Electronic Colloquium on Computational Complexity, Report No. 84 (2025)

Closure under factorization from a result of Furstenberg

Somnath Bhattacharjee * Mrinal Kumar † Shanthanu S. Rai†
Varun Ramanathan† Ramprasad Saptharishi† Shubhangi Saraf ‡

Abstract

We show that algebraic formulas and constant-depth circuits are *closed* under taking factors. In other words, we show that if a multivariate polynomial over a field of characteristic zero has a small constant-depth circuit or formula, then all its factors can be computed by small constant-depth circuits or formulas respectively.

Our result turns out to be an elementary consequence of a fundamental and surprising result of Furstenberg from the 1960s, which gives a non-iterative description of the power series roots of a bivariate polynomial. Combined with standard structural ideas in algebraic complexity, we observe that this theorem yields the desired closure results.

As applications, we get alternative (and perhaps simpler) proofs of various known results and strengthen the quantitative bounds in some of them. This includes a unified proof of known closure results for algebraic models (circuits, branching programs and VNP), an extension of the analysis of the Kabanets-Impagliazzo hitting set generator to formulas and constant-depth circuits, and a (significantly) simpler proof of correctness as well as stronger guarantees on the output in the subexponential time deterministic algorithm for factorization of constant-depth circuits from a recent work of Bhattacharjee, Kumar, Ramanathan, Saptharishi & Saraf.

*University of Toronto, Canada. Email: somnath.bhattacharjee@mail.utoronto.ca. Research partially supported by an NSERC Discovery Grant

†Tata Institute of Fundamental Research, Mumbai, India. Email: {mrinal, shanthanu.rai, varun.ramanathan, ramprasad}@tifr.res.in. Research supported by the Department of Atomic Energy, Government of India, under project number RTI400112, and in part by Google and SERB Research Grants.

‡University of Toronto, Canada. Email: shubhangi.saraf@utoronto.ca. Research partially supported by the McLean Award and an NSERC Discovery Grant.

ISSN 1433-8092

Contents

1 Introduction

This paper studies the following fundamental question — do all factors of "succinctly represented" polynomial have "succinct representations"? The answer to this question could depend on the particular model of representation. For instance, if the size of the representation is just the sum of monomials, then there are classical examples to show that s-sparse polynomials could have factors that are $s^{\Omega(\log s)}$ sparse. Are other "natural models" of computation for polynomials *closed* under taking factors?

The decade of 1980s witnessed remarkable progress for this problem for the representation of general algebraic circuits. A sequence of results [Kal82, Kal85, GK85], culminating in the celebrated results of Kaltofen [Kal89] and Kaltofen & Trager [KT88], showed that for any n-variate polynomial of degree d computable by a size s algebraic circuit, all its factors have algebraic circuits of size $\text{poly}(s, d, n)$ as well. Not only that, there are randomized algorithms that take a circuit for f and output circuits for all the irreducible factors of f together with their multiplicities, in $\text{poly}(s, d, n)$ time.[1] The fact that (low degree) algebraic circuits have this highly non-trivial property is perhaps one of the strongest pieces of evidence of this model being innately natural when studying computational questions about polynomials.

Do similar "closure" results hold for other natural subclasses of algebraic circuits? Indeed, such closure results for algebraic models under polynomial factorization appear to be rare. For instance, even though the last three decades or so of research in algebraic complexity has brought intense focus on the study of algebraic models, the only models where we know such closure results are for the class VNP (Chou, Kumar & Solomon [CKS19]), bounded individual degree constant-depth circuits (works of Dvir, Shpilka & Yehudayoff [DSY09] and Oliveira [Oli16]), bounded individual degree sparse polynomials [BSV20] and algebraic branching programs (Sinhababu & Thierauf [ST20]).

In addition to their considerable inherent interest, closure results for polynomial factorization for various algebraic models are also closely tied to the questions of hardness-randomness trade-offs for these algebraic models, as well as the complexity of derandomizing polynomial factorization for these models. For example, a fundamental result of Kabanets & Impagliazzo [KI04] shows that sufficiently strong lower bounds for algebraic circuits for explicit polynomial families implies quasipolynomial time deterministic PIT algorithms for these circuits. This result crucially relies on the closure of algebraic circuits under factorization, and thus is not readily applicable for models such as formulas or constant-depth circuits. If these models were indeed closed under taking factors, then we perhaps *only* need strong enough lower bounds for these models to derandomize PIT for these models.

Similarly, a result of Kopparty, Saraf & Shpilka [KSS15] shows that given a deterministic

[1] Strictly speaking, we need the underlying field to be rationals for this version of the result, but something very similar is true for finite fields as well.

PIT algorithm for algebraic circuits, one can derandomize Kaltofen's factorization algorithm for algebraic circuits. The main technical bottleneck for extending such a connection to other models such as formulas or constant-depth circuits is again the absence of a closure result for these models!

In the last few years, we have had partial progress on showing such closure results for models like formulas and constant-depth algebraic circuits, and indeed even these partial results have led to extremely interesting consequences for derandomizing PIT and polynomial factorization for these models. Specifically, Chou, Kumar & Solomon [CKS19] showed low (but growing) degree factors of polynomials with small constant-depth circuits have non-trivially small constant-depth circuits, and used this to conclude that if we had superpolynomial lower bounds for constant-depth circuits, we would get subexponential time deterministic PIT for such circuits. Quite remarkably, such lower bounds were proved by Limaye, Srinivasan & Tavenas [LST21] a few years ago, and these lower bounds yielded non-trivial deterministic PIT for constant-depth circuits due to the results in [CKS19]. Similarly, in recent years, we have seen steady progress on the question of derandomizing polynomial factorization for constant-depth circuits [KRS23, KRSV24, DST24], including a recent result of Bhattacharjee, Kumar, Ramanathan, Saptharishi & Saraf [BKR+25] that gives deterministic subexponential time algorithms for factorization of constant-depth circuits over fields of characteristic zero. Once again, this result crucially relies on both the results and the techniques in the partial closure result of Chou, Kumar & Solomon [CKS19].

Thus, while the question of closure under factorization of models like formulas and constant-depth circuits has remained open, partial progress on this problem, e.g. in [CKS19] has already had some fascinating consequences. Given this, it seems conceivable that complete closure-under-factorization results for these models would not only be of inherent interest on their own, they might also yield quantitative improvements for some of the aforementioned applications.

1.1 Our results

The main result in this paper is that constant-depth algebraic circuits and algebraic formulas are closed under taking factors over fields of zero or sufficiently large characteristic. More formally, we have the following theorem.

Theorem 1.1 (Closure under factorization)**.** *Let $\mathbb{F}$ be a field of characteristic zero, f be a polynomial on n variables of degree d over $\mathbb{F}$ and g be a factor of f. Then, the following are true.*

If f can be computed by an algebraic circuit of size s and depth Δ over $\mathbb{F}$, then g can be computed by an algebraic circuit of size $\mathrm{poly}(s, d, n)$ and depth $\Delta + O(1)$ over $\mathbb{F}$.

If f can be computed by an algebraic formula of size s over $\mathbb{F}$, then g can be computed by an algebraic formula of size $\mathrm{poly}(s, d, n)$ over $\mathbb{F}$.

Remark 1.2. *While the statement above is stated for fields of characteristic zero, it is also true for fields of sufficiently large characteristic (depending on the degree d) via the same proof.* ◊

Over fields of small positive characteristic, we have the following weaker version of the above theorem.

Theorem 1.3 (Closure over finite fields of positive characteristic)**.** *Let $\mathbb{F}_q$ be a field of positive characteristic p, f be an n variate polynomial of degree d over $\mathbb{F}_q$ and g be a factor of f, such that the largest power of g that divides f is $p^\ell \cdot e$ where $\gcd(p, e) = 1$.*

If f can be computed by an algebraic circuit of size s and depth Δ over $\mathbb{F}_q$, then g^{p^ℓ} can be computed by an algebraic circuit of size $\mathrm{poly}(s, d, n)$ and depth $\Delta + O(1)$ over the algebraic closure of $\mathbb{F}_q$.

If f can be computed by an algebraic formula of size s over $\mathbb{F}_q$, then g^{p^ℓ} can be computed by an algebraic formula of size $\mathrm{poly}(s, d, n)$ over the algebraic closure of $\mathbb{F}_q$.

Remark 1.4. *Theorem 1.3 is weaker than Theorem 1.1 in two aspects - (a) we only have a circuit/formula for a power of the factor g and not g itself and (b) the circuit/formula for the power of g is over the algebraic closure of the base field.*

For the rest of the paper, we just focus on the case of fields of characteristic zero. The same ideas essentially extend to the case of fields of positive characteristic with small technical changes and we discuss this case in the appendix for completeness. ◊

To an extent, Theorem 1.1 answers some very natural open questions asked in recent years in the polynomial factorization literature (over fields of sufficiently large or zero characteristic). This includes the question of natural subclasses of algebraic circuits being closed under taking factors (Questions 1.2 and 3.1 in [FS15] and also an open question in [KSS15]) and the question of proving a non-trivial upper bound on the complexity of the factors of sparse polynomials in any algebraic model (Question 1.3 in the survey of Forbes & Shpilka [FS15]). As we discuss in more detail in Section 1.1.1, when combined with the ideas in the recent work of Bhattacharjee et al. [BKR+25], Theorem 1.1 gives an efficient and deterministic reduction from the question of deterministic factoring and in particular deterministic irreducibility testing of polynomials computed by formulas and constant-depth circuits to the question of blackbox deterministic PIT for formulas and constant-depth circuits respectively. The general question of relationship between derandomization of polynomial factorization for algebraic models and PIT (both in the whitebox and blackbox settings) for them was mentioned as an open problem (Question 4.1) in [FS15]. For general algebraic circuits, such a result was shown by Kopparty, Saraf & Shpilka [KSS15].

As alluded to in the introduction, since the closure result for general circuits has many interesting applications, perhaps one can expect some applications of the closure result in Theorem 1.1. This indeed turns out to be the case.[2] We now discuss these applications.

[2] As we will see in the proofs, many of these applications are in fact a consequence of the intermediate statements in the proof of Theorem 1.1.

1.1.1 Applications

A unified proof of closure results: The proof of Theorem 1.1 is essentially a single unified proof of most of the closure results for factorization that we know. The proof extends over all algebraic models with some simple properties — models should support operations such as taking products and sums, extracting homogeneous components, interpolation, etc., without significant cost. Algebraic circuits, branching programs, formulas, constant-depth circuits and exponential sums over algebraic circuits (polynomials in the class VNP) are robust enough to satisfy these properties, and hence closure for them follows from the proof of Theorem 1.1. We stress the fact that almost nothing changes in the argument as we try to infer the closure of these models under factorization.

In addition to these closure results, we also get an alternative and perhaps slightly simpler proof that shows that factors of degree d of a size s circuit (of potentially exponential degree) is in the border of a circuit of size $\mathrm{poly}(s, d)$. This border version of Kaltofen's factor conjecture was originally proved by Bürgisser [Bür04]. The proof in this paper seems to differ from the original proof conceptually, and the appearance of the notion of border complexity here happens fairly naturally.

Improved hardness-randomness trade-offs: Theorem 1.1 immediately implies that the hardness-randomness trade-off of Kabanets and Impagliazzo [KI04] also holds for models like formulas and constant-depth circuits. In particular, exponential lower bounds for constant-depth circuits for explicit polynomial families imply quasipolynomial time deterministic PIT for constant-depth circuits.

Previously, only weaker statements of this form were known. Chou, Kumar & Solomon [CKS19] showed that hardness of low-degree explicit polynomial families for constant-depth circuits gives non-trivial PIT for such circuits, and Andrews & Forbes [AF22] constructed an alternative way of using hardness of the symbolic determinant to get non-trivial PIT for these circuits. It is unclear to us if either of these routes implies a quasipolynomial time deterministic PIT, when the hardness assumption is somewhat stronger (and yet weak enough that we do not get hardness for general algebraic circuits from the depth reduction results). For instance, one concrete conclusion that can be obtained from Theorem 1.1 here is that if we have an explicit n-variate degree n polynomial family P_n, such that any depth Δ circuit for P_n has size n^{n^ε} for any $\varepsilon > 0$, then we have deterministic quasipolynomial time PIT for circuits of size $\mathrm{poly}(n)$ and depth $\Delta + O(1)$.

Deterministic factorization of constant-depth circuits: A recent work of Bhattacharjee et al. [BKR$^+$25] gave a deterministic subexponential time algorithm for factoring constant-depth circuits. Theorem 1.1 and the techniques therein improve the results in [BKR$^+$25] in a few aspects.

- *A simpler and modular proof of correctness* - the proof of correctness of the algorithms in

[BKR+25] turns out to be fairly technical. At a high level, [KSS15] show that Kaltofen's result can be derandomized if we can solve PIT for certain identities that are built using the various factors of the circuit. Even though it was not known that factors of constant-depth circuits have constant-depth circuits, [BKR+25] managed to show that the hitting set generator obtained by combining the lower bounds in [LST21] and the hardness-randomness trade-offs in [CKS19] (which are typically intended to be used against constant-depth circuits) do preserve nonzeroness of these identities that do not appear to be constant depth. This part of the proof relies on finer details of the structure of power series roots obtained via Newton Iteration, and the structure of the specific hitting set generator.

The techniques in the proof of Theorem 1.1 give a clean and direct proof of the correctness of the algorithms in [BKR+25], and essentially demystefy and give a more satisfying reason for why the [LST21] plus [CKS19] hitting set generator works in the algorithm in [BKR+25]—factors of constant-depth circuits *are* indeed constant depth, and hence so are the relevant identities involved in the above sketch. Thus, any hitting set generator for constant-depth circuits will preserve irreducibility and factorization pattern of such circuits as well.

- *Constant-depth circuits for factors as output* - The algorithm in [BKR+25] outputs polynomial size circuits for the irreducible factors of the input circuit. However, these output circuits need not be of constant depth (naturally, since it was not even known if there exist constant-depth circuits for them). Theorem 1.1 implies that these output polynomials have small constant-depth circuits. We show that the ideas in the proof of Theorem 1.1 can be combined with the algorithm in [BKR+25] to output constant-depth circuit for all the factors.

Additionally, stronger lower bounds for constant-depth circuit would translate to faster deterministic algorithms for polynomial identity testing, and this, in turn, would translate to faster deterministic factorization algorithms for constant-depth circuits.

Blackbox PIT and deterministic factorization: A result of Kopparty, Saraf and Shpilka [KSS15] showed that deterministic PIT algorithms for general circuits (in both the blackbox and the whitebox models) implies a deterministic factorization algorithm for such circuits. We can now extend this connection in the blackbox setting to models such as constant-depth circuits or formulas. However, at the moment, it is unclear to us if such a conclusion also follows from whitebox derandomization of PIT for these models.

