Constant-Depth Circuits for Polynomial GCD over Any Characteristic

Somnath Bhattacharjee
University of Toronto, Canada

Mrinal Kumar
Tata Institute of Fundamental Research, Mumbai, India

Shanthanu S. Rai
Tata Institute of Fundamental Research, Mumbai, India

Varun Ramanathan
University of Copenhagen, Denmark

Ramprasad Saptharishi
Tata Institute of Fundamental Research, Mumbai, India

Shubhangi Saraf
University of Toronto, Canada

Abstract

We show that the GCD of two univariate polynomials can be computed by (piece-wise) algebraic circuits of constant depth and polynomial size over any sufficiently large field, regardless of the characteristic. This extends a recent result of Andrews & Wigderson who showed such an upper bound over fields of zero or large characteristic.

Our proofs are based on a recent work of Bhattacharjee, Kumar, Rai, Ramanathan, Saptharishi & Saraf that shows closure of constant depth algebraic circuits under factorization. On our way to the proof, we show that any n-variate symmetric polynomial P that has a small constant depth algebraic circuit can be written as the composition of a small constant depth algebraic circuit with elementary symmetric polynomials. This statement is a constant depth version of a result of Bläser & Jindal, who showed this for algebraic circuits of unbounded depth. As an application of our techniques, we also strengthen the closure results for factors of constant-depth circuits in the work of Bhattacharjee et al. over fields for small characteristic.

2012 ACM Subject Classification Theory of computation → Algebraic complexity theory

Keywords and phrases algebraic circuits, polynomial greatest common divisor, symmetric polynomials, finite fields, constant-depth circuits

Digital Object Identifier 10.4230/LIPIcs.CCC.2026.16

Related Version *Previous Version*: https://arxiv.org/abs/2506.23220v1
Previous Version: https://eccc.weizmann.ac.il/report/2025/085/

Funding *Somnath Bhattacharjee*: Research partially supported by an NSERC Discovery Grant.
Mrinal Kumar: Research Supported by the Department of Atomic Energy, Government of India, under project number RTI4014, and in part by research grants from Google Research, Premji Invest and SERB.
Shanthanu S. Rai: Research Supported by the Department of Atomic Energy, Government of India, under project number RTI4014, TCS Research Fellowship and in part by research grants from Google Research, Premji Invest and SERB.
Varun Ramanathan: This work was conducted primarily while Varun Ramanathan was a doctoral student at TIFR Mumbai, and was supported by the Department of Atomic Energy, Government of India, under project number RTI4014, and in part by research grants from Google Research, Premji Invest and SERB. Final revisions were completed as a postdoctoral researcher at the University of Copenhagen, Denmark, supported by the European Research Council (ERC) under grant agreement no. 101125652 (ALBA).

41st Computational Complexity Conference (CCC 2026).
Editor: Dana Moshkovitz; Article No. 16; pp. 16:1–16:21

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Ramprasad Saptharishi: Research Supported by the Department of Atomic Energy, Government of India, under project number RTI4014, and in part by research grants from Google Research, Premji Invest and SERB.
Shubhangi Saraf: Research partially supported by the McLean Award and an NSERC Discovery Grant.

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 organizers 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 for helpful discussions on the complexity of symmetric polynomials.

1 Introduction

A popular theme in the research on algebraic computation is that of parallelization. A typical result of this flavor shows that a seemingly sequential computational task can also be performed by algebraic circuits of surprisingly small depth. One of the earliest results of this nature is a result of Csanky [6] which shows that the $n \times n$ symbolic determinant polynomial can be efficiently computed by algebraic circuits of depth $O(\log^2 n)$. Subsequently, a significant generalization of this result by Valiant, Skyum, Berkowitz and Rackoff [11] showed that any algebraic circuit of size and degree $\text{poly}(n)$ can be converted to a circuit of depth $O(\log^2 n)$ with a $\text{poly}(n)$ blow up in size. The decades of 1980s saw many such surprising parallelizability results for algebraic problems, and we refer an interested reader to the detailed introduction in the work of Andrews & Wigderson [1] for further pointers to this literature.

Our focus in this work is to study the parallelizability of GCD computation of two univariate polynomials. More formally, we are interested in the following problem – can the GCD of two polynomials be computed by a polynomial size constant depth algebraic circuit that takes the coefficients of the polynomials as input? The aforementioned work of Andrews & Wigderson [1] showed that over fields of characteristic zero or sufficiently large characteristic, this is indeed the case, i.e. the GCD of two univariate polynomials can be computed (piece-wise) by constant depth algebraic circuits of polynomial size. Here, the "piece-wise" part is to account for the fact that the GCD of two polynomials is not a rational function of their coefficients. We define this model of piece-wise algebraic circuits formally in Definition 19, but for this introduction, we encourage the reader to think of this as a standard algebraic circuit. In addition to the GCD, Andrews & Wigderson showed a similar constant depth upper bound for many closely related algebraic problems like the LCM, the resultant, the discriminant, the inverse of the Sylvester matrix and computing square-free decomposition of a polynomial.

The key technical fact underlying the results in [1] was that many of these algebraic problems have a relatively natural constant depth algebraic circuit if the input polynomials were given via their roots, and not their coefficients. Moreover, these circuits really compute functions that are symmetric in the roots of the inputs polynomials. Thus, if one could somehow compute these symmetric functions of roots via a constant depth algebraic circuit that takes as input the coefficients of the polynomials (and not their roots), we would be done. Over fields of zero or sufficiently large characteristic, Andrews & Wigderson showed that this is indeed the case. Their primary technical ingredient for this is the fact that over

such fields, the classical Newton identities can be implemented efficiently: there are constant-depth algebraic circuits of polynomial size that transform the first n power-sum symmetric polynomials into the n elementary symmetric polynomials, and vice versa. Unfortunately, this last fact is not true when the underlying field has small (non-zero) characteristic. In fact, over such fields, power-sum symmetric polynomials of degree $1, 2, \ldots, n$ are not algebraically independent, while the elementary symmetric polynomials are, and hence we cannot expect to compute the n elementary symmetric polynomials from the first n power-sum symmetric polynomials. While this specific issue is fixable to an extent by considering power-sum symmetric polynomials of carefully chosen degrees (e.g. over fields of characteristic 2, we take the first n power-sums of odd degree), it is unclear if the transformation from this set of power-sums to elementary symmetric polynomials can be done efficiently in constant depth. This technical issue obstructs the natural attempt at extending the results in [1] to finite fields of small characteristic and necessitates the need for some new technical ideas.

In this work, we show that the GCD of two univariates can indeed be computed by polynomial size constant depth algebraic circuits (again, piece-wise) over all fields of polynomially large size, but arbitrary characteristic. We now discuss our results in more detail, before moving on to an overview of the proofs.

1.1 Our results

▶ **Theorem 1.** *There is an absolute constant $c \in \mathbb{N}$ such that the following is true.*

For all $n \in \mathbb{N}$ and any field $\mathbb{F}$ with at least n^c elements, the gcd *of two univariates over $\mathbb{F}$ of degree at most n can be computed piece-wise as a ratio of algebraic circuits of size* $\mathrm{poly}(n)$ *and depth $O(1)$.*

A more detailed version of the above theorem statement is provided in Theorem 20. The above (piece-wise) circuit yields an efficient parallel algorithm for computing the GCD of two polynomials. More precisely, there is an algebraic PRAM algorithm with polynomially many processors and constant number of rounds that takes the coefficient vectors of two univariate polynomials as input and computes their GCD. Each processor is allowed to perform basic field operations and branching instructions (based on if a variable is 0 or not).

Our proof techniques also extend to some of the other results in [1] and show that these continue to be true over all large enough fields. Perhaps the most interesting among these is that there is a constant depth algebraic circuit of $\mathrm{poly}(n)$ size that takes as input the coefficients of two univariates of degree n and computes their resultant. From Theorem 1, we also get a piece-wise small constant-depth circuit for the LCM of two polynomials using the fact that $\mathsf{LCM}(f, g) = \frac{f \cdot g}{\gcd(f,g)}$. However, our techniques do not extend to many of the other results in [1], most notably that of computing square-free decomposition and results on arbitrary manipulation of the multiplicities of roots (Section 9 in the full version of [1]).

On our way to the proof of Theorem 1, we prove the following theorem that might be independently interesting.

▶ **Theorem 2** (Complexity of symmetric polynomials)**.** *Let $n, d, s \in \mathbb{N}$ be parameters, and let $\mathbb{F}$ be a sufficiently large field, i.e. $|\mathbb{F}| \geq (nds)^c$ for an absolute constant $c > 0$. Let $P(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a symmetric polynomial on n variables of degree d and let $Q(\mathbf{z}) \in \mathbb{F}[\mathbf{z}]$ be the unique*[1] *degree d polynomial such that $P(\mathbf{x}) = Q(\mathrm{Esym}_1(\mathbf{x}), \ldots, \mathrm{Esym}_n(\mathbf{x}))$. Then, the following are true:*

[1] The existence and uniqueness of such a Q follows from the Fundamental Theorem of Symmetric Polynomials (see Theorem 11).

- *If P is computable by a circuit of size s and depth Δ, then Q can be computed by a circuit of size* $\mathrm{poly}(s, d, n)$ *and depth* $\Delta + O(1)$.
- *If P is computable by an algebraic formula of size s, then Q can be computed by a formula of size* $\mathrm{poly}(s, d, n)$.

In [3], Bläser & Jindal showed that if an n-variate symmetric polynomial P can be computed by an algebraic circuit of size s and degree d, then it can be expressed as the composition of a circuit Q of size and degree $\mathrm{poly}(s, d, n)$ with elementary symmetric polynomials. However, if P is assumed to be computable by a constant-depth circuit or a formula, it is unclear from the proof in [3] whether Q also has a small constant-depth circuit or a formula respectively. Thus, Theorem 2 is a version of the result of [3] for constant depth algebraic circuits and formulas. Showing such a statement for constant-depth circuits was mentioned as an open problem in [1] and Theorem 2 answers this question affirmatively. A partial generalization of [3] to formulas and constant-depth formulas was shown in [4, Lemma 1.4] in the context of understanding the complexity of Schur polynomials.

