^{2l-i}x_4^{l-j}$.

Hence, $\text{LM}\Big(\text{coeff}_{x_1^i x_3^{2l+j}}(f)\Big) = x_2^{2l-i}x_4^{l-j}((x_2 - x_4)^l)$ where $i, j \in [0, l]$ and these have non-zero coefficients. Also, these have all distinct l^2 many distinct individual degrees and hence, using Lemma 9 and then using Corollary 5, we get

$$w(f) \geq \dim\Big(\{\text{coeff}_{x_1^i x_3^{2l+j}}(f) \text{ where } i, j \in [0, l]\}\Big) = l^2$$

. □

Proof of Theorem (3.3.2). We assume for simpler calculations that $n = 2k$. Then we look at the dimension of the following coefficient space of f. We first define the monomial set with respect to which we look at the partial derivatives.

$$\Delta = \Big\{ \prod_{i=1}^{k} x_{2i-1}^{(2i-2)l+j_i} \text{ where } i, j_1, ..., j_k \in [0, l) \Big\}$$

We define the set of odd numbered variables to be $O = \{1, 3, ..., n-1\}$. Let $\mathbf{x}_O$ be the restriction of $\mathbf{x}$ to the odd numbered variables. So we want to calculate the dimension of the following set

$$\partial_\Delta(f) = \Big\{ \partial_m(f)|_{\mathbf{x}_O = \mathbf{0}} \Big| m \in \Delta \Big\}$$

Now, $|\partial_\Delta(f)| \leq l^k$. We claim that the equality holds.

Now we assume there is a *deg-lex* monomial ordering on the $m \in \partial_\Delta(f)$ with $x_2 > x_4 > ... > x_2k$ and we look at the leading monomial with respect to this monomial ordering.

We first give an algorithm to find a **leading monomial** with respect to this ordering.

Preprocessing Step: Since, $(x_{2i} - x_{2j})^l$ divides all of the partial derivatives, so we can remove them and look at the partial derivatives on the remaining f. Now, if the leading monomial must have non-zero coefficient, $((x_{2i-1} - x_{2j-1}) \nmid LM(\partial_m(f)))$. For each x_{2i-1}, $x_{2i-1}^{(i-1)l}$ must be used up. So we define

$$\Delta' = \Big\{ \prod_{i=1}^{k} x_{2i-1}^{(i-1)l+j_{2i-1}} \text{ where } i, j_1, ..., j_k \in [0, l) \Big\}$$

So we define

$$f' = \frac{f}{\prod_{i<j\in[k]}(x_{2i} - x_{2j})^l (x_{2i-1} - x_{2j-1})^l}$$

and now we want to compute.

$$\partial_{\Delta'}(f') = \Big\{ \partial_m(f')|_{\mathbf{x}_O=\mathbf{0}} \Big| m \in \Delta' \Big\}$$

Algorithm 1: Algorithm to find leading monomial

Result: This algorithm finds the leading monomial in the processed polynomial

$i = k$, Leadmo $= 1$, $f = \prod_{i<j\in[n]}(x_i - x_j)^l$, mon $= \prod_{j=1}^{k}(x_{2j-1})^l$,$f' = 0$;

while $i > 0$ **do**

  mon $= \frac{\text{mon}}{(x_{2i-1})^l}$;

  $f' = \partial_{mon}(f)$;

  Leadmo $=$ Leadmo $\times$ LM$(\partial_{x_{2i-1}^{(i-1)l+j_{2i-1}}}(f')|_{x_{2i-1}=0})$;

  $i = i - 1$;

end

Return: Leadmo

Proof of Correctness: This returns the leading monomial for the given Δ' because we maintain the invariant of finding the leading monomial at every step.

Let $m \in \Delta$ with parameters j_{2i-1} where $i \in [k]$.

$$\mathrm{LM}(\partial_m(f')) = \prod_{i=1}^{k} x_{2i}^{(k-(i-1))l-j_{n-(2i-1)}}$$

Now, it is obvious, that for different sets of j_i, we get different leading monomials. (Just look at every degree of x_2, it only depends on j_{n-1} and this happens for every $i = 2j$). Hence using Corollary (5)

$$\left|\partial_\Delta(f)\right| \geq \left|\left\{\mathrm{LM}(\partial_m(f))\middle| m \in \Delta\right\}\right| \geq l^k$$

Now we know from Lemma (9)

$$w(f) \geq \left|\partial_\Delta(f)\right| \geq l^k$$

□

Chapter Four

Stronger Cone-Size Hypothesis

We will be using the famous sparse PIT map as our shift. Structure Lemma (22) will describe the form of such a shifted polynomial. Then, we intend to use the special properties of the sparse PIT map in proving the width conjecture(3), to achieve the desired bound of hitting set.

Recall that in sparse PIT map, one constructs a weight function $w : \bar{x} \to \mathbb{N}$ such that when x_i is replaced with $t^{w(i)}$, it gives distinct weights to all the monomials, thus keeping the polynomial non-zero after substitution.

Here we again restate this lemma that we mentioned and proved in the preliminaries section.

Lemma 6: Let M be the set of all monomials in n variables $\bar{x} = \{x_1, x_2, \ldots, x_n\}$ with maximum individual degree d. For any value s, there is a polynomial-time constructible set of $N := ns\log(d+1)$ weight functions from $\bar{x}$ to $[2NlogN]$, such that for any set $A \subseteq M^2$ of s pairs of monomials, at least one of the weight functions w separates all the pairs in A; i.e., for all $(m, m') \in A, w(m) \neq w(m')$.