Randomized algorithms for factorization: By sampling random points instead of using a deterministic PIT algorithm in the aforementioned reduction, we also obtain an efficient randomized algorithm that takes a polynomial with a small formula / constant-depth circuit as input and outputs small formulas / constant-depth circuits for all the irreducible factors of the input, along with their multiplicities.

We now move on to a discussion of the main techniques.

1.2 Proof overview

The difficulty of proving closure for weaker models

Before discussing the main techniques in the proof of Theorem 1.1, we start with a brief discussion about the technical difficulty of proving close results for models like constant-depth circuits and formulas.

One of the main ingredients of all the multivariate factorization algorithms and closure results is the notion of Newton Iteration or one of its variants (e.g. Hensel Lifting). For this discussion, we confine ourselves to Newton iteration. Let us assume that the input polynomial P is of the form $P(\mathbf{x}, y) = (y - f(\mathbf{x})) \cdot Q(\mathbf{x}, y)$ for some unknown f and Q, with $f(\mathbf{0}) = 0$ and $\partial_y P(\mathbf{0}, 0) \neq 0$. Let $f_0(\mathbf{x}) = f(\mathbf{0})$. We first start by observing that $f_0(\mathbf{x}) = 0$ satisfies $P(\mathbf{x}, f_0(\mathbf{x})) = 0 \bmod \langle \mathbf{x} \rangle$ and iteratively define

$$f_{i+1} = f_i - \frac{P(\mathbf{x}, f_i)}{\partial_y P(\mathbf{0}, 0)} \bmod \langle \mathbf{x} \rangle^{i+1},$$

and show that f_i satisfies $P(\mathbf{x}, f_i(\mathbf{x})) = 0 \bmod \langle \mathbf{x} \rangle^{i+1}$, and (in some sense) is unique. Once i exceeds $\deg(f)$, then uniqueness would guarantee that $f_i(\mathbf{x})$ is (essentially) $f(\mathbf{x})$.

This is a strategy that works for general circuits, but since this process is quite sequential and involves successive composition of P with itself, we were unable to show that f_i's were computable by constant-depth circuits even if P was. Similar issues also arise in algorithms that use Hensel lifting. Almost all closure results [Kal89, KT88, Bür04, DSY09, KSS15, Oli16, CKS19, ST20, DSS22] follow the above overall sketch.

One way of getting around these issues would be to argue about the structure of these approximations of these power series roots directly without relying on the explicit iterative process used to construct them. This is essentially how the proof of Theorem 1.1 proceeds. We now discuss this in more detail.

Main ideas

All our results stem from the following fundamental (and surprising) result of Furstenberg from the 1960s, that essentially gives a "closed form" expression for the Newton iteration process, which we state now. The following statement is a special case of Theorem 3.1 for roots of multiplicity 1.

Theorem 1.5 ([Fur67]). *Let $\mathbb{F}$ be an arbitrary field and let $P(t, y) \in \mathbb{F}[t, y]$ be a polynomial and $\varphi(t) \in \mathbb{F}[\![t]\!]$*

with $\varphi(0) = 0$ be a power series satisfying $P(t,y) = (y - \varphi(t)) \cdot Q(t,y)$ and $Q(0,0) \neq 0$. Then

$$\varphi = \mathscr{D}\left(\frac{y^2 \cdot \partial_y P(ty,y)}{P(ty,y)}\right),$$

where, the diagonal *operator $\mathscr{D}$ operates on a bivariate power series $F(t,y) = \sum_{i\geq 0, j\geq 0} F_{i,j} t^i y^j$ as follows:*

$$\mathscr{D}(F)(t) := \sum_{i\geq 0} F_{i,i} \cdot t^i.$$

Semantically, under some mild conditions, the above theorem *almost* gives a way of writing a power series root of a bivariate polynomial P as a ratio of two polynomials whose complexity is close to that of the complexity of P. Here, the *almost* part hides the complexity of computing the diagonal of a power series.

It is worth stressing that the proof of the above result is completely elementary, and a full proof is provided in Section 3 for completeness. Also, via standard transformations, the above can be simplified to the following corollary.

Corollary 1.6. *Let $\mathbb{F}$ be an arbitrary field and let $P(t,y) \in \mathbb{F}[t,y]$ be a polynomial and $\varphi(t) \in \mathbb{F}[\![t]\!]$ with $\varphi(0) = 0$ be a power series satisfying $P(t,y) = (y - \varphi(t)) \cdot Q(t,y)$ with $Q(0,0) = 1$. Then*

$$\begin{aligned}\varphi(t) &= \sum_{m\geq 1} [y^{m-1}] \left\{(1 - \partial_y P(t,y)) \cdot (y - P(t,y))^m\right\} \\ &= \sum_{m\geq 1} \frac{1}{m} \cdot [y^{m-1}] \left\{(y - P(t,y))^m\right\} \quad \textit{(over char. 0 fields)}\end{aligned}$$

where $[y^a]\{G(t,y)\}$ refers to the coefficient of y^a in the polynomial $G(t,y)$.

Remark. *A reader familiar with techniques in enumerative combinatorics / generating functions might notice similarities with the classical* Lagrange inversion formula*, and that is indeed the case. The above can also be derived from the Lagrange inversion formula. A more elaborate discussion on this connection is provided in Section 1.2.1.* ◊

From Corollary 1.6, it almost immediately follows that truncations of power series roots of formulas and constant-depth circuits have small formulas and constant-depth circuits respectively over the closure of the base field. To go from power series roots (over the field closure) to general irreducible factors (over the base field) and to do so within constant depth (or formulas) requires some new observations on combining power series roots to get general irreducible factors. In particular, we rely on the fact that the transformation between elementary symmetric and power symmetric polynomials can be done within constant depth over fields of characteristic zero (or sufficiently large characteristic), as recently shown and crucially used in a work of Andrews and Wigderson [AW24].

Given the simplicity of ideas in the proofs in this paper, perhaps the main contribution of this work is to bring Theorem 1.5 to the attention of a theoretical computer science audience, and to notice its connections and consequences for some very natural questions in multivariate polynomial factorization and its applications.

1.2.1 Connections to the Lagrange Inversion Formula

Power series that are implicitly defined via functional equations have been a subject of intense study, especially in the area of enumerative combinatorics. A classical functional equation in this context is the following — given a power series $g(y)$, find a power series $\varphi(x)$ with $\varphi(0) = 0$ that satisfies the equation $x \cdot g(\varphi(x)) = \varphi(x)$. The Lagrange Inversion Formula (from the 18th century!), for formal power series, states that

$$\varphi(x) = \sum_{m \geq 1} \frac{1}{m} \cdot x^m \cdot [y^{m-1}] \{g(y)^m\}$$

is a solution to the above (and is also unique for nonzero g). (See [SW23] for a simple proof and its applications in enumerative combinatorics.)

To get this closer to the setup for approximate roots, suppose $P(x, y)$ is a polynomial with $\partial_y P(0, 0) = 1$. Then, if $\varphi(x)$ is a power series root satisfying $\varphi(0) = 0$ and $P(x, \varphi(x)) = 0$, then we have that $G(x, \varphi(x)) = \varphi(x)$ where $G(x, y) = y - P(x, y)$, which is very similar to the functional equation $x \cdot g(\varphi(x)) = \varphi(x)$ above. Unsurprisingly, the Lagrange Inversion Formula can be used to derive a closed form expression for $\varphi(x)$. In fact, this precise question is explicitly stated as an exercise[3] in Stanley's book [Sta99] on Enumerative Combinatorics!

Theorem 1.7 (Exercises 5.59 in [Sta99])**.** *Let $\mathbb{F}$ be a field of characteristic 0. Suppose $\varphi(x) \in \mathbb{F}[\![x]\!]$ with $\varphi(0) = 0$. Let $G(x, y) \in \mathbb{F}[\![x, y]\!]$ and φ satisfies the functional equation $G(x, \varphi) = \varphi$. Then,*

$$\varphi(x) = \sum_{m \geq 1} \frac{1}{m} \cdot [y^{m-1}] \{G(x, y)^m\}$$

(where $[y^a] \{P\}$ refers to the coefficient of y^a in P).

The above result appears to have been discovered several times (see [Sok09, Ges16] and references within for several avatars of the above statement) in the enumerative combinatorics literature with a different set of motivations and applications in mind, and appears to have evaded the gaze of the algebraic complexity theorists. The fact that the above can also be derived from Furstenberg's identity was also observed by Hu [Hu16]. Corollary 1.6 is an immediate consequence of the above by setting $G(x, y) = y - P(x, y)$.

[3] The book also provides solutions. For this setting, one could solve for $\varphi(t, x)$ satisfying $t \cdot G(x, \varphi(t, x)) = \varphi(t, x)$ via the Lagrange Inversion Formula to obtain $\varphi(t, x) = \sum_{m \geq 1} \frac{1}{m} \cdot t^m \cdot [y^{m-1}] \{G(x, y)^m\}$, and set $t = 1$.

Organization

The rest of this paper is organized as follows. We begin with some preliminaries in algebraic complexity in Section 2 (readers familiar with standard notions in algebraic complexity can safely skip this section). Section 3 presents the theorem of Furstenberg and alternate formulations, and Section 4 uses these to prove the structural closure results for power series roots and factors. Section 5 gives deterministic algorithms for computing factors of constant-depth circuits (and other natural subclasses of algebraic circuits). Section 6 presents other applications to some known structural results in the context of factorization. We discuss the extension of the closure results to finite fields of small characteristic in Appendix A.

For readers familiar with algebraic complexity, we suggest starting directly with Section 3, and referring to the preliminaries in Section 2 as and when necessary.

2 Preliminaries

Notation:

- We use bold-face letters (such as $\mathbb{F}, \mathbb{K}$) to denote fields. We use $\mathbb{F}[x]$ to denote the polynomial ring, $\mathbb{F}[\![x]\!]$ to the ring of formal power series, and $\mathbb{F}(\!(x)\!)$ refer to ring of Laurent series with respect to the variable x with coefficients from the field $\mathbb{F}$. We use $\overline{\mathbb{F}}$ to refer to the algebraic closure of the field $\mathbb{F}$.
- We use boldface letters such as $\mathbf{x}$ to refer to an order tuple of variables such as $(x_1, \ldots, x_n)$. The size of the tuple would usually be clear from context.
- For a polynomial $f(x) \in \mathbb{F}[x]$ (or more generally in $\mathbb{F}(\!(x)\!)$) and a monomial x^n, we use $[x^n]\{f\}$ to denote the coefficient of x^n in f. For multivariate polynomials such as $f(x, y)$, we will use $[x^n]\{f\}$ by interpreting $f(x) \in \mathbb{F}[y][x]$ and extracting the coefficient of x^n as a function of y.
- The notation $\mathrm{Hom}_d(F)$ refers to the degree d homogeneous part of F, and $\mathrm{Hom}_{\leq d}(F)$ refers to the sum of all homogeneous parts of F up to degree d (which is sometimes also referred to as 'truncating' the polynomial at degree d).
- The model of computation for multivariate polynomials would be the standard model of algebraic circuits (which are directed acyclic graphs with internal gates labelled by $+$ and $\times$, with leaves labelled by variables or field constants, with field constant on edges). The size of a circuit C, denoted by $\mathrm{size}(C)$, would be the number of wires in the circuit. The depth of the circuit, denoted by $\mathrm{depth}(C)$, would be the length of the longest path from root to a leaf node. For a polynomial $f(\mathbf{x})$, we shall use $\mathrm{size}(f)$ to denote the size of the smallest circuit that computes f.

We also briefly use the notion of *border computation* which is given by a circuit C with coefficients from $\mathbb{F}(\varepsilon)$ where ε is a formal variable. We shall say that the circuit C is a *border computation* for a polynomial $f(\mathbf{x})$ if $C = f(\mathbf{x}) + \varepsilon \cdot g(\mathbf{x}, \varepsilon)$ where $g(\mathbf{x}, \varepsilon) \in \mathbb{F}[\varepsilon][\mathbf{x}]$. We use $\overline{\mathrm{size}}(f)$ to denote the size of the smallest circuit that border computes f.

- A polynomial $f(\mathbf{x})$ is said to be *squarefree* if there is no non-constant polynomial $g(\mathbf{x})$ such that g^2 divides f. Extending this, if $f(\mathbf{x}) = g_1^{e_1} \cdots g_r^{e_r}$ (with each $e_i \geq 1$) is the factorization of the polynomial into distinct irreducibles, the *squarefree* part of f is given by $g_1 \cdots g_r$.
- A map $\mathscr{G} : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{w}]$ is said to be a *hitting-set generator* for a class $\mathscr{C}$ of polynomials if for every $F \in \mathscr{C}$ we have that $F \circ \mathscr{G} = 0$ implies $F = 0$. The degree of the generator is $\max_i \deg(\mathscr{G}(x_i))$. The hitting-set generator is said to be *explicit* if $\mathscr{G}$ can be computed efficiently.

2.1 Polynomial Identity Lemma

Lemma 2.1 (Polynomial Identity Lemma [GRS23, Lemma 9.2.2])**.** *Let $\mathbb{F}$ be an arbitrary field and let $P(x) \in \mathbb{F}[\mathbf{x}]$ be a nonzero n-variate polynomial of degree d. Let S be an arbitrary subset of $\mathbb{F}$. Then,*

$$\Pr_{\mathbf{a} \in S^n}[P(\mathbf{a}) = 0] \leq \frac{d}{|S|}$$

2.2 Interpolation and consequences

The following applications of polynomial interpolation to algebraic circuits is attributed to Michael Ben-Or.

Lemma 2.2 (Interpolation)**.** *Let R be a commutative ring that contains a field $\mathbb{F}$ of at least $d+1$ elements, and let $\alpha_0, \ldots, \alpha_d$ be distinct elements in $\mathbb{F}$. Then, for every $i \in \{0, \ldots, d\}$, there exists fields elements $\beta_{i0}, \ldots, \beta_{id}$ such that for any $f(t) = f_0 + f_1 t + \cdots + f_d t^d \in R[t]$ of degree at most d, we have*

$$[t^i]\{f\} = f_i = \beta_{i0} f(\alpha_0) + \cdots + \beta_{id} f(\alpha_d) \qquad \square$$

Corollary 2.3 (Standard consequences of interpolation)**.** *Let $\alpha_0, \ldots, \alpha_d$ be distinct elements in $\mathbb{F}$. Then,*

1. ***[Partial derivatives]** If $C(\mathbf{x}, y)$ has degree d in the variable y, then the i-th order partial derivative of C with respect to y can be expressed as an $\mathbb{F}[y]$-linear combination of $\{C(\mathbf{x}, \alpha_j) \,:\, j \in \{0, \ldots, d\}\}$. That is, there are polynomials $\mu_0(y), \ldots, \mu_d(y)$ (not depending on C) of degree at most d such that*

$$\partial_{y^i} C(\mathbf{x}, y) = \mu_0(y) \cdot C(\mathbf{x}, \alpha_0) + \cdots + \mu_d(y) \cdot C(\mathbf{x}, \alpha_d).$$

2. ***[Homogeneous components]** Let $C(\mathbf{x})$ be a degree d polynomial. Then, for any subset $\mathbf{x}_S \subseteq \mathbf{x}$ and any $i \in [d]$, the degree i homogeneous part of C with respect to $\mathbf{x}_S$, denoted by $\mathrm{Hom}_{\mathbf{x}_S, i}(C)$, can be*

expressed as

$$\mathrm{Hom}_{\mathbf{x}_S,i}(C) = \sum_{j=0}^{d} \beta_{i,j} \cdot C(\alpha_j \cdot \mathbf{x}_S, \mathbf{x}_{\overline{S}})$$

for some constants $\beta_{i,j} \in \mathbb{F}$ (not depending on C).