Our proofs are based on techniques that were recently used to show the closure of constant-depth circuits and formulas under factorization in [2] over fields of zero or large characteristic. These closure results, in turn, are a consequence of a classical theorem of Furstenberg [9] that gives a computationally simple and explicit expression for the power series roots of a bivariate polynomial (see Theorem 8 and Theorem 25).

Our proof techniques have another consequence of closure under factorization over fields of small characteristic. In [2], while the complete closure result was over fields of zero or large characteristic, a weaker statement was shown to be over fields of small characteristic. More specifically, it was shown (Theorem 1.3 in [2]) that if an n variate degree d polynomial f has a constant-depth circuit of size s over the field $\mathbb{F}_{p^k}$ (for a prime p) and g is a factor of f, then for some $i \in \mathbb{N}$, g^{p^i} has a constant-depth circuit of size $\mathrm{poly}(s, n, d)$ over the algebraic closure (or a very high degree extension) of $\mathbb{F}_{p^k}$. Using the techniques here, we show that the circuit for g^{p^i} is in fact over the field $\mathbb{F}_{p^k}$ itself, provided that this field is polynomially large. We refer to Theorem 22 for a formal statement of this result.

1.2 Proof overview

Proof of Theorem 2

We start with a discussion of the proof of Theorem 2. To this end, we begin with an outline of the proof in [3], who originally showed the version of Theorem 2 for general algebraic circuits. Here the authors consider the polynomial

$$F(\mathbf{z}, y) = y^n - z_1 y^{n-1} + \cdots + (-1)^n z_n - 1,$$

and think of the polynomial F as a univariate in y with coefficients from the ring $\mathbb{F}[\mathbf{z}]$. Since $F(\mathbf{0}, y) = y^n - 1$ has n distinct roots, namely the n^{th} roots of unity, we get by an application of Newton Iteration that each of these roots of $y^n - 1$ can be *lifted* to a unique power series (in $\mathbf{z}$) root of $F(\mathbf{z}, y)$. Let these power series roots be $A_1(\mathbf{z}), A_2(\mathbf{z}), \ldots, A_n(\mathbf{z})$. If $P(\mathbf{x})$ is a symmetric polynomial, then there is a unique n variate polynomial Q of degree at most $\deg(P)$ such that $P(\mathbf{x}) = Q(\mathrm{Esym}_1(\mathbf{x}), \ldots, \mathrm{Esym}_n(\mathbf{x}))$, where $\mathrm{Esym}_i(\mathbf{x})$ is the elementary symmetric polynomial of degree i in $\mathbf{x}$. Let us consider the substitution where the variable x_i is replaced by the power series $A_i(\mathbf{z})$. Note that $\mathrm{Esym}_j(A_1, \ldots, A_n)$ must equal z_j for every $j \in \{1, 2, \ldots, n-1\}$ and equals $z_n + (-1)^{n-1}$ for $j = n$. So, we get

$$P(A_1(\mathbf{z}), \ldots, A_n(\mathbf{z})) = Q(z_1, \ldots, z_{n-1}, z_n + (-1)^{n-1}).$$

We now note that since Q is a polynomial, it suffices to compute the power series $A_i(\mathbf{z})$ modulo $\langle \mathbf{z} \rangle^{\deg(Q)+1}$. Using the fact that if F has a small algebraic circuit, all its truncated power series roots have small circuits, and with a few small modifications, we get a small unbounded depth circuit for Q.

Given this outline, to prove Theorem 2, it would suffice to show that the truncated power series roots of $F(\mathbf{z}, y)$ have small constant-depth circuits, and we do exactly this. Note that F itself has a small constant-depth circuit. We now invoke the closure results for factors of constant-depth circuits shown in [2] (Theorem 4.1) to show that the truncations of the power series A_i can be computed by small constant-depth circuits. This fact together with the above outline gives us Theorem 2.

We note that even though the main closure results in [2] are for fields of zero or large characteristic, the instance we have at hand here is very special – we are trying to compute the power series roots of a polynomial that has no repeated roots. For such instances, essentially the arguments from fields of characteristic zero carry over as is to fields of small positive characteristic. This completes an outline of the proof of Theorem 2.

Computing the resultant in constant depth

To prove Theorem 1, we first observe that Theorem 2 can be generalized for polynomials that are bi-symmetric. More precisely, we observe that if a $2n$ variate polynomial $P(\mathbf{x}, \mathbf{y})$ is symmetric in $\mathbf{x}$ and in $\mathbf{y}$ (although it need not be symmetric in $\mathbf{x} \cup \mathbf{y}$ together), then P equals $Q(\mathrm{Esym}_1(\mathbf{x}), \ldots, \mathrm{Esym}_n(\mathbf{x}), \mathrm{Esym}_1(\mathbf{y}), \ldots, \mathrm{Esym}_n(\mathbf{y}))$ for a unique polynomial Q. We then show that if P has a small constant-depth circuit, then so does Q.

An almost immediate (and independently) interesting consequence of the above discussion is that the resultant of two monic univariates can be computed by a constant-depth circuit that takes the coefficients of the polynomials as inputs. If $f(x) = f_0 + f_1 x + \cdots + x^n$ and $g(x) = g_0 + g_1 x + \cdots + x^n$ are two univariates with roots denoted by $\{\alpha_1, \ldots, \alpha_n\}$ and $\{\beta_1, \ldots, \beta_n\}$ respectively, then their resultant R equals $\prod_{i,j \in [n]} (\alpha_i - \beta_j)$. We note that R is a polynomial that is symmetric in the roots of f and the roots of g, and has a small circuit of constant depth that takes these roots as inputs. Thus, from the generalization of Theorem 2 to bisymmetric polynomials, we get that there is a polynomial Q that can be computed by a small constant-depth circuit, such that Q composed with the elementary symmetric polynomials of $\{\alpha_1, \ldots, \alpha_n\}$ (i.e. the coefficients of f) and the elementary symmetric polynomials of $\{\beta_1, \ldots, \beta_n\}$ (i.e. the coefficients of g) equals R. Thus, we have a small constant-depth circuit for the resultant of two univariates that takes the coefficients as input.

Computing the GCD

It would be helpful to revisit the proof strategy of Andrews and Wigderson [1]. If the polynomials f and g are square-free, then the $\gcd(f, g)$ can be obtained by "filtering" from f only those roots of f that are also roots of g. One of the key aspects of [1] was to execute this "filtering" operation using the efficient transformation (via constant-depth circuits) between elementary and power-sum symmetric polynomials of the roots of f and g. In the more general setting when f and g are not square-free, using some additional ideas, [1] reduce to the square-free setting by first computing the *square-free decomposition*[2] and working with each part in the decomposition.

[2] i.e., computing $f_1, \ldots, f_d$ that are square-free and co-prime such that $f = \prod f_i^i$

In the setting of characteristic $p > 0$, we have to address two issues – (a) we do not have efficient transformations between power-sum and elementary symmetric polynomials, and (b) computing the square-free part of f is a more general operation than computing p^{th}-roots i.e. computing an efficient circuit for a polynomial f using an efficient circuit for f^p (something that we still do not know how to do, and would be very interesting). We give an alternate implementation of the filtering operation using Theorem 2 and also have a simple observation that allows us to completely bypass the calculation of the square-free part. We briefly elaborate on the key steps below.

- **An auxiliary polynomial containing $\gcd(f, g)$:** Our first observation is that, if we consider a *generic* linear combination of f and g given by $F(y, z) := f(y) + z \cdot g(y)$, then each root α of $\gcd(f, g)$ (from $\overline{\mathbb{F}}$) appears as a linear factor $(y - \alpha)$ of F with the *right* multiplicity. Furthermore, there is no other $\alpha \in \overline{\mathbb{F}}$ satisfying $F(\alpha, z) = 0$. Thus, if we could somehow *filter* the y-roots of $F(y, z)$ with their multiplicity, we will obtain $\gcd(f, g)$. This thereby avoids computing square-free components of f and g entirely.
- **Filtering roots of one polynomial using another, à la [1]:** Suppose $\{\sigma_1, \ldots, \sigma_m\}$ is the multi-set of roots of f over $\overline{\mathbb{F}}$ and $g(y) \in \mathbb{F}[y]$ (where $\mathbb{F}$ could be the field of rational functions over other variables), [1] define a "filter" operation to compute

$$\text{Filter}(f \mid \text{non-roots}(g)) = \prod_{i \in [m] \,:\, g(\sigma_i) \neq 0} (y - \sigma_i),$$

$$\text{Filter}(f \mid \text{roots}(g)) = \prod_{i \in [m] \,:\, g(\sigma_i) = 0} (y - \sigma_i) = \frac{f(y)}{\text{Filter}(f \mid \text{non-roots}(g))}.$$

  [1] show that the above can essentially be computed piece-wise via elementary symmetric polynomials of $\{g(\sigma_1), \ldots, g(\sigma_m)\}$ with the advice parameter being the degree of the filter. Rather than proceeding via power-sum symmetric polynomials (which requires large characteristic fields), we once again use Theorem 2 to express this as a constant-depth circuit over the coefficients of f and g. Theorem 1 now follows by interpreting $F(y, z)$ as a univariate in y over the larger field $\mathbb{F}(z)$, and noticing that

$$\gcd(f(y), g(y)) = \text{Filter}(F(y, z) \mid \text{roots}(g)) = \frac{F(y, z)}{\text{Filter}(F(y, z) \mid \text{non-roots}(g))}.$$

Organization of the paper

The rest of the paper is organized as follows. We start with some general notations and preliminaries in Section 2, followed by the proof of Theorem 2 in Section 3. In Section 4, we show that the resultant and GCD can be computed by constant-depth circuits and in Section 5, we use these techniques to strengthen the closure results for constant-depth circuits under factorization.