In other words, this lemma says that sparse PIT map **preserves $\mathbb{F}$-linear independence** of a given set of sparse number of monomials. Stated explicitly, it says:

Corollary 6. *Let $m_1, m_2, \ldots, m_s$ be a set of $\mathbb{F}$-linearly independent monomials, and let ϕ be the sparse PIT map for s-sparse polynomials as in Lemma(6). Then $\phi(m_1), \ldots, \phi(m_s)$ are also $\mathbb{F}$-linearly independent.*

Proof. Let $c_1, \ldots, c_s \in \mathbb{F}$. Then:

$$c_1 m_1 + \ldots + c_s m_s = 0 \Leftrightarrow c_1\phi(m_1) + \ldots + c_s\phi(m_s) = 0 \Rightarrow \forall i, c_i = 0$$

where the latter implication holds since, $\phi(m_1), \ldots, \phi(m_s)$ have distinct degree. □

The shift: Let $f(\bar{x}) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. Then we will shift f by $x_i \to x_i + t_i$ to get $f' \in \mathbb{F}[t][\bar{x}]$, where $t_1 = t, t_2 = t_2(t), t_3 = t_3(t), \ldots, t_n = t_n(t)$. The univariate map $(t_1, \ldots, t_n)$ is the sparse PIT map for some fixed sparsity (say) $\leq k$-sparse polynomials.

Cone size$\leq k$ Hypothesis: All cone size $\leq k$ monomials in $f' = f(\bar{x} + \bar{t})$ have coefficients $= 0$, where $\bar{t} = (t, t_2(t), \ldots, t_n(t))$.

Note that the coefficients of f' are univariate polynomials in t, and the hypothesis demands that $\forall \bar{e}$ such that $cs(\bar{e}) \leq k$, $\text{coeff}_{\bar{x}^{\bar{e}}}(f') = 0$. As, we saw in Exercise 3.1, the cone size hypothesis can also be written as :

$$\partial_{\bar{x}^{\bar{e}}} f(t, t_2(t), \ldots, t_n(t)) = 0, \quad \forall \bar{e}, cs(\bar{e}) \leq k$$

What can we say about a polynomial f which satisfies $cs \leq k$ hypothesis? We wish to explore the structure of f with the motive of getting a faster hitting set through that structure.

4.1 Structure Lemma

With the above definition of $cs \leq k$ hypothesis, we can prove the following lemma:

Lemma 22. *If f satisfies $cs \leq k$ hypothesis, then*

$$f \in \left\langle P_{\bar{e}} = \big(x_2 - t_2(x_1)\big)^{e_2}\big(x_3 - t_3(x_1)\big)^{e_3} \ldots \big(x_n - t_n(x_1)\big)^{e_n} \;\middle|\; cs(e_2 \cdots e_n) > k \right\rangle_{\mathbb{F}[x_1]}$$

Proof. The proof is by induction on k. Let

$$\mathcal{I} = \left\langle \big(x_2 - t_2(x_1)\big), \big(x_3 - t_3(x_1)\big), \ldots, \big(x_n - t_n(x_1)\big) \right\rangle_{\mathbb{F}[\bar{x}]}$$

Base Case: $(k = 1)$ $cs \leq 1$ hypothesis $\Rightarrow f(\bar{t}) = 0 \Rightarrow f \in \mathcal{I}$, where the cone-size of the exponents of generators is indeed > 1.

Inductive Case: By Induction hypothesis for $(k-1)$,

$$f \in \Big\langle P_{\bar{e}} \,|\, cs(\bar{e}) > (k-1) \Big\rangle$$

that is,

$$f = \sum_{\bar{e}: cs(\bar{e}) \geq k} a_{\bar{e}}(\bar{x}) P_{\bar{e}}$$

Now, f satisfies $cs \leq k$ hypothesis implies $\forall \bar{e} : cs(\bar{e}) \leq k, \partial_{\bar{x}^{\bar{e}}} f(\bar{t}) = 0$. In particular, for $a_{\bar{e}} P_{\bar{e}} = a_{\bar{e}}(\bar{x}) \cdot \big(x_2 - t_2(x_1)\big)^{e_2} \big(x_3 - t_3(x_1)\big)^{e_3} \ldots \big(x_n - t_n(x_1)\big)^{e_n}$ with $cs(e_2 \cdots e_n) = k$, we have

$$\partial_{x_2^{e_2} \ldots x_n^{e_n}} \Big(a_{\bar{e}} P_{\bar{e}}(\bar{t}) \Big) = 0 \Rightarrow a_{\bar{e}}(\bar{t}) = 0 \Rightarrow a_{\bar{e}} \in \mathcal{I}$$

Therefore f satisfies $cs \leq k$ hypothesis means,

$$\begin{aligned}
\partial_{x_2^{e_2} \ldots x_n^{e_n}} \big(f(\bar{t}) \big) = 0 \quad & \forall (e_2, \ldots, e_n) : cs(\bar{e}) = k \\
\partial_{x_2^{e_2} \ldots x_n^{e_n}} \Big(\sum_{\bar{e}: cs(\bar{e}) \geq k} a_{\bar{e}}(\bar{x}) P_{\bar{e}}(\bar{t}) \Big) = 0 \quad & \forall (e_2, \ldots, e_n) : cs(\bar{e}) = k \\
\sum_{\bar{e}: cs(\bar{e}) \geq k} \partial_{x_2^{e_2} \ldots x_n^{e_n}} \Big(a_{\bar{e}}(\bar{x}) P_{\bar{e}}(\bar{t}) \Big) = 0 \quad & \forall (e_2, \ldots, e_n) : cs(\bar{e}) = k \\
a_{\bar{e}}(\bar{x}) \in \mathcal{I} \quad & \forall (e_2, \ldots, e_n) : cs(\bar{e}) = k \\
f \in \Big\langle P_{\bar{e}} \,|\, cs(\bar{e}) > k \Big\rangle_{\mathbb{F}[\bar{x}]} &
\end{aligned}$$

