Power series in complexity: Algebraic Dependence, Factor Conjecture and Hitting Set for Closure of VP

A Thesis Submitted

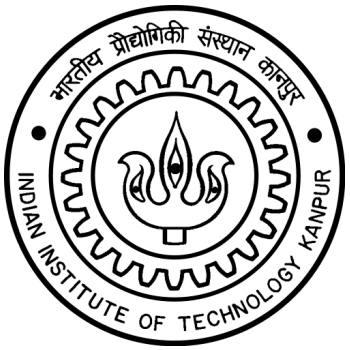
in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

by

Amit Kumar Sinhababu



to the

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

July, 2019

CERTIFICATE

It is certified that the work contained in the thesis entitled "Power series in complexity: Algebraic Dependence, Factor Conjecture and Hitting Set for Closure of VP", by "Amit Kumar Sinhababu", has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

Nitin Saxena
Department of Computer Science and Engineering
Indian Institute of Technology Kanpur

July, 2019

SYNOPSIS

Algebraic complexity is about studying polynomials from a computational viewpoint. In this thesis, we report progress on the following three problems in algebraic complexity. A common theme in our study is the use of formal power series concepts.

Testing Algebraic Dependence Over Finite Fields: Algebraic dependence is a natural generalization of linear dependence. Polynomials $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ are *algebraically dependent* if they satisfy a nonzero polynomial $A(f_1, \ldots, f_m) = 0$. If there is no such nonzero polynomial A, the given polynomials are called algebraically independent. A fundamental question is testing algebraic dependence of a given set of polynomials. In computer science, it has applications in derandomization of polynomial identity testing and explicit constructions of randomness extractors from polynomial sources.

Over the fields of characteristic zero or large, a classical criterion due to Jacobi (using the Jacobian matrix of the given polynomials) gives a randomized polynomial time algorithm to test algebraic dependence of polynomials. Over fields of small characteristic, the Jacobian criterion *fails*. A natural question is whether algebraic dependence can be tested in randomized polynomial time over small characteristic as well. In this thesis, we present two results that make progress on this question.

First, we give a natural generalization of the classical Jacobian criterion that works over all fields. Our proof exploits the Taylor expansion of a polynomial and can be interpreted using *Hasse derivatives*. Our criterion is efficient (randomized polynomial time) only under a given promise that *inseparability degree*, a relevant parameter here, is bounded by a small constant. Our criterion is stated using the concept of *functional dependence*, that asks, for

a given set of polynomials, whether one of the polynomials (after applying a random shift on the variables) can be *approximated* by a polynomial function of the other polynomials (shifted) in the set. This notion of functional dependence can be related to formal power series.

Using a different approach, we show that dependence testing over finite fields is in AM $\cap$ coAM. Prior to our work, it was known to be in $\mathrm{NP}^{\#\mathsf{P}}$, a class just below PSPACE and much higher than all of the polynomial hierarchy. Our result rules out the plausibility of NP-hardness of testing algebraic dependence over finite fields (unless polynomial hierarchy collapses) and brings us closer to the goal of finding a randomized polynomial time algorithm for the problem. Our proof uses elementary ideas of algebraic geometry and Goldwasser-Sipser set lowerbound protocol.

Towards Factor Conjecture: Multivariate polynomial factorization is a fundamental problem in computational algebra. A classical result of Kaltofen shows that factors of a polynomial given by size s arithmetic circuit of degree d can be computed by circuits of size bounded by $\mathrm{poly}(s, d)$. The factor conjecture states that any factor g of a polynomial f can be computed by circuits of size $\mathrm{poly}(s, d_g)$, where d_g is the degree of the factor g. In this thesis, we confirm the conjecture in the case when the degree of the squarefree part of the given polynomial is *low*.

A PSPACE Construction of Hitting Set for $\overline{\mathbf{VP}}$: Algebraic complexity class VP contains the families of low degree polynomials computed by small-sized circuits. The closure of the class VP, denoted as $\overline{\mathrm{VP}}$, contains families of low degree polynomials infinitesimally approximated by small-sized circuits. Over the field of reals, it contains the limits (in metric topology) of the polynomials in VP.

A *hitting set* $\mathcal{H}$ for a class of polynomials $\mathcal{C}$ is a set of points, such that any nonzero polynomial in the class C evaluates to a nonzero value in at least one point in $\mathcal{H}$. The existence of small-sized hitting sets for VP is known. The problem of giving an explicit construction of a hitting set for the class of VP is a challenging problem in complexity theory. It is expected that this problem can be solved in polynomial time, but the currently

known upper bound is PSPACE. In this thesis, we study the hitting set construction problem for $\overline{\text{VP}}$.

The question of explicit construction of a hitting set for $\overline{\text{VP}}$ has connections with fundamental problems in computational algebraic geometry. It can be directly solved in EXPSPACE and it was asked whether the complexity upper bound can be improved. Recently it was shown to be in PSPACE over characteristic zero fields using analytic concepts. We give a different approach of construction that works over arbitrary fields.

We reduce the problem of constructing a hitting set for $\overline{\text{VP}}$ to another problem: Approximate polynomials satisfiability. The well-known problem of polynomial system satisfiability checks for the existence of a common solution (over the closure of the underlying field) of a given system of polynomial equations. Approximate polynomials satisfiability asks for the existence of an infinite sequence of points such that all the polynomials in the given system, evaluated at that sequence, approach zero in the limit. Using classical algebraic geometry, this problem is equivalent to testing if the point of origin is in the Zariski closure of the image of the given polynomials. Testing if the origin is in closure of the image is equivalent to testing whether all the annihilating polynomials of the given polynomials have constant term zero. The latter problem can be solved in PSPACE over any field, thereby using our reduction, a hitting set for $\overline{\text{VP}}$ can be constructed in PSPACE.

Acknowledgements

I am fortunate to have Prof. Nitin Saxena as my advisor. I am thankful to him for many reasons, especially for hours of intense discussions interspersed with humor, for teaching me many nuggets of math and complexity, for being a source of ideas and for encouraging me to pursue my personal interests and approaches at my own pace. Even in the very beginning, when I barely knew anything, he made me feel confident. That I have made some progress over the years is owing to his patience and continued guidance.

This thesis is based on joint works with Anurag Pandey, Pranjal Dutta, Zeyu Guo, and my advisor. I am fortunate to have Anurag, Pranjal and Zeyu as friends and coauthors. Their clarity of thoughts, intuition and knowledge influenced and improved my way of thinking and understanding. This work would not have been possible without them.

I had the pleasure of interacting with quite a few brilliant researchers in our community. Many thanks to Sumanta Ghosh, Arpita Korwar, Rafael Oliveira, Prerona Chatterjee, Abhibhav Garg, Akash Jena, Tushant Mittal, Nikhil Balaji, Mrinal Kumar, Rohit Gurjar, Anamay Tengse, Pranav Bisht for helpful discussions. Special thanks to Ramprasad, for sharing many beautiful ideas through discussions and expositions. Interacting with him is always a joyful experience.

At IIT Kanpur, I was fortunate to experience the teaching of Prof. Manindra Agrawal, Prof. Sumit Ganguly and Prof. Surender Baswana. I am grateful to Prof. Agrawal for advising me in M.Tech. He encouraged to try challenging problems and showed how to explore a problem, starting from simple ideas. I want to thank Prof. Anil Seth, Prof. Piyush Kurur, Prof. Sanjeev Saxena and Prof. Shashank Mehta for the theory courses. Thanks to Prof. Satyadev Nandakumar, Prof. Raghunath Tewari and Prof. Rajat Mittal

for several discussions on various topics and helpful advice and feedback. Thanks to Prof. Somenath Biswas for encouraging us. Thanks to the staff in the CSE department and Hall-4 for always being helpful.

I thank the organizers of the workshops I participated: WACT 2016 in Tel Aviv, Mysore Park workshop in 2016 and NMI workshop in arithmetic circuit complexity at IMSc Chennai in 2017. Thanks to Microsoft Research India, Indian Association for Research in Computing Science, ACM India and Research-I foundation of Infosys for funding my travel to the conferences CCC 2018 and STOC 2018. Thanks to Prof. Thomas Thierauf for inviting to Ulm for Postdoc.

I am grateful to Prof. Sukanta Das and Prof. Arindam Biswas for motivating me to pursue higher studies and to Prof. Suryasarathi Barat for his courses on numerical analysis and automata theory that made me interested in theoretical computer science. In Shibpur, I had the privilege of meeting Prof. Asok Kumar Mallik. I am fortunate to get affection and motivation from him.

Thanks to my labmates (Diptarka, Sumanta, Ram, Garima, Keerti, Debarati, Rajendra, Mahesh, Prateek, Ras) for the wonderful time we had together. Many thanks to Ramranjan, Debojyoti, Sanghamitra, Diptarka, Biman, Niladri, Gogol, Jaydeep, Ankurji, Vikraman, Bikram, Anupam and Sudipto for the close friendship and support over the years. Special thanks to Sumanta for our decade-long friendship spanning the entire journey of Shibpur and Kanpur. I have learned many things from him. He introduced me to the joy of running. Thanks to Ashish Dwivedi and Arpita Korwar for giving me an immense amount of support, affection and helpful advice.

The love and support from my parents and my brother Arnab made it possible for me to pursue studies. I fondly dedicate this thesis to my father, whose curiosity and interest opened the wonderful world of science for me and to my late grandfather, who taught me how to read.

Contents

List of Publications

[PSS18] **Algebraic independence over positive characteristic: New criterion and applications to locally low-algebraic-rank circuits**
with Anurag Pandey and Nitin Saxena.
Computational Complexity, Volume 27, Issue 4 (also in MFCS 2016)

[DSS18] **Discovering the roots: Uniform closure results for algebraic classes under factoring**
with Pranjal Dutta and Nitin Saxena.
STOC 2018

[GSS18] **Algebraic dependencies and PSPACE algorithms in approximative complexity**
with Zeyu Guo and Nitin Saxena.
CCC 2018 (Invited in Theory of Computing)

Chapter 3 is based on [PSS18], Chapter 4 is based on a part of [GSS18], Chapter 5 is based on a part of [DSS18], Chapter 6 is based on a part of [GSS18]. Chapter 1, Chapter 2 and Chapter 7 reuses some parts of [PSS18], [GSS18] and [DSS18]. Some results from [DSS18] presented in Chapter 5 also appeared in the thesis [Dut18].

List of Figures

Chapter 1

Introduction

Polynomials are important in mathematics and computation. Diverse computational problems can be expressed via polynomials. In algebraic complexity theory, we analyze the *complexity* of computing a polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$. Here complexity is measured by the minimum number of additions and multiplications needed to compute the polynomial, starting from the variables $x_1, \ldots, x_n$ and freely using any constant from the underlying field $\mathbb{F}$. This notion of complexity is formalized via the model of *arithmetic circuits*.

An arithmetic circuit is a directed acyclic graph consisting of addition ($+$) and multiplication ($\times$) gates as nodes. It takes variables $x_1, \ldots, x_n$ and field constants as input nodes (leaves), and outputs a polynomial $f(x_1, \ldots, x_n)$ in the root node. The *size* of a circuit is defined to be the total number of edges in it. Arithmetic circuit is a succinct description of multivariate polynomials, as polynomials having many monomials (or of high degree) can sometimes be represented by small circuits. For example, the polynomial $\prod_{i=1}^{n}(1 + x_i)$ has 2^n many monomials, but it has a circuit of size $O(n)$.

The algebraic complexity class VP contains families of n-variate polynomials of low ($n^{O(1)}$) degree computed by small ($n^{O(1)}$) sized circuits. For example, the family of polynomials defined by the determinant of the symbolic matrix $X = (x_{ij})_{1 \le i,j \le n}$ is in VP [SY10]. For more on arithmetic circuits and various algebraic complexity classes, see the surveys [Mah14, SY10].

To deal with some problems on polynomials, we may need the more general object

known as *formal power series*. Formal power series can be informally described as a polynomial with infinitely many terms. For example, $\sum_{i\geq 0} x^i$ is a formal power series. Here x is a formal variable, we do not substitute it by any number.

To show the applicability of formal power series in algebraic complexity, we discuss the following well-known example. Suppose, in an arithmetic circuit, we allow *division* gates. A natural question is whether division gates add more power to the model of arithmetic circuits. Strassen [Str73] showed that if a polynomial $p(x_1, \ldots, x_n)$ of degree d can be computed by an arithmetic circuit (with division gates) of size s, we can compute the same polynomial by an arithmetic circuit of size $\text{poly}(n, s, d)$ that uses only addition and multiplication gates. Strassen's proof goes via computing a power series truncated up to some degree. See Lemma 2.8.3 for the details.

In the problems studied in this thesis, power series techniques turn out to be useful. The first problem is about testing whether there exists a nontrivial dependence, captured by a nonzero polynomial, between some polynomials over a finite field. If we ask a more general question, whether there exists a nontrivial relationship, captured by a formal power series, it gives a better understanding of the former question. The second problem we study is about multivariate polynomial factorization. Here also, if we extend the context to the formal power series ring and compute factors there, we get insights into the arithmetic circuit complexity of the factors. Finally, in the problem of hitting set construction for the closure of VP, formal power series concepts are used again.

Now, we discuss the problem of testing algebraic dependence of polynomials over finite fields. For a quick introduction to finite fields, see Section 2.2.2 in Chapter 2.

1.1 Algebraic Dependence of Polynomials

Algebraic dependence is a fundamental concept in algebra that captures the phenomena when a tuple of numbers or polynomials over a field satisfy a nonzero polynomial equation. Polynomials $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ are called *algebraically dependent* over field $\mathbb{F}$ 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 a polynomial A is called an *annihilating polynomial/annihilator* of $f_1, \ldots, f_m$. If no such nonzero polynomial exists, then the polynomials are called *algebraically independent* over $\mathbb{F}$.

For example, $f_1 = (x + y)$ and $f_2 = (x + y)^2$ are algebraically dependent over any field, and $y_1^2 - y_2$ is an annihilating polynomial of them. For any prime number p, the polynomials $x + y$ and $x^p + y^p$ are algebraically dependent over the fields of characteristic p (as $(x + y)^p = x^p + y^p$ in characteristic p). But they are algebraically independent over fields of characteristic zero (for example, $\mathbb{Q}$) and over fields of positive characteristic $\ell \neq p$. Polynomials $x_1, \ldots, x_n$ are examples of algebraically independent polynomials over any field.

Algebraic dependence is a natural generalization of linear dependence. Linear dependence implies algebraic dependence. Both algebraic dependence and linear dependence satisfy the *matroid* properties [Oxl06]. For example, if a set of polynomials is algebraically (linearly) independent, then any subset of them is algebraically (linearly) independent. Steinitz exchange lemma, the other key axiom of matroids, holds for both.

A transcendence base of a given set of polynomials is a maximal subset of algebraically independent polynomials. It follows from the matroid properties, that any two transcendence bases have the same cardinality. The *transcendence degree/algebraic rank* (trdeg or algRank) of a set of polynomials is defined as the cardinality of a transcendence base of the set. Any $n + 1$ vectors from $\mathbb{F}^n$ are always linearly dependent. Analogously, any $n + 1$ polynomials from $\mathbb{F}[x_1, \ldots, x_n]$ are always algebraically dependent. So the transcendence degree of a set of n-variate polynomials is at most n.

We are interested in the following two computational problems related to algebraic dependence.

Problem 1 (Algebraic Dependence Testing (AD($\mathbb{F}$)))**.** *Given polynomials $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$, test if they are algebraically dependent or not.*

Problem 2 (Computing an Annihilator)**.** *If the polynomials $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ are algebraically dependent, compute an annihilating polynomial of them.*

A direct approach of testing algebraic dependence of $f_1, \ldots, f_m$ goes via testing linear dependence of polynomials $\{f_1^{e_1} f_2^{e_2} \ldots f_m^{e_m}\}$ for all possible nonnegative integer exponents $e_1, \ldots, e_m$. If we know an upper bound on e_i's, this would give an algorithm to test for dependence and compute an annihilator if it exists. Perron [Per27, Pło05] gave an upper bound on the degree of an minimal annihilating polynomial, proving that it is bounded by the product of the degrees of the input polynomials. Note that this bound is exponential even for the case of n quadratic polynomials and there are examples known for which this bound is tight [Kay09].

Perron's bound reduces the problem of computing the annihilating polynomial to solving a system of exponentially many linear equations, where the coefficients of the annihilator are regarded as the unknown variables of the linear system. Solving a linear system is known to be in logspace-uniform NC [Csa75, BvzGH82, Mul87], which is contained in polylogarithmic space. Thus, solving a system of exponentially many linear equations is in PSPACE. So, algebraic independence testing (and computing an annihilator if it exists) over any field, can be solved in PSPACE.

Is there an efficient algorithm to compute an annihilator (represented by an arithmetic circuit)? Kayal [Kay09] proved the problem of deciding if the constant term of a minimal annihilating polynomial is zero is NP-hard. [Kay09] also showed that arithmetic circuit size of an annihilating polynomial may not be polynomial (wrt the input size) in general, otherwise the polynomial hierarchy collapses. Thus, computation of annihilating polynomial in general seems to be hard.

Surprisingly the decision question AD($\mathbb{F}$) is much more efficient over zero or large characteristic using a classical result, known as the Jacobian criterion [Jac41, BMS13] for testing algebraic dependence. The Jacobian criterion reduces algebraic dependence testing of $f_1, \ldots, f_m$ over $\mathbb{F}$ to linear dependence testing of the differentials $df_1, \ldots, df_m$ over $\mathbb{F}(x_1, \ldots, x_n)$, where df_i is the vector $\left(\frac{\partial f_i}{\partial x_1}, \ldots, \frac{\partial f_i}{\partial x_n}\right)$. The Jacobian matrix J of $f_1, \ldots, f_m$ has df_i in the i-th row.

The Jacobian criterion states that if the characteristic of the field is zero, or large enough (greater than the product of degrees of the given polynomials), then the algebraic

rank equals the linear rank (over $\mathbb{F}(x_1, \ldots, x_n)$) of the Jacobian matrix of the polynomials. This gives a simple randomized polynomial time algorithm for checking algebraic independence, as we can use random evaluations to compute the rank of the Jacobian matrix. It can be proved that the rank of the Jacobian matrix evaluated at a random point equals the rank of the Jacobian matrix with high probability using the Schwartz-Zippel-DeMillo-Lipton lemma [Sch80, DL78, Zip79].

If the polynomials are algebraically dependent, then their Jacobian matrix does not have full rank. 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. Similarly, the determinant of the Jacobian of x^p, y is zero, though x^p, y are independent. Note that if two algebraically independent polynomials over characteristic p have zero Jacobian determinant, then it does not imply that one or both of them has an exponent p. Consider, for example, $\{x^{p-1}y, xy^{p-1}\}$ over $\mathbb{F}_p$ for prime $p > 2$. There are infinitely many sets of algebraically independent polynomials, for which the Jacobian criterion fails. These examples come from inseparable algebraic field extensions. We give a quick informal introduction to the phenomena of inseparability here, as we need this concept to describe our results.

As mentioned before, any $n + 1$ polynomials $x_i, f_1, \ldots, f_n \in \mathbb{F}[x_1, \ldots, x_n]$ are always algebraically dependent. Let $A_i(y_0, y_1, \ldots, y_n)$ be an annihilating polynomial of minimal degree such that $A_i(x_i, f_1, \ldots, f_n) = 0$. If the polynomials $f_1, \ldots, f_n$ are algebraically independent, y_0 must occur in any annihilator $A_i(y_0, y_1, \ldots, y_n)$ of $x_i, f_1, \ldots, f_n$. Algebraically independent polynomials $f_1, \ldots, f_n$, where y_0 always occur with exponent p^k for some $k \geq 1$ in the annihilator of $x_i, f_1, \ldots, f_n$ are known as *inseparable* examples. Over fields of characteristic p, the Jacobian criterion fails for such polynomials.

If p^k is the highest exponent of p that divides every exponent of y_0 in $A_i(y_0, y_1, \ldots, y_n)$, we define the *inseparability degree* measure of $f_1, \ldots, f_n$ wrt to x_i to be p^k. The overall inseparability degree of $f_1, \ldots, f_n$ is the maximum of the inseparability degrees wrt x_i for $i \in [n]$. For more discussions on inseparability degree, see Section 2.4. Now we give a few examples to explain the definition. Let $f_1 = x_1^p, f_2 = x_2$. A minimal annihilator

of x_1, f_1, f_2 is $x_1^p - f$ and in this annihilator x_1 has exponent p. Inseparability degree of $\{f_1, f_2\}$ is p here. Inseparability degree of $f_1 = x_1^p, f_2 = x_2^{p^2}$ is p^2. When the inseparability degree is 1, it is called a *separable* case, where the Jacobian criterion always works. For example, the inseparability degree of $\{x_1, x_2\}$ is $p^0 = 1$.

For a fundamental problem like algebraic dependence testing, we would like to find an efficient (randomized poly-time) algorithm over small characteristic as well. Several papers [DGW09, Kay09, BMS13] posed this as an open question. [MSS14] gave a criterion, called *Witt-Jacobian*, that slightly improved the complexity of independence testing problem, over positive characteristic, from PSPACE to $\mathrm{NP}^{\#\mathrm{P}}$.

Before we discuss our results, we mention a few applications of algebraic dependence in computer science and other areas. Algebraic dependence is a fundamental concept in mathematics that appears in field theory, commutative algebra, algebraic geometry, invariant theory, theory of algebraic matroids. See [Ros15] for various applications of algebraic matroids in algebraic statistics and other areas. [L'v84] used annihilating polynomials in the analysis of program invariants of arithmetic circuits. To prove lower bounds on the formula size of determinant, [Kal85] used transcendence degree of polynomials as a tool.

[DGW09, Dvi09] constructed explicit deterministic randomness extractors for sources which are polynomial maps over finite fields. They used algebraic independence (and rank of the Jacobian) as a characterization of *entropy* of low degree polynomials. [DGRV11] gives a cryptography application. [BMS13, ASSS16] used the Jacobian criterion for designing deterministic polynomial time hitting sets for some special cases of the polynomial identity testing problem (PIT) and proving circuit lower bounds for restricted low depth circuits. Thus, an efficient criterion similar to Jacobian (in small characteristic) would generalize the PIT and lower bound results in [ASSS16], and explicit constructions of algebraic extractors in [DGW09] to arbitrary fields.

1.1.1 Our results on algebraic dependence over finite fields

We have two main results on algebraic dependence testing over finite fields. The first one is a generalization of the Jacobian criterion to arbitrary fields. This criterion is efficient

only in a special case. Our second result improves the complexity of algebraic dependence testing over finite fields, putting it in AM $\cap$ coAM.

Algebraic Dependence and Functional Dependence.

We mentioned earlier that transcendence basis is analogous to basis in linear algebra, but here is a difference between them. If polynomials $f_1, \ldots, f_m$ are linearly dependent, for all i, each of f_i can be expressed as a linear combination of a subset of the polynomials from $\{f_1, \ldots, f_m\}$ (forming the basis). The analogous property does not hold for algebraic dependence. If $f, g_1, \ldots, g_n$ are algebraically dependent, f may *not* be a polynomial function in $g_1, \ldots, g_n$, even if $g_1, \ldots, g_n$ is a transcendence basis. For example, x, x^2 are algebraically dependent and $\{x^2\}$ is a transcendence basis for $\{x, x^2\}$. But x is not a polynomial function in x^2, as a polynomial in x^2 cannot have a linear term like x.

Nevertheless, Kumar and Saraf [KS17] showed an analog, coming up with the notion of functional dependence. They showed that over fields of zero, or large characteristic, if polynomials $f, g_1, \ldots, g_n$ are algebraically dependent and $g_1, \ldots, g_n$ form a transcendence basis, then one can always *approximate* the polynomial f (shifted by a random point) as a polynomial function in the correspondingly shifted polynomials $g_1, \ldots, g_n$. The approximation is in the sense that we ignore the monomials with degree greater than the precision we want (in this case, the degree of f). For example, we can approximate $x + a$ as a polynomial in $(x + a)^2$ by truncating up to degree one: $x + a = \frac{(x+a)^2}{2a} + \frac{a}{2} \bmod x^2$.

Kumar and Saraf [KS17] asked if this structural property of functional dependence holds true for small characteristic. They also left open the question whether the converse of functional dependence is true, i.e. whether functional dependence implies algebraic dependence? We answer these questions, for arbitrary characteristic, in the affirmative.

Before stating the main theorems, we need the following notation. 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 $t \in \mathbb{N}$. For a polynomial $h \in \mathbb{F}[\mathbf{x}]$, $h^{\leq d}$ extracts out the degree $\leq d$ part of h. Note that by $h(g_1(\mathbf{x}), \ldots, g_n(\mathbf{x}))^{\leq d}$ we mean: first compute the composition $h(\mathbf{g}(\mathbf{x}))$ and then extract out the part with degree $\leq d$.