2 Preliminaries

Notation

- We use boldface letters such as $\mathbf{x}$ to refer to an ordered tuple of variables
- We use $\mathbb{F}, \mathbb{K}$ etc., to refer to fields, and use $\overline{\mathbb{F}}$ to refer to the algebraic closure of $\mathbb{F}$. Also, $\mathbb{F}[x]$ refers to the polynomial ring, $\mathbb{F}(x)$ to the field of fractions, $\mathbb{F}[\![x]\!]$ 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 will typically use $\mathbb{K}$ to denote the field of fractions $\mathbb{F}(\mathbf{x})$ for an underlying field $\mathbb{F}$ and some tuple of variables $\mathbf{x}$.

- Given a univariate polynomial $f(x) \in \mathbb{F}[x]$ (or more generally in the ring of formal Laurent series $\mathbb{F}(\!(x)\!)$), we denote by $[x^n]\{f\}$ the coefficient of the monomial x^n in f. In the case of a multivariate polynomial, such as $f(x, y)$, we interpret f as an element of $\mathbb{F}[y][x]$ – that is, as a univariate polynomial in x with coefficients in $\mathbb{F}[y]$ – and $[x^n]\{f\}$ then refers to the x^n coefficient viewed as a polynomial in 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. For any subset $\mathbf{x}_S \subseteq \mathbf{x}$ and any $i \in [d]$, the degree i homogeneous part of F with respect to $\mathbf{x}_S$, denoted by $\mathrm{Hom}_{\mathbf{x}_S,i}(F)$.
- For polynomials (or more generally, power series) $f(\mathbf{x})$ and $g(\mathbf{x})$, and $t \in \mathbb{N}$, we write $f(\mathbf{x}) \equiv g(\mathbf{x}) \bmod \langle \mathbf{x} \rangle^t$ to denote that $\mathrm{Hom}_{\leq t-1}(f) \equiv \mathrm{Hom}_{d \leq t-1}(g)$.
- The i^{th} elementary symmetric polynomial is sum of all multilinear monomials in $\mathbf{x}$ of degree i. Formally, the i^{th} elementary polynomial is $\mathrm{Esym}_i(\mathbf{x}) \coloneqq \sum_{S \subseteq [n]:|S|=i} \prod_{i \in S} x_i$.

2.1 Interpolation

The following uses of polynomial interpolation in the context of algebraic circuits are due to Michael Ben-Or.

▶ **Lemma 3** (Interpolation). *Let $\mathbb{F}$ be a subfield of a commutative ring R, and suppose that $\mathbb{F}$ contains at least $d+1$ distinct elements. Choose distinct points $\alpha_0, \ldots, \alpha_d \in \mathbb{F}$. Then, for any polynomial $f(t) = f_0 + f_1 t + \cdots + f_d t^d \in R[t]$ of degree at most d, and for each index $i \in \{0, \ldots, d\}$, there exist elements $\beta_{i0}, \ldots, \beta_{id} \in \mathbb{F}$ (depending only on i and the α_j's) such that the coefficient f_i can be recovered from the values of f at the interpolation points:*

$$f_i = \beta_{i0} f(\alpha_0) + \beta_{i1} f(\alpha_1) + \cdots + \beta_{id} f(\alpha_d).$$

▶ **Corollary 4** (Applications of interpolation). *Let $\alpha_0, \ldots, \alpha_d$ be distinct elements of the field $\mathbb{F}$. The following statements are immediate consequences of interpolation:*

1. ***[Extracting homogeneous parts]*** *Let $C(\mathbf{x})$ be a polynomial of total degree d, and let $\mathbf{x}_S \subseteq \mathbf{x}$ be a subset of the variables. For any $i \in \{0, \ldots, d\}$, the degree i homogeneous component of C with respect to the variables in $\mathbf{x}_S$, denoted $\mathrm{Hom}_{\mathbf{x}_S,i}(C)$, can be written 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 scalars $\beta_{i,j} \in \mathbb{F}$ that are independent of the polynomial C.
2. ***[Derivatives via evaluation]*** *Suppose $C(\mathbf{x}, y)$ is a polynomial with degree d in the variable y. Then the i^{th} partial derivative of C with respect to y can be expressed as an $\mathbb{F}[y]$-linear combination of the evaluations $C(\mathbf{x}, \alpha_j)$ for $j = 0, \ldots, d$. Specifically, there exist polynomials $\mu_0(y), \ldots, \mu_d(y) \in \mathbb{F}[y]$, each of degree at most d and independent of C, such that*

$$\partial_{y^i} C(\mathbf{x}, y) = \sum_{j=0}^{d} \mu_j(y) \cdot C(\mathbf{x}, \alpha_j).$$

Moreover, if C is computed by a circuit of size s and depth Δ, then both operations – extracting homogeneous components and computing partial derivatives – can be implemented by circuits of size $\mathrm{poly}(s, d)$ and depth at most $\Delta + O(1)$.

2.2 Resultants

▶ **Definition 5** (Sylvester Matrix and Resultant (for details, see [12, Chapter 6])). *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 $(a+b)\times(a+b)$ 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 $\mathrm{Syl}_z(P,Q)$.*

▶ **Fact 6** (see, e.g., Chapter 3 of [5]). *Suppose $f(z), g(z) \in \mathbb{F}[z]$ are monic polynomials, and the multiset of their roots over $\overline{\mathbb{F}}$ are $\alpha_1, \ldots, \alpha_a$ and $\beta_1, \ldots, \beta_b$. Then,*

$$\mathrm{Res}_z(f,g) = \prod_{i=1}^{a}\prod_{j=1}^{b}(\alpha_i - \beta_j).$$

2.3 Factorization of monic polynomials into power series

▶ **Lemma 7** (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 square-free. 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 $f(t, \varphi_\alpha(t)) = 0$.*

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 [7, Section 3] gives a formal proof of the above. We also note that the lemma can be applied in the multivariate setting by taking a multivariate polynomial in variables $x_1, x_2, \ldots, x_n, y$, replacing each x_i by $x_i \cdot t$ (for a new indeterminate t) and viewing the resulting polynomial as a bivariate in t and y with coefficients in the field $\mathbb{F}(x_1, x_2, \ldots, x_n)$.

2.4 Furstenberg's theorem for roots of multiplicity one

We define the following *diagonal* operator for bivariate power series $F(x,y) = \sum_{i,j} F_{i,j} x^i y^j \in \mathbb{F}[\![x,y]\!]$ as follows:

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

▶ **Theorem 8** ([9, Proposition 2]). *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 $P(t,y) = (y - \varphi(t)) \cdot Q(t,y)$. If $\varphi(0) = 0$ and $Q(0,0) \neq 0$, then*

$$\varphi = \mathcal{D}\left(\frac{y^2 \cdot \partial_y P(ty,y)}{P(ty,y)}\right) \tag{1}$$

Furstenberg [9] states the theorem for $P(t,y)$ being a polynomial and that will be sufficient for our purposes; the proof works even when $P(t,y)$ is a power series. The diagonal expression above can be simplified to a slightly more convenient expression for implicitly defined power series roots.

▶ **Corollary 9.** *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,*

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

Suppose $P(\mathbf{x}, y)$ is computable by a small constant-depth circuit. Using Corollary 9, [2] showed that truncations of power series roots (of multiplicity 1) of $P(\mathbf{x}, y)$ also have small constant-depth circuits.

▶ **Theorem 10** (Power series roots without multiplicity [2]). *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*

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

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

3 Complexity of Symmetric Polynomials

A polynomial $P(x_1, \ldots, x_n)$ is symmetric if for every permutation σ on n elements, $P(x_1, \ldots, x_n) = P(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$. The elementary symmetric polynomials $\{\mathrm{Esym}_i(\mathbf{x})\}_{i \in [n]}$ are ubiquitous examples of symmetric polynomials. Surprisingly, the *fundamental theorem of symmetric polynomials* states that every symmetric polynomial P can be uniquely expressed as polynomial expressions in $\{\mathrm{Esym}_i(\mathbf{x})\}_{i \in [n]}$.

▶ **Theorem 11** (Fundamental theorem of symmetric polynomials). *Let $P(\mathbf{x})$ be a symmetric polynomial on n variables of degree d. Then there is a unique polynomial $Q(y_1, \ldots, y_n)$ such that $P(\mathbf{x}) = Q(\mathrm{Esym}_1(\mathbf{x}), \mathrm{Esym}_2(\mathbf{x}), \ldots, \mathrm{Esym}_n(\mathbf{x}))$. Moreover, the degree of Q is at most d.*

Given this fact, it is natural to ask if complexity of the polynomial Q is related to the complexity of P. Bläser and Jindal [3] showed that if P is computable by a small algebraic circuit, then Q is also computable by a small algebraic circuit.

▶ **Theorem 12** (Theorem 4 in [3]). *Let $P(\mathbf{x}) \in \mathbb{C}[\mathbf{x}]$ be a symmetric polynomial on n variables of degree d computed by a circuit C of size s. Let $Q(\mathbf{x}) \in \mathbb{C}[\mathbf{x}]$ be the unique polynomial such that $P(\mathbf{x}) = Q(\mathrm{Esym}_1(\mathbf{x}), \ldots, \mathrm{Esym}_n(\mathbf{x}))$. Then, $Q(\mathbf{x})$ is computable by a circuit of size $\mathrm{poly}(s, d, n)$ over $\mathbb{C}$.*

One of the main steps in their proof is to compute the truncations of power series roots using Newton Iteration. Based on the ideas from [2], we would expect that replacing Newton Iteration with Furstenberg's theorem (Theorem 8) should give us an analogous result over constant-depth circuits. Indeed, this turns out to be the case. We now prove the constant-depth version of their result.