In particular, if C is computable by a size s, depth Δ circuit, then all the above operations yield a circuit of size $\mathrm{poly}(s,d)$ *and depth $\Delta + O(1)$.* □

Observe that even if $d \ll \deg(f)$, interpolating the coefficient of t^d in $f(t)$ requires $\deg(f)+1$ many evaluations. However, we can express the coefficient of t^d in $f(t)$ as the *limit (or border)* of a sum of just $d+1$ evaluations of f. The following lemma states this formally.

Lemma 2.4 (Border Interpolation)**.** *Let R be a commutative ring that contains a field $\mathbb{F}$ of at least $d+1$ elements, and let $\alpha_0, \ldots, \alpha_d$ be distinct elements in $\mathbb{F}$. Then, there exists fields elements $\beta_0, \ldots, \beta_d$ such that for any $f(t) \in R[\![t]\!]$, we have*

$$[t^d]\,\{f\} = \frac{1}{\varepsilon^d} \cdot (\beta_0 f(\varepsilon\alpha_0) + \cdots + \beta_d f(\varepsilon\alpha_d)) + O(\varepsilon).$$

Proof. Suppose $f = f_0 + f_1 t + f_2 t^2 + \cdots$. Define $f'(t) = f_0 + f_1 t + \cdots + f_d t^d$ and $f''(t) = f - f'$. Then, applying Lemma 2.2 for $g(t) = f'(\varepsilon \cdot t)$ and $i = d$, we get constants $\beta_0, \ldots, \beta_d$ such that

$$\beta_0 f'(\varepsilon\alpha_0) + \cdots + \beta_d f'(\varepsilon\alpha_d) = \beta_0 g(\alpha_0) + \cdots + \beta_d g(\alpha_d) = [t^d]\,\{g\} = \varepsilon^d \cdot [t^d]\,\{f'\} = \varepsilon^d f_d$$

On the other hand, $f''(\varepsilon\alpha) = O(\varepsilon^{d+1})$ for any constant α. Therefore, since $f = f' + f''$, we have

$$\frac{1}{\varepsilon^d} \cdot (\beta_0 f(\varepsilon\alpha_0) + \cdots + \beta_d f(\varepsilon\alpha_d)) = [t^d]\,\{f'\} + O(\varepsilon^{d+1-d}) = f_d + O(\varepsilon).$$ □

2.3 Resultant, Discriminant and Gauss Lemma

We now recall some definitions that are standard in the factorization literature. For more details, we encourage the readers to refer to von zur Gathen and Gerhard's book on computer algebra [vzGG13].

Definition 2.5 (Sylvester Matrix and Resultant)**.** *Let $\mathbb{F}$ be a field. Let $P(z)$ and $Q(z)$ be polynomials of degree $a \geq 1$ and $b \geq 1$ in $\mathbb{F}[z]$. Define a linear map $\Gamma_{P,Q} : \mathbb{F}^a \times \mathbb{F}^b \to \mathbb{F}^{a+b}$ that takes polynomials $A(z)$ and $B(z)$ in $\mathbb{F}[z]$ of degree $a-1$ and $b-1$ respectively, and maps them to $AP + BQ$, a polynomial of degree $a+b-1$.*
The Sylvester matrix of P and Q, denoted by $\mathrm{Syl}_z(P,Q)$*, is defined to be the $(a+b) \times (p+q)$ matrix for the linear map $\Gamma_{P,Q}$.*
The Resultant of P and Q, denoted by $\mathrm{Res}_z(P,Q)$*, is the determinant of* $\mathrm{Syl}_z(P,Q)$*.* ◊

Definition 2.6 (Discriminant)**.** *Let $P(z)$ be a polynomial over a field $\mathbb{F}$. The Discriminant of P, denoted by $\mathrm{Disc}_z(P)$, is defined as the resultant of P and $\frac{\partial P}{\partial z}$.* $\lozenge$

Lemma 2.7 (Resultant and GCD [vzGG13, Corollary 6.20])**.** *Let $\mathscr{R}$ be a unique factorization domain, and let $P(z), Q(z) \in \mathscr{R}[z]$ be polynomials of degree $p \geq 1$ and $q \geq 1$ respectively. Then, $\mathrm{Res}_z(P,Q) = 0 \iff \deg_z(\gcd(P,Q)) \geq 1$. Moreover, there exist polynomials A and B of degree $q-1$ and $p-1$ such that $AP + BQ = \mathrm{Res}_z(P,Q)$.*

Lemma 2.8 (Discriminant and squarefreeness[DSS22, Lemma 12])**.** *Let $\mathscr{R}$ be a unique factorization domain, and let $P(z) \in \mathscr{R}[z]$ be a polynomial of degree at least 1. Then, $\mathrm{Disc}_z(P) = 0$ if and only if P is squarefree i.e. every irreducible factor of P has multiplicity one.*

Lemma 2.9 (Gauss Lemma [vzGG13, Section 6.2, Corollary 6.10])**.** *Let $\mathscr{R}$ be a unique factorization domain and let $\mathscr{K}$ be its field of fractions. Let $P(z) \in \mathscr{R}[z]$ be a monic polynomial. Then, $P(z)$ is irreducible in $\mathscr{R}[z]$ if and only if $P(z)$ is irreducible in $\mathscr{K}[z]$. In particular, the factorization of a monic polynomial $P(z)$ into its irreducible factors in $\mathscr{R}[z]$ is identical to its factorization into irreducible factors in $\mathscr{K}[z]$.*

2.4 Reducing to factorizing bivariate polynomials

For many of the factorization applications, it would be convenient to reduce the problem to a *bivariate* setting. The following definition and subsequent lemmas formalize the precise transformation and their properties. In what follows, it would be convenient to imagine the set of variables instead as $(\mathbf{x}, y)$ (that is, calling the last variable as y since it plays a slightly different role). We sometimes abuse notation to refer to a set $\mathbf{x}$ of variables as $(\mathbf{x}, y)$ by reusing the same names by artificially introducing a variable.

Definition 2.10 (Valid pre-processing maps for factorization)**.** *Let $\mathbf{a}, \mathbf{b} \in \mathbb{F}^{|\mathbf{x}|}$ and let $\mathbb{K} = \mathbb{F}(\mathbf{x})$. The homomorphism $\Psi_{\mathbf{a},\mathbf{b}} : \mathbb{F}[\mathbf{x}, y] \to \mathbb{K}[t, y]$ given by*

$$\Psi_{\mathbf{a},\mathbf{b}} : x_i \mapsto t x_i + a_i y + b_i, \quad \text{for all } i,$$
$$\Psi_{\mathbf{a},\mathbf{b}} : y \mapsto y$$

is said to be a valid pre-processing *map for a polynomial $F(\mathbf{x}, y) \in \mathbb{F}[\mathbf{x}, y]$ if $G(t, y) = \Psi_{\mathbf{a},\mathbf{b}}(F)$ satisfies the following properties:*

1. *The coefficient of $y^{\deg(F)}$ in G is nonzero.*
2. *If $\tilde{G}$ be the squarefree part of G, then $\tilde{G}(0, y)$ must be squarefree as well.* $\lozenge$

Lemma 2.11 (Recovering factors from pre-processed factors)**.** *If $G(t, y) = \Psi_{\mathbf{a},\mathbf{b}}(F(\mathbf{x}, y))$ for a valid pre-processing map, then $G(0, y)$ is a univariate polynomial in y over the base field $\mathbb{F}$. Furthermore, the factors of $G(t, y) \in \mathbb{F}(\mathbf{x})[t, y]$ are in one-to-one correspondence with the factors of $F(\mathbf{x}, y)$, with the inverse map $\Psi^{-1}_{\mathbf{a},\mathbf{b}}$ given by $x_i \mapsto x_i - a_i y - b_i$ and $t \mapsto 1$.*

Proof. For any choice of $\mathbf{a}, \mathbf{b}$, the map $\psi_{\mathbf{a},\mathbf{b}} : x_i \mapsto x_i + a_i y + b_i$ is an automorphism of the polynomial ring $\mathbb{F}[\mathbf{x}, y]$ that keeps non-constant polynomials non-constant. Therefore, for any polynomial $F(\mathbf{x}, y)$, the factors of $F(\mathbf{x}, y)$ are in one-to-one correspondence with the factors of $\psi_{\mathbf{a},\mathbf{b}}(F(\mathbf{x}, y))$.

If $\Psi_{\mathbf{a},\mathbf{b}}$ is a valid pre-processing map for $F(\mathbf{x}, y)$ and $G(t, y) = \Psi_{\mathbf{a},\mathbf{b}}(F) \in \mathbb{F}[\mathbf{x}][t, y]$ then by Definition 2.10 Item 1 we have that $G(t, y)$ is monic in y.[4] By Gauss' lemma, the $G(t, y)$ is reducible over $\mathbb{F}(\mathbf{x})$ if and only if $G(t, y)$ is reducible over $\mathbb{F}[\mathbf{x}]$. If $G(t, y) = G_1(t, y) \cdot G_2(t, y)$ is a non-trivial factorization, we get a non-trivial factorization of $\psi_{\mathbf{a},\mathbf{b}}(F(\mathbf{x}, y))$ by setting $t = 1$ since $G_i(t, y)$ is monic for $i = 1, 2$ and will remain non-trivial under the substitution $t = 1$.

Thus, the factors of $G(t, y)$ are in one-to-one correspondence with the factors of $F(\mathbf{x}, y)$, and we can easily obtain one from the other. □

Lemma 2.12. *Let $F(\mathbf{x}, y)$ be a nonzero polynomial of degree d. Suppose $\mathbf{a}, \mathbf{b} \in \mathbb{F}^n$ (where $n = |\mathbf{x}|$) satisfy the following properties:*

1. $\mathrm{Hom}_d(F)(\mathbf{a}) \neq 0$,
2. $\mathrm{Disc}_y(\tilde{F}(a_1 y + b_1, \ldots, a_n y + b_n, y)) \neq 0$ *where $\tilde{F}$ is the squarefree part of F and $\mathrm{Disc}_y(F)$ denotes the discriminant of F with respect to y.*

Then, $\Psi_{\mathbf{a},\mathbf{b}}$ is a valid pre-processing map for $F(\mathbf{x}, y)$.

Proof. It is easy to observe that the coefficient of y^d in $G(t, y) = \Psi_{\mathbf{a},\mathbf{b}}(F)$ is precisely $\mathrm{Hom}_d(F)(\mathbf{a})$. Thus, this implies Definition 2.10 Item 1.

By Lemma 2.11, if $\tilde{G}$ is the squarefree part of G, then $\tilde{G} = \Psi_{\mathbf{a},\mathbf{b}}(\tilde{F})$. Therefore, $\tilde{G}(0, y)$ will be squarefree if and only if $\tilde{F}(a_1 y + b_1, \ldots, a_n y + b_n, y)$ is square free. Hence, $\mathrm{Disc}_y(\tilde{F}(a_1 y + b_1, \ldots, a_n y + b_n, y)) \neq 0$ implies Definition 2.10 Item 2. □

Remark 2.13. *A consequence of the above lemma, along with the Polynomial Identity Lemma (Lemma 2.1) is that, for any polynomial $F(\mathbf{x}, y)$, the map $\Psi_{\mathbf{a},\mathbf{b}}$ when $\mathbf{a}, \mathbf{b}$ is picked at random is a valid pre-processing map. In particular, valid pre-processing maps always exist.* ◊

2.5 Factorization of monic polynomials into power series

The following are some standard facts about power series roots of polynomials under modest conditions.

Lemma 2.14 (Factorization into power series)**.** *Let $f(t, y) \in \mathbb{K}[t, y]$ be a polynomial that is monic in y such that $f(0, y)$ is squarefree. For each $\alpha \in \overline{\mathbb{K}}$ (the algebraic closure) such that $f(0, \alpha) = 0$, there is a unique power series $\varphi_\alpha(t) \in \overline{\mathbb{K}}[\![t]\!]$ satisfying $\varphi_\alpha(0) = \alpha$ such that $f(t, \varphi_\alpha(t)) = 0$.*

[4] Here we just mean that the coefficient lies in $\mathbb{F}$.

In fact, the polynomial $f(t,y)$ factorizes in $\overline{\mathbb{K}}[\![t]\!][y]$ as

$$f(t,y) = \prod_{\alpha \in Z} (y - \varphi_\alpha(t))$$

where Z is the set of roots of $f(0,y)$ in $\overline{\mathbb{K}}$.

The above lemma is essentially folklore and [DSS22, Section 3] gives a formal proof of the above.

2.6 Some known results on algebraic circuits

Theorem 2.15 (Computing squarefree decomposition (Theorem I.9 in [AW24]))**.** *Let $\mathbb{F}$ be a field of characteristic zero. Given an algebraic circuit of size s and depth Δ computing a polynomial degree d polynomial $f(\mathbf{x})$, consider the (unique) sequence of polynomials $f_1, \ldots, f_d \in \mathbb{F}[\mathbf{x}]$ such that $\gcd(f_i, f_j) = 1$ and $f = \prod_{i=1}^{n} f_i^i$. Then, each f_i can be computed using size $\mathrm{poly}(s,d)$ and depth $\Delta + O(1)$ circuits. In particular, the squarefree part of f (which is equal to $f_1 \cdots f_n$) is computable by size $\mathrm{poly}(s,d)$ and depth $\Delta + O(1)$ circuits.*

Furthermore, the circuits can be computed in polynomial time given access to an oracle solving PIT for polynomial size constant-depth circuits (assuming Δ is constant). □

Remark 2.16. *The model of constant depth algebraic circuits as defined in [AW24] are allowed to have division gates. However, an inspection of their proof of Theorem 2.15 reveals that an oracle access to a PIT algorithm for division free constant depth algebraic circuits is sufficient to obtain an efficient algorithm that takes as input a constant depth circuit (without divisions) and outputs constant depth algebraic circuits (again, without division gates) for its squarefree decomposition.*

Moreover, the same argument holds for algebraic formulas, provided we have a PIT oracle for such formulas. ◊

We also make use of known hitting set generators for polynomial size constant-depth circuits. Although [LST21, Corollary 5] is stated as a whitebox PIT, this also provides an explicit hitting set generator. The statement below is an alternate generator construction from a result of Andrews and Forbes [AF22].