First, we show algebraic dependence implies functional dependence.

Theorem (Algebraic Dependence to Functional Dependence, Theorem 3.2.3)**.** *If transcendence degree 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,* $\forall i \in [m]$*,* $f_i(\mathbf{x}+\mathbf{a})^{\leq t} = h_i(g_1(\mathbf{x}+\mathbf{a}), \ldots, g_k(\mathbf{x}+\mathbf{a}))^{\leq t}$*.*

We also show the following converse.

Theorem (Algebraic independence to functional independence, Theorem 3.2.7)**.** *Let* $\mathbf{f} \subset \mathbb{F}_q[\mathbf{x}]$ *be algebraically independent polynomials (wlog* n*-variate* n *polynomials) with inseparable degree* p^i*. Then, for all* $t \geq p^i$*, for random* $\mathbf{a} \in \overline{\mathbb{F}}_q^n$*,* $f_n(\mathbf{x}+\mathbf{a})^{\leq t}$ *cannot be written as* $h(f_1(\mathbf{x}+\mathbf{a}), \ldots, f_{n-1}(\mathbf{x}+\mathbf{a}))^{\leq t}$*, for any* $h \in \overline{\mathbb{F}}_q[Y_1, \ldots, Y_{n-1}]$*.*

Generalization of the Jacobian criterion.

Our theorems (Theorem 3.2.7 and Theorem 3.2.3) can be seen as a natural generalization of the Jacobian criterion. In the Jacobian criterion, we take t to be 1. Let $\mathbb{F}$ be any field of characteristic zero. Let $f_1, f_2, \ldots, f_n$ be a set of algebraically independent polynomials. Then, for a random $\boldsymbol{a}$, there is no polynomial $h(y_1, \ldots, y_{n-1})$ such that $f_n(\boldsymbol{x}+\boldsymbol{a})^{\leq 1} = h(f_1(\boldsymbol{x}+\boldsymbol{a}), \ldots, f_{n-1}(\boldsymbol{x}+\boldsymbol{a}))^{\leq 1}$. Conversely, if f_n depends on $f_1, f_2, \ldots, f_{n-1}$, then for a random $\boldsymbol{a}$, there is a polynomial h such that $f_n(\boldsymbol{x}+\boldsymbol{a})^{\leq 1} = h(f_1(\boldsymbol{x}+\boldsymbol{a}), \ldots, f_{n-1}(\boldsymbol{x}+\boldsymbol{a}))^{\leq 1}$. The above two statements are direct consequences of the Jacobian criterion.

Algebraic independence testing using our criterion.

Suppose we are given polynomials (of total circuit size s) $f_1, \ldots, f_n \in \mathbb{F}_q[x_1, \ldots, x_n]$ and a positive integer k, that is a power of p, such that either of the following two cases holds: (1) $f_1, \ldots, f_n$ are algebraically dependent or (2) $f_1, \ldots, f_n$ are algebraically independent. Moreover, for $\forall i \in [n]$, the inseparable degree with respect to x_i, is bounded by k.

Using Theorem 3.2.7 and Theorem 3.2.7, we give an algorithm to separate the two cases in randomized poly$(s, \binom{k+n}{n})$-time. This is efficient if k is a small constant. See Section 3.3 for the details.

Interpretation using Hasse Derivative and Taylor shift. A different interpretation of our criterion can be given via Hasse derivatives (Jacobian criterion uses only first order

partial derivatives, whereas our criterion uses higher order Hasse derivatives). Here we give an overview of that interpretation.

To get a Jacobian like criterion in the inseparable case, we want an operator that has nice properties analogous to partial derivatives, but does not map x^p to zero. This leads to Hasse derivatives, a variant of usual higher order derivatives, that were originally defined by Hasse and Schmidt [HS37], and independently by Teichmüller [Tei36]. The k-th order Hasse derivative wrt a variable x is simply the k-th order usual derivative (wrt the variable x) divided by $k!$. Note that this is well defined even if k is divisible by characteristic of the field p. The p-th order Hasse derivative of x^p wrt x is 1.

In general, Hasse derivatives can be defined via multivariate Taylor expansion of a polynomial (around a generic point $\mathbf{z} = (z_1, \ldots, z_n)$). The coefficients of monomials (in $\mathbf{x} = (x_1, \ldots, x_n)$) in the shifted polynomial $f(x_1 + z_1, \ldots, x_n + z_n)$ are Hasse derivatives (see the Definition 2.6.1). To generalize the Jacobian criterion, we work with the following generalization of differential operator Df. We use the operator $\mathcal{H}_k$, on $\mathbb{F}_p[\mathbf{x}]$, for all $k \geq 1$. $\mathcal{H}_k f(\mathbf{x})$ contains all the terms, of the polynomial $f(\mathbf{x} + \mathbf{z}) - f(\mathbf{x}) \in \mathbb{F}_p(\mathbf{x})[\mathbf{z}]$, that are of degree (wrt $\mathbf{z}$) $\leq k$. Note that $\mathcal{H}_1$ operator is same as differential operator D, which is $Df = (\partial f/\partial x_1)z_1 + \cdots + (\partial f/\partial x_n)z_n$.

Now, we want to relate the algebraic independence of $\mathbf{f} := \{f_1, \ldots, f_n\}$ with the $\mathbb{F}_p(\mathbf{x})$-linear independence of $\mathcal{H}_k\mathbf{f} := \{\mathcal{H}_k f_1, \ldots, \mathcal{H}_k f_n\}$ for a large enough k ($\geq$ inseparability degree). We can not directly reduce to linear independence of $\mathcal{H}_k\mathbf{f} := \{\mathcal{H}_k f_1, \ldots, \mathcal{H}_k f_n\}$, we have to go modulo a subspace generated by t-wise products of the set $\{\mathcal{H}_k f_1, \ldots, \mathcal{H}_k f_n\}$ for all $2 \leq t \geq k$ (additionally we have to go modulo $\langle \mathbf{z} \rangle^{k+1}$, essentially truncating the polynomials up to degree k). For the precise statement, see Section 3.4 in Chapter 3.

Algebraic dependence over finite fields in AM ∩ coAM.

Our generalization of the Jacobian criterion does not improve the complexity status of dependence testing over small characteristic fields. Can it happen that the problem is NP-hard over small characteristic? In [GSS18], we improve the complexity of the problem, putting it in AM ∩ coAM. This result rules out the possibility of the problem being NP-hard or coNP hard (unless polynomial hierarchy collapses).

In Chapter 4, we prove the theorem.

Theorem (Theorem 4.2.1)**.** *Algebraic dependence testing of polynomials in* $\mathbb{F}_q[\mathbf{x}]$ *is in* $AM \cap coAM$.

Arthur-Merlin class (AM) is a randomized generalization of the class NP [AB09]. AM ∩ coAM is the class of decision problems for which both YES and NO answers can be verified by an AM protocol. It can be thought of as a randomized version of NP ∩ coNP. If such a problem is NP-hard or coNP-hard (even under random reductions) then the polynomial hierarchy collapses to its second-level [Sch88]. See Section 2.10.1 for more discussions on the AM class.

We use a standard AM protocol, the Goldwasser-Sipser set lower bound method, to separate the cases when a set S (whose membership can be tested in NP) has *small* or *large* cardinality. The set S in our context is the image of the given polynomial map (for the AM protocol) and preimage set corresponding to a random point in the image (for the co-AM) protocol. See Chapter 4 for the details.

1.2 Factor Conjecture

A basic theorem in algebra states that every polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ can be uniquely decomposed as a product of finitely many irreducible polynomials over $\mathbb{F}$. The standard proof of this theorem does not solve the algorithmic problem of polynomial factorization, which is: Given a polynomial, compute all its irreducible factors (with the corresponding multiplicities).

Kaltofen [Kal82] gave a polynomial-time algorithm to compute factors of bivariate polynomials using Hensel lifting (and univariate polynomial factoring). The algorithm of Kaltofen generalizes to polynomials with number of variables $n \geq 2$ and is efficient if the input polynomial is given in the *dense* representation (where all the coefficients corresponding to every monomial in n variables of degree $\leq d$ are given) and the output factors are also in dense representation. Dense representation is not a space efficient way to represent multivariate polynomials. A better way is to use sparse representation (where

the polynomial is given as a sum of monomials with nonzero coefficients) or the more succinct representation of arithmetic circuits. The *size* of a polynomial in a particular representation corresponds to the size needed to encode it. For the sparse representation, size is the sparsity of the polynomial (the number of monomials with nonzero coefficients in it) and for arithmetic circuits, size is the number of edges in the circuit. There are various other representations, for example, arithmetic formula, which is a special kind of arithmetic circuit having the structure of a directed tree.

Given a representation of multivariate polynomials, we are interested in the following two questions.

1. Closure question/ Factor size upper bound: Given a polynomial (in n variables) of size s and degree d in some representation, give an upper bound of the size of the all the factors in the same representation.

2. Algorithmic question: Given a polynomial of size s in some representation, output (in the same representation) all the irreducible factors (with the corresponding multiplicities).

Note that the first question is nontrivial for the sparse representation. The polynomial $x^d - 1$ has sparsity 2, but its factor $1 + x + \ldots + x^{d-1}$ has sparsity d. The polynomial $\prod_{i=1}^{n}(x_i^d - 1)$ has sparsity 2^n, but its factor $\prod_{i=1}^{n}(1 + x_i + x_i^2 + ... + x_i^{d-1})$ has sparsity $d^n = (2^n)^{\log d}$. Thus, the sparsity of the factors may not be polynomially upper bounded by the sparsity of the polynomial.

However, a remarkable result of Kaltofen [Kal89] shows that if a polynomial is given by an arithmetic circuit of size s, the factors have circuits of size polynomially bounded by s and the total degree of the input polynomial. Furthermore, Kaltofen in [Kal89] gave a randomized polynomial time algorithm to output the irreducible factors representation by circuits. This result is often stated as VP is (uniformly) closed under factors i.e. if a polynomial family $\{f_n(x_1, \ldots, x_n)\}_n$ is in VP, then a polynomial family $\{g_n(x_1, \ldots, x_n)\}_n$ is also in VP, where g_n is an arbitrary factor of f_n.

Closure under factoring has important applications in algebraic complexity, especially in hardness versus randomness tradeoff results [KI03, AV08, DSY09, AGS18, KST19]. It

is natural to ask whether classes other than VP are also closed under factoring. In VP, the polynomials are of low degree. In general, the degree of a polynomial computed by an arithmetic circuit can be exponential in the size of the circuit. For example, the polynomial x^{2^s} has a size s circuit, obtained by repeated squaring. What can be said about complexity of factors of polynomials computed by arithmetic circuits of high (exponential in size) degree? Note that Kaltofen's result [Kal89] gives an upper bound that is polynomially bounded in size, number of variables and the *degree* of the polynomial to be factored. Can we give a better bound, that is polynomially bounded only in the circuit size of the polynomial to be factored? This question was answered in negative by Lipton and Stockmeyer [LS78]. [LS78] proved there *exist* factors of $x^{2^n}-1$ that requires size $\geq \Omega(\frac{2^{n/2}}{\sqrt{n}})$. Recall that $x^{2^n}-1 = \prod_{j=1}^{2^n}(x-\zeta^j)$, where ζ denotes 2^n-th root of unity. $x^{2^n}-1$ has a circuit of size $O(n)$, but [LS78] showed that a random factor (of exponentially high degree) $\prod_{i\in S}(x-\zeta^i)$ where $S \subset [2^n]$, require an exponentially large circuit.

Now we shift our focus to the *low* degree (polynomially bounded wrt input size) factors. In the case of univariate polynomials, a factor of degree d has a size $O(d)$ arithmetic circuit. Can we say that the low degree factors of high degree circuits (of small size) computing multivariate polynomials have small size? This question was asked by Kaltofen [Kal87]. Bürgisser [Bür13, Conjecture 8.3] posed the above question as *factor conjecture*.

> *Suppose g is a factor of a polynomial $f \in \mathbb{F}[x_1,\ldots,x_n]$. If f has an arithmetic circuit of size s, g has an arithmetic circuit of size poly(s,n,d_g), where d_g is the total degree of g.*

Kaltofen [Kal87] proved that the factor conjecture is true when f is of a very special form: $f = g^e$ for some $g \in \mathbb{F}[x_1,\ldots,x_n]$. Kaltofen [Kal87] also showed that if $f = g^e h$ has an arithmetic circuit of size s, where g and h are coprime, then g has an arithmetic circuit of size poly(s,n,d_g,e), where d_g is the total degree of g. Thus, the factors with low multiplicity, have small circuits. In an interesting result [Bür04], Bürgisser proved that *approximative complexity* of the factors is polynomially bounded in the degree of the factor and size of the circuit. See Section 1.3 and Section 2.8.1, for an introduction to

approximative algebraic complexity.

1.2.1 Our result on factor conjecture

If polynomial f can be completely factorized as $\prod_{i=1}^{m} f_i^{e_i}$, where f_i are irreducible polynomials, $\prod_{i=1}^{m} f_i$ is called the squarefree part/radical of the polynomial. Note that the degree of the squarefree part can be exponentially smaller than the original polynomial. We give further evidence towards factor conjecture by proving the following theorem that relates the circuit complexity of the squarefree part of a polynomial and the polynomial itself.

In Chapter 5, we prove the following result. The proof of this given in Chapter 5 is different from the one in [DSS18, Theorem 1].

Theorem (Theorem 5.3.1)**.** *Every factor of a polynomial computed by size s circuit has circuits of size polynomially bounded by s and degree of the squarefree part of the polynomial.*

1.3 Explicit Hitting Set Construction for Closure of VP

A set of points $\mathcal{H}$ is called a *hitting set* for a set of polynomials $\mathcal{C}$ if any nonzero polynomial in $\mathcal{C}$ evaluates to a nonzero value in at least one point in $\mathcal{H}$. For example, any set of $d+1$ distinct numbers is a hitting set for the set containing all univariate polynomials of degree $\leq d$.

The notion of hitting set is important in the problem of black-box Polynomial Identity Testing (PIT): Given an arithmetic circuit of size s computing a polynomial $P(x_1, \ldots, x_n)$ of degree at most d, test if it computes the zero polynomial. In black-box PIT, the algorithm is allowed to use the circuit only to evaluate it at some points. In white-box PIT, the algorithm is allowed to access the internal structure of the circuit given. The famous Schwartz-Zippel-DeMillo-Lipton lemma [Sch80, Zip79, DL78] gives a randomized polynomial time black-box algorithm for PIT. A challenging question in complexity theory is *derandomization* of PIT, i. e. giving a deterministic polynomial time algorithm.

It can be proved that the black-box PIT for a class of polynomials $\mathcal{C}$ is equivalent to the problem of constructing a hitting set for $\mathcal{C}$ [For14]. It was shown by [HS80] that poly(n, s) sized hitting sets *exist* for the class of n-variate polynomials computed by size s arithmetic circuits over arbitrary fields. [HS80] proved that a *random* set of sufficiently large cardinality is a hitting set for the above class with high probability. The question of interest now is to give a deterministic polynomial time construction of a hitting set. We formalize the question below.

Problem 3 (Explicit construction of small hitting set)**.** *Given inputs n, s, d (in unary), compute a set of points of size* $\mathsf{poly}(n, s)$ *and bit complexity bounded by* $\mathsf{poly}(n, s, d)$*, that is guaranteed to be a hitting set for the class of size s circuits computing n-variate polynomials of degree d.*

We expect this problem to be in P, but currently the best unconditional upper bound of this problem is PSPACE. Note that this does not directly follow from the trivial derandomization of black-box PIT in PSPACE. In the hitting set construction problem, we are in interested in constructing an explicit hitting set of poly(n, s) size. The hitting set we get from the direct PSPACE derandomization of PIT is of exponential size $((d + 1)^n)$.

The PSPACE construction for VP goes via trying all possible candidate hitting sets and verifying if a candidate set is a hitting set for the class of size s circuits computing n-variate polynomials of degree d. As there are infinitely many polynomials in this class (over $\mathbb{Q}$), we can not go over all the polynomials to check whether or not one of them evaluates to zero at all the points in a given set. This verification problem can be solved via checking the existence of a solution (over the algebraic closure of the base field) of a system of polynomial equations. The latter problem is solved via Hilbert's Nullstellensatz (HN), that reduces to checking whether 1 is in the ideal generated by the given polynomials. Using effective Nullstellensatz [Kol88], this can be tested in PSPACE by solving a system of exponentially many linear equations in exponentially many unknowns[1].

Now we come to the explicit construction of hitting set for the approximative closure of VP. First, we discuss the notion of approximation in algebraic complexity. A polynomial

[1]Over the field of complex, assuming GRH, Hilbert's Nullstellensatz is in AM [Koi96].

$p(x_1, \ldots, x_n)$ over an algebraically closed field $\overline{\mathbb{F}}$ is said to be *infinitesimally approximated* by a circuit C of size s, if the circuit C computes a polynomial of the form $p + \epsilon p_1 + \epsilon^2 p_2 + \ldots + \epsilon^m p_m$, where the circuit C uses constants from the rational function field $\overline{\mathbb{F}}(\epsilon)$ (for example, $\frac{1}{\epsilon}$ can be used as a constant) and $p_1, \ldots, p_m \in \overline{\mathbb{F}}[x_1, \ldots, x_n]$. Over the field $\mathbb{R}$, if $\epsilon \to 0$, then the polynomial computed by the approximating circuit C tends to p (the polynomial it approximates).

The size of an approximating circuit can be *much smaller* than size of any circuit exactly computing the polynomial. In general, we can convert an approximative circuit to a circuit exactly computing the polynomial with exponential blow-up in size [Bür04] and it is an open question whether this exponential bound can be improved.

Given the notion of approximation, one can define approximative closure of any algebraic complexity class. The closure of the class VP (denoted as $\overline{\text{VP}}$) contains families of low degree ($n^{O(1)}$) polynomials approximated by small sized ($n^{O(1)}$) circuits. Now the hitting set construction problem for $\overline{\text{VP}}$ asks to output an explicit set of points that is a hitting set for the class of n-variate polynomials of degree at most d, infinitesimally approximated by size s arithmetic circuits. The above problem is interesting as natural questions like explicit construction of the normalization map (in Noether's Normalization Lemma) reduce to the construction of a hitting-set of $\overline{\text{VP}}$ [Mul17].

The proof [HS80] that poly(n, s) sized hitting sets *exist* for the class of n-variate polynomials of degree at most d computed by size s arithmetic circuits extends to the the class of n-variate polynomials of degree at most d, infinitesimally approximated by size s arithmetic circuits. So the hitting set construction algorithm iterates over all candidate hitting sets and verifies if a candidate set is indeed a hitting set. Here also the verification can be reduced to checking for solution of a system of polynomial equations, but the number of equations are exponentially many. So the trivial upper bound of this problem is EXPSPACE. Mulmuley [Mul17, Mul12] asked whether this EXPSPACE bound can be improved. This was recently shown to be in PSPACE, over the fields of real and complex numbers, by Forbes and Shpilka [FS18]. Their proof technique is analytic (uses the existential theory of reals) and does not apply to finite fields. In [GSS18], we give a

PSPACE construction over arbitrary fields.

1.3.1 Our results

In Chapter 6, we show how to reduce the hitting set construction problem to a new problem: approximate polynomials satisfiability. See Theorem 6.2.2 in Chapter 6.

A polynomial system with no solution may have an *approximate* solution in the following sense. The system $x = xy - 1 = 0$ has no solution. However, it has an approximate solution $\{x = \epsilon, y = 1/\epsilon\}$, in the sense that if $\epsilon \to 0$, both the polynomials $\to 0$.

This motivates the following problem, where we check whether there exist Laurent polynomials (polynomials in ϵ and ϵ^{-1}) that can be plugged in the given polynomials so that the evaluated polynomials are in the ideal generated by ϵ (polynomials in ϵ with constant term zero).

Problem (Approximate polynomials satisfiability (APS))**.** *Given $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$, determine if $f_1 = \ldots = f_m = 0$ have a common approximate solution, i.e., $\exists x_1, \ldots, x_n \in \overline{\mathbb{F}}[\epsilon, \epsilon^{-1}]$ such that $f_i(x_1, \ldots, x_n) \in \epsilon\overline{\mathbb{F}}[\epsilon]$ for $i = 1, \ldots, m$.*

Note that the *exact* version of APS is the well-known problem of polynomial system satisfiability— Does there exist $\beta \in \overline{\mathbb{F}}^n$ such that for all i, $f_i(\beta) = 0$? Hilbert's Nullstellensatz (HN) says that there is no solution to this system iff 1 is in the ideal of $f_1, \ldots, f_n$. If the system $\mathbf{f} = \{f_1, \ldots, f_m\}$ has an exact solution, then it is trivially in APS. But the converse is not true. As we have seen, $\{x, xy - 1\}$ is in APS, but there is no exact solution in $\overline{\mathbb{F}}$. Also, the instance $\{x, x + 1\}$ is neither in APS nor has an exact solution.

Using classical ideas from algebraic geometry, it can be proved that APS is equivalent to the following problem.

Problem (AnnAtZero)**.** *Test if the constant term of every annihilator, of a set of polynomials (computed by algebraic circuits) $\mathbf{f} = \{f_1, \ldots, f_m\}$, is zero.*

AnnAtZero is known to be NP-hard [Kay09]. The NP-hardness of APS follows from its equivalence with AnnAtZero. In [GSS18], we give a PSPACE algorithm for AnnAtZero, thereby also solving APS in PSPACE. If the ideal of the annihilating polynomials is

principal (generated by a single polynomial), AnnAtZero is in directly in PSPACE, as we need to check only one annihilating polynomial's constant term. The idea of annihilators may not be principal always. In [GSS18], we show AnnAtZero in PSPACE by giving a reduction to the principal ideal case. The proof of correctness of this reduction is presented in [GSS18].

Using the PSPACE algorithm for AnnAtZero, finally we get the following result.

Theorem (Theorem 6.2.4)**.** *There is a PSPACE algorithm that (given input n, s, r in unary & suitably large $\mathbb{F}_q$) outputs a set, of points from $\mathbb{F}_q^n$ of size $\mathsf{poly}(ns, \log qr)$, that hits all n-variate degree-r polynomials over $\overline{\mathbb{F}}_q$ that can be infinitesimally approximated by size s circuits.*

1.4 Organization of the thesis

In Chapter 2, we present some basic concepts used in the thesis. In Chapter 3, we give a generalized Jacobian criterion for testing algebraic dependence over finite fields. In Chapter 4, we show algebraic dependence testing over finite fields is in AM $\cap$ coAM.

In Chapter 5, we present our results on factor conjecture. In Chapter 6, we present a reduction of hitting set construction for polynomials with small approximative circuits to the problem approximate polynomials satisfiability.

Part I

Algebraic Dependence

Chapter 2

Preliminaries

In this chapter, we will present a few basic results on algebraic dependence that we use in this thesis. We also give some definitions and notations.

2.1 Notations

A few notations that we use in all the chapters are the following. We use the standard notations for the field of reals ($\mathbb{R}$), the field of complex numbers ($\mathbb{C}$), the set of integers ($\mathbb{Z}$) and the set of natural numbers ($\mathbb{N}$). A list of other notations is as followed.

- $\boldsymbol{x}$ denotes the variables $x_1, \ldots, x_n$.
- $\boldsymbol{x}^{\boldsymbol{a}}$ denotes the monomial $\prod_{i=1}^{n} x_i^{a_i}$.
- $\mathbf{f}(\mathbf{x})$ denotes the polynomials $f_1, \ldots, f_n \in \mathbb{F}[x_1, \ldots, x_n]$.
- $[n]$ denotes the set $\{1, 2, \ldots, n\}$.
- $\mathbb{F}$ is an arbitrary field. $\mathbb{A} := \overline{\mathbb{F}}$ is its algebraic closure.
- $\mathbb{F}_q$ is a finite field of size q and characteristic $p \geq 2$.
- F/K denote a field extension, where F is the bigger field.
- $\mathbb{F}[x_1, \ldots, x_n]$ is the polynomial ring in $x_1, \ldots, x_n$ over $\mathbb{F}$.

- $\mathbb{F}[f_1, \ldots, f_n]$ is the subring of $\mathbb{F}[x_1, \ldots, x_n]$ generated by $f_1, \ldots, f_n$.
- $\mathbb{F}[[x_1, \ldots, x_n]]$ denotes the formal power series ring over field $\mathbb{F}$.
- $\mathbb{F}(x_1, \ldots, x_n)$ denote the field of rational functions in n variables over the field $\mathbb{F}$.
- A polynomial of degree d can be decomposed into its homogeneous components, $f(\boldsymbol{x}) = f_0(\boldsymbol{x}) + f_1(\boldsymbol{x}) + \cdots + f_d(\boldsymbol{x})$, where all monomials in $f_i(\boldsymbol{x})$ has degree i. A polynomial is called homogeneous if all its monomials have same degree.
- If $f(\boldsymbol{x})$ is a polynomial or power series, $f(\boldsymbol{x}) \bmod \langle \boldsymbol{x} \rangle^t$ means $f(\boldsymbol{x})$ up to degree $(t-1)$ part. If the homogeneous decomposition of $f(\boldsymbol{x})$ is $f_0(\boldsymbol{x}) + f_1(\boldsymbol{x}) + \cdots + f_d(\boldsymbol{x})$, then $f(\boldsymbol{x}) \bmod \langle \boldsymbol{x} \rangle^t = f_0(\boldsymbol{x}) + f_1(\boldsymbol{x}) + \cdots + f_{t-1}(\boldsymbol{x})$.