▶ **Theorem 13** (Complexity of symmetric polynomials). *Let $n, d, s \in \mathbb{N}$ be parameters, and let $\mathbb{F}$ be a polynomially large field, i.e. $|\mathbb{F}| \geq (nds)^c$ for an absolute constant $c > 0$. Let $P(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a symmetric polynomial on n variables of degree d computed by a circuit C of size s and depth Δ. If $Q(\mathbf{z}) \in \mathbb{F}[\mathbf{z}]$ is the unique degree d polynomial such that $P(\mathbf{x}) = Q(\mathrm{Esym}_1(\mathbf{x}), \ldots, \mathrm{Esym}_n(\mathbf{x}))$, then $Q(\mathbf{z})$ is computable by a circuit of size $\mathrm{poly}(s, d, n)$ and depth $\Delta + O(1)$ over $\mathbb{F}$.*

Proof. Let $\alpha_1, \ldots, \alpha_n$ be n distinct elements from $\mathbb{F}$ and β_i denote $\mathrm{Esym}_i(\alpha_1, \ldots, \alpha_n)$ for each $i \in [n]$. Define $F(\mathbf{z}, y)$ as:

$$F(\mathbf{z}, y) := y^n - (z_1 + \beta_1)y^{n-1} + \cdots + (-1)^{n-1}(z_{n-1} + \beta_{n-1})y + (-1)^n(z_n + \beta_n).$$

$F[\mathbf{z}, y]$ has a circuit of size $O(n^2)$ and depth 2. Note that $F(\mathbf{0}, y) = y^n + \sum_{i=1}^{n} (-1)^i y^{n-i}\beta_i = \prod_{i\in[n]}(y - \alpha_i)$. Since $\alpha_1, \ldots, \alpha_n$ are distinct, we have that $F(\mathbf{0}, y)$ is square-free and hence (by Lemma 7) $F(\mathbf{z}, y)$ factorizes as

$$F(\mathbf{z}, y) = \prod_{i=1}^{n}(y - A_i(\mathbf{z}))$$

where $A_i(\mathbf{z}) \in \mathbb{F}[\![\mathbf{z}]\!]$ with $A_i(\mathbf{0}) = \alpha_i$ for each $i \in [n]$. Furthermore, $\mathrm{Esym}_k(A_1(\mathbf{z}), \ldots, A_n(\mathbf{z})) = z_k + \beta_k$.

We now apply Theorem 10 on $F(\mathbf{z}, y)$ to get size $\mathrm{poly}(n)$ and depth $O(1)$ circuits $\tilde{A}_1, \ldots, \tilde{A}_n$ such that $\tilde{A}_i(\mathbf{z}) = A_i(\mathbf{z}) \bmod \langle \mathbf{z} \rangle^{d+1}$ for each $i \in [n]$. This implies that for each $i \in [n]$,

$$\mathrm{Esym}_i(\tilde{A}_1, \ldots, \tilde{A}_n) = z_i + \beta_i \bmod \langle \mathbf{z} \rangle^{d+1}.$$
$$\therefore \quad P(\tilde{A}_1, \ldots, \tilde{A}_n) = P(A_1, \ldots, A_n) \bmod \langle \mathbf{z} \rangle^{d+1}$$

Since $P(\mathbf{x}) = Q(\mathrm{Esym}_1(\mathbf{x}), \ldots, \mathrm{Esym}_n(\mathbf{x}))$,

$$\begin{aligned} P(A_1, \ldots, A_n) &= Q(\mathrm{Esym}_1(A_1, \ldots, A_n), \ldots, \mathrm{Esym}_n(A_1, \ldots, A_n)) \\ &= Q(z_1 + \beta_1, \ldots, z_n + \beta_n) \end{aligned}$$

where degree of Q is at most d from Theorem 11. Therefore,

$$Q(z_1 + \beta_1, \ldots, z_n + \beta_n) = P(A_1, \ldots, A_n) = \mathrm{Hom}_{\le d}\, P(\tilde{A}_1, \ldots, \tilde{A}_n)$$

Recall that each $\tilde{A}_i$ is computable by a size $\mathrm{poly}(n)$ and depth $O(1)$ circuit, and P is computable by a size $\mathrm{poly}(s, d, n)$ and depth Δ circuit. Thus, we obtain a $\mathrm{poly}(s, d, n)$ and depth $\Delta + O(1)$ circuit for $Q(\mathbf{z} + \boldsymbol{\beta})$, and thus for $Q(\mathbf{z})$ as well (by shifting $\mathbf{z} \leftarrow \mathbf{z} - \boldsymbol{\beta}$). ◀

3.1 Multi-symmetric polynomials

▶ **Definition 14** (Multi-symmetric polynomials). *Let $\mathbf{x} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_k$ be a partition of the set of variables $\mathbf{x}$. A polynomial $f(\mathbf{x}) = f(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ is said to be* multi-symmetric *with respect to the above partition if for all $i \in [k]$ we have that $f(\mathbf{x})$ is symmetric with respect to the part $\mathbf{x}_i$. In other words, for every choice of permutations $\pi_1, \ldots, \pi_k$ on $\mathbf{x}_1, \ldots, \mathbf{x}_k$, we have $f(\pi_1(\mathbf{x}_1), \ldots, \pi_k(\mathbf{x}_k)) = f(\mathbf{x}_1, \ldots, \mathbf{x}_k)$.*

The following is a natural extension of Theorem 11. For purely the ease of exposition, we assume the size of each part in the partition is the same; the proof would readily extend to the general case with mere notational changes.

▶ **Theorem 15** (Fundamental theorem for multi-symmetric polynomials). *Let $\mathbf{x} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_k$ be a partition of the variables with each $|\mathbf{x}_i| = n$, and let $P(\mathbf{x})$ be a multi-symmetric polynomial with respect to the above partition of degree d. Then there is a unique polynomial $Q(\mathbf{y}_1, \ldots, \mathbf{y}_k)$ (where $\mathbf{y}_i = (y_{i1}, \ldots, y_{in})$) such that $P(\mathbf{x}) = Q(\overline{\mathrm{Esym}}(\mathbf{x}_1), \ldots, \overline{\mathrm{Esym}}(\mathbf{x}_k))$ where*

$$\overline{\mathrm{Esym}}(\mathbf{x}_i) \coloneqq (\mathrm{Esym}_1(\mathbf{x}_i), \ldots, \mathrm{Esym}_n(\mathbf{x}_i)).$$

Moreover, the degree of Q is at most d.

Proof. The proof is just a simple induction on k. For $k = 1$, the claim follows from Theorem 11.

Assuming that the theorem is true for $k-1$, and write $P(\mathbf{x})$ as

$$P(\mathbf{x}) = \sum_{\mathbf{e}\geq 0} [\mathbf{x}_k^{\mathbf{e}}]\{P\} \cdot \mathbf{x}_k^{\mathbf{e}}.$$

We will say $\mathbf{e}_i \sim \mathbf{e}_j$ if there is some permutation π on $|\mathbf{x}_k|$ elements such that $\pi(\mathbf{e}_i) = \mathbf{e}_j$; let $\mathcal{E}$ be a distinct set of representatives of the equivalence classes defined by $\sim$. Note that if $\mathbf{e}_i \sim \mathbf{e}_j$, then $[\mathbf{x}_k^{\mathbf{e}_i}]\{P\} = [\mathbf{x}_k^{\mathbf{e}_j}]\{P\}$. Therefore, if we define $H_{\mathbf{e}^*}(\mathbf{x}_k)$ for an $\mathbf{e}^* \in \mathcal{E}$ as

$$H_{\mathbf{e}^*}(\mathbf{x}_k) \coloneqq \sum_{\mathbf{e}' \,:\, \mathbf{e}' \sim \mathbf{e}^*} \mathbf{x}_k^{\mathbf{e}'}$$

then,

$$\begin{aligned} P(\mathbf{x}) &= \sum_{\mathbf{e}^* \in \mathcal{E}} [\mathbf{x}_k^{\mathbf{e}^*}]\{P\} \cdot H_{\mathbf{e}^*}(\mathbf{x}_k) \\ &= \sum_{\mathbf{e}^* \in \mathcal{E}} [\mathbf{x}_k^{\mathbf{e}^*}]\{P\} \cdot \tilde{Q}_{\mathbf{e}^*}(\overline{\mathrm{Esym}}(\mathbf{x}_k)) \qquad \text{(by Theorem 11)} \end{aligned}$$

for some $\tilde{Q}_{\mathbf{e}^*}$ of degree at most $|\mathbf{e}^*|$. Since P is multi-symmetric, we have that $[\mathbf{x}_k^{\mathbf{e}}]\{P\}$ is multi-symmetric with respect to $\mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_{k-1}$. By induction, there exists polynomials $Q_{\mathbf{e}}$ of degree at most $d - |\mathbf{e}|$ such that

$$\begin{aligned} &[\mathbf{x}_k^{\mathbf{e}}]\{P\} = Q_{\mathbf{e}}(\overline{\mathrm{Esym}}(\mathbf{x}_1), \ldots, \overline{\mathrm{Esym}}(\mathbf{x}_{k-1})). \\ \implies &P(\mathbf{x}) = \sum_{\mathbf{e}^* \in \mathcal{E}} Q_{\mathbf{e}^*}(\overline{\mathrm{Esym}}(\mathbf{x}_1), \ldots, \overline{\mathrm{Esym}}(\mathbf{x}_{k-1})) \cdot \tilde{Q}_{\mathbf{e}^*}(\overline{\mathrm{Esym}}(\mathbf{x}_k)) \\ &\qquad =: Q(\overline{\mathrm{Esym}}(\mathbf{x}_1), \ldots, \overline{\mathrm{Esym}}(\mathbf{x}_k)) \end{aligned}$$

for some $Q(y_{11}, \ldots, y_{kn})$ of degree at most d. ◀

▶ **Theorem 16** (Complexity of multi-symmetric polynomials). *Let $\mathbb{F}$ be a polynomially large field with respect to parameters $n, d, s \in \mathbb{N}$, i.e. $|\mathbb{F}| \geq (nds)^c$ for an absolute constant $c > 0$. Let $\mathbf{x} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_k$ be a partition of the variables, with each $|\mathbf{x}_i| = n$, and let $P(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a multi-symmetric polynomial with respect to the partition of degree d computed by a circuit C of size s and depth Δ. If $Q(z_{11}, \ldots, z_{kn}) \in \mathbb{F}[\mathbf{z}]$ is the unique degree d polynomial such that*

$$P(\mathbf{x}) = Q(\overline{\mathrm{Esym}}(\mathbf{x}_1), \ldots, \overline{\mathrm{Esym}}(\mathbf{x}_k))$$

where $\overline{\mathrm{Esym}}(\mathbf{x}_i) \coloneqq (\mathrm{Esym}_1(\mathbf{x}_i), \ldots, \mathrm{Esym}_n(\mathbf{x}_i))$, then $Q(\mathbf{z})$ is computable by a circuit of size $\mathrm{poly}(s, d, n)$ and depth $\Delta + O(1)$ over $\mathbb{F}$.

