Deterministic identity testing paradigms for bounded top-fanin depth-4 circuits

Pranjal Dutta
Chennai Mathematical Institute, India (& CSE, IIT Kanpur)

Prateek Dwivedi
Dept. of Computer Science & Engineering, IIT Kanpur

Nitin Saxena
Dept. of Computer Science & Engineering, IIT Kanpur

Abstract

Polynomial Identity Testing (PIT) is a fundamental computational problem. The famous depth-4 reduction (Agrawal & Vinay, FOCS’08) has made PIT for depth-4 circuits, an enticing pursuit. The largely open special-cases of sum-product-of-sum-of-univariates (Σ\[k\]ΠΣ∧) and sum-product-of-constant-degree-polynomials (Σ\[k\]ΠΣΠ\[δ\]), for constants k, δ, have been a source of many great ideas in the last two decades. For eg. depth-3 ideas (Dvir & Shpilka, STOC’05; Kayal & Saxena, CCC’06; Saxena & Seshadhri, FOCS’10, STOC’11); depth-4 ideas (Beecken, Mittmann & Saxena, ICALP’11; Saha, Saxena & Seshadhri, Comput.Compl.’13; Forbes, FOCS’15; Kumar & Saraf, CCC’16); geometric Sylvester-Gallai ideas (Kayal & Saraf, FOCS’09; Shpilka, STOC’19; Peleg & Shpilka, CCC’20, STOC’21). We solve two of the basic underlying open problems in this work.

We give the first polynomial-time PIT for Σ\[k\]ΠΣ∧. Further, we give the first quasipolynomial time blackbox PIT for both Σ\[k\]ΠΣ∧ and Σ\[k\]ΠΣΠ\[δ\]. No subexponential time algorithm was known prior to this work (even if k = δ = 3). A key technical ingredient in all the three algorithms is how the logarithmic derivative, and its power-series, modify the top Π-gate to ∧.

2012 ACM Subject Classification Theory of computation → Algebraic complexity theory

Keywords and phrases Polynomial identity testing, hitting set, depth-4 circuits

Digital Object Identifier 10.4230/LIPIcs.CCC.2021.11

Funding Pranjal Dutta: Google Ph. D. Fellowship
Nitin Saxena: DST (DST/SJF/MSA-01/2013-14) and N. Rama Rao Chair

Acknowledgements Pranjal thanks CSE, IIT Kanpur for the hospitality.

1 Introduction: PIT & beyond

Algebraic circuits are natural algebraic analog of boolean circuits, with the logical operations being replaced by + and × operations over the underlying field. The study of algebraic circuits comprise the large study of algebraic complexity, mainly pioneered (and formalized) by Valiant [87]. A central problem in algebraic complexity is an algorithmic design problem, known as Polynomial Identity Testing (PIT): given an algebraic circuit C over a field F and input variables x_1, ..., x_n, determine whether C computes the identically zero polynomial. PIT has found numerous applications and connections to other algorithmic problems. Among the examples are algorithms for finding perfect matchings in graphs [59, 62, 24], primality testing [4], polynomial factoring [52, 19], polynomial equivalence [21], reconstruction algorithms [48, 83, 44] and the existence of algebraic natural proofs [16, 53]. Moreover, efficient design of PIT algorithms is intrinsically connected to proving strong lower bounds [39, 1, 42, 23, 29, 17, 20]. Interestingly, PIT also emerges in many fundamental results in complexity theory such as IP = PSPACE [82, 60], the PCP theorem [10, 11], and the overarching Geometric Complexity Theory (GCT) program towards P ≠ NP [64, 63, 32, 41].
11.2 Bounded Depth-4 identity testing paradigms

There are broadly two settings in which the PIT question can be framed. In the whitebox setup, we are allowed to look inside the wirings of the circuit, while in the blackbox setting we can only evaluate the circuit at some points from the given domain. There is a very simple randomized algorithm for this problem - evaluate the polynomial at a random point from a large enough domain. With very high probability, a nonzero polynomial will have a nonzero evaluation; this is famously known as the Polynomial Identity Lemma [66, 18, 89, 81]. It has been a long standing open question to derandomize this algorithm.

For many years, blackbox identity tests were only known for depth-2 circuits (equivalently sparse polynomials) [13, 49]. In a surprising result, Agrawal and Vinay [7] showed that a complete derandomization of blackbox identity testing for just depth-4 algebraic circuits \((\Sigma \Pi \Sigma \Pi)\) already implies a near complete derandomization for the general PIT problem. More recent depth reduction results [50, 36], and the bootstrapping phenomenon [2, 55, 34, 9] show that even PIT for very restricted classes of depth-4 circuits (even depth-3) would have very interesting consequences for PIT of general circuits. These results make the identity testing regime for depth-4 circuits, a very meaningful pursuit.

Three PITs in one-shot. Following the same spirit, here we solve three important (and open) PIT questions. We give the first deterministic polynomial-time whitebox PIT algorithm for the bounded sum-of-product-of-sum-of-univariates circuits \((\Sigma^k \Pi \Sigma \land)\) [71, Open Prob. 2]; polynomials computed by these circuits are of the form \(\Sigma_{i \in [k]} \Pi_j (g_{ij1}(x_1) + \cdots + g_{ijn}(x_n))\) (Theorem 1). In fact, we also design the first quasipolynomial-time blackbox PIT algorithm for the same model (Theorem 2a). To the best of our knowledge, no subexponential time algorithm was known prior to this work. A similar technique also gives a quasipolynomial-time blackbox PIT algorithm for the bounded top and bottom fanin circuits \((\Sigma^k \Pi \Sigma \Pi^\delta)\) (where \(k\) and \(\delta\) are constants), see Theorem 2b. These circuits compute polynomials of the form \(\Sigma_{i \in [k]} \Pi_j g_{ij}(x)\), where \(\deg(g_{ij}) \leq \delta\). Even \(\delta = 2\) was a challenging open problem [56, Open Prob. 2].

Prior works on the related models. In the last two decades, there has been a surge of results on identity testing for restricted classes of bounded depth algebraic circuits (eg. ‘locally’ bounded independence, bounded read/occur, bounded variables). There have been numerous results on PIT for depth-3 circuits with bounded top fanin (known as \(\Sigma^k \Pi \Sigma\)-circuits). Divir and Shpilka [22] gave the first quasipolynomial-time deterministic whitebox algorithm for \(k = O(1)\), using rank based methods, which finally lead Karnin and Shpilka [45] to design algorithm of same complexity in the blackbox setting. Kayal and Saxena [47] gave the first polynomial-time algorithm for the bounded top and bottom fanin circuits \((\Sigma^k \Pi \Sigma \Pi^{\delta})\). Later, a series of works in [78, 79, 80, 5] generalized the model and gave \(\nu^{O(k)}\)-time algorithm when the algebraic rank of the product polynomials are bounded.

There has also been some progress on PIT for restricted classes of depth-4 circuits. A quasipolynomial-time blackbox PIT algorithm for multilinear \(\Sigma^k \Pi \Sigma \Pi\)-circuits was designed in [43], which was further improved to a \(\nu^{O(k^\delta)}\)-time deterministic algorithm in [74]. A quasipolynomial blackbox PIT was given in [12, 56] when algebraic rank of the irreducible factors in each multiplication gate as well as the bottom fanin are bounded. Further interesting restrictions like sum of product of fewer variables, and more structural restrictions have been exploited, see [28, 6, 25, 61, 57]. Some progress has also been made for bounded top-fanin and bottom-fanin depth-4 circuits via incidence geometry [35, 84, 68]. In fact, very recently, [69] gave a polynomial-time blackbox PIT for \((\Sigma^k \Pi \Sigma \Pi^{\Omega(1)})\)-circuits.

Why were the problems open? As mentioned above, people have studied depth-4 PIT only with extra restrictions, mostly due to the limited applicability of the existing techniques: they were tailor-made for the specific models and do not generalize. Eg, the
previous methods handle $\delta = 1$ (i.e. linear polynomials at the bottom) or $k = 2$ (via factoring, [71]). While $k = 2$ to 3, or $\delta = 1$ to 2 (i.e. ‘linear’ to ‘quadratic’) already demands a qualitatively different approach.

Whitebox $\Sigma[k]\Pi\Sigma\land$ model generalizes the famous bounded-top-fanin-depth-3 $\Sigma[k]\Pi\Sigma$ of [47]; but their Chinese Remaindering (CR) method, loses applicability and thus fails to solve even a slightly more general model. The blackbox setting involved similar ‘certifying path’ ideas [71] which could be thought of as general CR. It comes up with an ideal $I$ such that $f \neq 0 \mod I$ and finally preserves it under a constant-variate linear map. The preservation gets harder (for both $\Sigma[k]\Pi\Sigma\land$ and $\Sigma[k]\Pi\Sigma\Pi^{[\delta]}$) due to the increased non-linearity of the ideal $I$ generators. Intuitively, larger $\delta$, via ideal-based routes, brings us to the Gröbner basis method (which is doubly-exponential-time in $n$) [88]. We know that ideals even with 3-generators (analogously $k = 4$) already capture the whole ideal-membership problem [73]. The algebraic-geometric approach to $\Sigma[k]\Pi\Sigma\Pi^{[\delta]}$ has been explored in [12, 35, 61, 33].

The families which satisfy a certain Sylvester–Gallai configuration (called SG-circuits) is the harder case which is conjectured to have constant transcendence degree [35, Conj. 1]. Non-SG circuits is the case where the nonzeroness-certifying-path question reduces to radical-ideal non-membership questions [30]. This is really a variety question where one could use algebraic-geometry tools to design a poly-time blackbox PIT. In fact, very recently, Guo [33] gave a $s^{6k}$-time PIT by constructing explicit variety evasive subspace families. Unfortunately, this is not the case in the ideal non-membership; this scenario makes it much harder to solve $\Sigma[k]\Pi\Sigma\Pi^{[\delta]}$. From this viewpoint, radical-ideal-membership explains well why the intuitive $\Sigma[k]\Pi\Sigma$ methods do not extend to $\Sigma[k]\Pi\Sigma\Pi^{[\delta]}$.

Interestingly, Forbes [25] found a quasipolynomial-time PIT for $\Sigma\land\Pi\Sigma^{[\delta]}$ using shifted-partial derivative techniques; but it naively fails when one replaces the $\land$-gate by $\Pi$ (the ‘measure’ becomes too large). The ‘duality trick’ [75] completely solves whitebox PIT for $\Sigma\land\Sigma\land$, by transforming it to a read-once oblivious ABP (ROABP); but it is inapplicable to our models with the top $\Pi$-gate (due to large waring rank and ROABP-width). A priori, our models are incomparable to ROABP, and thus, the famous PIT algorithms for ROABP [28, 27, 37] are not expected to help either.

Similarly, a naive application of the ‘Jacobian’ + ‘certifying path’ technique [5] fails for our models because it is difficult to come up with a faithful map (for constant-variate reduction). Kumar and Saraf [56] crucially used that the computed polynomial has low individual degree (such that [23] can be invoked), while in [57] they exploits the low algebraic rank of the polynomials computed below the top $\Pi$-gate. Neither of them hold, in general, for our models. Very recently, Peleg and Shpilka [69] gave a poly-time blackbox PIT for $\Sigma[3]\Pi\Sigma\Pi^{[2]}$, via incidence geometry (e.g. Edelstein-Kelly theorem involving ‘quadratic’ polynomials), by solving [35, Conj. 1] for $k = 3, \delta = 2$. The method seems very strenuous to generalize even to ‘cubic’ polynomials ($\delta = 3 = k$).

**PIT for other models.** Blackbox PIT algorithms for many restricted models are known. Egs. ROABP related models [70, 40, 3, 37, 27, 8], log-variate circuits [26, 14], certain non-commutative models [31, 58]. We refer to [85, 76, 64, 77, 54, 72] for detailed surveys on PIT and related topics.

### 1.1 Our results: An analytic detour to three PITs

Though some attempts have been made to solve PIT for $\Sigma[k]\Pi\Sigma\land$, no subexponential time PIT for $k \geq 3$ even in the whitebox settings is known, see [71, Open Prob. 2]. Our first result exactly addresses this problem and designs a polynomial-time algorithm (Algorithm 1). The technique (we call it DiDI-paradigm, Sec. 1.2) used is very analytic (& ‘non-ideal’ based).
11.4 Bounded Depth-4 identity testing paradigms

Throughout the paper, we will work with $\mathbb{F} = \mathbb{Q}$, though all the results hold for field of large characteristic.

► Theorem 1 (Whitebox $\Sigma \Pi \Sigma \land$ PIT). There is a deterministic, whitebox $s^{O(k^3)}$-time PIT algorithm for $\Sigma^{[k]} \Pi \Sigma \land$ circuits of size $s$, over $\mathbb{F}[x]$. (See Algorithm 1.)

► Remark. 1. Case $k \leq 2$ can be solved by invoking [71, Thm.5.2]; but $k \geq 3$ was open.
2. Our technique necessarily blows up the exponent exponentially in $k$. In particular, it would be interesting to design a subexponential time algorithm when $k = \Theta(\log s)$.
3. It is not clear if the current technique gives PIT for $\Sigma^{[k]} \Pi \Sigma \land^{[2]}$ circuits, i.e. sum of bivariate polynomials computed and fed into the top product gate.