2.2 Basics of Algebra

We use the Schwartz-Zippel-DeMillo-Lipton lemma [Sch80, Zip79, DL78] several times in this thesis.

Lemma 2.2.1 (Schwartz-Zippel-DeMillo-Lipton lemma)**.** *Let $P(x_1, \ldots, x_n)$ be a nonzero polynomial of degree d over a field $\mathbb{F}$. Let S be a finite subset of $\mathbb{F}$. If $(\alpha_1, \ldots, \alpha_n)$ is picked randomly from S^n,* $\Pr[P(\alpha_1, \ldots, \alpha_n) = 0] \leq \frac{d}{|S|}$.

This lemma has a simple proof using induction (on number of variables). The base case follows from the fact that a univariate polynomial of degree d has at most d roots. For the details, see [AB09, Lemma A.36].

2.2.1 Formal Power Series

The ring of formal power series in $\boldsymbol{x}$ contains expressions of the form $\sum_{\boldsymbol{a} \in \mathbb{N}^n} \alpha_{\boldsymbol{a}} \boldsymbol{x}^{\boldsymbol{a}}$. Infinitely many $\boldsymbol{a}$ (of different degree) with nonzero coefficients are allowed. Addition and multiplication are defined like polynomials.

In the ring of formal power series, $1 - x$ can be inverted ($\sum_{i \geq 0} x^i$ is the inverse). The following lemma gives the criterion for invertibility.

Lemma 2.2.2 (Existence of inverse)**.** *A power series over a ring is invertible iff its constant term is invertible in that ring.*

Proof. Suppose a power series f is invertible and g is its inverse. The product of the constant terms of f and g has to be 1. Thus, the constant term of f is invertible.

Conversely, suppose the constant term of a power series f is invertible. Assume wlog that the constant term of f is 1. Now, $1 - f$ is constant free and $\sum_{i \geq 0}(1-f)^i$ is a formal power series. It is easy to see that $(1-(1-f))^{-1} = \sum_{i \geq 0}(1-f)^i$.

□

2.2.2 Basics of finite fields

Finite field $\mathbb{F}_p$ can be modeled as a set of numbers $0, 1, \ldots, p-1$, on which addition, subtraction, multiplication and division are defined. Addition and multiplication are done modulo p, where p is a prime number. It can be seen that modulo a prime p, any number from the set $\{1, \ldots, p-1\}$, has a unique multiplicative inverse, thus division is well-defined. For the formal definition and the axioms of the structure of field, see [LN94].

The characteristic of a field is the smallest positive number (if it exists) n, such that $\underbrace{1+1+\ldots+1}_{n\text{-times}} = 0$, where 1 is the multiplicative identity and 0 is the additive identity of the field. If no such positive number exists, for example in the fields of rationals, reals or complex numbers, the characteristic is defined to be zero. The characteristic of $\mathbb{F}_p$ is p. There are infinitely many finite fields of characteristic p. For any positive integer n, there is a field of size p^n, denoted as $\mathbb{F}_q$ where $q = p^n$ that contains the field $\mathbb{F}_p$ as a subfield. The algebraic closure of $\mathbb{F}_p$ is an infinite field of characteristic p. In characteristic zero, any derivative of a nonconstant polynomial is always nonzero. But in characteristic p, derivative of x^p is zero (as px^{p-1} is zero modulo p). In characteristic p, the binomial theorem takes a simpler form: $(x+y)^p = x^p + y^p$. For more details on finite fields, see the book [LN94].

2.3 Basic properties of algebraic dependence

Here we give a few propositions we use in our results. For more on algebraic dependence, see [Mit13].

For the sake of testing algebraic independence of a given set of polynomials, we can assume without loss of generality that the number of variables is equal to the number of polynomials, using Lemma 2.3.1 and Lemma 2.3.4. The following lemma is a well-known basic fact in algebra.

Lemma 2.3.1 (Extra polynomials)**.** *If $m > n$ then any $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ are algebraically dependent.*

This can be proved using the matroid structure of algebraic dependence. See [ER93] for a proof that algebraic dependence satisfies the matroid properties. In linear algebra, any two basis of a finite dimensional vector space have same cardinality. Similarly it can be proved, any two transcendence bases of a set of polynomials have same cardinality. Clearly $\{x_1, \ldots, x_n\}$ is a transcendence basis of $\mathbb{F}[x_1, \ldots, x_n]$. If $n+1$ polynomials from $\mathbb{F}[x_1, \ldots, x_n]$ are algebraically independent, they would be part of a transcendence basis of cardinality more than n. That would lead to contradiction.

For a different proof of this fact using a dimension counting argument, see [DGW09, Theorem 3.3]. We just sketch the main idea of the proof. We want to show that there exists a nonzero annihilating polynomial $A(y_1, \ldots, y_{n+1})$ of total degree D such that $A(f_1, \ldots, f_{n+1}) = 0$. From the equation, $A(f_1, \ldots, f_{n+1}) = 0$, we get a system of homogeneous linear equations (with coefficients of A as unknown variables). If this system has a nontrivial solution, A is an annihilating polynomial. We can show that for a large enough D, the number of unknowns (degree of freedom) exceeds the number of equations (constraints).

If we count the number of unknowns, we get $\binom{D+n+1}{n+1}$. If we count the number of equations, we get an upper bound of $\binom{dD+n}{n}$, where d is the maximum degree of the given polynomials. It can be shown that if $D > (n+1)d^n$, $\binom{D+n+1}{n+1} > \binom{dD+n}{n}$.

This argument can be used to give the following degree bound ([DGW09, Theorem

3.3]) on an annihilating polynomial of a set of algebraically dependent polynomials.

Lemma 2.3.2 (Degree bound of annihilator)**.** *Let $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ be a set of algebraically dependent polynomials of degree at most d. Then, there is an annihilating polynomial of them of degree at most $(n+1)d^n$.*

The following degree bound is from [BMS13].

Lemma 2.3.3 (Degree bound [BMS13])**.** *Let $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ be a set of algebraically dependent polynomials of degree at most d. Assume the transcendence degree of $f_1, \ldots, f_m$ is r. Then, there is an annihilator of $f_1, \ldots, f_m$ of degree at most d^r.*

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 2.3.4 (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.*

For a proof, see [BMS13, Lemma 16].

Can it happen that a given set of polynomials are algebraically independent over some field but become dependent over some extension of the field? The answer is no. For testing algebraic independence over a field, it suffices to work over the algebraic closure.

Lemma 2.3.5 (Closed field)**.** *Consider polynomials $\mathbf{f}(\mathbf{x})$ over any field $\mathbb{F}$. Their transcendence degree remains invariant if we go 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 [Pło05] 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. □

2.4 Inseparability & separating transcendence basis

For this section, let $\mathbb{E} \supseteq \mathbb{F}$ be fields. The failure of the Jacobian criterion over finite fields can be explained using the fundamental concept of inseparability from Galois theory [Isa94].

Definition 2.4.1. *An $f \in \mathbb{F}[x]$ is* separable *if it has no multiple roots (i.e. all the roots are distinct) in its splitting field (where it completely factorizes into linear factors).*

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. A field extension $\mathbb{E}/\mathbb{F}$ is said to be *algebraic* if every element $\alpha \in \mathbb{E}$ is root of a univariate polynomial (called minimal polynomial) over $\mathbb{F}$. 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 the 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.2.3.1, 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 $\mathbf{f}$. Note that p^m will divide the degree of the extension $\mathbb{F}(\mathbf{x})/\mathbb{F}(\mathbf{f})$, and could be as large.

Corollaries 3.2.10 and 3.2.11 to our main theorems relate the failure of the Jacobian criterion to the inseparability of the extension $\mathbb{F}(\mathbf{x})/\mathbb{F}(\mathbf{f})$.

The notion of *perfect fields* is relevant to the above discussion. A field $\mathbb{F}$ is called perfect if every irreducible polynomial in $\mathbb{F}[x]$ is separable. If a field has characteristic zero, it is perfect. All the finite fields and algebraically closed fields are perfect. In characteristic p, a field is perfect iff every element of the field is a power of p. For an example of a field which is not perfect, take $\mathbb{F}_p(x)$. The polynomial $x^p - y \in \mathbb{F}_p(x)[y]$ is irreducible, but not separable.

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

Definition 2.4.2 (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 separating transcendence bases exist for the perfect fields.

Theorem 2.4.3. *Consider a finite set of polynomials $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$. If $\mathbb{F}$ is a finite field (or, 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 (or algebraically closed) field. [Kna07, Thm.7.20] shows that the extension $\mathbb{F}(\mathbf{f})/\mathbb{F}$ is separably generated. Furthermore, [Kna07, 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.

2.5 Proof of the Jacobian criterion

The central object related to the testing of algebraic independence of polynomials is the Jacobian matrix.

Definition 2.5.1 (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 [Jac41]. For a proof, see [BMS13].

Theorem 2.5.2 (Jacobian criterion for algebraic dependence)**.** *Let* $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$ *be a finite set of polynomials of degree at most* d*, and* $\mathrm{trdeg}_{\mathbb{F}} \mathbf{f} \leq r$*. If* $\mathrm{char}(\mathbb{F}) = 0$*, or* $\mathrm{char}(\mathbb{F}) > d^r$*, then* $\mathrm{trdeg}_{\mathbb{F}} \mathbf{f} = \mathrm{rank}_{\mathbb{F}(\mathbf{x})} \mathcal{J}_{\mathbf{x}}(\mathbf{f})$*.*

Here we give a proof for the case when the number of polynomials equals the number of variables. We use the differential operator $\mathcal{H}_1 f = (\partial f/\partial x_1) z_1 + \cdots + (\partial f/\partial x_n) z_n$. Note that $\mathcal{H}_1(fg) = f\mathcal{H}_1(g) + g\mathcal{H}_1 f$ (product rule). Let $\langle \mathcal{H}_1 f_1 \ldots, \mathcal{H}_1 f_n \rangle_{\mathbb{F}(\mathbf{x})}$ denote the $\mathbb{F}(\mathbf{x})$-linear span of the polynomials $\mathcal{H}_1 f_1 \ldots, \mathcal{H}_1 f_n$.

Theorem 2.5.3 (Jacobian rephrased)**.** *Let* $f_1, \ldots, f_n \in \mathbb{F}[x_1, \ldots, x_n]$ *be algebraically independent polynomials such that the extension* $\mathbb{F}(x_1, \ldots, x_n)/\mathbb{F}(f_1, \ldots, f_n)$ *is separable. Then* $\mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_n$ *are* $\mathbb{F}(\mathbf{x})$*-linearly independent. Conversely, if* $f_1, \ldots, f_n \in \mathbb{F}[x_1, \ldots, x_n]$ *are algebraically dependent, then* $\mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_n$ *are* $\mathbb{F}(\mathbf{x})$*-linearly dependent.*

Proof. Algebraic independence of $f_1, \ldots, f_n$ together with Lemma 2.3.1 asserts the existence of the minimal annihilating polynomial $A_j \in \mathbb{F}_p[y_0, y_1, \ldots, y_n]$ for the polynomials $x_j, f_1, \ldots, f_n$, for all $j \in [n]$. Now, the separability of the extension $\mathbb{F}(\mathbf{x})/\mathbb{F}(\mathbf{f})$ implies that $\partial A_j / \partial y_0 \neq 0$, for all $j \in [n]$. We start with the equation

$$A_1(x_1, f_1, \ldots, f_n) = \sum_{\mathbf{e}_\ell} \alpha_{\mathbf{e}_\ell} \cdot x_1^{e_{0\ell}} f_1^{e_{1\ell}} \cdots f_n^{e_{n\ell}} = 0.$$

We apply the operator $\mathcal{H}_1$ on the above equation, and use the product rule of $\mathcal{H}_1$ to get

$$\mathcal{H}_1(A_1(x_1, f_1, \ldots, f_n)) = \sum_{\mathbf{e}_\ell} \alpha_{\mathbf{e}_\ell} e_{0\ell} x_1^{e_{0\ell}-1} \cdot z_1 \cdot (f_1^{e_{1\ell}} \cdots f_n^{e_{n\ell}}) + \sum_{\mathbf{e}_\ell} \alpha_{\mathbf{e}_\ell} x_1^{e_{0\ell}} \mathcal{H}_1(f_1^{e_{1\ell}} \cdots f_n^{e_{n\ell}}) = 0.$$

Note that separability implies the presence of at least one $e_{0\ell}$ in A_1 which is not a multiple of p. So, we have at least one non-zero summand in the first sum. Also, $\sum_{\mathbf{e}_\ell} \alpha_{\mathbf{e}_\ell} e_{0\ell} x_1^{e_{0\ell}-1} \cdot (f_1^{e_{1\ell}} \cdots f_n^{e_{n\ell}})$ cannot be zero, else we get a lower degree annihilating polynomial of $x_1, \mathbf{f}$, which contradicts the minimality of A_1. Using this observation and the fact that the product rule on $\mathcal{H}_1(f_1^{e_{1\ell}} \cdots {f_n}^{e_{n\ell}})$ breaks it into an $\mathbb{F}(\mathbf{x})$-linear combination of $\mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_n$, we deduce that $z_1 \in \langle \mathcal{H}_1 f_1 \ldots, \mathcal{H}_1 f_n \rangle_{\mathbb{F}(\mathbf{x})}$.

Similar operation on $A_2, \ldots, A_n$ gives $z_2, \ldots, z_n \in \langle \mathcal{H}_1 f_1 \ldots, \mathcal{H}_1 f_n \rangle_{\mathbb{F}(\mathbf{x})}$. So, we have n linearly independent elements $z_1, \ldots, z_n$ in $\langle \mathcal{H}_1 f_1 \ldots, \mathcal{H}_1 f_n \rangle_{\mathbb{F}(\mathbf{x})}$. Thus, we get that $\text{rank}\langle \mathcal{H}_1 f_1 \ldots, \mathcal{H}_1 f_n \rangle_{\mathbb{F}(\mathbf{x})} \geq n$.

The converse can be proved similarly, applying $\mathcal{H}_1$ operator on the annihilator. Let $A(y_1, \ldots, y_n)$ be a polynomial of smallest degree such that $A(f_1, \ldots, f_n) = 0$. Using the product rule of $\mathcal{H}_1$, we get that $\sum_{i=1}^{n} \partial_i A(f_1, \ldots, f_n) \mathcal{H}_1 f_i = 0$. Over characteristic zero, if $\partial_i A(f_1, \ldots, f_n) = 0$, then there is an annihilator of smaller degree. Over the finite fields, if $\partial_i A(f_1, \ldots, f_n) = 0$ for all i, then A is a power of p. That contradicts the smallest degree assumption.

□

The failure of the Jacobian happens in inseparable cases.

Theorem 2.5.4 (Jacobian fails for inseparable)**.** *For algebraically independent polynomials $f_1, \ldots f_n \in \mathbb{F}_p[x_1, \ldots, x_n]$ with the extension $\mathbb{F}_p(\mathbf{x})/\mathbb{F}_p(\mathbf{f})$ being inseparable, we have $\mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_n$ $\mathbb{F}(\mathbf{x})$-linearly dependent.*

Proof. Algebraic independence of $f_1, \ldots, f_n$ asserts the existence of minimal annihilating polynomial $A_j \in \mathbb{F}_p[y_0, y_1, \ldots, y_n]$ for the polynomials $x_j, f_1, \ldots, f_n$, for all $j \in [n]$. Now the inseparability of the extension $\mathbb{F}_p(x_1, \ldots, x_n)/\mathbb{F}_p(f_1, \ldots, f_n)$ implies that there exists at least one j such that A_j lives in $\mathbb{F}_p[y_0^p, y_1, \ldots, y_n]$. Thus, we have

$$A_j(x_j, f_1, \ldots, f_n) = \sum_{\mathbf{e}_\ell} \alpha_{\mathbf{e}_\ell} \cdot (x_j^p)^{e_{0\ell}} f_1^{e_{1\ell}} \cdots f_n^{e_{n\ell}} = 0.$$

Next, we apply the operator $\mathcal{H}_1$ on the above equation, which gives

$$\sum_{\mathbf{e}_\ell} \alpha_{\mathbf{e}_\ell} \cdot (x_j^p)^{e_{0\ell}} \cdot \mathcal{H}_1(f_1^{e_{1\ell}} \cdots f_n^{e_{n\ell}}) = 0$$

as $\mathcal{H}_1(x_j^p)^{e_{0\ell}} = 0$, i.e. $\mathcal{H}_1$ treats p-powers as constants. Now, because of the product rule of $\mathcal{H}_1$, this gives an $\mathbb{F}_p(\mathbf{x})$-linear dependence of $\mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_n$. (Note that we also used the nontriviality of A_j to get this linear dependence.) Hence the Jacobian criterion fails! □

2.6 Taylor expansion and Hasse derivatives

We apply shift (or translation) of variables to our polynomials. We view this as writing the Taylor expansion of a polynomial $f(\mathbf{x})$ at a "generic" point $\mathbf{z}$ [For14, Sec.C.1]. A second view is that of computing the Hasse-Schmidt higher derivatives of f at the point $\mathbf{z}$ [For15, DKSS13]. A third view is seeing the shifted polynomial as a Hasse-Schmidt differential [Tra98]. We collect these equivalent viewpoints in a single definition.

Definition 2.6.1 (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 $\frac{1}{\ell_1! \cdots \ell_n!} \frac{\partial^{(\ell_1+\cdots+\ell_n)} f}{\partial x_1^{\ell_1} \cdots \partial x_n^{\ell_n}}(\mathbf{z})$.

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 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 $\langle \mathbf{x} \rangle_R^{t+1}$.

2.7 Leading monomials and algebraic independence

A monomial/term ordering σ is a total order on the monomials in $\mathbb{F}[x_1, \ldots, x_n]$, that satisfies the property: if $m_1 \preceq_\sigma m_2$, then for any other monomial m_3, $m_1 m_3 \preceq_\sigma m_2 m_3$. The leading monomial/term of a polynomial wrt σ is the monomial that is greatest wrt the order σ. Graded lexicographic monomial ordering is an monomial ordering such that if two monomials have different degree, the one with smaller degree is considered to be smaller and in case two monomials have same degree, lexicographic order is followed.

The following standard lemma give a sufficient condition for algebraic independence via leading monomials.

Lemma 2.7.1. *[KR05, 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 [Kem10, 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 ,

$$\begin{aligned} m'(k_1, \ldots, k_n) &\preceq_\sigma m'(LM(f_1), \ldots, LM(f_n)) \\ &\preceq_\sigma m(LM(f_1), \ldots, LM(f_n)) \,. \end{aligned}$$

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 $f_1, \ldots, f_n$. □

2.8 Basic results in algebraic complexity

We use the following standard result to truncate a polynomial given by a circuit up to some degree.

Lemma 2.8.1 (Homogenization [Str73])**.** *If f is a polynomial given by a size s circuit, all the homogeneous components of f up to degree r, can be computed by a size $O(r^2 s)$ circuit.*

The idea is to create a new circuit where we keep track of all the homogeneous components separately and simulate the original circuit's additions and multiplications. For a proof, see [SY10, Thm.2.2].

We use the following result from [Kal87, Thm.1] to compute k-th derivative of an arithmetic circuit.

Lemma 2.8.2 (k-th derivative computation)**.** *Let a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ be given by an arithmetic circuit of size s. Then, for any $i \in [n]$, $\frac{\partial^k f}{\partial {x_i}^k}$ can be computed by a circuit of size $k^2 s$.*

The key idea here is to use Leibniz product rule of k-th order derivative inductively, keeping track of all the derivatives up to order k of the polynomials computed at each gate.

We use the following classical result to compute the power series approximation of a rational function and to eliminate division gates from an arithmetic circuit.

Lemma 2.8.3 (Eliminating division [Str73])**.** *Let $f, g \in \mathbb{F}[\boldsymbol{x}]$ be two degree-d polynomials, each computed by a circuit of size-s with $g(\mathbf{0}) \neq 0$. Then $f/g \bmod \langle \boldsymbol{x} \rangle^{d+1}$ can be computed by $\mathsf{poly}(s, d)$ size circuit.*

For a complete proof, see [SY10, Thm.2.11]. Here we give a proof sketch.

Proof. By appropriate normalization, assume wlog that $g(\mathbf{0}) = 1$. As $1 - g$ is constant free, we have the following power series identity in $\mathbb{F}[[\boldsymbol{x}]]$:

$$f/g = f/(1 - (1 - g)) = f + f(1 - g) + f(1 - g)^2 + f(1 - g)^3 + \cdots .$$

For all $d \geq 0$, taking modulo $\langle \boldsymbol{x} \rangle^{d+1}$ in both sides, the left hand side equals the right hand side of the above identity.

Now, if we want to compute $f/g \bmod \langle \boldsymbol{x} \rangle^{d+1}$, we compute the circuit of $f + f(1-g) + f(1-g)^2 + \cdots + f(1-g)^d$ and finally truncate up to degree d using homogenization (use Lemma 2.8.1). It is easy to see that the size of the final circuit is $\mathrm{poly}(s, d)$.

□

Remark. It may happen that $g(\mathbf{0}) = 0$, thus $1/g$ does not exist in $\mathbb{F}[[\boldsymbol{x}]]$, yet f/g exists and is a polynomial of degree d. In that case, we use the following normalization trick. We shift the polynomials f, g by a random point $\boldsymbol{\alpha} \in \mathbb{F}^n$. Using Schwartz-Zippel-DeMillo-Lipton lemma, the constant term of $g(\boldsymbol{x} + \boldsymbol{\alpha})$ is non-zero with high probability. Now, we can compute $f(\boldsymbol{x} + \boldsymbol{\alpha})/g(\boldsymbol{x} + \boldsymbol{\alpha})$ using Lemma 2.8.3. Finally, we compute the polynomial f/g by applying the inverse of the shift $\boldsymbol{x} \mapsto \boldsymbol{x} - \boldsymbol{\alpha}$.

Lemma 2.8.4 (Circuit division elimination)**.** *Let f be a polynomial of degree d computed by a circuit (with division gates) of size s. Then, f can be computed by $\mathsf{poly}(sd)$ size circuit.*

Here we just give the basic idea. See [SY10, Thm.2.12] for a full proof.

Proof idea. If there are many division gates in the circuit, we preprocess the circuit in the following way. We create a new circuit, separately keeping track of the numerator and the denominator computed at each gate. Now, in the new circuit, we simulate the addition, multiplication and division gates of the original circuit, by appropriately adding edges. This pre-processing incurs poly(sd) blow up of size. Finally, to remove the single division gate at the top, we use Lemma 2.8.3.

□

2.8.1 Approximative complexity and Closure of VP

In computer science, the notion of approximative complexity emerged in the context of tensors for matrix multiplication (the notion of border rank, see [LL89, BCS13] and references therein). [Bür04] used this concept in the context of arithmetic circuits. Approximative closure of algebraic complexity classes is of great interest in the geometric complexity theory program (see [GMQ16]).

The notion of border complexity or approximative complexity can be defined in many ways. Over fields like $\mathbb{R}$ or $\mathbb{C}$, it can be defined using *metric topology*. A polynomial $f(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]$ has approximative/border circuit complexity at most s if there is a sequence of polynomials $\{f_n(\boldsymbol{x})\}_n$ such that the sequence *converge* to f (coefficient-wise) in the limit and for all n, $f_n(\boldsymbol{x})$ can be computed by an arithmetic circuit of size at most s. This gives a definition of closure (under Euclidean metric topology) of VP. A polynomial family $\{f_n(\boldsymbol{x})\}_n$ is in $\overline{\text{VP}}$ if and only if $\exists f_{n_i}(\boldsymbol{x})$ such that $\lim_{i\to\infty} f_{n_i} = f_n$ and $\{f_{n_i}(\boldsymbol{x})\}_n$ is in VP. Thus, $\overline{\text{VP}}$ can be stated as the class containing all the limits of the polynomial families in the VP.