Proof. The proof is almost a direct extension of Theorem 13. As in the proof of Theorem 13, for each $i \in [k]$ define the polynomials for $\mathbf{z}_i = (z_{i,1}, \ldots, z_{i,n})$:

$$F_i(\mathbf{z}_i, y) \coloneqq y^n - (z_{i,1} + \beta_1)y^{n-1} + \cdots + (-1)^{n-1}(z_{i,n-1} + \beta_{n-1})y + (-1)^n(z_{i,n} + \beta_n)$$

where $\beta_i = \mathrm{Esym}_i(\alpha_1, \ldots, \alpha_n)$ for some distinct choice of $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$. By Lemma 7, since $F_i(\mathbf{0}, y) = \prod_{i=1}^n (y - \alpha_i)$ is square-free, we have that $F_i(\mathbf{z}_i, y)$ factors

$$F_i(\mathbf{z}_i, y) = \prod_{j=1}^{n} (y - A_{ij}(\mathbf{z}_i))$$

where $A_{ij}(\mathbf{z}_i) \in \mathbb{F}[\![\mathbf{z}_i]\!]$ and $A_{ij}(\mathbf{0}) = \alpha_j$. Note that $\mathrm{Esym}_r(A_{i1}, \ldots, A_{in}) = z_{i,r} + \beta_r$. By Theorem 10, we have circuits $\tilde{A}_{ij}$ of size $\mathrm{poly}(n)$ and depth $O(1)$ such that

$$\tilde{A}_{ij}(\mathbf{z}_i) = A_{ij}(\mathbf{z}_i) \bmod \langle \mathbf{z}_i \rangle^{d+1}.$$

From Theorem 16, degree of Q is at most d. Thus,

$$\begin{aligned} P(\tilde{A}_{11}(\mathbf{z}_1), \ldots, \tilde{A}_{kn}(\mathbf{z}_k)) &= P(A_{11}(\mathbf{z}_1), \ldots, A_{kn}(\mathbf{z}_k)) \bmod \langle \mathbf{z} \rangle^{d+1} \\ &= Q(\overline{\mathrm{Esym}}(A_{11}, \ldots, A_{1n}), \ldots, \overline{\mathrm{Esym}}(A_{k1}, \ldots, A_{kn})) \\ &= Q(\mathbf{z}_1 + \boldsymbol{\beta}, \ldots, \mathbf{z}_k + \boldsymbol{\beta}) \\ \implies Q(\mathbf{z}_1 + \boldsymbol{\beta}, \ldots, \mathbf{z}_k + \boldsymbol{\beta}) &= \mathrm{Hom}_{\leq d}\, P(\tilde{A}_{11}(\mathbf{z}_1), \ldots, \tilde{A}_{kn}(\mathbf{z}_k)) \end{aligned}$$

By Corollary 4, the RHS is computable by a $\mathrm{poly}(n, d, s)$-size depth $\Delta + O(1)$ circuit, and by un-translating by $\boldsymbol{\beta}$ we also obtain a such a circuit for $Q(\mathbf{z}_1, \ldots, \mathbf{z}_k)$. ◀

Note that the additive $O(1)$ increase in depth does not depend on k (the number of sets in the partition of variables) since we are simultaneously inverting the elementary symmetric polynomials for each set of variables by plugging in A_{ij} for each x_{ij}.

4 Computing resultants and GCD

4.1 A constant-depth circuit for the resultant over all fields

An almost immediate corollary of Theorem 16 is a polynomial size constant-depth circuit for the resultant of two polynomials over any large enough field.

▶ **Corollary 17** (Computing resultants in constant depth). *Let $\mathbb{F}$ be any large enough field (i.e., $|\mathbb{F}| \geq d^c$ for some absolute constant $c > 0$). Given two degree d univariate polynomials $f(y)$ and $g(y)$ that are monic in y and provided as a list of coefficients, we can compute the resultant of the two polynomials, $\mathrm{Res}_y(f, g)$, via a $\mathrm{poly}(d)$-size depth $O(1)$ circuit over the coefficients of f and g over the field $\mathbb{F}$.*

Proof. Let $f(y) = y^{d_1} - f_1 y^{d_1 - 1} + \cdots + (-1)^{d_1} f_0$ and $g(y) = y^{d_2} - g_1 y^{d_2 - 1} + \cdots + (-1)^{d_2} g_0$ and let $d = \max(d_1, d_2)$. If the multiset of roots (in $\overline{\mathbb{F}}$) of f and g are given by $\{\sigma_1, \ldots, \sigma_{d_1}\}$ and $\{\tau_1, \ldots, \tau_{d_2}\}$, then by Fact 6 we have

$$\mathrm{Res}_y(f, g) = \prod_{i=1}^{d_1} \prod_{j=1}^{d_2} (\sigma_i - \tau_j).$$

The RHS is clearly a bi-symmetric polynomial with respect to $\{\sigma_1, \ldots, \sigma_{d_1}\} \sqcup \{\tau_1, \ldots, \tau_{d_2}\}$ and is a $\mathrm{poly}(d)$ sized depth 2 circuit over the σ's and τ's. By Theorem 15, there exists a unique Q of degree at most $d_1 d_2$ such that

$$\mathrm{Res}_y(f, g) = Q(\overline{\mathrm{Esym}}(\sigma_1, \ldots, \sigma_{d_1}), \overline{\mathrm{Esym}}(\tau_1, \ldots, \tau_{d_2})).$$

Note that $\mathrm{Esym}_r(\sigma_1, \ldots, \sigma_{d_1})$ is just f_r and $\mathrm{Esym}_r(\tau_1, \ldots, \tau_{d_2})$ is just g_r. Thus, by Theorem 16, Q is also computable by a circuit of size $\mathrm{poly}(d)$ and depth $O(1)$ over the coefficients $f_1, \ldots, f_{d_1}, g_1, \ldots, g_{d_2}$ over the field $\mathbb{F}$. ◀

If $f(\mathbf{x}, y), g(\mathbf{x}, y)$ are multivariate polynomials that are monic in y and computable by size s depth Δ circuits, then the coefficients of monomials in y is also computable by size $\mathrm{poly}(s, n, d)$, depth $\Delta + O(1)$ circuits and thus $\mathrm{Res}_y(f, g)$ is also computable by a size $\mathrm{poly}(s, n, d)$ depth $\Delta + O(1)$ circuit.

The resultant is the case of computing the product of g's evaluations on the roots of f. A more general statement is the following lemma.

▶ **Lemma 18.** *Let $\mathbb{F}$ be any large enough field (i.e., $|\mathbb{F}| \geq (sd_1d_2n)^c$ for some absolute constant $c > 0$). Suppose $f(\mathbf{x}, y), g(\mathbf{x}, y) \in \mathbb{F}[\mathbf{x}, y]$ are two n-variate polynomial that are monic in y, of degree d_1 and d_2 respectively (in y), that are given by size s, depth Δ circuits. Let $\{\sigma_1, \ldots, \sigma_{d_1}\}$ be the multi-set of y-roots of f over $\overline{\mathbb{F}(\mathbf{x})}$. Then, for any $1 \leq r \leq d_1$, we can obtain a circuit of size $\mathrm{poly}(s, n, d_1, d_2)$ and depth $\Delta + O(1)$ for*

$$\mathrm{Esym}_r(\{g(\mathbf{x}, \sigma_1), \ldots, g(\mathbf{x}, \sigma_{d_1})\}) \in \mathbb{F}[\mathbf{x}].$$

Proof. Let $d = \max(d_1, d_2)$. Note that $\mathrm{Esym}_r(y_1, \ldots, y_m) = [t^r]\{(1 + ty_1) \cdots (1 + ty_m)\}$. By Corollary 4, it is computable by a size $\mathrm{poly}(r, m)$ depth 3 circuit over $\mathbb{F}$.

