On Polynomial Identity Testing for depth-4 bounded top & bottom fanin circuits

Devansh Shringi
Advisor: Prof. Nitin Saxena

May 22, 2020

Contents

1 Notation and Definitions

We use $\mathbb{F}$ to represent fields. We will mainly be working with function fields, and the variables will be denoted mainly by $x_1, \ldots, x_n$ with n denoting the number of indeterminants. $\mathbb{F}[x_1, \ldots, x_n]$ will denote the polynomial ring over $\mathbb{F}$. Polynomials are usually denoted by $f_1, \ldots, f_m$, with m denoting the number of polynomials. Arithmetic circuits are the most natural and standard model for computing polynomials and we will be using these to represent polynomials. We use the following definition of Arithmetic circuits

Definition 1.1. (Arithmetic circuits) An arithmetic circuit C over the field $\mathbb{F}$ and the set of variables $x_1, \ldots, x_n$ is a directed acyclic graph as follows. The vertices of C are called gates. Every gate of C of in-degree 0 is labelled by either a variable or a field element. Every other gate is labelled by either $+$ or $\times$. An edge is labelled with field constants, which is 1 by default.

An arithmetic circuit computes a polynomial in a natural way : An input gate labeled by $\alpha \in \mathbb{F} \cup \{x_1, \ldots, x_n\}$ computes the polynomial α. A product gate (gate with label $\times$) computes the product of the polynomials computed by its children. Similarly a sum gate (gate with label $+$) computes the sum of the polynomials computed by its children. An example of an arithmetic circuit is given below.

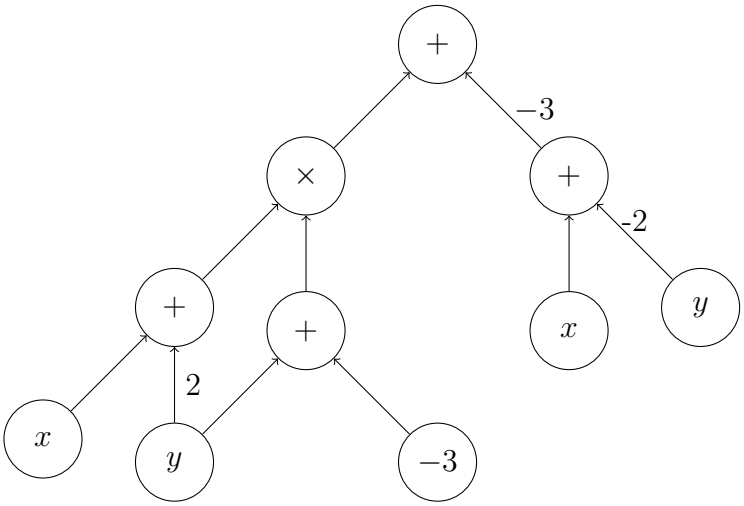

Figure 1: Circuit computing $xy + 2y^2$

We define the *size* of an arithmetic circuit to be the number of edges in the graph. We define the *depth* of a gate to be the length of the longest directed path to it. The depth of a circuit is the maximal depth of a gate in it. We refer to the input degree of a gate as its *fanin*, and output degree of a gate as its *fanout*. We can also see an arithmetic circuit as layers of $+$ and $\times$ gates, as consecutive $+$ or consecutive $\times$ gates can be combined just by increasing the fanin. Hence, a depth3 circuit is frequently represented as $\Sigma\Pi\Sigma$ or $\Pi\Sigma\Pi$. The fanins are often written in superscript, for eg., $\Sigma^k\Pi\Sigma$ represents depth3 circuit with top gate fanin k.

As we saw in the model of arithmetic circuits, the two main resources are size and depth.

Based on size, we define class VP as the family of circuits $\{C_n\}$ computing polynomials such that n is number of variables , degree and size of the circuit is bounded by $poly(n)$. The class is the arithmetic analog of P. For more details in the algebraic complexity area, refer to the survey [SY10].

We will further need a few definitions of a few terms which we give below.

Definition 1.2. (Ideal) The ideal generated by $f_1, \ldots, f_k$ is the set $\{\sum_i h_i \cdot f_i : h_1, \ldots, h_k \in \mathbb{F}[x_1, \ldots, x_n]\}$ and denoted by $\langle f_1, \ldots, f_k \rangle$.

We will denote the quotient ring of an ideal by $\mathbb{F}[x_1, \ldots, x_n]/I$. We will use I_d to denote the polynomials in I of degree d.

Definition 1.3. (Radical) The radical of an ideal I is the set $\{g \in \mathbb{F}[x_1, \ldots, x_n] : g^e \in I \text{ for some integer } e \geq 1\}$ and is denoted by $\sqrt{I}$.

There many different definitions of *variety* but we will be using the following in this report

Definition 1.4. (Variety) The variety of a set of polynomials $f_1, \ldots, f_k \in \mathbb{F}[x_1, \ldots, x_n]$ is the set of all their common zeroes in $\mathbb{F}^n$,i.e. the set $\{(a_1, \ldots, a_n) \in \mathbb{F}^n : f_1(a_1, \ldots, a_n) = \ldots = f_k(a_1, \ldots, a_n) = 0\}$. It is denoted by $V(f_1, \ldots, f_k)$.

2 Introduction

2.1 The Problem

In this report we will be studying the problem of Polynomial Identity Testing (PIT). It is the problem in which we are given a polynomial as an arithmetic circuit $C(\mathbf{x})$, over a ring R, to efficiently test whether the C is identically zero. In this report we will focus on the case of R being a field. By efficient we mean the algorithm should run in $poly(size(C))$ many $\mathbb{F}$ operations. The problem is trivial if the polynomial is given as a vector of coefficients, for which we only need to check if any of the coefficient is non-zero. It also has an easy solution for univariate polynomials, which requires it to be evaluated at $degree + 1$ many points, and it is an identity iff all the evaluations are zero. This method doesn't work for multivariate polynomials, as there can be infinite solution for a simple bivariate polynomial (eg. $xy = 0$ over $\mathbb{R}$). There are 2 versions to this problem, Blackbox PIT and Whitebox PIT. In the Blackbox version, we are only allowed evaluations of C at points from $\mathbb{F}^n$, and cannot look inside the computations at inner gates. In the whitebox version we have access to the inner gates of C. Let us give formal definition of the problem

Definition 2.1.1. [For14] **(Polynomial Identity Testing)** Let $\mathcal{C}$ be a class of circuits having $size \leq s$, which computes polynomials in $\mathbb{F}[x_1, \ldots, x_n]$ of degree $< d$. The PIT problem for this class $\mathcal{C}$ asks for a deterministic algorithm to test whether a polynomial f_C, computed by a circuit $\mathbb{C} \in \mathcal{C}$, is identically zero or not. The algorithm is considered efficient if it uses only $poly(s, n, d)\mathbb{F}$ operations.

Definition 2.1.2. [For14] **(Hitting Set)** Let $\mathcal{C}$ be a class of circuits having $size \leq s$, which compute polynomials in $\mathbb{F}[x_1, \ldots, x_n]$ of degree $< d$. A Hitting set $\mathcal{H} \subseteq \mathbb{F}^n$ for the circuit class $\mathcal{C}$ is a set of points such that if a circuit $C \in \mathcal{C}$ computes a non-zero polynomial f_C, then $\exists(\alpha_1, \ldots, \alpha_n) \in \mathcal{H}$ such that $f(\alpha_1, \ldots, \alpha_n) \neq 0$.

From it's definition giving a $poly(s, n, d)$ sized Hitting set for a circuit class $\mathcal{C}$, gives an efficient blackbox PIT for $\mathcal{C}$. It is notable that the problem of PIT has a very simple and elegant randomized solution thanks to the PIT lemma.

Lemma 2.1.3. (PIT Lemma) *(Schwartz-Zippel[Sch80])* Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a non-zero polynomial of total degree $d \geq 0$. Let S be any finite subset of $\mathbb{F}$, and let $\alpha_1, \ldots, \alpha_n$ be elements selected independently, uniformly and randomly from S. Then,

$$Pr_{\alpha_1,\ldots,\alpha_n \in S}[f(\alpha_1, \ldots, \alpha_n) = 0] \leq \frac{d}{|S|}$$

The above lemma can be easily proved inductively, with the base case being the univariate case. This puts PIT in $coRP$. The problem of derandomizing PIT, so as to put it in P is still open. One trivial derandomization is to check $(d+1)^n$ many points, but is inefficient. It is formally stated below

Lemma 2.1.4. (Combinatorial Nullstellensatz) *[AT99]* Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a non-zero polynomial of individual degree d. Let S be a set of distinct values of size $> d$. Then, there exists $(\alpha_1, \ldots, \alpha_n) \in S^n$ such that $f(\alpha_1, \ldots, \alpha_n) \neq 0$.