Next, we go to the blackbox setting and address two models of interest, namely — $\Sigma^{[k]} \Pi \Sigma \land$ and $\Sigma^{[k]} \Pi \Sigma \Pi^{[k]}$, where $k, \delta$ are constants. The prior best algorithms were exponential-time in $s$. Our work builds on previous ideas for unbounded top fanin — (1) Jacobian [5], (2) the known blackbox PIT for $\Sigma \land \Sigma \land$ and $\Sigma \land \Sigma \Pi^{[k]}$ [37, 25] — maneuvering with an analytic approach, via power-series, which unexpectedly reduces the top $\Pi$-gate to a $\land$-gate.

► Theorem 2 (Blackbox PIT for depth-4). (a) There is a deterministic $s^{O(k \log \log s)}$-time blackbox PIT algorithm for $\Sigma^{[k]} \Pi \Sigma \land$ circuits of size $s$, over $\mathbb{F}[x]$.
(b) There is a $s^{O(\delta^2 k \log s)}$-time blackbox PIT algorithm for $\Sigma^{[k]} \Pi \Sigma \Pi^{[k]}$ circuits of size $s$, over $\mathbb{F}[x]$.

► Remark. 1. Thm. 2a has a better dependence on $k$, but worse on $s$, than Thm. 1. Our results are quasipoly-time even up to $k, \delta = \text{poly}(\log s)$.
2. Thm. 2a is better than Thm. 2b, because $\Sigma \land \Sigma \land$ has a faster algorithm than $\Sigma \land \Sigma \Pi^{[k]}$.
3. Even for $\Sigma^{[k]} \Pi \Sigma \land$ and $\Sigma^{[k]} \Pi \Sigma \Pi^{[k]}$ models, we leave the poly-time blackbox question open.

1.2 Proof ideas: A technical synopsis

In this section, we overview the proof of Theorems 1-2. Both the proofs are analytic, i.e. they use logarithmic derivative, and its power-series expansion; greatly transforming the respective models. The first proof is inductive, while the second is a one-shot proof. We remark that in both the cases, we essentially reduce to the well-known ‘wedge’ models, that have unbounded top fanin, yet for which PITs are known. This reduction is unforeseeable and quite ‘powerful’.

The analytic tool that we use, appears in algebra (and complexity theory) through the formal power series ring $R[[x_1, \ldots, x_n]]$ (in short $R[[x]]$), see [65, 86, 19]. The advantages of the ring $R[[x]]$ are many. They usually emerge because of the inverse: $(1 - x_1)^{-1} = \sum_{i \geq 0} x_1^i$, which does not make sense in $R[x]$, but valid in $R[[x]]$. Other analytic tools used are inspired from Wronskian (aka linear dependence) [51, Thm.7] [46], jacobian (aka algebraic dependence) [12, 5, 67], and logarithmic derivative operator $\text{dlog}_{z_1}(f) = (\partial_{z_1} f)/f$.

Moreover, we will be working with the division operator (eg. $R(z_1)$, over a certain ring $R$). The divisions do not come for ‘free’— they require invertibility with respect to $z_1$ throughout (again landing us in $R[[z_1]]$, see Lem. 17). We define class $C/D := \{f/g \mid f \in C, D \ni g \neq 0\}$, for circuit classes $C, \mathcal{D}$, (similarly $C : \mathcal{D}$ denotes the class taking respective products).

The DiDi-technique [Idea of Theorem 1]. The proof of Thm. 1 is recursive and uses a novel technique that we introduce in this work, called DiDi (D= Divide, D=Derive, I=Induct). We illustrate it in $k = 3$, which generalizes to any $k$.

Before going into the technicalities, we want to convey that $k = 3$ is perhaps the first non-trivial case-study. While $k = 1$ is the simplest case (follows directly using sparse-PIT
hitting set \([49]\), \(k = 2\) invokes a strong irreducibility property \([71, \text{Thm. 5.2}]\); and neither of them work for \(k \geq 3\).

The case \(k = 3\) asks to check whether \(T_1 + T_2 + T_3 = 0\), where \(T_i \in \Pi \Sigma \land\) of \(\deg < d\). We apply a homomorphism \(\Phi : \mathbb{F}[x] \rightarrow \mathbb{F}[x, z_1, z_2]\) such that \(x_i \mapsto z_1 \cdot x_i + \Psi(x_i)\) where \(\Psi\) is another homomorphism. The map \(\Psi : \mathbb{F}[x] \rightarrow \mathbb{F}[z_2]\) is a sparse-PIT map s.t. \(\Psi(T_i) \neq 0\) for non-zero \(T_i\), using \([49]\), which ensures that the degree of \(z_2\) is polynomially bounded (Theorem 10). Think of the variable \(z_1\) as a degree-counter which also helps later to derive (the second ‘D’ of DiDi). Observe that \(\Phi\) is a nonzeroness preserving 1-1 map:

\[
T_1 + T_2 + T_3 \neq 0 \iff \Phi(T_1) + \Phi(T_2) + \Phi(T_3) \neq 0.
\]

Denote \(R := \mathbb{F}(z_2)[z_1]/(z_1^d)\). We divide (first ‘D’ of DiDi), by \(\Phi(T_3)\), and derive, wrt \(z_1\), to conclude that \(T_1 + T_2 + T_3 = f\) over \(\mathbb{F}[x]\) implies

\[
\partial_{z_1} \left( \frac{\Phi(T_1)}{\Phi(T_3)} \right) + \partial_{z_1} \left( \frac{\Phi(T_2)}{\Phi(T_3)} \right) = \partial_{z_1} \left( \frac{\Phi(f)}{\Phi(T_3)} \right) \quad \text{over } \mathbb{R}(x).
\]

Denote \(\tilde{T}_1 := \partial_{z_1} (\Phi(T_1)/\Phi(T_3))\) and \(\tilde{T}_2 := \partial_{z_1} (\Phi(T_2)/\Phi(T_3))\). Moreover, \(\partial_{z_1} (\Phi(f)/\Phi(T_3)) = 0\), over \(\mathbb{R}(x)\), if and only if either (1) \(\Phi(f)/\Phi(T_3)\) is \(z_1\)-free, in which case it is an element of \(\mathbb{F}(z_2)\), this can be easily argued by substituting \(z_1 = 0\) in the map \(\Phi\); or (2) \(\text{val}_{z_1}(\partial_{z_1} (\Phi(f)/\Phi(T_3))) \geq d\), which is a contradiction since it implies \(\text{val}_{z_1}(\Phi(f)) > d + 1\). Here, \(\text{val}_{z_1}(\cdot)\) denotes the valuation i.e. the maximum power of \(z_1\) dividing it (which easily extends to fractions via \(\text{val}_{z_1}(p/q) := \text{val}_{z_1}(p) − \text{val}_{z_1}(q)\)). Whenever we talk about val, think of working over \(\mathbb{F}(z_2, x)(z_1)\); which is a ring notion that helps us computationally, and we track the degree of \(z\). This discussion summarizes a crucial fact:

\[
T_1 + T_2 + T_3 \neq 0 \iff \tilde{T}_1 + \tilde{T}_2 \neq 0 \quad \text{over } \mathbb{R}(x), \quad \text{or} \quad \Phi(f) \bigg|_{\Phi(T_3) = 0} \in \mathbb{F}(z_2)\setminus\{0\}.
\]

We remark that the \(z_1 = 0\) substitution is a natural condition as the derivation forgets the \((\mod z_1)\)-part. At the core, theidea is really ‘primal’: if a polynomial \(g(x) \neq 0\), then either its derivative \(g'(x) \neq 0\), or its constant-term \(g(0) \neq 0\) (note: \(g(0) = g(\mod x)\)).

\(\text{Note that, the } z_1 = 0 \text{ substitution part is easy by poly-degree restriction on } z_2. \text{ If it is already } \neq 0, \text{ we are done, otherwise we need to check } \tilde{T}_1 + \tilde{T}_2 \neq 0. \text{ Rewrite } \tilde{T}_1 \text{ as } \Phi(T_1)/\Phi(T_3) \cdot \text{dlog}_{z_1}(\Phi(T_1)/\Phi(T_3))\text{, where dlog denotes the logarithmic-derivative, i.e. } \text{dlog}_{z_1}(\cdot) = \partial_{z_1}(\cdot)/\cdot)\text{.}

\text{Convert top } \Pi \text{ to } \land: \text{ version 1. The map } \Phi \text{ ensures that } \Phi(T_3) \text{ is a unit over } \mathbb{R}. \text{ A calculation shows that the action } \text{dlog}(\Sigma \land) \text{ is in } \Sigma \land / \Sigma \land \in \Sigma \land \Sigma \land, \text{ over } \mathbb{R}[x] \text{ (Claim 4). This crucially uses the inverse identity:}

\[
1 - a \cdot z_1 = 1 + a \cdot z_1 + \ldots + a^{d-1} \cdot z_1^{d-1} \quad \text{over } \mathbb{R}[x],
\]

for \(a \in \mathbb{R}[x]\). Since, dlog is additive over a product (Sec. 2), the action puts dlog(\(\Pi \Sigma \land / \Pi \Sigma \land\)) in \(\sum \text{dlog}(\Sigma \land), \text{ so in } \Sigma \land \Sigma \land. \text{ Thus, both } \tilde{T}_1 \text{ and } \tilde{T}_2 \text{ are of the bloated form } (\Pi \Sigma \land / \Pi \Sigma \land \cdot (\Sigma \land \Sigma \land), \text{ over } \mathbb{R}(x).

\text{Invertibility. The crucial point is that the } \Pi \Sigma \land\text{-circuits are still invertible over } \mathbb{R}[x] \text{ as: dlog newly introduces only } \Sigma \land \Sigma \land, \text{ while the } \Pi \Sigma \land\text{-parts get multiplied by the } \Pi \Sigma \land \text{ within } T_i\text{'s, which are invertible by } \Psi. \text{ Thus, such } (\Pi \Sigma \land) \big|_{z_1 = 0} \in \mathbb{F}(z_2)\setminus\{0\}; \text{ which will be useful later.}

\text{Bloated } k = 2 \text{ model. Is the newly ‘reduced’ model similar to } k = 2 \text{ base-case? It is a more general expression } (\Pi \Sigma \land / (\Pi \Sigma \land \cdot (\Sigma \land \Sigma \land)) + (\Pi \Sigma \land / (\Pi \Sigma \land \cdot (\Sigma \land \Sigma \land))). \text{ Let } \tilde{T}_1 + \tilde{T}_2 = f_1, \text{ over } \mathbb{R}(x). \text{ We know that } f_1 \neq 0 \text{ (by hypothesis). We again apply ‘Divide and Derive’ of}
11.6 Bounded Depth-4 identity testing paradigms

DiDi; here we divide with the \( \tilde{T}_1 \) where \( \text{val}_{z_1} \) is minimal. Wlog, \( \text{val}_{z_1}(\tilde{T}_2) =: v \), is the minimal valuation. Of course, \( 0 \leq v < d \) (strict because of \( \Psi \)). Let us define \( R' := \mathbb{F}(z_2)[z_1]/(z_1^{d-v-1}) \).

Then, \( (\tilde{T}_1/\tilde{T}_2) + 1 = f_1/\tilde{T}_2 \) over \( R'(x) \). This is well-defined as the division is being done by the minimum valuation (Lemma 17); thus after derivation, the modulus goes from \( z_1^d \) to \( z_1^{d-v-1} \) which is well-defined over \( R'(x) \). However, if we derive: \( \partial_{z_1}(f_1/\tilde{T}_2) =: f_2 \) may become \( 0 \) over \( R'(x) \). That could happen if and only if:

1. Either, \( f_1/\tilde{T}_2 \) is \( z_1 \)-free; in that case
   \[
   \left. \frac{f_1}{\tilde{T}_2} \right|_{z_1 = 0} = \left( \frac{\tilde{T}_1}{\tilde{T}_2} + 1 \right) \bigg|_{z_1 = 0} \in \mathbb{F}(z_2) \cdot \frac{\Sigma \wedge \Sigma \wedge}{\Sigma \wedge \Sigma \wedge} + 1.
   \]

   This is easy to test using \( \Sigma \wedge \Sigma \wedge \) whitebox PIT (Lemma 18) (we keep track of the circuit-size respectively the degree of \( z_2 \) and ensure them polynomially bounded),

2. Or, \( \text{val}_{z_1}(f_2) \geq d - v - 1 \implies \partial_{z_1}(f_1/\tilde{T}_2) = z_1^{d-v-1} \cdot p \), for some \( p \in R'(x) \) s.t. \( \text{val}_{z_1}(p) \geq 0 \); this further implies \( p \in \mathbb{F}(z_2)[[z_1]] \) (Lemma 17). Thus \( \text{val}_{z_1}(f_1/\tilde{T}_2) \geq d - v \implies f_1 = 0 \), over \( R(x) \), a contradiction.

Thus, we check the easy condition (1). If the \( z_1 = 0 \) substitution outputs 0, we need to check whether other monomials of \( z_1 \) in \( f_2 \) survive. This suffices to conclude \( f \neq 0 \).

Thankfully \( f_2 = \partial_{z_1}(\tilde{T}_1/\tilde{T}_2) \) is now a \( (\Pi \Sigma \wedge /\Pi \Sigma \wedge) \cdot (\Sigma \wedge \Sigma \wedge /\Sigma \wedge \Sigma \wedge) \) circuit over \( R'(x) \). This is the same analysis as above that converts top \( \Pi \) to \( \wedge \). Except, we may not need to remove \( \Sigma \wedge \Sigma \Lambda \) from the denominator; so we work with this fractional bloated model. (Note: the reciprocal may not be in the polynomial ring \( R'[x] \), but only in the ring \( R'(x) \).)