Theorem 2.17 (Explicit hitting sets for constant-depth circuits (Theorem 6.8 in the full version of [AF22]))**.** *Let $\mathbb{F}$ be a field of characteristic zero. For every $k \in \mathbb{N}$, there is a hitting set generator $\mathscr{G}_k : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{w}]$ with $|\mathbf{w}| = n^{1/2^k + o(1)}$ and for the class of $\mathrm{poly}(n)$-size depth $\Delta \leq o(\log\log\log n)$ circuits.* □

3 Explicit formulas for implicitly defined power series

For a bivariate power series $F(x,y) = \sum_{i,j} F_{i,j} x^i y^j \in \mathbb{F}[\![x,y]\!]$, define the *diagonal* operator $\mathscr{D}(F)(t)$ as

$$\mathscr{D}(F)(t) = \sum_{i \geq 0} F_{i,i} \cdot t^i.$$

The following result of Furstenberg [Fur67] allows us to express power series roots as a diagonal of a rational function. We present Furstenberg's proof to keep the exposition self-contained.

Theorem 3.1 ([Fur67, Proposition 2])**.** *Let $\mathbb{F}$ be an arbitrary field. Let $P(t,y) \in \mathbb{F}[\![t,y]\!]$ be a power series and $\varphi(t) \in \mathbb{F}[\![t]\!]$ be a power series satisfying*

$$P(t,y) = (y - \varphi(t))^e \cdot Q(t,y)$$

for some $e \geq 1$ that is invertible in $\mathbb{F}$. If $\varphi(0) = 0$ and $Q(0,0) \neq 0$, then

$$\varphi = \mathscr{D}\left(\frac{y^2 \cdot \partial_y P(ty,y)}{e \cdot P(ty,y)}\right) \tag{3.2}$$

Remark. *Although the [Fur67, Proposition 2] originally stated it for $P(t,y)$ being a polynomial and for $e = 1$, it can be seen to readily extend to $P(t,y)$ being a power series as well (cf. [Hu16]) and for any e that is invertible in $\mathbb{F}$.* ◊

Proof. By scaling P if required, we may assume without loss of generality that $Q(0,0) = 1$.

$$\begin{aligned}
P(t,y) &= (y - \varphi(t))^e Q(t,y) \\
\implies \frac{\partial_y P(t,y)}{P(t,y)} &= \frac{e}{y - \varphi(t)} + \frac{\partial_y Q(t,y)}{Q(t,y)} \\
\implies \frac{y^2 \cdot \partial_y P(ty,y)}{e \cdot P(ty,y)} &= \frac{y^2}{y - \varphi(ty)} + \frac{y^2 \cdot \partial_y Q(ty,y)}{e \cdot Q(ty,y)} \\
\implies \mathscr{D}\left(\frac{y^2 \cdot \partial_y P(ty,y)}{e \cdot P(ty,y)}\right) &= \mathscr{D}\left(\frac{y^2}{y - \varphi(ty)}\right) + \mathscr{D}\left(\frac{y^2 \cdot \partial_y Q(ty,y)}{e \cdot Q(ty,y)}\right).
\end{aligned}$$

We first deal with the second summand. Since $Q(0,0) = 1$, we can write $Q(ty,y)$ as $(1 - \tilde{Q}(ty,y))$ for with $\tilde{Q}$ satisfying $\tilde{Q}(0,0) = 0$. Thus, the second summand, as a power series expression, becomes

$$\mathscr{D}\left(\frac{y^2 \partial_y Q(ty,y)}{e \cdot Q(ty,y)}\right) = \mathscr{D}\left(y^2 \cdot e^{-1} \cdot \partial_y Q(ty,y) \left(\sum_{i=0}^{\infty} \tilde{Q}(ty,y)^i\right)\right)$$

In $\partial_y Q(ty,y) \cdot (\sum_{i=0}^{\infty} \tilde{Q}(ty,y)^i)$, every monomial has y-degree at least as large as the t-degree (since

we replaced t by ty). Multiplying by y^2 ensures that there are no monomials with t-degree equal to y-degree. Hence, $\mathscr{D}\left(\frac{y^2\partial_y Q(ty,y)}{eQ(ty,y)}\right) = 0$.

For the first summand, since $\varphi(0) = 0$, we have $\varphi(ty)$ is divisible by y and hence

$$\mathscr{D}\left(\frac{y}{1-(\varphi(ty)/y)}\right) = \mathscr{D}\left(y\sum_{i\geq 0}\left(\frac{\varphi(ty)}{y}\right)^i\right) = \mathscr{D}\left(\sum_{i=0}^{\infty}\left(\frac{(\varphi(ty))^i}{y^{i-1}}\right)\right).$$

Observe that for every $i \geq 0$, $m_i := (\varphi(ty))^i/y^{i-1}$ satisfies $\deg_t(m_i) = \deg_y(m_i) + i - 1$, which means that $\deg_t(m_i) = \deg_y(m_i)$ if and only if $i = 1$. Thus, $\mathscr{D}\left(\frac{y}{1-(\varphi(ty)/y)}\right) = \mathscr{D}(\varphi(ty)) = \varphi$. □

The diagonal expression above can be simplified to a slightly more convenient expression for implicitly defined power series roots (which we state below for the case of roots of multiplicity one).

Corollary 3.3. *Let $\mathbb{F}$ be an arbitrary field. Let $P(t,y) \in \mathbb{F}[\![t,y]\!]$ and $\varphi(t) \in \mathbb{F}[\![t]\!]$ such that $\varphi(0) = 0$, $P(t,\varphi(t)) = 0$ and $\partial_y P(0,0) = \alpha \neq 0$. Then,*

$$\varphi(t) = \sum_{m\geq 1}\frac{1}{\alpha^{m+1}}\cdot[y^{m-1}]\left\{\partial_y P(t,y)\cdot(\alpha y - P(t,y))^m\right\}.$$

For characteristic zero fields, the following is an alternate expression

$$\varphi(t) = \sum_{m\geq 1}\frac{1}{m\cdot\alpha^m}\cdot[y^{m-1}]\left\{(\alpha y - P(t,y))^m\right\}.$$

Proof. By scaling P if required, assume $\partial_y P(0,0) = 1$. Since $P(0,0) = 0$, this implies that $\frac{P(ty,y)}{y}$ is a polynomial with constant term equal to 1, and is thus invertible as a power series. By Theorem 3.1, we have

$$\begin{aligned}\varphi(t) &= \mathscr{D}\left(\frac{y\cdot\partial_y P(ty,y)}{P(ty,y)/y}\right) = \mathscr{D}\left(\frac{y\cdot\partial_y P(ty,y)}{1-(1-\frac{P(ty,y)}{y})}\right)\\ &= \mathscr{D}\left(\sum_{m\geq 0} y\cdot\partial_y P(ty,y)\cdot\left(1-\frac{P(ty,y)}{y}\right)^m\right)\end{aligned}$$

The term corresponding to $m = 0$ is just $y\cdot\partial_y P(ty,y)$ and consists only of monomials where the y-degree is greater than the t-degree and hence does not contribute any diagonal terms. Hence,

$$\begin{aligned}\varphi(t) &= \mathscr{D}\left(\sum_{m\geq 1} y\cdot\partial_y P(ty,y)\cdot\left(1-\frac{P(ty,y)}{y}\right)^m\right)\\ \implies [t^n](\varphi) &= [t^n y^n]\left\{\sum_{m\geq 1} y\cdot\partial_y P(ty,y)\cdot\left(1-\frac{P(ty,y)}{y}\right)^m\right\}\end{aligned}$$

$$= [t^n y^n] \left\{ \sum_{m \geq 1} y^{1-m} \cdot \partial_y P(ty, y) \cdot (y - P(ty, y))^m \right\}$$
$$= \sum_{m \geq 1} [t^n y^{n+m-1}] \left\{ \partial_y P(ty, y) \cdot (y - P(ty, y))^m \right\}$$

For any Laurent series $R(t, y)$, we have $[t^i y^j] \{R(t, y)\} = [t^i y^{i+j}] \{R(ty, y)\}$. Hence, we have

$$[t^n](\varphi) = \sum_{m \geq 1} [t^n y^{m-1}] \left\{ \partial_y P(t, y) \cdot (y - P(t, y))^m \right\}$$
$$= [t^n] \left\{ \sum_{m \geq 1} [y^{m-1}] \left\{ \partial_y P(t, y) \cdot (y - P(t, y))^m \right\} \right\}$$

which completes the proof of the first expression (the proof for $\partial_y P(0, 0) = \alpha \neq 1$ follows similarly).

For the expression over characteristic zero fields, let $G(t, y) = y - P(t, y)$. Then

$$[t^n] \{\varphi(t)\} = \sum_{m \geq 1} [y^{m-1}] \left\{ (1 - \partial_y G(t, y)) \cdot G(t, y)^m \right\}$$
$$= \sum_{m \geq 1} \left([y^{m-1}] \left\{ G(t, y)^m \right\} - [y^{m-1}] \left\{ \partial_y G(t, y) \cdot G(t, y)^m \right\} \right)$$
$$= \sum_{m \geq 1} \left([y^{m-1}] \left\{ G(t, y)^m \right\} - \left(\frac{1}{m+1} \right) [y^{m-1}] \left\{ \partial_y G(t, y)^{m+1} \right\} \right)$$

For any power series $R(x, y)$ and any $k > 0$, note that $[y^{k-1}] \left\{ \partial_y R(x, y) \right\} = k \cdot [y^k] \{R(x, y)\}$. Hence,

$$[t^n] \{\varphi(t)\} = \sum_{m \geq 1} \left([y^{m-1}] \left\{ G(t, y)^m \right\} - \left(\frac{m}{m+1} \right) [y^m] \left\{ G(t, y)^{m+1} \right\} \right)$$
$$= \sum_{m \geq 1} \frac{1}{m} \cdot [y^{m-1}] \left\{ G(t, y)^m \right\}. \qquad \square$$

4 Closure under taking factors

In this section, we use the techniques discussed in Section 3 to prove closure results (Theorem 1.1).

We start by proving upper bounds on the complexity of power series roots, followed by a proof of upper bounds on general factors.

4.1 Complexity of power series roots

Theorem 4.1 (Power series roots without multiplicity)**.** *Let $P(\mathbf{x}, y) \in \mathbb{F}[\mathbf{x}, y]$ be a polynomial computed by a circuit C, and let $\varphi(\mathbf{x}) \in \mathbb{F}[\![\mathbf{x}]\!]$ be a power series satisfying $\varphi(\mathbf{0}) = 0$, $P(\mathbf{x}, \varphi(\mathbf{x})) = 0$ and $\partial_y P(\mathbf{0}, 0) \neq$*

0. *Then, for any $d \in \mathbb{N}$, there is a circuit C' computing* $\mathrm{Hom}_{\leq d}[\varphi]$ *such that*

$$\mathrm{size}(C') \leq \mathrm{poly}(d, \mathrm{size}(C))$$

$$\mathrm{depth}(C') \leq \mathrm{depth}(C) + O(1)$$

Proof. This theorem follows almost immediately from Corollary 3.3 along with some standard ideas in factorization literature.

For simplicity, by scaling the polynomial P if required, we may assume without loss of generality that $\partial_y P(\mathbf{0}, 0) = 1$. We first perform the standard transformation of replacing each x_i by $t \cdot x_i$ and work over the field $\mathbb{K} := \mathbb{F}(\mathbf{x})$. We define $\hat{P}(t, y) := P(t \cdot \mathbf{x}, y) \in \mathbb{K}[t, y]$ and $\hat{\varphi}(t) := \varphi(t \cdot \mathbf{x}) \in \mathbb{K}[\![t]\!]$. Thus, we maintain the conditions $\hat{\varphi}(0) = 0$, $\hat{P}(t, \hat{\varphi}(t)) = 0$ and $\partial_y \hat{P}(0, 0) = 1$. We can now apply Corollary 3.3 to get

$$\hat{\varphi}(t) = \sum_{m \geq 1} [y^{m-1}] \left\{ \partial_y \hat{P}(t, y) \cdot (y - \hat{P}(t, y))^m \right\}.$$

Note that, since $\hat{P}(0, 0) = 0$ and $\partial_y \hat{P}(0, 0) = 1$, we have that every monomial of $y - \hat{P}(t, y)$ is either divisible by t or by y^2. Therefore,

$$\begin{aligned} \mathrm{Hom}_{\leq d}[\hat{\varphi}(t)] &= \mathrm{Hom}_{\leq d}\left[\sum_{m \geq 1} [y^{m-1}] \left\{ \partial_y \hat{P}(t, y) \cdot (y - \hat{P}(t, y))^m \right\} \right] \\ &= \mathrm{Hom}_{\leq d}\left[\sum_{m=1}^{2d} [y^{m-1}] \left\{ \partial_y \hat{P}(t, y) \cdot (y - \hat{P}(t, y))^m \right\} \right] \end{aligned}$$

since every monomial of $(y - \hat{P}(t, y))^\ell = (t \cdot A + y^2 \cdot B)^\ell$ either has t-degree at least $\ell/2$, or y-degree at least ℓ and thus terms with $m > 2d$ have no contribution to the LHS.

The expression is clearly in $\mathbb{F}[\mathbf{x}, t]$ and not just $\mathbb{K}[t]$. Setting $t = 1$ helps us retrieve $\mathrm{Hom}_{\leq d}[\varphi(\mathbf{x})]$ since the transformation $x_i \mapsto t \cdot x_i$ ensures that each monomial has the same $\mathbf{x}$-degree and t-degree in $\varphi(t \cdot \mathbf{x})$. The depth and size bounds follow via interpolation and homogenization (Lemma 2.2 and Corollary 2.3). □

4.2 Complexity of general factors

We now proceed to prove our closure result for general factors. If $f(t, y)$ is a polynomial that is monic in y and is divisible by $g(t, y)$, then over the algebraic closure $\overline{\mathbb{F}}$ we can express $g(0, y) = \prod_{i=1}^{\ell} (y - \alpha_i)$. Thus, a simple proof to obtain a constant-depth circuit for $g(t, y)$ *over the algebraic closure* $\overline{\mathbb{F}}$ would be to just consider

$$g(t, y) = \mathrm{Hom}_{\leq d}\left(\prod_{i=1}^{\ell} (y - C_{\alpha_i}(t)) \right)$$

where $C_{\alpha_i}(t)$ is the root of the power series lifted from α_i obtained via Theorem 4.1. (We need a few additional ingredients, such as reducing to the square-free case and taking a suitable shift)

To obtain a circuit for the factors over the base field, we require a few more modifications that we now describe (and obtaining a circuit over the base field would also be essential for computing the factors algorithmically). We would require the following statement from the recent work of Andrews and Wigderson [AW24] to prove our closure result for general factors.

