Algebraic independence over positive characteristic: New criterion and applications to locally low algebraic rank circuits

Anurag Pandey[1], Nitin Saxena[2], and Amit Sinhababu[2]

1 **MPI for Informatics & Saarland University, Department of Computer Science,** apandey@mpi-inf.mpg.de

2 **Department of CSE, Indian Institute of Technology Kanpur,** {nitin, amitks}@cse.iitk.ac.in

Abstract

The motivation for this work comes from two problems– test algebraic independence of arithmetic circuits over a field of small characteristic, and generalize the structural property of algebraic dependence used by (Kumar, Saraf CCC'16) to arbitrary fields.

It is known that in the case of zero, or large characteristic, using a classical criterion based on the Jacobian, we get a randomized poly-time algorithm to test algebraic independence. Over small characteristic, the Jacobian criterion fails and there is no subexponential time algorithm known. This problem could well be conjectured to be in RP, but the current best algorithm puts it in $\mathrm{NP}^{\#\mathsf{P}}$ (Mittmann, Saxena, Scheiblechner Trans.AMS'14). Currently, even the case of two bivariate circuits over $\mathbb{F}_2$ is open. We come up with a natural generalization of Jacobian criterion, that works over all characteristic. The new criterion is efficient if the underlying *inseparable degree* is promised to be a constant. This is a modest step towards the open question of fast independence testing, over finite fields, posed in (Dvir, Gabizon, Wigderson FOCS'07).

In a set of linearly dependent polynomials, any polynomial can be written as a linear combination of the polynomials forming a basis. The analogous property for algebraic dependence is false, but a property approximately in that spirit is named as "functional dependence" in (Kumar, Saraf CCC'16) and proved for zero or large characteristic. We show that functional dependence holds for *arbitrary* fields, thereby answering the open questions in (Kumar, Saraf CCC'16). Following them we use the functional dependence lemma to prove the first exponential lower bound for *locally low algebraic rank* circuits for *arbitrary* fields (a model that strongly generalizes homogeneous depth-4 circuits). We also recover their quasipoly-time hitting-set for such models, for fields of characteristic smaller than the ones known before.

Our results show that approximate functional dependence is indeed a more fundamental concept than the Jacobian as it is field independent. We achieve the former by first picking a "good" transcendence basis, then translating the circuits by new variables, and finally approximating them by truncating higher degree monomials. We give a tight analysis of the "degree" of approximation needed in the criterion. To get the locally low algebraic rank circuit applications we follow the known shifted partial derivative based methods.

1998 ACM Subject Classification I.1 Symbolic and Algebraic Manipulation, F.2 Analysis of Algorithms and Problem Complexity

Keywords and phrases independence, transcendence, finite field, Hasse-Schmidt, Jacobian, differential, inseparable, circuit, identity testing, lower bound, depth-4, shifted partials

Digital Object Identifier 10.4230/LIPIcs.MFCS.2016.75

1 Introduction

Algebraic dependence is a fundamental concept in algebra that captures algebraic/polynomial relationship of objects like numbers, polynomials, rational functions or power series, over some field. Here we define algebraic dependence of polynomials, since in this work we deal only with polynomials. Polynomials $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ are called *algebraically dependent* over field k if and only if there exists a nonzero polynomial $A(y_1, \ldots, y_m) \in \mathbb{F}[y_1, \ldots, y_m]$ such that $A(f_1, \ldots, f_m) = 0$ and such an A is called an *annihilating polynomial* of $f_1, \ldots, f_m$. If no such nonzero polynomial A exists, the given polynomials are called *algebraically independent* over k.

For example, $f_1 = (x+y)^2$ and $f_2 = (x+y)^3$ are algebraically dependent over any field, as $y_1^3 - y_2^2$ is an annihilating polynomial. Polynomials $x+y$ and x^p+y^p are dependent over $\mathbb{F}_p$, but independent over $\mathbb{Q}$. Monomials $x_1, \ldots, x_n$ are examples of algebraically independent polynomials over any field.

Algebraic dependence can be viewed as a generalization of linear dependence as the former captures algebraic relationships of any degree, whereas the latter captures linear relationships. Algebraic dependence shares a few combinatorial properties (known as matroid properties [38]) with linear dependence. For example, if a set of polynomials are algebraically independent then any subset of them are algebraically independent. The *transcendence degree* (trdeg or algRank) of a set of polynomials is defined as the maximal number of algebraically independent polynomials and it is well defined thanks to the matroid properties. The concepts of rank and basis in linear algebra have analogs here as transcendence degree and *transcendence basis* respectively.

The concept of algebraic independence is useful in several areas of mathematics: field theory, commutative algebra, algebraic geometry, invariant theory, theory of algebraic matroids. It has found interesting applications in computer science as well. For example, [35] used algebraic dependence in analysis of program invariants of arithmetic straight line programs. To prove lower bounds on the formula size of determinant, [25] also used transcendence degree as a tool. [10, 12] constructed explicit deterministic randomness extractors for sources which are polynomial maps over finite fields. [11] gives a cryptography application, using algebraic characterization of entropy of low degree polynomials. [6, 2, 34] used it for designing faster deterministic hitting-sets for some interesting cases of the polynomial identity testing problem (PIT) and proving circuit lower bounds. [8] used algebraic independence of polynomials to show the hardness of a parameterized counting problem.

An example relevant to computer science is to compute the "entropy" of a given polynomial map $\phi : (x_1, \ldots, x_n) \mapsto (f_1, \ldots, f_n)$ in the space $\mathbb{F}_q^n$, where q is a power of $p = 2$ (more, generally, p grows as a polynomial in the input size). This turns out to be a question of computing the transcendence degree of the polynomials $f_1, \ldots, f_n$ [10]. For constant p, there are no good methods known. Our work improves the state of the art in this regime.

To discuss the complexity of algebraic independence testing, we have to specify the representation of input polynomials. An *arithmetic circuit* is a directed acyclic graph consisting of addition ($+$) and multiplication ($\times$) gates as nodes, takes variables $x_1, \ldots, x_n$ and field constants as input (leaves), and outputs a polynomial $f(x_1, \ldots, x_n)$. This is a succinct representation of multivariate polynomials, as polynomials of high degree (or having many monomials) can be represented by small circuits.

Perron [39, 40] gave a bound on the degree of the minimal annihilating polynomial, proving that it is bounded by the product of the degrees of the input polynomials. This bound was subsequently slightly improved in [26, 6]. Perron's bound gives us the brute-force approach. It reduces the problem of computing annihilating polynomial to solving

an exponential sized system of linear equations and this can be done in PSPACE. Thus, PSPACE is the "trivial" complexity upper bound for algebraic independence testing, over any field.

The degree bound on the minimal annihilating polynomial happens to be tight. We can give examples of n quadratic polynomials, such that the degree of their minimal annihilating polynomial is 2^n [26]. There is a hardness result known [26], that shows that computing even the constant term of the annihilating polynomial is NP-hard, and that annihilating polynomial is not of polynomial size in general, unless the polynomial hierarchy collapses.

It turns out that the decision version, i.e. checking if the polynomials are algebraically independent, is much more efficient over zero or large characteristic, even when the polynomials are succinctly represented as circuits. The key idea is a classical result, known as the Jacobian criterion [23, 6]. It says that if the characteristic of the field is zero, or large enough (compared to the product of degrees of the given polynomials), then the transcendence degree equals the linear rank of the Jacobian matrix of the polynomials. This leads to a simple randomized poly-time algorithm for checking algebraic independence, as we can get the circuits of the partial derivatives efficiently [5] and then use random evaluations to compute the rank of the Jacobian matrix. This final step of randomized evaluation is possible due to the Schwartz-Zippel-DeMillo-Lipton lemma [44, 9, 49].

One direction of the Jacobian criterion (if the polynomials are algebraically dependent, then their Jacobian matrix is not full rank) holds true for any characteristic. But the converse fails if the characteristic is small compared to the product of the degrees of input polynomials. For example, x^p is algebraically independent of $\mathbb{F}_p$, yet its derivative vanishes. We remark here that if two algebraically independent polynomials over characteristic p have zero Jacobian, then it does not mean that one of them is a p power. Consider, for example, $\{x^{p-1}y, xy^{p-1}\}$ over $\mathbb{F}_p$ for prime $p > 2$.

There are infinitely many input instances (set of polynomials), where the Jacobian criterion fails, i.e. Jacobian vanishes even though the given polynomials are independent. Those instances can be characterized by the notion of *inseparable extension*, that appears in Galois theory, and is formally defined in Sec.2.1. For example, the field extension $\mathbb{F}_p(x)/\mathbb{F}_p(x^p)$ has inseparable degree p as that many *conjugates* of $\sqrt[p]{x^p}$ in the splitting field are equal. This is a hard algebraic situation with no good geometric interpretation. Such behavior is absent over zero characteristic fields. So, positive characteristic requires inventing new concepts.

Naturally, we would like to come up with an efficient (randomized poly-time) algorithm over small characteristic. Though the failure of Jacobian criterion over small characteristic is known for long [15, 18], owing to the interest in algebraic independence from computer science perspective, several recent papers [10, 26, 6] posed the complexity status of this problem (whether it is in RP) as an open question. One curious aspect is that this problem is one of the rare ones in computer science where the gap between the known time complexity (EXP) and the conjectured one (RP) is that stark!

Talking about the two degrees. Let us consider a case where Jacobian criterion fails and certifying independence gets tricky. Let m_1m_2 be coprime to p, and $f_1 = x_1^{pm_1}, f_2 = x_2^{m_2}$. It is easy to deduce that the degree of the extension $\mathbb{F}_p(x_1, x_2)/\mathbb{F}_p(f_1, f_2)$ is pm_1m_2. In fact, the degree of the annihilating polynomial of $\{x_1, f_1, f_2\}$ (resp. $\{x_2, f_1, f_2\}$) is pm_1 (resp. m_2). However, the inseparable degree of the extension is only p, as the former annihilating polynomial (i.e. $y_1^{pm_1} - y_2$) is a polynomial in y_1^p but not in $y_1^{p^2}$. Thus, there are cases when the inseparable degree can be much smaller, even $O(1)$, compared to the extension degree. Notice that, in general, the inseparable degree is a p-power that divides the extension degree, which in turn is upper bounded by $\prod_i \deg(f_i)$ (by Perron's bound)– usually an exponen-

tially large parameter. The methods developed in this work only depend on the underlying inseparable degree, thus, our algorithm is expected to be much better than brute-force (in many cases).

A criterion that works for all characteristic for a natural problem like testing algebraic independence would be mathematically interesting. Computational implications of an efficient Jacobian like criterion would include a possible generalization (to small characteristic) of PIT or lower bound results [2], and algebraic extractors or entropy concepts [10].

Work done in case of finite fields. [37] gave a criterion that works over all fields, which they named *Witt-Jacobian* criterion. One key idea of the Witt-Jacobian criterion is to lift the input polynomials from characteristic $p \geq 2$ to a field of p-adics, which is zero characteristic. Witt-Jacobian polynomial can be seen as a scaled up p-adic lift of Jacobian polynomial and the criterion involves checking certain monomials (degeneracy testing; which looks hard) rather than zero testing. The main object underlying the proof is the de Rham-Witt pro-complex; a tool from modern algebraic-geometry (an excellent survey is [21]).

Witt-Jacobian criterion improved the complexity of independence testing problem, over positive characteristic, from PSPACE to $\mathrm{NP}^{\#\mathsf{P}}$. In the hierarchy of complexity classes, $\mathrm{NP}^{\#\mathsf{P}}$ is far above RP; thus there is a huge gap between what we have and what we want.

Partial derivative (defined as formal operators on polynomials), that played a key role in Jacobian criterion, behaves strangely over positive characteristic. Though it satisfies the usual rules of derivatives like linearity, product rule and chain rule, one important difference here is the fact that a non-constant polynomial can have a zero derivative. Another difference is that the higher derivatives of order $k \geq p$ are zero for all polynomials over characteristic p. Hasse-Schmidt derivatives are variants of usual derivatives, that were originally defined by [20], and independently by [47], to tackle this problem. In computer science literature, Hasse derivatives were used recently in coding theory (see [13] and the references therein), and PIT or lower bounds via generalized versions of shifted partial derivatives [17, 16].

Background on PIT and circuit lower bounds. The problems of derandomization of PIT and proving lower bounds, for explicit family of polynomials, are two fundamental questions in complexity theory. The question of PIT asks to test whether the polynomial computed by an arithmetic circuit is identically zero. This question can be studied in two settings. In the whitebox setting we are allowed to see inside the circuit, whereas in the blackbox setting we can only evaluate the circuit at some field points. The problem of blackbox PIT is equivalent to the problem of designing *hitting-sets* efficiently. Hitting-set is defined as follows. Let $\mathcal{C}$ be a class of polynomials in N variables over a field $\mathbb{F}$. Then, a set $\mathcal{H} \subseteq \mathbb{F}^N$ is called a *hitting-set* for the class $\mathcal{C}$, if for every nonzero polynomial $C \in \mathcal{C}$, there exists an $x \in \mathcal{H}$ such that $C(x) \neq 0$. PIT has a randomized poly-time algorithm, thanks to Schwartz-Zippel-DeMillo-Lipton lemma [44, 49, 9]. Derandomization of PIT is an outstanding open question in complexity theory with several implications, including proving arithmetic circuit lower bounds (refer to [4] & the survey [45]).

In the world of arithmetic complexity, we have strong structural results like *depth reductions* [19, 4]. These results show that strong enough lower bound, or PIT, results for homogeneous depth-4 (or general depth-3) circuits would give us exponential lower bounds and quasipoly-time derandomized PIT for general circuits (up to VP). Recent years have seen a fast growth in papers giving lower bound and PIT results for several special cases of small depth arithmetic circuits [42, 43]. Although there are strong (almost exponential, [33, 27]) lower bounds for homogeneous depth-4 circuits, the best known lower bounds for non-homogeneous depth-4 circuits are only superlinear (see [41] & the references therein).

Circuits with locally low algebraic rank. Kumar & Saraf [34] defined a *locally low*

algebraic rank circuit of degree n in N variables over $\mathbb{F}$, denoted $\Sigma\Gamma^{(k)}\Sigma\Pi^d$, as: $C = \sum_{i\in[T]} \Gamma_i(Q_{i1},\ldots,Q_{it})$, where Q_{ij} is a sparse polynomial (all monomials are given explicitly) of degree at most d, algRank of $\{Q_{ij} \mid j \in [t]\}$ is at most k, and Γ_i is an arbitrary t-variate polynomial, for $i \in [T]$.