If $\{\tau_1, \ldots, \tau_{d_2}\}$ is the multi-set of y-roots of g, then

$$\mathrm{Esym}_r(\{g(\mathbf{x}, \sigma_1), \ldots, g(\mathbf{x}, \sigma_{d_1})\}) = \mathrm{Esym}_r\left(\left\{\prod_{j=1}^{d_2}(\sigma_i - \tau_j) \;:\; i \in [d_1]\right\}\right)$$

is clearly a bi-symmetric polynomial with respect to $\{\sigma_1, \ldots, \sigma_{d_1}\} \sqcup \{\tau_1, \ldots, \tau_{d_2}\}$. Thus, if $f(\mathbf{x}, y) = y^{d_1} + f_1 y^{d_1 - 1} + \cdots + f_{d_1}$ and $g(\mathbf{x}, y) = y^{d_2} + g_1 y^{d_2 - 1} + \cdots + g_0$, then each f_i, g_i are the elementary symmetric polynomials of $\{\sigma_1, \ldots, \sigma_{d_1}\}$ and $\{\tau_1, \ldots, \tau_{d_2}\}$, and Corollary 4 shows that we can obtain circuits of size $\mathrm{poly}(s, d)$ and depth $\Delta + O(1)$ for them. By Theorem 16, we can express the above as $Q(f_1, \ldots, f_{d_1}, g_1, \ldots, g_{d_2})$ where $f(\mathbf{x}, y) = y^{d_1} + f_1 y^{d_1 - 1} + \cdots + f_{d_1}$ and $g(\mathbf{x}, y) = y^{d_2} + g_1 y^{d_2 - 1} + \cdots + g_0$ with Q computable by a size $\mathrm{poly}(s, n, d)$, depth $\Delta + O(1)$ circuit. ◀

The above lemma can also be easily generalized to compute elementary symmetric polynomials of a *rational function* $g(x)/h(x)$ evaluated at the roots of $f(x)$, as shown in Lemma 29.

4.2 Computing the GCD of two polynomials

Unlike operations such as computing the resultant of two polynomials, the GCD of two polynomials $f(y), g(y)$ is not a continuous function of the coefficients and therefore cannot be expressed as a polynomial (or even a rational) function of the coefficients of f and g. Andrews and Wigderson [1] show that it can be expressed as a *piece-wise rational function* of the coefficients of f, g where we have a sequence of rational functions computed by small algebraic circuits over the coefficients of f and g such that one of them computes $\gcd(f, g)$, and we know which of them via appropriate *zero-tests*.

▶ **Definition 19** (Piece-wise rational circuit families)**.** *A* piece-wise rational circuit *for a function $P : \mathbb{F}^n \to \mathbb{F}^m$ is a family of "computation circuits" $\{A_1, \ldots, A_s : \mathbb{F}^n \to \mathbb{F}^m\}$, $\{B_1, \ldots, B_s : \mathbb{F}^n \to \mathbb{F}^m\}$ and "test circuits" $\{T_1, \ldots, T_s : \mathbb{F}^n \to \mathbb{F}\}$ such that for every $\mathbf{a} \in \mathbb{F}^n$ we have*

$$P(\mathbf{a}) = \frac{A_r(\mathbf{a})}{B_r(\mathbf{a})}$$

where the advice parameter $r = \max\{k \in \mathbb{N} \;:\; T_k(\mathbf{a}) \neq 0\}$. *The size of the family is said to be the sum of the sizes of A_i's, B_i's and T_i's, and the depth of the family is maximum depth of any A_i, B_i, T_i that constitute the family.*

Situations where $P(\mathbf{a}) \in \mathbb{F}[y]$ (such as computing the GCD of two polynomials given by a list of coefficients) can be equivalently viewed as a multi-output function that outputs the list of coefficients of the polynomial $P(\mathbf{a})$. For a more detailed discussion on this perspective and the definition of *piece-wise rational* AC^0 *circuits*, we refer the reader to [1, Section 2 in the full version].

The following is the analogue of the main result of [1] for computing the GCD of two polynomials over all sufficiently large fields expressed in terms of piece-wise rational circuit families.

▶ **Theorem 20** (Computing GCD of polynomials in terms of coefficients). *Let $\mathbb{K}$ be any polynomially large field (i.e., $|\mathbb{K}| \geq (d_1 d_2)^c$ for some absolute constant $c > 0$). There are explicit families of algebraic circuits $\{P_{d_1,d_2,k}\}$, $\{Q_{d_1,d_2,k}\}$, $\{T_{d_1,d_2,k}\}$ of size $\mathrm{poly}(d_1, d_2)$ and depth $O(1)$ such that, for every $f(y), g(y) \in \mathbb{K}[y]$ that are monic in y with $\deg(f) = d_1, \deg(g) = d_2$, we have*

$$\gcd(f(y), g(y)) = \frac{P_{d_1,d_2,r}(\overline{\mathrm{coeff}}(f), \overline{\mathrm{coeff}}(g))}{Q_{d_1,d_2,r}(\overline{\mathrm{coeff}}(f), \overline{\mathrm{coeff}}(g))}$$

where $r = \max\left\{k \leq \mathrm{poly}(d_1, d_2) \ : \ T_{d_1,d_2,k}(\overline{\mathrm{coeff}}(f), \overline{\mathrm{coeff}}(g)) \neq 0\right\}$. (Here will use $\overline{\mathrm{coeff}}(f)$ to refer to the vector of coefficients of f.)

In fact, the "advice" r above is essentially just $\deg(\gcd(f, g))$. The key idea is to consider the polynomial $F(y, z) = f(y) + z \cdot g(y) = \gcd(f, g) \cdot F'(y, z)$ and somehow "filter out" the $F'(y, z)$. We now describe how to accomplish this.

Let $f(y) \in \mathbb{K}[y]$ be a monic polynomial that factorizes over $\overline{\mathbb{K}}$ as $(y - \sigma_1)^{e_1} \cdots (y - \sigma_m)^{e_m}$, with $e_i \geq 1$. For any polynomial $g(y) \in \mathbb{K}[y]$, we define $\mathrm{Filter}(f \mid \mathrm{roots}(g))$ and $\mathrm{Filter}(f \mid \text{non-roots}(g))$ as

$$\mathrm{Filter}(f \mid \text{non-roots}(g)) = \prod_{\substack{i \in [m] \\ g(\sigma_i) \neq 0}} (y - \sigma_i)^{e_i},$$

$$\mathrm{Filter}(f \mid \mathrm{roots}(g)) = \prod_{\substack{i \in [m] \\ g(\sigma_i) = 0}} (y - \sigma_i)^{e_i} = \frac{f}{\mathrm{Filter}(f \mid \text{non-roots}(g))}.$$

▶ **Lemma 21** (Computing Filter from coefficients). *Let $\mathbb{K}$ be any large enough field (i.e., $|\mathbb{K}| \geq (d_1 d_2)^c$ for some absolute constant $c > 0$). There are explicit families of circuits $\{A_{d_1,d_2,k}\}$, $\{B_{d_1,d_2,k}\}$, $\{T_{d_1,d_2,k}\}$ of size $\mathrm{poly}(d_1, d_2)$ and depth $O(1)$ such that, for every pair of monic polynomial $f, g \in \mathbb{K}[y]$ with $\deg f = d_1$ and $\deg g = d_2$, we have*

$$\mathrm{Filter}(f \mid \text{non-roots}(g)) = \frac{A_{d_1,d_2,r}(\overline{\mathrm{coeff}}(f), \overline{\mathrm{coeff}}(g))}{B_{d_1,d_2,r}(\overline{\mathrm{coeff}}(f), \overline{\mathrm{coeff}}(g))}$$

where the advice parameter $r = \max\left\{k \leq \mathrm{poly}(d_1, d_2) \ : \ T_{d_1,d_2,k}(\overline{\mathrm{coeff}}(f), \overline{\mathrm{coeff}}(g)) \neq 0\right\}$.

Proof. Let $f(y) = \prod_{i=1}^{m} (y - \sigma_i)^{e_i}$ with $\sigma_i \in \overline{\mathbb{K}}$ and $e_i \geq 1$, and let Σ be the multiset of roots of $f(y)$ in $\overline{\mathbb{K}}$. If $r = \deg(\mathrm{Filter}(f \mid \text{non-roots}(g)))$, then

$$\mathrm{Esym}_r\left(\{g(\sigma) \ : \ \sigma \in \Sigma\}\right) = \prod_{i \in [m] \,:\, g(\sigma_i) \neq 0} g(\sigma_i)^{e_i} \neq 0,$$

$$\mathrm{Esym}_k\left(\{g(\sigma) \ : \ \sigma \in \Sigma\}\right) = 0 \quad \text{for all } k > r$$

as any subset of size greater than r among Σ must necessarily contain a σ such that $g(\sigma) = 0$. By Lemma 18, the above is expressible as circuit of size $\mathrm{poly}(d_1, d_2)$ and depth $O(1)$ over the coefficients of f and g and that gives us the claimed family $\{T_{d_1,d_2,k}\}$.

To obtain the filter as a ratio of algebraic circuits, define the polynomial $h(x,y) = (y-x)\cdot g(x)$ for a fresh variable x. Note that, for any $\sigma \in \overline{\mathbb{K}}$, we have $h(\sigma, y) = 0$ if and only if $g(\sigma) = 0$. Therefore,

$$\begin{aligned}\mathrm{Esym}_r\left(\{h(\sigma,y) \;:\; \sigma\in\Sigma\}\right) &= \prod_{i\in[m]\,:\,g(\sigma_i)\neq 0} \left((y-\sigma_i)\cdot g(\sigma_i)\right)^{e_i} \\ &= \left(\prod_{i\in[m]\,:\,g(\sigma_i)\neq 0} (y-\sigma_i)^{e_i}\right)\cdot \mathrm{Esym}_r\left(\{g(\sigma) \;:\; \sigma\in\Sigma\}\right) \\ \therefore \mathrm{Filter}(f \mid \text{non-roots}(g)) &= \frac{\mathrm{Esym}_r\left(\{h(\sigma,y) \;:\; \sigma\in\Sigma\}\right)}{\mathrm{Esym}_r\left(\{g(\sigma) \;:\; \sigma\in\Sigma\}\right)}\end{aligned}$$

Once again by Lemma 18, both the numerator and denominator on the RHS can be expressed as circuits of size $\mathrm{poly}(d_1,d_2)$ and depth $O(1)$ over the coefficients of f and g and that gives us the circuit families $\{A_{d_1,d_2,k}\}$ and $\{B_{d_1,d_2,k}\}$. ◀

We are now ready to prove Theorem 20.