Theorem 4.2 (Theorem I.8 of [AW24])**.** *Let $\mathbb{F}$ be a field of zero or large characteristic. Suppose $f, g, h \in \mathbb{F}[z]$ with $\deg(f), \deg(g), \deg(h) \leq d$. Suppose $\alpha_1, \ldots, \alpha_d \in \overline{\mathbb{F}}$ be the roots of $f(z)$ with multiplicity, with $h(\alpha_i) \neq 0$ for all i. Then, for any $r \in [d]$, $\mathbf{Esym}_r(\frac{g(\alpha_1)}{h(\alpha_1)}, \ldots, \frac{g(\alpha_d)}{h(\alpha_d)})$ can be computed by a circuit of size $\mathrm{poly}(d)$ and depth $O(1)$ over the coefficients of f, g and h.*

It is worth stressing that the above circuit is over the base field $\mathbb{F}$, and *does not* take the α_i's as input; the inputs are just the coefficients of f, g and h which come from the base field.

Theorem 4.3 (General factors of algebraic circuits)**.** *Let $\mathbb{F}$ be a field of zero or large enough characteristic, and let $P(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a polynomial on n variables of degree d computed by a circuit C of size s and depth Δ. Then, any factor $g(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ of P is computable by a circuit of size $\mathrm{poly}(s, d, n)$ and depth $\Delta + O(1)$ over $\mathbb{F}$.*

Proof. We may assume that we work with the squarefree part of $P(\mathbf{x})$, which is also computable by a circuit of size $\mathrm{poly}(s, d, n)$ and depth $\Delta + O(1)$ by appealing to Theorem 2.15. By reusing symbols, let us assume that $P(\mathbf{x})$ is squarefree.

By interpreting $P(\mathbf{x})$ as an element of $\mathbb{F}[\mathbf{x}, y]$, let $\tilde{P}(t, y) = \Psi(P(\mathbf{x}, y))$ for some valid pre-processing map Ψ (recall Definition 2.10, and that they always exist (Remark 2.13)). Note that $\tilde{P}(0, y)$ is squarefree, and we have that $\tilde{P}(t, y) \in \mathbb{F}[\mathbf{x}][t, y]$ is computable by a size $\mathrm{poly}(s, d, n)$, depth $\Delta + O(1)$ circuit over $\mathbb{F}$. By Lemma 2.11, it suffices to show that an arbitrary factor $g(t, y) \in \mathbb{F}[\mathbf{x}][t, y]$ of $\tilde{P}(t, y)$ is computable by a $\mathrm{poly}(s, d, n)$ size, depth $\Delta + O(1)$ circuit over the field $\mathbb{F}$.

Define the Laurent series $\hat{R}(z) \in \mathbb{F}[\mathbf{x}, t]((z))$, and its truncated rational function $R(z) \in \mathbb{F}[\mathbf{x}, t](z)$ as follows:

$$\hat{R}(z) = z + \sum_{m \geq 1} \left(\frac{1}{\partial_y \tilde{P}(0, z)} \right)^{m+1} \cdot [y^{m-1}] \left\{ \partial_y \tilde{P}(t, y + z) \cdot (y \cdot \partial_y \tilde{P}(0, z) - \tilde{P}(t, y + z))^m \right\}$$

$$R(z) = z + \sum_{m=1}^{2d+2} \left(\frac{1}{\partial_y \tilde{P}(0, z)} \right)^{m+1} \cdot [y^{m-1}] \left\{ \partial_y \tilde{P}(t, y + z) \cdot (y \cdot \partial_y \tilde{P}(0, z) - \tilde{P}(t, y + z))^m \right\}$$

Note that we have that the rational function $R(z)$ can be easily expressed as $\frac{R_{\mathrm{num}}(z)}{R_{\mathrm{denom}}(z)}$ where each $R_{\mathrm{num}}(z)$ and $R_{\mathrm{denom}}(z)$ are both computable by a $\mathrm{poly}(s, d, n)$ sized depth $\Delta + O(1)$ circuits since

$\tilde{P}(t,y)$ is given by a $\mathrm{poly}(s,d,n)$ sized depth $\Delta + O(1)$ circuit (using Lemma 2.2 and Corollary 2.3). Furthermore, $R_{\mathrm{denom}}(z) = (\partial_y \tilde{P}(0,z))^{2d+3}$ and since $\tilde{P}(0,y)$ is square free we have $R_{\mathrm{denom}}(\alpha)$ is nonzero for every root α of $\tilde{P}(0,y)$.

For any $\alpha \in \overline{\mathbb{F}}$ such that $\tilde{P}(0,\alpha) = 0$, note that $\hat{R}(\alpha)$ is in fact an element of $\mathbb{F}[\mathbf{x}][\![t]\!]$ and $R(\alpha)$ is an element of $\mathbb{F}[\mathbf{x}][t]$. Also, if $\lambda(t,y) = (y \cdot \partial_y \tilde{P}(0,\alpha) - \tilde{P}(t, y+\alpha))$, then $\lambda(0,0) = \partial_y \lambda(0,0) = 0$ and hence every monomial of $\lambda(t,y)$ is divisible by either t or y^2. Therefore, we have

$$R(\alpha) = \hat{R}(\alpha) \bmod t^{d+1}.$$

For any such $\alpha \in \overline{\mathbb{F}}$, by Corollary 3.3 (applied to $\varphi_\alpha(t) - \alpha$ being a root of $\tilde{P}(t, y+\alpha)$), we have $\varphi_\alpha(t) := \hat{R}(\alpha) \in \mathbb{F}[\mathbf{x}][\![t]\!]$ as the unique power series such that $\varphi_\alpha(0) = \alpha$ and $\tilde{P}(t, \varphi_\alpha(t)) = 0$.

Thus, if $\tilde{P}(0,y) = \prod_{i=1}^{r}(y - \alpha_i)$ for $\alpha_i \in \overline{\mathbb{F}}$, and $g(0,y) = \prod_{\alpha \in S}(y-\alpha)$ for some subset $S \subset \{\alpha_1, \ldots, \alpha_r\}$, then the power series roots of $g(t,y)$ are precisely $\varphi_\alpha(t)$ for $\alpha \in S$ and hence

$$\begin{aligned}
g(t,y) &= \prod_{\alpha \in S}(y - \varphi_\alpha(t)) = \prod_{\alpha \in S}(y - \hat{R}(\alpha)). \\
&= \prod_{\alpha \in S}(y - R(\alpha)) \bmod t^{d+1} \\
\implies g(t,y) &= \mathrm{Hom}_{\le d}\left(g'(t,y)\right) \\
\text{where} \quad g'(t,y) &= \prod_{\alpha \in S}(y - R(\alpha)).
\end{aligned}$$

Each coefficient of any y^i in $g'(t,y)$ is an appropriate elementary symmetric polynomial of the set $\left\{ \frac{R_{\mathrm{num}}(\alpha)}{R_{\mathrm{denom}}(\alpha)} \ : \ \alpha \in S \right\}$. Since $g(0,y) = \prod_{\alpha \in S}(y-\alpha)$, the elementary symmetric polynomials of the set $\{\alpha \ : \ \alpha \in S\}$ are just the coefficients of $g(0,y)$, which are just elements of the field $\mathbb{F}$. By Theorem 4.2, the elementary symmetric polynomials of the set $\left\{ \frac{R_{\mathrm{num}}(\alpha)}{R_{\mathrm{denom}}(\alpha)} \ : \ \alpha \in S \right\}$ can computed as $\mathrm{poly}(s,d,n)$ sized depth $\Delta + O(1)$ circuits. Therefore, we have a similar circuit for $g'(t,y)$ and hence also for $g(t,y)$ (by Corollary 2.3). □

Therefore, over any characteristic zero (or large enough characteristic) field, classes of algebraic circuits such as VP, VNP, algebraic branching programs, algebraic formulas, constant-depth circuits are all closed under taking factors. (See Corollary 5.7 for a slightly more detailed statement.)

5 Deterministic algorithms for factorization

As discussed in Section 1.1.1, our closure results lead to a clean proof of correctness for the results of [BKR+25], which gave deterministic subexponential time algorithms to output efficient circuits (of potentially unbounded depth) for each of the factors of constant-depth circuit. Moreover, we can output constant-depth circuits for each of the factors.

The core of the algorithm is the following lemma whose proof we will defer to the end of the section. A version of this lemma also appeared in [BKR+25] but only for a specifically chosen hitting set generator. The statement below is more general, and the proof is much cleaner.

Lemma 5.1 (Irreducibility preservation)**.** *Let $\mathbb{F}$ be a field of zero (or large enough) characteristic. Let $F(\mathbf{x}, t, y) \in \mathbb{F}[\mathbf{x}, t, y]$ be a nonzero degree d polynomial that is computable by a circuit of size s and depth Δ. Suppose F is monic in y, with the property that $F(\mathbf{x}, 0, y) = F(\mathbf{0}, 0, y) \in \mathbb{F}[y]$ (i.e., every monomial divisible by an x_i is also divisible by t), and $F(\mathbf{0}, 0, y)$ is squarefree.*

Let $\mathscr{G} : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{w}]$ be a hitting set generator for the class of size $\mathrm{poly}(s, d)$, depth $\Delta + O(1)$ circuits.

Then, for every irreducible factor $G(\mathbf{x}, t, y)$ of $F(\mathbf{x}, t, y)$ we have that $G \circ \mathscr{G} \in \mathbb{F}[\mathbf{w}, t, y]$ is also irreducible.

We now proceed with the main theorem of this section.

Theorem 5.2. *Let $\mathbb{F}$ be the field of rational numbers. Fix any constant $\Delta \in \mathbb{N}$ and $\varepsilon > 0$. Then, there is a deterministic algorithm $\mathscr{A}_{\Delta,\varepsilon}$ that takes as input a size s depth-Δ circuit for a degree d polynomial $P(\mathbf{x}) \in \mathbb{F}[x_1, \ldots, x_n]$ and outputs circuits of size $\mathrm{poly}(s, d)$ and depth $\Delta + O(1)$ for each irreducible factor $g(\mathbf{x})$ of $P(\mathbf{x})$, along with their multiplicities. Moreover, $\mathscr{A}_{\Delta,\varepsilon}$ runs in time $\mathrm{poly}(s, d)^{O(n^\varepsilon)}$.*

Remark 5.3. *The theorem continues to be true over any field of zero or sufficiently large characteristic assuming that we have an efficient deterministic algorithm to factor univariates over this field.* ◊

For any bivariate degree d polynomial $F(t, y) \in \mathbb{F}[\mathbf{x}][t, y]$ that is monic in y, we define the polynomial $R_F(z) \in \mathbb{F}[\mathbf{x}, t][z]$ (as in the proof of Theorem 4.3), as follows

$$R_F(z) = z + \sum_{m=1}^{2d+2} \left(\frac{1}{\partial_y F(0, z)} \right)^{m+1} \cdot [y^{m-1}] \left\{ \partial_y F(t, y + z) \cdot (y \cdot \partial_y F(0, z) - F(t, y + z))^m \right\}.$$

Proof of Theorem 5.2. Let $P(\mathbf{x})$ be an n-variate degree d polynomial given by a circuit of size s and depth Δ. We may assume without loss of generality that $P(\mathbf{x})$ is squarefree (as the squarefree component[5] can be extracted using Theorem 2.15). We outline the rough steps of the algorithm $\mathscr{A}_{\Delta,\varepsilon}$ below and elaborate on the correctness.

1. **(Pre-processing)** Build a circuit C of size $\mathrm{poly}(s)$ and depth $\Delta + O(1)$ for $F(\mathbf{x}, t, y) = \Psi_{\mathbf{a},\mathbf{b}}(P) \in \mathbb{F}[\mathbf{x}][t, y]$ where $\Psi_{\mathbf{a},\mathbf{b}}$ is a valid pre-processing map for $P(\mathbf{x})$.

[5] In fact, the algorithm from Theorem 2.15 outputs the *squarefree decomposition* of the polynomial. The squarefree decomposition of a polynomial $P(\mathbf{x})$ is a sequence of polynomials $(P_1, \ldots, P_r)$ such that each P_i is a product of exactly those irreducible factors of P that have multiplicity i in its factorization. In particular, Theorem 2.15 immediately gives us the multiplicity of each factor that we obtain from the rest of the algorithm. Since our candidate factors from the algorithm have constant-depth circuits, we can also run a divisibility test on powers of each candidate factor to compute their multiplicities.

2. **(Variable reduction)** For a generator $\mathscr{G} : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{w}]$ for size $\text{poly}(s)$, depth $\Delta + O(1)$ circuits, define the polynomial $\tilde{F}(\mathbf{w}, t, y) := F(\mathbf{x}, t, y) \circ \mathscr{G}$. (Instantiating with the generator in Theorem 2.17, we may assume $|\mathbf{w}| = n^{\varepsilon}$)

3. **(Factorizing variable-reduced polynomial)** Factorize $\tilde{F}(\mathbf{w}, t, y)$ into irreducibles as $\tilde{F}(\mathbf{w}, t, y) = \tilde{G}_1(\mathbf{w}, t, y) \cdots \tilde{G}_k(\mathbf{w}, t, y)$. Use interpolation to compute the coefficients of $g_i(y) = \tilde{G}_i(\mathbf{w}, 0, y) = \tilde{G}_i(\mathbf{0}, 0, y)$ for all $i \in [k]$.

4. **(Building the factors)** For each $j \in [r]$, if $S_j \subset \overline{\mathbb{F}}$ are the roots of $g_j(y)$ in the algebraic closure, define the polynomial

$$G_j(\mathbf{x}, t, y) := \text{Hom}_{\leq d}\left(\prod_{i \in S_j}(y - R_F(\alpha_i))\right)$$

   From the coefficients of $g_j(y)$, use Theorem 4.2 to compute the coefficients (as elements of $\mathbb{F}[\mathbf{x}, t]$) of $G_j(\mathbf{x}, t, y)$ via $\text{poly}(s)$ size depth $\Delta + O(1)$ circuits.

5. **(Undo pre-processing and return)** Return $\left\{\Psi^{-1}_{\mathbf{a},\mathbf{b}}(G_j) \ : \ j \in [k]\right\}$.

We will justify correctness for each of the above steps.