Finally, identity testing of \( (\Pi \Sigma \wedge /\Pi \Sigma \wedge) \cdot (\Sigma \wedge \Sigma \wedge /\Sigma \wedge \Sigma \wedge) \), over \( R'(x) \) is easy: (1) \( \Sigma \wedge \Sigma \wedge \) is closed under coefficient extraction with respect to \( z_1 \) (Lemma 13), (2) whitebox identity testing is in \( \mathcal{P} \) for both \( \Pi \Sigma \wedge \) (Theorem 10) and \( \Sigma \wedge \Sigma \wedge \) (convert it to an ROABP using [75] and invoke [70], see Lemma 18), (3) the degree of \( z_1, z_2 \) respectively circuit-size remain polynomially bounded.

For general induction, our bloated model is \( \Sigma^{[k]}(\Pi \Sigma \wedge /\Pi \Sigma \wedge) \cdot (\Sigma \wedge \Sigma \wedge /\Sigma \wedge \Sigma \wedge) \).

More work shows that it is closed under DiDi-technique. This is primarily what makes our polynomial-time algorithm possible. For details, refer to Section 3.1 and Algorithm 1

**Jacobian hits again [Idea of Theorem 2].** Suppose we want to test \( T_1 + \ldots + T_k \neq 0 \), where \( T_i \in \Pi \Sigma \Pi^{[d]} \) (resp. \( \Pi \Sigma \Lambda \)). We associate a famous polynomial— the Jacobian \( J(T_1, \ldots, T_r) \) (see Sec. 2). It captures the algebraic independence of \( T_1, \ldots, T_r \), assuming this to be a transcendence basis of the \( T_i \)’s (see Fact 2). If we could find an \( r \)-variate linear map \( \Phi \), that keeps \( T_1, \ldots, T_r \) algebraically independent, then \( \Phi(T_1), \ldots, \Phi(T_r) \) are again algebraically independent and it can be shown that for any \( C: C(T_1, \ldots, T_k) = 0 \iff C(\Phi(T_1), \ldots, \Phi(T_k)) = 0 \) (Fact 1). Such a map is called ‘faithful’ [5].

The overall idea is to find an explicit homomorphism \( \Psi : \mathbb{F}[x] \rightarrow \mathbb{F}[x, z_1, z_2] \), and then fix \( x \) by a hitting-set \( H' \) to get a *composed* map \( \Psi' \) s.t. \( r_{\Psi(x)}(J_{\Phi}(T)) = r_{\Psi(x)}(J_{\Phi}(T)) \) (here \( J \) is the jacobian matrix and \( T = (T_1, \ldots, T_r) \)). Next, extend this map to \( \Phi : \mathbb{F}[x] \rightarrow \mathbb{F}[z, y, t] \) s.t. \( x_i \mapsto (\sum_{j=1}^k y_j t^{j-1}) + \Psi(x_t) \), which is faithful. The construction of the map \( \Psi' \) is crucial. We efficiently construct it by reducing \( \Psi(J_{x_t}(T)) \) to \( \Sigma \wedge \Sigma \Pi^{[d]} \) (resp. \( \Sigma \wedge \Sigma \) ) circuits, which have quasi-poly size hitting sets [25] (resp. Lemma 18).

---

1 This is a special case of \( \Sigma^{[k]} \Pi \Sigma \wedge \Sigma \wedge \) circuits; which is really depth-6.
Jacobian works. A priori, Jacobian is a difficult determinant to work with, and so is finding a faithful $\Phi$. However, for the special models (in this paper) we are able to design $\Phi$, mainly because of two reasons— (1) Jacobian being defined via partial derivatives, has a nice ‘linearizing effect’ on the top II-gates (that are only $r \leq k$ many), (2) Jacobian under a homomorphism $\Psi$ has a nice expression (think of this as a generalized $d\log$-expression):

$$\Psi(J_{x,\Psi}(T)) = \Psi(T_1 \cdots T_r) \cdot \sum_{g_1 \in L(T_1), \ldots, g_r \in L(T_r)} \frac{\Psi(J_{x,\Psi}(g_1, \ldots, g_r))}{\Psi(g_1 \cdots g_r)}. \quad (\text{see Eqn. 6})$$

Here, $L(T_i)$ denotes the multiset of sparse polynomials that constitutes $T_i$. We show: each $1/\Psi(\cdot)$ has a ‘small’ $\Sigma \land \Sigma \Pi^{[\ell]}$-circuit (respes. $\Sigma \land \Sigma \land$). The last point requires invertibility. Define, $\Psi : x_i \mapsto z_1 x_i + \Psi_1(x_i)$, where $\Psi_1(\cdot)$ is a sparse-PIT map s.t. $\Psi_1 : F[x] \rightarrow F[z_2]$ s.t. $\Psi_1(T_{i1}) \neq 0$. Under the $\Psi$, $T_i$ is a unit over ring $R := F[z_2][z_1]/(z_1^r)$, where $D$ is polynomially bounded. The idea behind the map is similar to that of Thm. 1. Next, we sketch why $\Psi(J_{x,\Psi}(T))$ has a $\Sigma \land \Sigma \Pi^{[\ell]}$ circuit (respes. $\Sigma \land \Sigma \land$) of size $s^{O(k)}$ over $R[x]$.

Convert top II to $\land$: version 2. The critical point is to show that $1/\Psi(g_1 \cdots g_k)$, over $R[x]$, where $g_i \in \Sigma \Pi^{[\ell]}$ (respes. $\Sigma \land$) has $s^{O(k)}$ size $\Sigma \land \Sigma \Pi^{[\ell]}$ (respes. $\Sigma \land \Sigma \land$) circuit (see Lem. 9): this again follows from the inverse identity Equation 1. We keep track of the degree of $z$ throughout, which eventually is bounded by $s^{O(k)}$. Thus, the $H'$ can be efficiently constructed from the hitting set of the respective models (of quasipolynomial size), see Thm. 24 and 18. The map $\Phi$ ultimately provides a hitting set for $T_1 + \ldots + T_k$, as we reduce to a PIT of a polynomial over ‘few’ (roughly equal to $k$) variables, yielding a QP-time algorithm.

It is important to note that there was no power series in $[5]$; this really empowers the jacobian technique as it now manifests new reduced models, for which a hitting-set is known. This technique is also inherently map-based. So, it requires a hitting-set and fails to give a poly-time whitebox PIT for the respective models. For the detailed proof, see Section 3.2.

2 Preliminaries

Before proving the results, we describe some of the assumptions and notations used throughout the paper. $x$ denotes $(x_1, \ldots, x_n)$. $[n]$ denotes $\{1, \ldots, n\}$.

Logarithmic derivative. Over a ring $R$ and a variable $y$, the logarithmic derivative $d\log_y : R[y] \rightarrow R(y)$ is defined as $d\log_y(f) := \partial_y f/f$: here $\partial_y$ denotes the partial derivative with respect to variable $y$. One important property of $d\log$ is that it is additive over a product as

$$d\log_y(f \cdot g) = \frac{\partial_y (f \cdot g)}{f \cdot g} = \frac{(f \cdot \partial_y g + g \cdot \partial_y f)}{f \cdot g} = d\log_y(f) + d\log_y(g).$$

We refer this effect as linearization of product.

Circuit size. Sparsity $sp(\cdot)$ refers to the number of nonzero monomials. In this paper, it is a parameter of the circuit size. In particular, $\text{size}(g_1 \cdots g_k) := \sum_{i \in [s]} (\text{sp}(g_i) + \text{deg}(g_i))$, for $g_i \in \Sigma \land \Sigma \Pi^{[\ell]}$. In whitebox settings, we also include the bit-complexity of the circuit (i.e. bit complexity of the constants used in the wires) in the size parameter. Some of the complexity parameters of a circuit are depth (number of layers), syntactic degree (the maximum degree polynomial computed by any node), fanin (maximum number of inputs to a node).

Hitting set. A set of points $H \subseteq \mathbb{F}^n$ is called a hitting-set for a class $C$ of $n$-variate polynomials if for any nonzero polynomial $f \in C$, there exists a point in $H$ where $f$ evaluates
to a nonzero value. A \( T(n) \)-time hitting-set would mean that the hitting-set can be generated in time \( T(n) \), for input size \( n \).

**Valuation.** Valuation is a map \( \text{val}_y : R[y] \to \mathbb{Z}_{\geq 0} \), over a ring \( R \), such that \( \text{val}_y(p) \) is defined to be the maximum power of \( y \) dividing the element. It can be easily extended to fraction field \( R(y) \), by defining \( \text{val}_y(p/q) := \text{val}_y(p) - \text{val}_y(q) \); where it can be negative.

**Field.** We denote the underlying field as \( \mathbb{F} \) and assume that it is of characteristic 0. All our results hold for other fields (eg. \( \mathbb{Q}_p, \mathbb{F}_p \)) of large characteristic (see Remarks in Section 3.1-3.2).

**Jacobian.** The Jacobian of a set of polynomials \( f = \{ f_1, \ldots, f_m \} \) in \( \mathbb{F}[x] \) is defined to be the matrix \( J_f(f) := (\partial_{x_j}(f_i))_{m \times n} \). Let \( S \subseteq x = \{ x_1, \ldots, x_n \} \) and \( |S| = m \). Then, polynomial \( J_S(f) \) denotes the minor (i.e. determinant of the submatrix) of \( J_f(f) \), formed by the columns corresponding to the variables in \( S \). For its useful properties, see Appendix C.

## 3 Proof of the main theorems

This section proves Theorems 1-2. The proofs are self contained and we assume for the sake of simplicity that the underlying field \( \mathbb{F} \) has characteristic 0. When this is not the case, we discuss the corresponding required characteristic as remarks after the respective proofs.

### 3.1 Proof of Theorem 1

As seen in Section 1.2, we will induct over the bloated model which naturally generalizes \( \Sigma \Pi \Sigma \land \) circuits and works ideally under the DiDi-techniques. Formally, we define it below.

▶ **Definition 3.** We call a circuit \( C \in \text{Gen}(k, s) \), over \( R(x) \), for any ring \( R \), with parameter \( k \) and size-\( s \), if \( C \in \Sigma[k](\Pi \Sigma \land / \Pi \Sigma \land \cdot (\Sigma \land \Sigma \land / \Sigma \land \Sigma \land) \). It computes \( f \in \mathbb{R}(x) \), if \( f = \sum_{i=1}^k T_i \), where

1. \( T_i =: (U_i/V_i) \cdot (P_i/Q_i) \), for \( U_i, V_i \in \Pi \Sigma \land \), and \( P_i, Q_i \in \Sigma \land \Sigma \land \),
2. size\((T_i) = \text{size}(U_i) + \text{size}(V_i) + \text{size}(P_i) + \text{size}(Q_i) \), and size\((f) = \sum_{i \in [k]} \text{size}(T_i) \).

Eg. Size-\( s \) \( \Sigma[k]\Pi \Sigma \land \) -circuit \( \in \text{Gen}(k, s) \). We will design a recursive algorithm.

**Proof of Theorem 1.** Begin with \( T_{i,0} := f_i \) and \( f_0 := f \) where \( T_{i,0} \in \Pi \Sigma \land ; \sum_{i} T_{i,0} = f_0 \), and \( f_0 \) has size \( \leq s \). Assume \( \text{deg}(f) < d \leq s \); we keep the parameter \( d \) separately, to help optimize the complexity later. In every recursive call we work with \( \text{Gen}(\cdot, \cdot) \) circuits. As the input case, define \( U_{i,0} := T_{i,0} \) and \( V_{i,0} := P_{i,0} := Q_{i,0} := 1 \). Further define a 1-1 homomorphism \( \Phi : \mathbb{F}[x] \rightarrow \mathbb{F}[x, z_1, z_2] \) such that \( x_i \rightarrow z_1 \cdot x_i + \Psi(x_i) \). Here, \( \Psi : \mathbb{F}[x] \rightarrow \mathbb{F}[z_2] \) is a sparse-PIT map [49] s.t. \( \Psi(U_{i,0}) \neq 0, \forall i \in [k] \) (Theorem 10). Invertibility implies that \( f_0 = 0 \iff \Phi(f_0) = 0 \). Further, the degree bound of \( z_2 \) on \( \Phi(T_{i,0}) \) is \( \text{poly}(s) \). The algorithm is recursive, and first reduces the identity testing from top-fanin \( k \) to \( k - 1 \). Note: \( k = 1 \) is trivial.

0-th step. Efficient reduction from \( k \) to \( k - 1 \). By assumption, \( \sum_{i=1}^k T_{i,0} = f_0 \) and
\( T_{k,0} \neq 0 \). Apply \( \Phi \) both sides. Then divide and derive:
\[
\sum_{i \in [k]} T_{i,0} = f_0 \iff \sum_{i \in [k]} \Phi(T_{i,0}) = \Phi(f_0)
\]
\[
\iff \sum_{i \in [k-1]} \Phi(T_{i,0}) + 1 = \frac{\Phi(f_0)}{\Phi(T_{k,0})}
\]
\[
\iff \sum_{i \in [k-1]} \partial_{z_i} \left( \frac{\Phi(T_{i,0})}{\Phi(T_{k,0})} \right) = \partial_{z_k} \left( \frac{\Phi(f_0)}{\Phi(T_{k,0})} \right)
\]
\[
\iff \sum_{i = 1}^{k-1} \Phi(T_{i,0}) \cdot \text{dlog} \left( \frac{\Phi(T_{i,0})}{\Phi(T_{k,0})} \right) = \partial_{z_k} \left( \frac{\Phi(f_0)}{\Phi(T_{k,0})} \right). 
\]  
(2)

Define the following:
- \( R_1 := f(z_2)[z_1]/\langle z_1^d \rangle \). Note that, Eqn.(2) holds over \( R_1(x) \).
- \( \bar{T}_{i,1} := \Phi(T_{i,0})/\Phi(T_{k,0}) \cdot \text{dlog}(\Phi(T_{i,0})/\Phi(T_{k,0})), \forall i \in [k-1] \).
- \( f_1 := \partial_{z_k}(\Phi(f_0)/\Phi(T_{k,0})), \) over \( R_1(x) \).