We will look in this report at a special case where the circuits will be have depth 4. As per results in [AV08], solving the problem for depth4 circuits gives us solution for PIT of all circuits in VP. Therefore, we look at an even more restricted case with the top and bottom fanin are also just $O(1)$. The depth 4 circuits can be of two types $\Sigma\Pi\Sigma\Pi$ and $\Pi\Sigma\Pi\Sigma$. The case of $\Pi\Sigma\Pi\Sigma$ is reduced to simply checking the smaller $\Sigma\Pi\Sigma$ circuits which multiply in the final multiplication gate, which itself is a different problem of depth-3 circuits. Hence, we look at only inputs of the form $\Sigma^k\Pi\Sigma\Pi^r$, where $k \geq 3 = O(1)$ and $r \geq 2 = O(1)$. The case of $k = 2$ is solved in whitebox case by division of the common factors in both terms and just comparing the left constants (as $\mathbb{F}[x_1, \ldots, x_n]$ is Unique Factorization Domain). The Blackbox case for the problem is open for even $k = 2$. While, the case of $r = 1$ is the case of depth-3 constant top fanin which is discussed in section 3.1. Let us give a formal definition of the bounded case:

Consider the input circuit be C such that it has a form $C = \Sigma^k\Pi\Sigma\Pi^r$, i.e. the circuit has alternate $+$ and $\times$ gates where the fanin of the top $+$ gate is $\leq k$ and the fanin of the bottom $\times$ gate is $\leq r$. Such a circuit C computes the polynomial of the form

$$C(x_1, \ldots, x_n) = \sum_{i=1}^{k} T_i = \sum_{i=1}^{k} \prod_{j=1}^{d_i} L_{ij}$$

where d_i is the fanin of the i^{th} $\times$ gate on the second level. The circuit is said to be *simple* if $gcd(T_1, \ldots, T_k) = 1$. It is *minimal* if no proper subset of $T's$ sum upto 0, i.e. for every $\phi \subset A \subset [k] : \sum_{i \in A} T_i \neq 0$. We assume the input circuit to be simple and minimal as if it's not simple it breaks into 2 different smaller simpler circuits by taking out the gcd from all terms, and if it's not minimal we can decrease the top fan-in. To even simplify the circuits further we assume that it is a *homogeneous* circuit, i.e. all the $T_i's$ are homogeneous of the same degree (and therefore L_{ij} are homogeneous).

2.2 The Motivation

The problem of PIT has many applications like the problem of deciding existence of perfect matching in a graph efficiently can be seen as a question of finding efficient PIT algorithm for the determinant polynomial of the graph's Tutte matrix. The idea of PIT was also very useful in the proof of the complexity result $IP = PSPACE$. Even the problem of Primality testing was solved by working with a PIT formulation. It was observed that a positive integer n is prime iff $(x+1)^n = (x^n+1)(mod n)$, which can be considered as checking $P(x) = (x+1)^n - (x^n+1)$ to be identity over $\mathbb{Z}/n\mathbb{Z}$. The problem of derandomizing PIT has relations to complexity results like $PIT \in P \implies NEXP \not\subseteq P/poly$ or $VP \neq VNP$. For more details into PIT and it's application, look at the surveys [Sax09], [Sax14] and [SY10].

The depth reduction results in algebraic complexity have brought the computation of any polynomial in VP to computation by circuits of just depth3. The case of depth3 bounded top fanin case has been solved for both whitebox version in [KS07] and blackbox in [SS12]. Also by results in [AV08], an efficient hsg for $\Sigma^s \wedge^{\omega(1)} \Sigma^s \Pi^{O(\log s)}$ gives an $n^{O(\log n)}$-hsg for VP. Thus, this is one of the open cases near the general case but is also close to already solved cases.

3 Previous Work

We will look at the case of depth3 circuits with bounded top fanin which have been solved in both blackbox and whitebox versions, as it is closely related to our case of bounded top and bottom fan-in depth 4 circuits (if we fix the bottom $fanin = 1$, it becomes the depth 3 case). After that we will have look at the work done in [Gup14] where he proposes conjectures about Sylvester-Gallai Type theorems that can solve the problem completely. We will also look in the work done in [Shp19] in proving one of the weaker conjectures.

3.1 The Depth 3 Circuits

A lot of work has been done for this case, which eventually was solved with a Blackbox poly-time algorithm in [SS12]. We will sketch here the basic idea of the proof.

The basic idea is to first prove that if $C \neq 0$, then there exists an ideal I with generators of $k-1$ elements chosen from $L'_{ij}s$ such $C \not\equiv 0 \mod I$. Then we will use a variable reduction that conserves the ideal membership, and decreases the number of variables to k, after which the identity testing can be done by brute-forcing PIT lemma. We assume the input circuit to be simple and minimal. That means we can assume the terms $T_1, T_2, \ldots, T_k$ are linearly independent.Let us assume $C \neq 0$, then we consider $C \mod \langle T_1 \rangle$. Since the terms $T_1, T_2, \ldots, T_k$ are linearly independent, $C \not\equiv 0 \mod \langle L_{11}^{e11} L_{12}^{e12} \ldots L_{1d_1}^{e1d_1} \rangle$. By Chinese remainder theorem, we have $C \not\equiv 0 \mod \langle L_{11}^{e11} \rangle$ or $\langle L_{12}^{e12} \rangle$ or $\ldots$ or $\langle L_{1d_1}^{e1d_1} \rangle$. Say $f_1 = L_{1p_1}^{e1p_1}$ is one such polynomial whose ideal doesn't contain C. This means $T_2 + T_3 + \ldots + T_k \not\equiv 0 \mod \langle f_1 \rangle$. Now in this smaller problem we consider the coprime factors of T_2 modulo f_1. Again by Chinese remainder theorem there has to be one polynomial f_2 in these coprime factors such that $T_3 + \ldots + T_k \not\equiv 0 \mod \langle f_1, f_2 \rangle$. And so on, till only the last term remains, that is for the selected $f_1, \ldots, f_{k-1}$ we have $T_k \not\equiv 0 \mod \langle f_1, f_2, \ldots, f_{k-1} \rangle$. This selection of polynomials is termed a *path*. We will formally define these terms

Definition 3.1.1. A *Path* ($\overline{p}$) with respect to an ideal is a sequence of terms $\{p_1, p_2, \ldots, p_b\}$ (these are products of linear forms) with the property that each p_i divides T_i, and each p_i is a **node** of T_i with respect to the ideal $\langle I, p_1, \ldots, p_{i-1} \rangle$. p_i is a node when some non-zero constant multiple of p_i is identical to a power-of-a-linear-form $\mod radsp(\langle I, p1, p2, ..., p_{i-1} \rangle$, where *radsp* is the ideal generated by the set of all the linear polynomials that divide $p_j, j \in [i-1]$ and the generators of I.This means p_1 is a node of T_1 wrt $\langle I \rangle$, p_2 is node of T_2 wrt $\langle I, p_1 \rangle$, and so on.

The above definition with the chinese remaindering idea discussed above is formalized as the following theorem in [SS12][Theorem 25]

Theorem 3.1.2. (Certificate for non-identity) Let I be an ideal generated by some multiplication terms. Let $C = \sum_{i \in [k]} T_i$ be a depth-3 circuit that is nonzero $\mod I$. Then $\exists i \in [k-1]$ such that $\sum_{j \in [i]} T_j \mod I$ has a path $\overline{p}$ satisfying $C \equiv \alpha \cdot T_{i+1} \not\equiv 0 (\mod I + \langle \overline{p} \rangle)$ for some $\alpha \in \mathbb{F}^*$.

For better understanding, we give an example (from notes of CS748-IITK) of the above concept.

Consider the input

$$C = \underbrace{x_1^2 x_3 x_4}_{T_1} - \underbrace{x_2(x_2+2x_1)(x_3-x_1)(x_4+x_2-x_1)}_{T_2} + \underbrace{(x_2+x_1)^2(x_3+4x_1)(x_4+x_2)}_{T_3}$$

which is $\Sigma^3\Pi^4\Sigma^4$ circuit. We as discussed go modulo T_1, and as the terms are linearly independent

$C \not\equiv 0 \mod \langle x_1^2 x_3 x_4 \rangle \implies C \not\equiv 0 \mod \langle x_1^2 \rangle$ *or* $\langle x_3 \rangle$ *or* $\langle x_4 \rangle$

Let us say we chose $f_1 = x_1^2$. We have $C \not\equiv 0 \mod \langle x_1^2 \rangle \implies T_2 + T_3 \not\equiv 0 \mod \langle x_1^2 \rangle$. Here we consider the coprime factors of T_2 modulo f_1, which becomes the set $S = \{x_2(x_2 + 2x_1), (x_3 - x_1), (x_4 + x_2 - x_1)\}$. Again, we have

$$C \not\equiv 0 \mod \langle f_1 \rangle + \langle \text{ one of } S \rangle$$

Let's say we chose $f_2 = x_3 - x_1$. Now all we have to do is check for all such possibilities of f_1, f_2, that $T_3 \equiv 0 \mod \langle f_1, f_2 \rangle$, and if it is non-zero for any one, we have the non-identity certifying path, otherwise we are sure by the chinese remaindering discussion above that $C = 0$.