Pre-processing: By Corollary 2.3, the highest degree homogeneous part of $P(\mathbf{x})$ is also computable by size $\text{poly}(s)$, depth $\Delta + O(1)$ sized circuits, and by [AW24] we have that $\text{Disc}_y(P)$ is computable by a size $\text{poly}(s)$, depth $\Delta + O(1)$ circuit. Thus, by Lemma 2.12, any hitting set for size $\text{poly}(s)$, depth $\Delta + O(1)$ circuits may be used to compute a valid pre-procesing map $\Psi_{\mathbf{a},\mathbf{b}}$.

For what follows, let $F(\mathbf{x}, t, y) = \Psi_{\mathbf{a},\mathbf{b}}(P)$.

Variable reduction: Note that the polynomial $F(\mathbf{x}, t, y)$ satisfy the requirements of Lemma 5.1. Thus, by Lemma 5.1, we have that $\mathscr{G}$ preserves the irreducibility of the irreducible factors of $F(\mathbf{x}, t, y)$.

Factorizing variable-reduced polynomial: Once we have a variable reduced polynomial, any off-the-shelf factorization algorithm (such as [Lec07]) may be employed to factorize $F \circ \mathscr{G}$ in time $\text{poly}((sd)^{|\mathbf{w}|})$. Computing the coefficients of $g_k(y)$ can be done via interpolation (Lemma 2.2).

Building the factors: Let $\tilde{F}(\mathbf{w}, t, y) = F \circ \mathscr{G} = \tilde{G}_1 \cdots \tilde{G}_k$. By Lemma 5.1, we have that $F = G_1 \cdots G_k$ is the decomposition of F into irreducibles with $\tilde{G}_j(\mathbf{w}, t, y) = G_j \circ \mathscr{G}$. Consider an arbitrary irreducible factor $G_j(\mathbf{x}, t, y)$ of F with coefficients over the field $\mathbb{F}$. By Lemma 5.4, we have that

$G_j(\mathbf{x}, t, y) = Q_U(\mathbf{x}, t, y)$ for an appropriate set $U \subseteq [r]$. Since the generator is only applied to the $\mathbf{x}$ variables, we have that $\tilde{G}_j(\mathbf{w}, 0, y) = G_j(\mathbf{x} \circ \mathscr{G}, 0, y) = G_j(\mathbf{0}, 0, y) = \prod_{i \in U}(y - \alpha_i)$. Since we have already computed $\tilde{G}_j$, we have the coefficients of $G_j(\mathbf{x}, 0, y) = G_j(\mathbf{0}, 0, y)$ which are the elementary symmetric polynomials of $\{\alpha_i \,:\, i \in U\}$. As in the proof of Theorem 4.3, we can use Theorem 4.2 to compute a $\mathrm{poly}(s)$ size, depth $\Delta + O(1)$ circuit for G_j.

Undo pre-processing: Now that we have obtained the irreducible factors of $F(\mathbf{x}, t, y) = \Psi_{\mathbf{a},\mathbf{b}}(P(\mathbf{x}))$, Lemma 2.11 provides the inverse transformation to obtain the corresponding factors of $P(\mathbf{x})$.

Running time: Finding the right pre-processing map $\Psi_{\mathbf{a},\mathbf{b}}$ using $\mathscr{G}$ takes time $\mathrm{poly}(s, d)^{O(n^\varepsilon)}$. The factorization of the variable-reduced polynomial also runs in time $\mathrm{poly}(s, d)^{O(n^\varepsilon)}$. Rest of the steps take time $\mathrm{poly}(s, d)$. Thus, the total running time is $\mathrm{poly}(s, d)^{O(n^\varepsilon)}$.

This completes the proof correctness of Theorem 5.2 modulo the proof of Lemma 5.1. □

5.1 Proof of Lemma 5.1

At the core of the algorithm of Bhattacharjee et al [BKR+25] was a method to characterize variable reductions that preserve the factorization structure of the polynomial $F(\mathbf{x}, t, y)$. Recall that $F(\mathbf{x}, t, y)$ is monic in y and $F(\mathbf{0}, 0, y)$ is a squarefree.

Lemma 5.4 (Lemma 8.3 in [BKR+25])**.** *Let $\{\alpha_1, \ldots, \alpha_r\}$ be the roots of $F(\mathbf{0}, 0, y)$ in $\overline{\mathbb{F}}$. For a subset $S \subset [r]$, define*

$$Q_S(\mathbf{x}, t, y) = \mathrm{Hom}_{\leq d}\left(\prod_{i \in S}(y - R_F(\alpha_i))\right)$$

Then the factors of $F(\mathbf{x}, t, y)$ over the field $\mathbb{F}$ is exactly the same as

$$\mathscr{F} = \left\{ Q_U(\mathbf{x}, t, y) \;:\; \begin{array}{c} U \subseteq [r] \text{ where } Q_U(\mathbf{x}, t, y) \in \overline{\mathbb{F}}[\mathbf{x}, t, y] \text{ divides } F \text{ and} \\ Q_U(\mathbf{x}, 0, y) = Q_U(\mathbf{0}, 0, y) = \prod_{i \in U}(y - \alpha_i) \in \mathbb{F}[y] \end{array} \right\}. \qquad \square$$

Let $G(\mathbf{x}, t, y) \in \mathbb{F}[\mathbf{x}, t, y]$ be an arbitrary irreducible factor of $F(\mathbf{x}, t, y)$. By the above lemma, there exists some $U \subseteq [d]$ such that $G(\mathbf{x}, t, y) = Q_U(\mathbf{x}, t, y)$ with $Q_U(\mathbf{0}, 0, y)$ having coefficients in $\mathbb{F}$. Note that, since $F(\mathbf{0}, 0, y)$ is squarefree, distinct elements of $\mathscr{F}$ have distinct set of roots when $\mathbf{x}, t$ are set to zero. Since $G(\mathbf{x}, t, y)$ is irreducible, the set U is minimal in the sense that for every $\emptyset \neq U' \subsetneq U$, we have that $Q_{U'}(\mathbf{x}, t, y)$ has coefficients outside $\mathbb{F}$ or does not divide $F(\mathbf{x}, t, y)$.

For the sake of contradiction, assume that $\tilde{G}(\mathbf{w}, t, y) := G \circ \mathscr{G}$ is reducible and $h(\mathbf{w}, t, y) \in \mathbb{F}[\mathbf{w}, t, y]$ is a non-trivial factor of $\tilde{G}$. Then, $h(\mathbf{0}, 0, y)$ divides $\tilde{G}(\mathbf{0}, 0, y) = G(\mathbf{0}, 0, y)$ and the set of

roots of $h(\mathbf{0}, 0, y)$ in $\overline{\mathbb{F}}$ is $\{\alpha_i : i \in U'\}$ for some $\emptyset \neq U' \subsetneq U$.

Consider the polynomial $Q_{U'}(\mathbf{x}, t, y)$. As in the proof of Theorem 4.3, since $h(\mathbf{0}, 0, y)$ has coefficients in $\mathbb{F}$ and the $Q_{U'}$ is symmetric with respect to the set $\{\alpha_i : i \in U'\}$, we have that $Q_{U'}(\mathbf{x}, t, y)$ has all coefficients in $\mathbb{F}$ as well. Hence, as argued above, $Q_{U'}(\mathbf{x}, t, y)$ does not divide $F(\mathbf{x}, t, y)$.

However, applying Lemma 5.4 for $\tilde{G}(\mathbf{w}, t, y)$, we have

$$h(\mathbf{w}, t, y) = \mathrm{Hom}_{\leq d}\left(\prod_{i \in U'}(y - R_{\tilde{G}}(\alpha_i))\right)$$

As $F(\mathbf{0}, 0, y) = \tilde{F}(\mathbf{0}, 0, y)$ is squarefree and $\tilde{F}(\mathbf{0}, 0, \alpha_i) = 0$ for each $\alpha_i \in U'$, by Lemma 2.14 there is a unique power series root $\tilde{\varphi}_i(\mathbf{w}, t)$ for $\tilde{F}$ mod t^{d+1} that satisfies $\tilde{F}(\mathbf{w}, t, \tilde{\varphi}_i) = 0$ and $\tilde{\varphi}_i(0) = \alpha_i$. Note that both $R_{\tilde{G}}(\alpha_i)$ and $R_F(\alpha_i) \circ \mathscr{G}$ satisfy these properties. Hence, by the uniqueness of the power series modulo t^{d+1}, we have

$$\begin{aligned}
& R_F(\alpha_i) \circ \mathscr{G} = R_{\tilde{G}}(\alpha_i) \bmod t^{d+1} \\
\implies\ & h(\mathbf{w}, t, y) = \mathrm{Hom}_{\leq d}\left(\prod_{i \in U'}(y - R_{\tilde{G}}(\alpha_i))\right) \\
& = \mathrm{Hom}_{\leq d}\left(\prod_{i \in U'}(y - R_F(\alpha_i) \circ \mathscr{G})\right) \\
& = \mathrm{Hom}_{\leq d}\left(\prod_{i \in U'}(y - R_F(\alpha_i))\right) \circ \mathscr{G} = Q_{U'} \circ \mathscr{G}.
\end{aligned}$$

Therefore, we have that $Q_{U'}(\mathbf{x}, t, y) \in \mathbb{F}[\mathbf{x}, t, y]$ does not divide $F(\mathbf{x}, t, y)$ but $h(\mathbf{w}, t, y) = Q_{U'} \circ \mathscr{G} \in \mathbb{F}[\mathbf{w}, t, y]$ does divide $\tilde{F}(\mathbf{w}, t, y) = F(\mathbf{w}, t, y) \circ \mathscr{G}$. It turns out that divisiblity testing of a pair of polynomials can be reduced to an appropriate polynomial identity test. This reduction was first observed by Forbes [For15] and then crucially used in deterministic factorization algorithms [KRS23, KRSV24, DST24, BKR+25]. We give below a lemma from [BKR+25] that implements the reduction in [For15] via the results of [AW24].

Lemma 5.5. *(Lemma 8.9 in [BKR+25]) Let $D \geq t \geq 0$ be integer parameters. Let $\mathbb{F}$ be any field of characteristic zero or large enough. Then, there is a constant-depth $\mathrm{poly}(D, t)$-sized circuit $\mathrm{DivTest}_{D,t}$ on $D + t + 1$ variables, that takes $(D + t)$ inputs labelled $f_0, \ldots, f_{D-1} \in \mathbb{F}$ and $g_0, \ldots, g_{t-1} \in \mathbb{F}$ respectively, such that*

$$\mathrm{DivTest}_{D,t}(y, f_0, \ldots, f_{D-1}, g_0, \ldots, g_{t-1}) = 0$$

if and only if the polynomial $f(y) = f_0 + f_1 y + \cdots + f_{D-1} y^{D-1} + y^D$ divides the polynomial $g(y) =$

$g_0 + g_1 y + \cdots + g_{t-1} y^{t-1} + y^t$. □

Define $C(\mathbf{x}, t, y) := \mathrm{DivTest}(y, \overline{\mathrm{coeff}}_y(F(\mathbf{x}, t, y)), \overline{\mathrm{coeff}}_y(Q_{U'}(\mathbf{x}, t, y)))$ where $\overline{\mathrm{coeff}}_y(F)$ refers to the vector of coefficients when F is interpreted as a univariate in y (with coefficients involving the other variables). By Lemma 5.5, we have that $C(\mathbf{x}, t, y)$ is a nonzero polynomial since $Q_{U'}$ does not divide F but $C \circ \mathscr{G}$ is zero since $Q_{U'} \circ \mathscr{G}$ divides $F \circ \mathscr{G}$. But C is a circuit of size $\mathrm{poly}(s)$ and depth $\Delta + O(1)$ and hence this violates the assumption that $\mathscr{G}$ is a hitting set generator for this class. Hence, we must have that $G \circ \mathscr{G}$ continues to be irreducible for every irreducible factor $G(\mathbf{x}, t, y)$ of $F(\mathbf{x}, t, y)$. □(Lemma 5.1)

Remark 5.6. *In [BKR+25], the polynomials $R_F(\alpha_i)$ were instead replaced by truncated power series obtained via Newton Iteration, and therefore it was not known if the polynomials $Q \in \mathscr{F}$ are computable by constant-depth circuits. As a consequence, [BKR+25] could not provide a polynomial size constant depth upper bound for the above circuit C. Thus, [BKR+25] involved a fairly delicate argument to show that the [LST21]+[KI04]+[CKS19] generator maintains the nonzeroness of these non-divisibility identity tests that arises from approximate power series roots obtained via Newton Iteration.*

With Theorem 4.1, we now can argue that the circuit C above is indeed a polynomial size constant-depth circuit and hence any *generator for this class of circuits would preserve the factorization of $F(\mathbf{x}, t, y)$.* ◊

5.2 Deterministic factorization from hitting-set generators

The algorithm in Theorem 5.2 and its analysis via Lemma 5.1 proves a more general statement.

Corollary 5.7. *(Informal) Let $\mathbb{F}$ be a field of characteristic zero or large enough, and let $\mathscr{C}$ be a* robust enough *class of circuits that is $\mathscr{C}$ is closed under small sums and products, substitution by sparse polynomials (thereby admitting interpolation). Consider the larger class $\mathscr{C}'$ computing polynomials of the form $F(g_1, \ldots, g_m)$ where F is computable by a $\mathrm{poly}(m)$-sized constant-depth circuit, and each $g_i \in \mathscr{C}$.*

If we have a blackbox PIT for the class $\mathscr{C}'$ running in time $T(n)$, then we have a deterministic $T(\mathrm{poly}(n))$-time algorithm to factorize polynomials from the class $\mathscr{C}$.

This, for natural subclasses of algebraic circuits — such as algebraic formulas, algebraic branching programs, algebraic circuits, etc. — we have a reduction from factorization to polynomial identity testing. This generalizes the result of Kopparty, Saraf and Shpilka [KSS15] who established this connection for the class of general algebraic circuits. Further, by solving each PIT instance by random sampling (and using Lemma 2.1), we get an efficient randomized algorithm that takes a polynomial from $\mathscr{C}$ as input and outputs circuits in $\mathscr{C}$ for each irreducible factor, where $\mathscr{C}$ is some robust enough class of polynomials.

However, Theorem 5.2 appears to require *blackbox* PITs for the class $\mathscr{C}'$, whereas [KSS15] established such connections even in the whitebox setting. It is an intriguing open question if efficient whitebox algorithms for PIT of $\mathscr{C}'$ would imply efficient deterministic factoring algorithms for $\mathscr{C}$.

6 Other applications

6.1 Hardness-randomness trade-offs for constant-depth circuits

An immediate consequence of the closure theorems is that we get better hardness-randomness trade-offs for constant-depth circuits directly from the Kabanets-Impagliazzo hitting-set generator [KI04].