Note that we are still not done, as in the proof above the ideal is over $\mathbb{F}[x_1, x_2, \ldots, x_n]$, while in the lemma statement we want ideal over $\mathbb{F}[x_1]$. However, due to the specific structure of $P_{\bar{e}}$, there is a simple trick: In $a_{\bar{e}}(\bar{x})$, replace x_i by $\big(x_i - t_i(x_1)\big) + t_i(x_1)$, $\forall i \in \{2, \ldots, n\}$. Using binomial expansion, we can push the $(x_i - t_i(x_1))^d$ terms into $P_{\bar{e}}$, as $cs(\bar{e})$ will only increase. We might get extra summands but that does not matter. Thus, we will still get the required form

$$f = \sum_{\bar{e}: cs(\bar{e}) \geq k} a'_{\bar{e}}(x_1) Q_{\bar{e}}$$

. □

4.2 Width Conjecture

The benefit of having the above structural form of f will be justified by the following conjecture:

Conjecture 3. *If* $f \in \left\langle P_{\bar{e}} = \big(x_2 - t_2(x_1)\big)^{e_2} \dots \big(x_n - t_n(x_1)\big)^{e_n} \;\middle|\; cs(\bar{e}) > k \right\rangle_{\mathbb{F}[x_1]}$, *then ROABP width* $w(f) > k$, *where* $\bar{t}$ *is the sparse PIT map for* $\leq k$ *sparse polynomials.*

Our first step will be to show that $P_{\bar{e}} = \big(x_2 - t_2(x_1)\big)^{e_2} \dots \big(x_n - t_n(x_1)\big)^{e_n}$, where $cs(\bar{e}) > k$ has $w(P_{\bar{e}}) > k$. We shall then progress towards showing the same lower bound for $\sum_{\bar{e}:cs(\bar{e})>k} a_{\bar{e}} P_{\bar{e}}$, where $a_{\bar{e}} \in \mathbb{F}$. The final nail in the coffin is to then show the width lower bound for $\sum_{\bar{e}:cs(\bar{e})>k} a_{\bar{e}}(x_1) P_{\bar{e}}$, where $a_{\bar{e}}(x_1) \in \mathbb{F}[x_1]$, which is expected to be the most difficult part. The difficulty lies in ruling out possible width reduction due to possible cancellations.

Hitting set overview: Suppose we are given an input polynomial f as a black-box ROABP of size k (width $\leq k$), and number of variables $n = O(\log k)$. We know that number of $cs \leq k$ monomials when $n = O(\log k)$ is $poly(k)$ (Lemma 2.5). Hence, in polynomial time we will test whether the coefficients of $cs \leq k$ monomials in $f(x_1 + t, x_2 + t_2(t), \dots, x_n + t_n(t))$ are zero, which can be done in polynomial time as they are polynomially many and each coefficient is a univariate polynomial in t. If the input polynomial f was non-zero, then one of these coefficients must be non-zero. Because otherwise, f satisfies $cs \leq k$ hypothesis, which by Structure Lemma (22) and Width Conjecture(3) will contradict the width of input polynomial.

4.3 Single summand

The first natural and necessary step to proving Width Conjecture(3) would be to first prove the width lower bound for a single summand.

Lemma 23. *For $P_{\bar{e}} = a_{\bar{e}}(x_1) \cdot \big(x_2 - t_2(x_1)\big)^{e_2} \dots \big(x_n - t_n(x_1)\big)^{e_n}$, ROABP width $w(P_{\bar{e}}) > k$, where $cs(\bar{e}) > k$ and $\bar{t}$ is sparse PIT map for $k+1$ sparse polynomials.*

Proof. Let $g = \big(x_2 - t_2(x_1)\big)^{e_2} \dots \big(x_n - t_n(x_1)\big)^{e_n}$. Consider the set of univariate polynomials obtained by taking all the coefficients of g with respect to x_2 to x_n, that is:

$$\text{coeff}(\bar{x}^*_{2,n})(g) = t_2(x_1)^{\leq e_2} t_3(x_1)^{\leq e_3} \dots t_n(x_1)^{\leq e_n}$$

Since, we have $|\text{coeff}(\bar{x}^*_{2,n})(g)| = cs(\bar{e}) > k$, we will consider any subset of coefficients of size $k+1$. By 6, the polynomials in that set will be $\mathbb{F}$ linearly independent, as they will have distinct x_1 degrees. Therefore, for the complete set,

$$rank_{\mathbb{F}}\Big(\text{coeff}(\bar{x}^*_{2,n})(g)\Big) > k$$

By Nisan's width criterion, this implies $w(g) > k$. For $f = a_{\bar{e}}(x_1) \cdot g$, the argument is same as above, since $rank_{\mathbb{F}}\big(a_{\bar{e}} \cdot \text{coeff}(\bar{x}^*_{2,n})(g)\big) = rank_{\mathbb{F}}\big(\text{coeff}(\bar{x}^*_{2,n})(g)\big)$. □

4.4 A (probable) step towards the width conjecture

Lemma 24 (Degree-Width)**.** *Let*

$$f = a_1 \cdot (x_2 - t_2(x_1))^{e_2} \cdot g_1 + a_2 \cdot (x_2 - t_2)^{e_2'} \cdot g_2 + a_3 \cdot (x_2 - t_2)^{e_2''} \cdot g_3 + \cdots$$

where $a_i \in \mathbb{F}$, $e_2 \geq e_2' \geq e_2'' \cdots$, f has arbitrary number of summands and each g_i contains the remaining product for that summand (For example $g_1 = (x_3 - t_3(x_1))^{e_3} \cdots (x_n - t_n(x_1))^{e_n}$). Then f has width, $w(f) > e_2$.