The *size* of C comprises N, n, T and the maximum sparsity of Q_{ij}'s. Note that $k \leq N$, and we will be interested in the cases when kd is somewhat restricted.

Interestingly, $\Sigma\Gamma^{(n)}\Sigma\Pi$ subsumes homogeneous depth-4 circuits computing a degree n polynomial, as for homogeneous circuits $k \leq t \leq n$ and Γ_i is merely the product gate. Since this class includes non-homogeneous circuits as well (where t can be arbitrarily larger than k, n), it can be seen as a significant generalization of homogeneous depth-4.

This model subsumes certain other interesting models that were studied by [17, 2, 6] in the context of lower bounds and PIT. Invariably, their methods need to assume that $\mathbb{F}$ has characteristic zero or exponentially large (since partial derivatives are involved). Our goal in this paper is to overcome such restrictions.

1.1 Our contribution and relation with previous works

Broadly, in this paper, we prove two main technical theorems, one about the algebraically dependent polynomials and the other about algebraically independent polynomials. We apply these two theorems to obtain an algebraic independence testing algorithm, an arithmetic circuit lower bound over arbitrary field and a PIT algorithm (over fields of characteristic larger than the individual-degree of the polynomial). We now describe each of the results.

Algebraic dependence to approximate functional dependence. We show that over arbitrary fields, algebraic dependence of polynomials $f_1,\ldots,f_m$ imply the existence of a transcendence basis such that all the polynomials $f_1,\ldots,f_m$ can be obtained (upto a random shift and a truncation) as a polynomial function of the basis elements (Thm.10). Essentially, to obtain the desired polynomial, say f_k, we truncate a polynomial function in the elements of the basis upto the degree of f_k. This generalizes the functional dependence result of [34, Lem.3.1] which asserted the same over fields of zero (or large) characteristic.

We use a proof approach which is different from [34] to achieve the more general results. In the case of fields of zero characteristic, the subtle strength that this functional dependence property possesses is that *any* transcendence basis serves the purpose, which in general is false over positive characteristic. Our result explains this subtlety using the concept of *separating transcendence basis* from Galois theory (Sec.2.1). With this, a simple algebraic manipulation on the annihilating polynomial, and subspace of polynomial products (Lem.12), yields a functional dependence up to *any* desired degree of approximation. (This is a bit simpler than the approach of [34, Lem.2.4] where they approximate the roots of any multivariate polynomial using [14, Lem.3.1]. Such methods also appear in classical analysis under *Implicit Function Theorems*, see [31].)

Eg. $\{x_1, x_2, x_1x_2^2\}$ are algebraically dependent over $\overline{\mathbb{F}}_2$. Pick random field elements a_1, a_2. The shifted polynomials are $\{x_1+a_1, x_2+a_2, (x_1+a_1)(x_2^2+a_2^2)\}$. Clearly, (x_2+a_2) is not a function of the other two modulo the ideal $\langle \mathbf{x}\rangle^2$. However, (x_1+a_1) is trivially a function of the other two, namely, $(x_1+a_1) \equiv a_2^{-2}\cdot(x_1+a_1)(x_2^2+a_2^2) \mod \langle \mathbf{x}\rangle^2$.

Algebraically independent polynomials - Criterion. The above example shows that over fields of positive characteristic, an approximate functional dependence may exist even in the case of algebraically independent polynomials. We overcome this issue and show that the independence can be captured by truncating the polynomial function in the basis elements upto a precise parameter, i.e. if we choose the truncation point to be greater than that parameter, then algebraically independent polynomials *cannot* exhibit functional

dependence (Thm.13). This parameter is actually the *inseparable degree* of an appropriate field extension, which is a well studied concept in Galois theory (Sec.2.1).

Continuing the above example– $\{x_1, x_1x_2^2\}$ are algebraically independent over $\overline{\mathbb{F}}_2$. Pick random field elements a_1, a_2. The shifted polynomials are $\{x_1 + a_1, (x_1 + a_1)(x_2^2 + a_2^2)\}$. It can be verified that neither is a polynomial function of the other modulo the ideal $\langle \mathbf{x} \rangle^3$. This becomes a certificate of algebraic independence. (Note that the inseparable degree of $\mathbb{F}_2(x_1, x_2)/\mathbb{F}_2(x_1, x_1x_2^2)$ is 2.)

When the inseparable degree is 1 (which means a *separable extension*), then looking at the truncation upto the linear term of shifted basis elements would suffice. So, our result implies that *separable extension* is precisely the case when the Jacobian works (an exposition can be found in the full version). For higher inseparable degree t, our result can be reinterpreted as giving a Jacobian like result: algebraically independent polynomials have $\mathbb{F}(\mathbf{z})$-linearly independent higher differentials (Sec.2.2), modulo a carefully chosen subspace $\mathcal{U}_t$ (Rmk.11). This follows by considering the Taylor series, around a "generic" point $\mathbf{z}$, whence, the functional independence of polynomials shifted by $\mathbf{z}$, implies the linear independence of shifted polynomials modulo $\mathcal{U}_t$. As shifted polynomials contain all the Hasse-Schmidt higher derivatives (wrt $\mathbf{x}$ and evaluated at the point $\mathbf{z}$), we deduce their $\mathbb{F}(\mathbf{z})$-linear independence modulo $\mathcal{U}_t$.

Again, a key technical lemma used in finishing the proof is Lem.12 (*subspace reduction*), which concerns the ideal theoretic properties of the subspace $\mathcal{U}_t$. Basically, it helps us prove that if $\{h_1, \ldots, h_n\}$ are polynomials with their degree($\leq t$)-part having algebraically independent leading monomials, and g_n functionally depends on $\{g_1, \ldots, g_{n-1}\}$ (with truncation beyond t), then some h_i is functionally independent of $\{g_1, \ldots, g_n\}$.

Application 1: Testing algebraic independence. An easy consequence of Thm.10 and Thm.13 is that we get a randomized poly-time algorithm for testing algebraic independence of polynomials over finite fields (say, $\mathbb{F}_q$ of characteristic p) in the cases when the inseparable degree is constant. Since the latter is a p-power (Sec.2.1), our algorithm is interesting when p is a constant. (Whenever required, we can assume wlog that the input is n circuits in n variables over an algebraically closed field; see full version for simple proofs.)

▶ **Theorem 1** (Independence testing). *For circuits $\mathbf{f} \in \mathbb{F}_q[\mathbf{x}]$, we have a randomized poly(s, $\binom{t+n}{n}$)-time algebraic independence testing algorithm, where the inseparable degree of the field extension $\mathbb{F}_q(\mathbf{x})/\mathbb{F}_q(\mathbf{f})$ is t (assuming $\mathbf{f}$ algebraically independent) & input size is s.*

This covers a lot of interesting cases as the inseparable degree can be quite small even in case of polynomials with exponential degree. As a simple example, take two bivariate circuits of exponential degree over $\mathbb{F}_2$. Suppose they are independent and their Jacobian is nonzero. Now if we square any one of these two, then Jacobian would fail as the inseparable degree becomes 2. Previously known algorithms cannot deal with even such a simple case, whereas we easily handle the case by trying our test with $t = 2$. In general, the inseparable degree is upper bounded by Perron's degree bound (product of degrees of given polynomials, [40]), so in the worst-case our algorithm is exponential-time. (Witt-Jacobian criterion [37] is exponential-time in all cases.) We illustrate the overall idea, and its comparison with Jacobian criterion, in the figure in the conclusion (Sec.4).

An interesting by-product of the algorithm is that it computes the inseparable degree, of the given independent polynomials, in the same time.

Application 2: Lower bound for locally low algebraic rank circuits. Using the functional dependence result, we give an explicit family of polynomials in VNP of degree n in N variables, where $N = n^{O(1)}$ such that any $\Sigma\Gamma^{(n)}\Sigma\Pi$ circuit computing it has size

$N^{\Omega(\sqrt{n})}$. We obtain this lower bound over *arbitrary* fields. This generalizes the lower bound of [34, Thm.1.4] which itself was a strong generalization of the shifted partials based homogeneous depth-4 lower bounds [27] and Jacobian based lower bounds [2] (all over zero or large characteristic). Since our functional dependence generalizes the key technical ingredient of [34] to arbitrary fields, we are able to get the same lower bound (for the same model and hard polynomial family) over arbitrary fields. Formally,

▶ **Theorem 2.** *Let $\mathbb{F}$ be any field. There exists a family $\{P_n\}$ of polynomials in VNP, such that P_n is a polynomial of degree n in $N = n^{O(1)}$ variables with $0,1$ coefficients, and for any $\Sigma\Gamma^{(k)}\Sigma\Pi$ circuit C, if $k \leq n$ and if C computes P_n over $\mathbb{F}$, then Size$(C) \geq N^{\Omega(\sqrt{n})}$.*

Remark: As remarked by [34], the above model is challenging even for $k = 2$ (& was open before us for small characteristic fields). Also, the proof goes through for any $k = n^{O(1)}$, as long as one picks N as an appropriately large polynomial in n.

The proof of this theorem closely follows [34], and is sketched in the full version.

Application 3: Hitting-set for $\Sigma\Gamma^{(k)}\Sigma\Pi^d$ circuits. We show that for any size-s circuit $C \in \Sigma\Gamma^{(k)}\Sigma\Pi^d$, where $k, d = \text{polylog}(s)$, over fields of characteristic $p >$ individual-degree(C), there exists a quasipoly(s)-time hitting-set.

▶ **Theorem 3.** *Let $\mathbb{F}$ be any field of characteristic p. There exists an $\exp(\log^{O(1)} s)$-time constructible hitting-set $\mathcal{H} \subseteq \overline{\mathbb{F}}^N$ for size-s circuit $C \in \Sigma\Gamma^{(k)}\Sigma\Pi^d$ with $kd = \log^{O(1)} s$, assuming $p >$ individual-degree(C) or $p = 0$.*

Again, the proof follows [34]. For PIT, algebraic rank based models have already been considered by [6, 2, 34]. Our result generalizes some of these results to smaller positive characteristic (only requiring $p >$ individual-degree(C)). The previous results required $p > d^k$, which is super-polynomial in the above regime. Our inability to remove this restriction lies in the nature of shifted partials [17, Lem.4.13]. Eg. the dimension of shifted partials of a p-power monomial $x_1^{p^{e_1}} \cdots x_n^{p^{e_n}}$ is not that large over $\mathbb{F}_p$.

2 Preliminaries: Jacobi, Galois and Hasse-Schmidt

We define the central object related to the testing of algebraic independence is the Jacobian.

▶ **Definition 4** (Jacobian)**.** The *Jacobian* of polynomials $\mathbf{f} = \{f_1, \cdots, f_m\}$ in $\mathbb{F}[x_1, \cdots, x_n]$ is the matrix $\mathcal{J}_{\mathbf{x}}(\mathbf{f}) = (\partial_{x_j} f_i)_{m \times n}$, where $\partial_{x_j} f_i := \partial f_i / \partial x_j$.

We state the classical Jacobian criterion [23, 6].

▶ **Lemma 5** (Jacobian criterion)**.** *Let $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$ be a finite set of polynomials of degree at most d, and $\text{trdeg}_{\mathbb{F}} \mathbf{f} \leq r$. If $\text{char}(\mathbb{F}) = 0$, or $\text{char}(\mathbb{F}) > d^r$, then $\text{trdeg}_{\mathbb{F}} \mathbf{f} = \text{rank}_{\mathbb{F}(\mathbf{x})} \mathcal{J}_{\mathbf{x}}(\mathbf{f})$.*

Previously, we saw some examples of polynomials over fields of smaller characteristic where the Jacobian *fails*. Here is another nontrivial example: $\mathbf{f} = \{x_1^2 x_2 + x_1^3,\ x_1 x_2^2 + x_1 x_2^5\}$ in $\mathbb{F}_3[x_1, x_2]$ is a set of algebraically independent polynomials, but $\text{rank}_{\mathbb{F}_3(\mathbf{x})} \mathcal{J}_{\mathbf{x}}(\mathbf{f}) = 1$, and hence the criterion fails.

2.1 Inseparability & separating transcendence basis

For this section, let $\mathbb{E} \supseteq \mathbb{F}$ be fields. Failure of the Jacobian criterion can be explained using the fundamental concept of inseparability from Galois theory [22].

▶ **Definition 6.** An $f \in \mathbb{F}[x]$ is *separable* if it has no multiple roots in its splitting field.

It is easy to prove that– For an irreducible f, separability is implied by the non-zeroness of $\partial_x f$. Thus, if $\mathrm{char}(\mathbb{F}) = 0$, then any irreducible polynomial is separable. It further implies that if $\mathrm{char}(\mathbb{F}) = p > 0$ then, an irreducible f is separable if and only if $f \notin \mathbb{F}[x^p]$. We have this notion of separability in case of field extensions as well. An algebraic extension $\mathbb{E}/\mathbb{F}$ is said to be *separable* if every element $\alpha \in \mathbb{E}$ has a minimal polynomial over $\mathbb{F}$ that is separable.

For polynomials $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$, we deal with the extension $\mathbb{F}(x_1, \ldots, x_n)/\mathbb{F}(f_1, \ldots, f_m)$. This extension is algebraic iff $\mathrm{trdeg}(\mathbf{f}) = n$ (by Lem.19, every x_j depends on $\mathbf{f}$). In which case, the extension $\mathbb{F}(\mathbf{x})/\mathbb{F}(\mathbf{f})$ is separable iff the minimal polynomial of x_j over $\mathbb{F}(\mathbf{f})$ is separable, for all $j \in [n]$. The latter, clearly, is the case when $\mathrm{char}(\mathbb{F}) = 0$. When $\mathrm{char}(\mathbb{F}) = p > 0$, the extension is inseparable if there exists $j \in [n]$, such that the minimal polynomial of x_j over $\mathbb{F}(\mathbf{f})$ lives in $\mathbb{F}(\mathbf{f})[y^p]$. Thus for every x_j, we have an m_j such that $x_j^{p^{m_j}}$ has a separable minimal polynomial over $\mathbb{F}(\mathbf{f})$.

The *inseparable degree of the extension* $\mathbb{F}(\mathbf{x})/\mathbb{F}(\mathbf{f})$ is defined as the minimum p^m such that the minimal polynomial of $x_j^{p^m}$ over $\mathbb{F}(\mathbf{f})$ is separable, for all $j \in [n]$. We also associate this inseparable degree with the set $\mathbf{f}$.