The closure can be also taken under Zariski topology. Recall that a subset of $\mathbb{C}^n$ is Zariski closed if it is a zero set of a set of polynomials. The algebraic closure of a field contains all the roots of all the univariate polynomials over the field. Similarly, if a given set of points S vanishes on a given set of polynomials P, if we add all the common roots of the polynomials in P to the set S, we get a closed set $\overline{S}$, which is defined to be the closure of

S. A n-variate polynomial of degree d can be seen as a coefficient vector containing $\binom{n+d}{d}$ many coefficients. If we take the closure of the set containing all the coefficient vectors of the polynomials in VP, we get the class $\overline{\text{VP}}$. Note that over $\mathbb{C}$, if a set is Zariski-closed, it is also closed under Euclidean topology and the two definitions of $\overline{\text{VP}}$ coincide [Bür04].

The following definition (equivalent with the previous definitions) of approximation makes sense over any field and can be used to define closure of VP. Here, ϵ is a formal variable and $\mathbb{F}(\epsilon)$ is the rational function field of ϵ. For an algebraic complexity class C, the approximative closure of C is defined as follows [BIZ18, Defn.2.1].

Definition 2.8.5 (Approximative closure of a class [BIZ18])**.** *Let C be an algebraic complexity class over field $\mathbb{F}$. A family $\{f_n\}$ of polynomials from $\mathbb{F}[\boldsymbol{x}]$ is in the* class $\overline{C}(\mathbb{F})$ *if there are polynomials $f_{n;i}$ and a function $t : \mathbb{N} \mapsto \mathbb{N}$ such that g_n is in the class C over the field $\mathbb{F}(\epsilon)$ with $g_n(\boldsymbol{x}) = f_n(\boldsymbol{x}) + \epsilon f_{n;1}(\boldsymbol{x}) + \epsilon^2 f_{n;2}(x) + \ldots + \epsilon^{t(n)} f_{n;t(n)}(\boldsymbol{x})$.*

Whether VP is a proper subset of $\overline{\text{VP}}$ is a big open question. See [GMQ16] for more on this question.

2.9 Basic definitions from algebraic geometry

We work with commutative rings in this thesis. A subset I of a ring R is an *ideal* if it forms a subgroup of the additive group of ring R (thus, it is closed under addition and subtraction) and for every element b in R and $a \in I$, their product ab belong to I. An ideal I is generated by the elements $f_1, \ldots, f_m$ if and only if every element $f \in I$ can be expressed as $f = \sum_{i=1}^{m} f_i g_i$, where $g_i \in R$.

An ideal generated by a single element of the corresponding ring is called a *principal* ideal. An ideal of R that is not a proper subset of any ideal other than R itself, is called a *maximal* ideal. An ideal P of ring R is called a *prime ideal* if $\forall a, b \in R$, $ab \in P \implies a \in P$ or $b \in P$, and P is not equal to R. The *radical* of an ideal is an ideal such that an element x is in the radical iff x^k (for some positive integer k) is in the ideal. A radical ideal is an ideal which is equal to its radical.

Let $\mathbb{A} := \overline{\mathbb{F}}$ be the algebraic closure of a field $\mathbb{F}$. For $d \in \mathbb{N}^+$, write $\mathbb{A}^d$ for the *d-dimensional affine space* over $\mathbb{A}$. It is defined to be the set $\mathbb{A}^d$, equipped with the *Zariski topology*, defined as follows: A subset S of $\mathbb{A}^d$ is *closed* iff it is the set of common zeros of some subset of polynomials in $\mathbb{A}[X_1, \ldots, X_d]$. For other subsets S it makes sense to consider the *closure* $\overline{S}$— the smallest closed set containing S. Set S is *dense* in $\mathbb{A}^d$ if $\overline{S} = \mathbb{A}^d$. Complement of closed sets are called *open*.

A closed set is called a *hypersurface* (resp. *hyperplane*) if it is definable by a single polynomial (resp. single linear polynomial).

Closed subsets of $\mathbb{A}^d$ are also called *algebraic sets* or *zero sets*. An algebraic set is *irreducible* if it cannot be written as the union of finitely many proper algebraic sets. An irreducible algebraic subset of an affine space is also called an *affine variety*.

An algebraic set V can be uniquely represented as the union of finitely many varieties, and these varieties are called the *irreducible components* of V.

The ideal-variety correspondence connects algebra and geometry. Given a variety V, $I(V)$ denotes the ideal of the polynomials that vanish on the points in the variety. Given an ideal I, $V(I)$ denote the variety that contains zeroes of the polynomials from I. Affine zero sets (resp. varieties) are in 1-1 correspondence with *radical* (resp. *prime*) ideals. If $I_1 \subseteq I_2$ are ideals, then variety V_1 corresponding to I_1 contains the variety V_2 corresponding to I_2. If variety V_1 is contained in variety V_2, $I(V_2) \subseteq I(V_1)$.

The *dimension* of a variety V is defined to be the largest integer m such that there exists a chain of varieties $\emptyset \subsetneq V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_m = V$. More generally, the dimension of an algebraic set V, denoted by $\dim V$, is the maximal dimension of its irreducible components. For example, we have $\dim \mathbb{A}^d = d$. The dimension of the empty set is -1 by convention. One dimensional varieties are called *curves*.

The *degree* of a variety V in $\mathbb{A}^d$ is the number of intersections of V with a general/random affine subspace of dimension $d - \dim V$. More generally, we define the degree of an algebraic set V, denoted by $\deg(V)$, to be the sum of the degrees of its irreducible components.

2.10 Basics of complexity theory

For the formal definitions of the complexity classes mentioned in this thesis (NC, P, NP, RP, AM, PH, PSPACE), see [AB09]. Here we give a brief overview of the class AM.

2.10.1 Complexity class AM

Arthur-Merlin protocols, introduced by Babai [Bab85], is a special type of interactive proof system in which Arthur is the randomized poly-time verifier and Merlin is the all-powerful prover. They have only constantly many rounds of exchange. Arthur first sends the input string x and results of random coin tosses (a random string y) to Merlin. After Merlin sends back a string z which is a proof/witness corresponding to (x, y), Arthur verifies it (decides whether to accept x or not) in polynomial time. AM is the class of decision problems for which the yes answer can be verified by such a protocol. For the input strings in the language (YES instances), Merlin should be able to send a proof such that Arthur accepts with probability (over the random choices of Arthur) at least 2/3 and if the input string is not in the language (NO instances), Arthur rejects with probability (over the random choices of Arthur) at least 2/3, irrespective of Merlin's message. The class AM contains interesting problems like verifying if two graphs are non-isomorphic. A language is in coAM iff its complement language is in AM. If a language is in AM $\cap$ coAM, it can not be NP or coNP hard, otherwise the polynomial hierarchy collapses to the second level [Sch88]. See [Sax06, Proposition 2.5] for a proof of this. See [KS06] for a few natural algebraic problems in AM $\cap$ coAM.

In the Chapter 4, an AM protocol will be used to distinguish whether a set S is 'small' or 'large'. This is done using the Goldwasser-Sipser [GS86] set lowerbound method:

Lemma 2.10.1 (Goldwasser-Sipser [GS86])**.** *Let $m \in \mathbb{N}$ be given in binary. Suppose S is a set whose membership can be tested in nondeterministic polynomial time and its size is promised to be either $\leq m$ or $\geq 2m$. Then, the problem of deciding whether $|S| \stackrel{?}{\geq} 2m$ is in AM.*

See, [AB09, Chapter 8] or [Sax06, Proposition 2.4] for a complete proof.

Chapter 3

Algebraic Dependence and Functional Dependence

Abstract

This chapter is based on joint work [PSS18] with Pandey and Saxena.

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 [KS17] and proved for zero or large characteristics. We show that functional dependence holds for *arbitrary* fields, thereby answering the open questions in [KS17]. We also show that functional dependence is equivalent to algebraic dependence. A consequence of our characterization of algebraic dependence 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. This can be seen as a natural generalization of the classical Jacobian criterion.

3.1 Introduction

In this chapter, 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.

Algebraic dependence to approximate functional dependence. We show that over arbitrary fields, algebraic dependence of polynomials $f_1, \ldots, f_m$ implies 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 (Theorem 3.2.3). 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 [KS17, Lem.3.1] which asserted the same over fields of zero (or large) characteristic.

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 (taking x_1 and $x_1x_2^2$) 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 (Theorem 3.2.7). This parameter is actually the *inseparable degree* of an appropriate field extension, which is a well studied concept in Galois theory (Sec.2.4).

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 seen 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. We show that *separable extension* is precisely the case when the Jacobian works (Corollary 3.2.10). 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.2.6), modulo a carefully chosen subspace $\mathcal{U}_t$ (Remark 3.2.2). 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 $\mathcal{U}_t$. We give a possible interpretation of the main criterion using Hasse-Schmidt differentials and matrices in Sec.3.4. We illustrate the overall idea, and its comparison with Jacobian criterion, in Fig.3.1.

The algorithm we get using our criterion is efficient only under the promise that the inseparability degree is bounded by a small constant. Note that there are situations where the inseparability degree is quite small compared to the product of the degrees of the input polynomials. Let m_1 and m_2 be integers such that m_1m_2 is coprime to p, and let $f_1 = x_1^{pm_1}, f_2 = x_2^{m_2}$. 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. 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 exponentially large parameter (in terms of input size).

Analytic Dependence and Algebraic Dependence. Functional dependence can also be stated using formal power series, relating to the notion of analytic dependence [Abh56,

AM76]. A polynomial f is *analytically dependent* on polynomials $g_1, \ldots, g_n$ if there exists a formal power series $A \in \mathbb{F}[[x_1, \ldots, x_n]]$ such that $f = A(g_1, \ldots, g_n)$. Without loss of generality, we assume that the constant terms of the given polynomials are zero, so that the above equality is well defined in $\mathbb{F}[[x_1, \ldots, x_n]]$. Our main theorems (Theorem. 3.2.3 and Theorem. 3.2.7) show that algebraic dependence and analytic dependence are equivalent for polynomials over arbitrary fields.

3.2 Main structure theorems

We use the following standard notation:

1. 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 $R = S$.)

2. 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. Note that by $h(g_1(\mathbf{x}), \ldots, g_n(\mathbf{x}))^{\leq d}$ we would mean that: first compute the composition $h(\mathbf{g}(\mathbf{x}))$ and then extract out the degree$\leq d$ part.

3. 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.

4. If $\mathcal{I}$ is an ideal of a ring, the t-th power of it, denoted as $\mathcal{I}^t$, is generated by all the products of any t (may not be distinct) elements from $\mathcal{I}$.

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

In the following section, we prove a few technical lemmas that we use in the proofs of our main theorems. One may want to skip some of the details in the first reading and move to Section 3.2.2.

3.2.1 Technical lemmas

We will use $\mathbf{z}$ as a formal variable (n-tuple) and can fix it later to a suitable constant $\mathbf{a}$. We consider the ring $R := \overline{\mathbb{F}}(\mathbf{z})[\mathbf{x}]$ and its ideal $\mathcal{I}_0 := \langle \mathbf{x} \rangle_R$. This ideal contains all the constant-free linear polynomials (linear forms) in $x_1, \ldots, x_n$. 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, $f(\mathbf{x}+\mathbf{z}) = f(\mathbf{z}) + \mathcal{H}_t f(\mathbf{x})$).

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

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

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 3.2.1 (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 $\mathcal{Q}_{t+i-1}$. □

Lemma 3.2.1 tells us that due to the filtration (filtering out the higher degree terms) in $\mathcal{Q}_t$, some of these terms will be equivalent to terms involving $\mathcal{H}_r$ with $r < t$.

Remark 3.2.2. *In $\mathcal{Q}_t$, this is the same subspace as $\langle \mathcal{H}_t f_i, \mathcal{H}_t g_1, \ldots, \mathcal{H}_t g_k \rangle^2_{\mathbb{F}(\mathbf{z})} + \cdots + \langle \mathcal{H}_t f_i, \mathcal{H}_t g_1, \ldots, \mathcal{H}_t g_k \rangle^t_{\mathbb{F}(\mathbf{z})}$ by Lemma 3.2.1.*

3.2.2 Functional dependence for algebraically dependent polynomials

We show that algebraic dependence implies functional dependence over arbitrary fields (to arbitrary degree of approximation t).

Theorem 3.2.3 (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, $\forall i \in [m]$, $f_i(\mathbf{x}+\mathbf{a})^{\leq t} = h_i(g_1(\mathbf{x}+\mathbf{a}), \ldots, g_k(\mathbf{x}+\mathbf{a}))^{\leq t}$.*

Remark 3.2.4. *Clearly, $\overline{\mathbb{F}}^n$ is an infinite space. What we mean here by a* random **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 section merely existence of* **a** *is needed. Section 3.3.1 will discuss the estimate.*

Pf. of Theorem 3.2.3. 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 Theorem 2.4.3 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

$$\begin{aligned}
&\mathcal{H}_t g_0/g_0(\mathbf{z}) \cdot \left(\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} e_{0\ell} g_0^{e_{0\ell}} \cdots g_k^{e_{k\ell}} \right) \\
&+ \sum_{j \in [k]} \mathcal{H}_t g_j/g_j(\mathbf{z}) \cdot \left(\sum_{\mathbf{e}_\ell} a_{\mathbf{e}_\ell} e_{j\ell} g_0^{e_{0\ell}} \cdots g_k^{e_{k\ell}} \right) \\
&+ (\text{higher terms with } \sum_j s_j \geq 2) \equiv 0 \,. \qquad (3.1)
\end{aligned}$$

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.3.1. 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 $\mathcal{U}_t$ (Remark 3.2.2) 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 our subspace reduction lemma (Lemma 3.2.5) which gives 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 Remark 3.2.2) 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 prove the subspace reduction lemma, that essentially shows that if $\mathcal{H}_r f_n$

depends on higher order terms (in the sense of Eqn.3.1) then it can be "dropped" from the ideal manipulations.

Lemma 3.2.5 (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". In general, one can think of the lemma as replacing $\mathcal{H}_t f_n$ by u everywhere.

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 (Lemma 3.2.1 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 Lemma 3.2.6) 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 (Lemma 3.2.1, with $t = r$ and $i = \ell + 1 - r$) to get, in $\mathcal{Q}_\ell$,

$$\langle \mathcal{H}_r f_1, \ldots \mathcal{H}_r f_n \rangle^{\ell+1-r}_{\mathbb{F}(\mathbf{z})} \subseteq \langle v_1^{q_1} \cdots v_r^{q_r} \mid \sum_{j \in [r]} q_j = \ell + 1 - r \, , \, q_j \geq 0, \mathbf{v} \rangle_{\mathbb{F}(\mathbf{z})} \tag{3.2}$$

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 subspaces $\langle \mathcal{H}_r f_1, \ldots \mathcal{H}_r f_n \rangle^s_{\mathbb{F}(\mathbf{z})}$ in the RHS of Eqn.3.2. This allows us to rewrite Eqn.3.2 as (recall Remark 3.2.2),

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

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. □

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 3.2.6 (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 $\langle \mathbf{x} \rangle_R^{r+1}$. □

3.2.3 Algebraic independence: Criterion

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

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 3.2.7 (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 $\mathbf{a} \in \overline{\mathbb{F}}_q^n$, $f_n(\mathbf{x}+\mathbf{a})^{\leq t}$ cannot be written as $h(f_1(\mathbf{x}+\mathbf{a}), \ldots, f_{n-1}(\mathbf{x}+\mathbf{a}))^{\leq t}$, 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 $\mathbf{a} \in \overline{\mathbb{F}}_q^n$, $f_j(\mathbf{x}+\mathbf{a})^{\leq t}$ can be written as $h_{jt}(f_1(\mathbf{x}+\mathbf{a}), \ldots, f_{j-1}(\mathbf{x}+\mathbf{a}), f_{j+1}(\mathbf{x}+\mathbf{a}), \ldots, f_n(\mathbf{x}+\mathbf{a}))^{\leq t}$, 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 the case of prime characteristic we get the above statement, while in the zero characteristic case one should set the inseparable degree $= 1$ to read the above statement. The meaning of 'random $\mathbf{a}$' was explained in Remark 3.2.4.

Proof idea- By the hypothesis we have that each monomial $x_j^{p^i}$, $j \in [n]$, 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 Theorem 3.2.3 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 we give an argument using monomial ordering that there exists no functional dependence among $\mathbf{f}(\mathbf{x}+\mathbf{z})$ (Lemma 3.2.8).

Pf. of Theorem 3.2.7. [$t < p^i$ *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 Theorem 3.2.3.

The descent principle (Lemma 3.2.6) 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 Lemma 2.3.1). 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.3.1 (Sec.3.2.2) to get that for some j', $f_{j'}(\mathbf{x}+\mathbf{z})^{\leq p^i-1} = h_{j'}(f_1(\mathbf{x}+\mathbf{z}), \ldots, f_{j'-1}(\mathbf{x}+\mathbf{z}), f_{j'+1}(\mathbf{x}+\mathbf{z}), \ldots, f_n(\mathbf{x}+\mathbf{z}))^{\leq p^i-1}$ 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 Lemma 3.2.6 implies that proving the theorem for $t = p^i$ suffices. For contradiction, assume that (wlog) $f_n(\mathbf{x}+\mathbf{z})^{\leq p^i}$ can be written as $h(f_1(\mathbf{x}+\mathbf{z}), \ldots, f_{n-1}(\mathbf{x}+\mathbf{z}))^{\leq p^i}$ 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.2.4) 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

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

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 Remark 3.2.2, 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 (Lemma 3.2.5), 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) Lemma 3.2.8, and hence finishes the proof. (One can easily see that we get functional independence for random fixing of $\mathbf{z}$ in the space $\overline{\mathbb{F}}_q^n$.) □

Now we show that n 'pure' monomials cannot functionally depend on $< n$ polynomials. This is at the heart of our criterion.

Lemma 3.2.8 (Impossible containment)**.** *Let $\mathbb{F}$ be any field. Consider the subspace $\mathcal{V}'_t := \langle \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1} \rangle_{\mathbb{F}(\mathbf{z})} + \ldots + \langle \mathcal{H}_1 f_1, \ldots, \mathcal{H}_1 f_{n-1} \rangle^t_{\mathbb{F}(\mathbf{z})}$ of $\mathcal{Q}_t$, for $t \geq 1$. Then, $\{x_1^t, \ldots, x_n^t\} \not\subseteq \mathcal{V}'_t$.*

Proof. Remark 3.2.2 suggests that $\mathcal{V}'_t$ equals the subspace $\langle \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1} \rangle_{\mathbb{F}(\mathbf{z})} + \cdots + \langle \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1} \rangle^t_{\mathbb{F}(\mathbf{z})}$ 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:

$$x_i^t + \alpha_i \in \langle \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1} \rangle_{\mathbb{F}(\mathbf{z})} + \cdots + \langle \mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1} \rangle^t_{\mathbb{F}(\mathbf{z})} ,$$

for some $\alpha_i \in \langle \mathbf{x} \rangle^{t+1}_R$, for all $i \in [n]$. This simply means $x_i^t + \alpha_i = P_i(\mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1})$, 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 [CLO07, 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 (Lemma 2.7.1), we get that $\mathrm{trdeg}_{\mathbb{F}(\mathbf{z})}\{x_i^t + \alpha_i \mid i \in [n]\} = n$. On the other hand, clearly,

$$\mathrm{trdeg}_{\mathbb{F}(\mathbf{z})}\{P_i(\mathcal{H}_t f_1, \ldots, \mathcal{H}_t f_{n-1}) \mid i \in [n]\} \leq n-1.$$

This makes the containment impossible. □

Remark 3.2.9. *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 can see the classical Jacobian criterion as a special case of Theorems 3.2.3 and 3.2.7.

3.2.4 Recovering the classics

As a corollary of Theorem 3.2.3 and Theorem 3.2.7, we get the classical Jacobian criterion for the separable case (i.e. inseparable degree $= p^0 = 1$).

Corollary 3.2.10 (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 Theorem 3.2.7 helps us in characterizing the failure of the Jacobian.

Corollary 3.2.11 (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.*

Now we describe the algorithm to test algebraic independence using our criterion.

3.3 Application: Algebraic independence testing algorithm

An easy consequence of Theorem 3.2.3 and Theorem 3.2.7 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.

Theorem 3.3.1 (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) and* s *is the total input size.*

Algorithm idea: The criterion (by Theorems 3.2.3 & 3.2.7) 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.

3.3.1 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 [Mit13, 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 [AL86]. 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 $\mathbb{F}_q^n$ [Sch80, DL78, Zip79].

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.

3.3.2 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 [Str73, SY10, 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.

3.3.3 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.

Figure 3.1: Our criterion

	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 $\mathrm{poly}\binom{n+p^e}{n}$-time algorithm

3.3.4 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 $\mathrm{poly}(s, \binom{t+n}{n})$-time.

Thus, we have a $\mathrm{poly}(s, \binom{t+n}{n})$-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 Theorem 3.3.1.

3.4 Interpretation of the criterion via Hasse-Schmidt differential

Motivated by the classical Jacobian criterion, we try to generalize the criterion to positive characteristic. Jacobian criterion can be proved using the differential operator (denoted by H_1) defined as follows,

$$H_1 := \frac{\partial}{\partial x_1} z_1 + \cdots + \frac{\partial}{\partial x_n} z_n \,.$$

H_1 acts on $f \in \mathbb{F}[\mathbf{z}]$ as $H_1 f = (\partial f/\partial x_1) z_1 + \cdots + (\partial f/\partial x_n) z_n$.

The key idea of Jacobian criterion is that the differential operator can reduce the nonlinear problem of testing algebraic independence to linear algebra (testing linear independence over function fields) in zero, or large characteristic. Formally, polynomials $f_1, \ldots, f_n$ are algebraically dependent iff polynomials $H_1 f_1, \ldots H_1 f_n$ are linearly dependent over the function field $\mathbb{F}(x_1, \ldots, x_n)$. If we see the proof of Jacobian, we observe that four properties of differential operators (acting on some annihilating polynomial) crucially help to achieve this reduction: linearity, product rule (*Leibniz* rule: $H_1(fg) = fH_1g + gH_1f$), the fact that differentiating a polynomial reduces its degree, and differentiation does not make a nonconstant polynomial zero. The last property fails over small characteristic, as we know that the partial derivatives vanish for p-powers over characteristic p. This leads to failure of Jacobian criterion.

Let us start with the question of finding analogs of differential operator in positive characteristic that may help to reduce our problem to linear algebra. In our search for nice derivative-like operator which may not kill p-powers, and that satisfies an analogous product rule, Hasse derivative (Eqn.3.3) naturally comes up. Although derivative of x^p with respect to x is zero over characteristic p, the p-th order Hasse derivative of x^p is 1 (essentially, differentiate it p-times over $\mathbb{Q}$, divide by $p!$, and return back to $\mathbb{F}_p$). Hasse derivatives also satisfy a product rule analogous to the product rule (Sec.2.6) of higher order derivatives. Hasse derivatives can be used to construct higher order differentials and such differentials are known as Hasse-Schmidt higher order differentials. For example, we can define a second order differential operator using the second order Hasse-Schmidt derivatives: $H_2 := H_1^2/2!$, with its action on f given as

$$H_2 f := \frac{H_1^2}{2!} f = \sum_{i,j \in [n]} \frac{1}{2!} \frac{\partial^2 f}{\partial x_i \partial x_j} z_i z_j \,.$$

More generally, we can define:

$$\begin{aligned} H_k f :=& \frac{H_1^k}{k!} f \\ =& \sum_{k_1 + \cdots + k_n = k} \frac{1}{(k_1)!(k_2)! \ldots (k_n)!} \cdot \frac{\partial^k f}{\partial x_1^{k_1} \partial x_2^{k_2} \cdots \partial x_n^{k_n}} \cdot z_1^{k_1} \cdots z_n^{k_n} \,. \end{aligned} \tag{3.3}$$

So, the natural choice of differential operator for getting a Jacobian like criterion over characteristic p would be Hasse-Schmidt derivation of order p^i, for a large enough i. Hasse-Schmidt operator gets us through one direction (algebraic dependence implies linear dependence), but this operator vanishes for polynomials of degree $< p^i$, so it cannot be used to certify independence in all cases. This problem can be avoided by considering instead the sum of Hasse-Schmidt higher order differentials up to k, where $k \geq p^i$. Let us call this operator $\mathcal{H}_k$ (Sec.2.6).

But taking higher derivative does not immediately give a generalization of Jacobian. A major problem is, one wants to reduce the question of deciding algebraic independence of n polynomials to deciding the linear independence of n polynomials/vectors. As in the case of the proof of Jacobian criterion, ideally we want a bijection between algebraic dependencies amongst the polynomials and the linear dependencies of the corresponding vectors. But simple counting shows that there are n first derivatives in n variables, but the number of higher derivatives is $> n$. We can fix this issue by reducing algebraic dependence of n polynomials to linear dependence of n polynomials *modulo* a carefully chosen subspace (by stuffing higher-products of the higher differentials in that subspace). This gets implemented by working modulo $\mathcal{U}_t$ in Sec.3.2.2.