Theorem 6.1 (Hardness-randomness for constant-depth circuits)**.** *Let $\mathbb{F}$ be any field of characteristic 0 or large enough. Fix any $\Delta > 0$. Suppose there is an explicit family $\{f_m(x_1, \ldots, x_m)\}_{m \geq 0}$ of polynomials with $\deg(f_m) \leq m$ that requires depth Δ circuits of size $B(m)$ to compute them. Then, there is a family $\{\mathscr{H}_n\}$ of explicit hitting sets for the class polynomial size circuits of depth at most $\Delta - O(1)$ such that*

$$|\mathscr{H}_n| = n^{O((B^{-1}(n))^2 / \log n)}.$$

In particular,

- *If $B(m) = 2^{\Omega(m)}$, then $|\mathscr{H}_n| = n^{O(\log n)}$.*
- *If $B(m) = 2^{m^\varepsilon}$ for some $\varepsilon > 0$, then $|\mathscr{H}_n| = n^{O(\log n)^c}$ for some $c > 0$.*
- *If $B(m) = m^{\omega(1)}$, then $|\mathscr{H}_n| \leq n^{O(n^\varepsilon)}$ for every $\varepsilon > 0$.*

Proof sketch. The proof is exactly the same as the standard Kabanets-Impagliazzo generator for general circuits, except that instead of using Kaltofen's result for closure of general circuits under roots, we use Theorem 4.1 instead. □

6.2 Border version of the factor conjecture

Another consequence of the techniques in this paper is a conceptually simpler alternative proof of a result of Bürgisser that shows that *low degree* factors of polynomials with small circuits (but potentially exponentially high degree) are in the border of small circuits. We recall the formal theorem and discuss its proof below.

Theorem 6.2 (Bürgisser [Bür04])**.** *Assume that $\mathrm{char}(\mathbb{F}) = 0$. Suppose $P(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ is computable by a size s circuit (of possibly exponential degree) and g is a factor of P. Then, the $\overline{\mathrm{size}}(g) = \mathrm{poly}(s, \deg(g))$.*

Proof. As in the previous cases, it would suffice to prove a $\overline{\mathrm{size}}$ upper bound for truncated power series roots of P. By working with a suitable shift, and the substitution $x_i \mapsto t x_i$, we can assume we have $P(t, y) \in \mathbb{K}[t, y]$ where $\mathbb{K} = \mathbb{F}(\mathbf{x})$ with $\varphi(t) \in \mathbb{K}[\![t]\!]$ satisfying $\varphi(0) = 0$ and

$$P(t, y) = (y - \varphi(t))^e \cdot (1 + Q(t, y))$$

with $Q(0, 0) = 0$.

By Theorem 3.1, we have

$$\varphi(t) = \mathscr{D}\left(\frac{y^2 \cdot \partial_y P(ty,y)}{e \cdot P(ty,y)}\right) = \mathscr{D}\left(\frac{y \cdot \partial_y P(ty,y)/y^{e-1}}{e \cdot P(ty,y)/y^e}\right)$$
$$\implies [t^n]\{\varphi(t)\} = [t^n y^n]\left\{\frac{y \cdot \partial_y P(ty,y)/y^{e-1}}{e \cdot P(ty,y)/y^e}\right\}$$

Since $P(t,y) = (y-\varphi(t))^e \cdot (1+Q(t,y))$, we have

$$\frac{P(ty,y)}{y^e} = \left(1 - \frac{\varphi(ty)}{y}\right)^e \cdot (1+Q(ty,y)) = 1 - R(t,y)$$

for some $R(t,y) \in \mathbb{K}[\![t,y]\!]$ with $R(0,0) = 0$.

Given a circuit C for $P(t,y)$, we now have circuits $C_1(t,y), C_2(t,y)$ of $O(s)$ size such that C_1 has a single division by y^{e-1} computing $\partial_y P(ty,y)/(e \cdot y^{e-1})$, and C_2 has a single division by y^e and computes $R(t,y)$. Therefore,

$$\begin{aligned}
[t^n y^n]\left\{\frac{y \cdot \partial_y P(ty,y)/y^{e-1}}{e \cdot P(ty,y)/y^e}\right\} &= [t^n y^n]\left\{\frac{y \cdot C_1}{1 - C_2}\right\} \\
&= [t^n y^n]\left\{y \cdot C_1 \cdot \left(1 + C_2 + C_2^2 + \cdots\right)\right\} \\
&= [t^n y^n]\left\{y \cdot C_1 \cdot \left(1 + C_2 + C_2^2 + \cdots + C_2^{2n}\right)\right\} \\
&=: [t^n y^n]\{C_{3,n}(t,y)\}
\end{aligned}$$

Note that $C_{3,n}$ is a circuit of size $\mathrm{poly}(s,n)$ with divisions only by powers of y. By the border interpolation (Lemma 2.4), we can choose nonzero $\alpha_0^{(n)}, \ldots, \alpha_n^{(n)}, \beta_0^{(n)}, \ldots, \beta_n^{(n)} \in \mathbb{F}(\varepsilon)$ such that

$$\sum_{i=0}^{n} \beta_i^{(n)} \cdot C_{3,n}(t, \alpha_i^{(n)}) = [y^n]\{C_{3,n}(t,y)\} + O(\varepsilon)$$

and the LHS is now a division-free circuit $C_{4,n}(t)$ of size $\mathrm{poly}(s,n)$. Once again,

$$\begin{aligned}
\sum_{i=0}^{n} \beta_i^{(n)} \cdot C_{4,n}(\alpha_i^{(n)}) &= \sum_{i,j=0}^{n} \beta_i^{(n)} \beta_j^{(n)} \cdot C_{3,n}(\alpha_i^{(n)}, \alpha_j^{(n)}) \\
&= [t^n]\{C_{4,n}(t)\} + O(\varepsilon) = [t^n y^n]\{C_{3,n}(t,y)\} + O(\varepsilon)
\end{aligned}$$

Thus, we have a circuit $C_5(t) \in \mathbb{K}(\varepsilon)[t]$ of size $\mathrm{poly}(n,s)$ defined by

$$C_5(t) := \sum_{r=1}^{n} t^r \cdot \left(\sum_{i,j=0}^{r} \beta_i^{(r)} \beta_j^{(r)} \cdot C_{3,r}(\alpha_i^{(r)}, \alpha_j^{(r)})\right)$$

such that $C_5(t) = \mathrm{Hom}_{\leq n}(\varphi(t)) + O(\varepsilon)$. Therefore, $\overline{\mathrm{size}}(\mathrm{Hom}_{\leq n}(\varphi(t))) = \mathrm{poly}(s, n)$. □

7 Open questions

We conclude with some open problems.

1. **Whitebox PIT to deterministic factorization:** Kopparty, Saraf and Shpilka [KSS15] showed that efficient algorithms for PIT for the class of general circuits leads to efficient deterministic factorization of general circuits, and this connection is for both the whitebox and the blackbox setting for PITs. Although Corollary 5.7 extends the blackbox connection to other natural subclasses of circuits (such as formulas, branching programs, constant-depth circuits), establishing a similar connection in the whitebox setting remains open.

2. **Computing p-th roots of circuits:** One of the simplest-to-state open problems in the area of factorization of algebraic circuits is the following — over a characteristic p field, if a polynomial f^p is an n-variate, degree d polynomial computed by a $\mathrm{poly}(n, d)$-sized circuit, is f also computable by a $\mathrm{poly}(n, d)$-sized circuit? The answer to this question is unknown even for the setting of general algebraic circuits.

Acknowledgements: The discussions leading to this work started when a subset of the authors were at the workshop on Algebraic and Analytic Methods in Computational Complexity (Dagstuhl Seminar 24381) at Schloss Dagstuhl, and continued when they met again during the HDX & Codes workshop at ICTS-TIFR in Bengaluru. We are thankful to the organisers of these workshops and to the staff at these centers for the wonderful collaborative atmosphere that facilitated these discussions.

Varun Ramanathan is grateful to Srikanth Srinivasan and Amik Raj Behera at the University of Copenhagen, and Nutan Limaye and Prateek Dwivedi at ITU Copenhagen, for the helpful discussions on polynomial factorization.

References

[AF22] Robert Andrews and Michael A. Forbes. Ideals, determinants, and straightening: proving and using lower bounds for polynomial ideals. In *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, pages 389–402. ACM, 2022.

[AW24] Robert Andrews and Avi Wigderson. Constant-Depth Arithmetic Circuits for Linear Algebra Problems. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 2367–2386, 2024.

[BKR+25] Somnath Bhattacharjee, Mrinal Kumar, Varun Ramanathan, Ramprasad Saptharishi, and Shubhangi Saraf. Deterministic factorization of constant-depth algebraic circuits in subexponential time. *CoRR*, abs/2504.08063, 2025. Pre-print available at `arXiv:2504.08063`.

[BSV20] Vishwas Bhargava, Shubhangi Saraf, and Ilya Volkovich. Deterministic Factorization of Sparse Polynomials with Bounded Individual Degree. *J. ACM*, 67(2):8:1–8:28, 2020.

[Bür04] Peter Bürgisser. The Complexity of Factors of Multivariate Polynomials. *Found. Comput. Math.*, 4(4):369–396, 2004.

[CKS19] Chi-Ning Chou, Mrinal Kumar, and Noam Solomon. Closure Results for Polynomial Factorization. *Theory of Computing*, 15(13):1–34, 2019.

[DSS22] Pranjal Dutta, Nitin Saxena, and Amit Sinhababu. Discovering the Roots: Uniform Closure Results for Algebraic Classes Under Factoring. *J. ACM*, 69(3), June 2022.

[DST24] Pranjal Dutta, Amit Sinhababu, and Thomas Thierauf. Derandomizing Multivariate Polynomial Factoring for Low Degree Factors, 2024. Pre-print available at `arXiv:2411.17330`.

[DSY09] Zeev Dvir, Amir Shpilka, and Amir Yehudayoff. Hardness-Randomness Tradeoffs for Bounded Depth Arithmetic Circuits. *SIAM J. Comput.*, 39(4):1279–1293, 2009.

[For14] Michael Andrew 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 *Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, FOCS '15, page 451–465, USA, 2015. IEEE Computer Society.

[FS15] Michael A. Forbes and Amir Shpilka. Complexity Theory Column 88: Challenges in Polynomial Factorization1. *SIGACT News*, 46(4):32–49, dec 2015.

[Fur67] Harry Furstenberg. Algebraic functions over finite fields. *Journal of Algebra*, 7(2):271–277, 1967.

[Ges16] Ira M. Gessel. Lagrange inversion. *Journal of Combinatorial Theory, Series A*, 144:212–249, 2016. Fifty Years of the Journal of Combinatorial Theory.

[GK85] J. von zur Gathen and E. Kaltofen. Factoring Sparse Multivariate Polynomials. *Journal of Computer and System Sciences*, 31(2):265–287, 1985.

[GRS23] Venkatesan Guruswami, Atri Rudra, and Madhu Sudan. *Essential coding theory*. 2023.

[Hu16] Yining Hu. Explicit Implicit Function Theorem for All Fields, 2016. Pre-print available at `arXiv:1611.01415`.

[Kal82] Erich L. Kaltofen. A Polynomial-Time Reduction from Bivariate to Univariate Integral Polynomial Factorization. In *23rd Annual Symposium on Foundations of Computer Science, Chicago, Illinois, USA, 3-5 November 1982*, pages 57–64. IEEE Computer Society, 1982.

[Kal85] Erich Kaltofen. Polynomial-Time Reductions from Multivariate to Bi- and Univariate Integral Polynomial Factorization. *SIAM Journal of Computing*, 14(2):469–489, 1985.

[Kal89] Erich Kaltofen. Factorization of Polynomials Given by Straight-Line Programs. In *Randomness and Computation*, pages 375–412. JAI Press, 1989.

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing Polynomial Identity Tests Means Proving Circuit Lower Bounds. *Computational Complexity*, 13(1-2):1–46, 2004. Preliminary version in the *35th Annual ACM Symposium on Theory of Computing (STOC 2003)*.

[KRS23] Mrinal Kumar, Varun Ramanathan, and Ramprasad Saptharishi. Deterministic Algorithms for Low Degree Factors of Constant Depth Circuits. *CoRR*, abs/2309.09701, 2023. Pre-print available at `arXiv:2309.09701`.

[KRSV24] Mrinal Kumar, Varun Ramanathan, Ramprasad Saptharishi, and Ben Lee Volk. Towards Deterministic Algorithms for Constant-Depth Factors of Constant-Depth Circuits, 2024. Pre-print available at `arXiv:2403.01965`.

[KSS15] Swastik Kopparty, Shubhangi Saraf, and Amir Shpilka. Equivalence of Polynomial Identity Testing and Polynomial Factorization. *Computational Complexity*, 24(2):295–331, 2015. Preliminary version in the *29th Annual IEEE Conference on Computational Complexity (CCC 2014)*.

[KT88] Erich L. Kaltofen and Barry M. Trager. Computing with Polynomials Given By Black Boxes for Their Evaluation: Greatest Common Divisors, Factorization, Separation of Numerators and Denominators. In *29th Annual Symposium on Foundations of Computer Science, White Plains, New York, USA, 24-26 October 1988*, pages 296–305. IEEE Computer Society, 1988.

[Lec07] Grégoire Lecerf. Improved dense multivariate polynomial factorization algorithms. *Journal of Symbolic Computation*, 42(4):477–494, 2007.

[LST21] Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. Superpolynomial Lower Bounds Against Low-Depth Algebraic Circuits. In *Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2021)*, pages 804–814. IEEE, 2021. Preliminary version in the Electronic Colloquium on Computational Complexity (ECCC), Technical Report TR21-081.

[Oli16] Rafael Oliveira. Factors of low individual degree polynomials. *Comput. Complex.*, 25(2):507–561, 2016.

[Sok09] Alan D. Sokal. A ridiculously simple and explicit implicit function theorem, 2009. Pre-print available at `arXiv:0902.0069`.

[ST20] Amit Sinhababu and Thomas Thierauf. Factorization of Polynomials Given By Arithmetic Branching Programs. In *35th Computational Complexity Conference (CCC 2020)*, volume 169 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 33:1–33:19, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[Sta99] Richard Stanley. *Enumerative Combinatorics - Volume 2*. Cambridge University Press, 1999.

[SW23] Erlang Surya and Lutz Warnke. Lagrange Inversion Formula by Induction. *The American Mathematical Monthly*, 130(10):944–948, 2023. Pre-print available at `arXiv:2305.17576`.

[vzGG13] Joachim von zur Gathen and Jürgen Gerhard. *Modern Computer Algebra*. Cambridge University Press, 3 edition, 2013.