Proof of Theorem 20. Without loss of generality, let us assume that $d_1 := \deg f > \deg g =: d_2$ (by working with $f, g-f$ if their degrees were the same). Let $\{\sigma_1,\ldots,\sigma_m\} \in \overline{\mathbb{K}}$ be the union of roots of f and g, and let

$$\begin{aligned} f(y) &= (y-\sigma_1)^{e_1}\cdots(y-\sigma_m)^{e_m}, \\ g(y) &= (y-\sigma_1)^{e_1'}\cdots(y-\sigma_m)^{e_m'} \\ \implies \gcd(f,g) &= (y-\sigma_1)^{e_1^*}\cdots(y-\sigma_m)^{e_m^*} \quad \text{where } e_i^* = \min(e_i,e_i').\end{aligned}$$

Consider the polynomial $F(y,z) = f(y) + z\cdot g(y)$ for a fresh variable z, which is monic in y since $\deg f > \deg g$. The key observation is that $\gcd(f,g) = \mathrm{Filter}(F(y,z) \mid \mathrm{roots}(g))$ and the theorem essentially follows from Lemma 21 and this observation. We formalize this below.

Since $F(y,z)$ is clearly divisible by $\gcd(f,g)$, we can write $F(y,z)$ as

$$F(y,z) = \gcd(f,g)\cdot F'(y,z)$$

for some $F'(y,z) \in \mathbb{K}[y,z]$. Note that, for every σ_i, we have $F'(\sigma_i,z) \neq 0$ (for otherwise $(y-\sigma_i)^{e_i^*+1}$ divides both f and g). Therefore,

$$\begin{aligned} &\mathrm{Filter}(F(y,z) \mid \text{non-roots}(g)) = F'(y,z) \\ \implies &\mathrm{Filter}(F(y,z) \mid \mathrm{roots}(g)) = \frac{F(y,z)}{\mathrm{Filter}(F(y,z) \mid \text{non-roots}(g))} = \gcd(f,g).\end{aligned}$$

By Lemma 21, we have circuit families $\{A_{d_1,d_2,k}\}$, $\{B_{d_1,d_2,k}\}$, $\{T_{d_1,d_2,k}\}$ such that

$$\begin{aligned} \mathrm{Filter}(F(y,z) \mid \text{non-roots}(g)) &= \frac{A_{d_1,d_2,r}(\overline{\mathrm{coeff}_y}(F), \overline{\mathrm{coeff}}(g))}{B_{d_1,d_2,r}(\overline{\mathrm{coeff}_y}(F), \overline{\mathrm{coeff}}(g))} \\ \text{where } r &= \max\left\{k \le \mathrm{poly}(d_1,d_2) \;:\; T_{d_1,d_2,k}(\overline{\mathrm{coeff}_y}(F), \overline{\mathrm{coeff}}(g)) \neq 0\right\} \\ \implies \gcd(f,g) &= \mathrm{Filter}(F(y,z) \mid \mathrm{roots}(g)) \\ &= \frac{F(y,z)\cdot B_{d_1,d_2,r}(\overline{\mathrm{coeff}_y}(F), \overline{\mathrm{coeff}}(g))}{A_{d_1,d_2,r}(\overline{\mathrm{coeff}_y}(F), \overline{\mathrm{coeff}}(g))}.\end{aligned}$$

Since each coefficient of F can be expressed as a polynomial function of the coefficients of f and g, the above circuit families $\{A_{d_1,d_2,k}\}$, $\{B_{d_1,d_2,k}\}$, $\{T_{d_1,d_2,k}\}$ can also be expressed as $\mathrm{poly}(d_1, d_2)$ size depth $O(1)$ circuit families over the coefficients of f and g. It can also be checked that the advice parameter $r = d_1 - \deg(\gcd(f, g))$.

As stated, although $\gcd(f(y), g(y)) \in \mathbb{K}[y]$ and is thus independent of the variable z, the RHS of the above equations consist of numerators and denominators that depend on z. To eliminate z from the RHS, we can remove any powers of z dividing the numerator and denominator and then set z to zero. This can be achieved by expanding the test circuits to also find the smallest i such that $[z^i]\left\{A_{d_1,d_2,r}(\overline{\mathrm{coeff}_y}(F), \overline{\mathrm{coeff}}(g))\right\} \neq 0$ and set the numerator and denominator circuits as

$$P_{d_1,d_2,(r,i)}(\overline{\mathrm{coeff}}(f), \overline{\mathrm{coeff}}(g)) := [z^i]\left\{F(y,z) \cdot B_{d_1,d_2,r}(\overline{\mathrm{coeff}_y}(F), \overline{\mathrm{coeff}}(g))\right\},$$
$$Q_{d_1,d_2,(r,i)}(\overline{\mathrm{coeff}}(f), \overline{\mathrm{coeff}}(g)) := [z^i]\left\{A_{d_1,d_2,r}(\overline{\mathrm{coeff}_y}(F), \overline{\mathrm{coeff}}(g))\right\}$$

respectively. Thus, overall, we obtain size $\mathrm{poly}(d_1, d_2)$ and depth $O(1)$ circuits over the coefficients of f and g as claimed by the theorem. ◀

5 Improved closure results for fields of small characteristic

In this section, we use the Theorem 2 to prove the following closure result for constant-depth circuits over fields of characteristic p. This statement is similar to a result in [2] with one key difference – the circuit for $g(\mathbf{x})^{p^\ell}$ in Theorem 22 is over the underlying field itself and not over the algebraic closure (or a high degree extension) of the base field $\mathbb{F}_q$, provided q is polynomially large. This extends the results of Kaltofen [10] for fields of characteristic p for natural subclasses of algebraic circuits such as constant-depth circuits, formulas, branching programs etc. The key technical ingredient in the proof again is the use of Theorem 2.

▶ **Theorem 22** (Complexity of factors). *Let $\mathbb{F}_q$ be a field of positive characteristic p that is polynomially large (i.e., there is an absolute constant c such that $|\mathbb{F}| \geq (snd)^c$). 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 $\mathbb{F}_q$.*

We start with a few preliminaries necessary for the proof of the theorem. We shall interpret $P(\mathbf{x})$ as an element of $\mathbb{F}[\mathbf{x}, y]$, and apply a map $\Psi : x_i \mapsto tx_i + a_i y + b_i$ on $P(\mathbf{x})$ for randomly chosen a_is and b_is, to ensure certain properties[3]. Thus, without loss of generality, we can work with a bivariate polynomial $P(t, y)$ over a field that contains the other variables. Some preliminaries will also be stated for bivariate polynomials.

5.1 Hasse derivatives

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 [8, Appendix C].

▶ **Definition 23** (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 $F(t, y + z)$.*

[3] We expand on this in the beginning of Subsection 5.4. For more details, please refer to the preliminaries of [2].

▶ **Lemma 24** (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,*

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

5.2 Furstenberg's theorem over small characteristic fields

While Theorem 8 states Furstenberg's theorem for power series roots that appear as a linear factor with multiplicity one, there is a version of the theorem that works even when the roots appear with arbitrary multiplicity e (see [2, Theorem 3.1]) if e is invertible in the field $\mathbb{F}$. Over a field of characteristic p, p might divide e and thus e might not be invertible. The following version of Furstenberg's theorem over small characteristic is very similar to Theorem 8, 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 25** (Furstenberg's theorem over small characteristic fields (Theorem A.3 in [2])). *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

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

We can further simplify the expression in Theorem 25 to get a version of Corollary 9 over small characteristic fields.

▶ **Corollary 26** (Corollary A.5 in [2]). *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,

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

The following claim would be helpful to express $\alpha = Q(0,0)$ as an appropriate Hasse derivative of P.

▷ Claim 27. Suppose $P(t,y) = (y - \varphi(t))^{p^\ell e} \cdot Q(t,y) \in \mathbb{F}[\![t,y]\!]$ with $\gcd(p,e) = 1$, $\varphi(0) = 0$ and $Q(0,0) = \alpha \neq 0$. Then,

$$\alpha = \mathrm{D}_y^{(p^\ell \cdot e)} P(0,0)$$

Proof. Let $F(t,y) = (y-\varphi(t))^{p^\ell \cdot e}$. Then, by Lemma 24,

$$\mathrm{D}_y^{(p^\ell \cdot e)} P(t,y) = \sum_{i+j=p^\ell \cdot e} \mathrm{D}_y^{(i)} F(t,y) \cdot \mathrm{D}_y^{(j)} Q(t,y)$$

$$\implies \mathrm{D}_y^{(p^\ell \cdot e)} P(0,0) = \sum_{i+j=p^\ell \cdot e} \mathrm{D}_y^{(i)} F(0,0) \cdot \mathrm{D}_y^{(j)} Q(0,0)$$

Note that $\mathrm{D}_y^{(i)} F = [z^i]\left\{(y^{p^\ell} + z^{p^\ell} - \varphi^{p^\ell})^e\right\}$ and is a nonzero polynomial only for i of the form $p^\ell \cdot e'$ for some $0 \le e' \le e$, and $\mathrm{D}_y^{(p^\ell e)} F = 1$. Furthermore, for every $e' < e$, we have $\mathrm{D}_y^{(p^\ell e')} F$ is divisible by $(y-\varphi)^{p^\ell}$ and hence $\mathrm{D}_y^{(p^\ell e')} F(0,0) = 0$ for all $e' < e$. Therefore,

$$\begin{aligned}\mathrm{D}_y^{(p^\ell \cdot e)} P(0,0) &= \sum_{i+j=p^\ell \cdot e} \left(\mathrm{D}_y^{(i)} F(0,0)\right) \cdot \left(\mathrm{D}_y^{(j)} Q(0,0)\right) \\ &= \left(\mathrm{D}_y^{(p^\ell e)} F(0,0)\right) \cdot \left(\mathrm{D}_y^{(0)} Q(0,0)\right) = Q(0,0) = \alpha. \end{aligned}$$ ◁

5.3 Complexity of power series roots over small characteristic fields

Using Theorem 25 and Corollary 26, we get the following analogue of Theorem 10 over arbitrary fields of small characteristic.