Now we give a version of generalized Jacobian criterion described in terms of Hasse-Schmidt differentials. We describe the notation first. $\mathcal{H}_i f(\mathbf{x})$ contains all the terms, of the polynomial $f(\mathbf{x}+\mathbf{z}) - f(\mathbf{x}) \in \mathbb{F}_p(\mathbf{x})[\mathbf{z}]$, that are of degree (wrt $\mathbf{z}$) $\leq i$.

Let $S \supseteq R$ be a ring extension over a field $\mathbb{F}$, and let $v_1, \ldots, v_m \in S$. Then the *R-module* $\langle v_1, \ldots, v_m \rangle_R$ is simply the set of linear combinations of the v_i's where the coefficients come from R. It is also a vector space over $\mathbb{F}$. We extend this notation to powers ($r \geq 1$):

$$\langle v_1, \ldots, v_m \rangle_R^r := \left\{ \sum_{\alpha_i \in R} \alpha_i \cdot v_1^{q_1} \cdots v_m^{q_m} \mid q_1 + \cdots + q_m = r,\ q_j \geq 0 \right\}.$$

Using the above notation, we define $\mathcal{U}_k$, a subspace of $\mathbb{F}_p(\mathbf{x})[\mathbf{z}]$.

$$\mathcal{U}_k := \langle \mathcal{H}_{k-1}\mathbf{f}\rangle^2_{\mathbb{F}_p(\mathbf{x})} + \cdots + \langle \mathcal{H}_1\mathbf{f}\rangle^k_{\mathbb{F}_p(\mathbf{x})} + \langle \mathbf{z}\rangle^{k+1}_{\mathbb{F}_p(\mathbf{x})[\mathbf{z}]},$$
$$\mathcal{U}_1 := \langle \mathbf{z}\rangle^2_{\mathbb{F}_p(\mathbf{x})[\mathbf{z}]}.$$

Now we are ready to state the theorems.

Theorem 3.4.1 (Dependent)**.** *Let $\mathbf{f} \subset \mathbb{F}_p[\mathbf{x}]$ be a set of n n-variate polynomials. If $\mathbf{f}$ is algebraically dependent, then $\mathcal{H}_k\mathbf{f}$ is $\mathbb{F}_p(\mathbf{x})$-linearly dependent modulo the subspace $\mathcal{U}_k$, for all $k \geq 1$.*

Theorem 3.4.2 (Independent)**.** *If $\mathbf{f} \subset \mathbb{F}_p[\mathbf{x}]$ is algebraically independent with inseparable degree (of the field extension $\mathbb{F}_p(\mathbf{x})/\mathbb{F}_p(\mathbf{f})$) equal to p^i, then $\mathcal{H}_k\mathbf{f}$ is $\mathbb{F}_p(\mathbf{x})$-linearly independent modulo the subspace $\mathcal{U}_k$, for $k \geq p^i$.*

Moreover, $\mathcal{H}_k\mathbf{f}$ is $\mathbb{F}_p(\mathbf{x})$-linearly dependent modulo $\mathcal{U}_k$, for every $1 \leq k < p^i$.

These two theorems together give a Jacobian like criterion for testing algebraic independence of n polynomials, assuming we are given a *promise* that the inseparable degree is bounded by p^i. The proofs of these two theorems are essentially same as the the proofs of Theorem 3.2.3 and Theorem 3.2.7.

Matrix version of the criterion. From the above two theorems, we can get a Jacobian like matrix J with entries in $\mathbb{F}_p[\mathbf{x}]$. The rows of the matrix J are indexed by $\mathcal{H}_k\mathbf{f}$ and the columns of the matrix are indexed by the monomials in $\mathbf{z}$ with degree $\leq k$. The entry, in the i-th row and the column indexed by a monomial m, is the coefficient of monomial m in the polynomial $\mathcal{H}_k f_i$. Clearly, the entries of J are order $\leq k$ Hasse-Schmidt derivatives of f_i's.

The subspace $\mathcal{V}_k$ of the ambient vector space $\mathbb{F}_p(\mathbf{x})^{\binom{n+k}{k}}$ is defined by taking all the coefficient vectors (corresponding to the coefficients of $\mathbf{z}$-monomials of degree $\leq k$) of the polynomials in $\mathcal{U}_k$.

It can be seen directly from our theorems that *the row-span of the matrix J has full rank modulo the subspace $\mathcal{V}_k$ iff the polynomials $\mathbf{f}$ are algebraically independent* (assuming that inseparability degree is $\leq k$).

3.5 Discussions

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 t. Bringing down the complexity from randomized poly(s, $\binom{t+n}{n}$) to randomized poly(s, n, t) can be the next target.

The complexity of computing annihilating polynomials is well understood, but the complexity of computing an approximate functional relation is not yet clear. Is there a randomized poly(s, n, t)-time algorithm to compute an approximate functional dependence up to precision t, given a set of dependent polynomials?

Following Kumar and Saraf [KS17], another interpretation of our criterion can be given via implicit function theorem/ Newton iteration (see Lemma 5.1.1). Given a polynomial $p(x_1, \ldots, x_n)$, let us view it as $c_0 + c_1 x_n + \ldots + c_d {x_n}^d$, where the coefficients $c_i \in \mathbb{F}[x_1, \ldots, x_{n-1}]$. A polynomial $g(x_1, \ldots, x_{n-1})$ is a root of $p(x_1, \ldots, x_n)$ wrt x_n if $p(x_1, \ldots, x_{n-1}, g) = 0$. Using Newton iteration, a beautiful lemma of [DSY09] showed that $g^{\leq t}$ (for all nonnegative integer t) can be written as a truncated polynomial $F_t(c_0, \ldots, c_d)^{\leq t}$ if the polynomial p satisfies a non-degeneracy condition: $\partial_{x_n} p(x_1, \ldots, x_n)$ evaluates to a nonzero constant at $(\mathbf{0}, g(\mathbf{0}))$. If $A(f_1, \ldots, f_n) = 0$ for a set of polynomials $f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})$, f_n can be seen as a root of $A(f_1, \ldots, f_{n-1}, y)$ wrt y. Now using Lemma 3.1 of [DSY09], one of the polynomials f_i can be approximately written as a function of the other polynomials if the annihilator A satisfies the above non-degeneracy condition. Over characteristic zero, it is easy to see that an annihilator of minimal degree satisfies the non-degeneracy condition if the variables of the given polynomials are shifted by a random point $\mathbf{a}$. Using the ideas presented in this chapter, we can generalize this to positive characteristic.

It is known that algebraic matroids over fields of positive characteristic may not have any linear representation (unlike the zero characteristic case, where using the Jacobian criterion any algebraic matroid can be represented by a linear matroid) [Lin85]. Does this suggest that algebraic dependence of polynomials may not be directly reducible to linear dependence over finite fields?

Finally, we mention an interesting application of our criterion. [BMS13] defined transcendence degree preserving homomorphisms as *faithful homomorphism* and showed that this concept has applications in the derandomization of polynomial identity testing problem. Over characteristic zero or large, constructing faithful homomorphism boils down to preserving the rank of the Jacobian. For several interesting special cases, [ASSS16] gave such constructions. Over small characteristic, they do not work due to failure of the Jacobian. Using our criterion, Chatterjee and Saptharishi [CS18] gave an explicit construction of faithful maps for some special classes over small characteristic. Similar to our criterion, the construction of [CS18] is efficient only in the bounded inseparability degree setting. Improving the construction of [CS18] is another problem of interest.

Chapter 4

Algebraic Dependence over Finite Fields is in AM ∩ coAM

Abstract

This chapter is based on joint work with Guo and Saxena [GSS18].

In this chapter, we give a different approach of testing algebraic dependence over finite fields, using basic algebraic geometry. We show: For finite fields, algebraic dependence testing (AD($\mathbb{F}_q$)) is in AM ∩ coAM. This result vastly improves the current best upper bound known for AD($\mathbb{F}_q$)– from being 'outside' the polynomial hierarchy (namely NP$^{\#\mathsf{P}}$ [MSS14]) to 'lower' than the second-level of polynomial hierarchy (namely AM ∩ coAM). Our result rules out the possibility of the problem's NP-hardness, under standard complexity theory assumptions.

4.1 Proof overview

Now, we present a different approach of testing algebraic dependence over finite fields using some basic tools from algebraic geometry. We view the given polynomials $f_1, \ldots, f_n \in \mathbb{F}[x_1, \ldots, x_n]$ as a polynomial map $f : \mathbb{F}^n \to \mathbb{F}^n$. This map is defined as: $a \mapsto (f_1(a), \ldots,$

$f_n(a)$), where a denote a point $(a_1, \ldots, a_n)$.

Note that if $f_1, \ldots, f_n$ are algebraically dependent and A is an annihilator of them, then for any point $a \in \mathbb{F}^n$, $A(f_1(a), \ldots, f_n(a)) = 0$. Thus, the image of the polynomial map f is contained in the zero set of the polynomial $A(y_1, \ldots, y_n)$. By the Schwartz-Zippel-DeMillo-Lipton lemma [AB09, A.36], the size of this set is *small*. On the contrary, the image of the polynomial map defined by algebraically independent polynomials is large. For example, the image of algebraically independent polynomials $x_1, \ldots, x_n$ is the full ambient space $\mathbb{F}^n$. Exploiting the gap of the size of the image set in these two cases, we can show algebraic dependence testing is in AM.

For the coAM protocol, we study the cardinality of a different set: the set of preimage for a random point in the image of the polynomial map. In the case of algebraically independent polynomials, the size of this set is small. On the other hand, as the image of algebraically dependent polynomials is small, the preimage set is large for most points.

The main idea here can be expressed using basic algebraic geometry extending an intuition from linear algebra. If the given polynomials are all linear, then for any point a in the image of the given linear polynomials, dimension (Image) + dimension (Preimage of a)=n. This is known as the rank-nullity theorem in linear algebra. In algebraic geometry, this relation generalizes to polynomials of higher degree as well, but now it holds only for a *random/generic* point in the image (not for all the points in the image) [Eis13, Section 14.3]. The transcendence degree is equal to the dimension of the variety corresponding to the closure of the image of $f_1, \ldots, f_n$. If the polynomials are independent, the dimension of the image is n. Thus the dimension of the preimage for a random point is zero. If the polynomials are dependent, for a random point, the dimension of the preimage is at least one. Exploiting the gap of the size of the preimage set (for a point in the image randomly picked by Arthur) in these two cases, we can show algebraic dependence testing is in coAM.

Remark: Note that there is no general relationship between the transcendence degree of $f_1, \ldots, f_n$ and the dimension of the zero set corresponding to the system of equations $\{f_1 = 0, \ldots, f_n = 0\}$[MSS14]. Dimension of the zero set V_1 corresponding to the equation

$x_1 = 0$ is same as dimension of the zero set V_2 corresponding to the equations $\{x_1 = x_1x_2 = 0\}$. But transcendence degree of $\{x_1\}$ is one, whereas the transcendence degree of $\{x_1, x_1x_2\}$ is two.

It is known that computing the dimension of the zero set of $f_1, \ldots, f_n$ (or the decision version whether the dimension is at least d for a given number d) is NP-hard [Koi97]. But, we show that computing the dimension of the closure of $\mathrm{Im}(f)$ is unlikely to be NP-hard. One advantage in our problem is that we could sample a random point in the set $\mathrm{Im}(f)$. In contrast, it is not clear how to sample a random point in the zero set of $f_1, \ldots, f_n$.

4.2 Proof of the main result

Now, we will prove the main result of this chapter.

Theorem 4.2.1. *Algebraic dependence testing of polynomials (given as arithmetic circuits) in $\mathbb{F}_q[\mathbf{x}]$ is in AM $\cap$ coAM.*

Given $f_1, \ldots, f_m \in \mathbb{F}_q[x_1, \ldots, x_n]$, we want to decide if they are algebraically dependent. We could assume, with some preprocessing (Lemma 2.3.1, Lemma 2.3.4), that $m = n$. So, we assume the input instance to be $\mathbf{f} := \{f_1, \ldots, f_n\}$ with nonconstant polynomials.

In the following, let $D := \prod_{i\in[n]} \deg(f_i) > 0$ and $D' := \max_{i\in[n]} \deg(f_i) > 0$. Let d be a positive integer and $q' = q^d$. The value of d will be determined later. Let $f : \mathbb{F}_{q'}^n \to \mathbb{F}_{q'}^n$ be the polynomial map $a \mapsto (f_1(a), \ldots, f_n(a))$. For $b = (b_1, \ldots, b_n) \in \mathbb{F}_{q'}^n$, denote by N_b the size of the preimage $f^{-1}(b) = \left\{\mathbf{x} \in \mathbb{F}_{q'}^n \mid f(\mathbf{x}) = b\right\}$.

Define $\mathbb{A} := \overline{\mathbb{F}}_q$ and $\overline{N}_b := \#\{\mathbf{x} \in \mathbb{A}^n \mid f_i(\mathbf{x}) = b_i, \text{ for all } i \in [n]\}$ which might be ∞. Let $Q \in \mathbb{F}_q[y_1, \ldots, y_n]$ be a nonzero annihilator, of minimal degree, of $f_1, \ldots, f_n$. If it exists then $\deg(Q) \leq D$ by Perron's bound.

4.2.1 AM protocol

First, we study the independent case.

Lemma 4.2.2 (Dimension zero preimage)**.** *Suppose* $\mathbf{f}$ *are independent. Then* $\overline{N}_{f(a)}$ *is* finite *for all but at most* (nDD'/q')*-fraction of* $a \in \mathbb{F}_{q'}^n$*.*

Proof. For $i \in [n]$, let $G_i \in \mathbb{F}_q[z, y_1, \ldots, y_n]$ be the annihilator of $\{x_i, f_1, \ldots, f_n\}$. We have $\deg(G_i) \leq D$ by Perron's bound. Consider $a \in \mathbb{F}_{q'}^n$ such that $G_i'(z) := G_i(z, f_1(a), \ldots, f_n(a)) \in \mathbb{F}_q[z]$ is a nonzero polynomial for every $i \in [n]$. We claim that $\overline{N}_{f(a)}$ is finite for such a.

To see this, note that for any $b = (b_1, \ldots, b_n) \in \mathbb{A}^n$ satisfying the equations $f_i(b) = f_i(a)$, $i \in [n]$, we have

$$0 = G_i(b_i, f_1(b), \ldots, f_n(b)) = G_i(b_i, f_1(a), \ldots, f_n(a)) = G_i'(b_i), \quad \forall i \in [n]\,.$$

Hence, each b_i is a root of G_i'. It follows that $\overline{N}_{f(a)} \leq \prod_{i\in[n]} \deg(G_i') < \infty$, as claimed.

It remains to prove that the number of $a \in \mathbb{F}_{q'}^n$ satisfying $G_i' = 0$, for some index $i \in [n]$, is bounded by $nDD'q'^{-1} \cdot q'^n$. Fix $i \in [n]$. Suppose $G_i = \sum_{j=0}^{d_i} G_{i,j} z^j$, where $d_i := \deg_z(G_i)$ and $G_{i,j} \in \mathbb{F}_q[y_1, \ldots, y_n]$, for $0 \leq j \leq d_i$. The leading coefficient G_{i,d_i} is a nonzero polynomial. As $f_1, \ldots, f_n$ are algebraically independent, the polynomial $G_{i,d_i}(f_1, \ldots, f_n) \in \mathbb{F}_q[x_1, \ldots, x_n]$ is also nonzero. Its degree is $\leq D' \deg(G_{i,d_i}) \leq D' \deg(G_i) \leq DD'$. By the Schwartz-Zippel-DeMillo-Lipton lemma, for all but at most (DD'/q')-fraction of $a \in \mathbb{F}_{q'}^n$, we have $G_{i,d_i}(f_1(a), \ldots, f_n(a)) \neq 0$ which implies

$$G_i'(z) = G_i(z, f_1(a), \ldots, f_n(a)) = \sum_{j=0}^{d_i} G_{i,j}(f_1(a), \ldots, f_n(a)) z^j \neq 0\,.$$

The claim now follows from the union bound. □

We need the following affine version of Bézout's Theorem. Its proof can be found in [Sch95, Theorem 3.1].

Theorem 4.2.3 (Bézout's)**.** *Let* $g_1, \ldots, g_n \in \mathbb{A}[x_1, \ldots, x_n]$*. Then the number of common zeros of* $g_1, \ldots, g_n$ *in* $\mathbb{A}^n$ *is either infinite, or at most* $\prod_{i\in[n]} \deg(g_i)$*.*

Combining Lemma 4.2.2 with Bézout's Theorem, we obtain

Lemma 4.2.4 (Small preimage)**.** *Suppose* $\mathbf{f}$ *are independent. Then* $N_{f(a)} \leq D$ *for all but at most* (nDD'/q')*-fraction of* $a \in \mathbb{F}_{q'}^n$*.*

Next, we study the dependent case (with an annihilator Q).

Lemma 4.2.5 (Large preimage)**.** *Suppose* $\mathbf{f}$ *are dependent. Then for* $k > 0$*, we have* $N_{f(a)} > k$ *for all but at most* (kD/q')*-fraction of* $a \in \mathbb{F}_{q'}^n$*.*

Proof. Let $\mathrm{Im}(f) := f(\mathbb{F}_{q'}^n)$ be the image of the map. Note that Q vanishes on all the points in $\mathrm{Im}(f)$. So, $|\mathrm{Im}(f)| \leq Dq'^{n-1}$ by the Schwartz-Zippel-DeMillo-Lipton lemma [AB09, A.36].

Let $B := \{b \in \mathrm{Im}(f) : N_b \leq k\}$ be the "bad" images. We can estimate the bad domain points as,

$$\#\{a \in \mathbb{F}_{q'}^n : N_{f(a)} \leq k\} \;=\; \#\{a \in \mathbb{F}_{q'}^n : f(a) \in B\} \;\leq\; k|B| \leq k|\mathrm{Im}(f)| \;\leq\; kDq'^{n-1}\,.$$

which proves the lemma. □

Theorem 4.2.6 (AM)**.** *Testing algebraic dependence of* $\mathbf{f}$ *is in AM.*

Proof. Fix $q' = q^d > 4nDD' + 4kD$ and $k := 2D$. Note that d will be polynomial in the input size. For an $a \in \mathbb{F}_{q'}^n$, consider the set $f^{-1}(f(a)) := \{\mathbf{x} \in \mathbb{F}_{q'}^n \mid f(\mathbf{x}) = f(a)\}$.

By Lemma 4.2.4 and Lemma 4.2.5: When Arthur picks a randomly, with high probability, $|f^{-1}(f(a))| = N_{f(a)}$ is more than $2D$ in the dependent case while $\leq D$ in the independent case. Note that an upper bound on $\prod_{i\in[n]} \deg(f_i)$ can be deduced from the size of the input circuits for f_i's; thus, we know D. Moreover, containment in $f^{-1}(f(a))$ can be tested in P. Thus, by Lemma 2.10.1, $\mathrm{AD}(\mathbb{F}_q)$ is in AM. □

4.2.2 coAM protocol

We first study the independent case.

Lemma 4.2.7 (Large image)**.** *Suppose* $\mathbf{f}$ *are independent. Then* $N_b > 0$ *for at least* $(D^{-1} - nD'q'^{-1})$*-fraction of* $b \in \mathbb{F}_{q'}^n$*.*

Proof. Let $S := \{a \in \mathbb{F}_{q'}^n : N_{f(a)} \leq D\}$. Then $|S| \geq (1 - nDD'q'^{-1}) \cdot q'^n$ by Lemma 4.2.4. As every $b \in f(S)$ has at most D preimages in S under f, we have $|f(S)| \geq |S|/D \geq (D^{-1} - nD'q'^{-1}) \cdot q'^n$. This proves the lemma since $N_b > 0$ for all $b \in f(S)$. □

Next, we study the dependent case.

Lemma 4.2.8 (Small image)**.** *Suppose* $\mathbf{f}$ *are dependent. Then* $N_b = 0$ *for all but at most* (D/q')*-fraction of* $b \in \mathbb{F}_{q'}^n$*.*

Proof. By definition: $N_b > 0$ iff $b \in \mathrm{Im}(f) := f(\mathbb{F}_{q'}^n)$. It was shown in the proof of Lemma 4.2.5 that $|\mathrm{Im}(f)| \leq Dq'^{n-1}$. The lemma follows. □

Theorem 4.2.9 (coAM)**.** *Testing algebraic dependence of* $\mathbf{f}$ *is in coAM.*

Proof. Fix $q' = q^d > D(2D + nD')$. Note that d will be polynomial in the input size. For $b \in \mathbb{F}_{q'}^n$, consider the set $f^{-1}(b) := \{\mathbf{x} \in \mathbb{F}_{q'}^n \mid f(\mathbf{x}) = b\}$ of size N_b.

Define $S := \mathrm{Im}(f)$. Note that: $b \in \mathbb{F}_{q'}^n$ has $N_b > 0$ iff $b \in S$. Thus, by Lemma 4.2.7 (resp. 4.2.8), $|S| \geq (D^{-1} - nD'q'^{-1})q'^n > 2Dq'^{n-1}$ (resp. $|S| \leq Dq'^{n-1}$) when $\mathbf{f}$ are independent (resp. dependent). Note that an upper bound on $\prod_{i\in[n]} \deg(f_i)$ can be deduced from the size of the input circuits for f_i's; thus, we know Dq'^{n-1}. Moreover, containment in S can be tested in NP. Thus, by Lemma 2.10.1, $\mathrm{AD}(\mathbb{F}_q)$ is in coAM. □

Proof of Theorem 4.2.1. The statement directly follows from Theorem 4.2.6 and Theorem 4.2.9. □

4.3 Discussions

We give protocols to distinguish the two cases, whether the dimension of the closure of the image is n or less. Can we extend the ideas further to give a randomized polynomial-time algorithm? An intermediate problem (suggested by Ilya Volkovich) is to prove the result $\mathrm{AD}(\mathbb{F}) \in \mathrm{SBP} \cap \mathrm{coSBP}$. Here SBP stands for "small bounded-error probability" and is a subclass of AM introduced in [BGM06].

As discussed in [DGW09, Dvi09, DGRV11], transcendence degree is a measure of *entropy* of a polynomial source $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$, where $\mathbb{F}_q$ is a finite field. A polynomial source [BSG13] can be seen as a random variable $f(U)$, where $f : \mathbb{F}_q^r \to \mathbb{F}_q^n$ is a polynomial map and U is a random variable uniformly distributed over $\mathbb{F}_q^r$. The problem of approximating entropy of polynomial sources over finite fields is another related problem of interest [DGRV11].

Part II

Factor Conjecture

Chapter 5

Towards Factor Conjecture

Abstract

The factor conjecture states that low degree factors of high degree polynomials computed by small-sized arithmetic circuits have small-sized circuits. In this chapter, we give a simple proof of this conjecture for a special case. If the polynomial $f = \prod_{i=1}^{m} f_i^{e_i}$ is given by a size s circuit, we show that the squarefree part $\prod_{i=1}^{m} f_i$ can be computed by a circuit of $\text{poly}(s, d)$ size, where d is the degree of the squarefree part. This shows that the factor conjecture is true when squarefree part of the polynomial has low degree. We assume that the underlying field has characteristic zero.

This chapter is partially based on joint work with Dutta and Saxena [DSS18]. The proof of Theorem 5.2.1 presented in this chapter also appeared in the thesis [Dut18]. The proof of the main result in this chapter (Theorem 5.3.1) differs from the proof given in [DSS18, Dut18].

5.1 Preliminaries

First, we present some preliminary results. Recall that the notation $\boldsymbol{x}$ stands for the variables $x_1, \ldots, x_n$. $\mathbf{0}$ denote the point $(0, \ldots, 0)$. $I := \langle \boldsymbol{x} \rangle$ is the ideal of the polynomial

ring $\mathbb{F}[x_1, \ldots, x_n]$ generated by $x_1, \ldots, x_n$. $\langle \boldsymbol{x} \rangle^d$ is the d-th power of ideal I. $\langle \boldsymbol{x} \rangle^d$ is generated by all monomials of degree d. If $f(\boldsymbol{x})$ is a polynomial or power series, $f(\boldsymbol{x})$ mod $\langle \boldsymbol{x} \rangle^d$ equals part of $f(\boldsymbol{x})$ up to degree $(d-1)$.

The following classical result ([BCS13, Theorem.2.31]) says power series roots of a polynomial can be approximated to arbitrary degree using Newton iteration. It is a key tool used in this chapter.