**Definability of \( T_{i,1} \) and \( f_1 \).** It is easy to see that these are well-defined terms. Here, we emphasize that we do not exactly compute/store \( T_{i,1} \) as a fraction where the degree in \( z_1 \) is \(< d \); instead it is computed/stored as an element in \( F(z_2)(z_1, x) \), where \( z_1 \) is a formal variable. Formally, we compute \( T_{i,1} \in F(z_2)(z_1, x) \), such that \( \bar{T}_{i,1} = T_{i,1} \), over \( R_1(x) \). We keep track of the degree of \( z_1 \) and \( z_2 \) in \( T_{i,1} \). Thus, \( \sum_{i \in [k-1]} T_{i,1} = f_1 \), over \( R_1(x) \).

**The ‘iff’ condition.** Equality in Eqn. (2) over \( R_1(x) \) is one-sided; however we want a \( \iff \) condition to efficiently reduce the identity testing. Note that \( f_1 \neq 0 \) implies \( \text{val}_{z_1}(f_1) < d = d_1 \). By assumption, \( \Phi(T_{k,0}) \) is invertible over \( R_1(x) \). Further, \( f_1 = 0 \), over \( R_1(x) \), implies –

1. Either, \( \Phi(f_0)/\Phi(T_{k,0}) \) is \( z_1 \)-free. This implies \( \Phi(f_0)/\Phi(T_{k,0}) \in F(z_2)(x) \), which further implies it is in \( F(z_2) \), because of the map \( \Phi \) (\( z_1 \)-free implies \( x \)-free, by substituting the \( z_1 = 0 \)). Also, note that \( f_0, T_{k,0} \neq 0 \) implies \( \Phi(f_0)/\Phi(T_{k,0}) \) is a nonzero element in \( F(z_2) \). Thus, it suffices to check whether \( \Phi(f_0)|_{z_1=0} = \Psi(f_0) \) is non-zero or not. Further, the degree of \( z_2 \) in \( \Phi(f_0) \) is polynomially bounded.

2. Or, \( \partial_{z_1}(\Phi(f_0)/\Phi(T_{k,0})) = z_1^{d_1} \cdot p \) where \( p \in F(z_2)(z_1, x) \) s.t. \( \text{val}_{z_1}(p) \geq 0 \). By simple power series expansion, one can conclude that \( p \in F(z_2, x)[[z_1]] \) (Lemma 17). Hence,
\[
\Phi(f_0)/\Phi(T_{k,0}) = z_1^{d_1+1} \cdot q \text{ where } q \in F(z_2, x)[[z_1]], \text{i.e.}
\]
\[
\Phi(f_0)/\Phi(T_{k,0}) \in \langle z_1^{d_1+1} \rangle F(z_2, x)[[z_1]] \Rightarrow \text{val}_{z_1}(\Phi(f_0)) \geq d + 1,
\]
a contradiction.

Conversely, it is obvious that \( f_0 = 0 \) implies \( f_1 = 0 \). Thus, we have proved the following
\[
\sum_{i \in [k]} T_{i,0} \neq 0 \text{ over } F[x] \iff \sum_{i \in [k-1]} T_{i,1} \neq 0 \text{ over } R_1(x), \text{ or, } 0 \neq \Phi(f_0)|_{z_1=0} \in F(z_2).
\]
Eventually, we show that \( T_{i,1} \in (\Pi \Sigma \Lambda / \Pi \Sigma \Lambda) \cdot (\Sigma \Lambda \Sigma \Lambda / \Sigma \Lambda \Sigma \Lambda), \) over \( R_1(x) \), with polynomial blowup in size (Claim 4). So, the above circuit is in \( \text{Gen}(k-1, \cdot), \) over \( R_1(x) \), which we recurse on to finally give the identity testing. The 1-th step is a bit more tricky:

**Induction step.** Assume that we are in the \( j \)-th step \( (j \geq 1) \). Our induction hypothesis assumes –
1. \( \sum_{i \in [k-j]} T_{i,j} = f_j \), over \( R_j(x) \), where \( R_j := \mathbb{F}(z_2)[z_1]/(z_1^{d_j}) \), and \( T_{i,j} \neq 0 \).
2. Here, \( T_{i,j} := (U_{i,j}/V_{i,j}) \cdot (P_{i,j}/Q_{i,j}) \), where \( U_{i,j}, V_{i,j} \in \Pi \Sigma \Lambda \), and \( P_{i,j}, Q_{i,j} \in \Sigma \Lambda \), each in \( R_j[x] \). Think of them being computed as \( \mathbb{F}(z_2)(z_1, x) \), with the degrees being tracked.

Wlog, assume that \( val_{z_1}(T_{k-j}) \) is the minimal among all \( T_{i,j} \)’s.

3. \( val_{z_1}(T_{i,j}) \geq 0, \forall i \in [k-j] \). Moreover, \( U_{i,j}, V_{i,j} \in \mathbb{F}(z_2) \setminus \{0\} \) (similarly \( V_{i,j} \)).

4. \( f \neq 0 \), over \( \mathbb{F}[x] \) \( \iff \) \( f_j \neq 0 \), over \( R_j(x) \), or, \( \bigvee_{i=0}^{j-1} (f_i/T_{k-i}) |_{z_1=0} \neq 0 \), over \( \mathbb{F}(z_2)(x) \).

We follow the 0-th step, without applying any further homomorphism. Note that the ‘or condition’ in the last hypothesis is similar to the \( j = 0 \) case except that there is no \( \Phi \): this is because \( \Phi(f_0) |_{z_1=0} \neq 0 \iff \Phi(f_0/T_{k,0}) |_{z_1=0} \neq 0 \). This condition just separates the derivative from the constant-term (as pointed in Section 1.2).

Let \( \text{val}_{z_1}(P_{i,j}/Q_{i,j}) =: v_{i,j}, \text{ for } i \in [k-j] \). Note that
\[
\min_i \text{val}_{z_1}(T_{i,j}) = \text{min}_{i,j} \text{val}_{z_1}(P_{i,j}/Q_{i,j}) = v_{k-j,j}
\]
since \( \text{val}_{z_1}(U_{i,j}) = \text{val}_{z_1}(V_{i,j}) = 0 \) (else we reorder). We remark that \( 0 \leq v_{i,j} < d_j \) for all \( i \)’s in \( j \)-th step: upper-bound is strict, since otherwise \( T_{i,j} = 0 \) over \( R_j(x) \).

**Min val computation is easy.** Finding this min val is easy, as we can compute \( \text{val}_{z_1}(P_{i,j}) \) and \( \text{val}_{z_1}(Q_{i,j}), \forall i \in [k-j] \). To compute val, note that \( \text{coef}_{z_1}^{i,j}(P_{i,j}) \) and \( \text{coef}_{z_1}^{i,j}(Q_{i,j}) \) are in \( \Sigma \Lambda \) as well, over \( \mathbb{F}(z_2)[x] \) (Lemma 13). We can keep track of \( z_1 \) degree and thus interpolate to find the minimum \( e < d_j \) such that it is \( \neq 0 \) (implying it to be the respective val).

**Efficient reduction from \( k-j \) to \( k-j-1 \).** Similar to the 0-th step, we divide and derive:
\[
\sum_{i \in [k-j]} T_{i,j} = f_j \iff \sum_{i \in [k-j-1]} \frac{T_{i,j}}{T_{k-j,j}} + 1 = f_j/T_{k-j,j}
\]
\[
\iff \sum_{i \in [k-j-1]} z_1(T_{i,j}/T_{k-j,j}) = \partial z_1(f_j/T_{k-j,j})
\]
\[
\iff \sum_{i=1}^{k-j-1} T_{i,j}/T_{k-j,j} \cdot \text{dlog}(T_{i,j}/T_{k-j,j}) = \partial z_1(f_j/T_{k-j,j}) \quad (3)
\]

Define the following:
- \( R_{j+1} := \mathbb{F}(z_2)[z_1]/(z_1^{d_{j+1}}), \text{ where } d_{j+1} := d_j - v_{k-j,j} - 1 \).
- \( \tilde{T}_{i,j+1} := T_{i,j}/T_{k-j,j} \cdot \text{dlog}(T_{i,j}/T_{k-j,j}), \forall i \in [k-j-1] \).
- \( f_{j+1} := \partial z_1(f_j/T_{k-j,j}), \text{ over } R_{j+1}(x) \).

**Definability of \( T_{i,j+1} \) and \( f_{j+1} \).** By the minimal valuation assumption, it follows that \( \text{val}(f_j) \geq v_{k-j,j} \), and thus \( \tilde{T}_{i,j+1} \) and \( f_{j+1} \) are all well-defined over \( R_{j+1}(x) \). Note that, Eqn. (3) holds over \( R_{j+1}(x) \) as \( d_{j+1} < d_j \) (because, whatever identity holds true \( z_1^{d_j} \) must hold \( \text{mod } z_1^{d_{j+1}} \) as well). Hence, we must have \( \sum_{i=1}^{k-j-1} \tilde{T}_{i,j+1} = f_{j+1} \), over \( R_{j+1}(x) \) [proving induction hypothesis (1)].

Similarly, we emphasize on the fact that we do not exactly compute \( \tilde{T}_{i,j+1} \odot z_1^{d_{j+1}} \) instead it is computed as a fraction in \( \mathbb{F}(z_2)(z_1, x) \), with formal \( z_1 \). Formally, we compute/store \( T_{i,j+1} \in \mathbb{F}(z_2)(z_1, x) \), such that \( \tilde{T}_{i,j+1} = T_{i,j+1} \), over \( R_{j+1}(x) \). We keep track of the degree of \( z_1 \) and \( z_2 \) in \( T_{i,j+1} \). Also, by definition, \( \text{val}_{z_1}(T_{i,j+1}) \geq 0 \) (as we divide by the min val).
Thus, we have proved that with only polynomial blowup in size. The above Eqn. (3) pioneers to reduce from \( k - j \)-summands to \( k - j - 1 \). But we want a \( \iff \) condition to efficiently reduce the identity testing. If \( f_{j+1} \neq 0 \), then \( \text{val}_{z_{1}}(f_{j+1}) < d_{j+1} \). Further, \( f_{j+1} = 0 \), over \( R_{j+1}(x) \) implies –

1. Either, \( F_{j}/T_{k-j} \) is \( z_{1} \)-free. This implies it is in \( \mathbb{F}(z_{2})(x) \). Now, if indeed \( f_{0} \neq 0 \), then the computed \( T_{i,j} \) as well as \( f_{j} \) must be non-zero over \( \mathbb{F}(z_{2})(z_{1}, x) \), by induction hypothesis (as they are non-zero over \( R_{j}(x) \)). However,

\[
\left( \frac{T_{i,j}}{T_{k-j,j}} \right) \bigg|_{z_{1}=0} = \left( \frac{U_{i,j} \cdot V_{k-j,j}}{U_{k-j,j} \cdot V_{i,j}} \right) \bigg|_{z_{1}=0} \cdot \left( \frac{P_{i,j} \cdot Q_{k-j,j}}{P_{k-j,j} \cdot Q_{i,j}} \right) \bigg|_{z_{1}=0} \in \mathbb{F}(z_{2}) \cdot \left( \frac{\Sigma \land \Sigma \land \Sigma \land \Sigma}{\Sigma \land \Sigma \land \Sigma} \right).
\]

Thus,

\[
\frac{f_{j}}{T_{k-j,j}} \in \sum \mathbb{F}(z_{2}) \cdot \left( \frac{\Sigma \land \Sigma \land \Sigma \land \Sigma}{\Sigma \land \Sigma \land \Sigma} \right) \in \left( \frac{\Sigma \land \Sigma \land \Sigma \land \Sigma}{\Sigma \land \Sigma \land \Sigma} \right).
\]

Here we crucially use that \( \Sigma / \Sigma \land \Sigma \land \Sigma \land \Sigma \) is closed under multiplication (Lemma 15). We show that the degree of \( z_{2} \) (in denominator and numerator) in each \( T_{i,j}/T_{k-j} \) is poly-bounded. Thus, this identity testing can be done in poly-time (Lemma 18). For, detailed time-complexity and calculations, see Claim 4 and its subsequent paragraph.

Or, \( \partial_{z_{1}}(f_{j}/T_{k-j,j}) = z_{1}^{d_{j}+1} \cdot p \), where \( p \in \mathbb{F}(z_{2})(z_{1}, x) \) s.t. \( \text{val}_{z_{1}}(p) \geq 0 \). By a simple power series expansion, one concludes that \( p \in \mathbb{F}(z_{2}, x) \land (z_{1}) \) (Lemma 17). Hence, one concludes that

\[
\frac{f_{j}}{T_{k-j,j}} \in \left( z_{1}^{d_{j}+1} \right) \mathbb{F}(z_{2}, x) \land (z_{1}) \implies \text{val}_{z_{1}}(f_{j}) \geq d_{j},
\]

i.e. \( f_{j} = 0 \), over \( R_{j}(x) \).