In the case when $\mathbf{f}$ are algebraically dependent, we would like to use a "good" transcendence basis. This is captured by:

▶ **Definition 7** (Separating transcendence basis)**.** A field extension $\mathbb{E}/\mathbb{F}$ is called *separably generated* if there exists an algebraically independent set (i.e. transcendence basis) $S = \{f_1, \ldots, f_r\} \subset \mathbb{E}$ such that $\mathbb{E}/\mathbb{F}(S)$ is algebraic and separable.

S is called a *separating transcendence basis* of $\mathbb{E}/\mathbb{F}$.

It is a classical result that such bases exist for fields that we are interested in.

▶ **Theorem 8.** *Consider a finite set of polynomials $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$. If $\mathbb{F}$ is a finite field (resp. an algebraically closed field) then there exists a separating transcendence basis, of $\mathbb{F}(\mathbf{f})/\mathbb{F}$, in $\mathbf{f}$.*

In case $\mathbb{F}$ is a zero characteristic field then every *transcendence basis of $\mathbf{f}$ is a separating one of the extension $\mathbb{F}(\mathbf{f})/\mathbb{F}$.*

Proof. It is clear that if $\mathbb{F}$ has characteristic zero then there is no possibility of inseparability.

Let $\mathbb{F}$ be a finite (resp. algebraically closed) field. [30, Thm.7.20] shows that the extension $\mathbb{F}(\mathbf{f})/\mathbb{F}$ is separably generated. Furthermore, [30, Thm.7.18] shows that $\mathbf{f}$ contains a subset that is a separating transcendence basis of the extension. ◀

Examples. Extension $\mathbb{F}_3(x^3)/\mathbb{F}_3$ has $\{x^3\}$ as a separating transcendence basis. Consider the two transcendence bases of the extension $\mathbb{F}_3(x^2, x^3)/\mathbb{F}_3$ – $\{x^3\}$ and $\{x^2\}$. The latter is a separating transcendence basis, but the former is not.

2.2 Taylor expansion at $\mathbf{z}$, higher derivatives & differentials

We consider the application of shift (or translation) to our polynomials. We view this as writing the Taylor expansion of a polynomial $f(\mathbf{x})$ at a "generic" point $\mathbf{z}$ [16, Sec.C.1]. A second view is that of computing the Hasse-Schmidt higher derivatives of f at the point $\mathbf{z}$ [17, 13]. A third view is seeing the shifted polynomial as a Hasse-Schmidt differential [48]. We collect these equivalent viewpoints in a single definition.

▶ **Definition 9** (Formal shift)**.** We see $f(\mathbf{x} + \mathbf{z})$ as a polynomial in $R := \mathbb{F}_p(\mathbf{z})[\mathbf{x}]$ where the variables $x_1, \ldots, x_n$ are *shifted* respectively by the function field elements $z_1, \ldots, z_n$.

Now the coefficient of $m := x_1^{\ell_1} \cdots x_n^{\ell_n}$ in the Taylor-series expansion of $f(\mathbf{x} + \mathbf{z})$ can be written as $\frac{1}{\ell_1! \cdots \ell_n!} \frac{\partial^{(\ell_1 + \cdots + \ell_n)} f}{\partial x_1^{\ell_1} \ldots \partial x_n^{\ell_n}}(\mathbf{z})$.

This is called the *Hasse-Schmidt derivative of f wrt m evaluated at the point* $\mathbf{z}$. It can be denoted, by some abuse of notation, as $\partial_m f(\mathbf{x})|_{\mathbf{z}}$.

Finally, we can see the formal shift as a *Hasse-Schmidt differential*, namely, $f(\mathbf{x}+\mathbf{z}) = \sum_m m \cdot \partial_m f(\mathbf{x})|_{\mathbf{z}}$ (sum over all monomials m in $\mathbf{x}$).

Example. We have $\partial^2 x^2/\partial x^2 = 0$ over $\mathbb{F}_2$, but $\partial^2 x^2/2!\partial x^2 = 1$. Thus, Hasse-Schmidt derivatives offer a natural solution to this vanishing problem.

This connection between the shifts and Hasse-Schmidt higher derivatives/ differentials is what motivated us to search for the right framework to study algebraic independence.

Now the Jacobian criterion is given in terms of the first order derivatives of the polynomials and the failure of Jacobian essentially exposes the inability of first order derivative in capturing independence. Intuitively, it seems that going to higher derivatives may help. The above connection points out that perhaps we need to look at higher degree terms (wrt $\mathbf{x}$) of $f(\mathbf{x}+\mathbf{z})$ to get an algebraic independence criterion in cases where Jacobian fails. Eventually, we will see that the intuition is indeed true.

Operator $\mathcal{H}$. For notational convenience, we define the non-constant part of $f(\mathbf{x}+\mathbf{z})$ up to degree$\leq t$ wrt $\mathbf{x}$, as $\mathcal{H}_t f := f^{\leq t}(\mathbf{x}+\mathbf{z}) - f(\mathbf{z})$.

This is easier to work with when we do manipulations modulo the ideal $\langle \mathbf{x} \rangle_R^{t+1}$.

3 Main structure theorems

We use the following standard notation in the paper:

1. $\mathbb{F}$ is an arbitrary field. $\overline{\mathbb{F}}$ is its algebraic closure.
2. $\mathbb{F}_q$ is a finite field of size q and characteristic $p \geq 2$.
3. Let $R \supseteq S$ be a commutative ring extension over a field $\mathbb{F}$, let $v_1, \ldots, v_m \in R$ and $r \geq 1$. Then $\langle v_1, \ldots, v_m \rangle_S^r$ is simply the set of all S-linear combinations of products $v_{i_1} \cdots v_{i_r}$ (i_j's in $[m]$). It is both an S-module and an $\mathbb{F}$-vector space. (It is an ideal when $R = S$.)
4. For a polynomial $h \in \mathbb{F}[\mathbf{x}]$, $h^{\leq d}$ extracts out the degree$\leq d$ part of h and returns it as an element in $\mathbb{F}[\mathbf{x}]$ again.
5. For a polynomial $h \in \mathbb{F}[\mathbf{x}]$, $h^{[\leq d]}$ extracts out the degree$\leq d$ part of h and returns it as a $d+1$ tuple, where for $i \in [0 \ldots d]$, i-th entry of the tuple contains $h^{=i}$ which is defined as the homogeneous component of h of degree i.

3.1 Functional dependence for algebraically dependent polynomials

A fact about linear independence is that if $f_1, \ldots, f_m \in \mathbb{F}[\mathbf{x}]$ are linearly dependent, it also implies that every polynomial can be written as a linear combination of the polynomials in the basis. The question is whether the same can be extended to algebraic dependence: Does algebraic dependence imply that all the polynomials can be written as a function of the polynomials in the transcendence basis? It was shown in [34, Lem.3.1] that it is indeed true (approximately) over fields of zero (and large) characteristic.

We generalize the property using a different proof approach and show that algebraic dependence implies functional dependence over arbitrary fields (to arbitrary degree of approximation t).

▶ **Theorem 10** (Functional dependence over arbitrary fields). *Let $\mathbf{f} = \{f_1, \ldots, f_m\} \subset \mathbb{F}[x_1, \ldots, x_n]$ be a set of polynomials, where $\mathbb{F}$ is any field, and $t \in \mathbb{N}$. If trdeg of $\{f_1, \ldots, f_m\}$ is k, then there exist algebraically independent $\{g_1, \ldots, g_k\} \subset \mathbf{f}$, such that for random $\mathbf{a} \in \overline{\mathbb{F}}^n$, there are polynomials $h_i \in \overline{\mathbb{F}}[Y_1, \ldots, Y_k]$ satisfying, $\forall i \in [m]$, $f_i^{\leq t}(\mathbf{x}+\mathbf{a}) = h_i^{\leq t}(g_1(\mathbf{x}+\mathbf{a}), \ldots, g_k(\mathbf{x}+\mathbf{a}))$.*

▶ Remark. Clearly, $\overline{\mathbb{F}}^n$ is an infinite space. What we mean here by a *random* $\mathbf{a}$ is "random point in any sufficiently large, but finite, subset of the space". It will be clear from the proof that it would suffice to sample from any set of size at most exponential in the input size. We skip the detailed estimate as in this paper merely existence of $\mathbf{a}$ is needed.

We will use $\mathbf{z}$ as a formal variable (n-tuple) and can fix it later to a suitable constant $\mathbf{a}$. To prove the theorem, we consider the ring $R := \overline{\mathbb{F}}(\mathbf{z})[\mathbf{x}]$ and its ideal $\mathcal{I}_0 := \langle \mathbf{x} \rangle_R$. The ideal collects the non-constant linear polynomials. Now, define the ideal $\mathcal{I}_t := \mathcal{I}_0^{t+1}$ and the quotient algebra $\mathcal{Q}_t := R/\mathcal{I}_t$, i.e. we are filtering out, or *truncating*, all the terms of degree $> t$. Now $\mathcal{Q}_t$ can also be seen as a finite $\binom{n+t}{n}$ dimensional vector space over $\overline{\mathbb{F}}(\mathbf{z})$ whose basis is monomials in $\mathbf{x}$ of degree at most t. In our theorems and proofs, most of the operations happen in this quotient ring $\mathcal{Q}_t$ for increasing t's.

In our analysis, we plan to use the shifting of the variables in the evaluated annihilating polynomial of $\{f_i, g_1, \ldots, g_k\}$, and it is clear that on applying the shifts, we will end up having terms of the form $(\mathcal{H}_t f_i)^{j_0}(\mathcal{H}_t g_1)^{j_1} \cdots (\mathcal{H}_t g_k)^{j_k}$ (recall that in Q_t, $f(\mathbf{x}+\mathbf{z}) = f(\mathbf{z}) + \mathcal{H}_t f(\mathbf{x})$). Now, note that due to the filtration in $\mathcal{Q}_t$, some of these terms will be equivalent to terms involving $\mathcal{H}_r$ with $r < t$. We consider an appropriate subspace $\mathcal{U}_t \subset \mathcal{Q}_t$ generated by such "higher" products, which we formally define as: $\mathcal{U}_1 := \{0\}$ and

$$\mathcal{U}_t := \langle \mathcal{H}_{t-1} f_i, \mathcal{H}_{t-1} g_1, \ldots, \mathcal{H}_{t-1} g_k \rangle^2_{\overline{\mathbb{F}}(\mathbf{z})} + \cdots + \langle \mathcal{H}_1 f_i, \mathcal{H}_1 g_1, \ldots, \mathcal{H}_1 g_k \rangle^t_{\overline{\mathbb{F}}(\mathbf{z})}\,, \quad t \geq 2.$$

▶ Remark 11. In $\mathcal{Q}_t$, observe that, this is the same subspace as $\langle \mathcal{H}_t f_i, \mathcal{H}_t g_1, \ldots, \mathcal{H}_t g_k \rangle^2_{\overline{\mathbb{F}}(\mathbf{z})} + \cdots + \langle \mathcal{H}_t f_i, \mathcal{H}_t g_1, \ldots, \mathcal{H}_t g_k \rangle^t_{\overline{\mathbb{F}}(\mathbf{z})}$

Pf. of Thm.10. Consider the set $\mathbf{f} := \{f_1, \ldots, f_m\} \subset \mathbb{F}[\mathbf{x}]$ with algebraic rank k. If we work over $\overline{\mathbb{F}}$, then Thm.8 guarantees the existence of a separating transcendence basis $\{g_1, \ldots, g_k\} \subseteq \mathbf{f}$. Let $g_0 := f_i$ for a fixed $i \in [m]$. Now we consider the separable annihilating polynomial $A(\mathbf{y}) = \sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \mathbf{y}^{\mathbf{e}_\ell}$ of the set $\mathbf{g} := \{g_0, g_1, \ldots, g_k\}$, and $a_{\mathbf{e}_\ell}$'s are in $\overline{\mathbb{F}}$ ($\mathbf{e}_\ell$ is a $(k+1)$-tuple $(e_{j\ell} \mid j \in [0 \ldots k])$). Thus, $A(\mathbf{g}) = \sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \prod_{j=0}^k g_j(\mathbf{x})^{e_{j\ell}} = 0$. We now apply the formal shift $\mathbf{x} \mapsto \mathbf{x} + \mathbf{z}$ to get $A(g_0(\mathbf{x}+\mathbf{z}), \ldots, g_k(\mathbf{x}+\mathbf{z})) = 0$, i.e. $\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \prod_j g_j(\mathbf{x}+\mathbf{z})^{e_{j\ell}} = 0$.

We now study this relation in the algebra $\mathcal{Q}_t$. By Taylor series expansion, we know that $f(\mathbf{x}+\mathbf{z}) \equiv f(\mathbf{z}) + \mathcal{H}_t f(\mathbf{x})$ in $\mathcal{Q}_t$, so we get $\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \prod_j (g_j(\mathbf{z}) + \mathcal{H}_t g_j)^{e_{j\ell}} \equiv 0$. The binomial expansion gives a compact expression:

$$\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \sum_{\mathbf{0} \leq \mathbf{s} \leq \mathbf{e}_\ell} \binom{\mathbf{e}_\ell}{\mathbf{s}} \cdot (\mathcal{H}_t \mathbf{g})^{\mathbf{s}} \cdot \mathbf{g}^{\mathbf{e}_\ell - \mathbf{s}} \equiv 0\,.$$

Note that the contribution by $\mathbf{s} = \mathbf{0}$ terms sum up to $\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \prod_{j=0}^k g_j(\mathbf{z})^{e_{j\ell}}$ which is zero. This implies that an $\overline{\mathbb{F}}(\mathbf{z})$-linear combination of the products of the form $(\mathcal{H}_t g_0)^{s_0} \cdots (\mathcal{H}_t g_k)^{s_k}$, $\sum_j s_j \geq 1$, vanishes in $\mathcal{Q}_t$. Now the key step is to separate out the terms *linear* in $\mathcal{H}_t g_j$ and switch the sums, to obtain

$$\mathcal{H}_t g_0 \cdot g_0(\mathbf{z})^{-1} \left(\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \cdot e_{0\ell} g_0^{e_{0\ell}} \cdots g_k^{e_{k\ell}} \right) + \sum_{j \in [k]} \mathcal{H}_t g_j \cdot g_j(\mathbf{z})^{-1} \left(\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \cdot e_{j\ell} g_0^{e_{0\ell}} \cdots g_k^{e_{k\ell}} \right)$$
$$+ (\text{higher terms with } \sum_j s_j \geq 2) \equiv 0\,. \qquad (1)$$

Further, we argue using the minimality and separability of A (in terms of the first variable) that the "linear" term $\mathcal{H}_t g_0$ in the vanishing sum above has a non-zero coefficient: as it would either mean a lower degree annihilating polynomial $A := \sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} e_{0\ell} y_0^{e_{0\ell}-1} \cdot y_1^{e_{1\ell}} \cdots y_k^{e_{k\ell}}$

i.e. contradicting the minimality, or that all the $e_{0\ell}$'s are divisible by p (when $\mathbb{F}$ has characteristic p) which means that f_i does not depend separably on $\{g_1, \ldots, g_k\}$; which contradicts the fact that $\{g_1, \ldots, g_k\}$ is a separating transcendence basis.