Proof. Ignoring the constants and signs (they will not affect proof), let us consider a subset

of $coeff(x_2^*)(f)$:

$$\begin{aligned}
\frac{\partial_{x_2^{e_2}}}{e_2!} f_{|x_2=0} &= g_1 \\
\frac{\partial_{x_2^{e_2-1}}}{(e_2-1)!} f_{|x_2=0} &= \binom{e_2}{1} \cdot t_2 \cdot g_1 \\
&\vdots \\
\frac{\partial_{x_2^{e_2'}}}{e_2'!} f_{|x_2=0} &= \binom{e_2}{e_2-e_2'} \cdot t_2^{e_2-e_2'} \cdot g_1 + g_2 \\
&\vdots \\
\frac{\partial_{x_2^{e_2''}}}{e_2''!} f_{|x_2=0} &= \binom{e_2}{e_2-e_2''} \cdot t_2^{e_2-e_2''} \cdot g_1 + \binom{e_2'}{e_2'-e_2''} \cdot t_2^{e_2'-e_2''} \cdot g_2 + g_3 \\
&\vdots \\
f_{|x_2=0} &= t_2^{e_2} \cdot g_1 + t_2^{e_2'} \cdot g_2 + t_2^{e_2''} \cdot g_3
\end{aligned}$$

Let us look into the $\mathbb{F}$ linear rank of these coefficients via a matrix equation.

$$\bar{\alpha} \cdot \begin{bmatrix}
1 & 0 & 0 & \cdots \\
\binom{e_2}{1} t_2 & 0 & 0 & \cdots \\
 & \vdots & & \\
\binom{e_2}{e_2'} t_2^{e_2-e_2'} & 1 & 0 & \cdots \\
 & \vdots & & \\
\binom{e_2}{e_2''} t_2^{e_2-e_2''} & \binom{e_2'}{e_2''} t_2^{e_2'-e_2''} & 1 & \\
 & \vdots & & \\
t_2^{e_2} & t_2^{e_2'} & t_2^{e_2''} & \cdots
\end{bmatrix} \cdot \begin{pmatrix} g_1 \\ g_2 \\ g_3 \\ \vdots \end{pmatrix} = 0 \tag{4.1}$$

Above $\bar{\alpha} \in \mathbb{F}^{(e_2+1)}$. Let us index the columns by C_1, C_2 and so on.

In order to show $w(f) > e_2$, we need to prove that 4.1 cannot hold. Note that g_i, g_j are $\mathbb{F}[x_1]$ independent unless they have the same exponent $(e_3, \cdots, e_n)$

Let us look at the case when all g_i's are $\mathbb{F}[x_1]$ independent. Let us rewrite 4.1 as:

$$\sum_i \langle \bar{\alpha}, C_i \rangle . g_i = 0 \tag{4.2}$$

where $\langle \bar{\alpha}, C_i \rangle$ is the inner product between $\bar{\alpha}$ and C_i which is the i-th column vector. Observe that $\langle \bar{\alpha}, C_i \rangle \in \mathbb{F}[x_1]$. Also, $\langle \bar{\alpha}, C_1 \rangle \neq 0$, since the terms in C_1 are distinct degree non-zero monomials in x_1 and are hence $\mathbb{F}$ linearly independent. And since g_1 is $\mathbb{F}[x_1]$ independent of other g_i's, $\langle \bar{\alpha}, C_1 \rangle$ cannot be cancelled. Therefore, 4.2 cannot be true.

Now let us see why 4.1 cannot hold even when g_i's are $\mathbb{F}[x_1$ dependent. First, observe that if g_i and g_j are $\mathbb{F}[x_1]$ dependent, then they are $\mathbb{F}$ dependent. Since the $(e_3, \cdots, e_n)$ exponents will be same, g_j will just be a constant multiple of g_i. Also, one of the following **two conditions** must always hold:

1. g_i and g_j are $\mathbb{F}[x_1]$ independent.

2. C_i and C_j (the corresponding columns) are distinct.

Both the conditions cannot be simultaneously violated, since otherwise

$$(x_2 - t_2)^{e_i} g_i + (x_2 - t_2)^{e_j} g_j = (x_2 - t_2)^{e_i} g_i + (x_2 - t_2)^{e_i} (c.g_i)$$

which can then be clubbed into a single summand $c'(x_2 - t_2)^{e_i} g_i$ and hence is a contradiction (c and c' are constants).

Therefore, if g_1 and g_i are dependent, then

$$\langle \bar{\alpha}, C_1 \rangle g_1 + \langle \bar{\alpha}, C_i \rangle g_i = \langle \bar{\alpha}, C_1 + c \cdot C_i \rangle g_1$$

Since $C_1 \neq C_i$, the leading monomial t^{e_2} in C_1 cannot be cancelled, hence $\langle \bar{\alpha}, C_1 + c \cdot C_i \rangle \neq 0$. Therefore all the g_i's which are $\mathbb{F}[x_1]$ dependent on g_1 can be clubbed together in Equation 4.2, and we still get the same form as before:

$$\Big\langle \bar{\alpha}, \big(C_1 + \sum_{j=1}^{l} c_j C_{i_j}\big) \Big\rangle . g_1 + \langle \bar{\alpha}, C_2 \rangle . g_2 + \langle \bar{\alpha}, C_3 \rangle . g_3 + ...$$

where g_1 is $\mathbb{F}[x_1]$ independent of g_2, g_3 and rest of the g_i's in the sum, implying that the sum is non-zero, as earlier.

Thus, there is no non-zero $\bar{\alpha} \in \mathbb{F}^{e_2+1}$ satisfying Equation 4.1, which means

$$rank_{\mathbb{F}}(\text{coeff}_{x_2^{\leq\infty}}(f)) > e_2$$

which implies that $w(f) > e_2$, as required. □

Chapter Five

A simpler proof of cone-size concentration for BIWA

In this section, we give a different and a simpler proof for the fact that polynomials when shifted by a basis isolating weight assignment have cone-size concentration. This is already implied from [FGS18] who show that polynomials when shifted by a basis isolating weight assignment has a cone-closed basis over $\mathbb{F}(t)$ (Theorem 5) This implies that the shifted polynomial has cone-size concentration.(Lemma 17)

[Gur+15] proved that a polynomial in $\mathbb{A}_k[\mathbf{x}]$ when shifted by a Basis Isolating Weight Assignment gives $\log(k+1)$-support concentration. We improve on the proof of this result, especially by the improvement of a combinatorial lemma gives us a simpler proof for the fact that the polynomial shifted by a Basis Isolating Weight Assignment has cs $\leq k$ concentration. This is strictly better than the log-support concentration result.