The idea of the whitebox algorithm in [KS07] was to use a linear map that would take these $f_1, \ldots, f_{k-1}$ to $x_1, \ldots, x_{k-1}$, and multiply out the terms with only these variables as all other terms will be non-zero divisors with respect to this ideal, in $poly(d^k)$ operations. Further they looked at all the d^{k-1} possibilities for $f_1, \ldots, f_{k-1}$. Thus, it gave a $poly(n, d^k)$ time whitebox algorithm. For the blackbox case, it uses the fact that for the product of linear forms(T) will lie in the ideal generated by linear forms only when one of the factors of T will lie in the radsp of the generators. The rank of the set S_0 of the linear polynomials that divide the nodes in the path $\overline{p}$ is $< k$(since the path length is below k). T_{i+1} factors into atmost d linear polynomials, denoted by S_1. So, if we apply a variable reduction map that preserves the rank of each of the d sets $S_0 \cup \{l\}, l \in S_1$, we will ensure that the element from S_1 if it's not in the radsp of nodes in S, it will not be there after application of the map, hence preserving ideal non-membership. For this we will look into the

Definition 3.1.3. Vandermonde map We define a homomorphism Ψ_β, for a $\beta \in \mathbb{F}$, as:

$$\forall i \in [n], \Psi_\beta : x_i \to \sum_{j=1}^{k} \beta^{ij} y_j$$

and $\Psi_\beta(\alpha) = \alpha$ for all $\alpha \in \mathbb{F}$.

We have the nice property of Ψ_β, which allows us to preserve rank

Lemma 3.1.4. (Ψ_β preserves k-rank) Let S be a subset of linear forms in $\mathbb{F}[x_1, \ldots, x_n]$ with $rank(S) \leq k$, and $|\mathbb{F}| > nk^2$. Then $\exists \beta \in \mathbb{F}$, $rank(\Psi_\beta(S)) = rank(S)$.

The above means that the number of "bad values" of β, i.e. values for which the rank is not preserved is $\leq nk^2$. So, if we use more than nk^2 values for β one of them will definitely preserve the rank for S. As discussed earlier, the $rank(S_0 \cup \{l\}, l \in S_1) \leq k$ means that we can use Ψ_β to preserve the sets ranks, and therefore preserve ideal non-membership. Finally, we have achieved a map Ψ_β by looking at $> nk^2$ values for β, which reduces the number of variables from n to k, and also preserves ideal non-membership for any path of length less than $k-1$, meaning it preserves the non-zeroness of the circuit. Once this is done we can simply use a brute-force hitting set to give a $poly(n, d^k)$ blackbox PIT algorithm.

3.2 The Sylvester-Gallai Approach

Sylvester-Gallai theorem is a famous theorem in incidence geometry, which is stated below

Theorem 3.2.1. Given a finite number of non-collinear points S in the plane $\mathbb{R}^2$, there always exists a line which passes through exactly two points in S.

The above has a simple proof from geometry. It has a higher dimensional generalization that says:

Theorem 3.2.2. Let S be a finite set of points spanning an affine space $V \subseteq \mathbb{R}^n$ such that $dim(V) \geq 2t$. Then, there exists $(t+1)$ points in S that span a t dimensional affine space $H \subset V$ such that $|H \cap S| = t+1$.

The case of depth3 PIT algorithm led to development of lot of new techniques. Karin and Shplika showed that if we have a rank bound of $R(k,d)$ for minimal,simple $\Sigma\Pi\Sigma(n,k,d)$ identities then we have a blackbox PIT algorithm with $poly(n, d^{R(k,d)})$ many field operations. By *rank* of a depth3 circuit we mean the dimension of the vector space spanned by the linear polynomials that appear int he multiplication terms. In [DS07] it was showed that a minimal and simple depth-3 identity has rank at most $\log^k d$. Sylvester-Gallai theorems were used in attempts to get a good bound on the rank of identities. In the case of depth3 circuits The space S is the set of all linear forms that appear in the circuit, hence $dim(V) = rank(C)$. Kayal and Saraf [KS09] used Theorem 3.2.2 to get a bound on the rank of minimal,simple $\Sigma\Pi\Sigma(n,k,d)$ identity.

We will look at the work done by Gupta in [Gup14] to solve the case of PIT of depth 4 bounded top and bottom fanin circuits. It gives solution to the special case when one of the terms T_i doesn't lie in the radical generated by other terms . For the other cases, which is referred to as the *Sylvester-Gallai* configuration, he proposes conjectures for higher degree polynomials with bounds on transcendence degree similar to the results known for linear polynomials and their rank. We first define the Sylvester-Gallai configuration of the PIT depth4 case.

Definition 3.2.3. (SG-$\Sigma^k\Pi\Sigma\Pi^r$ circuits A simple, minimal, homogeneous $\Sigma^k\Pi\Sigma\Pi^r$ circuit is SG if

$$\forall i \in [k] \bigcap_{j \in [k] \setminus \{i\}} V(T_j) \subseteq V(T_i)$$

By Hilbert's Nullstellensatz, over $\mathbb{C}$ this equivalent to

$$\forall i \in [k] \qquad T_i \in \sqrt{\langle T_1, \ldots, T_{i-1}, T_{i+1}, \ldots, T_k \rangle}$$

Gupta in his paper gave an algorithm to identify if a depth4 circuit is an SG circuit or not for circuits that work in the Complex field($\mathbb{C}$). If a circuit is an identity the sum of all terms is zero,so we know $\forall i \in [k], T_i \in \langle T_1, \ldots, T_{i-1}, T_{i+1}, \ldots, T_k \rangle$, and hence definitely lie in it's radical. Thus, all Identities are SG circuits, but the reverse is not true. Gupta gives an algorithm to identify the cases of non-identity in which one of the terms doesn't lie in the radical generated by other terms, i.e. without the loss of generality we have $T_k \notin \sqrt{\langle T_1, \ldots, T_{k-1} \rangle}$. The proof is based on the following proposition:

Proposition 3.2.4. Let $P_1, \ldots, P_d, L_1, \ldots, L_k \in \mathbb{C}[x_1, \ldots, x_n]$ be homogeneous and degree of each L_i is at most r. Then,

$$P_1 \cdots P_d \in \sqrt{\langle L_1, \ldots, L_k \rangle} \iff \exists \{i_1, \ldots, i_{r^k}\} \subseteq [d] : P_{i_1} \cdots P_{i_{r^k}} \in \sqrt{\langle L_1, \ldots, L_k \rangle}$$

Proof: The reverse direction of the proof is obvious from the definition of radical of an ideal. For the forward direction assume, $V_1 \cup V_2 \cup \ldots \cup V_t$ be the minimal decomposition of $V(L_1, \ldots, L_k)$ where V_i's are irreducible. Then by Nullstellensatz,

$$V_1 \cup V_2 \cup \ldots \cup V_t \subseteq V(P_1) \cup \ldots \cup V(P_d)$$

As V_i's are irreducible, we have that for each i there is an $i_j \in [d]$ such that $V_i \subseteq V(P_{i_j})$. Therefore

$$V(L_1, \ldots, L_k) = V_1 \cup V_2 \cup \ldots \cup V_t \subseteq V(P_{i_1}) \cup \ldots \cup V(P_{i_t})$$

which by Nullstellensatz means $P_{i_1} \cdots P_{i_t} \in \sqrt{\langle L_1, \ldots, L_k \rangle}$. The number of irreducible components of a variety is bounded by it's cumulative degree which is the sum of all it's irreducible components. By Bezout's Theorem, cumulative degree of $V(L_1, \ldots, L_k)$ is at most $\Pi_i deg(L_i)$. Hence, $t \leq r^k$. In simpler words the proposition says that if a product of polynomials lies inside the radical generated by k polynomials of degree atmost r, then the product of the elements of a subset of size r^k from the product will also lie in radical. This enables us to create whitebox algorithm straight by checking all r^k subsets, which since both r and k are constant will be polynomial in d and also there products degree will be small$(= r^{k+1})$. For the blackbox algorithm of the same Gupta proves that radical non-membership is preserved under random linear projections, which allows to decrease the number of variables and hence gives us a poly sized hitting set.

For the SG-circuit, he proposes that the transcendence degree $(trdeg)$ is small$(O(1))$. We state the conjecture below

Conjecture 3.2.5. Let $T_1, \ldots, T_k$ be finite sets of irreducible homogeneous polynomials in $\mathbb{C}[x_1, \ldots, x_n]$ of degree $\leq r$ st. $\cap_i T_i = \phi$ and for every $k-1 L_1, \ldots, L_{k-1}$, each from a distinct set T_j being the remaining set st. $T_j \in \sqrt{\langle L_1, \ldots, L_{k-1} \rangle}$. Then, $trdeg_{\mathbb{C}}(\cup_i T_i) \leq \lambda(k, r)$ for some function λ.

The above is true for $r = 1$ and was first proved in [KS09]. He further proposed simpler conjectures to solve in which he proposed instead of product, individual polynomials to lie in the radical, and another one in which he took out the elements being from distinct set(colored version) condition. He says that these could function as stepping stones in proving the main conjecture.