Thus, we get that $\mathcal{H}_t g_0$ lives in the $\overline{\mathbb{F}}(\mathbf{z})$-linear span of $\mathcal{H}_t g_1, \ldots, \mathcal{H}_t g_k$ modulo the subspace generated by the higher terms of the summation in Eqn.1. So, $\mathcal{H}_t g_0$ lives in the $\overline{\mathbb{F}}(\mathbf{z})$-linear span of $\mathcal{H}_t g_1 \ldots, \mathcal{H}_t g_k$ modulo the subspace $\mathcal{U}_t$ (Rmk.11) in $\mathcal{Q}_t$.

We got $\mathcal{H}_t f_i \in \langle \mathcal{H}_t g_1, \ldots, \mathcal{H}_t g_k \rangle_{\overline{\mathbb{F}}(\mathbf{z})} + \mathcal{U}_t$. Now, we are in a position to apply Lem.12, which essentially says that if $\mathcal{H}_r f_n$ depends on higher order terms (in the sense of Eqn.1) then it can be "dropped" from the ideal manipulations. Thus, we get that $\mathcal{H}_t f_i \in \langle \mathcal{H}_t g_1, \ldots, \mathcal{H}_t g_k \rangle_{\overline{\mathbb{F}}(\mathbf{z})}$ $+\langle \mathcal{H}_{t-1} g_1, \ldots, \mathcal{H}_{t-1} g_k \rangle^2_{\overline{\mathbb{F}}(\mathbf{z})} + \cdots + \langle \mathcal{H}_1 g_1, \ldots, \mathcal{H}_1 g_k \rangle^t_{\overline{\mathbb{F}}(\mathbf{z})}$. The latter (by Rmk.11) is exactly $\langle \mathcal{H}_t g_1, \ldots, \mathcal{H}_t g_k \rangle_{\overline{\mathbb{F}}(\mathbf{z})} + \langle \mathcal{H}_t g_1, \ldots, \mathcal{H}_t g_k \rangle^2_{\overline{\mathbb{F}}(\mathbf{z})} + \cdots + \langle \mathcal{H}_t g_1, \ldots, \mathcal{H}_t g_k \rangle^t_{\overline{\mathbb{F}}(\mathbf{z})}$.

This implies $f_i(\mathbf{x}+\mathbf{z}) \in \langle 1, g_1(\mathbf{x}+\mathbf{z}), \ldots, g_k(\mathbf{x}+\mathbf{z}) \rangle^t_{\overline{\mathbb{F}}(\mathbf{z})}$ in $\mathcal{Q}_t$, which yields the approximate functional dependence around a generic point $\mathbf{z}$.

Fixing $\mathbf{z}$ (avoiding some bad choices that make certain $\mathbf{z}$-polynomials in the above proof zero) to an element $\mathbf{a} \in \overline{\mathbb{F}}^n$ finishes the proof. ◀

We now formally state our subspace reduction lemma:

▶ **Lemma 12** (Subspace reduction). *Let $\mathbb{F}$ be any field, $R := \mathbb{F}(\mathbf{z})[\mathbf{x}]$, $\mathcal{Q}_r := R/\langle \mathbf{x} \rangle^{r+1}$ for $r \geq 1$, and $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$. Define $\mathcal{U}_1 = \mathcal{V}_1 = \{0\}$, and for $u \in \langle \mathbf{x} \rangle_R$, $r \geq 2$, define the subspaces (in the quotient algebra $\mathcal{Q}_r$),*

$$\mathcal{U}_r := \langle \mathcal{H}_{r-1} f_1, \ldots, \mathcal{H}_{r-1} f_n \rangle^2_{\mathbb{F}(\mathbf{z})} + \cdots + \langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_n \rangle^r_{\mathbb{F}(\mathbf{z})} ,$$
$$\mathcal{V}_r := \langle \mathcal{H}_{r-1} f_1, \ldots, \mathcal{H}_{r-1} f_{n-1}, u \rangle^2_{\mathbb{F}(\mathbf{z})} + \cdots + \langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_{n-1}, u \rangle^r_{\mathbb{F}(\mathbf{z})} .$$

If $\mathcal{H}_t f_n \in \langle \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1}, u \rangle_{\mathbb{F}(\mathbf{z})} + \mathcal{U}_t$, then $\mathcal{U}_t \subseteq \mathcal{V}_t$ (for any $t \in \mathbb{N}$).

Remark: If $u = 0$ then the lemma "reduces" the n polynomial generators, of the subspace $\mathcal{U}_t$, by one. Hence, the name "subspace reduction". A simple inductive proof of the lemma is given in the full version.

3.2 Algebraically independent polynomials: Criterion

Having proved the functional dependence for algebraically dependent polynomials, one naturally asks whether a converse exists (for arbitrary fields? to what degree?). We will characterize this completely.

It's all about the inseparable degree- We show that if f is algebraically independent of $\{g_1, \ldots, g_k\}$ then, under a random shift, f cannot be written as a function of $\{g_1, \ldots, g_k\}$ when chosen to truncate at (or beyond) the inseparable degree of the extension $\mathbb{F}_q(\mathbf{x})/\mathbb{F}_q(f, g_1, \ldots, g_k)$. Moreover, for each truncation at lower degrees we get functional dependence.

▶ **Theorem 13** (Algebraic to functional independence). *Let $\mathbf{f} \subset \mathbb{F}_q[\mathbf{x}]$ be algebraically independent polynomials (wlog n-variate n polynomials) with inseparable degree p^i. Then,*

1. *for all $t \geq p^i$, for random $\mathbf{a} \in \overline{\mathbb{F}}_q^n$, $f_n^{\leq t}(\mathbf{x}+\mathbf{a})$ cannot be written as $h^{\leq t}(f_1(\mathbf{x}+\mathbf{a}), \ldots, f_{n-1}(\mathbf{x}+\mathbf{a}))$, for any $h \in \overline{\mathbb{F}}_q[Y_1, \ldots, Y_{n-1}]$.*
2. *for all $1 \leq t < p^i$, $\exists j \in [n]$, for random $\mathbf{a} \in \overline{\mathbb{F}}_q^n$, $f_j^{\leq t}(\mathbf{x}+\mathbf{a})$ can be written as $h_{jt}^{\leq t}(f_1(\mathbf{x}+\mathbf{a}), \ldots, f_{j-1}(\mathbf{x}+\mathbf{a}), f_{j+1}(\mathbf{x}+\mathbf{a}), \ldots, f_n(\mathbf{x}+\mathbf{a}))$, for some $h_{jt} \in \overline{\mathbb{F}}_q[\mathbf{Y}]$.*

Remark. Our proof works for any field $\mathbb{F}$ (manipulate in $\overline{\mathbb{F}}$). In case of characteristic $p \geq 2$ we get the above statement and in characteristic zero use inseparable degree $= 1$.

Proof idea: By the hypothesis we have that each variable $x_j^{p^i}$, $j \in [n]$, algebraically depends on $\mathbf{f}$ with a *separable* annihilating polynomial over $\mathbf{F}_q$. Consider ring $R := \overline{\mathbb{F}}_q(\mathbf{z})[\mathbf{x}]$.

The basic idea is to consider the minimal annihilating polynomial A_j of $\{x_j^{p^i}, \mathbf{f}\}$ and formally shift the relevant polynomials by $\mathbf{z}$. From the proof of Thm.10 we get a functional dependence of $x_j^{p^i}$ on $\mathbf{f}(\mathbf{x}+\mathbf{z})$ up to any degree t.

Interestingly, when we take $t < p^i$ the monomial $x_j^{p^i}$ vanishes mod $\langle \mathbf{x} \rangle^{t+1}$. This means that the above yields, in fact, a functional dependence among $\mathbf{f}(\mathbf{x}+\mathbf{z})$.

On the other hand, for $t \geq p^i$, we get a nontrivial functional dependence of $x_j^{p^i}$ on $\mathbf{f}(\mathbf{x}+\mathbf{z})$, for all $j \in [n]$. In this case, one can give an argument using monomial ordering that there exists no functional dependence among $\mathbf{f}(\mathbf{x}+\mathbf{z})$.

We can see that the classical Jacobian criterion as a special case of Theorems 10 and 13. The detailed discussions and missing proofs are given in the full version.

4 Conclusion

We give a criterion for testing algebraic independence over positive characteristic, in the spirit of Jacobian criterion, that works for any field. Its complexity is parameterized by the inseparable degree bound. It is also strong enough to give the inseparable degree at the same time. We give applications to locally low algebraic rank circuits in the cases that were open before.

	Jacobian Criterion	**Our Criterion**
The approach:	reduces algebraic independence to linear independence testing	reduces algebraic independence to linear independence testing
Related "approximate" shift :	$\mathbf{f}(\mathbf{x}) \mapsto \mathbf{f}(\mathbf{x}+\mathbf{z}) \mod \langle \mathbf{x} \rangle^2_{\mathbb{F}(\mathbf{z})[\mathbf{x}]}$	$\mathbf{f}(\mathbf{x}) \mapsto \mathbf{f}(\mathbf{x}+\mathbf{z}) \mod \mathcal{U}_t$
Vectors for $\mathbb{F}(\mathbf{z})$-dependence:	$\mathcal{H}_1\mathbf{f} \mod \mathcal{U}_1$	$\mathcal{H}_t\mathbf{f} \mod \mathcal{U}_t$
Certifies alg.independence if:	$\mathbb{F}(\mathbf{x})/\mathbb{F}(\mathbf{f})$ is separable	separable or inseparable $\mathbb{F}(\mathbf{x})/\mathbb{F}(\mathbf{f})$
Efficiency in $\mathrm{char}(\mathbb{F}) = 0$:	randomized poly-time algorithm	$t = 1$, (same as Jacobian criterion)
Efficiency in $\mathrm{char}(\mathbb{F}) = p$, inseparable degree $\leq p^e$:	fails	randomized poly$\binom{n+p^e}{n}$-time algorithm

The main open problem is to investigate whether we can improve the criterion to get a randomized poly-time algorithm for circuits over a finite field. We mention a few special cases based on different restrictions on input. None of these cases are (efficiently) solved by presently known techniques.

- the polynomials are *supersparse*, i.e. sparse polynomials with possibly exponential degree.
- two bivariate circuits, with an exponentially large inseparable degree, over $\mathbb{F}_2$.
- n quadratic polynomials over $\mathbb{F}_2$.

Our hitting-set result, for locally low algebraic rank circuits, still has a mild assumption on the characteristic. Can this be eliminated?

5 Acknowledgements

We thank Manindra Agrawal, Rohit Gurjar & Arpita Korwar for the insightful discussions and encouragement. We thank Markus Bläser and the anonymous reviewers for the elaborate suggestions to improve the draft. NS acknowledges the support from DST/SJF/MSA-01/2013-14 and SB/FTP/ETA-177/2013.

References

1 L. M. Adleman and H. W. Lenstra. Finding irreducible polynomials over finite fields. In *STOC*, pages 350–355, 1986.

2 M. Agrawal, C. Saha, R. Saptharishi, and N. Saxena. Jacobian hits circuits: Hitting-sets, lower bounds for depth-D occur-k formulas & depth-3 transcendence degree-k circuits. In *Proceedings of the 44th ACM Symposium on Theory of Computing (STOC)*, pages 599–614, 2012. (In SICOMP special issue).

3 Manindra Agrawal, Chandan Saha, and Nitin Saxena. Quasi-polynomial hitting-set for set-depth-δ formulas. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 321–330. ACM, 2013.

4 Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS*, pages 67–75, 2008.

5 W. Bauer and V. Strassen. The complexity of partial derivatives. *Theoretical Computer Science*, 22(3):317–330, 1983.

6 M. Beecken, J. Mittmann, and N. Saxena. Algebraic Independence and Blackbox Identity Testing. *Inf. Comput.*, 222:2–19, 2013. (Conference version in ICALP 2011).

7 David A Cox, John Little, and Donal O'Shea. Ideals, varieties, and algorithms. undergraduate texts in mathematics, 2007.

8 Radu Curticapean. Counting matchings of size k is #W[1]-hard. In *Automata, Languages, and Programming*, pages 352–363. Springer, 2013.

9 Richard A DeMillo and Richard J Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7(4):193–195, 1978.

10 Z. Dvir, A. Gabizon, and A. Wigderson. Extractors and rank extractors for polynomial sources. *Comput. Complex.*, 18(1):1–58, 2009. (Conference version in FOCS 2007).

11 Z. Dvir, D. Gutfreund, G.N. Rothblum, and S.P. Vadhan. On approximating the entropy of polynomial mappings. In *Innovations in Computer Science (ICS)*, pages 460–475, 2011.

12 Zeev Dvir. Extractors for varieties. In *Proceedings of the 24th IEEE Conference on Computational Complexity (CCC)*, pages 102–113, 2009.

13 Zeev Dvir, Swastik Kopparty, Shubhangi Saraf, and Madhu Sudan. Extensions to the method of multiplicities, with applications to kakeya sets and mergers. *SIAM Journal on Computing*, 42(6):2305–2328, 2013. (Preliminary version in FOCS'09).

14 Zeev Dvir, Amir Shpilka, and Amir Yehudayoff. Hardness-randomness tradeoffs for bounded depth arithmetic circuits. *SIAM Journal on Computing*, 39(4):1279–1293, 2009.

15 Richard Ehrenborg and Gian-Carlo Rota. Apolarity and canonical forms for homogeneous polynomials. *European Journal of Combinatorics*, 14(3):157–181, 1993.

16 Michael A Forbes. *Polynomial identity testing of read-once oblivious algebraic branching programs*. PhD thesis, Massachusetts Institute of Technology, 2014.

17 Michael A Forbes. Deterministic divisibility testing via shifted partial derivatives. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 451–465. IEEE, 2015.