5.1 Introduction

Recall that a polynomial $A(x)$ over an $\mathbb{F}$-algebra $\mathbb{A}$ is called low-support concentrated if its low-support coefficients span all its coefficients. We use the quasi-polynomial size hitting-set for ROABPs given by Agrawal et al. [Agr+15]. Their hitting-set is based on a basis isolating

weight assignment which we define next

Definition 10 (Basis Isolating Weight Assignment)**.** *A weight assignment w is called a* ***basis isolating weight assignment*** *for a polynomial $P(x) \in \mathbb{F}[\boldsymbol{x}]$, if there exists a set of monomials B such that*

- *The coefficients of the monomials in B form a basis for $sp(P)$*
- *weights of all the monomials in B are distinct*
- *For all $m \in supp(P) \setminus B$, $coeff_m(P) \in span_{\mathbb{F}}\{coeff_{m'}(P) | m' \in B, w(m') < w(m)\}$*

Let $A'(\mathbf{x})$ be the shifted polynomial

$$A'(\mathbf{x}) = A(\mathbf{x} + t^{\mathbf{w}}) = A(x_1 + t^{w_1}, ..., x_n + t^{w_n})$$

We will prove that A' has cone-size concentration

The coefficients of A' are $\mathbb{F}[t]$-linear combinations of the coefficients of A. So we get the following equations

$$\text{coeff}_{A'}(\mathbf{x}^{\mathbf{a}}) = \sum_{\mathbf{b} \in M} \binom{\mathbf{b}}{\mathbf{a}} t^{\mathbf{w}(\mathbf{b}-\mathbf{a})}.\text{coeff}_A(\mathbf{x}^{\mathbf{b}}) \tag{5.1}$$

where $\binom{\mathbf{b}}{\mathbf{a}} = \prod_{i=1}^{n} \binom{b_i}{a_i}$ for any $\mathbf{a}, \mathbf{b} \in \mathbb{N}^n$.

This equation can be further expressed in terms of matrices. Let C be the coefficient matrix of A, i.e. the $M \times [k]$ matrix with $\text{coeff}_A(\mathbf{x}^{\mathbf{a}})$ as rows. Let C' be the coefficient matrix of A', i.e. the $M \times [k]$ matrix with $\text{coeff}_{A'}(\mathbf{x}^{\mathbf{a}})$ as rows.

$$C(\mathbf{a}, \cdot) = \text{coeff}_A(\mathbf{x}^{\mathbf{a}})^T \tag{5.2}$$

Let T be the $M \times M$ transfer matrix denoted by

$$T(\mathbf{a}, \mathbf{b}) = \binom{\mathbf{b}}{\mathbf{a}}$$

and let D be the $M \times M$ diagonal matrix denoted by

$$D(\mathbf{a}, \mathbf{a}) = t^{\mathbf{w}(\mathbf{a})}$$

From the equation of the coefficients of A' in terms of A, we get that

$$C' = D^{-1}TDC \tag{5.3}$$

5.2 Isolation to Concentration

We will prove that A' has cone-size concentration. There is a standard method of proving this. We follow the ideas from [Gur+15] that show low support concentration. We write the coefficients of A' as linear combinations of the coefficients of A. Since they are linear combinations, there exists a transfer matrix $D^{-1}TD$. To study the coefficients of the cs $\leq k$ monomials of the shifted polynomial, we truncate the matrices in Equation 5.3 appropriately. Then we prove a combinatorial lemma regarding the rank of the coefficient matrix and plug that into the proof, to get cone-size concentration. We now present the statement more formally.

Theorem 11. *Let $A(\boldsymbol{x}) \in \mathbb{A}_k[\boldsymbol{x}]$. Let $\boldsymbol{w}$ be a basis-isolating weight assignment for $A(\boldsymbol{x})$. Then $A(\boldsymbol{x} + t^{\boldsymbol{w}})$ is cs $\leq k$ concentrated.*

Proof. Let $A'(\mathbf{x}) = A(\mathbf{x} + t^{\mathbf{w}})$. Now we consider equation (5.3) with respect to monomials with cone-size $\leq k$. We define $M_k \coloneqq \{\mathbf{a} \in M | cs(\mathbf{a}) \leq k\}$. We define matrices

C'_k : $M_k \times [k]$ sub-matrix of C' that contains coefficients of A' of cone-size $\leq k$

T_k : $M_k \times M$ sub-matrix of M restricted to the rows $\mathbf{a} \in M_l$

D_k : $M_k \times M_k$ sub-matrix of D restricted to the rows and columns of M_k

To show that A' is cs $\leq k$-concentrated, we need to prove that $\text{rank}(C'_k) = \text{rank}(C)$. By equation (5.3), we know that $C'_k = D_k^{-1}T_kDC$. Now D_k has full rank and hence D_k^{-1} as well. Hence, we just need to show that $\text{rank}(T_kDC) = \text{rank}(C)$.

W.l.o.g. we assume that the order of the rows and columns in all the above matrices that are indexed by M or M_k is according to increasing weight $\mathbf{w}(\mathbf{a})$ of the indices a. The rows with the same weight can be arranged in an arbitrary order.