A Extending closure results to fields of small characteristic

The closure results (Theorem 4.1, Theorem 4.3) over zero or large characteristic fields extend to small characteristic fields, with some caveats. Suppose $P(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ (where $\mathrm{char}(\mathbb{F}) = p$) has a constant-depth circuit, and let $g(\mathbf{x})$ be any irreducible factor of $P(\mathbf{x})$ with multiplicity $p^\ell e$ satisfying $\gcd(p, e) = 1$. Then, we show that $g(\mathbf{x})^{p^\ell}$ has a constant-depth circuit over $\overline{\mathbb{F}}$, the algebraic closure of $\mathbb{F}$. These results follow due to a version of Furstenberg's theorem over small characteristic fields, which we state and prove below. Note that the case of roots of multiplicity 1 already follows from the original version of Furstenberg's theorem. In the following theorem we show how to extend it to higher-order multiplicity roots.

A.1 Furstenberg's theorem over small characteristic fields

We shall work with the notion of Hasse derivative, which is the standard alternative to partial derivatives in the small characteristic setting. We state the definition and the product rule for Hasse derivatives. For more details, we recommend the reader to refer to [For14, Appendix C].

Definition A.1 (Hasse derivatives)**.** *The* Hasse Derivative of order i *of* $F(t,y) \in \mathbb{F}[t,y]$ *with respect to* y*, denoted as* $\mathrm{D}_y^{(i)}(F)$*, is defined as the coefficient of* z^i *in the polynomial* $F(t, y+z)$*.* $\lozenge$

Lemma A.2 (Product rule for Hasse derivatives)**.** *Let* $G(t,y), H(t,y) \in \mathbb{F}[t,y]$ *be bivariate polynomials and let* $k \geq 0$*. Then,*

$$\mathrm{D}_y^{(k)}(GH) = \sum_{i+j=k} \mathrm{D}_y^{(i)}(G) \cdot \mathrm{D}_y^{(j)}(H)$$

The following version of Furstenberg's theorem over small characteristic is very similar to Theorem 3.1, with some key differences. The theorem expresses an appropriate power of a power series root of a polynomial as a diagonal of a rational expression involving the polynomial and its derivatives.

Theorem A.3 (Furstenberg's theorem over small characteristic fields)**.** *Let* $\mathbb{F}$ *be a field of characteristic* p*. Let* $P(t,y) \in \mathbb{F}[\![t,y]\!]$ *be a power series and* $\varphi(t) \in \mathbb{F}[\![t]\!]$ *be a power series satisfying*

$$P(t,y) = (y - \varphi(t))^{p^\ell e} \cdot Q(t,y)$$

for some $\ell \geq 0, e \geq 1$ *such that* $\gcd(p,e) = 1$*. If* $\varphi(0) = 0$ *and* $Q(0,0) \neq 0$*, then*

$$\varphi^{p^\ell} = \mathscr{D}\left(\frac{y^{2p^\ell} \cdot \mathrm{D}_y^{(p^\ell)}(P)(ty,y)}{e \cdot P(ty,y)}\right) \tag{A.4}$$

Proof. Firstly, observe that

$$\begin{aligned}\mathrm{D}_y^{(j)}((y-\varphi(t))^{p^\ell e}) &= [z^j]\left\{(y+z-\varphi(t))^{p^\ell e}\right\} \\ &= [z^j]\left\{(y^{p^\ell} + z^{p^\ell} - \varphi(t)^{p^\ell})^e\right\}\end{aligned}$$

Hence, $\mathrm{D}_y^{(j)}((y-\varphi(t))^{p^\ell e}) = 0$ for all $0 < j < p^\ell$, and

$$\mathrm{D}_y^{(p^\ell)}((y-\varphi(t))^{p^\ell e}) = e \cdot (y-\varphi(t))^{p^\ell(e-1)}$$

By applying product rule for Hasse derivatives (Lemma A.2), $\mathrm{D}_y^{(p^\ell)}(P)(t,y)$ simplifies to

$$\begin{aligned}
\mathrm{D}_y^{(p^\ell)}(P)(t,y) &= \sum_{i+j=p^\ell} \mathrm{D}_y^{(i)}((y-\varphi(t))^{p^\ell e}) \cdot \mathrm{D}_y^{(j)}(Q)(t,y) \\
&= \mathrm{D}_y^{(p^\ell)}((y-\varphi(t))^{p^\ell e}) \cdot \mathrm{D}_y^{(0)}(Q)(t,y) + \mathrm{D}_y^{(0)}((y-\varphi(t))^{p^\ell e}) \cdot \mathrm{D}_y^{(p^\ell)}(Q)(t,y) \\
&= e \cdot (y-\varphi(t))^{p^\ell(e-1)} \cdot Q(t,y) + (y-\varphi(t))^{p^\ell e} \cdot \mathrm{D}_y^{(p^\ell)}(Q)(t,y)
\end{aligned}$$

Following along the lines of proof of Theorem 3.1,

$$\begin{aligned}
P(t,y) &= (y-\varphi(t))^{p^\ell e} Q(t,y) \\
\implies \frac{\mathrm{D}_y^{(p^\ell)}(P)(t,y)}{P(t,y)} &= \frac{e}{(y-\varphi(t))^{p^\ell}} + \frac{\mathrm{D}_y^{(p^\ell)}(Q)(t,y)}{Q(t,y)} \\
\implies \frac{y^{2p^\ell} \cdot \mathrm{D}_y^{(p^\ell)}(P)(ty,y)}{e \cdot P(ty,y)} &= \frac{y^{2p^\ell}}{(y-\varphi(ty))^{p^\ell}} + \frac{y^{2p^\ell} \cdot \mathrm{D}_y^{(p^\ell)}(Q)(ty,y)}{e \cdot Q(ty,y)} \\
\implies \mathscr{D}\left(\frac{y^{2p^\ell} \cdot \mathrm{D}_y^{(p^\ell)}(P)(ty,y)}{e \cdot P(ty,y)}\right) &= \mathscr{D}\left(\frac{y^2}{y-\varphi(ty)}\right)^{p^\ell} + \mathscr{D}\left(\frac{y^{2p^\ell} \cdot \mathrm{D}_y^{(p^\ell)}(Q)(ty,y)}{e \cdot Q(ty,y)}\right).
\end{aligned}$$

As in the proof of Theorem 3.1, the second term in the RHS is zero and

$$\mathscr{D}\left(\frac{y^2}{y-\varphi(ty)}\right) = \varphi(t) \implies \mathscr{D}\left(\frac{y^2}{y-\varphi(ty)}\right)^{p^\ell} = (\varphi(t))^{p^\ell} \qquad \square$$

We can further simplify the expression in Theorem A.3 to get a version of Corollary 3.3 over small characteristic fields.

Corollary A.5 (Corollary 3.3 for small characteristic)**.** *Let $P(t,y), Q(t,y) \in \mathbb{F}[\![t,y]\!]$ and $\varphi(t) \in \mathbb{F}[\![t]\!]$ satisfy*

$$P(t,y) = (y-\varphi(t))^{p^\ell e} \cdot Q(t,y)$$

with $\gcd(p,e) = 1$, $\varphi(0) = 0$ and $Q(0,0) = \alpha \neq 0$. Then,

$$\varphi(t)^{p^\ell} = \sum_{m \geq 0} [y^{p^\ell(e(m+1)-2)}] \left\{ \frac{\mathrm{D}_y^{(p^\ell)}(P)(t,y)}{e \cdot \alpha^{m+1}} \Big(\alpha y^{p^\ell \cdot e} - P(t,y) \Big)^m \right\}.$$

Moreover,

$$\mathrm{Hom}_{\leq d}[\varphi(t)^{p^\ell}] = \mathrm{Hom}_{\leq d}\left[\sum_{m \geq 0}^{2e(d+p^\ell)} [y^{p^\ell(e(m+1)-2)}] \left\{ \frac{\mathrm{D}_y^{(p^\ell)}(P)(t,y)}{e \cdot \alpha^{m+1}} \Big(\alpha y^{p^\ell \cdot e} - P(t,y) \Big)^m \right\} \right].$$

Proof. By dividing P by α, let us assume without loss of generality that $Q(0,0) = 1$.

$$\begin{aligned}
\varphi^{p^\ell} &= \mathscr{D}\left(\frac{y^{2p^\ell} \cdot \mathrm{D}_y^{(p^\ell)}(P)(ty,y)}{e \cdot P(ty,y)}\right) \\
[t^k]\left\{\varphi^{p^\ell}\right\} &= [t^k y^k]\left\{\frac{y^{2p^\ell} \cdot \mathrm{D}_y^{(p^\ell)}(P)(ty,y)}{e \cdot P(ty,y)}\right\} = [t^k y^k]\left\{\frac{y^{p^\ell(2-e)} \cdot \mathrm{D}_y^{(p^\ell)}(P)(ty,y)}{e \cdot P(ty,y)/y^{p^\ell \cdot e}}\right\} \\
&= [t^k y^k]\left\{\frac{y^{p^\ell(2-e)} \cdot \mathrm{D}_y^{(p^\ell)}(P)(ty,y)}{e} \sum_{m\geq 0}\left(1 - \frac{P(ty,y)}{y^{p^\ell \cdot e}}\right)^m\right\} \\
&= [t^k y^k]\left\{\sum_{m\geq 1}\frac{y^{p^\ell(2-e(m+1))} \cdot \mathrm{D}_y^{(p^\ell)}(P)(ty,y)}{e}\left(y^{p^\ell \cdot e} - P(ty,y)\right)^m\right\} \\
&= \sum_{m\geq 0}[t^k y^{k-p^\ell(2-e(m+1))}]\left\{\frac{\mathrm{D}_y^{(p^\ell)}(P)(ty,y)}{e}\left(y^{p^\ell \cdot e} - P(ty,y)\right)^m\right\} \\
&= \sum_{m\geq 0}[t^k y^{p^\ell(e(m+1)-2)}]\left\{\frac{\mathrm{D}_y^{(p^\ell)}(P)(t,y)}{e}\left(y^{p^\ell \cdot e} - P(t,y)\right)^m\right\} \\
&= [t^k]\left\{\sum_{m\geq 0}[y^{p^\ell(e(m+1)-2)}]\left\{\frac{\mathrm{D}_y^{(p^\ell)}(P)(t,y)}{e}\left(y^{p^\ell \cdot e} - P(t,y)\right)^m\right\}\right\} \\
\therefore \varphi^{p^\ell} &= \sum_{m\geq 0}[y^{p^\ell(e(m+1)-2)}]\left\{\frac{\mathrm{D}_y^{(p^\ell)}(P)(t,y)}{e}\left(y^{p^\ell \cdot e} - P(t,y)\right)^m\right\}
\end{aligned}$$

The first part of the corollary follows from the above statement for the case of $Q(0,0) = 1$.

To obtain a finite expression for $\mathrm{Hom}_{\leq d}\, \varphi^{p^\ell}$, observe that in

$$y^{p^\ell e} - P(t,y) = y^{p^\ell \cdot e} - \left((y^{p^\ell} - \varphi(t^{p^\ell}))^e \cdot (Q(t,y))\right)$$

every monomial in t has degree at least p^ℓ. Furthermore, setting $t = 0$ reduces the above expression to $y^{p^\ell e} - y^{p^\ell e} \cdot Q(0,y)$ which is divisible by $y^{p^\ell e+1}$ (since $Q(0,0) = 1$). Therefore,

$$y^{p^\ell \cdot e} - P(t,y) = t^{p^\ell} \cdot A + y^{p^\ell e+1} \cdot B$$

for some $A, B \in \mathbb{F}[\![t,y]\!]$. Therefore, every monomial in $\left(y^{p^\ell \cdot e} - P(t,y)\right)^m$ with t-degree at most d has y-degree at least $(m - \frac{d}{p^\ell}) \cdot (p^\ell e + 1)$, which is greater than $p^\ell(e(m+1) - 2)$ when $m > d(e + \frac{1}{p^\ell}) + p^\ell(e-2)$. Thus, for $m \geq 2e(d + p^\ell)$, there is no term with t-degree at most d and

y-degree $p^\ell(e(m+1)-2)$. Therefore,

$$\mathrm{Hom}_{\le d}[\varphi(t)^{p^\ell}] = \mathrm{Hom}_{\le d}\left[\sum_{m\ge 0}^{2e(d+p^\ell)} [y^{p^\ell(e(m+1)-2)}]\left\{\frac{\mathrm{D}_y^{(p^\ell)}(P)(t,y)}{e\cdot\alpha^{m+1}}\Big(\alpha y^{p^\ell\cdot e} - P(t,y)\Big)^m\right\}\right]. \quad \square$$

A.2 Complexity of power series roots and factors over $\overline{\mathbb{F}_q}$

Using Theorem A.3 and Corollary A.5, we get the following analogue of Theorem 4.1 over arbitrary fields of small characteristic.

Theorem A.6 (Power series roots with multiplicity over small characteristic)**.** *Let $\mathbb{F}$ be a field of positive characteristic p. Suppose $P(\mathbf{x},y)\in\mathbb{F}[\mathbf{x},y]$ is a polynomial computed by a circuit C, and $\varphi(\mathbf{x})\in\mathbb{F}[\![\mathbf{x}]\!]$ is a power series satisfying $P(\mathbf{x},y) = (y-\varphi(\mathbf{x}))^{p^\ell e}\cdot Q(\mathbf{x},y)$ where $\varphi(\mathbf{0})=0$, $\gcd(p,e)=1$ and $Q(\mathbf{0},0)\neq 0$. Then, for any $d\in\mathbb{N}$, there is a circuit C' over $\overline{\mathbb{F}}$ computing $\mathrm{Hom}_{\le d}\left[\varphi^{p^\ell}\right]$ such that*

$$\mathrm{size}(C') \le \mathrm{poly}(d,\mathrm{size}(C))$$

$$\mathrm{depth}(C') \le \mathrm{depth}(C) + O(1)$$

Almost immediately, a similar statement follows for all factors of constant-depth circuits. Given a polynomial $P(\mathbf{x})$, we apply a valid pre-processing map (Definition 2.10) to get $P(t,y)$ that is monic in y and has the property that different power-series roots have different constant terms (a random shift suffices). We then apply Theorem A.6 to get small constant-depth circuits over $\overline{\mathbb{F}}$ for appropriate powers of each of the power-series roots, and then combine them (followed by a truncation) to get the following theorem.

Theorem A.7 (Complexity of factors over $\overline{\mathbb{F}_q}$)**.** *Let $\mathbb{F}_q$ be a field of positive characteristic p. Let $P(\mathbf{x})\in\mathbb{F}_q[\mathbf{x}]$ be a polynomial on n variables of degree d computed by a circuit C of size s and depth Δ. Further, let $g(\mathbf{x})$ be a factor of $P(\mathbf{x})$ with multiplicity $p^\ell\cdot e$ where $\gcd(p,e)=1$. Then, $g(\mathbf{x})^{p^\ell}$ is computable by a circuit of size $\mathrm{poly}(s,d,n)$ and depth $\Delta+O(1)$ over $\overline{\mathbb{F}_q}$.*

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il