18 Krister Forsman. Two themes in commutative algebra: Algebraic dependence and kähler differentials. 1992.

19 Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic circuits: A chasm at depth three. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 578–587, 2013.

20 Helmut Hasse and Friedrich K. Schmidt. Noch eine begründung der theorie der höheren differentialquotienten in einem algebraischen funktionenkörper einer unbestimmten. (nach einer brieflichen mitteilung von f.k.schmidt in jena). *Journal für die reine und angewandte Mathematik*, 177:215–223, 1937.

21 L. Illusie. Crystalline cohomology. In *Proc. Sympos. Pure Math.*, volume 55, pages 43–70, 1994. Motives (Seattle, WA, 1991).

22 I Martin Isaacs. *Algebra: a graduate course*, volume 100. American Mathematical Soc., 1994.

23 C. G. J. Jacobi. De determinantibus functionalibus. *J. Reine Angew. Math.*, 22(4):319–359, 1841.

24 Nathan Jacobson. *Lectures in Abstract Algebra: III. Theory of Fields and Galois Theory*, volume 32. Springer Science & Business Media, 2012.

25 K. A. Kalorkoti. A Lower Bound for the Formula Size of Rational Functions. *SIAM J. Comp.*, 14(3):678–687, 1985. (Conference version in ICALP 1982).

26 N. Kayal. The Complexity of the Annihilating Polynomial. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity (CCC)*, pages 184–193, 2009.

27 Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An exponential lower bound for homogeneous depth four arithmetic formulas. In *Foundations of Computer Science (FOCS), IEEE 55th Annual Symposium on*, pages 61–70. IEEE, 2014.

28 Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 146–153. ACM, 2014.

29 Gregor Kemper. *A course in Commutative Algebra*, volume 256. Springer Science & Business Media, 2010.

30 Anthony W Knapp. *Advanced algebra*. Springer Science & Business Media, 2007.

31 Steven G Krantz and Harold R Parks. *The implicit function theorem: history, theory, and applications*. Springer Science & Business Media, 2012.

32 Martin Kreuzer and Lorenzo Robbiano. *Computational commutative algebra 2*. Springer Science & Business Media, 2005.

33 Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. In *Foundations of Computer Science (FOCS), IEEE 55th Annual Symposium on*, pages 364–373. IEEE, 2014.

34 Mrinal Kumar and Shubhangi Saraf. Arithmetic circuits with locally low algebraic rank. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:194, 2015. (To appear in CCC 2016).

35 M.S. L'vov. Calculation of invariants of programs interpreted over an integrality domain. *Cybernetics and Systems Analysis*, 20:492–499, 1984.

36 Johannes Mittmann. *Independence in Algebraic Complexity Theory*. PhD thesis, Universitäts-und Landesbibliothek Bonn, 2013.

37 Johannes Mittmann, Nitin Saxena, and Peter Scheiblechner. Algebraic independence in positive characteristic: A p-adic calculus. *Transactions of the American Mathematical Society*, 366(7):3425–3450, 2014.

38 James G Oxley. *Matroid theory*, volume 3. Oxford university press, 2006.

39 O. Perron. *Algebra I (Die Grundlagen)*. W. de Gruyter, Berlin, 1927.

40 A. Płoski. Algebraic Dependence of Polynomials After O. Perron and Some Applications. In *Computational Commutative and Non-Commutative Algebraic Geometry*, pages 167–173. 2005.

41 Ran Raz. Elusive functions and lower bounds for arithmetic circuits. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 711–720. ACM, 2008.

42 Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity. 2016. https://github.com/dasarpmar/lowerbounds-survey/.

43 Nitin Saxena. Progress on polynomial identity testing-ii. In *Perspectives in Computational Complexity*, pages 131–146. Springer, 2014.

44 J.T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, 1980.

45 A. Shpilka and A. Yehudayoff. Arithmetic Circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, 2010.

46 Volker Strassen. Vermeidung von divisionen. *Journal für die reine und angewandte Mathematik*, 264:184–202, 1973.

47 Oswald Teichmüller. Differentialrechnung bei charakteristik p. *Journal für die reine und angewandte Mathematik*, 175:89–99, 1936.

48 William Nathaniel Traves. Differential operators and Nakai's conjecture. 1998.

49 Richard Zippel. *Probabilistic algorithms for sparse polynomials*. Springer, 1979.

A Main structure theorems

A.1 Proof of Thm.10

This section proves several technical properties about the arithmetic modulo $\mathcal{U}_t$. The most important of these is:

LEMMA 12 (RESTATED). *Let $\mathbb{F}$ be any field, $R := \mathbb{F}(\mathbf{z})[\mathbf{x}]$, $\mathcal{Q}_r := R/\langle\mathbf{x}\rangle^{r+1}$ for $r \geq 1$, and $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$. Define $\mathcal{U}_1 = \mathcal{V}_1 = \{0\}$, and for $u \in \langle\mathbf{x}\rangle_R$, $r \geq 2$, define the subspaces (in the quotient algebra $\mathcal{Q}_r$),*

$$\mathcal{U}_r := \langle \mathcal{H}_{r-1}f_1, \ldots, \mathcal{H}_{r-1}f_n\rangle^2_{\mathbb{F}(\mathbf{z})} + \cdots + \langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_n\rangle^r_{\mathbb{F}(\mathbf{z})}\,,$$

$$\mathcal{V}_r := \langle \mathcal{H}_{r-1}f_1, \ldots, \mathcal{H}_{r-1}f_{n-1}, u\rangle^2_{\mathbb{F}(\mathbf{z})} + \cdots + \langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_{n-1}, u\rangle^r_{\mathbb{F}(\mathbf{z})}\,.$$

If $\mathcal{H}_t f_n \in \langle \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1}, u\rangle_{\mathbb{F}(\mathbf{z})} + \mathcal{U}_t$, then $\mathcal{U}_t \subseteq \mathcal{V}_t$ (for any $t \in \mathbb{N}$).

Proof. We prove the lemma using induction on t.

Base Case ($t = 2$): By definition, $\mathcal{U}_2 = \langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_n\rangle^2_{\mathbb{F}(\mathbf{z})}$. Now, from the hypothesis, we have that, in $\mathcal{Q}_1$: $\langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_n\rangle_{\mathbb{F}(\mathbf{z})} \subseteq \langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_{n-1}, u\rangle_{\mathbb{F}(\mathbf{z})}$.

Apply the powering (Lem.15 with $t = 1, i = 2$) to get, in $\mathcal{Q}_2$, $\langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_n\rangle^2_{\mathbb{F}(\mathbf{z})} \subseteq \langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_{n-1}, u\rangle^2_{\mathbb{F}(\mathbf{z})}$. So, $\mathcal{U}_2 \subseteq \mathcal{V}_2$ and the base case is true.

Induction Step: The induction hypothesis is that the lemma holds for all $t < \ell$. To prove the lemma for $t = \ell$, we take $\mathcal{Q}_\ell$ and its subspace $\mathcal{U}_\ell$, and consider its general summand $\langle \mathcal{H}_r f_1, \ldots, \mathcal{H}_r f_n\rangle^{\ell+1-r}_{\mathbb{F}(\mathbf{z})}$ from the above sum of subspaces ($r \in [\ell - 1]$). We try to show the containment of this summand in a desired subspace. Firstly, note that the dependence hypothesis (with Lem.16) gives, in $\mathcal{Q}_r$,

$$\langle \mathcal{H}_r f_1, \ldots, \mathcal{H}_r f_n\rangle_{\mathbb{F}(\mathbf{z})} \subseteq \langle \mathcal{H}_r f_1, \ldots, \mathcal{H}_r f_{n-1}, u\rangle_{\mathbb{F}(\mathbf{z})} + \mathcal{U}_r\,.$$

By the induction hypothesis on $\mathcal{U}_r$, $r < \ell$, we get, in $\mathcal{Q}_r$,

$$\langle \mathcal{H}_r f_1, \ldots, \mathcal{H}_r f_n\rangle_{\mathbb{F}(\mathbf{z})} \subseteq \langle \mathcal{H}_r f_1, \ldots, \mathcal{H}_r f_{n-1}, u\rangle_{\mathbb{F}(\mathbf{z})} + \cdots + \langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_{n-1}, u\rangle^r_{\mathbb{F}(\mathbf{z})}\,.$$

Apply the powering (Lem.15, with $t = r$ and $i = \ell + 1 - r$) to get, in $\mathcal{Q}_\ell$,

$$\langle \mathcal{H}_r f_1, \ldots \mathcal{H}_r f_n\rangle^{\ell+1-r}_{\mathbb{F}(\mathbf{z})} \subseteq \langle v_1^{q_1} \cdots v_r^{q_r} \mid q_1 + \cdots + q_r = \ell + 1 - r\,,\, q_j \geq 0, \mathbf{v}\rangle_{\mathbb{F}(\mathbf{z})} \tag{2}$$

where we consider all the possible $v_j \in \langle \mathcal{H}_{r-j+1} f_1, \ldots \mathcal{H}_{r-j+1} f_{n-1}, u\rangle^j_{\mathbb{F}(\mathbf{z})}$ for $j \in [r]$. Now observe that, for any f, $\mathcal{H}_1 f, \ldots, \mathcal{H}_r f, u$ are all in $\langle\mathbf{x}\rangle_R$.

So, the least degree term (wrt variables $\mathbf{x}$) of the above product $v_1^{q_1} \cdots v_r^{q_r}$ would have degree at least $s := q_1 + 2q_2 + \cdots + rq_r$. In $\mathcal{Q}_\ell$, only the terms with degree $\leq \ell$ survive.

This restricts s in the range: $\ell + 1 - r \leq s \leq \ell$ and we only need to consider the corresponding $r + 1$ subspaces $\langle \mathcal{H}_r f_1, \ldots \mathcal{H}_r f_n\rangle^s_{\mathbb{F}(\mathbf{z})}$ in the RHS of Eqn.2. This allows us to rewrite Eqn.2 as (recall Rmk.11),

$$\langle \mathcal{H}_r f_1, \ldots \mathcal{H}_r f_n\rangle^{\ell+1-r}_{\mathbb{F}(\mathbf{z})} \subseteq \langle \mathcal{H}_r f_1, \ldots, \mathcal{H}_r f_{n-1}, u\rangle^{\ell+1-r}_{\mathbb{F}_p(\mathbf{z})} + \cdots + \langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_{n-1}, u\rangle^\ell_{\mathbb{F}(\mathbf{z})}\,.$$

Hence, we now have the desired containment for a general summand of $\mathcal{U}_\ell$. Since in $\mathcal{U}_\ell$, r is in the range $[\ell-1]$, we get that, in $\mathcal{Q}_\ell$,

$$\mathcal{U}_\ell \subseteq \langle \mathcal{H}_{\ell-1}f_1, \ldots, \mathcal{H}_{\ell-1}f_{n-1}, u\rangle^2_{\mathbb{F}(\mathbf{z})} + \cdots + \langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_{n-1}, u\rangle^\ell_{\mathbb{F}(\mathbf{z})}\,.$$

This proves $\mathcal{U}_\ell \subseteq \mathcal{V}_\ell$, finishing the induction step. ◀

Now we easily generalize a structural property of $\Sigma\Gamma^{(k)}\Sigma\Pi^d$ circuits [34, Lem.3.5], which will be used in the lower bound and PIT applications later.

▶ **Corollary 14** (Rewrite $\Sigma\Gamma^{(k)}\Sigma\Pi^d$). *Let $\mathbb{F}$ be an arbitrary field. Let $C = \sum_{i=1}^T F_i(Q_{i1}, \ldots, Q_{it})$ be a $\Sigma\Gamma^{(k)}\Sigma\Pi^d$ circuit in $\mathbb{F}[x_1, \ldots, x_N]$ of degree n, with $\mathcal{B}_i := \{Q_{i1}, \ldots, Q_{ik}\}$ be a separating transcendence basis of $\{Q_{i1}, \ldots, Q_{it}\}$, for all $i \in [T]$. Then, for random $\mathbf{a} \in \overline{\mathbb{F}}^N$, there exist polynomials F_i' in variables at most $k(d+1)$ over $\overline{\mathbb{F}}$ such that*

$$C(\mathbf{x}+\mathbf{a}) = \sum_{i=1}^T F_i'(\, Q_{i1}^{[\leq d]}(\mathbf{x}+\mathbf{a}), \ldots, Q_{ik}^{[\leq d]}(\mathbf{x}+\mathbf{a})\,)\,.$$

Proof. This follows from our functional dependence result (Thm.10), and the univariate interpolation trick from [34, Cor.3.4]: From the representation $f(\mathbf{x}+\mathbf{a}) = h^{\leq d}(g_1(\mathbf{x}+\mathbf{a}), \ldots, g_k(\mathbf{x}+\mathbf{a}))$ one can get an h' and an absolute representation $f(\mathbf{x}+\mathbf{a}) = h'(g_1^{[\leq d]}(\mathbf{x}+\mathbf{a}), \ldots, g_k^{[\leq d]}(\mathbf{x}+\mathbf{a}))$, for $d \geq \text{degree}(f)$. Applying this idea on each Q_{ij} gives the desired result. ◀

We now prove a (standard) property of ideal *powering* in a filtration. Essentially, one needs a "lower accuracy" $a_1, \ldots, a_i \in \mathcal{Q}_j$ to compute their product $a_1 \cdots a_i$.

▶ **Lemma 15** (Powers in filtration). *Recall the algebras $R := \mathbb{F}(\mathbf{z})[\mathbf{x}]$ and $\mathcal{Q}_t$, $t \geq 1$. If, for $j \in [i]$, $b_j \in \langle \mathbf{x}\rangle_R$ and $a_j \equiv b_j$ in $\mathcal{Q}_t$, then $a_1 \cdots a_i \equiv b_1 \cdots b_i$ in $\mathcal{Q}_{t+i-1}$, for $i \geq 1$.*

Proof. The congruence $a_j \equiv b_j$ in $\mathcal{Q}_t$ implies that $a_j - b_j$ is a polynomial $\alpha_j(\mathbf{x})$ in $\mathcal{I}_0^{t+1}$. We write it as $a_j = b_j + \alpha_j(\mathbf{x})$, and take the product on both sides. This yields $\prod_j a_j = \prod_j (b_j + \alpha_j)$ which is contained in $\prod_j b_j + \mathcal{I}_0^{t+1} \cdot \mathcal{I}_0^{i-1}$, which is in $\prod_j b_j + \mathcal{I}_0^{i-1+t+1}$ [$\because \mathcal{I}_0$ is an ideal of R, and each b_j is in $\mathcal{I}_0$]. In other words, $\prod_j a_j \in \prod_j b_j + \mathcal{I}_0^{i+t}$.