Lemma 5.1.1 (Implicit Function Theorem [BCS13])**.** *Let $P(\boldsymbol{x}, y) \in \mathbb{F}(\boldsymbol{x})[y]$, $P'(\boldsymbol{x}, y) = \frac{\partial P(\boldsymbol{x}, y)}{\partial y}$ and $\mu \in \mathbb{F}$ be such that $P(\mathbf{0}, \mu) = 0$ but $P'(\mathbf{0}, \mu) \neq 0$. Then, there is a unique power series S such that $S(\mathbf{0}) = \mu$ and $P(\mathbf{0}, S) = 0$.*

Furthermore, the power series root S can be approximated to arbitrary degree using Newton iteration. Let $y_0 = \mu$ and $\forall t \geq 0$,

$$y_{t+1} = y_t - \frac{P(\boldsymbol{x}, y_t)}{P'(\boldsymbol{x}, y_t)}.$$

Then, $\forall t \geq 0$,

$$S \equiv y_t \bmod \langle \boldsymbol{x} \rangle^{2^t}.$$

Proof. The proof is by induction on t. Let $y_0 := \mu$. Thus, base case is true. Define $y_{t+1} := y_t - \frac{P(\boldsymbol{x}, y_t)}{P'(\boldsymbol{x}, y_t)}$. As, $y_t \equiv y_{t-1} \bmod \langle \boldsymbol{x} \rangle^{2^{t-1}} \implies y_t(\mathbf{0}) = \mu$. Hence $P'(\boldsymbol{x}, y_t)|_{\boldsymbol{x}=\mathbf{0}} = P'(\mathbf{0}, \mu) \neq 0$ and so $P'(\boldsymbol{x}, y_t)$ is invertible in the power series ring. So, $y_{t+1} \in \mathbb{F}[[\boldsymbol{x}]]$.

Now, use Taylor expansion to complete the induction step.

$$\begin{aligned} P(\boldsymbol{x}, y_{t+1}) &= P\left(\boldsymbol{x},\, y_t - \frac{P(\boldsymbol{x}, y_t)}{P'(\boldsymbol{x}, y_t)}\right) \\ &= P(\boldsymbol{x}, y_t) - P'(\boldsymbol{x}, y_t)\frac{P(\boldsymbol{x}, y_t)}{P'(\boldsymbol{x}, y_t)} + \frac{P''(\boldsymbol{x}, y_t)}{2!}\left(\frac{P(\boldsymbol{x}, y_t)}{P'(\boldsymbol{x}, y_t)}\right)^2 - \ldots \\ &= 0 \bmod \langle \boldsymbol{x} \rangle^{2^{t+1}}. \end{aligned}$$

Thus, $P(\boldsymbol{x}, y_{t+1}) \equiv 0 \bmod \langle \boldsymbol{x} \rangle^{2^{t+1}}$ and $y_{t+1} \equiv y_t \bmod \langle \boldsymbol{x} \rangle^{2^t}$.

□

Remark. In the above theorem, we crucially need that $P(\boldsymbol{x}, y)$ is a polynomial in y, not a power series. Thus, we can evaluate $P(\boldsymbol{x}, y)$ at a power series with a nonzero constant term. In [BCS13, Theorem.2.31], the theorem is stated in full generality where $P(\boldsymbol{x}, y)$

can be a power series in y. In that case, μ must be zero.

Lemmas for pre-processing:

A polynomial is squarefree if it has no factor with multiplicity greater than one. The following standard lemma gives a criterion of a polynomial to be squarefree (over characteristic zero).

Lemma 5.1.2 (Squarefree criterion)**.** *Let $f \in \mathbb{F}(\boldsymbol{x})[y]$ be a polynomial with $deg_y(f) \geq 1$. f is squarefree if and only if f and $f' := \partial_y f$ are coprime.*

Proof. We show that there does not exist $g \in \mathbb{F}(\boldsymbol{x})[y]$ with positive degree in y such that $g \mid \gcd_y(f(\boldsymbol{x}, y), f'(\boldsymbol{x}, y))$. Suppose g is an irreducible polynomial with positive degree in y that divides both $f(\boldsymbol{x}, y)$ and $f'(\boldsymbol{x}, y)$. If $f(\boldsymbol{x}, y) = gh$, then $f'(\boldsymbol{x}, y) = gh' + g'h$. Thus, $g \mid g'h$. As g is irreducible and $\deg_y(g') < \deg_y(g)$, we get that $g \mid h$. Hence, $g^2 \mid f$. This contradicts the hypothesis that f is squarefree. □

We need the following property of resultants. Given two polynomials $f(\boldsymbol{x}, y)$ and $g(\boldsymbol{x}, y)$, the resultant of f and g wrt y is the determinant of the Sylvester matrix. For the definition of the Sylvester matrix, see [LN97].

Proposition 5.1.3 (Resultant and gcd)**.** *1. Let $f, g \in \mathbb{F}[\boldsymbol{x}, y]$ be polynomials with positive degree in y. Then, $Res_y(f, g) = 0 \iff f$ and g have a common factor in $\mathbb{F}[\boldsymbol{x}, y]$ which has positive degree in y.*

2. There exists $u, v \in \mathbb{F}[\boldsymbol{x}]$ such that $uf + vg = Res_y(f, g)$.

A proof of this can be found in [VZGG13, Sec.6].

The following standard lemma states that coprimality (wrt a variable) of two polynomials are preserved under random projections of the remaining variables.

Lemma 5.1.4 (Coprimality)**.** *Let $f, g \in \mathbb{F}(\boldsymbol{x})[y]$ be coprime polynomials wrt y (& nontrivial in y). Then, for $\boldsymbol{\beta} \in_r \mathbb{F}^n$, $f(\boldsymbol{\beta}, y)$ and $g(\boldsymbol{\beta}, y)$ are coprime (& nontrivial in y).*

Proof. Let $f = \sum_{i=1}^{d} f_i y^i$ and $g = \sum_{i=1}^{e} g_i y^i$. Choose a random $\boldsymbol{\beta} \in_r \mathbb{F}^n$. Then, by Proposition 5.1.3 & Schwartz-Zippel-DeMillo-Lipton lemma , $f_d \cdot g_e \cdot \mathrm{Res}_y(f, g)$ at $\boldsymbol{x} = \boldsymbol{\beta}$ is nonzero. This implies that $\mathrm{Res}_y(f(\boldsymbol{\beta}, y), g(\boldsymbol{\beta}, y)) \neq 0$. This implies, by Proposition 5.1.3, $f(\boldsymbol{\beta}, y)$ and $g(\boldsymbol{\beta}, y)$ are coprime. □

Transforming the polynomial into a monic polynomial: A polynomial $f(\boldsymbol{x}, y)$ is called monic wrt y if the coefficient of the highest degree (degree wrt y) monomial in y is a constant. The given polynomial to be factored $f(\boldsymbol{x})$ may not be monic wrt any of the variables $x_1, \ldots, x_n$. We can transform to a monic polynomial by applying an invertible linear transformation on the variables. After getting the factors of the monic polynomial, we get the factors of the original polynomial by applying the inverse of the transformation.

Lemma 5.1.5 (Transform to monic)**.** *For a polynomial $f(\boldsymbol{x})$ of total degree $d \geq 0$ and random $\alpha_i \in_r \mathbb{F}$, the transformed polynomial $g(\boldsymbol{x}, y) := f(\boldsymbol{\alpha} y + \boldsymbol{x})$ has a nonzero constant as coefficient of y^d, and degree wrt y is d.*

Proof. The transformation we apply is $x_i \mapsto x_i + \alpha_i y$ where $i \in [n]$. Now, we write $f = \sum_{|\boldsymbol{\beta}|=d} c_{\boldsymbol{\beta}} \boldsymbol{x}^{\boldsymbol{\beta}} +$ lower degree terms . Coefficient of y^d in g is $\sum_{|\boldsymbol{\beta}|=d} c_{\boldsymbol{\beta}} \boldsymbol{\alpha}^{\boldsymbol{\beta}}$. For a random $\boldsymbol{\alpha}$ this coefficient will not vanish with high probability (using the Schwartz-Zippel-DeMillo-Lipton lemma). As it is the highest degree monomial in g, $\deg_y(g) = \deg(f) = d$ and g is monic wrt y. □

Kaltofen proved that the factor conjecture is true when f is of a very special form: $f = g^e$ for some $g \in \mathbb{F}[x_1, \ldots, x_n]$.

Theorem 5.1.6 ([Kal87])**.** *Suppose $f = g^e$ has a circuit of size s. Then g has a circuit of size poly(s, d_g) where d_g is the degree of g.*

This theorem can be proved by reducing factoring to root finding using an auxiliary equation $T^e - g^e = 0$. Note that the equation $P(T) = T^e - g^e = 0$ has a root $T = g$ of multiplicity one. We can inductively build a circuit for the root g using Lemma 5.1.1. See Theorem 2 in [Kal87] for the details of the circuit construction.

5.2 Factorization via power series root approximation

Most of the works in multivariate polynomial factorization use the concept of Hensel lifting. A closely related (and mathematically equivalent [VzG84]) approach is via approximation of the power series roots of the polynomial to be factored. The idea of computing factors via finding approximate power series root was used in bivariate factoring [Kal82] and

for general circuit factoring [Bür04]. Recently, [Oli16] showed how to reduce polynomial factoring (of bounded individual degree polynomials given by low depth circuits) to finding approximations of power series roots. In [DSS18], we give another reduction of factoring to root approximation and show its applications in proving various closure under factoring results for different classes. Here, we closely follow the exposition of [DSS18].

The following basic theorem in algebra relates factors and roots of a polynomial. A polynomial $f(x)$ over a field $\mathbb{F}$ has a factor $x - a$ iff $f(a) = 0$. This generalizes this to multivariate polynomials as well: $y - g(\boldsymbol{x})$ is a factor of $f(\boldsymbol{x}, y)$ iff $f(\boldsymbol{x}, g(\boldsymbol{x})) = 0$.

Thus, if there is a factor of the form $(y - g(\boldsymbol{x}))$, it can be found by finding roots of the polynomial $f(\boldsymbol{x}, y)$ with respect to y. But, a polynomial may not have any factor of the form $y - g(\boldsymbol{x})$, where $g(\boldsymbol{x})$ is a polynomial. Nevertheless, we show that after applying a random shift on the variables, a polynomial can be written as product of factors of the form $y - g(\boldsymbol{x})$, where $g(\boldsymbol{x})$ is a formal power series in $x_1, \ldots, x_n$. We formalize this in the following theorem.

Theorem 5.2.1 (Power Series Complete Split)**.** *Let $f \in \mathbb{F}[\boldsymbol{x}]$ and d_0 be the degree of the squarefree part of f. Consider $\alpha_i, \beta_i \in_r \mathbb{F}$ and the map $\tau : x_i \mapsto \alpha_i y + x_i + \beta_i$, $i \in [n]$, where y is a new variable. Then, over $\mathbb{F}[[\boldsymbol{x}]]$, $f(\tau\boldsymbol{x}) = k \cdot \prod_{i\in[d_0]}(y - g_i)^{\gamma_i}$, where $k \in \mathbb{F}^*$, $\gamma_i > 0$, and $g_i(\mathbf{0}) := \mu_i$. Moreover, μ_i's are distinct nonzero field elements.*

Remark. Note, that a polynomial $P \in \mathbb{F}(\boldsymbol{x})[y]$ can be completely factored in the algebraic closure of $\mathbb{F}(\boldsymbol{x})$. For example, take $P(x, y) = y^2 - x$. The roots of this polynomial are $\sqrt{x}$ and $-\sqrt{x}$. Theorem 5.2.1 says, after a random linear transformation, P completely splits over $\mathbb{F}[[\boldsymbol{x}]]$. For example, shifted $P(x + 1, y) = y^2 - (1 + x)$ has roots $\sqrt{(1 + x)}$ and $-\sqrt{(1 + x)}$. Now, $\sqrt{(1 + x)}$ can be written as a formal power series in x (recall Newton's binomial theorem with fractional exponent).

Proof. Let the complete irreducible factorization of f be $\prod_{i\in[m]} f_i^{e_i}$. We apply a random linear map τ so that f become monic in y (Lemma 5.1.5), this implies that all factors also would become monic. As the map τ is invertible, the factors $\tilde{f}_i := f_i(\tau\boldsymbol{x})$ remain irreducible.

As $f_i(\tau\boldsymbol{x})$ is irreducible, applying Lemma 5.1.4, we see $\tilde{f}_i(\mathbf{0}, y) = f_i(\boldsymbol{\alpha}y + \boldsymbol{\beta})$ and

$\partial_y \tilde{f}_i(\mathbf{0}, y)$ remain coprime. Thus, using Lemma 5.1.2, $\tilde{f}_i(\mathbf{0}, y)$ is square free.

Now, we can write $\tilde{f}_1(\mathbf{0}, y)$ as $\prod_{i=1}^{\deg(f_1)}(y - \mu_{1,i})$ for distinct nonzero field elements $\mu_{1,i}$.

Using classical Newton Iteration (Lemma 5.1.1) on $\tilde{f}_1(\boldsymbol{x}, y)$, we can write $\tilde{f}_1(\boldsymbol{x}, y)$ as a product of power series $\prod_{i=1}^{\deg(f_1)}(y - g_{1,i})$, with $g_{1,i}(\mathbf{0}) := \mu_{1,i}$. Similarly, each $f_i(\tau\boldsymbol{x})$ can be factored into linear factors in $\mathbb{F}[[\boldsymbol{x}]][y]$.

As f_i's are irreducible coprime polynomials, by Lemma 5.1.4, it is clear that $\forall i \in [m]$, $\tilde{f}_i(\mathbf{0}, y)$ are mutually coprime. In other words, $\mu_{j,i}$ are distinct and they are $\sum_i \deg(f_i) = d_0$ many. Hence, $f(\tau\boldsymbol{x})$ can be completely factored as $\prod_{i\in[m]} f_i(\tau\boldsymbol{x})^{e_i} = \prod_{i\in[d_0]}(y - g_i)^{\gamma_i}$, with $\gamma_i > 0$ and the field constants $g_i(\mathbf{0})$ being distinct. □

The following proposition is a classical theorem, that says factorization over power series ring is unique.

Proposition 5.2.2. *[ZS75, Chap.VII] Power series ring $\mathbb{F}[[\boldsymbol{x}]]$ is a unique factorization domain (UFD), and so is $\mathbb{F}[[\boldsymbol{x}]][y]$.*

Computing a factor via power series root approximation: Suppose g is a non-trivial factor of the polynomial f. If the degree of the input polynomial is d, the degree of g would be $\leq (d-1)$. Assume without loss of generality (Theorem 5.2.1) that the polynomial f completely splits into factors of the form $(y - r_i(\boldsymbol{x}))$ where $r_i(\boldsymbol{x})$ is a formal power series. Now, some of these factors combine to give g. Assume, we know all the power series roots of g. If $g = \prod_{i=1}^{m}(y - r_i(\boldsymbol{x}))$. If we approximate r_i's up to the precision equal to the degree of g (at most $d-1$), we can recover g completely by multiplying these m factors and truncating up to degree of g. Thus, to prove size upper bound of factors, we need to only bound the circuit size of the approximate roots.

How do we algorithmically find the right subset of roots that correspond to g? The brute force method considers all possible combinations. An efficient approach is via finding a minimal polynomial of the approximate root using linear algebra. This is a classical idea from [Kal82], see [Bür04, DSS18] for the details.

5.2.1 VP closed under factors: a proof via power series roots.

There are several proofs of closure of VP under factoring. The first proof by Kaltofen [Kal89] used Hilbert's irreducibility theorem and gave an efficient algorithm to compute the factors. There is another proof by Kaltofen [Kal87], where he uses a technique called single factor Hensel lifting. See [Bür04, Thm.1.2] for another exposition of this proof. Inspired by Kaltofen and Trager's [KT90] famous black-box factoring result, Bürgisser [Bür13, Thm.2.21] has given an elegant proof. Recently, [CKS19] has given another proof using Newton iteration for several variables.

Using the reduction of factoring to power series approximation, we can give a short proof of the fact that VP is closed under factors.

Theorem 5.2.3 ([Kal87])**.** *Let $f = g^e h$ be given by a size s arithmetic circuits. Assume g, h are coprime and let d_g be the degree of g. Now, g can be computed by size poly (e, d_g, s) sized circuit.*

Proof. Suppose $f(\boldsymbol{x}) = g^e h$ is a polynomial of degree d, given by a circuit of size s. After applying a random shift, we can assume that the polynomial has become monic and completely splits into power series roots. Using Theorem 5.2.1, $\tilde{f}(\boldsymbol{x}, y) = f(\tau\boldsymbol{x}) = k \cdot \prod_{i=1}^{d_0}(y - g_i)^{\gamma_i}$, with $g_i(\mathbf{0}) := \mu_i$ being distinct. For the sake of simplicity, we rename $\tilde{f}(\boldsymbol{x}, y)$ to $f(\boldsymbol{x}, y)$ and rename $\tilde{g}(\boldsymbol{x}, y)$ to $g(\boldsymbol{x}, y)$.

Now, g is a factor of multiplicity e. All the power series roots of g have multiplicity e. If a power series root has multiplicity greater than one, we can not approximate it using the classical Newton iteration formula, as the denominator become non invertible. To handle this, we do the following preprocessing.

If we take $(e-1)$-th derivative of $\tilde{f}(\boldsymbol{x}, y)$ with respect to y, we get a polynomial which has all the power series roots of g, but with multiplicity one. Using Lemma 2.8.2, we compute the circuit of $(e-1)$-th derivative of f, incurring size blow-up of $e^2 s$. Henceforth, we work with this circuit (for simplicity, we rename the polynomial to f again) to get approximations of the power series roots of the factor g.

Now we show, if a power series root g_i of a polynomial f of circuit size s has multiplicity one, it can be approximated up to degree δ, by an arithmetic circuit of size poly(s, δ).

Using Newton iteration formula, $g_{i,t+1} := g_{i,t} - \frac{f}{\partial_y f}\big|_{y=g_{i,t}} \bmod \langle \boldsymbol{x} \rangle^{2^t}$, where $g_{i,t} \equiv g_i \bmod \langle \boldsymbol{x} \rangle^{2^t}$. We compute $g_{i,t}$'s incrementally, $0 \leq t \leq \log \delta + 1$, by a circuit with division gates (keeping track of numerator and denominator separately).

Suppose $g_{i,t}$ has a circuit of size s_t. Now, from the Newton iteration formula, $s_{t+1} = s_t + \text{poly}(s)$. Thus, $g_{i,\delta}$ can be computed by a circuit (with division gates) of size $\text{poly}(s, \log \delta)$. Finally, we use Strassen's division elimination and homogenization (Lemma 2.8.3) to compute a circuit of size $\text{poly}(s, \delta)$ without any division gate. Using the above, all power series roots of g up to degree of g, $(g_i^{\leq d_g})$ can be computed by circuits of size $\text{poly}(e, d_g, s)$.

To get the circuit of the factor $g = \prod_i (y - g_i)$ of degree d_g, we have to multiply all the circuits of $(y - g_i^{\leq d_g})$ and finally truncate (using Lemma 2.8.1) up to degree d_g. The final size of the circuit of the factor is $\text{poly}(e, d_g, s)$.

□

Remark. To compute roots of multiplicity e, we take derivatives of order $(e-1)$ to reduce to the case of multiplicity one. Taking e-th derivative using Lemma 2.8.2 causes a blow-up of size by a factor of $\text{poly}(e)$. It is unlikely that e-th derivative of a size s circuit can be computed by a $\text{poly}(s, \log e)$ sized circuit, otherwise, VP=VNP [Kal87].

Newton iteration with multiplicity: The following formula is known as *generalized/modified Newton Iteration*, as it works with any multiplicity $e > 0$. When $e = 1$, this is same as classical Newton iteration. See [DSS18] for a proof.

Lemma 5.2.4. *If $f(\boldsymbol{x}, y) = (y-g)^e h$, where $h|_{y=g} \neq 0 \bmod \langle \boldsymbol{x} \rangle$ and $e > 0$, then the power series for g can be approximated by the recurrence:*

$$y_{t+1} := y_t - e \cdot \frac{f}{\partial_y f}\bigg|_{y=y_t} \tag{5.1}$$

where $y_t \equiv g \bmod \langle \boldsymbol{x} \rangle^{2^t}$.

When $e \geq 2$, the denominator $\partial_y f|_{y=y_t}$ is zero $\bmod \langle \boldsymbol{x} \rangle$, thus, its reciprocal does not exist. However, the ratio $(f/\partial_y f)\big|_{y=y_t}$ does exist in $\mathbb{F}[[\boldsymbol{x}]]$.

Can we use modified Newton Iteration to prove factor conjecture? The answer is yes, if we can solve the following problem of modular division [DSS18, Theorem 2]. The *modular*

division problem is to show that if f/g is defined in $\mathbb{F}[[\boldsymbol{x}]]$, where polynomials f and g can be computed by a circuit of size s, then $f/g \bmod \langle \boldsymbol{x} \rangle^d$ can be computed by a circuit of size poly(sd). Note that if g is invertible in $\mathbb{F}[[\boldsymbol{x}]]$, then the question of modular division can be solved using Strassen's division elimination [Str73]. But, here we have to handle the case when g is not invertible in $\mathbb{F}[[\boldsymbol{x}]]$ (though f/g is well-defined). If we shift the numerator and denominator, we would get $f(\boldsymbol{x}+\boldsymbol{a})/g(\boldsymbol{x}+\boldsymbol{a}) \bmod \langle \boldsymbol{x} \rangle^d$, but we do not know how to recover $f(\boldsymbol{x})/g(\boldsymbol{x}) \bmod \langle \boldsymbol{x} \rangle^d$ from it (unless f/g is a polynomial).

To handle factors with high multiplicity, Bürgisser [Bür04] worked with the perturbed polynomial $F(\boldsymbol{x}, y) = f(\boldsymbol{x}, y+\epsilon) - f(0, \epsilon)$ instead of the original polynomial $f(\boldsymbol{x}, y)$. Note that $(\mathbf{0}, 0)$ is a simple root of $F(\boldsymbol{x}, y)$. Now, the partial derivative $\partial_y F(\mathbf{0}, 0)$ does not vanish over $\mathbb{F}(\epsilon)$. So, a power series root of F can be approximated iteratively by Newton iteration. This would give an approximative circuit of the factors. See [Bür04] for the details.

5.3 Factors of arithmetic circuits of low degree radical

Now we prove the main theorem of this chapter. We assume the characteristic of the underlying field to be zero.

Theorem 5.3.1. *Every factor of a polynomial computed by size s circuit has circuits of size polynomially bounded by s and degree of the squarefree part of the polynomial.*

Proof. Suppose $f = \prod_{i=1}^m f_i^{e_i}$ is the complete irreducible factorization of f. The squarefree part of f is $\prod_{i=1}^m f_i$. Now, $\partial_{x_1} f = \prod_{i=1}^m f_i^{e_i-1} u$, where $u = \sum_{i=1}^m e_i f_1 \cdots f_{i-1} \cdot (\partial_{x_1} f_i) \cdot f_{i+1} \cdots f_m$. We multiply this derivative by a new variable z. Now, if we take the polynomial $F := f + z\partial_{x_1} f$, it would be factorized as $\prod_{i=1}^m f_i^{e_i-1}(\prod_{i=1}^m f_i + z \cdot u)$. Note that the factor $G := (\prod_{i=1}^m f_i + z \cdot u)$ has multiplicity one. Also note $\prod_{i=1}^m f_i^{e_i-1}$ and $(\prod_{i=1}^m f_i + z \cdot u)$ are coprime. So, invoking Theorem 5.2.3 on F, the factor G can be computed by a poly(s, d') sized circuit, where d' is the degree of the squarefree part. After finding the circuit of $\prod_{i=1}^m f_i + z \cdot u$, we put z to zero to get a circuit for the squarefree part $\prod_{i=1}^m f_i$. After we get the circuit of the squarefree part, we can compute any irreducible factor of f in

poly(s, d') by factoring the squarefree part using Theorem 5.2.3.

□

Remark. In [DSS18], we proved Theorem 5.3.1 using a different technique that we call allRootsNI (recursive root finding using matrices). In allRootsNI, we simultaneously find the approximations of all the power series roots g_i of $f(\tau(\boldsymbol{x}))$. For all i, from approximations of g_i up to degree $\delta - 1$ (denoted $g_i^{<\delta}$), we calculate approximations of g_i up to degree δ.

Assume, the power series split of the shifted polynomial is: $f(\boldsymbol{x}, y) = \prod_i (y - g_i)^{\gamma_i}$ and d_0 is the degree of the radical. Now, we analyze the logarithmic derivative identity: $(\partial_y f)/f = \sum_i \gamma_i/(y - g_i)$. We reduce the above identity modulo $I^{\delta+1}$, where $I := \langle \boldsymbol{x} \rangle$ and $\mu_i := g_i(\mathbf{0}) \equiv g_i \mod I$.