Conversely, \( f_{j} = 0 \), over \( R_{j}(x) \), implies

\[
\text{val}_{z_{1}}(f_{j}) \geq d_{j} \implies \text{val}_{z_{1}} \left( \frac{f_{j}}{T_{k-j,j}} \right) \geq d_{j} - v_{k-j,j} - 1 \implies f_{j+1} = 0 \), over \( R_{j+1}(x) \).
\]

Thus, we have proved that \( \sum_{i \in [k-j]} T_{i,j} \neq 0 \) over \( R_{j}(x) \) if

\[
\sum_{i \in [k-j]} T_{i,j+1} \neq 0 \text{ over } R_{j+1}(x) , \text{ or } , 0 \neq \left( \frac{f_{j}}{T_{k-j,j}} \right) \bigg|_{z_{1}=0} \in \mathbb{F}(z_{2})(x).
\]

Therefore induction hypothesis (4) holds. All we need to show is hypothesis (2) and second part of (3). This part is involved in the size-analysis and dlog-computation, discussed below.

**Invertibility of \( \Pi \Sigma \land \Sigma \)-circuits.** Before going into the size analysis, we want to remark that the dlog computation plays a crucial role here. The action \( \text{dlog}(\Sigma / \Sigma \land \Sigma) \in \Sigma / \Sigma \land \Sigma \land \Sigma \land \Sigma \), is of poly-size (Lemma 16). What is the action on \( \Pi \Sigma \land \Sigma \)? dlog distributes the product additively, so it suffices to work with dlog(\( \Sigma \land \Sigma \)); and we show that dlog(\( \Sigma \land \Sigma \)) \in \( \Sigma / \Sigma \land \Sigma \land \Sigma \land \Sigma \) of poly-size. Assuming these, we simplify

\[
\frac{T_{i,j}}{T_{k-j,j}} = \frac{U_{i,j} \cdot V_{k-j,j}}{U_{k-j,j} \cdot V_{i,j}} \cdot \frac{P_{i,j} \cdot Q_{k-j,j}}{P_{k-j,j} \cdot Q_{i,j}}
\]

and its dlog. Thus, using Eq. (3), \( U_{i,j} \cdot V_{k-j,j} \) grows to \( U_{i,j} \cdot V_{k-j,j} \) (and similarly \( V_{i,j+1} \)). This also means: \( U_{i,j+1} \big|_{z_{1}=0} \in \mathbb{F}(z_{2}) \setminus \{0\} \) (proving hypothesis (3), second part).

**Size analysis.** We will show that \( T_{i,j+1} \in (\Pi \Sigma / \Pi \Sigma \land \Sigma \land \Sigma / \Sigma \land \Sigma \land \Sigma \), over \( R_{j+1}(x) \), with only polynomial blowup in size. Let \( \text{size}(T_{i,j}) \leq s_{j} \), for \( i \in [k-j] \), and \( j \in [k] \). Note that, by assumption, \( s_{0} \leq s \).
11:12 Bounded Depth-4 identity testing paradigms

\[ \text{Claim 4 (Final size). } T_{1,k-1} \in (\prod \Sigma \land / \prod \Sigma \land ) \cdot (\Sigma \land \Sigma \land / \Sigma \land \Sigma \land ) \text{ of size } s^{O(k^7)}, \text{ over } R_{k-1}(x). \]

**Proof.** Steps \( j = 0 \) and \( j > 0 \) are slightly different because of the \( \Phi \). However the main idea of using power-series is the same which eventually shows that \( d \log(\Sigma \land) \in \Sigma \land \Sigma \land \).

We first deal with \( j = 0 \). Let \( A = z_1 \cdot B = \Phi(g) \in \Sigma \land \), for some \( A \in \mathbb{F}[z_2] \) and \( B \in R_1(x) \).

Note that \( A \neq 0 \) because of the map \( \Phi \). Further, size(\( B \)) \( \leq O(d \cdot \text{size}(g)) \), as a single monomial of the form \( x^e \) can produce \( d + 1 \)-many monomials. Over \( R_1(x) \),

\[
d \log(\Phi(g)) = - \frac{\partial_z (B \cdot z_1)}{A(1 - \frac{B}{A} \cdot z_1)} = - \frac{\partial_z (B \cdot z_1)}{A} \cdot \sum_{i=0}^{d_1-1} \left( B \cdot z_1^i \right)^i.
\]

(4)

\( B' \) has a trivial \( \land \Sigma \land \)-circuit of size \( O(d \cdot \text{size}(g)) \). Also, \( \partial_z (B \cdot z_1) \) has a \( \land \Sigma \land \)-circuit of size at most \( O(d \cdot \text{size}(g)) \). Using waring identity (Lemma 14), we get that each \( \partial_z (B \cdot z_1) \cdot (B/A)^i \cdot z_1^i \) has size \( O(i \cdot d \cdot \text{size}(g)) \), over \( R_1(x) \). Summing over \( i \in [d_1 - 1] \), the overall size is at most \( O(d^2 \cdot d \cdot \text{size}(g)) = O(d^3 \cdot \text{size}(g)) \), as \( d_0 = d_1 = d \).

For the \( j \)-th step, we emphasize that the degree could be larger than \( d \). Assume that syntactic degree of denominator and numerator of \( T_{i,j} \) (each in \( \mathbb{F}[x, z] \)) are bounded by \( D_j \) (it is not \( d_j \) as seen above; this is to save on the trouble of mod-computation at each step).

Of course, \( D_0 < d \leq s \).

For \( j > 0 \), the above summation in Equation 4 is over \( R_j(x) \). However the degree could be \( D_j \) (possibly more than \( d_j \)) of the corresponding \( A \) and \( B \). Thus, the overall size after the power-series expansion would be \( O(D_j^2 \cdot d \cdot \text{size}(g)) \).

Using Lemma 16, we can show that \( \text{dlog}(P_{i,j}) \in \Sigma \land \Sigma \land / \Sigma \land \Sigma \land \) (similarly for \( Q_{i,j} \)), of size \( O(D_j^2 \cdot s_j) \). Also \( \text{dlog}(U_{i,j} \cdot V_{k-j,j}) \in \sum \text{dlog}(\Sigma \land) \), i.e. sum of action of \( \text{dlog} \) on \( \Sigma \land \) (since \( \text{dlog} \) linearizes product) and it can be computed by the above formulation. Thus, \( \text{dlog}(T_{i,j}/T_{k-j,j}) \) is a sum of \( 4 \)-many \( \Sigma \land \Sigma \land / \Sigma \land \Sigma \land \) of size at most \( O(D_j^2 \cdot s_j) \) and \( 1 \)-many \( \Sigma \land \Sigma \land \) of size \( O(D_j^2 d \cdot s_j) \) (from the above power-series computation) [Note: we summed up the \( \Sigma \land \Sigma \land \)-expressions from \( \text{dlog}(\Sigma \land) \) together]. Additionally the syntactic degree of each denominator and numerator (of the \( \Sigma \land \Sigma \land / \Sigma \land \Sigma \land \)) is \( O(D_j) \). We rewrite the 4 expressions (each of \( \Sigma \land \Sigma \land / \Sigma \land \Sigma \land \) and express it as a single \( \Sigma \land \Sigma \land / \Sigma \land \Sigma \land \) using waring identity (Lemma 15), with the size blowup of \( O(D_j^2 \cdot s_j^2) \); here the syntactic degree blowup to \( O(D_j) \).

Finally we add the remaining \( \Sigma \land \Sigma \land \) circuit (of size \( O(D_j^2 s_j^2) \)) and degree \( O(d D_j) \) to get \( O(s_j^2 D_j^2 d) \). To bound this, we need to understand the degree bound \( D_j \).

Finally we need to multiply \( T_{i,j}/T_{k-j,j} \in (\prod \Sigma \land / \prod \Sigma \land \cdot (\Sigma \land \Sigma \land / \Sigma \land \Sigma \land ) \) where each \( \Sigma \land \Sigma \land \) is a product of two \( \Sigma \land \Sigma \land \) expression of size \( s_j \) and syntactic degree \( D_j \); clubbed together owing a blowup of \( O(D_j \cdot s_j^2) \). Hence multiplying it with \( \Sigma \land \Sigma \land / \Sigma \land \Sigma \land \) expression obtained from \text{dlog} computation above gives size blowup of \( s_{j+1} \cdot s_j^2 \cdot D_j^{O(1)} \cdot d \).

Computing \( T_{i,j}/T_{k-j,j} \) increases the syntactic degree ‘slowly’; which is much less than the size blowup. As mentioned before, the deg-blowup in \text{dlog}-computation is \( O(d D_j) \) and in the clearing of four expressions, it is just \( O(D_j) \). Thus, \( D_{j+1} = O(d D_j) \Rightarrow D_j = O^{(i)}(j) \).

The recursion on the size is \( s_{j+1} = s_j^2 \cdot D_j^{O(1)} \). Using \( d \leq s \) we deduce, \( s_j = (sd)O^{(i)}(j) \), In particular, \( s_{k-1}, \) size after \( k - 1 \) steps is \( s^{O(k^7)} \). This computation quantitatively establishes induction hypothesis (2). □

**Final time complexity.** The above proof actually shows that \( T_{1,k-1} \) has a ‘bloated’ circuit of size \( s^{O(k^7)} \) over \( R_{k-1}(x) \); and that the degree bound on \( z_2 \) and \( z_1 \) (over \( \mathbb{F}[z_2][z_1, x] \)), keeping denominator and numerator ‘in place’ is \( D_{k-1} = d^{O(k)} \). We note that whitebox PIT for both \( \prod \Sigma \land \land \Sigma \land \Sigma \land \land \) is in poly-time (using Thm. 10 & Lem. 18 respectively), and the proof above is constructive: we calculate \( U_{i,j+1} \) (and other terms) from \( U_{i,j} \) explicitly. Thus, this part can be done in \( s^{O(k^7)} \) time.
What remains is to test the \( z_1 \) = 0-part of induction hypothesis (4); it could short-circuit the recursion much before \( j = k - 1 \). As we mentioned before, in this case, we need to do a PIT on \( \Sigma \wedge \Sigma \wedge \) only. At the \( j \)-th step, when we substitute \( z_1 = 0 \), the size of each \( T_{i,j} \) can be at most \( s_j \) (by definition). We need to do PIT on a simpler model: \( \sum_{i,j} F(z_j) \cdot (\Sigma \wedge \Sigma \wedge /\Sigma \wedge \Sigma \wedge \). We can clear out and express this as a single \( \Sigma \wedge \Sigma \wedge /\Sigma \wedge \Sigma \wedge \) expression; with a size blowup of \( s_j^{O(k-j)} \leq (sd)^{O(j(k-j)^{d})} \). Further, use the fact that \( \max_{j \in [k-1]} j(k-j)^{d} = (k-1)^{d-1} \) (see Lemma 19). The degree bound on \( z_2 \) remains as before. Finally, use Lemma 18 for the base-case whitebox PIT. Thus, the final time complexity is \( s^{O(k \cdot \gamma)} \).

Here we also remark that in \( z_1 = 0 \) substitution, \( \Sigma \wedge \Sigma \wedge /\Sigma \wedge \Sigma \wedge \) may be undefined. However, we keep track of \( z_1 \) degree of numerator and denominator, which will be polynomially bounded as seen in the discussion above. We can easily interpolate and cancel the \( z_1 \) power to make it work.

**Bit complexity.** It is routine to show that the bit-complexity is really what we claim. Initially, the given circuit has bit-complexity \( s \). The main blowup happens due to the dlog-computation which is a poly-size blowup. We also remark that while using Lemma 15 (using Lemma 14), we may need to go to a field extension of at most \( s^{O(k)} \) (because of the \( \varepsilon(i) \) and correspondingly the constants \( \gamma_{(2)}, \ldots, \gamma(k) \), but they still are \( s^{O(k)} \)-bits). Also, Theorem 10 and Lemma 18 computations blowup bit-complexity polynomially. This concludes the proof.

**Remark.** 1. The above method does not give whitebox PIT (in poly-time) for \( \Sigma^{[k]} \Pi \Sigma \Pi^{[d]} \), as we donot know poly-time whitebox PIT for \( \Sigma \wedge \Sigma \wedge \). However, the above methods do show that whitebox-PIT for \( \Sigma^{[d]} \Pi \Sigma \Pi^{[d]} \) polynomially reduces to whitebox-PIT for \( \Sigma \wedge \Sigma \wedge \).

2. DiDI-technique can be used to give whitebox PIT for the general bloated model \text{Gen}(k, s).

3. The above proof works when the characteristic is \( \geq d \). This is because the nonzeroness remains preserved after derivation wrt \( z_1 \).

### 3.2 Proof of Theorem 2

Here we prove Theorem 2b only. The proof technique of part (a) has analogous calculations (using bottoms \( \Sigma \wedge \) instead of \( \Sigma \Pi^{[d]} \)); see Appendix D. The main idea is to use the Jacobian [5]. In fact, it solves a more general model than \( \Sigma^{[k]} \Pi \Sigma \Pi^{[d]} \).

**Transcendence basis.** Polynomials \( T_1, \ldots, T_m \) are called algebraically dependent if there exists a nonzero annihilator \( A \) s.t. \( A(T_1, \ldots, T_m) = 0 \). Transcendence degree is the size of the largest subset \( S \subseteq \{T_1, \ldots, T_m\} \) that is algebraically independent. Then \( S \) is called a transcendence basis.

**Problem 1.** Let \( \{T_i \mid i \in [m]\} \) be \( \Pi \Sigma \Pi^{[d]} \) circuits of (syntactic) degree at most \( d \) and size \( s \). Let the transcendence degree of \( T_i \)'s, \( \text{trdeg}_x(T_1, \ldots, T_m) = k \ll s \). Further, \( C(x_1, \ldots, x_m) \) be a circuit of (size + deg) \( < s' \). Design a blackbox-PIT algorithm for \( C(T_1, \ldots, T_m) \).