Hence, $\prod_j a_j \equiv \prod_j b_j$ in $\mathcal{Q}_{t+i-1}$. ◀

The following lemma implies that proving the linear independence for truncation t suffices to prove it for every truncation above t. Moreover, it also implies that proving the dependence for truncation t suffices to prove it for every truncation below t.

▶ **Lemma 16** (Descent). *If $\mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_n$ are $\mathbb{F}(\mathbf{z})$-linearly dependent modulo $\mathcal{U}_t$, then $\mathcal{H}_r f_1, \ldots, \mathcal{H}_r f_n$ are $\mathbb{F}(\mathbf{z})$-linearly dependent modulo $\mathcal{U}_r$, for all $r \in [t]$.*

Proof. If we see the linear dependence of $\mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_n$ modulo $\mathcal{U}_t$ in the quotient ring $\mathcal{Q}_r$ instead (i.e. reduce modulo $\langle \mathbf{x}\rangle_R^{r+1}$), then we get the dependence of $\mathcal{H}_r f_1, \ldots, \mathcal{H}_r f_n$ modulo $\mathcal{U}_r$. This is true since $\mathcal{H}_t f = \mathcal{H}_r f +$ (degree$> r$)-terms in $\mathbf{x}$, and $\mathcal{Q}_r$ filters out $\langle \mathbf{x}\rangle_R^{r+1}$. ◀

A.2 Proof of Thm.13

Theorem 13 (restated). *Let $\mathbf{f} \subset \mathbb{F}_q[\mathbf{x}]$ be algebraically independent polynomials (wlog n-variate n polynomials) with inseparable degree p^i. Then,*

1. *for all $t \geq p^i$, for random $\mathbf{a} \in \overline{\mathbb{F}}_q^n$, $f_n^{\leq t}(\mathbf{x}+\mathbf{a})$ cannot be written as $h^{\leq t}(f_1(\mathbf{x}+\mathbf{a}), \ldots, f_{n-1}(\mathbf{x}+\mathbf{a}))$, for any $h \in \overline{\mathbb{F}}_q[Y_1, \ldots, Y_{n-1}]$.*

2. *for all* $1 \leq t < p^i$, $\exists j \in [n]$, *for random* $\mathbf{a} \in \overline{\mathbb{F}}_q^n$, $f_j^{\leq t}(\mathbf{x}+\mathbf{a})$ *can be written as* $h_{jt}^{\leq t}(f_1(\mathbf{x}+\mathbf{a}),\ldots,f_{j-1}(\mathbf{x}+\mathbf{a}),f_{j+1}(\mathbf{x}+\mathbf{a}),\ldots,f_n(\mathbf{x}+\mathbf{a}))$, *for some* $h_{jt} \in \overline{\mathbb{F}}_q[\mathbf{Y}]$.

Proof. [$t < p^i$ *part.*] We first prove the dependence part of the theorem. We use the shifts on the annihilating polynomial of the algebraically dependent set $\{x_j, \mathbf{f}\}$ and then argue about desired dependence by making use of the arguments used in the proof of Thm.10.

The descent principle (Lem.16) implies that we need to prove it only for $t = p^i - 1$. Algebraic independence of $\mathbf{f}$ asserts the existence of the minimal annihilating polynomial $A_j \in \mathbb{F}_q[y_0, y_1, \ldots, y_n]$ for the polynomials $\{x_j, \mathbf{f}\}$, for all $j \in [n]$ (because of Lem.19). Now the inseparable degree of the extension $\mathbb{F}_q(\mathbf{x})/\mathbb{F}_q(\mathbf{f})$ being p^i implies that there exists a j such that A_j lives in $\mathbb{F}_q[y_0^{p^i}, y_1, \ldots, y_n]$ but not in $\mathbb{F}_q[y_0^{p^{i+1}}, y_1, \ldots, y_n]$. Let us fix that j. Thus, we have $A_j(x_j, \mathbf{f}) = \sum_{\mathbf{e}_\ell} \alpha_{\mathbf{e}_\ell} \cdot (x_j^{p^i})^{e_{0\ell}} f_1^{e_{1\ell}} \cdots f_n^{e_{n\ell}} = 0$, where $\alpha_{\mathbf{e}_\ell} \in \mathbb{F}_q$.

Next we apply the shift and note that truncating $A_j(x_j, \mathbf{f})$ at degree $\leq p^i - 1$ is same as looking at $A_j(x_j, \mathbf{f})$ in $\mathcal{Q}_{p^i-1}$. In $\mathcal{Q}_{p^i-1}$, the above equation gives us $\sum_{\mathbf{e}_\ell} \alpha_{\mathbf{e}_\ell} \cdot (z_j^{p^i})^{e_{0\ell}} \cdot f_1^{e_{1\ell}}(\mathbf{x}+\mathbf{z}) \cdots f_n^{e_{n\ell}}(\mathbf{x}+\mathbf{z}) \equiv 0$, since in $\mathcal{Q}_{p^i-1}$, $(x_j + z_j)^{p^i} \equiv z_j^{p^i}$.

We can now repeat the arguments used in Eqn.1 (Sec.A.1) to get that for some j', $f_{j'}^{\leq p^i-1}(\mathbf{x}+\mathbf{z}) = h_{j'}^{\leq p^i-1}(f_1(\mathbf{x}+\mathbf{z}),\ldots,f_{j'-1}(\mathbf{x}+\mathbf{z}),f_{j'+1}(\mathbf{x}+\mathbf{z}),\ldots,f_n(\mathbf{x}+\mathbf{z}))$ for some $h_{j'} \in \mathbb{F}_q[Y_1,\ldots,Y_{n-1}]$ to finish the proof of the dependence part of the theorem.

[$t \geq p^i$ *part.*] Next, we prove the independence part of the theorem which gives us the independence testing criterion, and we do it by contradiction. The contrapositive of Lem.16 implies that proving the theorem for $t = p^i$ suffices. For contradiction, assume that (wlog) $f_n^{\leq p^i}(\mathbf{x}+\mathbf{z})$ can be written as $h^{\leq p^i}(f_1(\mathbf{x}+\mathbf{z}),\ldots,f_{n-1}(\mathbf{x}+\mathbf{z}))$ for some $h \in \mathbb{F}_q[Y_1,\ldots,Y_{n-1}]$ which implies that the non-constant part of $f_n(\mathbf{x}+\mathbf{z})$ $\mathbb{F}_q(\mathbf{z})$-linearly depends on the non-constant parts of $f_1(\mathbf{x}+\mathbf{z}),\ldots,f_{n-1}(\mathbf{x}+\mathbf{z})$ modulo the subspace $\mathcal{U}_{p^i}$. Thus, $\mathcal{H}_{p^i} f_n$ $\mathbb{F}_q(\mathbf{z})$-linearly depends on $\mathcal{H}_{p^i} f_1, \ldots, \mathcal{H}_{p^i} f_{n-1}$ modulo the subspace $\mathcal{U}_{p^i}$.

We are given that the inseparable degree of the field extension $\mathbb{F}_q(\mathbf{x})/\mathbb{F}_q(\mathbf{f})$ is p^i. This by the definition of inseparable degree (Sec.2.1) implies that the minimal annihilating polynomial $A_j \in \mathbb{F}_q[y_0,\ldots,y_n]$ of $\{x_j^{p^i}, \mathbf{f}\}$ is separable with respect to y_0, for all j, i.e. the derivative of A_j does not vanish with respect to y_0.

Let us consider such an $A_j = \sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \mathbf{y}^{\mathbf{e}_\ell}$. We begin by applying the variable shift as we did in the dependent case, and get that $A_j((x_j + z_j)^{p^i}, \mathbf{f}(\mathbf{x}+\mathbf{z})) \equiv 0$ in $\mathcal{Q}_{p^i}$. Now Taylor expansion allows us to write $f(\mathbf{x}+\mathbf{z})$ as $f(\mathbf{z}) + \mathcal{H}_{p^i} f(\mathbf{x})$ in $\mathcal{Q}_{p^i}$ (i.e sum of constant terms and non-constant terms of degree $\leq p^i$). Using this, we expand the congruence as $\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \cdot (z_j^{p^i} + x_j^{p^i})^{e_{0\ell}} \cdot (f_1(\mathbf{z}) + \mathcal{H}_{p^i} f_1)^{e_{1\ell}} \cdots (f_n(\mathbf{z}) + \mathcal{H}_{p^i} f_n)^{e_{n\ell}} \equiv 0$.

Note that $(z_j^{p^i} + x_j^{p^i})^{e_{0\ell}} \equiv z_j^{p^i e_{0\ell}} + e_{0\ell} \cdot z_j^{p^i(e_{0\ell}-1)} x_j^{p^i}$. Using this, we further expand to,

$$\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \cdot \left(z_j^{p^i e_{0\ell}} + e_{0\ell} \cdot z_j^{p^i(e_{0\ell}-1)} x_j^{p^i} \right) \cdot (f_1(\mathbf{z}) + \mathcal{H}_{p^i} f_1)^{e_{1\ell}} \cdots (f_n(\mathbf{z}) + \mathcal{H}_{p^i} f_n)^{e_{n\ell}} \equiv 0.$$

Observe that $x_j^{p^i} \cdot \mathcal{H}_{p^i} f_\ell \equiv 0$ in $\mathcal{Q}_{p^i}$, for $\ell \in [n]$. Thus, the above equation reduces to

$$\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} \cdot z_j^{p^i e_{0\ell}} \cdot (f_1(\mathbf{z}) + \mathcal{H}_{p^i} f_1)^{e_{1\ell}} \cdots (f_n(\mathbf{z}) + \mathcal{H}_{p^i} f_n)^{e_{n\ell}} + x_j^{p^i} \cdot \sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} e_{0\ell} \cdot z_j^{p^i(e_{0\ell}-1)} \cdot \mathbf{f}^{\mathbf{e}_\ell} \equiv 0.$$

Thus, an $\mathbb{F}_q(\mathbf{z})$-linear combination of $x_j^{p^i}$ and the products of the form $(\mathcal{H}_{p^i} f_1)^{t_1} \cdots (\mathcal{H}_{p^i} f_n)^{t_n}$ vanishes in $\mathcal{Q}_{p^i}$.

By the separability of A_j at least one $e_{0\ell}$ is not a multiple of p. Now having shown that there is at least one non-zero term in the sum $\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} e_{0\ell} \cdot (z_j^{p^i})^{e_{0\ell}-1} \cdot \mathbf{f}^{\mathbf{e}_\ell}$, we argue that the

overall sum cannot be zero. This follows immediately from the minimality of A_j again since the zero sum would imply the existence of an annihilating polynomial with degree less than the degree of A_j. Thus, we get that $x_j^{p^i}$ lives in the subspace generated by the terms of the form $(\mathcal{H}_{p^i}f_1)^{t_1}\cdots(\mathcal{H}_{p^i}f_n)^{t_n}$, with $\sum_j t_j \geq 1$. (Note that the $\mathbf{x}$-free terms cancel out.)

We write the above subspace as $\langle\mathcal{H}_{p^i}\mathbf{f}\rangle_{\mathbb{F}_q(\mathbf{z})} + \langle\mathcal{H}_{p^i}\mathbf{f}\rangle^2_{\mathbb{F}_q(\mathbf{z})} + \cdots + \langle\mathcal{H}_{p^i}\mathbf{f}\rangle^{p^i}_{\mathbb{F}_q(\mathbf{z})}$ which, by Rmk.11, is the same as the subspace $\langle\mathcal{H}_{p^i}\mathbf{f}\rangle_{\mathbb{F}_q(\mathbf{z})} + \mathcal{U}_{p^i} =: \mathcal{U}'_{p^i}$. Using the assumption of the linear dependence of $\mathcal{H}_{p^i}\mathbf{f}$ modulo $\mathcal{U}_{p^i}$, and subspace reduction (Lem.12), we get that $x_j^{p^i}$ lives in $\mathcal{U}'_{p^i} = \mathcal{V}'_{p^i} := \langle\mathcal{H}_{p^i}f_1,\ldots,\mathcal{H}_{p^i}f_{n-1}\rangle_{\mathbb{F}_q(\mathbf{z})} + \mathcal{V}_{p^i}$, where $\mathcal{V}_{p^i} := \langle\mathcal{H}_{p^i}f_1,\ldots,\mathcal{H}_{p^i}f_{n-1}\rangle^2_{\mathbb{F}_q(\mathbf{z})} + \cdots + \langle\mathcal{H}_{p^i}f_1,\ldots,\mathcal{H}_{p^i}f_{n-1}\rangle^{p^i}_{\mathbb{F}_q(\mathbf{z})}$.

On repeating this for all the A_j's, we get that $\{x_1^{p^i},\ldots,x_n^{p^i}\} \subseteq \mathcal{V}'_{p^i}$. This contradicts (the impossible containment) Lem.17, and hence finishes the proof. (One can easily see that we get functional independence for random fixing of $\mathbf{z}$ in the space $\overline{\mathbb{F}}_q^n$.) ◀

We use the above notation and any field.

▶ **Lemma 17** (Impossible containment). *Let $\mathbb{F}$ be any field. Consider the subspace $\mathcal{V}'_t := \langle\mathcal{H}_tf_1,\ldots,\mathcal{H}_tf_{n-1}\rangle_{\mathbb{F}(\mathbf{z})}+\ldots+\langle\mathcal{H}_1f_1,\ldots,\mathcal{H}_1f_{n-1}\rangle^t_{\mathbb{F}(\mathbf{z})}$ of $\mathcal{Q}_t$, for $t \geq 1$. Then, $\{x_1^t,\ldots,x_n^t\} \not\subseteq \mathcal{V}'_t$.*

Proof. Rmk.11 suggests that $\mathcal{V}'_t$ equals the subspace $\langle\mathcal{H}_tf_1,\ldots,\mathcal{H}_tf_{n-1}\rangle_{\mathbb{F}(\mathbf{z})}+\cdots+\langle\mathcal{H}_tf_1,\ldots,\mathcal{H}_tf_{n-1}\rangle^t_{\mathbb{F}(\mathbf{z})}$ in $\mathcal{Q}_t$.