Now, we would get the following.

$$\frac{\partial_y f}{f} = \sum_{i=1}^{d_0} \frac{\gamma_i}{y - g_i} \equiv \sum_{i=1}^{d_0} \frac{\gamma_i}{y - g_i^{<\delta}} + \sum_{i=1}^{d_0} \frac{\gamma_i \cdot g_i^{=\delta}}{(y - \mu_i)^2} \mod I^{\delta+1}.$$

The above is a linear equation in terms of the d_0 unknowns $g_i^{=\delta}$ (γ_i, μ_i's are known.) By fixing y to d_0 different elements c_i in $\mathbb{F}$, $i \in [d_0]$, we get a linear system with a unique solution for the unknowns. By solving this system, we get all the $g_i^{=\delta}$. From the approximations of these power series roots, we can show that the factors can be computed by poly(s, d_0) sized circuits. See [DSS18, Theorem 1] for the details.

5.4 Discussions

Factor conjecture has interesting implications in algebraic complexity. See [Bür13, Bür04] for some consequences of factor conjecture, especially in relating decision and computation in algebraic complexity.

Suppose a polynomial f has a circuit of size s. Can we say that the squarefree part of f has poly(s) sized circuit [DSS18]? If this is true, the factor conjecture directly follows.

Another question relevant here is about arithmetic circuit complexity of high degree GCD of a set of polynomials. Kaltofen proved [Kal87, Theorem 4] if $f_1, \ldots, f_m \in$

$\mathbb{F}[x_1, \ldots, x_n]$ are polynomials given by a size s arithmetic circuit and g is the GCD of $f_1, \ldots, f_m$ and d be the total degree of g, then g can be computed by size $\text{poly}(s, d)$ arithmetic circuit. Kaltofen asked [Kal87, Problem 4] whether g can be computed by a circuit of size $\text{poly}(s)$. An affirmative answer to this question would show that even high degree GCD of polynomials given by a small-sized circuit can be computed by a small-sized circuit.

There are several other open questions on closure of different classes under factoring. Is it true that all factors of a arithmetic formula of size s can be computed by an arithmetic formula of size polynomially bounded in s? In [DSS18], we show that the reduction to power series root approximation gives a $\text{poly}(s, d^{\log d})$ size upper bound for the formula size of factors of a polynomial of degree d computed by a formula of size s. Finally, it is unknown whether VP over positive characteristic is closed under factors [Bür13]. In fact, the following question is open [KSS15]. Given a circuit of a polynomial $f(x_1, \ldots, x_n)$ over $\mathbb{F}_p$, where $f(x_1, \ldots, x_n) = g^p$, for some polynomial $g(x_1, \ldots, x_n)$, it is not known whether g has a circuit of size $\text{poly}(s, d)$, where d is the degree of f.

Part III

Hitting Set for $\overline{\mathrm{VP}}$

Chapter 6

A PSPACE Construction of Hitting Set for $\overline{\text{VP}}$

Abstract

This chapter is based on joint work with Guo and Saxena [GSS18].

The hitting set construction problem for $\overline{\text{VP}}$ asks to construct a small hitting set for the class of n-variate polynomials of degree d, approximated by size s circuits. In this chapter, we show how this problem can be reduced to a problem, approximate polynomials satisfiability, that can be seen as a generalization of the well-known problem of polynomial system satisfiability (Hilbert's Nullstellensatz). Using basic algebraic geometry, this reduces to the problem of testing if all the annihilators of the polynomials in the system have zero in the constant term. Using a result from [GSS18] that solves the latter problem in PSPACE, we solve the hitting set construction problem for $\overline{\text{VP}}$ in PSPACE. Mulmuley [Mul17] showed that the hitting set construction for $\overline{\text{VP}}$ is in EXPSPACE, whereas for VP it is in PSPACE. Forbes and Shpilka [FS18] showed hitting set construction can be done in PSPACE for $\overline{\text{VP}}$ as well, but their method is analytic and applicable only for fields like rationals or

reals. Our approach is algebraic and applicable to arbitrary fields.

6.1 PSPACE Construction of Hitting Set for VP

First, we discuss the explicit hitting set construction problem for the class VP. Let $\mathcal{C}(n, s, d)$ be the set of all n-variate degree $\leq d$ polynomials computable by size $\leq s$ arithmetic circuits. We want to find a hitting set of poly(n, s) size for this class.

A counting argument (using Schwartz-Zippel-DeMillo-Lipton lemma and an upper bound of number of size s circuits of fan-in 2 over a finite field) shows that hitting sets of size poly(n, s) *exist* for the class $\mathcal{C}(n, s, d)$ over finite fields of size $\geq d^2$. See [For14, Lemma 3.2.14] for a proof. This proof shows a *random* set of size slightly more than s is a hitting set for the class of size s circuits, with high probability. Over infinite fields, there are infinitely many circuits of same size, so the counting argument does not work. Nevertheless, counting and comparing the *dimension* and *degree* of the variety containing the coefficient vectors of the polynomials in the class $\mathcal{C}(n, s, d)$, [HS80] showed a random set (each element picked independently and uniformly at random from a large enough fixed set) of size poly(n, s) is a hitting set for the class $\mathcal{C}(n, s, d)$.

For the hitting-set construction problem, it suffices to focus only on homogeneous polynomials (given a black-box of a polynomial, one can efficiently compute its homogeneous components in a black-box way [SY10]). Homogeneous polynomials are known to be computable by homogeneous circuits, where each gate computes a homogeneous polynomial [SY10].

Given inputs n, s, r (in unary), there is a PSPACE construction [FS18, Mul17] of a set of points of size poly(n, s) and bit complexity bounded by poly(n, s, r), that is guaranteed to be a hitting set for the class of size s homogeneous circuits computing n-variate homogeneous polynomials of degree r. The construction has the following two main steps.

Step 1 (Guess): Enumerate over all possible candidate hitting sets of size bounded by $(ns)^c$ for some large enough constant c (each element of the set is from $[r^2]^n$).

Step 2 (Verify): Check if a candidate set is a hitting set for the class. Is there a circuit of size s computing a nonzero polynomial of degree r that evaluates to zero at all points in the candidate set?

Going over all possible hitting sets can be done in PSPACE. The challenging step is to certify that a candidate set is a hitting set for the given class. If the polynomials are over some finite field, one can check if a set is hitting set by evaluating all polynomials in the class, at all the points in the candidate set. This is not possible for infinite fields like $\mathbb{Q}$ or $\mathbb{C}$. Here we need the idea of universal circuits. An universal circuit $\Psi(\mathbf{y}, \mathbf{x})$ is a homogeneous circuit in n essential variables $\mathbf{x}$ and $s' := O(sr^4)$ auxiliary variables $\mathbf{y}$ of size $O(sr^4)$. By appropriately fixing the auxiliary variables to constants, it can output any homogeneous n-variate polynomial of degree r which is computable by a size s circuits. See [Raz08, SY10] for a universal circuit construction.

Now, we show how the hitting set verification problem can be reduced to polynomial system satisfiability using universal circuits. Given a candidate hitting set $\{\mathbf{x_1}, \ldots, \mathbf{x_m}\}$, we check the existence for solutions (over the closure of the underlying field) for the following system of polynomial equations. For all $i \in [m]$, $\Psi(\mathbf{x_i}, \mathbf{a}) = 0$ and $\Psi(\mathbf{b}, \mathbf{a}) = 1$. If there is a solution $(\mathbf{b}, \mathbf{a})$ to this system, then there is a nonzero polynomial $\Psi(\mathbf{x}, \mathbf{a})$ that vanishes at all points in the candidate hitting set. On the contrary, if the candidate set is indeed a hitting set, there would be no solution of this system. Checking whether a solution exists for a system of polynomials is in PSPACE using an effective version of Hilbert's Nullstellensatz [Kol88] and solving a linear system of exponentially many unknowns and equations.

Finally, note that in the above construction, we actually construct hitting set for a bigger class than the class of size s homogeneous circuits computing n-variate homogeneous polynomials of degree r. We give hitting set for the polynomials that are projections of a $O(sr^4)$ sized universal circuit (fixing the auxiliary variables by all possible constants). Every possible projection of the universal circuit $\Psi(\mathbf{y}, \mathbf{x})$ may not be computable by a size s circuit.

The above construction does not directly generalize to the closure of VP. There are

several issues. Firstly, as noted by Forbes and Shpilka [FS18], a hitting set for a class of polynomials, may not be a hitting set for the closure of that class. Secondly, in the verification step, this approach leads to a system of exponentially many polynomial equations (as the best bound known for converting an approximative circuit to a circuit exactly computing the polynomial is exponential). We show a natural approach to circumvent this problem in the following section.

6.2 PSPACE Construction of Hitting Set for $\overline{\mathrm{VP}}_{\mathbb{A}}$

Assumptions on the field. The results in this section are valid for any field, though we describe it only for the closure of finite fields. Suppose p is a prime. Define $\mathbb{A} := \overline{\mathbb{F}}_p$. We want to find hitting-sets for certain polynomials in $\mathbb{A}[x_1, \ldots, x_n]$. Fix a p-power $q \geq \Omega(sr^6)$, for the given parameters s, r. Assume that $p \nmid (r+1)$. Also, fix a model for the finite field $\mathbb{F}_q$ [AL86].

First, we define the notion of 'infinitesimally approximating' a polynomial by a small circuit.

Approximative closure of VP. [BIZ18] A family $(f_n|n)$ of polynomials from $\mathbb{A}[\mathbf{x}]$ is in the *class* $\overline{\mathrm{VP}}_{\mathbb{A}}$ if there are polynomials $f_{n,i}$ and a function $t : \mathbb{N} \mapsto \mathbb{N}$ such that g_n has a poly(n)-size poly(n)-degree algebraic circuit, over the field $\mathbb{A}(\epsilon)$, computing $g_n(\mathbf{x}) = f_n(\mathbf{x}) + \epsilon f_{n,1}(\mathbf{x}) + \epsilon^2 f_{n,2}(\mathbf{x}) + \ldots + \epsilon^{t(n)} f_{n,t(n)}(\mathbf{x})$. That is, $g_n \equiv f_n \bmod \epsilon\mathbb{A}[\epsilon][\mathbf{x}]$.

The smallest possible circuit size of g_n is called the *approximative complexity* of f_n, namely $\overline{\mathrm{size}}(f_n)$.

Hitting-set for $\overline{\mathrm{VP}}_{\mathbb{A}}$. Given functions $s = s(n)$ and $r = r(n)$, a finite subset $\mathcal{H} \subseteq \mathbb{A}^n$ is called a *hitting-set* for degree-r polynomials of approximative complexity s, if for every such nonzero polynomial f: $\exists \mathbf{v} \in \mathcal{H}$, $f(\mathbf{v}) \neq 0$.

Heintz and Schnorr's [HS80] proof of the existence of small hitting set for VP uses concepts from algebraic geometry, essentially counting dimension and degree of the variety containing coefficient vectors of all low degree polynomials computed by small circuits. As varieties are closed under Zariski topology, it already contains the polynomials in $\overline{\mathrm{VP}}_{\mathbb{A}}$.

Thus, Heintz and Schnorr's [HS80] result directly extend to degree-r $\overline{\text{size}}$-s polynomials.

Lemma 6.2.1. [HS80, Theorem 4.4] *There exists a hitting-set $\mathcal{H} \subseteq \mathbb{F}_q^n$ of size $O(s^2n^2)$ (assuming $q \geq \Omega(sr^2)$) that hits all nonzero degree-r n-variate polynomials in $\mathbb{A}[\mathbf{x}]$ that can be infinitesimally approximated by size-s algebraic circuits.*

Ultimately, we are interested in computing such a hitting-set in poly$(s, \log r, \log q)$-time. Here we give a PSPACE explicit construction.

Note that for the hitting-set construction problem, it suffices to focus only on homogeneous polynomials. They are known to be computable by homogeneous circuits, where each gate computes a homogeneous polynomial [SY10]. We use universal circuits [Raz08, SY10] that can simulate any circuit of size-s computing a degree-r homogeneous polynomial in $\mathbb{A}(\epsilon)[x_1, \ldots, x_n]$. Recall that the *universal circuit* $\Psi(\mathbf{y}, \mathbf{x})$ is a circuit in n essential variables $\mathbf{x}$ and $s' := O(sr^4)$ auxiliary variables $\mathbf{y}$. The variables $\mathbf{y}$ are the ones that one can specialize in $\mathbb{A}(\epsilon)$, to compute a specific polynomial in $\mathbb{A}(\epsilon)[x_1, \ldots, x_n]$. Every specialization gives a homogeneous degree-r $\overline{\text{size}}$-s' polynomial. Moreover, the set of these polynomials is closed under constant multiples [FS18, Theorem 2.2].

Note that there is a hitting-set, with $m := O(s'^2n^2)$ points in $\mathbb{F}_q^n$ ($\because q \geq \Omega(s'r^2)$), for the set of polynomials $\mathcal{P}$ infinitesimally approximated by the specializations of $\Psi(\mathbf{y}, \mathbf{x})$ [HS80].

In Chapter 1, we already defined the problem of approximate polynomials satisfiability (APS). Now, we give a criterion to decide whether a candidate set is a hitting-set by reducing the problem to an APS instance.

Theorem 6.2.2 (hitting-set criterion)**.** *Set $\mathcal{H} =: \{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq \mathbb{F}_q^n$ is* not *a hitting-set for the family of polynomials $\mathcal{P}$ infinitesimally approximated by the specializations of $\Psi(\mathbf{y}, \mathbf{x})$ iff there is a satisfying assignment $(\alpha, \beta) \in \mathbb{A}(\epsilon)^{s'} \times \mathbb{A}(\epsilon)^n$ such that:*

(1) $\forall i \in [n]$, ${\beta_i}^{r+1} - 1 \in \epsilon\mathbb{A}[\epsilon]$, where r is the degree of the specializations of $\Psi(\mathbf{y}, \mathbf{x})$,

(2) $\Psi(\alpha, \beta) - 1 \in \epsilon\mathbb{A}[\epsilon]$, and

(3) $\forall i \in [m]$, $\Psi(\alpha, \mathbf{v}_i) \in \epsilon\mathbb{A}[\epsilon]$.

Remark– The above criterion holds for algebraically closed fields $\mathbb{A}$ of *any* characteristic. Thus, it reduces those hitting-set verification problems to APS as well.

Proof. First we show that: $\exists x \in \mathbb{A}(\epsilon)$, $x^{r+1}-1 \in \epsilon\mathbb{A}[\epsilon]$ implies $x \in \mathbb{A}[[\epsilon]] \cap \mathbb{A}(\epsilon)$ (= rational functions defined at $\epsilon = 0$).

Claim 6.2.3. *$\exists x \in \mathbb{A}(\epsilon)$, $x^{r+1}-1 \in \epsilon\mathbb{A}[\epsilon]$ implies $x \in Z_{r+1} + \epsilon\mathbb{A}[[\epsilon]]$, where Z_{r+1} is the set of $(r+1)$-th roots of unity in $\mathbb{A}$.*

Proof. Recall the formal power series $\mathbb{A}[[\epsilon]]$ and its group of units $\mathbb{A}[[\epsilon]]^*$. Note that for any polynomial $a = \left(\sum_{i_0 \le i \le d} a_i \epsilon^i\right)$ with $a_{i_0} \neq 0$, the inverse $a^{-1} = \epsilon^{-i_0} \cdot \left(\sum_{i_0 \le i \le d} a_i \epsilon^{i-i_0}\right)^{-1}$ is in $\epsilon^{-i_0} \cdot \mathbb{A}[[\epsilon]]^*$. This is just a consequence of the identity $(1-\epsilon)^{-1} = \sum_{i \ge 0} \epsilon^i$. In other words, any rational function $a \in \mathbb{A}(\epsilon)$ can be written as an element in $\epsilon^{-i}\mathbb{A}[[\epsilon]]^*$, for some $i \ge 0$. Thus, write x as $\epsilon^{-i} \cdot (b_0 + b_1\epsilon + \cdots)$ for $i \ge 0$ and $b_0 \in \mathbb{A}^*$. This gives

$$x^{r+1} - 1 \;=\; \epsilon^{-i(r+1)}(b_0 + b_1\epsilon + b_2\epsilon^2 + \cdots)^{r+1} \;-\; 1\,.$$

For this to be in $\epsilon\mathbb{A}[\epsilon]$, clearly i has to be 0 (otherwise, $\epsilon^{-i(r+1)}$ remains uncanceled); implying that $x \in \mathbb{A}[[\epsilon]]$.

Moreover, we deduce that $b_0^{r+1} - 1 = 0$. Thus, condition (1) implies that b_0 is one of the $(r+1)$-th roots of unity $Z_{r+1} \subseteq \mathbb{A}$ (recall that, since $p \nmid (r+1)$, $|Z_{r+1}| = r+1$). Thus, $x \in Z_{r+1} + \epsilon\mathbb{A}[[\epsilon]]$. □

[⇒]: Suppose $\mathcal{H}$ is not a hitting-set for $\mathcal{P}$. Then, there is a specialization $\alpha \in \mathbb{A}(\epsilon)^{s'}$ of the universal circuit such that $\Psi(\alpha, \mathbf{x})$ computes a polynomial in $\mathbb{A}[\epsilon][\mathbf{x}] \setminus \epsilon\mathbb{A}[\epsilon][\mathbf{x}]$, but still 'fools' $\mathcal{H}$, that is, $\forall i \in [m]$, $\Psi(\alpha, \mathbf{v}_i) \in \epsilon\mathbb{A}[\epsilon]$. What remains to be shown is that conditions (1) and (2) can be satisfied too.

Consider the polynomial $g(\mathbf{x}) := \Psi(\alpha, \mathbf{x})\big|_{\epsilon=0}$. It is a nonzero polynomial, in $\mathbb{A}[\mathbf{x}]$ of degree-r, that 'fools' $\mathcal{H}$. By the Schwartz-Zippel-DeMillo-Lipton lemma, there is a $\beta \in Z_{r+1}^n$ such that $a := g(\beta)$ is in $\mathbb{A}^*$. Clearly, $\beta_i^{r+1} - 1 = 0$, for all i. Consider $\psi' := a^{-1} \cdot \Psi(\alpha, \mathbf{x})$. Note that $\psi'(\beta) - 1 \in \epsilon\mathbb{A}[\epsilon]$, and $\psi'(\mathbf{v}_i) \in \epsilon\mathbb{A}[\epsilon]$ for all i. Moreover, the normalized polynomial $\psi'(\mathbf{x})$ can easily be obtained from the universal circuit Ψ by

changing one of the coordinates of α (For example, the incoming wires of the root of the circuit). This means that the three conditions (1)-(3) can be simultaneously satisfied by (some) $(\alpha', \beta) \in \mathbb{A}(\epsilon)^{s'} \times Z_{r+1}^n$.

[⇐]: Suppose the satisfying assignment is $(\alpha, \beta') \in \mathbb{A}(\epsilon)^{s'} \times \mathbb{A}(\epsilon)^n$. As shown in Lemma 6.2.3, condition (1) implies: $\beta'_i \in Z_{r+1} + \epsilon\mathbb{A}[[\epsilon]]$ for all $i \in [n]$. Let us define $\beta_i := \beta'_i\big|_{\epsilon=0}$, for all $i \in [n]$; they are in $Z_{r+1} \subseteq \mathbb{A}$. By Condition (3): $\forall i \in [m]$, $\Psi(\alpha, \mathbf{v}_i) \in \epsilon\mathbb{A}[\epsilon]$.

Since any rational function $a \in \mathbb{A}(\epsilon)$ can be written as an element in $\epsilon^{-i}\mathbb{A}[[\epsilon]]^*$, for some $i \geq 0$, we get that $\Psi(\alpha, \mathbf{x})$ is in $\epsilon^{-j}\mathbb{A}[[\epsilon]][\mathbf{x}]$, for some $j \geq 0$. Expand the polynomial $\Psi(\alpha, \mathbf{x})$, wrt ϵ, as:

$$g_{-j}(\mathbf{x})\epsilon^{-j} + \cdots + \epsilon^{-2}g_{-2}(\mathbf{x}) + g_{-1}(\mathbf{x})\epsilon^{-1} + g_0(\mathbf{x}) + \epsilon g_1(\mathbf{x}) + \epsilon^2 g_2(\mathbf{x}) + \dots .$$

Let us study Condition (2). If for each $0 \leq \ell \leq j$, polynomial $g_{-\ell}(\mathbf{x})$ is zero, then $\Psi(\alpha, \beta')\big|_{\epsilon=0} = 0$ contradicting the condition. Thus, we can pick the largest $0 \leq \ell \leq j$ such that the polynomial $g_{-\ell}(\mathbf{x}) \neq 0$.

Note that the normalized circuit $\epsilon^\ell \cdot \Psi(\alpha, \mathbf{x})$ equals $g_{-\ell}$ at $\epsilon = 0$. This means that $g_{-\ell} \in \mathcal{P}$, and it is a nonzero polynomial fooling $\mathcal{H}$. Thus, $\mathcal{H}$ cannot be a hitting-set for $\mathcal{P}$ and we are done. □

Using Theorem 6.2.2 and assuming Theorem 6.3.3, we get the following result.

Theorem 6.2.4. *There is a PSPACE algorithm that (given input n, s, r in unary & suitably large $\mathbb{F}_q$) outputs a set, of points from $\mathbb{F}_q^n$ of size poly$(nsr, \log q)$, that hits all n-variate degree-r polynomials over $\overline{\mathbb{F}}_q$ that can be infinitesimally approximated by size s circuits.*

Proof. Given a prime p and parameters n, r, s in unary (wlog $p \nmid (r+1)$), fix a field $\mathbb{F}_q$ with $q \geq \Omega(sr^6)$. Fix the universal circuit $\Psi(\mathbf{y}, \mathbf{x})$ with n essential variables $\mathbf{x}$ and $s' := \Omega(sr^4)$ auxiliary variables $\mathbf{y}$. Fix $m := \Omega(s'^2n^2)$.

For every subset $\mathcal{H} =: \{\mathbf{v}_1, \dots, \mathbf{v}_m\} \subseteq \mathbb{F}_q^n$ solve the APS instance described by Conditions (1)-(3) in Theorem 6.2.2. These are $(n + m + 1)$ algebraic circuits of degree poly$(srn, \log p)$ and a similar bitsize. Using the PSPACE algorithm for AnnAtZero test, it can be solved in poly$(srn, \log p)$-space as followed.

The number of subsets $\mathcal{H}$ is q^{nm}. So, in poly$(nm \log q)$-space we can go over all of them. If APS fails on one of them (say $\mathcal{H}$) then we know that $\mathcal{H}$ is a hitting-set for $\mathcal{P}$. Since Ψ is universal, for homogeneous degree-r $\overline{\text{size}}$-s polynomials in $\mathbb{A}[\mathbf{x}]$, we output $\mathcal{H}$ as the desired hitting-set.

□

6.3 APS, Origin in closure and Annihilating at zero

First, we give some notations used. Let $\mathbb{A}$ be the algebraic closure of $\mathbb{F}$. Write $f : \mathbb{A}^n \to \mathbb{A}^m$ for the polynomial map sending a point $x = (x_1, \ldots, x_n) \in \mathbb{A}^n$ to $(f_1(x), \ldots, f_m(x)) \in \mathbb{A}^m$. For a subset S of an affine or projective space, $\overline{S}$ denotes its Zariski closure in that space. We use O to denote the origin $\mathbf{0}$ of an affine space. $\text{Im}(f)$ denote the image of the polynomial map f.

6.3.1 APS and AnnAtZero

The problem $O \in \text{Im}(f)$ is same as checking the existence of a solution $x \in \mathbb{A}^n$ to $f_i = 0$, for $i \in [m]$. It is well known in algebraic geometry that checking whether $O \in \overline{\text{Im}(f)}$ is equivalent to the existence of an "approximate solution" $x \in \mathbb{A}[\epsilon, \epsilon^{-1}]^n$, which is a tuple of Laurent polynomials in a formal variable ϵ.

Theorem 6.3.1 (Approx. wrt ϵ)**.** *$O \in \overline{\text{Im}(f)}$ iff there exists $x = (x_1, \ldots, x_n) \in \mathbb{A}(\epsilon)^n$ such that $f_i(x) \in \epsilon\mathbb{A}[\epsilon]$, for all $i \in [m]$. Moreover, when such x exists, it may be chosen such that*

$$x_i \in \epsilon^{-\Delta}\mathbb{A}[\epsilon] \cap \epsilon^{\Delta'}\mathbb{A}[\epsilon^{-1}] = \left\{ \sum_{j=-\Delta}^{\Delta'} c_j \epsilon^j : c_j \in \mathbb{A} \right\}, \quad i \in [n],$$

where $\Delta := \prod_{i\in[m]} \deg(f_i) > 0$ and $\Delta' := (max_{i\in[m]} \deg(f_i)) \cdot \Delta > 0$.

Its proof is essentially given in [LL89]. See also, [BCS13, Lemma 20.28], [GP18, Page 37:53], [GSS18].

Now, APS can be reduced to testing the constant term of annihilators using the concept of ideal-variety correspondence.

Lemma 6.3.2 (O in the closure)**.** *The constant term of every annihilator for* $\mathbf{f}$ *is zero iff* $O \in \overline{\mathrm{Im}(f)}$.