3.2.1 Sylvester Gallai Type theorems for Quadratic Polynomials

Some of the conjectures proposed in [Gup14] have been proved for the case of $r = 2$ and $k = 3$, but not conjecture 3.2.5, and hence the case of PIT for even $\Sigma^3\Pi\Sigma\Pi^2$ remains open. The following theorems related to the original conjecture were proved in [Shp19]. These instead of working with the trdeg, prove the linear rank of SG circuit to be small, which is a stronger result. But these require atleast a quadratic polynomial to lie in the radical to get the bounds, instead of the products, which is a strong assumption as an counterexample can be constructed which disproves it's existence.

Theorem 3.2.1.1. Let $\{Q_1, Q_2, \ldots, Q_m\}$ be m homogeneous quadratic polynomials over $\mathbb{C}$ such that each Q_is either irreducible or a square of a linear function. Assume further that for every $i \neq j, \exists k \notin \{i, j\}$ such that whenever Q_i and Q_j vanish Q_k vanishes as well$(V(Q_i, Q_j) \subseteq V(Q_k) \implies Q_k \in \sqrt{\langle Q_i, Q_j \rangle}$. Then the linear span of Q_i's have dimension $O(1)$.

He also proved a colored version of this theorem, which is given below. Note the square of linear function allows to handle linear factors as well as keeping them homogeneous.

Theorem 3.2.1.2. Let T_1, T_2 and T_3 be finite sets of homogeneous quadratic polynomials over $\mathbb{C}$ satisfying the following properties:

- Each $Q \in \cup_i T_i$ is either irreducible or a square of a linear function.
- No two polynomials are multiples of each other (i.e., every pair is linearly independent).
- For every two polynomials Q_1 and Q_2 from distinct sets there is a polynomial Q_3 in the third set so that $Q_3 \in \sqrt{\langle Q_1, Q_2 \rangle}$

For the proof of the above the theorems, the following structural theorem is very important. After that it is just a case by case analysis to get a bound on the linear rank of the cases created by the structure theorem.

Theorem 3.2.1.3. If $Q \in \sqrt{\langle Q_1, Q_2 \rangle}$, then one of the following cases hold:

1. Q is in the linear span of Q_1, Q_2
2. There exists a non trivial linear combination of the form $\alpha Q_1 + \beta Q_2 = l^2$ where l is a linear function
3. There exist two linear functions l_1 and l_2 such that when setting $l_1 = l_2 = 0$ we get that Q, Q_1 and Q_2 vanish.

The proof of this theorem involves changing the radical membership from 2 basis elements to a single basis formed by resultant of Q_1, Q_2 and look at all the cases how Q can lie in that radical. A stronger structural theorem is suggested in a talk by Shir Peleg that deal with the product of quadratics lying in the radical, and hence proving theorem 3.2.1.1 in the case when product is in the radical. But still for a solution to the $\Sigma^3\Pi\Sigma\Pi^2$ case, it

requires a colored version of this theorem, which is not known.

Further this style of proof is very tough to extend in both cases of $k > 3$ and $r > 3$ as both will require new structure theorems for each value. Hence to solve the case of $\Sigma^{O(1)}\Pi\Sigma\Pi^{O(1)}$ it seems an unlikely method.

4 The main sub-problem

As we saw the Sylvester-Gallai approach is hard to extend for larger k and r, We will focus on extending the approach used for depth3 circuits in [SS12] to depth4 circuits. A major difference between the two cases is that for the case of an ideal generated by linear forms, for the product to lie in the ideal, atleast one of the factors should lie in the ideal too. Hence a variable reduction preserving k rank space($k-1$ ideal generators, 1 linear factor) was sufficient, and hence gave a homomorphism that preserved ideal membership. To extend, this approach we need to understand more about the when a product lies in the ideal generated by polynomials of higher degree. One popular method for such ideal membership, is of using the **Gröbner's basis**, but is used mainly when polynomial to be tested is given in dense form (that is as a coefficient vector).

The main problem that we will focus in this report from now on, will be to find an efficient way to solve the ideal membership of a polynomial given as a product of degree δ polynomials, in the ideal generated by 2 degree δ irreducible polynomials.

Problem 4.1. Let $P_1, \ldots, P_d, L_1, L_2 \in \mathbb{F}[x_1, \ldots, x_n]$ be homogeneous irreducible polynomials of degree $\delta = O(1)$. Design an efficient algorithm to check if

$$P_1 P_2 \cdots P_d \in \langle L_1, L_2 \rangle$$

We aim to extend this to ideals with constant number of generators. We possibly aim to create a variable reduction similar to Vandermonde map, which preserves this ideal membership. Such a solution should allow us to solve the case when $k = 3$ and $r = O(1)$, with the restriction that the factors in the terms are degree $= r$ with multiplicity 1.

5 Computing The Gröbners Basis

Everything in this section(5) is based on reading from [BFS15].We start with the definition of a monomial ordering.

Definition 5.1. A monomial ordering is a total order on monomials that is compatible with the product and such that every nonempty set has a smallest element for the order. Such an ordering is graded if monomials of different degrees are ordered according to their degree.

Such an ordering allows us to have a leading term even in the case of multivariate polynomials. The $LT(f)$ of a polynomial f corresponds to the term(i.e. monomial multiplied with a non-zero constant) that has the largest monomial according to the given ordering. For this part we will mainly use the *grevlex* ordering which is graded ordering. The order between monomials of same degree , consider $x_\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$ and $x_\beta = x_1^{\beta_1} \dots x_n^{\beta_n}$, is given by $x_\alpha \succ x_\beta$ when the last nonzero element of $(\alpha_1 - \beta_1, \dots, \alpha_n - \beta_n)$ is negative. Thus the order among the monomials of degree d is

$$x_1^d \succ x_1^{d-1}x_2 \succ \dots \succ x_2^d \succ x_1^{d-1}x_3 \succ x_1^{d-2}x_2x_3 \succ x_1^{d-2}x_3^2 \succ \dots \succ x_n^d$$

We can now look at the definition of the Gröbner's basis, and some of it's properties that are useful to us.

Definition 5.2. A **Gröbner's basis** of an Ideal I for a given monomial ordering is a set of generators of I such that the leading monomial ideal $\langle LT(I) \rangle$, which is the ideal generated by the monomials $LT(f)$, $f \in I$.

To understand why Gröbner's basis is important to ideal membership, we define the weak remainder which can be seen as the local optimum of the division algorithm, i.e. we will reduce any polynomial as far as possible, by canceling out the monomials divisible by leading terms of generators.

Definition 5.3. Consider $f, h_1, \dots, h_m \in \mathbb{F}[x_1, \dots, x_n]$. The *weak remainder* of f with respect to the h_i is the polynomial $r = f - \sum q_i h_i$ such that no monomial of r is divisible by any $LT(h_i)$.

It should be noted that when we use the usual division algorithm for reduction, the remainder we get is the weak remainder, which may not be the remainder by the ideal. A simple example which can be seen in the uni-variate case itself is when the polynomial considered is of degree lower than the generators. Consider $f = x$ and $h_1 = x(x+1), h_2 = x(x+2)$. Clearly $\langle h_1, h_2 \rangle = \langle x \rangle$, which means the remainder is zero, but the weak remainder is x, i.e. the division algorithm gets stuck. In come Gröbner's basis, for which we can show the weak remainder is unique.

Lemma 5.4. Let $g_1, \dots, g_t$ be a Gröbner's basis for the ideal $J = \langle g_1, \dots, g_t \rangle$. Then for any f, the weak remainder wrt the $g_i's$ is unique.

Proof: Suppose $f = r + \sum q_i g_i = r' + \sum q_i' g_i$ are two weak remainder decomposition of f. Then $r - r' = \sum (q_i - q_i')g_i \in J$. As the g_i are a Gröbner basis, it follows that if $r - r'$ is non-zero then its leading monomial of $r - r'$ must be divisible by some $LT(g_i)$. But the monomials of $r - r'$ are a subset of the union of the monomials of r and r', and none of those monomials are divisible by any g_i. Thus, it follows that $r - r'$ must be zero, so $r = r'$. $\square$

In this part we will look on the part of the problem of computing the gröbner's basis of an ideal of the form $\langle L_1, L_2, \ldots, L_m \rangle$ where $L_1, L_2, \ldots, L_m$ are irreducible polynomials which form a regular sequence of degree δ, with *delta* and m both are $O(1)$. We start with the general computation of gröbner's basis for an ideal, using the Macaulay's matrix, it's running complexity which comes in terms of a bound on the degree of elements **D** inside the gröbner's basis. After this, we show a constant bound on D for the ideal generated by a regular sequence which allows the computation to occur in polynomial time. Then we present the final matrix F5 algorithm which allows computation of gröbner's basis without useless reductions to zero .

Another definition that we will need is of regular system of polynomials

Definition 5.5. The polynomial system $(f_1, f_2, \ldots, f_m)$ is regular if $\forall i \in [m], f_i$ is not a zero-divisor in the quotient ring $k[x_1, \ldots, x_n]/\langle f_1, \ldots, f_{i-1} \rangle$. This means if $\exists g$ such that $gf_i \in \langle f_1, f_2, \ldots f_{i-1} \rangle$, then $g \in \langle f_1, f_2, \ldots f_{i-1} \rangle$.