Intuitively, these n 'pure' monomials $x_1^t,\ldots,x_n^t$ should not all appear in the subspace $\mathcal{V}'_t$ as it has merely $n-1$ many "key" generators. However, assume for the sake of contradiction that $\{x_1^t,\ldots,x_n^t\} \subseteq \mathcal{V}'_t$. We rewrite this in absolute terms (in R) as:

$$x_i^t + \alpha_i \in \langle\mathcal{H}_tf_1,\ldots,\mathcal{H}_tf_{n-1}\rangle_{\mathbb{F}(\mathbf{z})} + \cdots + \langle\mathcal{H}_tf_1,\ldots,\mathcal{H}_tf_{n-1}\rangle^t_{\mathbb{F}(\mathbf{z})} ,$$

for some $\alpha_i \in \langle\mathbf{x}\rangle_R^{t+1}$, for all $i \in [n]$. This simply means $x_i^t + \alpha_i = P_i(\mathcal{H}_tf_1,\ldots,\mathcal{H}_tf_{n-1})$, for some polynomial $P_i \in \mathbb{F}(\mathbf{z})[Y_1,\ldots,Y_{n-1}]$ of degree at most t, for $i \in [n]$. Notice that the degree of α_i (in $\mathbf{x}$) is $\geq t+1$. Thus, by choosing a graded lexicographic monomial ordering (see [7, Pg.58]) in which lower degree terms lead, we get the leading monomials of the set $\{x_i^t + \alpha_i \mid i \in [n]\}$ to be $\{x_1^t,\ldots,x_n^t\}$.

Now, using the fact that the algebraic independence of leading monomials imply the algebraic independence of the corresponding polynomials (Lem.18), we get that $\mathrm{trdeg}_{\mathbb{F}(\mathbf{z})}\{x_i^t + \alpha_i | i \in [n]\} = n$. On the other hand, clearly, $\mathrm{trdeg}_{\mathbb{F}(\mathbf{z})}\{P_i(\mathcal{H}_tf_1,\ldots,\mathcal{H}_tf_{n-1}) | i \in [n]\} \leq n-1$. This makes the containment impossible. ◀

▶ Remark. The proof works if we replace the n pure monomials by any polynomials whose leading monomials are algebraically independent and appear in degree$\leq t$ part (under some strict monomial ordering in which lower degree terms lead).

We give the following for the sake of completeness.

▶ **Lemma 18.** *[32, Prop.6.6.11] Let $f_1,\ldots,f_n \in \mathbb{F}[x_1,\ldots,x_n]$ be non-zero polynomials. If under some (strict) monomial ordering σ, leading monomials of $f_1,\ldots,f_n$ are algebraically independent over $\mathbb{F}$, then $f_1,\ldots,f_n$ are algebraically independent over $\mathbb{F}$.*

Proof. Let us fix the monomial ordering σ, and let the leading monomials of $f_1,\ldots,f_n$ wrt σ be $LM(f_1),\ldots,LM(f_n)$ respectively (they uniquely exist as σ is strict and total). By the hypothesis the leading monomials are algebraically independent.

Recall that for $h_1, h_2 \in \mathbb{F}[x_1,\ldots,x_n]$, the LM operator has the properties (eg. [29, Sec.9.1]):

- $LM(h_1 \cdot h_2) = LM(h_1) \cdot LM(h_2)$,
- $LM(h_1 + h_2) \preceq_\sigma \max\{LM(h_1), LM(h_2)\}$.

We use the above two properties to prove the lemma. Consider any nonzero polynomial $g \in \mathbb{F}[y_1, \ldots, y_n]$, and let m be the monomial in the support of g such that $m(LM(f_1), \ldots, LM(f_n))$ is maximal with respect to σ. Hence, for any monomial m' in the support of g, and any monomial k_i in the support of f_i ,

$$m'(k_1, \ldots, k_n) \preceq_\sigma m'(LM(f_1), \ldots, LM(f_n)) \preceq_\sigma m(LM(f_1), \ldots, LM(f_n)) .$$

In this case the last inequality cannot be equality, unless $m' = m$. Otherwise, $m' - m$ is the annihilating polynomial of the leading monomials, contradicting the hypothesis.

This proves that the monomial $m(LM(f_1), \ldots, LM(f_n))$ cannot cancel with other monomials in $g(\mathbf{f}(\mathbf{x}))$. This implies that there is no nonzero annihilating polynomial for $f_1, \ldots, f_n$. ◀

A.3 Technical lemmas

For completeness, we present standard results that entail that for our main Theorems (Thm.10 and Thm.13) it suffices to study the case of n polynomials in n variables over an algebraically closed field. The first lemma handles the case when the polynomials are more than the number of variables.

▶ **Lemma 19** (Extra polys). *If $m > n$ then any $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ are algebraically dependent.*

Proof. It is proved in [40, 18] and books on field theory [24]. ◀

The next lemma deals with the case when the variables are more than the number of polynomials. We can use this lemma to project n variables to a random m dimensional subspace (over a large enough field extension L of $\mathbb{F}$) in our input polynomials. Thus, in case $n > m$, we reduce to the case of m polynomials with m variables.

▶ **Lemma 20** (Extra variables). *Let $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ with $m < n$ and the transcendence degree of the set $\{f_1 \ldots, f_m\}$ be r. Then, there exists a linear map $\phi : L[x_1, \ldots, x_n] \mapsto L[y_1, \ldots, y_m]$ such that $trdeg_L\{\phi(f_1), \ldots, \phi(f_m)\}$ is also r.*

Proof. Roughly, the idea is to consider the annihilating polynomial A_S of $\{\mathbf{x}_S, \mathbf{f}\}$, and study the action of a 'random' linear ϕ on it. For a proof refer to [6, Theorem 4]. ◀

For algebraic independence over a field, it suffices to work over the algebraic closure.

▶ **Lemma 21** (Closed field). *Consider polynomials $\mathbf{f}(\mathbf{x})$ over any field $\mathbb{F}$. Their trdeg remains invariant if we move from $\mathbb{F}$ to any algebraic extension.*

Proof. Let $B = \{g_1, \ldots, g_r\}$ be a transcendence basis of $\mathbf{f}$ over $\mathbb{F}$. Let us move to the algebraic closure $\overline{\mathbb{F}}$. Clearly, any $f_i \in \mathbf{f}$ continues to be algebraically dependent on B as the original annihilating polynomial works.

Suppose polynomials in B become algebraically dependent over $\overline{\mathbb{F}}$. Then, by Perron's bound [40] we know that $\{\mathbf{g}^{\mathbf{e}} \mid |\mathbf{e}| \leq \prod_i \deg(f_i)\}$ has to be $\overline{\mathbb{F}}$-linearly dependent. But these polynomials are in $\mathbb{F}[\mathbf{x}]$, so they must be $\mathbb{F}$-linearly dependent, implying that B is algebraically dependent over $\mathbb{F}$. This contradiction proves the lemma. ◀

A.4 Recovering the classics

As a corollary of Thm.10 and Thm.13, we get the classical Jacobian criterion for the separable case (i.e. inseparable degree $= p^0 = 1$).

▶ **Corollary 22** (Jacobian rephrased). *Let $\mathbb{F}$ be any field. Let $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$ be such that the field extension $\mathbb{F}(\mathbf{x})/\mathbb{F}(\mathbf{f})$ is separable, then the linear terms (in $\mathbf{x}$) of $f_1(\mathbf{x}+\mathbf{z}), \ldots, f_n(\mathbf{x}+\mathbf{z})$ are $\mathbb{F}(\mathbf{z})$-linearly dependent iff $f_1, \ldots, f_n$ are algebraically dependent.*

The dependence part of Thm.13 helps us in characterizing the failure of the Jacobian.

▶ **Corollary 23** (Jacobian fails for inseparable). *For algebraically independent polynomials $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$ such that the field extension $\mathbb{F}(\mathbf{x})/\mathbb{F}(\mathbf{f})$ is inseparable, the linear terms (in $\mathbf{x}$) of $f_1(\mathbf{x}+\mathbf{z}), \ldots, f_n(\mathbf{x}+\mathbf{z})$ are $\mathbb{F}(\mathbf{z})$-linearly dependent.*

Thus, Jacobian being zero implies that either the n-variate n polynomials are algebraically dependent, or *they are independent but inseparable.*

B Application 1: Algebraic indepedence testing algorithm

THEOREM 1 (RESTATED). *For circuits $\mathbf{f} \in \mathbb{F}_q[\mathbf{x}]$ we have a randomized poly(s, $\binom{t+n}{n}$)-time algebraic independence testing algorithm, where the inseparable degree of the field extension $\mathbb{F}_q(\mathbf{x})/\mathbb{F}_q(\mathbf{f})$ is t (assuming $\mathbf{f}$ algebraically independent) and s is the total input size.*

Algorithm idea: The criterion (by Theorems 10 & 13) essentially involves testing $\mathcal{H}_t f_n \equiv 0$ modulo the subspace $\mathcal{V}'_t := \langle 1, \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1} \rangle^t_{\mathbb{F}_q(\mathbf{z})}$ in $\mathcal{Q}_t$, where t is the inseparable degree of the field extension $\mathbb{F}_q(\mathbf{x})/\mathbb{F}_q(\mathbf{f})$. (In fact, one needs to check whether $\mathcal{H}_t f_j$ functionally depends on the remaining $n-1$ polynomials, for all $j \in [n]$.) Implementing the criterion involves three main steps:

Step 1: Computing the arithmetic circuits for $\mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_n$ in $\mathcal{Q}_t$ using the fact that $\mathcal{H}_t f = f(\mathbf{x}+\mathbf{z}) - f(\mathbf{z})$ in $\mathcal{Q}_t$.

Step 2: Computing the arithmetic circuits for the basis vectors generating the subspace $\mathcal{V}'_t$ in $\mathcal{Q}_t$.

Step 3: Testing the nonzeroness of $\mathcal{H}_t f_n$ modulo the linear space $\mathcal{V}'_t$ given its basis vectors as circuits, in $\mathcal{Q}_t$.

A subroutine that we use several times in our algorithm computes a basis of a given subspace, over the field $\mathbb{F}(\mathbf{z})$, generated by given arithmetic circuits in $\mathbb{F}[\mathbf{z}][\mathbf{x}]$. Let us call this subroutine BASIS.

B.1 The subroutine BASIS

Suppose we are given m circuits $a_1, \ldots, a_m \in \mathbb{F}_q[\mathbf{z}][\mathbf{x}]$ and we want to compute a basis B of the subspace generated by $a_1, \ldots, a_m$ over $\mathbb{F}_q(\mathbf{z})$. Let d be a degree bound (wrt $\mathbf{x}, \mathbf{z}$), and s a size bound, for these circuits.

We invoke the *Alternant criterion* as proven in [36, Lem.3.1.2]. It says that– If $a_1, \ldots, a_m$ are $\mathbb{F}_q(\mathbf{z})$-linearly independent, then for "random" points $\alpha_i, i \in [m]$, in $\mathbb{F}_q^n$, $\det(a_j(\alpha_i)) \neq 0$. For this to work we need $q > 2dm$. Note that such a field extension $\mathbb{F}_q/\mathbb{F}_p$ can be constructed in polylog(dm)-time by [1]. Once we have fixed the $\mathbf{x}$ variables we still have to test $\det(a_j(\alpha_i)) \neq 0$. This we can do by, again, randomly fixing the $\mathbf{z}$ variables to a single point in $\mathbb{F}_q^n$ [44, 9, 49].

Moreover, to compute a basis B we merely have to find a *column-basis* of the matrix $(a_j(\alpha_i))_{i,j}$. This can be done by basic linear algebra (using minors and random evaluations

as above), in randomized poly$(sm \log d)$-time. So BASIS runs in randomized poly-time in the input size.

B.2 Computing the arithmetic circuits for $\mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_n$

Recall that $\mathcal{H}_t f = f(\mathbf{x}+\mathbf{z}) - f(\mathbf{z})$ in $\mathcal{Q}_t$. Since $\mathcal{H}_t f$ is nothing but the non-constant part of the shifted f, truncated at degree t, we can get the circuit for $\mathcal{H}_t f$ by shifting the variables of $f(\mathbf{x})$ and using standard circuit reductions.

Given an arithmetic circuit for $f(\mathbf{x})$, we easily get the circuit for $f(\mathbf{x}+\mathbf{z})$. Now to get the terms with degree $\leq t$ wrt $\mathbf{x}$, from the above circuit, use Strassen's homogenization technique [46, 45, Thm.2.2] which gives a homogeneous circuit of size $O(t^2 s)$ computing the homogeneous parts of $\mathcal{H}_t f$ upto degree t.

B.3 Computing the basis vectors of $\mathcal{V}'_t$

Recall that $\mathcal{V}'_t$ is generated as $\langle 1, \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1} \rangle^t_{\mathbb{F}(\mathbf{z})}, t \geq 1$, in $\mathcal{Q}_t$. Now, having computed the circuits for $\mathcal{H}_t f_j$ in $\mathcal{Q}_t$, we compute the generators for $\mathcal{V}'_t$ iteratively.

We first compute the linear basis $\mathcal{B}_1$ of the set, of above computed circuits $\{1, \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1}\}$, using the subroutine BASIS.

Next, we multiply every element of the obtained basis to every element of the set $\{1, \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1}\}$ in $\mathcal{Q}_t$ and compute the basis $\mathcal{B}_2$ of the corresponding set of products obtained.

We repeat the procedure and multiply every element of $\mathcal{B}_2$ to every element of $\{1, \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1}\}$ and compute the basis to obtain $\mathcal{B}_3$, and so on.

Clearly, the size of the intermediate basis $\mathcal{B}_i$ remains bounded by the dimension of $\mathcal{Q}_t$ which is $\binom{n+t}{n}$. Further, we only need to go up to $i \leq t$.

Hence, we compute the final basis, using BASIS, in randomized poly$(s, \binom{n+t}{n})$-time.

B.4 Testing nonzeroness modulo the subspace $\mathcal{V}'_t$

We now test nonzeroness of $\mathcal{H}_t f_n$ modulo $\mathcal{V}'_t$. This is simply the question of computing the dimension of the subspace spanned by $\{\mathcal{H}_t f_n\} \cup \mathcal{B}_t$ and the one by $\mathcal{B}_t$, and checking whether the difference is 1. Clearly, BASIS can be used to do this in randomized poly$(s, \binom{t+n}{n})$-time.

Thus, we have a poly$(s, \binom{t+n}{n})$-time randomized algorithm for testing algebraic independence, where t upper bounds the inseparable degree of the field extension $\mathbb{F}_q(\mathbf{x})/\mathbb{F}_q(\mathbf{f})$ and s is the input size. This finishes the proof of Thm.1.