Trivially, \( \Sigma^{[d]} \Pi \Sigma \Pi^{[d]} \) is a very special case of the above setting. Let \( T := \{T_1, \ldots, T_m\} \). Let \( T_k := \{T_1, \ldots, T_k\} \) be a transcendence basis. For \( T_i = \prod_j g_{ij} \), we denote the set \( L(T_i) := \{g_{ij} \mid j \} \).

We want to find an explicit homomorphism \( \Psi : F[x] \to F[x, z_1, z_2] \) s.t. \( \Psi(J_x(T')) \) is of a ‘nice’ form. In the image we fix \( x \) suitably, to get a composed map \( \Psi' : F[x] \to F[z_1, z_2] \) s.t. \( \text{rk}_{F[x]}(J_x(T) = \text{rk}_{F[z]}(J_x(T')) \). Then, we can extend this map to \( \Phi : F[x] \to F[z, y, t] \) s.t. \( x_i \mapsto (\sum_{j=1}^{k} y_j f^{(j)}_{x_i}) + \Psi'(x_i) \), which is faithful [5, Lemma 2.7]; see Lemma 21. We show that the map \( \Phi \) can be efficiently constructed using a scaling and shifting map (\( \Psi \)) which is
eventually fixed by the hitting set \((H' \text{ defining } \Psi')\) of a \(\Sigma \land \Sigma \Pi^[[d]]\) circuit. Overall, \(\Phi(f)\) is a \(k + 3\)-variate polynomial for which a trivial hitting set exists.

Wlog, \(J_x(T)\) is full rank with respect to the variable set \(x_k = (x_1, \ldots, x_k)\). Thus, by assumption, \(J_{x_k}(T_k) \neq 0\) (for notation, see Section 2). We want to construct a \(\Psi\) s.t. \(\Psi(J_{x_k}(T_k))\) has an ‘easier’ PIT. We have the following identity [5, Eqn. 3.1], from the linearity of the determinant, and the simple observation that \(\partial_x(T_i) = T_i \left( \sum_j \partial_x(g_{ij})/g_{ij} \right)\),

\[
J_{x_k}(T_k) = \sum_{g_k \in L(T_k),} \left( T_1, \ldots, T_k \right) \cdot J_{x_k}(g_1, \ldots, g_k).
\]

The homomorphism \(\Psi\). Define \(\Psi : F[x] \rightarrow F[x, z_1, z_2]\) as \(x_i \mapsto z_i \cdot x_i + \Psi_1(x_i)\), where \(\Psi_1 : F[x] \rightarrow F[z_2]\), is a sparse-PIT map. The importance of \(\Psi_1\) is to ensure that \(\Psi_1(g) \neq 0\), \(\forall g \in \bigcup_r L(T)\). As \(\deg(g) \leq \delta\), \(sp(g) \leq \left( \frac{n + \delta}{\delta} \right)\). Thus, [49] (Theorem 10) gives the upper bound:

\[
de\lbrack \Psi(g) \rbrack \leq \delta \cdot \left( \left\lceil \frac{n + \delta}{\delta} \right\rceil \cdot n \cdot \log \delta \right)^2 =: D_1.
\]

Denote the ring \(R[x]\) where \(R := F[z_2][z_1]/\langle z_1^P \rangle\), and \(D := k \cdot (d - 1) + 1\). Being 1-1, \(\Psi\) is clearly a non-zero preserving map. Moreover,

\(\triangleright\) Claim 5. \(J_{x_k}(T_k) = 0 \iff \Psi(J_{x_k}(T_k)) = 0\), over \(R[x]\).

\textbf{Proof.} As \(\deg(T_i) \leq d\), each entry of the matrix can be of degree at most \(d - 1\); therefore \(\deg(J_{x_k}(T_k)) \leq k(d - 1) = D - 1\). Thus, \(\deg_{z_2}(\Psi(J_{x_k}(T_k))) < D\). Hence, the conclusion. \(\triangleright\)

Eqn. (5) implies that

\[
\Psi(J_{x_k}(T_k)) = \Psi(T_1, \ldots, T_k) \cdot \sum_{g_k \in L(T_k),} \frac{\Psi(J_{x_k}(g_1, \ldots, g_k))}{\Psi(g_1, \ldots, g_k)}.
\]

As \(T_i\) has product fanin \(s\), the top-fanin in the sum in Eqn. (6) can be at most \(s^k\). Then define,

\[
\tilde{F} := \sum_{g_k \in L(T_k),} \frac{\Psi(J_{x_k}(g_1, \ldots, g_k))}{\Psi(g_1, \ldots, g_k)}, \text{ over } R[x].
\]

\textbf{Well-definability of } \(\tilde{F}\). Note that,

\[
\Psi(z_1) \equiv \Psi_1(z_1) \mod z_1 \neq 0 \implies 1/\Psi(g_1, \ldots, g_k) \in F(z_2)[[x, z_1]].
\]

Thus, \(RHS\) is an element in \(F(z_2)[[x, z_1]]\) and taking \(\mod z_1^P\) it is in \(R[x]\). We remark that instead of minimally reducing \(\mod z_1^P\), we will work with an \(F \in F(z_2)[z_1, x]\) such that \(F = \tilde{F}\) over \(R[x]\). Further, we ensure that the degree of \(z\) is polynomially bounded.

\(\triangleright\) Claim 6. Over \(R[x]\), \(\Psi(J_{x_k}(T_k)) = 0 \iff F = 0\).

\textbf{Proof sketch.} This follows from the invertibility of \(\Psi(T_1, \ldots, T_k)\) in \(R[x]\). \(\triangleright\)

\textbf{The hitting set } \(H'\). By \(J_{x_k}(T_k) \neq 0\), and Claims 5-6, we have \(F \neq 0\) over \(R[x]\). We want to find \(H' \subseteq \mathbb{F}^n\), s.t. \(\Psi(J_{x_k}(T_k)))_{x = \alpha} \neq 0\), for some \(\alpha \in H'\) (which will ensure the rank-preservation). Towards this, we will show (below) that \(F\) has \(s^{O(\delta k)}\)-size \(\Sigma \land \Sigma \Pi^[[d]]\)-circuit over \(R[x]\). Next, Theorem 24 provides the hitting set \(H'\) in time \(s^{O(\delta^7 k \log s)}\).
Claim 7 (Main size bound). \( F \in R[x] \) has \( \Sigma^\wedge \Sigma[\Pi[\delta]] \)-circuit of size \( (s3^\delta)^O(k) \).

The proof studies the two parts of Eqn. (7)—

1. The numerator \( \Psi(J_{x_k}(g_1, \ldots, g_k)) \) has \( O(3^\delta 2^k ks) \)-size \( \Sigma^\wedge \Sigma[\Pi[\delta-1]] \)-circuit (see Lemma 8), and
2. \( 1/\Psi(g_1 \cdots g_k) \), for \( g_i \in L(T_i) \) has \( (s3^\delta)^O(k) \)-size \( \Sigma^\wedge \Sigma[\Pi[\delta]] \)-circuit; both over \( R[x] \) (see Lemma 9).

Lemma 8 (Numerator size). \( \Psi(J_{x_k}(g_1, \ldots, g_k)) \in \Sigma^\wedge \Sigma[\Pi[\delta-1]] \) of size \( O(3^\delta 2^k ks) =: s_2 \).

Proof sketch. One can show that \( J_{x_k}(g_1, \ldots, g_k) \in [s3^\delta][\Pi[k]][\Pi[\delta-1]] \) of size \( O(ks) \), where \( g_i \in L(T_i) \) (Claim 22); this basically follows from the determinant expansion which has fanin \( k! \) and the degree at the bottom is \( \leq \delta - 1 \) because of the derivative. Moreover, for a \( g \in [\Pi[\delta-1]] \), we have \( \Psi(g) \in [\Pi[\delta-1]] \) of size at most \( 3^\delta \cdot \text{size}(g) \), over \( R[x] \) (Claim 23): this follows from the fact that \( x^g \) (where \( |e_0| \leq \delta \), after shift, can produce at most \( \prod (e_i + 1) \leq e^\delta \) many monomials (for large \( n \)). Combining these, one concludes \( \Psi(J_{x_k}(g_1, \ldots, g_k)) \in [s3^\delta][\Pi[k]][\Pi[\delta-1]] \), of size \( O(3^\delta 2^k ks) \). We convert the \( \land \) gate to \( \wedge \) gate using warping identity (Lemma 14) which blowup the size by a multiple of \( 2^k - 1 \). Thus, \( \Psi(J_{x_k}(g_1, \ldots, g_k)) \in [\Sigma^\wedge \Sigma[\Pi[\delta-1]] \) of size \( O(3^\delta 2^k ks) \).

By power series expansion of expressions like \( 1/(1 - a \cdot z_1) \), one can conclude that \( 1/\Psi(g) \) has a small \( \Sigma^\wedge \Sigma[\Pi[\delta]] \)-circuit, which would further imply the same for \( 1/\Psi(g_1 \cdots g_k) \) (see below).

Lemma 9 (Denominator size). Let \( g_i \in L(T_i) \). Then, \( 1/\Psi(g_1 \cdots g_k) \) can be computed by a \( \Sigma^\wedge \Sigma[\Pi[\delta]] \)-circuit of size \( s_1 := (s3^\delta)^O(k) \), over \( R[x] \).

Proof. Let \( g \in L(T_i) \) for some \( i \). Assume, \( \Psi(g) = A - z_1 \cdot B \), for some \( A \in \mathbb{F}[z_2] \) and \( B \in R[x] \) of degree \( \delta \), with \( \text{size}(B) \leq 3^\delta \cdot s \), from Claim 23. Note that, over \( R[x] \),

\[
\frac{1}{\Psi(g)} = \frac{1}{A(1 - B/A \cdot z_1)} = \frac{1}{A} \sum_{i=0}^{D-1} \left(\frac{B}{A}\right)^i \cdot z_1^i.
\]

As, \( \text{size}(B^i) \) has a trivial \( \wedge \Sigma[\Pi[\delta]] \)-circuit (over \( R[x] \)) of size \( \leq 3^\delta \cdot s + i \); summing over \( i \in [D - 1] \), the overall size is at most \( D \cdot 3^\delta \cdot s + O(D^2) \). As \( D < k \cdot d \), we conclude that \( 1/\Psi(g) \) has \( \Sigma^\wedge \Sigma[\Pi[\delta]] \) of size \( \text{poly}(s \cdot k \cdot d3^\delta) \), over \( R[x] \). Multiplying \( k \)-many such products directly gives an upper bound of \( (s \cdot 3^\delta)^O(k) \), using Lemma 15 (basically, warping identity).

Proof of Claim 7. Combining Lemmas 8-9, observe that \( \Psi(J_{x_k}(g_1, \ldots, g_k))/\Psi(g_1 \cdots g_k) \) has \( \Sigma^\wedge \Sigma[\Pi[\delta]] \)-circuit of size at most \( (s_1 \cdot s_2)^2 = (s \cdot 3^\delta)^O(k) \), over \( R[x] \), using Lemma 15. Summing up at most \( s^k \) many terms (by defn. of \( F \)), the size still remains \( (s \cdot 3^\delta)^O(k) \).

Degree bound. As, syntactic degree of \( T_i \) are bounded by \( d \), and \( \Psi \) maintain \( \text{deg}_r = \text{deg}_z \), we must have \( \text{deg}_r(\Psi(J_{x_k}(g_1, \ldots, g_k))) = \text{deg}_r(J_{x_k}(g_1, \ldots, g_k)) \leq D - 1 \). Similarly, by assumption \( \text{deg}_z(\Psi(g)) \leq D_1 := \text{poly}(n^\delta) \), and thus \( \text{deg}_z(\Psi(J_{x_k}(g_1, \ldots, g_k))) \leq D_1 \cdot k \). Note that, Lemma 8 actually works over \( \mathbb{F}[x, z] \) and thus there is no additional degree-blow up (in \( z \)). However, there is some degree blowup in Lemma 9, due to Eqn. (8).

Note that Eqn. (8) shows that over \( R[x] \),

\[
\frac{1}{\Psi(g)} = \left(\frac{1}{A}\right) \cdot \left(\sum_{i=0}^{D-1} A^{D-1-i} z_1^i \cdot B^i\right) =: \frac{p(x, z)}{q(z_2)},
\]
where \( q(z_2) = A^D \). We think of \( p \in \mathbb{F}[x, z] \) and \( q \in \mathbb{F}[z_2] \). It follows that \( \deg_{z_2}(q) \leq D_1 \cdot D \). Also, \( \deg_{z_2}(\Psi(q)) \leq \delta \) implies \( \deg_{z_2}(p) \leq \deg_{z_2}((B z_1)^{D-1}) \leq \delta \cdot (D - 1) \). Since, \( \deg_{z_2}(\Psi(q)) \leq D_1 \), by assumption, \( \deg_{z_2}(p) \leq \max, \deg_{z_2}(A^{D-1}, B_i) \leq D_1 \cdot (D - 1) \).