Now, recall that $\mathbf{w}$ is a basis isolating weight assignment. Hence, there exists a set $S \subseteq M$ such that the coefficients $\text{coeff}_A(\mathbf{b})$, for $b \in S$, span all coefficients $\text{coeff}_A(\mathbf{a})$, for $a \in M$. In terms of the coefficient matrix C, for any $a \in M$ we can write

$$C(\mathbf{a}, \cdot) \in \text{span}\{C(\mathbf{b}, \cdot) | b \in S \textbf{ and } \mathbf{w}(\mathbf{b}) \leq \mathbf{w}(\mathbf{a})\} \tag{5.4}$$

Let $S = \{s_1, ..., s_{k'}\}$ such that $k' \leq k$. Let C_0 be the $k' \times k$ sub-matrix of C, such that $C_0(i, \cdot) = C(s_i, \cdot)$. Now by (5.4), we know for all $\mathbf{a} \in M$, there exists $\gamma_{\mathbf{a}} = \{\gamma_{\mathbf{a},1}, ..., \gamma_{\mathbf{a},k'}\} \in \mathbb{F}^{k'}$ such that

$$C(\mathbf{a}, \cdot) = \sum_{j=1}^{k'} \gamma_{\mathbf{a},j} C_0(j, \cdot) \tag{5.5}$$

Now, we take $\Gamma = (\gamma_{\mathbf{a},j})$ be the $M \times [k']$ matrix and this gives us the equations

$$C = \Gamma C_0 \tag{5.6}$$

Observe that the s_i-th row of Γ is simply e_i. By (5.4), $C(s_i, \cdot)$ is used to express $C(\mathbf{a}, \cdot)$ only when $\mathbf{w}(\mathbf{a}) > \mathbf{w}(s_i)$. Recall that the rows of the matrices indexed by M, like Γ, are in order the of increasing weight of the index. Therefore, when we consider the i-th column of Γ from the top, the entries are all zero down to row s_i where we hit on the one from e_i,which we can write as

$$\Gamma(s_i, i) = 1 \text{ and } \forall a \neq s_i,\ \mathbf{w}(\mathbf{a}) \leq \mathbf{w}(s_i) \implies \Gamma(\mathbf{a}, i) = 0 \tag{5.7}$$

Now, we want to show that $\text{rank}(T_k DC) = \text{rank}(C)$. Now, we know that $\text{rank}(C_0) = \text{rank}(C)$. Hence, we just need to show that $R \coloneqq T_k D\Gamma$ has full column rank k'.

Expanding the product by column, we get

$$R(\cdot, j) = \sum_{\mathbf{a} \in M} T_k(\cdot, \mathbf{a}) \Gamma(\mathbf{a}, j) t^{\mathbf{w}(\mathbf{a})} \tag{5.8}$$

Now, the term with the lowest degree in t in $R(\cdot, j)$ is $t^{\mathbf{w}(s_j)}$. By $\text{lc}(R(\cdot, j))$, we denote the coefficient of the lowest degree term in the polynomial $R(\cdot, j)$. Now, because $\Gamma(s_j, j) = 1$, we have

$$\text{lc}(R(\cdot, j)) = T_k(\cdot, s_j) \tag{5.9}$$

We denote the $M_k \times [k']$ matrix R_0 such that $R_0(\cdot, j) = T_k(\cdot, s_j)$. We will show that the columns of the matrix T_k are linearly independent.Therefore the k' columns of R_0 are linearly independent.

Hence, there are k' rows in R_0 such that its restriction to these rows, say R'_0 is a square matrix with $\det(R'_0) \neq 0$. Let R' be the restriction of R to the same rows. Now we can observe that $\mathrm{lc}(\det(R')) = \det(R'_0)$. This is because the lowest degree term in $\det(R')$ has degree $\sum_{j=1}^{k'} \mathbf{w}(s_j)$ and this can only be obtained if the degree $\mathbf{w}(s_j)$ term has been taken from the j-th column. Hence, we conclude that $\det(R') \neq 0$ and hence, R has full column rank. □

Theorem 12. *Let M be the set of all monomials of degree $\leq d$ in n variables and let $S \subseteq M$ be any set of k distinct monomials in M. Define M_k to be the set of all cone-size $\leq k$ monomials in M, $M_k = \{\bar{a} \in M \mid cs(\bar{a}) \leq k\}$. Then, the multinomial matrix T defined by $T(\bar{a}, \bar{b}) = \binom{\bar{b}}{\bar{a}}$, $\bar{a} \in M_k$ and $\bar{b} \in S$, is full rank.*

The proof of Theorem 12 is in the form of Lemma 25. The multinomial matrix T is exactly the transfer matrix of a k-sparse polynomial shifted by $\bar{1}$. Suppose T is not full column rank, then we will get a non trivial right null vector. Construct an adversarial sparse polynomial with the elements of null vector as its coefficients. Then this sparse polynomial will contradict Lemma 25.

Lemma 25. *Let $f(\bar{x}) \in \mathbb{F}[\bar{x}]$ be a non-zero polynomial of sparsity at most k. Then $f'(\bar{x}) = f(\bar{x} + \bar{1})$ has a monomial of cone-size $\leq k$ with non-zero coefficient.*

Proof. Proof is by induction on number of variables n.

Base Case: The $n = 1$ case is proved by Equation 5.10. Let $f = \beta_1 x^{j_1} + \ldots + \beta_k x^{j_k}$. For the sake of contradiction, suppose that all $cs \leq k$ monomials in $f' = f(\bar{x} + \bar{1})$ vanish, then we get $T\bar{\beta} = \bar{0}$, where T is the binomial matrix in Equation 5.10. We will show T is a square matrix with full rank, which means $\bar{\beta} = \bar{0}$, thus contradicting non-zeroness of f.

That T is full rank, is in itself an interesting fact. We shall prove it again here, for the sake of completeness. Here, T will be a square matrix of size $k \times k$, with rows indexed by $M_k = \{x^0, x^1, \ldots, x^{k-1}\}$ and columns indexed by $B = \{x^{j_1}, \ldots, x^{j_k}\}$. To show that T has full rank, we will prove that T does not have a non-zero left null vector. Let $\bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_k) \in \mathbb{F}^{1\times k}$ be an arbitrary non-zero vector. We need to show that $\bar{\alpha} \cdot T \neq \bar{0}$.