The regularity exists in our system exists as we consider the ideal generated by 2 irreducible deg δ forms, which form a regular system unless one of them is a multiple of the other. The later case is easy to handle as the ideal generated by these 2 polynomials is the same as the ideal generated by one polynomial in which presence of T_3 is easy to check as it will require one of the factors of T_3 to be present in the ideal, and since all factors are irreducibles of deg δ, one of them will be a scalar multiple of the generating polynomial. Thus, remains the case when the 2 generating polynomials form a regular system. Though as we increase the fanin of the top gate from 3, we loose this condition, and it is not guaranteed the "path" certifying non-zeroness will be a regular sequence. So for higher fanins this is not applicable.

5.1 Finding Gröbner's basis efficiently

It has been shown in [MM82] that the worst case complexity of gröbner's basis computation of the general case of polynomial ideals is doubly exponential in the number of variables. The problem of Ideal Membership is also known to EXPSPACE hard even for the case of 3 generators. Thus, at the first glance the computation of the gröbner's basis for memebership checking seems hard, but for the special case that we will be looking into of constant number of generators of constant degree bound which form a regular system this can be done efficiently.

The Macaulay's Matrix:

The matrix is defined for a particular system of polynomials $(f_1, \ldots, f_m)$ and degree d and denoted by $M_{d,m}$. In this the columns are indexed by the monomials of degree d according to grevlex monomial ordering. For each polynomial f_i of the system and each monomial

t of degree $d - d_i$, it contains one row whose entry in the column indexed by a monomial t' is the coefficient of t' in tf_i.

For eg. for a $d = 3$, and number of variables $n = 3(x, y, z)$, the set of monomials in grevlex ordering is $\{x^3, x^2y, xy^2, y^3, x^2z, xyz, zy^2, xz^2, yz^2, z^3\}$. Considering only one element of $f = x^2 - y^2 + 2yz - 2z^2$ in the system we add rows of xf, yf, zf in the Macaulay's matrix $M_{3,1}$ as vectors $[1, 0, -1, 0, 0, 2, 0, -2, 0, 0], [0, 1, 0, -1, 0, 0, 2, 0, -2, 0], [0, 0, 0, 0, 1, 0, -1, 0, 2, -2]$.

It is clear that the linear combination of rows of $M_{d,m}$ generate all the elements of degree d in the ideal generated by $(f_1, \ldots, f_m)$. Thus, the basis of this row space form the basis for the I_d, which can be computed using Gaussian elimination. Also, if we use a graded ordering for ordering the columns,and don't allow for column pivoting, the leading terms of the basis gives the ideal of $LT(I_d)$. Doing this upto D, which is the bound on degree of the elements of gröbner's basis, we get all the elements of the gröbner's basis. Thus, the problem reduces to doing Gaussian elimination efficiently D times, i.e. on the Macaulay's matrices $M_{d,m}$ for $min(d_1, d_2, \ldots, d_m) \leq d \leq D$.

It is shown that Gauss elimination can be done efficiently using matrix multiplication in [Sto00], with a complexity of $O(RCr^{\omega-2})$ where R is the number of rows, C is the number of columns, r is the rank of the matrix and ω is the exponent of matrix multiplication. The number of columns for $M_{d,m}$ is the number of monomials of degree d, which is $\binom{n+d-1}{d}$. The number of rows is bounded by mC_d as each polynomial f_i in the system has a row corresponding to it's multiplication by the monomials of degree $d - d_i$, whose number is less than the number of monomials of degree d,i.e. C_d. Similarly the rank of the matrix is upper bounded by the number of columns. Substituting these into the bound from [Sto00] we get complexity of Gauss elimination for the matrix $M_{d,m}$ is $O(mC_d^\omega)$.

The total complexity we get is

$$\sum_{d=min(d_1,\ldots,d_m)}^{D} O(mC_d^\omega) = O(m \sum_{d=min(d_1,\ldots,d_m)}^{D} C_d^\omega)$$

And since $C_d = \binom{n+d-1}{d}$ is an increasing function of d, we have

$$\sum_{d=min(d_1,\ldots,d_m)}^{D} C_d^\omega \leq D * C_D^\omega$$

giving the final bound of $O(mD\binom{n+D-1}{D}^\omega)$.

Proposition 5.1.1. Let $(f_1, \ldots, f_m)$ be a system of homogeneous polynomials in $\mathbb{F}[x_1, \ldots, x_n]$. The number of $\mathbb{F}$ operations required to compute a Gröbner's basis of the ideal generated by $(f_1, \ldots, f_m)$ for a graded monomial ordering upto degree D is bounded by

$$O\Big(mD\binom{n + D - 1}{D}^\omega\Big)$$

where ω is the exponent of matrix multiplication over $\mathbb{F}$.

A major benefit of viewing Gröbner's computation in terms of Macaulay's matrix is that one it allows us to view the computation at linear algebra operations and get bounds on special cases. A major issue is that for the computation to occur an upper bound on D needs to be known before we start the computation. Another benefit is it enables us to see the connection between Hilbert function of the ideal, and its gröbner's basis. We explore this idea further.

5.1.1 Relation to Hilbert Polynomial

The Hilbert function of an ideal I is defined by

$$HF_I(d) = dim(\mathbb{F}[x_1, \ldots, x_n]_d / I_d)$$

Where $\mathbb{F}[x_1, \ldots, x_n]_d$ denotes the space of polynomials over $\mathbb{F}$ with n variables and degree d, and I_d denotes the elements of ideal I with degree d. It comes from the definition of Hilbert function that it is equal to the dimension of $\mathbb{F}[x_1, \ldots, x_n]_d = \binom{n+d-1}{d}$ minus the rank of $M_{d,m}$. The generating function $H_I = \sum_{d \geq 0} HF_I(d) Z^d$ is called the Hilbert series of the ideal. If the Hilbert series of an ideal is a polynomial of degree D, then we have $HF_I(d \geq D) = 0$, that is all the monomials of $deg > D$ can be generated from the members of leading terms of the basis elements with $deg < D$, then we can say there are no polynomials required of $deg > D$ for the gröbner's basis. Hence showing that the Hilbert series for an ideal is a polynomial, puts a bound on the degree of the elements of the Gröbner's basis.

5.2 The bound on degree for regular sequence

Now we work to prove that the Gröbner's basis of an ideal generated by homogeneous $deg < \delta$ polynomials $(f_1, \ldots, f_m)$ contain elements of degree $D \leq \sum_{i=1}^{m}(d_i - 1) + 1$, which is also known as the *Macaulay's bound*, when the variables are in a Noether position with respect to these polynomials. We will start with a few lemma for the Hilbert series :

Lemma 5.2.1. Let $I \subset \mathbb{F}[x_1, \ldots, x_n]$ be an homogeneous ideal, and f be a homogeneous of degree $\delta \geq 1$. Then f is not a zero-divisor in $\mathbb{F}[x_1, \ldots, x_n]/I$ if and only if $H_{I+\langle f \rangle}(z) = (1 - z^\delta) H_I(z)$.

Proof Simply consider the dimensional relation between the kernel (K_d) and image of the application of multiplication by f from $\mathbb{F}[x_1, \ldots, x_n]_{d-\delta}/I_{d-\delta}$ to $\mathbb{F}[x_1, \ldots, x_n]_d/I_d$. The polynomials in this image are zero in the new ideal $(I + \langle f \rangle)$, hence need to be subtracted from dimension of $\mathbb{F}[x_1, \ldots, x_n]_d/I_d$. By rank nullity, the dimension of this image will be $dim(\mathbb{F}[x_1, \ldots, x_n]_{d-\delta}/I_{d-\delta}) - dim(K_d)$. Thus, we have

$$dim(\mathbb{F}[x_1, \ldots, x_n]_d/(I+\langle f \rangle)_d = dim(\mathbb{F}[x_1, \ldots, x_n]_d/I_d) - (dim(\mathbb{F}[x_1, \ldots, x_n]_{d-\delta}/I_{d-\delta}) - dim(K_d))$$

which is

$$HF_{I+\langle f\rangle}(d) = HF_I(d) - HF_I(d-\delta) + dim(K_d) \qquad d \in \mathbb{N}$$

Multiplying the above with z^d and summing over all d to create the Hilbert series of the ideals, we get

$$H_{I+\langle f\rangle}(z) = (1-z^\delta)H_I(z) + \sum_{d\geq 0} dim(K_d)$$

From the definition of zero-divisor and kernel of a map, f is a not a zero-divisor in $\mathbb{F}[x_1,\ldots,x_n]/I$ iff the multiplication kernel K_d has dimension 0, i.e. $dim(K_d) = 0$ for all $d \geq 0$.

Lemma 5.2.2. The system of homogeneous polynomials $(f_1,\ldots,f_m) \subset \mathbb{F}[x_1,\ldots,x_n]$ is regular iff if it's Hilbert Series is

$$H_I(z) = \frac{\prod_{j=1}^m (1-z^{d_j})}{(1-z)^n}$$

If $m = n$, then the sequence $(f_1,\ldots,f_n)$ is regular iff it's Hilbert series is a polynomial.