Finally, denote \( 1/\Psi(g_1 \cdots g_k) =: P_{g_1, \ldots, g_k} / Q_{g_1, \ldots, g_k} \), over \( \mathbb{R}[x] \). This is just multiplying \( k \)-many \( (p/q) \)’s; implying a degree blowup by a multiple of \( k \). In particular,

\[
\deg_{z_1}(P_{g_1, \ldots, g_k}) \leq \delta \cdot k \cdot (D - 1),
\]

\[
\deg_{z_2}(P_{g_1, \ldots, g_k}) \leq D_1 \cdot (D - 1) \cdot k,
\]

\[
\deg_{z_2}(Q_{g_1, \ldots, g_k}) \leq D_1 \cdot D \cdot k.
\]

Thus, in Eqn. (7), summing up \( s^k \)-many terms gives an expression (over \( \mathbb{R}[x] \)):

\[
F = \sum_{g_1 \in L(T_1), \ldots, g_k \in L(T_k)} \Psi(J_{g_1}(g_1, \ldots, g_k)) \cdot \left( \frac{P_{g_1, \ldots, g_k}}{Q_{g_1, \ldots, g_k}} \right) =: \frac{P(x, z)}{Q(z_2)}.
\]

Verify that \( Q \in \mathbb{F}[z_2] \) is of degree at most \( s^k \cdot D_1 \cdot D \cdot k = s^{O(k)} \cdot \text{poly}(n^\delta) \) (since \( k, d < s \)).

A similar bound also holds for \( \deg_{z_2}(P) \). The degree of \( z_1 \) also remains bounded by

\[
\max_{g_i \in L(T_i), i \in [k]} \deg_{z_1}(P_{g_1, \ldots, g_k}) + \delta k \leq \text{poly}(s).
\]

Using the degree bounds, we finally have \( P \in \mathbb{F}[x, z] \) as a \( \Sigma \land \Sigma \Pi \Sigma \Pi \)-circuit (over \( \mathbb{F}(z) \)) of size \( s^{O(\delta) \cdot (s^\delta)^{O(k)} = 3^{O(\delta k)} s^{O(k + \delta)} =: s_3} \).

We want to construct a set \( H' \subseteq \mathbb{F}^n \) such that the action \( P(H', z) \neq 0 \). Using [25] (Theorem 24), we conclude that it has \( s^{O(\delta \log s)} = s^{O(\delta^2 k \log s)} \) size hitting set which is constructible in a similar time. Hence, the construction of \( \Phi \) follows, making \( \Phi(f) \) a \( k + 3 \) variate polynomial. Finally, by the obvious degree bounds of \( y, z, t \) from the definition of \( \Phi \), we get the blackbox PIT algorithm with time-complexity \( s^{O(\delta^2 k \log s)} \); finishing Theorem 2b.

We could also give the final hitting set for the general problem.

**Solution to Problem 1.** We know that \( C(T_1, \ldots, T_m) = 0 \iff E := \Phi(C(T_1, \ldots, T_m)) = 0 \).

Since, \( H' \) can be constructed in \( s^{O(\delta^2 k \log s)} \)-time, it is trivial to find hitting set for \( E|_{H'} \) (which is just a \( k + 3 \)-variate polynomial with the aforementioned degree bounds). The final hitting set for \( E \) can be constructed in \( s^{O(k)} \cdot s^{O(\delta^2 k \log s)} \)-time.

**Remark.** 1. As Jacobian Criterion (Fact 2) holds when the characteristic is greater than \( d^\text{trdeg} \), it is easy to conclude that our theorem holds for all fields of char greater than \( d^k \).

2. The above proof gives an efficient reduction from blackbox PIT for \( \Sigma^{[k]} \Pi \Sigma \Pi^{[k]} \) circuits to \( \Sigma \land \Sigma \Pi \Sigma \Pi^{[k]} \) circuits. In particular, a poly-time hitting set for \( \Sigma \land \Sigma \Pi \Sigma \Pi^{[k]} \) circuits would put PIT for \( \Sigma^{[k]} \Pi \Sigma \Pi^{[k]} \) in \( \mathcal{P} \).

3. Also, DiDi-technique (of Theorem 1) directly gives a blackbox algorithm, but the complexity is exponentially worse (in terms of \( k \) in the exponent) for its recursive blowups.

**4 Conclusion**

This work introduces the powerful DiDi-technique and solves three open problems in PIT for depth-4 circuits, namely \( \Sigma^{[k]} \Pi \Sigma \Pi^{[k]} \) (blackbox) and \( \Sigma^{[k]} \Pi \Sigma \land \) (both whitebox and blackbox). Here are some immediate questions of interest which require rigorous investigation.
1. Can the exponent in Theorem 1 be improved to $O(k)$? Currently, it is exponential in $k$.
2. Can we improve Theorem 2b to $s^{O(\log \log s)}$ (like in Theorem 2a)?
3. Can we design a polynomial-time PIT for $\Sigma^k \Pi^k$?
4. Design a poly-time PIT for $\Sigma \land \Sigma^k \Pi^k$ circuits (i.e. unbounded top-fanin)?
5. Can we solve PIT for $\Sigma^k \Pi^k$ circuits in subexponential-time?
6. Can we design a subexponential-time PIT for rational functions of the form $\Sigma (1/\Sigma \land \Sigma)$ or $\Sigma (1/\Sigma \land \Sigma)$ (for unbounded top-fanin)?

References

Bounded Depth-4 identity testing paradigms


11:20 Bounded Depth-4 identity testing paradigms


Bounded Depth-4 identity testing paradigms


A Basic tools from algebraic complexity

There have been a lot of work on sparse-PIT, for details see [13, 49] and references therein. Eventually, there is a poly-time hitting set, for a proof see [76, Thm. 2.1]

Theorem 10 ([49]). Let \( p(x) \in \mathbb{F}[x] \) with individual degree at most \( d \) and sparsity at most \( m \). Then, there exists \( 1 \leq r \leq (mn \log d)^2 \), such that \( p(y^r, y^d, \ldots, y^{d^{m-1}}) \) \( \neq 0 \), mod \( y^r - 1 \).

An ABP is a layered directed acyclic graph with \( q+1 \) many layers of vertices \( \{V_0, \ldots, V_q\} \) and a source \( a \) and a sink \( b \) such that all the edges in the graph only go from \( a \) to \( V_0 \), \( V_i \) to \( V_{i+1} \) for any \( i \in [q] \), and \( V_q \) to \( b \). The edges have univariate polynomials as their weights. The ABP is said to compute the polynomial

\[
f(x) = \sum_{p \in \text{paths}(a,b)} \prod_{e \in p} W(e),
\]

where \( W(e) \) is the weight of the edge \( e \). The ABP has width-\( w \) if \( |V_i| \leq w, \forall i \in \{0, \ldots, q\} \). Formally, it computes polynomials of the form \( A^T (\prod_{i \in [q]} D_i) B \), where \( A, B \in \mathbb{F}^{w \times 1}[x] \), and \( D_i \in \mathbb{F}^{w \times w}[x] \), where entries are univariate polynomials.

Definition 11 (Read-once oblivious ABP (ROABP)). An ABP is called a read-once oblivious ABP (ROABP) if the edge weights are univariate polynomials in distinct variables across layers. Formally, there is a permutation \( \pi \) on the set \( [q] \) such that the entries in the \( i \)-th matrix \( D_i \) are univariate polynomials over the variable \( x_{\pi(i)} \), i.e., they come from the polynomial ring \( \mathbb{F}[x_{\pi(i)}] \).

A polynomial \( f(x) \) is said to be computed by width-\( w \) ROABPs in any order, if for every permutation \( \sigma \) of the variables, there exists a width-\( w \) ROABP in the variable order \( \sigma \) that computes the polynomial \( f(x) \). There have been quite a few results on blackbox PIT for ROABPs [28, 27, 37] and the current best known algorithm works in quasipolynomial time.

Theorem 12 ([37]). For \( n \)-variate, individual-degree-\( d \) polynomials computed by width-\( w \) ROABPs in any order, a hitting set of size \( (ndw)^{O(\log \log w)} \) can be constructed.

B Details for Section 3.1

Here is an important lemma which shows that coefficient of \( y^r \) of a polynomial \( f(x, y) \in \mathbb{F}[x, y] \), computed by a \( \Sigma \wedge \Sigma \wedge \) circuit, can be computed by a small \( \Sigma \wedge \Sigma \wedge \) circuit.

Lemma 13 (Coefficient extraction). Let \( f(x, y) \in \mathbb{F}[y][x] \) be computed by a \( \Sigma \wedge \Sigma \wedge \) circuit of size \( s \) and degree \( d \). Then, \( \text{coeff}_y(f) \in \mathbb{F}[x] \) can be computed by a small \( \Sigma \wedge \Sigma \wedge \) circuit of size \( O(sd) \), over \( \mathbb{F}[x] \).

Proof sketch. Let, \( f = \sum_{i=0}^d a_i \cdot g_i^{e_i} \). Of course, \( e_i \leq s \) and \( \deg_y(f) \leq d \). Thus, write \( f = \sum_{i=0}^d f_i \cdot y^{e_i} \), where \( f_i \in \mathbb{F}[x] \). We can interpolate on \( d+1 \)-many distinct points \( y \in \mathbb{F} \) and conclude that \( f_i \) has a \( \Sigma \wedge \Sigma \wedge \) circuit of size at most \( O(sd) \). □
The next identity gives us a way to write a product of a few powers as a sum of powers, using simple interpolation. For a more algebraic proof, see \cite[Proposition 4.3]{15}.

\begin{lemma} \textbf{(Waring Identity for a monomial)}. Let \( M = x_1^{b_1} \cdots x_k^{b_k} \), where \( 1 \leq b_1 \leq \ldots \leq b_k \), and roots of unity \( \mathcal{Z}(i) := \{ z \in \mathbb{C} : z^{b_i+1} = 1 \} \). Then,
\[
M = \sum_{\varepsilon(1) \leq \ldots \leq \varepsilon(k)} \gamma_{\varepsilon(2), \ldots, \varepsilon(k)} \cdot (x_1 + \varepsilon(2)x_2 + \ldots + \varepsilon(k)x_k)^d,
\]
where \( d := \deg(M) = b_1 + \ldots + b_k \), and \( \gamma_{\varepsilon(2), \ldots, \varepsilon(k)} \) are scalars (\( \rho_k(M) := \prod_{i=2}^{k}(b_i + 1) \) many).
\end{lemma}

\textbf{Remark}. We actually need not work with \( \mathbb{F} = \mathbb{C} \). We can go to a small extension (at most \( d^k \)), for a monomial of degree \( d \), to make sure that \( \varepsilon(i) \) exists.

The next lemma shows that \( \Sigma \land \Sigma \land \) is closed under multiplication.

\begin{lemma} \textbf{(Differentiation)}. Let \( f_i(x, y) \in \mathbb{F}[y][x] \), of syntactic degree \( \leq d_i \), be computed by a \( \Sigma \land \Sigma \land \) circuit of size \( s_i \), for \( i \in [k] \) (wrt \( x \)). Then, \( f_1 \cdots f_k \) has \( \Sigma \land \Sigma \land \) circuit of size \( O((d_2 + 1) \cdots (d_k + 1) \cdot s_1 \cdots s_k) \).
\end{lemma}

\textbf{Proof}. Let \( f_i = \sum_{j} f_{ij}^{e_{ij}} \); by assumption \( e_{ij} \leq d_i \) (by assumption). Using Lemma 14, \( f_{ij}^{e_{ij}} \cdots f_{jk}^{e_{jk}} \) has size at most \( (d_2 + 1) \cdots (d_k + 1) \cdot (\sum_{i \in [k]} \text{size}(f_{ij})) \), for indices \( j_1, \ldots, j_k \). Summing up for all \( s_1 \cdots s_k \) many products (atmost) gives the upper bound.

The next lemma shows that \( \Sigma \land \Sigma \land \) is closed under differentiation.

\begin{lemma} \textbf{(Valuation)}. Consider a polynomial \( f \in \mathbb{F}(x, y) \) such that \( \text{val}_y(f) \geq 0 \). Then, \( f \in \mathbb{F}(x)[[y]] \cap \mathbb{F}(x, y) \).
\end{lemma}

\textbf{Proof sketch}. Lemma 13 shows that each \( f_e \) has \( O(sd) \) size circuit where \( f = \sum_e f_e y^e \).

Doing this for each \( e \in [0, d] \) gives a blowup of \( O(sd^2) \).

The next lemma shows that non-negative valuation corresponds to a power-series.

\begin{lemma} \textbf{(PIT for \( \Sigma \land \Sigma \land \) circuits)}. Let \( P \in \Sigma \land \Sigma \land \) of size \( s \). Then, there exists a \( \text{poly}(s) \) (resp. \( s^{\Omega(\log \log s)} \)) time whitebox (resp. blackbox) PIT for the same.
\end{lemma}
11:24 Bounded Depth-4 identity testing paradigms

**Proof sketch.** We show that any $g(x)^e = (g_1(x_1) + \ldots + g_n(x_n))^e$, where $\deg(g_i) \leq s$ can be written as $\sum_{j=1}^n h_{j1}(x_1) \cdots h_{jn}(x_n)$, for some $h_{jl} \in \mathbb{F}[x]$ of degree at most $e$. Define, $G := (y + g_1) \cdots (y + g_n) - y^n$. In its $e$-th power, notice that the leading-coefficient is $\text{coef}_{y^n}(G^e) = g^e$. So, interpolate on $e(n-1)+1$ many points $(y = \beta_i \in \mathbb{F})$ to get

$$\text{coef}_{y^n}(G^e) = \sum_{i=1}^{e(n-1)+1} \alpha_i G^e(\beta_i).$$

Now, expand $G^e(\beta_i) = ((\beta_i+g_1) \cdots (\beta_i+g_n)-\beta_i^n)^e$, by binomial expansion (without expanding the inner $n$-fold product). The top-fanin can be atmost $s \cdot (e+1) \cdot (e(n-1)+1) = O(se^2n)$. The individual degrees of the intermediate univariates can be at most $e$. Thus, it can be computed by an ROABP (of any order) of size at most $O(s^2e^3n)$.