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_k \end{bmatrix} \begin{bmatrix} \binom{j_1}{0} & \binom{j_2}{0} & \cdots & \binom{j_k}{0} \\ \binom{j_1}{1} & \binom{j_2}{1} & \cdots & \binom{j_k}{1} \\ & & \ddots & \\ \binom{j_1}{k-1} & \binom{j_2}{k-1} & \cdots & \binom{j_k}{k-1} \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix} \tag{5.10}$$

To show that the above equation cannot hold, we construct an auxiliary polynomial

$$f(y) = \sum_{i=1}^{k} \alpha_i \frac{y(y-1)\ldots(y-i+2)}{(i-1)!}$$

Since α is a non-zero vector, $f(y)$ is a non-zero polynomial. Notice that f has degree bounded by $(k-1)$, and thus cannot have more than $(k-1)$ roots. But if equation 5.10 holds, then $f(y)$ has k roots, namely $j_1, \ldots, j_k$, which is a contradiction. Hence, T is full rank.

Induction Step: Rewrite f wrt last variable as

$$f = \sum_{i=0}^{d} g_i x_n^i$$

where $g_i \in \mathbb{F}[x_1, \ldots, x_{n-1}]$. Let m be the count of non-zero g_i's. Note that

$$sp(g_0) + \ldots + sp(g_d) \leq k$$

, where $sp(g)$ denotes sparsity of g. By averaging argument, there exists a $g_j, j \in [0, d]$ with $sp(g_j) \leq k/m$.

Applying induction hypothesis, there exists a monomial in $n-1$ variables, say $\bar{x}_{n-1}^{\bar{a}}$, in

$g'_j = g_j(\bar{x} + \bar{1})$ with non-zero coefficient and of cone-size $\leq k/m$. Now, we can write the shifted polynomial f' as

$$\begin{aligned} f'(\bar{x}) &= \sum_{i=0}^{d} g'_i(x_n + 1)^i \\ &= \bar{x}_{n-1}^{\bar{a}} \cdot \big(\sum_{i=0}^{d} c_i(x_n + 1)^i \big) + \text{ other monomials} \end{aligned}$$

Here, c_i's are field constants and number of non-zero c_i's are at most m. Also, $c_j \neq 0$, thus $\big(\sum_{i=0}^{d} c_i(x_n + 1)^i \big)$ is a non-zero univariate polynomial with sparsity bounded by m. By the base case, it has a univariate monomial $x_n^{j_n}$ of $cs \leq m$, that is $j_n \leq m - 1$. Thus, we have isolated the monomial $\bar{x}_{n-1}^{\bar{a}} x_n^{j_n}$ in f' with cone-size $\leq k/m \cdot (j_n + 1) \leq k$. □

Chapter Six

Conclusion and Future Work

In the third and fourth chapters, we tried to set up a new proof strategy for a poly-time PIT for log-variate ROABPs. We analyse the notion of cone-size hypothesis for the following shifts:

- Shift with $t = (t, t, t, ..., t)$
- Shift with $t = (t, t_2(t), ..., t_n(t))$ where t_i are sparse-PIT maps

For the shift with $t = (t, ..., t)$ we prove a weaker structural characterization for polynomials satisfying the cone-size hypothesis and we prove corresponding width lower bounds for polynomials satisfying the structural characterization. For constant-variate ROABPs, this helps us to give a poly(k) PIT algorithm.

We try to prove a stronger characterization for the polynomials that satisfy the cone-size hypothesis. We are able to prove this for the trivariate case. A future direction of research will be to extend the proof strategy to the general case. Replicating this proof idea becomes difficult because applying the atomic operation on a $P_{\mathbf{e}}$ which has cone-size $> k$ might create a $P_{\mathbf{e}'}$ which has cone-size $\leq k$. And the induction step fails. But there might be another interesting way to replace that by permutation operations and we will notice that their differences are also of the same form. That might help in proving the general case.

We show the width lower bound for a specific class of polynomials that satisfy this stronger

structural hypothesis. This proof strategy might be interesting in proving width lower bounds for polynomials satisfying this stronger structural hypothesis.

We show a similar structural characterization for polynomials that satisfy the cone-size hypothesis when the shift is by sparse-PIT maps. We prove a corresponding width result as well. A similar strategy can be used to first prove width lower bounds for the case when the span is over the underlying field $\mathbb{F}$ and then that can be extended to the general case (over $\mathbb{F}[x_1]$).

We also give a simpler proof for the interesting structural result for Basis Isolating Weight Assignments (BIWA) i.e a polynomial when shifted by a BIWA achieves cone-size concentration. This is achieved by proving an interesting combinatorial result regarding the rank of matrices of binomial coefficients.

References

[Agr+15] Manindra. Agrawal et al. "Hitting-Sets for ROABP and Sum of Set-Multilinear Circuits". In: *SIAM Journal on Computing* 44.3 (2015), pp. 669–697. DOI: 10.1137/140975103. eprint: https://doi.org/10.1137/140975103. URL: https://doi.org/10.1137/140975103.

[AGS18] Manindra Agrawal, Sumanta Ghosh, and Nitin Saxena. "Bootstrapping Variables in Algebraic Circuits". In: *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2018. Los Angeles, CA, USA: Association for Computing Machinery, 2018, pp. 1166–1179. ISBN: 9781450355599. DOI: 10.1145/3188745.3188762. URL: https://doi.org/10.1145/3188745.3188762.

[AKS02] Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. "Primes is in P". In: *Annals of Mathematics* 160 (Sept. 2002). DOI: 10.4007/annals.2004.160.781.

[ASS13] Manindra Agrawal, Chandan Saha, and Nitin Saxena. "Quasi-Polynomial Hitting-Set for Set-Depth- Formulas". In: *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing*. STOC '13. Palo Alto, California, USA: Association for Computing Machinery, 2013, pp. 321–330. ISBN: 9781450320290. DOI: 10.1145/2488608.2488649. URL: https://doi.org/10.1145/2488608.2488649.

[BS20] Pranav Bisht and Nitin Saxena. "Poly-time blackbox identity testing for sum of log-variate constant-width ROABPs". In: *Electronic Colloquium on Computational Complexity (ECCC)* 27 (2020), p. 42.