Proof: The $dim(\mathbb{F}[x_1,\ldots,x_n]_d)$ will be the number of n-variate monomials of degree d,i.e. $HF_{\langle\rangle}(d) = \binom{n+d-1}{d}$, which means the Hilbert series of the empty ideal is $H_{\langle\rangle}(z) = (1-z)^{-n}$. Now adding f_i into the ideal one-by-one and using lemma 1, we have the first part of the result.

For the second part, the idea is that in the proof of lemma 5.2.1 if it's a zero divisor, the dimension of the Kernel will also be added to the new Hilbert series implying for non-regular sequences $H_I(z) \geq \frac{\prod_{j=1}^m (1-z^{d_j})}{(1-z)^n}$ as for each new addition to the sequence the dimension of kernel increase the value of Hilbert function further. Since the increase in each value the above greater than equal to relation is coefficient by coefficient, with equality only when the complete sequence is regular. Now we will show that if the Hilbert series is a polynomial for $m = n$, then the equality holds for $z = 1$.For $m = n$, if $H_I(z)$ is a polynomial, then Bézout's bound([]) states that $H_I(1)$, which is the number of solutions of I in the algebraic closure of $\mathbb{F}$, is bounded by $\prod_{j=1}^n d_j$, i.e. $H_I(1) \leq \prod_{j=1}^n d_j$. The inequality $H_I(z) \geq \frac{\prod_{j=1}^m (1-z^{d_j})}{(1-z)^n}$ for $z = 1$ and $m = n$, this gives $H_I(1) \geq \prod_{j=1}^n d_j$, because $(1-z^d)/(1-z) = z^{d-1} + z^{d-2} + \ldots + 1$ evaluated at $z = 1$ is d. Hence combining the two gives us equality, implying $H_I(z) = \frac{\prod_{j=1}^m (1-z^{d_j})}{(1-z)^n}$, which means the sequence is regular by first part of the lemma. □

As a corollary of the above two we have:

Corollary 5.2.3. Let $(f_1,\ldots,f_n)$ be a regular system of homogeneous polynomials in $\mathbb{F}[x_1,\ldots,x_n]$, then the highest degree in the elements of a Gröbner basis for a graded ordering is bounded by Macaulay's bound.

Proof When $m = n$, the Hilbert series is a polynomial, whose degree is $\sum_{i\in[n]}(d_i - 1)$. This means that Hilbert coefficient is zero for the degree $\geq D = \sum_{i\in[n]}(d_i - 1) + 1$. As discussed in section 5.1.1,it represents that for degree $\geq D$, the Macaulay's matrix is full rank, this means that all the monomials of these degrees can be made from elements of ideal with lower degree, hence none of these degree monomials can come in the Gröbner's basis. □

Example Consider the system of 3 polynomials

$$f_1 = x^2 + y^2 - 2xz - 2yz + z^2 + h^2$$

$$f_2 = x^2 + xy + yz - z^2 - 2h^2$$

$$f_3 = x^2 - y^2 + 2yz - 2z^2$$

in $\mathbb{F}[x, y, z, h]$. From computations done in section 5.4, we get that the $LT(\langle f_1, f_2, f_3\rangle)$ is generated by monomials $x^2, xy, y^2, xz^2, yz^2, z^4$. From this we get that the $H_{\langle f_1,f_2\rangle}(t) = (1+t)^2/(1-t)^2 = (1-t^2)^2/(1-t)^4$ and $H_{\langle f_1,f_2,f_3\rangle}(t) = (1+t)^3/(1-t) = (1-t^2)^3/(1-t)^4$, which shows these systems are regular by lemma 5.2.2 .

5.2.1 The Noether position

The idea here is to get the result of corollary 5.2.3 from $m = n$ to general case. Here is where Noether position and grevlex ordering become important. We use the following definition of Noether position

Definition 5.2.1.1. The variables $(x_1, \ldots, x_m)$ are in Noether position with respect to the system $(f_1, \ldots, f_m)$ if their canonical images in $\mathbb{F}[x_1, \ldots, x_n]/\langle f_1, \ldots, f_m\rangle$ are algebraic integers over $\mathbb{F}[x_{m+1}, \ldots, x_n]$, i.e. $\forall i \in [m], x_i \in \mathbb{F}[x_1, \ldots, x_n]/\langle f_1, \ldots, f_m\rangle$, there exists a polynomial $g \in \mathbb{F}[x_i, x_{m+1}, \ldots, x_n] \cap \langle f_1, \ldots, f_m\rangle$ that is monic with respect to x_i. Also $\mathbb{F}[x_{m+1}, \ldots, x_n] \cap \langle f_1, \ldots, f_m\rangle = \langle 0\rangle$.

What it geometrically means is that the algebraic set defined by the system of polynomials has dimension $n - m$ and, in an algebraic closure of $\mathbb{F}$, the system has exactly the same number of solutions (counting multiplicity), irrespective the value of $(x_{m+1}, \ldots, x_n)$.

As explained in [Giu88] by Giusti , in a sufficiently large field, for regular systems, the variables can be put in Noether position by a generic linear change of variables.

The following proposition characterizes algebraically the Noether position property for homogeneous ideals.

Proposition 5.2.1.2. [LJ84] Let $(f_1, \ldots, f_m)$ be a system of homogeneous polynomials of $\mathbb{F}[x_1, \ldots, x_n]$, such that $\langle f_1, \ldots, f_m\rangle \neq \langle 1\rangle$. If the variables $(x_1, \ldots, x_m)$ are in Noether position with respect to the system $(f_1, \ldots, f_m)$, the the sequence $(f_1, \ldots, f_m, x_{m+1}, \ldots, x_n)$ is regular.

Proof: From definition 1, we have the variables $x_1, \ldots, x_m$ being in Noether position with respect to the system $(f_1, \ldots, f_m)$, for each $i \in [m]$, there exits a polynomial $g \in \mathbb{F}[x_i, x_{m+1}, \ldots, x_n] \cap \langle f_1, \ldots, f_m \rangle$ of degree $n_i \geq 1$ in x_i such that the coefficient of $x_i^{n_i}$ in g is 1. This means all monomials with degree of x_i greater than n_i can be formed using this g, which hold for $i \leq m$, while for the other variables all monomials with them can be trivially created in the ideal $\langle f_1, \ldots, f_m, x_{m+1}, \ldots, x_n \rangle$. Thus, all monomials of degree greater than the max value of n_i can be created, hence the value of hilbert function for these is 0. Hence, the hibert series for the ideal, i.e. $H_{\langle f_1, \ldots, f_m, x_{m+1}, \ldots, x_n \rangle}(z)$ is a polynomial which we know by lemma 5.2.2 means that the sequence $(f_1, \ldots, f_m, x_{m+1}, \ldots, x_n)$ is regular.

This final proposition gives us the very precise information about the structure of a grevlex Gröbner's basis in a Noether position.

Proposition 5.2.1.3. [LJ84] Let $(x_1, \ldots, x_n)$ be in the Noether position with respect to the homogeneous system $(f_!, \ldots, f_m)$. Let θ_m be a ring endomorphism of $\mathbb{F}[x_1, \ldots, x_n]$, such that $\theta_m(x_i) = x_i$ for $i \in \{1, \ldots, m\}$, while $\theta_m(x_i) = 0$ for $i > m$. Then for grevlex monomial ordering

$$LT(\langle f_1, \ldots, f_m \rangle) = LT(\theta_m(\langle f_1, \ldots, f_m \rangle)) \cdot \langle x_{m+1}, \ldots, x_n \rangle$$

Simply speaking, the leading terms of the elements of the reduced Gröbner's basis do not depend on the variables $(x_{m+1}, \ldots, x_n)$.

Proof: Let $I = \langle f_1, \ldots, f_m \rangle$. The inclusion of $LT(I) \supset LT(\theta_m(I)) \cdot \langle x_{m+1}, \ldots, x_n \rangle$ is easy to see from the fact that for the grevlex monomial ordering, when $\theta_m(f) \neq 0, LT(f) = LT(\theta_m(f))$, which is a direct consequence of definition of grevlex.