Now, if $f = \sum_{j \in [s]} f_j^i$ is computed by a $\Sigma \land \Sigma \land$ circuit of size $s$, then clearly, $f$ can also be computed by an ROABP (of any order) of size at most $O(s^6)$. So, the whitebox PIT follows from [70], while the blackbox PIT follows from Theorem 12.

For the time-complexity bound, we need optimization of the following function:

**Lemma 19.** Let $k \in \mathbb{N}$, and $h(x) := x(k-x)^7^e$. Then, $\max_{i \in [k-1]} h(i) = h(k-1)$.

**Proof sketch.** Differentiate to get $h'(x) = (k-x)^7^e - 7e x^7^e + x(k-x)(\log 7)^7^e = 7^e \cdot [x^2(\log 7) + x(k\log 7 - 2)]$. It vanishes at $x = \left(\frac{1}{2} - \frac{1}{\log 7}\right) + \sqrt{\left(\frac{1}{2} - \frac{1}{\log 7}\right)^2 - \frac{k}{\log 7}}$. Thus, $h$ is maximized at the integer $x = k-1$.

### C Details for Section 3.2

**Definition 20 (Faithful hom.).** $\Phi : \mathbb{F}[x] \rightarrow \mathbb{F}[y]$ is faithful for $T$ if $\text{trdeg}_\mathbb{F}(T) = \text{trdeg}_\mathbb{F}(\Phi(T))$.

The following fact about faithful maps is from [5, Thm. 2.4].

**Fact 1 (Faithful is useful).** For any $C \in \mathbb{F}[y_1, \ldots, y_m]$, $C(T) = 0 \iff C(\Phi(T)) = 0$.

Here is an important criterion about the jacobian matrix which basically shows that it preserves algebraic independence.

**Fact 2 (Jacobian criterion).** Let $f \subset \mathbb{F}[x]$ be a finite set of polynomials of degree at most $d$, and $\text{trdeg}_\mathbb{F}(f) \leq r$. If $\text{char}(\mathbb{F}) = 0$, or $\text{char}(\mathbb{F}) > d'$, then $\text{trdeg}_\mathbb{F}(f) = \text{rk}_\mathbb{F}(\mathcal{J}_x(f))$.

The following lemma (& the proof) is similar to [5, Lem. 2.7]. It is a recipe to ‘drastically’ reduce variables, if trdeg is small.

**Lemma 21 (Recipe for faithful maps).** Let $T \in \mathbb{F}[x]$ be a finite set of polynomials of degree at most $d$ and $\text{trdeg}_\mathbb{F}(T) \leq r$, and $\text{char}(\mathbb{F}) = 0$ or $d'$. Let $\Psi : \mathbb{F}[x,y] \rightarrow \mathbb{F}[z_1, z_2]$ such that $\text{rk}_\mathbb{F}(\mathcal{J}_x(T)) = \text{rk}_\mathbb{F}(\Psi(\mathcal{J}_x(T)))$.

Then, the map $\Phi : \mathbb{F}[x] \rightarrow \mathbb{F}[z, t, y]$, such that $x_i \mapsto (\sum_{j} y_{ij} t^j) + \Psi(x_i)$, is a faithful homomorphism for $T$. 
C.1 Technical Details for Theorem 2b

Claim 22. Let \( g_i \in L(T_i) \), where \( T_i \in \Pi\Sigma\Pi^{[d]} \) of size at most \( s \), then \( J_{x_k}(g_1, \ldots, g_k) \in \Sigma^{[k]}\Pi^{[k]}\Sigma\Pi^{[d-1]} \) of size \( O(k!ks) \).

Proof sketch. Each entry of the matrix has degree at most \( \delta - 1 \). Trivial expansion gives \( k! \) top-fanin where each product (of fanin \( k \)) has size \( \sum_i \text{size}(g_i) \). As, \( \text{size}(T_i) \leq s \), trivially each size(\( g_i \)) \( \leq s \). Therefore, the total size is \( k! \cdot \sum_i \text{size}(g_i) = O(k!ks) \).

Claim 23. Let \( g \in \Sigma^{[d]} \), then \( \Psi(g) \in \Sigma^{[d]} \) of size \( 3^\delta \cdot \text{size}(g) \) (for \( n \gg \delta \)).

Proof sketch. Each monomial \( x^e \) of degree \( \delta \), can produce \( \prod_i (e_i + 1) \leq ((\sum_i e_i + n)/n)^n \leq (\delta/n + 1)^n \)-many monomials, by AM-GM inequality as \( \sum_i e_i \leq \delta \). As \( \delta/n \to 0 \), we have \((1 + \delta/n)^n \to e^\delta \). As \( e < 3 \), the upper bound follows.

[25, Prop. 4.18] gave the first nontrivial PIT for \( \Sigma^{\land} \Sigma \Pi^{[d]} \) circuits:

\[ \text{Theorem 24 ([25]). There is a poly}(n, d, \delta \log s)\text{-explicit hitting set of size } (nd)^{O(\delta \log s)} \text{ for the class of } n\text{-variate, degree-(\leq d)} \text{ polynomials } f(x), \text{ computed by } \Sigma^{\land} \Sigma \Pi^{[d]} \text{-circuit of size } s. \]

D Proof sketch of Theorem 2a: Similar to Section 3.2

Similar to Theorem 2b, we generalize this theorem and prove for a much bigger class of polynomials.

Problem 2. Let \( \{T_i \mid i \in [m]\} \) be \( \Pi\Sigma\Pi \) circuits of (syntactic) degree at most \( d \) and size \( s \).

Let the transcendence degree of \( T_i \) ’s, trdeg_\( \Sigma \)(\( T_1, \ldots, T_m \)) := \( k \ll s \). Further, \( C(x_1, \ldots, x_m) \) be a circuit of size + degree < \( s \). Design a blackbox-PIT algorithm for \( C(T_1, \ldots, T_m) \).

It is trivial to see that \( \Sigma^{[k]}\Pi\Sigma \Pi \) is a very special case of the above settings. We will use the same idea (\& notation) as in Theorem 2b, using the Jacobian technique. The main idea is to come up with \( \Phi \) map, and correspondingly the hitting set \( H' \). If \( g \in L(T_i) \), then \( \text{size}(g) \leq O(dn) \). We also note that \( D_1 \), which is an upper bound on \( \deg_\Sigma \Psi(g) = \text{poly}(n, d) \) (Lemma 10). The \( D \) (and hence \( R[x] \)) remains as before. Claims 5-6 hold similarly. We will construct the hitting set \( H' \) by showing that \( F \) has a small \( \Sigma^{\land}\Sigma \Pi \) circuit over \( R[x] \).

Note that, Claim 22 remains the same for \( \Sigma^{\land}\Sigma \Pi \) (implying the same size blowup). However, Claim 23, the size blowup is \( \text{O}(d\text{size}(g)) \), because each monomial \( x^e \) can only produce \( d + 1 \) many monomials. Therefore, similar to Lemma 9, one can show that \( \Psi(J_{x_k}(g_1, \ldots, g_k)) \in \Sigma^{\land}\Sigma \Pi, \text{ of size } O(2^k k!ks) \). Similarly, the size in Lemma 8 can be replaced by \( s^{O(k)} \). Therefore, we get (similar to Claim 7):

Claim 25. \( F \in R[x] \) has \( \Sigma^{\land}\Sigma \Pi \) -circuit of size \( s^{O(k)} \).

Next, the degree bound also remains the same (except the parameter \( D_1 \) which is now \( \text{poly}(nd) \)). Following the same footsteps, it is not hard to see that the degree bound of \( z_2 \) on \( P \) and \( Q \), where \( F = P(x, z)/Q(z_2) \), is \( s^{O(k)} \text{poly}(nd) \), while degree bound on \( z_1 \) remains \( \text{poly}(k!s) \). Therefore, \( P \in F[x, z] \) has \( \Sigma^{\land}\Sigma \Pi \) -circuit of size \( s^{O(k)} \).

We want to construct a set \( H' \subseteq \mathbb{F}^n \) such that the action \( P(H', z) \neq 0 \). By Theorem 18, we conclude that it has \( s^{O(k \log \log s)} \) size hitting set which is constructible in a similar time.

Hence, the construction of map \( \Phi \) and the theorem follows (from \( z \)-degree bound).

Solution to Problem 2. We know that \( C(T_1, \ldots, T_m) = 0 \iff E := \Phi(C(T_1, \ldots, T_m)) = 0 \).

Since, \( H' \) can be constructed in \( s^{O(k \log \log s)} \) time, it is trivial to find hitting set for \( E|_{H'} \) (which is just a \( k + 3 \)-variate polynomial with the aforementioned degree bounds). The final hitting set for \( E \) can be constructed in \( s^{O(k)} \cdot s^{O(k \log \log s)} \) time.
Algorithm for Theorem 1

The whitebox PIT for Theorem 1, that is discussed in Section 3.1, appears (below) as Algorithm 1.
Algorithm 1 Whitebox PIT Algorithm for $\Sigma^k\Pi^\Sigma\Lambda$-circuits

Input: $f = T_1 + \ldots + T_k \in \Sigma^k\Pi^\Sigma\Lambda$, a whitebox circuit of size $s$ over $\mathbb{F}[x]$.

Output: 0, if $f \equiv 0$, and 1, if it is non-zero.

1. Let $\Psi: \mathbb{F}[x] \rightarrow \mathbb{F}[z_2]$, be a sparse-PIT map, using [49] (Theorem 10). Apply it on $f$
and check whether $\Psi(f) \equiv 0$. If non-zero, output 1 otherwise, apply
$\Phi: x_i \rightarrow z_1 \cdot x_i + \Psi(x_i)$ on $f$. Check $\sum_{i \in [k-1]} \partial_{z_1}(\Phi(T_i)/\Phi(T_k)) \equiv 0 \mod z_1^{d_1}$
(d_1 := s) as follows:
2. Consider each $T_{i,1} := \partial_{z_1}(\Phi(T_i)/\Phi(T_k))$ over $R_1(x)$, where $R_1 := \mathbb{F}(z_2)[z_1]/(z_1^{d_1})$.
Use dlog computation (Claim 4), to write each $T_{i,1}$ in a ‘bloated’ form as
$$(\Pi^\Sigma / \Pi^\Sigma\Lambda) \cdot (\Sigma\Lambda / \Sigma\Lambda\Lambda)$$.
3. for $j \leftarrow 1$ to $k-1$ do
4. Reduce the top-fanin at each step using ‘Divide & Derive’ technique. Assume
that at $j$-th step, we have to check the identity:
$$\sum_{i \in [k-j]} T_{i,j} \equiv 0 \text{ over } R_j(x)$$, where $R_j := \mathbb{F}(z_2)[z_1]/(z_1^{d_1})$,
each $T_{i,j}$ has a $(\Pi^\Sigma / \Pi^\Sigma\Lambda) \cdot (\Sigma\Lambda / \Sigma\Lambda\Lambda)$ representation and therein each
$\Pi^\Sigma\Lambda \mid z_1 = 0 \in \mathbb{F}(z_2) \setminus \{0\}$.
1. Compute $v_{k-j,j} := \min\{\text{val}_{z_1}(T_{i,j})\}$: by reordering it is for $i = k - j$. To compute $v_{k-j,j}$,
use coefficient extraction (Lemma 13) and $\Sigma\Lambda / \Sigma\Lambda\Lambda$ -circuit PIT (Lemma 18).
2. ‘Divide’ by $T_{k-j,j}$ and check whether
$$\left(\sum_{i \in [k-j-1]} \frac{T_{i,j}}{T_{k-j,j}} + 1\right) \bigg|_{z_1 = 0} \equiv 0.$$
Note: this expression is in $(\Sigma\Lambda / \Sigma\Lambda\Lambda)$. Use — (1) $\Pi^\Sigma\Lambda \mid z_1 = 0 \in \mathbb{F}(z_2)$, and (2)
closure of $\Sigma\Lambda\Lambda$ under multiplication. Finally, do PIT on this by Lemma 18.
3. If it is non-zero, output 1, otherwise ‘Derive’ wrt $z_1$ and ‘Induct’ on
$$\left(\sum_{i \in [k-j-1]} \partial_{z_1}(T_{i,j}/T_{k-j,j})\right) \equiv 0 \text{ over } R_{j+1}(x)$$
where $R_{j+1} := \mathbb{F}(z_2)[z_1]/(z_1^{d_1} - v_{k-j,j-1})$.
4. Again using dlog (Claim 4), show that $T_{i,j+1} := \partial_{z_1}(T_{i,j}/T_{k-j,j})$ has small
$(\Pi^\Sigma / \Pi^\Sigma\Lambda) \cdot (\Sigma\Lambda / \Sigma\Lambda\Lambda)$-circuit over $R_{j+1}(x)$. So call the algorithm on
$$\sum_{i \in [k-j-1]} T_{i,j+1} \equiv 0.$$
5. $j \leftarrow j + 1.$
6. At the end, $j = k - 1$. Do PIT (Lemma 18) on the single
$(\Pi^\Sigma / \Pi^\Sigma\Lambda) \cdot (\Sigma\Lambda / \Sigma\Lambda\Lambda)$-circuit, over $R_{k-1}(x)$. If it is zero, output 0
otherwise output 1.
7. Words of caution: Throughout the algorithm there are intermediate expressions to be
stored compactly. Think of them as ‘special’ circuits in $x$, but over the function-field
$\mathbb{F}(z)$. Keep track of their degrees wrt $z_1, z_2$; and that of the sizes of their fractions
represented in ‘bloated’ circuit form.