C Application 2: Exponential lower bounds

In this section, we prove

THEOREM 2 (RESTATED). *Let $\mathbb{F}$ be any field. There exists a family $\{P_n\}$ of polynomials in VNP, such that P_n is a polynomial of degree n in $N = n^{O(1)}$ variables with $0,1$ coefficients, and for any $\Sigma\Gamma^{(k)}\Sigma\Pi$ circuit C, if $k \leq n$ and if C computes P_n over $\mathbb{F}$, then Size$(C) \geq N^{\Omega(\sqrt{n})}$.*

We state the main lemmas following the notation of [34, Sec.4] and discuss proof ideas; the details are the same as those in [34]. The main reason why their lower bound result needed characteristic of the underlying field to be zero (or large enough) is due to the fact that their key lemma (algebraic dependence to functional dependence) worked only for those

characteristics. As we have generalized the key lemma to arbitrary fields, we are able to generalize their lower bound results to arbitrary fields as well.

All the recent arithmetic circuit lower bound proofs follow a common recipe with the following main steps. (Refer to the evolving survey [42].)

- Coming up with a complexity *measure* on polynomials that is sub-additive.
- Calculating an *upper bound* on the complexity measure of the family of circuits against which we would like to prove the lower bound.
- Calculating a *lower bound* on the complexity measure for the hard polynomial.
- Set appropriate parameters and compare these bounds using *binomial estimates*.

Following [34] we adopt the same strategy here.

C.1 The complexity measure: dimension of projected shifted partial derivatives space

The complexity measure used in [34] is dimension of projected shifted partial derivatives of a polynomial. This measure was used in [27] to prove a strong lower bound against homogeneous depth-4 circuits for zero or large characteristic. Later [33, 34] extended it to other models.

For a polynomial P and a monomial γ, $\frac{\partial P}{\partial \gamma}$ is the partial derivative of P with respect to γ. For a set of monomials $\mathcal{M}$, $\partial_{\mathcal{M}}(P)$ is the set of partial derivatives of P with respect to monomials in $\mathcal{M}$. Mult$[P]$ is the projection of P on the multilinear monomials in its support.

▶ **Definition 24** ($(\mathcal{M}, m)$-projected shifted partial derivatives [34])**.** For an N variate polynomial $P \in \mathbb{F}[X_1, \ldots, X_N]$, set of monomials $\mathcal{M}$ of degree r and a positive integer $m \geq 0$, the space of (M, m)-projected shifted partial derivatives of P is defined as

$$\langle \partial_{\mathcal{M}}(P) \rangle_m := \mathbb{F} - \text{span}\left(\text{Mult}\Big[\prod_{i \in S} X_i \cdot g \Big] \,:\, g \in \partial_{\mathcal{M}}(P), S \in \binom{[N]}{m} \right) .$$

The complexity measure we use is dimension of the projected shifted partial derivatives space. Formally, $\phi_{\mathcal{M},m}(P) := \text{Dim}(\langle \partial_{\mathcal{M}}(P)_m \rangle)$.

It is easy to check that the measure is subadditive.

The following lemma is used in the proof and it is easy to verify that it is valid for all characteristic. It gives an upper bound on the measure of the homogeneous component of a polynomial of low degree.

▶ **Lemma 25.** *[34, Lem.4.3] Let P be a polynomial of degree at most d. Then for every $0 \leq i \leq d$ and for all choice of parameters m, r and a set $\mathcal{M}$ of monomials of degree r,*

$$\phi_{\mathcal{M},m}(P^{=i}) \leq \phi_{\mathcal{M},m}(P) .$$

C.2 Target polynomials for the lower bound

The target polynomial family (in VNP) is a variant of Nisan-Wigderson polynomials – Nisan-Wigderson composed with linear forms. First, we give the definition of Nisan-Wigderson family of polynomials, which was first introduced in [28].

▶ **Definition 26** (Nisan-Wigderson family of polynomials, [34] Defn.4.5)**.** Let n, q, e be arbitrary parameters with q being a power of prime and $n, e \leq q$. We have some identification $[n] \subseteq \mathbb{F}_q$. The Nisan-Wigderson polynomial with parameters n, q, e, denoted by $\mathrm{NW}_{q,n,e}$ is defined as

$$\mathrm{NW}_{q,n,e}(\overline{X}) := \sum_{\substack{p(t)\in\mathbb{F}_q[t] \\ \deg(p)<e}} X_{1,p(1)} \cdots X_{n,p(n)} .$$

Note that it has arity equal to $N = nq$. Now we define the family of polynomials which is hard for the circuit model we consider. This is in VNP (by Valiant's criterion).

▶ **Definition 27** (Nisan-Wigderson composed with linear forms, [34] Defn.4.6)**.** Let $\delta \in (0, 1)$ be an arbitrary constant, and let $p = N^{-\delta}$. Let $\gamma = N/p$. The polynomial $\mathrm{NW} \circ \mathrm{Lin}_{q,n,e,p}$ is defined as

$$\mathrm{NW} \circ \mathrm{Lin}_{q,n,e,p} = \mathrm{NW}_{q,n.e}\Big(\sum_{i=1}^{\gamma} X_{1,1,i}, \sum_{i=1}^{\gamma} X_{1,2,i}, \ldots \sum_{i=1}^{\gamma} X_{n,q,i}\Big) .$$

This polynomial, of arity γN, behaves well under random restrictions on the variables. Let V be the set of variables in the polynomial $\mathrm{NW} \circ \mathrm{Lin}$. We define a distribution $\mathcal{D}_p$ over the subsets of V as follows. Each variable in V is independently kept alive with a probability $p = N^{-\delta}$.

We notice that [34, Lem.4.7] & [33, Sec.6] (lower bound on the dimension of projected shifted partial derivatives of NW) holds for any field $\mathbb{F}$ (unlike [27]).

▶ **Lemma 28** (NW lower bound, [34] Lem.4.7, [33] Sec.6)**.** *For every n and $r = O(\sqrt{n})$, there exists parameters q, e, ε such that $q = \Omega(n^2)$, $N = qn$ and $\varepsilon = \Theta(\frac{\log n}{\sqrt{n}})$ with $q^r \geq (1+\varepsilon)^{2(n-r)}$ and $q^{e-r} = (\frac{2}{1+\varepsilon})^{n-r} \cdot poly(q)$. For any n, q, e, r, ε satisfying the above constraints, for $m = \frac{N}{2}(1-\varepsilon)$, over* any *field $\mathbb{F}$, we have*

$$\phi(\mathrm{NW}_{q,n,e}) \geq \binom{N}{m+n-r} \cdot \exp(-O(\log^2 n)) .$$

Using the above lemma it can be shown that, with high probability, the measure of $\mathrm{NW} \circ \mathrm{Lin}$ remains high.

▶ **Lemma 29** ([34] Lem.4.8)**.** *With probability $1 - o(1)$ over $V \leftarrow \mathcal{D}_p$, there exist variables $V' \subseteq V$ with N elements such that $\phi(\mathrm{NW} \circ \mathrm{Lin}|_{V'}) \geq \binom{N}{m+n-r} \cdot \exp(-O(\log^2 n))$.*

The proof is given in [34] and, importantly, works for any field.

C.3 Measure upper bound for $\Sigma\Gamma^{(k)}\Sigma\Pi^d$

Observe that, by our functional dependence result (Thm.10), the relevant proof of [34, Lem.4.9] immediately extends over arbitrary fields. (A technical point is that one uses $\overline{\mathbb{F}}$ in their arguments.)

▶ **Lemma 30** (Measure upper bound, [34] Lem.4.9)**.** *Let m, r, s be parameters such that $m + rs \leq N/2$. Let M be any set of multilinear monomials of degree r. Let C be a $\Sigma\Gamma^{(k)}\Sigma\Pi^d$ circuit computing a homogeneous polynomial of degree n such that*

$$C = \sum_{i=1}^{T} C_i(Q_{i1}, Q_{i2}, \ldots, Q_{it})$$

where for each $i \in [T]$, C_i is an arbitrary polynomial in t variables, for each $(i,j) \in [T] \times [t]$, Q_{ij} is a homogeneous polynomial in N variables and for each $i \in [T]$, the algebraic rank of $\{Q_{ij} : j \in [t]\}$ is at most k. Let S_{ij} be the support of Q_{ij} and assume it to have monomials of support$\leq s$. If

$$\left| \bigcup_{i \in [T], j \in [t]} S_{ij} \right| \leq N^{\frac{\delta s}{2}}$$

then, with probability $1 - o(1)$ over $V \leftarrow \mathcal{D}_p$, for all subsets V' of V of size at most N,

$$\phi(C|_{V'}) \leq T \binom{k(n+1)+r}{r} \binom{N}{m+rs}.$$

The proof strategy is the same as [34]. The first step is using random restrictions to simplify the circuit into a circuit with bounded bottom support. This step is not sensitive to the characteristic or size of the underlying field.

The step crucial for us is their second step, where low algebraic rank is exploited in the rewriting (Cor.14). Here, we invoke functional dependence (Thm.10) to get the same upper bound.

The third step is to simply estimate the measure once t has been reduced to $k(n+1)$. We note that this part is purely combinatorial and field independent.

C.4 Wrapping up

Finally, assume that NW $\circ$ Lin has a circuit $C \in \Sigma\Gamma^{(k)}\Sigma\Pi^d$. Consider the degree-$n$ homogeneous part of the randomly shifted NW $\circ$ Lin polynomial (this gives back the original polynomial).

The above analysis (with Cor.14 & Lem.25) entails that: with a positive probability, there exists a subset V' of variables of size N so that simultaneously

$$\phi_{\mathcal{M},m}(C|_{V'}) \leq T \binom{k(n+1)+r}{r} \binom{N}{m+rs}$$

and

$$\phi_{\mathcal{M},m}(\mathrm{NW} \circ \mathrm{Lin}|_{V'}) \geq \binom{N}{m+n-r} \exp(-\log^2 n).$$

As C computes NW $\circ$ Lin ,

$$T \geq \frac{\binom{N}{m+n-r} \exp(-\log^2 n)}{\binom{k(n+1)+r}{r} \binom{N}{m+rs}}.$$

Setting appropriate parameters as in [34, Pg.21], we would get

$$T = N^{\Omega(\sqrt{n})}.$$

D Application 3: Quasipoly-time hitting-set

In this section, we prove

THEOREM 3 (RESTATED). *Let $\mathbb{F}$ be any field of characteristic p. There exists an $\exp(\log^{O(1)} s)$-time constructible hitting-set $\mathcal{H} \subseteq \overline{\mathbb{F}}^N$ for size-s circuit $C \in \Sigma\Gamma^{(k)}\Sigma\Pi^d$ with $kd = \log^{O(1)} s$, assuming $p >$ individual-degree(C) or $p = 0$.*

We only sketch the proof ideas, along the lines of [34, Sec.5]. Their main trick is the following. If we can prove that every nonzero polynomial P (of degree at most n and in N variables) in the class $\Sigma\Gamma^{(k)}\Sigma\Pi^d$ (of size s) has a monomial of low support (say, at most ℓ) then a hitting-set for the class can be easily constructed in poly$(s(nN)^\ell)$-time (see [3]).

This trick was combined with the shifted partials measure by Forbes [17] for interesting models, to get hitting-sets and also to solve circuit divisibility testing questions. Basically, he showed that a circuit with a low measure also has a low support *trailing monomial*. [34, Lem.5.2] proved the same for $\Sigma\Gamma^{(k)}\Sigma\Pi^d$ circuits. Albeit their proof requires the characteristic to be zero or super-polynomially large.

We extend [34, Lem.5.2] to fields of characteristic greater than the individual-degree of the circuit. To prove this, [17, Lem.4.18] is used which related shifted partials measure to the support of trailing monomial.

▶ **Lemma 31** ([17] Lem.4.18, [34] Lem.5.3)**.** *Let $\mathbb{F}$ be a field with characteristic p. Let $R(\overline{X})$ be a polynomial in $\mathbb{F}[\overline{X}]$ such that*

$$R(\overline{X}) = \sum_{i=1}^{T} F_i(Q_{i1}, Q_{i2}, \dots Q_{it})$$

and for each $(i,j) \in [T]\times[t]$, the degree of Q_{ij} is at most d. Let α be the trailing monomial of R. If $p = 0$ or $p >$ individual-degree(α), then the support of α is at most $2e^3 d(\ln T + t\ln 2t + 1)$ (e is Euler's constant).

Now, using our rewriting of $\Sigma\Gamma^{(k)}\Sigma\Pi^d$ circuits (Cor.14), we can generalize [34, Lem.5.2].

▶ **Lemma 32** (Trailing monomial has low support, [34] Lem.5.2)**.** *Let $\mathbb{F}$ be a field of characteristic p. Let P be a homogeneous polynomial of degree Δ in N variables such that P can be represented as*

$$P = \sum_{i=1}^{T} C_i(Q_{i1}, Q_{i2}, \dots, Q_{it})$$

such that the following are true.

For each $i \in [T]$, C_i is a polynomial in t variables.

For each $i \in [T]$ and $j \in [t]$, Q_{ij} is a polynomial of degree at most d in N variables.

For each $i \in [T]$, the algebraic rank of the set of polynomials $Q_{ij} : j \in [t]$ is at most k.

Let α be the trailing monomial of P. If $p = 0$ or $p >$ individual-degree(α), then α has support at most

$$2e^3 d \cdot (\ln(T(\Delta+1)) + (d+1)k\ln((d+1)k) + 1)\,.$$

Sketch of Proof. We want to show that

$$P = \sum_{i=1}^{T} C_i(Q_{i1}, Q_{i2}, \dots, Q_{it})$$

has a trailing monomial of low support.

The proof uses Cor.14, which we have shown for arbitrary fields (after a shift from $\overline{\mathbb{F}}^N$), to reduce t to $k(d+1)$. Now, invoke Lem.31, which requires a mildly large characteristic, to deduce that the trailing monomial has low support.

The proof also uses the fact that degree-Δ homogeneous component of the shifted $P(\mathbf{x})$ is P itself, and applies Lemmas 25 & 30, to upper bound the circuit's measure. ◀

Note that a non-homogeneous P can be first made homogeneous, for PIT purposes, and then apply the above.

Thus, if $dk = \text{polylog}(s)$ then we get a quasipoly-time hitting-set for $\Sigma\Gamma^{(k)}\Sigma\Pi^d$ circuits.