[DL78] Richard A. DeMillo and Richard J. Lipton. "A Probabilistic Remark on Algebraic Program Testing". In: *Inf. Process. Lett.* 7 (1978), pp. 193–195.

[FGS18] Michael A. Forbes, Sumanta Ghosh, and Nitin Saxena. "Towards Blackbox Identity Testing of Log-Variate Circuits". In: *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*. 2018, 54:1–54:16. DOI: 10.4230/LIPIcs.ICALP.2018.54. URL: https://doi.org/10.4230/LIPIcs.ICALP.2018.54.

[FGT16] Stephen Fenner, Rohit Gurjar, and Thomas Thierauf. "Bipartite Perfect Matching is in Quasi-NC". In: *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*. STOC '16. Cambridge, MA, USA: Associa-

tion for Computing Machinery, 2016, pp. 754–763. ISBN: 9781450341325. DOI: 10.1145/2897518.2897564. URL: https://doi.org/10.1145/2897518.2897564.

[For14a] Michael A. Forbes. "Polynomial identity testing of read-once oblivious algebraic branching programs". PhD thesis. Massachusetts Institute of Technology, Cambridge, MA, USA, 2014. URL: http://hdl.handle.net/1721.1/89843.

[For14b] Michael Andrew Forbes. *Polynomial Identity Testing of Read-Once Oblivious Algebraic Branching Programs*. 2014.

[FSS14] Michael A. Forbes, Ramprasad Saptharishi, and Amir Shpilka. "Hitting Sets for Multilinear Read-Once Algebraic Branching Programs, in Any Order". In: *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*. STOC '14. New York, New York: Association for Computing Machinery, 2014, pp. 867–875. ISBN: 9781450327107. DOI: 10.1145/2591796.2591816. URL: https://doi.org/10.1145/2591796.2591816.

[Gho19] Sumanta Ghosh. "Low Variate Polynomials: Hitting Set and Bootstrapping". PhD thesis. INDIAN INSTITUTE OF TECHNOLOGY KANPUR, 2019.

[Guo+19] Zeyu Guo et al. "Derandomization from Algebraic Hardness*: A borderless version". In: 2019.

[Gur+15] Rohit Gurjar et al. "Deterministic Identity Testing for Sum of Read-Once Oblivious Arithmetic Branching Programs". In: *Proceedings of the 30th Conference on Computational Complexity*. CCC '15. Portland, Oregon: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015, pp. 323–346. ISBN: 9783939897811.

[HS80] J. Heintz and C. P. Schnorr. "Testing Polynomials Which Are Easy to Compute (Extended Abstract)". In: *Proceedings of the Twelfth Annual ACM Symposium on Theory of Computing*. STOC '80. Los Angeles, California, USA: Association for Computing Machinery, 1980, pp. 262–272. ISBN: 0897910176. DOI: 10.1145/800141.804674. URL: https://doi.org/10.1145/800141.804674.

[KI03] Valentine Kabanets and Russell Impagliazzo. "Derandomizing Polynomial Identity Tests Means Proving Circuit Lower Bounds". In: *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*. STOC '03. San Diego, CA, USA: Association for Computing Machinery, 2003, pp. 355–364. ISBN: 1581136749. DOI: 10.1145/780542.780595. URL: https://doi.org/10.1145/780542.780595.

[KSS14] Swastik Kopparty, Shubhangi Saraf, and Amir Shpilka. "Equivalence of Polynomial Identity Testing and Deterministic Multivariate Polynomial Factorization". In: June 2014, pp. 169–180. ISBN: 978-1-4799-3626-7. DOI: 10.1109/CCC.2014.25.

[KST18] Mrinal Kumar, Ramprasad Saptharishi, and Anamay Tengse. "Near-optimal Bootstrapping of Hitting Sets for Algebraic Models." In: *arXiv: Computational Complexity* (2018).

[Lov79] László Lovász. "On determinants, matchings, and random algorithms". In: *FCT*. 1979.

[Mul12] Ketan Mulmuley. "Geometric Complexity Theory V: Efficient algorithms for Noether Normalization". In: *arXiv: Computational Complexity* (2012), pp. 225–309.

[Nis91] Noam Nisan. "Lower Bounds for Non-Commutative Computation". In: *Proceedings of the Twenty-Third Annual ACM Symposium on Theory of Computing*. STOC '91. New Orleans, Louisiana, USA: Association for Computing Machinery, 1991, pp. 410–418. ISBN: 0897913973. DOI: 10.1145/103418.103462. URL: https://doi.org/10.1145/103418.103462.

[RS04] Ran Raz and Amir Shpilka. "Deterministic polynomial identity testing in non commutative models". In: *Computational Complexity* 14 (2004), p. 2005.

[Sax08] Nitin Saxena. "Diagonal Circuit Identity Testing and Lower Bounds". In: *Automata, Languages and Programming*. Ed. by Luca Aceto et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 60–71. ISBN: 978-3-540-70575-8.

[Sch80] J. T. Schwartz. "Fast Probabilistic Algorithms for Verification of Polynomial Identities". In: *J. ACM* 27.4 (Oct. 1980), pp. 701–717. ISSN: 0004-5411. DOI: 10.1145/322217.322225. URL: https://doi.org/10.1145/322217.322225.

[Tut47] W. T. Tutte. "The Factorization of Linear Graphs". In: *Journal of The London Mathematical Society-second Series* (1947), pp. 107–111.

[Vai15] Rishabh Vaid. "Blackbox Identity Testing for Simple Depth 3 Circuits". PhD thesis. Indian Institute of Technology Kanpur, 2015.

[Zip79] Richard Zippel. "Probabilistic algorithms for sparse polynomials". In: *Symbolic and Algebraic Computation*. Ed. by Edward W. Ng. Berlin, Heidelberg: Springer Berlin Heidelberg, 1979, pp. 216–226. ISBN: 978-3-540-35128-3.