For the reverse direction, consider $f \in I$ and $M = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ be its leading monomial for the grevlex ordering. We have to prove there exists a g with the leading monomial $x_1^{\alpha_1} \cdots x_m^{\alpha_m}$. Let l be the largest index such that $x_l | M$. By definition of grevlex and that M is the leading monomial of f, there exists homogeneous polynomials $g_l, \ldots, g_n \in \mathbb{F}[x_1, \ldots, x_n]$, such that

$$f = x_l^{\alpha_l} g_l + x_{l+1} g_{l+1} + \ldots + x_n g_n, \qquad g_l \in \mathbb{F}[x_1, \ldots, x_l] \setminus \{0\} \text{ and } LT(g_l) = x_1^{\alpha_1} \cdots x_{l-1}^{\alpha_{l-1}}$$

By Proposition 5.2.1.2, the sequence $(f_1, \ldots, f_m, x_{m+1}, \ldots, x_n)$ is regular. If $l > m$, then $f \equiv x_l^{\alpha_l} g_l \equiv 0 \mod I + \langle x_{l+1}, \ldots, x_n \rangle$ and since from lemma 5.2.2 x_l is not a zero divisor in $\mathbb{F}[x_1, \ldots, x_n]/(I + \langle x_{l+1}, \ldots, x_n \rangle)$ we get that $g_l \equiv 0 \mod I + \langle x_{l+1}, \ldots, x_n \rangle$, and since g_l doesn't depend on the other variables $g_l \equiv 0 \mod I$. Hence starting from $f \in I$ such that $LT(f) \in \mathbb{F}[x_1, \ldots, x_l]$ with $l > m$, we obtain $g_l \in I$ such that $LT(f) = x_l^{\alpha_l} LT(g_l)$ and $LT(g_l \in \mathbb{F}[x_1, \ldots, x_{l-1}]$. By induction on l we can find a polynomial $g \in I$ such that $LT(f) = x_{m+1}^{\alpha_{m+1}} \cdots x_l^{\alpha_l} LT(g)$, and $LT(g) \in \mathbb{F}[x_1, \ldots, x_m]$. This proves the converse inclusion. □

Consider the ideals $I_1 = \langle f_1, \ldots, f_m \rangle$ and $I_2 = \langle f_1, \ldots, f_m, x_{m+1}, \ldots, x_n \rangle$. The difference between the terms in Gröbner's basis would be that terms in I_1's Gröbner's basis with $x_{m+1}, \ldots, x_n$ will be reduced in I_2. But proposition 5.2.1.3 tells us that the leading terms(for grevlex in Noether position) will not get canceled as they do not depend on these variables , so the degree of terms in Gröbner's basis of I_1 is bounded by degree of terms in Gröbner's basis of I_2, by definition of grevlex that decides leading terms. By corollary we know the bound on degree of terms in Gröbner's basis of I_2 to be $D = \sum_{i\in[m]}(d_i - 1) + \sum_{i\in[m+1,n]}(1-1)+1 = \sum_{i\in[m]}(d_i - 1) + 1$. Hence, the Macaulay's bound, $D = \sum_{i\in[m]}(d_i - 1) + 1$, bounds the degree of elements in Gröbner's basis of an ideal generated by a regular sequence $(f_1, \ldots, f_m)$.

Example Consider the system of polynomials with f_1, f_2 from example we saw earlier

$$f_1 = x^2 + y^2 - 2xz - 2yz + z^2 + h^2$$

$$f_2 = x^2 + xy + yz - z^2 - 2h^2$$

in $\mathbb{F}[x, y, z, h]$. The variables (x, y) are in Noether position wrt to the system (f_1, f_2). It can be seen that the ideal contains polynomials $2y^4 - 6y^3z + 12y^2z^2 + 7y^2h^2 - 8yz^3 - 20yzh^2 + 4z^2h^2 + 9h^4 \in \mathbb{F}[y, z, h]$ and $2x^4 + 2x^3z - 2x^2z^2 - 3x^2h^2 - 2xz^3 - 2xzh^2 + z^2h^2 + 4h^4 \in \mathbb{F}[x, z, h]$ which are monic polynomials wrt y and x.

On the other hand we can see that (y, z) are not in Noether position, as we can show $\langle f_1, f_2 \rangle \cap \mathbb{F}[y, x, h] = \langle 2y^3x + 4y^2x^2 - y^2h^2 + 8yx^3 - 8yxh^2 - 4x^2h^2 - h^4 \rangle$, which means there is no monic polynomial in y in the intersection.

5.3 The matrix F5 algorithm

We discussed in section 5.1, how we can get the Gröbner's basis by doing Gauss elimination on Macaulay's matrices $M_{d,m}$ till degree $d < D$, which becomes polynomial time if the upper bound on degree of elements in Gröbner's basis is $O(1)$, which we saw can be obtained when the polynomials of degree $O(1)$ form a regular sequence. The Faugère's F5 algorithm(2002)[Fau02] is designed so that it ensures that no "useless" reductions to 0 is performed when the input system is regular. The matrix version of this is variant whose analysis is easier to do, in which the using the *F5 criterion* only the required rows are added into the next degree Macaulay's matrix.

To understand the algorithm better, we will look at a bit different notation. In order to keep track of the polynomials that lead to the different rows of the matrices encountered during the algorithm , it is convenient to view a matrix (M) as a map $(s, t) \in S \times T \to M_{s,t} \in \mathbb{F}$ where S is a finite subset of $\mathbb{N} \times \mathcal{T}$ and T a finite subset of $\mathcal{T}$ ordered using a graded ordering, where $\mathcal{T}$ is the set of non-zero monomials in $x_1, \ldots, x_n$ of degree d(when used for Macaulay's matrix of degree d). The basic idea is to represent each row of a Macaulay's matrix with an index $s = (i, \tau)$ which represents sum of τf_i and some other "smaller" polynomials in the ideal, with s being the *signature* of the corresponding polynomial. Each row is viewed as a vector with entry t denoting the coefficient of monomial ranked t in the ordering, in

the polynomial which the row represents. We denote $\overline{M}_{d,i}$ the result of Gauss elimination applied to the matrix $M_{d,i}$.

The algorithm constructs matrices incrementally in the degree(from $min(d_1, \ldots, d_m)$ to D) and the number of polynomials (added in from Gauss elimination of previous degree Macaulay's matrices). We denote the current degree by d and the current number of polynomials with i.

The original F5 criterion translates to the following in the matrix version

Proposition 5.3.1. (F5 criterion) If t is the leading term of $Row(\overline{M}_{d-d_i,i-1}, s)$ where $s < (i, 1)$ then the row indexed by (i, t) belongs to the vector space generated by the rows of $M_{d,i}$ having smaller index.

Proof: We have $t \in LT(\langle f_1, \ldots, f_{i-1}\rangle_{d-d_i})$ which means there is some h such that $t = LT(h)$ with $h = \sum_{k=1}^{i-1} h_k f_k$. This implies that $tf_i = \sum_{k=1}^{i-1} f_i h_k f_k + (t - h)f_i$, where the first term belongs to $\langle Row(M_{d,i-1})\rangle$ and the last one is a linear combination of rows of $M_{d,i}$ having smaller index, as $LT(h) \succ LT(t - h)$. □

The above criteria let's us avoid the cases that are already checked in previous iterations making the number of rows smaller, and hence speeding up the algorithm. Thus, in the algorithm we donot add the rows that are already in the F5 criterion. The time complexity analysis is already given in section 5.1. We give the pseudocode for the algorithm now

Algorithm 1 Finding Gröbner's basis using matrix F5

Require: homogeneous polynomials $(f_1, \ldots, f_m)$ with degrees $d_1 \leq \ldots \leq d_m$; a maximal degree D

1: **function** MATRIXF5$(f_1, \ldots, f_m, D)$
2: $G = \phi$
3: **for** $d \leftarrow \{d_1, d_1 + 1, \ldots, D\}$ **do**
4: $M_{d,0} = \phi, \overline{M}_{d,0} = \phi$
5: **for** $i \leftarrow \{1, 2, \ldots, m\}$ **do**
6: **if** $d < d_i$ **then**
7: $M_{d,i} = M_{d,i-1}$;
8: **else**
9: **if** $d = d_i$ **then**
10: $M_{d_i,i} =$ add the new row f_i to $\overline{M}_{d_i,i-1}$ with index $(i, 1)$
11: **else**
12: $M_{d,i} = \overline{M}_{d,i-1}$
13: $F5criterion = LT(\overline{M}_{d-d_i,i-1})$
14: **for** $f \leftarrow Rows(M_{d-1,i}) \setminus Rows(M_{d-1,i-1})$ **do**
15: $(i, u) = index(f)$, with $u = x_{j_1} \cdots x_{j_{d-d_i-1}}$
16: and $1 \leq j_1 \leq \ldots \leq j_{d-d_i-1} \leq n$
17: **for** $j \leftarrow \{j_{d-d_i-1}, \ldots, n\}$ **do**
18: **if** $ux_j \notin F5criterion$ **then**
19: add the new row $x_j f$ with index (i, ux_j) in $M_{d,i}$
20: Compute $\overline{M}_{d,i}$ by Gauss elimination from $M_{d,i}$
21: Add to G all rows of $\overline{M}_{d,i}$ not reducible by $LT(G_i)$
22: Output G

Example Consider the system of 3 polynomials

$$f_1 = x^2 + y^2 - 2xz - 2yz + z^2 + h^2$$
$$f_2 = x^2 + xy + yz - z^2 - 2h^2$$
$$f_3 = x^2 - y^2 + 2yz - 2z^2$$

in $\mathbb{F}[x, y, z, h]$. As shown earlier this system is regular and the variables (x, y, z) are in Noether position with respect to the system. So we know $D = (2 - 1) * 3 + 1 = 4$. We will look at how the computation in matrix F5 works from this point.