Proof. Note that: $Q \in \mathbb{A}[Y_1, \ldots, Y_m]$ vanishes on $\mathrm{Im}(f)$ iff $Q(\mathbf{f})$ vanishes on $\mathbb{A}^n$, which holds iff $Q(\mathbf{f}) = 0$, that is, Q is an annihilator for $\mathbf{f}$. So $\overline{\mathrm{Im}(f)} = V(I)$, where the ideal $I \subseteq \mathbb{A}[Y_1, \ldots, Y_m]$ consists of the annihilators for $\mathbf{f}$. Also note that $\{O\} = V(\mathfrak{m})$, where $\mathfrak{m}$ is the maximal ideal $\langle Y_1, \ldots, Y_m \rangle$.

Let us study the condition $O \in \overline{\mathrm{Im}(f)}$. By the ideal-variety correspondence, $\{O\} = V(\mathfrak{m}) \subseteq \overline{\mathrm{Im}(f)} = V(I)$ is equivalent to $I \subseteq \mathfrak{m}$, that is, $Q \bmod \mathfrak{m} = 0$ for $Q \in I$. But $Q \bmod \mathfrak{m}$ is just the constant term of the annihilator Q. Hence, we have the equivalence. □

As an interesting corner case, the above lemma proves that whenever $\mathbf{f}$ are algebraically *independent*, we have $\mathbb{A}^m = \overline{\mathrm{Im}(f)}$. For example, $f_1 = X_1$ and $f_2 = X_1X_2 - 1$. But in the dependent cases, $\mathrm{Im}(f)$ is not necessarily closed in the Zariski topology. Consider the following example. Let $n = 2$, $m = 3$. Consider $f_1 = f_2 = X_1$ and $f_3 = X_1X_2 - 1$. The annihilators are multiples of $(Y_1 - Y_2)$, which means by Lemma 6.3.2 that $O \in \overline{\mathrm{Im}(f)}$. But there is no solution to $f_1 = f_2 = f_3 = 0$, i. e. , $O \notin \mathrm{Im}(f)$.

Is APS equivalent to a projective version of polynomials satisfiability? The following remark answers this question. Given a system of polynomial equations $f_1(x_1, \ldots, x_n) = 0, \ldots, f_m(x_1, \ldots, x_n) = 0$, we can always homogenize f_i's by introducing a new variable z. The question of *projective polynomials satisfiability* for a given system of equations is whether there is a nonzero solution (called a *projective solution*) to the homogenized system. We have seen that APS does not directly reduce to (affine) polynomials satisfiability, as there are unsatisfiable systems, For example, $\{x = 0,\ xy = 1\}$, that have approximate solutions. We note that APS does not directly reduce to projective polynomials satisfiability either: Consider the system $\{x + y = 0,\ x + y = 1\}$, corresponding to two parallel affine lines. It has no approximate solution, but the homogenized system $\{x + y = 0,\ x + y = z\}$ has a projective solution $(1, -1, 0)$.

Nonetheless, it is true that if the equations $f_1 = 0, \ldots, f_m = 0$ have an approximate

solution, then the homogenized equations $\hat{f}_1 = 0, \ldots, \hat{f}_m = 0$ have a projective solution (in $\mathbb{P}^n$). We sketch a proof for this fact: Suppose $a_1, \ldots, a_n \in \overline{\mathbb{F}}(\epsilon) \subseteq \overline{\mathbb{F}}((\epsilon))$ form an approximate solution of the original system. Let $a_{n+1} = 1$. Then $f_i(a_1, \ldots, a_n) = \hat{f}_i(a_1, \ldots, a_n, a_{n+1})$ for $i \in [m]$. Choose the smallest $k \in \mathbb{Z}$ such that $\epsilon^k a_i$ is in the ring of formal power series $\overline{\mathbb{F}}[[\epsilon]]$ for all $i \in [n+1]$. We have $k \geq 0$ as $a_{n+1} = 1$. For $i \in [n+1]$, let $\bar{a}_i \in \overline{\mathbb{F}}$ be the constant term of $\epsilon^k a_i$. Minimality of k guarantees that $\bar{a}_i \neq 0$ for some $i \in [n+1]$. Assigning $\bar{a}_1, \ldots, \bar{a}_{n+1}$ to $x_1, \ldots, x_n, z$ then gives a projective solution to the equations $\hat{f}_1 = 0, \ldots, \hat{f}_m = 0$.

6.3.2 PSPACE algorithm for APS.

The connection with annihilators lead to a PSPACE algorithm for APS. The annihilators of $\mathbf{f}$ constitute a prime ideal of the polynomial ring $\mathbb{F}[y_1, \ldots, y_m]$. This ideal is principal (generated by one polynomial) when trdeg of $\mathbf{f}$ is $m - 1$. This is a classical result in commutative algebra [Mat80, Theorem 47]. See also [Kay09, Lemma 7] for an exposition. In this case, we can decide if the constant term of the minimal annihilator is zero in PSPACE, as the *unique* annihilator (up to scaling) can be computed in PSPACE.

If the transcendence degree of $\mathbf{f}$ is less than $m - 1$, the ideal of the annihilators of $\mathbf{f}$ is no longer principal. Although the ideal is finitely generated, finding the generators of this ideal is computationally very hard. (For example, using Gröbner basis techniques, we can do it in EXPSPACE [DK15, Section 1.2.1].) In this case, can we decide if all the annihilators of $\mathbf{f}$ have constant term zero?

In [GSS18], we give a randomized reduction to the principal ideal case by reducing the number of polynomials from m to $k + 1$. Fix a finite subset $S \subseteq \mathbb{F}$, and choose $c_{i,j} \in S$ at random for $i \in [k+1]$ and $j \in [m]$. For this to work, we need a large enough S and $\mathbb{F}$. Now, for $i \in [k+1]$, let $g_i := \sum_{j=1}^{m} c_{i,j} f_j$. Let $\delta := (k+1)(\max_{i \in [m]} \deg(f_i))^k / |S|$. The following theorem shows that we can work with $g_1, \ldots, g_{k+1}$.

Theorem 6.3.3 (Random reduction)**.** *It holds, with probability $\geq (1 - \delta)$, that*

(1) the transcendence degree of $\mathbb{F}(g_1, \ldots, g_{k+1})/\mathbb{F}$ equals k, and

(2) the constant term of every annihilator for $g_1, \ldots, g_{k+1}$ is zero iff the constant term of every annihilator for $f_1, \ldots, f_m$ is zero.

See [GSS18, Theorem 4.5] for a proof using techniques from algebraic geometry.

Algorithm for APS.

Given an instance $\mathbf{f}$ of APS, we can first find the trdeg k. Fix a subset $S \subseteq \mathbb{A}$ to be larger than $2(k+1)(\max_{i\in[m]} \deg(f_i))^k$ (which can be scanned using only polynomial-space). Consider the points $((c_{i,j} \mid i \in [k+1], j \in [m])) \in S^{(k+1)\times m}$; for each such point define $\mathbf{g} := \left\{g_i := \sum_{j=1}^{m} c_{i,j} f_j \mid i \in [k+1]\right\}$. Compute the transcendence degree of $\mathbf{g}$, and if it is k then solve AnnAtZero for the instance $\mathbf{g}$. Output NO iff some $\mathbf{g}$ failed the AnnAtZero test.

All these steps can be achieved in space polynomial in the input size, using the uniqueness of the annihilator for $\mathbf{g}$ [Mat80, Theorem 47], Perron's degree bound [Pło05] and linear algebra [BvzGH82, Mul87]. [BvzGH82] showed that solving a system of linear equations reduces to computing rank of a matrix and [Mul87] gave a logspace-uniform NC algorithm for computing rank over arbitrary fields. NC is contained in polylogarithmic space. Thereby, solving a linear system (of size exponential wrt input size) is in PSPACE.

6.4 Discussions

The problem of hitting set construction for closure of VP has connections with other natural problems in computational algebraic geometry. As shown in [Mul17], the problem of constructing a normalizing map in Noether's Normalization Lemma (NNL) reduces to that of constructing hitting-sets for $\overline{\text{VP}}$ [Mul17, Theorem 4.5]. Approximate polynomials satisfiability may have further applications to problems in computational algebraic geometry and algebraic complexity. The null-cone problem defined in [BGO$^+$18] and border rank computation of a given tensor (over $\overline{\mathbb{F}}$) can be reduced to an APS instance and, hence, solved in PSPACE by the algorithm of [GSS18]. Another motivation for AnnAtZero or APS comes from the *geometric ideal proof system* in algebraic proof complexity, introduced in [GP18], where they study AnnAtZero for systems of polynomial equations corresponding

to Boolean tautologies. See [GP18, Appendix B] for more details.

We conclude with an open question. Polynomial system satisfiability or Hilbert's Nullstellensatz is in AM over fields of characteristic zero, assuming GRH [Koi96]. Can we solve APS (equivalently, AnnAtZero) in AM for fields of characteristic zero assuming GRH? [Kay09] asked this question for AnnAtZero. It seems plausible that deciding whether the constant term of an annihilator is zero is not as hard as computing the whole annihilator. Using our reduction, any improvement in the complexity of APS, would also improve the complexity status of hitting set construction for $\overline{\text{VP}}$.

Chapter 7

Conclusion

We give here the main open problems related to this thesis.

Algebraic dependence

We give two different criteria for testing algebraic dependence over finite fields. One is in the spirit of the Jacobian criterion, reducing the problem to linear dependence testing. The other uses elementary counting estimates and basic algebraic geometry. It is possible that extending these approaches would lead to better complexity bound for dependence testing over finite fields.

Our result on algebraic dependence testing in AM $\cap$ coAM gives further indication that a randomized polynomial time algorithm for the problem exists, but currently not even a sub-exponential time algorithm is known. Studying the following special case might be helpful to get an idea for designing better algorithms [PSS18].

Given quadratic polynomials $f_1, \ldots, f_n \in \mathbb{F}_2[x_1, \ldots, x_n]$, test if they are algebraically dependent in randomized polynomial time.

Factor conjecture

We give strong evidence that the factor conjecture is likely to be true. It is possible that our ideas in tackling the low degree radical case can be extended to prove the conjecture. Can we transform the input polynomial (without blowing-up the arithmetic circuit size)

with a factor/root of high multiplicity to another polynomial with the same factor/root of multiplicity one? Can we get a size poly(s, d) circuit computing the power series expansion (truncated up to degree d) of the rational function (given by a size s circuit) f/g, where g is not invertible? Positive answers to these questions would prove the factor conjecture.

Approximate Polynomials Satisfiability

As indicated in this thesis, approximate polynomials satisfiability, or equivalently testing zero-membership in the Zariski closure of the image, may have further applications to various problems in computational algebraic geometry and algebraic complexity.

Currently, we do not have good understanding of the closure of VP. We can not rule out the possibility that the closure of VP is same as VP. Better understanding of approximate polynomials satisfiability may lead to better understanding of the closure of VP. We end with an interesting open question.

Can we solve AnnAtZero (or APS) in AM for fields of characteristic zero assuming GRH [Kay09]? This would also imply a better hitting set construction for $\overline{\text{VP}}$.

Bibliography

[AB09] S. Arora and B. Barak. *Computational Complexity – A Modern Approach*. Cambridge University Press, 2009.

[Abh56] Shreeram Abhyankar. Two notes on formal power series. *Proceedings of the American Mathematical Society*, 7(5):903–905, 1956.

[AGS18] Manindra Agrawal, Sumanta Ghosh, and Nitin Saxena. Bootstrapping variables in algebraic circuits. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1166–1179. ACM, 2018.

[AL86] L. M. Adleman and H. W. Lenstra. Finding irreducible polynomials over finite fields. In *STOC*, pages 350–355, 1986.

[AM76] SS Abhyankar and TT Moh. On analytic independence. *Transactions of the American Mathematical Society*, 219:77–87, 1976.

[ASSS16] Manindra Agrawal, Chandan Saha, Ramprasad Saptharishi, and Nitin Saxena. Jacobian Hits Circuits: Hitting Sets, Lower Bounds for Depth-D Occur-k Formulas and Depth-3 Transcendence Degree-k Circuits. *SIAM J. Comput.*, 45(4):1533–1562, 2016. (Special issue for STOC 2011).

[AV08] 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.

[Bab85] László Babai. Trading group theory for randomness. In *Proceedings of the 17th Annual ACM Symposium on Theory of Computing*, STOC '85, pages 421–429. ACM, 1985.

[BCS13] Peter Bürgisser, Michael Clausen, and Amin Shokrollahi. *Algebraic complexity theory*, volume 315. Springer Science & Business Media, 2013.

[BGM06] Elmar Böhler, Christian Glaßer, and Daniel Meister. Error-bounded probabilistic computations between ma and am. *Journal of Computer and System Sciences*, 72(6):1043–1076, 2006.

[BGO+18] Peter Bürgisser, Ankit Garg, Rafael Oliveira, Michael Walter, and Avi Wigderson. Alternating minimization, scaling algorithms, and the null-cone problem from invariant theory. In *9th Innovations in Theoretical Computer Science Conference (ITCS 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

[BIZ18] Karl Bringmann, Christian Ikenmeyer, and Jeroen Zuiddam. On algebraic branching programs of small width. *Journal of the ACM (JACM)*, 65(5):32, 2018.

[BMS13] M. Beecken, J. Mittmann, and N. Saxena. Algebraic Independence and Blackbox Identity Testing. *Inf. Comput.*, 222:2–19, 2013. (Special issue for ICALP 2011).

[BSG13] Eli Ben-Sasson and Ariel Gabizon. Extractors for polynomial sources over fields of constant order and small characteristic. *Theory of Computing*, 9(1):665–683, 2013.

[Bür04] Peter Bürgisser. The complexity of factors of multivariate polynomials. *Foundations of Computational Mathematics*, 4(4):369–396, 2004. (Preliminary version in FOCS 2001).

[Bür13] Peter Bürgisser. *Completeness and reduction in algebraic complexity theory*, volume 7. Springer Science & Business Media, 2013.

[BvzGH82] Allan Borodin, Joachim von zur Gathem, and John Hopcroft. Fast parallel matrix and gcd computations. *Information and Control*, 52(3):241–256, 1982.

[CKS19] Chi-Ning Chou, Mrinal Kumar, and Noam Solomon. Closure of vp under taking factors: a short and simple proof. *arXiv preprint arXiv:1903.02366*, 2019.

[CLO07] David A Cox, John Little, and Donal O'Shea. Ideals, varieties, and algorithms. undergraduate texts in mathematics, 2007.

[CS18] Prerona Chatterjee and Ramprasad Saptharishi. Constructing faithful homomorphisms over fields of finite characteristic. *arXiv preprint arXiv:1812.10249*, 2018.

[Csa75] Laszlo Csanky. Fast parallel matrix inversion algorithms. In *16th Annual Symposium on Foundations of Computer Science (sfcs 1975)*, pages 11–12. IEEE, 1975.

[DGRV11] 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.

[DGW09] 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).

[DK15] Harm Derksen and Gregor Kemper. *Computational Invariant Theory*. Springer, 2015.

[DKSS13] 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. (Conference version in FOCS 2009).

[DL78] Richard A DeMillo and Richard J Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7(4):193–195, 1978.

[DSS18] Pranjal Dutta, Nitin Saxena, and Amit Sinhababu. Discovering the roots: Uniform closure results for algebraic classes under factoring. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1152–1165. ACM, 2018.

[DSY09] Zeev Dvir, Amir Shpilka, and Amir Yehudayoff. Hardness-randomness tradeoffs for bounded depth arithmetic circuits. *SIAM Journal on Computing*, 39(4):1279–1293, 2009. (Conference version in STOC 2008).

[Dut18] Pranjal Dutta. Discovering the roots: Unifying and extending results on multivariate polynomial factoring in algebraic complexity. Master's thesis, Chennai Mathematical Institute, 2018.

[Dvi09] Zeev Dvir. Extractors for varieties. In *Proceedings of the 24th IEEE Conference on Computational Complexity (CCC)*, pages 102–113, 2009.

[Eis13] David Eisenbud. *Commutative Algebra: with a view toward algebraic geometry*, volume 150. Springer Science & Business Media, 2013.

[ER93] Richard Ehrenborg and Gian-Carlo Rota. Apolarity and canonical forms for homogeneous polynomials. *European Journal of Combinatorics*, 14(3):157–181, 1993.

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

[For15] 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.

[FS18] Michael A Forbes and Amir Shpilka. A pspace construction of a hitting set for the closure of small algebraic circuits. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1180–1192. ACM, 2018.

[GMQ16] Joshua A Grochow, Ketan D Mulmuley, and Youming Qiao. Boundaries of vp and vnp. In *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.

[GP18] Joshua A. Grochow and Toniann Pitassi. Circuit complexity, proof complexity, and polynomial identity testing: The ideal proof system. *Journal of the ACM]*, 65(6):37, 2018.

[GS86] Shafi Goldwasser and Michael Sipser. Private coins versus public coins in interactive proof systems. In *Proceedings of the 18th Annual ACM Symposium on Theory of Computing*, STOC '86, pages 59–68. ACM, 1986.

[GSS18] Zeyu Guo, Nitin Saxena, and Amit Sinhababu. Algebraic dependencies and pspace algorithms in approximative complexity. In *Proceedings of the 33rd Computational Complexity Conference*, page 10. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018.

[HS37] 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.

[HS80] Joos Heintz and Claus-Peter Schnorr. Testing polynomials which are easy to compute. In *Proceedings of the twelfth annual ACM symposium on Theory of computing*, pages 262–272. ACM, 1980.

[Isa94] I Martin Isaacs. *Algebra: a graduate course*, volume 100. American Mathematical Soc., 1994.

[Jac41] C. G. J. Jacobi. De determinantibus functionalibus. *J. Reine Angew. Math.*, 22(4):319–359, 1841.

[Kal82] Erich Kaltofen. A polynomial-time reduction from bivariate to univariate integral polynomial factorization. In *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)*, pages 57–64. IEEE, 1982.

[Kal85] 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).

[Kal87] Erich Kaltofen. Single-factor hensel lifting and its application to the straight-line complexity of certain polynomials. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 443–452. ACM, 1987.

[Kal89] Erich Kaltofen. Factorization of polynomials given by straight-line programs. *Randomness and Computation*, 5:375–412, 1989.

[Kay09] N. Kayal. The Complexity of the Annihilating Polynomial. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity (CCC)*, pages 184–193, 2009.

[Kem10] Gregor Kemper. *A course in Commutative Algebra*, volume 256. Springer Science & Business Media, 2010.

[KI03] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 355–364. ACM, 2003.

[Kna07] Anthony W Knapp. *Advanced algebra*. Springer Science & Business Media, 2007.

[Koi96] Pascal Koiran. Hilbert's Nullstellensatz is in the polynomial hierarchy. *Journal of complexity*, 12(4):273–286, 1996.

[Koi97] Pascal Koiran. Randomized and deterministic algorithms for the dimension of algebraic varieties. In *Proceedings 38th Annual Symposium on Foundations of Computer Science*, pages 36–45. IEEE, 1997.

[Kol88] János Kollár. Sharp effective Nullstellensatz. *Journal of the American Mathematical Society*, 1(4):963–975, 1988.

[KR05] Martin Kreuzer and Lorenzo Robbiano. *Computational commutative algebra 2*. Springer Science & Business Media, 2005.

[KS06] Neeraj Kayal and Nitin Saxena. Complexity of ring morphism problems. *Computational Complexity*, 15(4):342–390, 2006.

[KS17] Mrinal Kumar and Shubhangi Saraf. Arithmetic circuits with locally low algebraic rank. *Theory of Computing*, 13(1):1–33, 2017.

[KSS15] Swastik Kopparty, Shubhangi Saraf, and Amir Shpilka. Equivalence of polynomial identity testing and polynomial factorization. *computational complexity*, 24(2):295–331, Jun 2015.

[KST19] Mrinal Kumar, Ramprasad Saptharishi, and Anamay Tengse. Near-optimal bootstrapping of hitting sets for algebraic circuits. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 639–646. Society for Industrial and Applied Mathematics, 2019.

[KT90] Erich Kaltofen and Barry M Trager. Computing with polynomials given by black boxes for their evaluations: Greatest common divisors, factorization, separation of numerators and denominators. *Journal of Symbolic Computation*, 9(3):301–320, 1990.

[Lin85] Bernt Lindström. On the algebraic characteristic set for a class of matroids. *Proceedings of the American Mathematical Society*, 95(1):147–151, 1985.

[LL89] Thomas Lehmkuhl and Thomas Lickteig. On the order of approximation in approximative triadic decompositions of tensors. *Theoretical Computer Science*, 66(1):1–14, 1989.

[LN94] Rudolf Lidl and Harald Niederreiter. *Introduction to finite fields and their applications*. Cambridge university press, 1994.

[LN97] Rudolph Lidl and Harald Niederreiter. *Finite Fields*. Cambridge University Press, Cambridge, UK, 1997.

[LS78] Richard J Lipton and Larry J Stockmeyer. Evaluation of polynomials with super-preconditioning. *Journal of Computer and System Sciences*, 16(2):124–139, 1978.

[L'v84] M.S. L'vov. Calculation of invariants of programs interpreted over an integrality domain. *Cybernetics and Systems Analysis*, 20:492–499, 1984.

[Mah14] Meena Mahajan. Algebraic complexity classes. In *Perspectives in Computational Complexity*, pages 51–75. Springer, 2014.

[Mat80] Hideyuki Matsumura. *Commutative Algebra*. Benjamin-Cummings Pub Co, 2nd edition edition, 1980.

[Mit13] Johannes Mittmann. *Independence in Algebraic Complexity Theory*. PhD thesis, Universitäts-und Landesbibliothek Bonn, 2013.

[MSS14] 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.

[Mul87] Ketan D. Mulmuley. A fast parallel algorithm to compute the rank of a matrix over an arbitrary field. *Combinatorica*, 7(1):101–104, 1987.

[Mul12] Ketan D. Mulmuley. Geometric complexity theory V: Equivalence between blackbox derandomization of polynomial identity testing and derandomization of Noether's normalization lemma. In *FOCS*, pages 629–638, 2012.

[Mul17] Ketan Mulmuley. Geometric complexity theory V: Efficient algorithms for Noether normalization. *Journal of the American Mathematical Society*, 30(1):225–309, 2017.

[Oli16] Rafael Oliveira. Factors of low individual degree polynomials. *Computational Complexity*, 25(2):507–561, 2016. (Special issue for CCC 2015).

[Oxl06] James G Oxley. *Matroid theory*, volume 3. Oxford university press, 2006.

[Per27] O. Perron. *Algebra I (Die Grundlagen)*. W. de Gruyter, Berlin, 1927.

[Pło05] A. Płoski. Algebraic Dependence of Polynomials After O. Perron and Some Applications. In *Computational Commutative and Non-Commutative Algebraic Geometry*, pages 167–173. IOS press, 2005.

[PSS18] Anurag Pandey, Nitin Saxena, and Amit Sinhababu. Algebraic independence over positive characteristic: New criterion and applications to locally low-algebraic-rank circuits. *computational complexity*, 27(4):617–670, 2018.

[Raz08] 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.

[Ros15] Zvi H Rosen. *Algebraic Matroids in Applications*. PhD thesis, University of California, Berkeley, 2015.

[Sax06] Nitin Saxena. *Automorphisms of rings and applications to complexity*. PhD thesis, PhD thesis, Indian Institute of Technology Kanpur, 2006.

[Sch80] J.T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, 1980.

[Sch88] Uwe Schöning. Graph isomorphism is in the low hierarchy. *Journal of Computer and System Sciences*, 37(3):312–323, 1988.

[Sch95] Joachim Schmid. On the affine Bezout inequality. *manuscripta mathematica*, 88(1):225–232, 1995.

[Str73] Volker Strassen. Vermeidung von divisionen. *Journal für die reine und angewandte Mathematik*, 264:184–202, 1973.

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends® in Theoretical Computer Science*, 5(3–4):207–388, 2010.

[Tei36] Oswald Teichmüller. Differentialrechnung bei charakteristik p. *Journal für die reine und angewandte Mathematik*, 175:89–99, 1936.

[Tra98] William Nathaniel Traves. *Differential operators and Nakai's conjecture*. PhD thesis, National Library of Canada= Bibliothèque nationale du Canada, 1998.

[VzG84] Joachim Von zur Gathen. Hensel and newton methods in valuation rings. *Mathematics of Computation*, 42(166):637–661, 1984.

[VZGG13] Joachim Von Zur Gathen and Jürgen Gerhard. *Modern computer algebra*. Cambridge university press, 2013.

[Zip79] Richard Zippel. *Probabilistic algorithms for sparse polynomials*. Springer, 1979.

[ZS75] Oscar Zariski and Pierre Samuel. Commutative algebra. vol. 1. with the cooperation of is cohen. corrected reprinting of the 1958 edition. *Graduate Texts in Mathematics*, 28, 1975.

Index