Degree 2: For degree 2 each polynomial is simply added as $(i, 1)$ to $M_{2,i}$ at line 10 in pseudocode. This when $i = m = 3$, gives $M_{2,3}$ to be

$$M_{2,3} = \begin{array}{c} \\ f_1 \\ f_2 \\ f_3 \end{array} \begin{array}{c} \begin{array}{cccccccccc} x^2 & xy & y^2 & xz & yz & z^2 & hx & yh & zh & h^2 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & -2 & -2 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -2 \\ 1 & 0 & -1 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

Gaussian reduction on the above gives and making the last row leading term coefficient positive gives us:

$$\overline{M}_{2,3} = \begin{array}{c} \\ f_1 \\ f_2 \\ f_3 \end{array} \begin{array}{c} \begin{array}{cccccccccc} x^2 & xy & y^2 & xz & yz & z^2 & hx & yh & zh & h^2 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & -2 & -2 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 2 & 3 & -2 & 0 & 0 & 0 & -3 \\ 0 & 0 & 2 & -2 & -4 & 3 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

As a result we add the polynomials $x^2+y^2-2xz-2yz+z^2+h^2, xy-y^2+2xz+3yz-2z^2-3h^2$ and $2y^2-2xz-4yz+3z^2+h^2$ to G, as it completes the degree 2 possible cases.

Degree 3: The number of columns in degree 3 matrices will be $\binom{n+d-1}{d} = \binom{4+3-1}{3} = \binom{6}{3} = 20$ whose matrix cannot fit. So we will look in terms of the signature of polynomials. The *F5criterion* will be empty as there is nothing of degree $d-d_i = 3-2 = 1$ in the generators. This means all the rows of the Macaulay's matrix will be added getting us $M_{3,3}$. The rows after Gauss elimination with there leading terms will be with f_1, f_2, f_3 being the ones of the $\overline{M}_{2,3}$ are

(ind)	$(1,x)$	$(1,y)$	$(1,z)$	$(1,h)$	$(2,x)$	$(2,y)$	$(2,z)$	$(2,h)$	$(3,x)$	$(3,y)$	$(3,z)$	$(3,h)$
(LT)	x^3	x^2y	x^2z	x^2h	y^3	xy^2	xyz	xyh	$\underline{xz^2}$	y^2z	$\underline{yz^2}$	y^2h

The underlined are the ones that cannot be reduced by elements by G till now. Hence, we add those polynomials corresponding to those rows. These are $4xz^2+3yz^2-2z^3+3xh^2+4yh^2+2zh^2$ and $3yz^2-6z^3+11xh^2-5yh^2-3zh^2$.

Degree 4: This is the first part where the F5 criterion comes into play. The set of *F5Criterion* is empty for $i=1$, but for $i=2$, it has x^2 and x^2, xy for $i=3$. This means when we are adding rows on line19 we will skip $(2,x^2)$, $(3,x^2)$ and $(3,xy)$. Note, these were the rows that would have been reduced to zero during Gausss elimination. So $\overline{M}_{4,3}$ will contain $\binom{4+4-1}{4} = \binom{7}{4} = 35$ columns and $3\binom{5}{2} - 3 = 27$ rows. All the rows are reducable using existing G except $(3,y^2)$ which adds $3z^4+4xzh^2+12yzh^2-7z^2h^2-12h^4$ to G. Thus, the final G we have is

$$\begin{aligned} G = \langle & x^2+y^2-2xz-2yz+z^2+h^2, xy-y^2+2xz+3yz-2z^2-3h^2, \\ & 2y^2-2xz-4yz+3z^2+h^2, 4xz^2+3yz^2-2z^3+3xh^2+4yh^2+2zh^2, \\ & 3yz^2-6z^3+11xh^2-5yh^2-3zh^2, 3z^4+4xzh^2+12yzh^2-7z^2h^2-12h^4 \rangle \end{aligned}$$

5.4 Implications for our case

We were looking at the problem of checking ideal membership of a product of constant degree polynomials, in an ideal generated by 2 degree $\delta = O(1)$ irreducible polynomials. As we saw earlier that the Gröbner's basis define a unique weak remainder for a given ideal, obtaining the Gröbner's basis allowed us to check ideal membership. 2 irreducible polynomials of same degree form a regular sequence unless they aren't just a scalar multiple of each other. By the proofs in section 5.2 above we know that the Gröbner's basis of these degree δ polynomials

can only have elements of degree $D = \sum_{i\in[2]}(\delta - 1) + 1 = 2\delta - 1$. This means that the space of monomials that make the gröbner's basis is small(polynomial in n), so that is what we aim at using. We also have seen that we can calculate the gröbner's basis itself too when the variables are in Noether position (which can be obtained by a generic linear change of variables) for grevlex monomial ordering.

6 Using the Degree Bound

We aim at using a variable reduction that preserves the monomial space of all monomials with degree less than $D = 2\delta - 1$, which should conserve the Gröbner's basis for these and hence conserve ideal membership. This variable reduction should also preserve non-zeroness on ideas similiar to the ones in [SS12] for depth3 circuits, giving us a polynomial size hitting set.

7 Conclusion and Future Scope

We looked at all the work done to solve the PIT problem for bounded top and bottom fanin depth4 circuits, and looked at a possible way of extending the work for bounded top fanin depth3 identity testing to depth4. The main idea was to find a homomorphism(variable reduction) which would preserve ideal membership and hence non-zeroness of the circuit. To see this we explored the Gröbner's basis for regular sequences, as Gröbner's basis are the main method when it comes to ideal membership with generators of degree > 1. We found that the degree of the elements in Gröbner's basis is bounded constantly for our case, and thus, the monomial space they lie in is known to be small($poly(n, \delta, m)$).

In future we explore to develop the homomorphism by conserving this small space. The next step to finding such homomorphism would be to try and solve the case when the path is not a regular sequence. The next step to it would be trying to solve the case when the multiplicity of factors in terms is not 1.

References

[AT99] Noga Alon and M Tarsi. Combinatorial nullstellensatz. *Combinatorics Probability and Computing*, 1999.

[AV08] Manindra Agrawal and V Vinay. Arithmetic circuits: A chasm at depth four. *In Foundations of Computer Science*, 2008.

[BFS15] Magali Bardet, Jean-Charles Faugère, and Bruno Savy. On the complexity of f5 gröbner's basis algorithm. *Journal of Symbolic Computation,Elsevier,2015,70,pp 49-70*, 2015.

[DS07] Z. Dvir and Amir Shpilka. Locally decodable code with 2 queries and polynomial identity testing for depth-3 circuits. *SIAM Journal on Computing, 36(5):1404–1434*, 2007.

[Fau02] Jean-Charles Faugère. A new efficient algorithm for computing gröbner's bases without reduction to zero(f_5). *ACM press,pp 75-83*, 2002.

[For14] Michael Andrew Forbes. A polynomial identity testing of read-once oblivious algebraic branching programs. *PhD thesis, Massachusetts Institute of Technology*, 2014.

[Giu88] Marc Giusti. Combinatorial dimension theory of algebraic varieties. *Journal of Symbolic Computation (1988) 6, 249-265*, 1988.

[Gup14] Ankit Gupta. Algebraic geometric techniques for depth-4 pit and sylvester-gallai conjectures for varieties. *Electronic Colloquium on Computational Complexity (ECCC),21:130*, 2014.

[KS07] Neeraj Kayal and Nitin Saxena. Polynomial identity testing for depth 3 circuits. *Computational Complexity, 16(2):115–138*, 2007.

[KS09] Neeraj Kayal and Shubhangi Saraf. Blackbox polynomial identity testingfor depth 3 circuits. *In Foundations of Computer Science, 2009. FOCS'09.50th Annual IEEE Symposium on, pages 198–207*, 2009.

[LJ84] M. Lejeune-Jalabert. Effectivé de calculs polynomiaux. *Université de Grenoble, cours de DEA 84-85*, 1984.

[MM82] E.W. Mayr and A.R. Meyer. The complexity of the word problems for commutative semigroups and polynomial ideals. *Advances in Mathematics 46(3),pg 305-329*, 1982.

[Sax09] Nitin Saxena. Progress on polynomial identity testing. *Bulletin of the EATCS, 99:49–79*, 2009.

[Sax14] Nitin Saxena. Progress on polynomial identity testing- ii. *In Perspectives in Computational Complexity, volume 26 of Progress in Computer Science and Applied Logic, pages 131–146. Springer International Publishing*, 2014.

[Sch80] Jacob T Schwart. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM (JACM)*, 1980.

[Shp19] Amir Shpilka. Sylvester-gallai type theorems for quadratic polynomials. *In Moses Charikar and Edith Cohen, editors,Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019, pages 1203–1214*, 2019.

[SS12] Nitin Saxena and C. Seshadhri. Blackbox identity testing for bounded top-fan in depth-3 circuits: The field doesn't matter. *SIAM Journal on Computing, 41(5):1285–1298*, 2012.

[Sto00] A. Storjohann. Algorithms for matrix canonical forms. *PhD thesis, Department of Computer Science, ETH Zurich*, 2000.

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science: Vol. 5*, 2010.