▶ **Theorem 28** (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 $\mathbb{F}$ computing $\mathrm{Hom}_{\le d}\left[\varphi^{p^\ell}\right]$ such that*

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

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

Proof. The theorem follows from Corollary 26 that asserts that

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

and Corollary 4 shows that the RHS is computable by a size $\mathrm{poly}(s,n,d)$ and depth $\Delta + O(1)$ circuit. ◀

5.4 Proof of Theorem 22

We will need the following generalization of Lemma 18, whose proof we will defer to the end of this section.

▶ **Lemma 29.** *Let $\mathbb{F}$ be any large enough field (i.e., $|\mathbb{K}| \ge (sdn)^c$ for some absolute constant $c > 0$). Let $f(\mathbf{x},y), g(\mathbf{x},y), h(\mathbf{x},y) \in \mathbb{F}[\mathbf{x},y]$ be polynomials that are monic in y of degrees d_f, d_g and d_h respectively, with $\gcd(f,h) = 1$, that are given by size s, depth Δ circuits; let $d = \max(d_f, d_g, d_h)$. If $\{\sigma_1, \ldots, \sigma_{d_f}\}$ be the multi-set of y-roots of f over $\overline{\mathbb{F}(\mathbf{x})}$, then for any $1 \le r \le d_f$, we can get circuits $A(\mathbf{x}), B(\mathbf{x})$ of size $\mathrm{poly}(s,n,d)$ and depth $\Delta + O(1)$ such that*

$$\mathrm{Esym}_r\left(\left\{\frac{g(\mathbf{x},\sigma_i)}{h(\mathbf{x},\sigma_i)} \;:\; i \in [d_f]\right\}\right) = \frac{A(\mathbf{x})}{B(\mathbf{x})}.$$

We are now ready to complete the proof of Theorem 22.

Proof of Theorem 22. By interpreting $P(\mathbf{x})$ as an element of $\mathbb{F}[\mathbf{x}, y]$, let $\tilde{P}(t, y) = \Psi(P(\mathbf{x}, y)) \in \mathbb{F}[\mathbf{x}][t, y]$ for a map $\Psi : x_i \mapsto tx_i + a_i y + b_i$ where the a_is and b_is are chosen at random. This will ensure that $\tilde{P}(t, y)$ is monic in y, the square-free part of $\tilde{P}(t, y)$ is also square-free at $t = 0$, and that $\tilde{P}(t, y) \in \mathbb{F}[\mathbf{x}][t, y]$ is also computable by a size $\text{poly}(s, d, n)$, depth $\Delta + O(1)$ circuit over $\mathbb{F}$. From Gauss' Lemma, we know that all the factors of $\tilde{P}$ can now be assumed to be monic in y without loss of generality, and viewing $\tilde{P}$ as a bivariate in t and y with coefficients in $\mathbb{F}(\mathbf{x})$ maintains the factorization pattern of the polynomial. Furthermore, the map can be inverted using the transformation $x_i \mapsto x_i - a_i y - b_i$ and $t \mapsto 1$ and this gives us a way to argue about factors of P using factors of $\tilde{P}$. Thus[4], it suffices to show that for an arbitrary factor $g(t, y) \in \mathbb{F}[\mathbf{x}][t, y]$ of $\tilde{P}(t, y)$, if the multiplicity of $g(t, y)$ is $p^\ell \cdot e$ with $\gcd(e, p) = 1$, then $g(t, y)^{p^\ell}$ is computable by a $\text{poly}(s, d, n)$ size, depth $\Delta + O(1)$ circuit over the field $\mathbb{F}$.

First we note that the existence of a power series factorization of $\tilde{P}(t, y)$ follows *almost* immediately from Lemma 7. We cannot apply Lemma 7 directly because the preprocessing map Ψ only ensures that the square-free part of $\tilde{P}(t, y)$ remains square-free when t is set to zero. Applying Lemma 7 on the square-free part of $\tilde{P}(t, y)$, denoted by $\hat{P}(t, y)$, gives us the factorization $\hat{P}(t, y) = \prod_{i \in [m]}(y - \varphi_i(t))$, where each $\varphi_i(t)$ has a distinct constant term $\varphi_i(0) \in \overline{\mathbb{F}}$.

Suppose $g(t, y)$ is a factor with multiplicity $p^\ell \cdot e$. For each power series y-root $\varphi(t)$ of $g(t, y)$, we can write $\tilde{P}(t, y) = (y - \varphi(t))^{p^\ell \cdot e} Q(t, y)$ for some $Q(t, y)$ that satisfies $Q(0, \varphi(0)) \neq 0$. Translating y to $y + \varphi(0)$, we can use Corollary 26 on $\tilde{P}(t, y + \varphi(0))$ to show that $\text{Hom}_{\leq d}[\varphi(t)^{p^\ell}]$ is equal to

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

where $\alpha = Q(0, \varphi(0)) \neq 0$, which is also equal to $\mathrm{D}_y^{(p^\ell \cdot e)} \tilde{P}(0, \varphi(0))$ by Claim 27. Thus, we define the rational function $R(z)$ as

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

If $\hat{P}(0, y) = \prod_{i=1}^m (y - \varphi_i(0))$, and $g(0, y)^{p^\ell} = \prod_{i \in S}(y - \varphi_i(0))^{p^\ell}$ for some subset $S \subset [m]$, then the power series roots of $g(t, y)$ are precisely $\varphi_j(t)$ for $j \in S$ and hence

$$\begin{aligned}
g(t, y)^{p^\ell} &= \prod_{j \in S} (y^{p^\ell} - \varphi_j(t)^{p^\ell}) = \prod_{j \in S} (y^{p^\ell} - R(\varphi_j(0))) \bmod t^{d+1} \\
\implies g(t, y)^{p^\ell} &= \text{Hom}_{\leq d}\left[g'(t, y)\right] \\
\text{where} \quad g'(t, y) &= \prod_{j \in S} (y^{p^\ell} - R(\varphi_j(0))) \\
&= \sum_{k=0}^{|S|} y^{p^\ell(|S|-k)} \cdot (-1)^{|S|-k} \cdot \text{Esym}_k(\{R(\varphi_j(0)) \ : \ j \in S\})
\end{aligned}$$

[4] For more details, we encourage the reader to refer to the preliminaries in [2].

Thus, it suffices to show that $\mathrm{Esym}_k(\{R(\varphi_j(0)) \ : \ j \in S\})$ for all $1 \le k \le d$ can be computed by a $\mathrm{poly}(s,n,d)$ size depth $\Delta + O(1)$ circuit as that would immediately yield similar size and depth bounds for $g(t,y)^{p^\ell}$ by Corollary 4.

Note that the rational function $R(z)$ can be easily expressed as $\frac{R_{\mathrm{num}}(z)}{R_{\mathrm{denom}}(z)}$ where $R_{\mathrm{num}}(z)$ 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 3 and Corollary 4). Furthermore, $R_{\mathrm{denom}}(z) = (\mathrm{D}_y^{(p^\ell \cdot e)}(\tilde{P})(0,z))^{2e(d+p^\ell)+1}$, so it remains nonzero element of $\overline{\mathbb{F}_q}$ when z is set to any root of $\tilde{P}(0,y)$ (Claim 27).

Note that $\mathrm{Esym}_k(\{\varphi_j(0) \ : \ j \in S\})$ is in $\mathbb{F}$ since they are the (signed) coefficients of $g(0,y)$. Thus, by Lemma 29, we can express $\mathrm{Esym}_k(\{R(\varphi_j(0)) \ : \ j \in S\})$ as a ratio of two $\mathrm{poly}(s,d,n)$ size depth $\Delta + O(1)$ circuits over $\mathbb{F}$. By eliminating the division gate using the standard technique of Strassen (which works in constant depth), we obtain a $\mathrm{poly}(s,d,n)$ size depth $\Delta + O(1)$ circuits over $\mathbb{F}$ for $\mathrm{Esym}_k(\{R(\varphi_j(0)) \ : \ j \in S\})$ for each $1 \le k \le d$, and thus a similar bound for $g(t,y)^{p^\ell}$. ◀

Proof of Lemma 29. Note that

$$\mathrm{Esym}_r\left(\frac{u_1}{v_1}, \ldots, \frac{u_{d_f}}{v_{d_f}}\right) = [t^r]\left\{(v_1 + tu_1)\cdots(v_{d_f} + tu_{d_f})\right\} \cdot \frac{1}{v_1 \ldots v_{d_f}}. \tag{3}$$

Let $\left\{\tau_1, \ldots, \tau_{d_g}\right\}, \{\rho_1, \ldots, \rho_{d_h}\} \subset \overline{\mathbb{F}(\mathbf{x})}$ be the multi-set of y-roots of g and h respectively. Thus, setting $u_i = g(\mathbf{x}, \sigma_i)$ and $v_i = h(\mathbf{x}, \sigma_i)$ in (3), we see that

$$\begin{aligned}\Gamma(\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\rho}) &= [t^r]\left\{(h(\mathbf{x},\sigma_1) + tg(\mathbf{x},\sigma_1))\cdots(h(\mathbf{x},\sigma_{d_f}) + tg(\mathbf{x},\sigma_{d_f}))\right\} \\ &= [t^r]\left\{\left(\prod_{j=1}^{d_h}(\sigma_1 - \rho_j) + t\prod_{k=1}^{d_g}(\sigma_1 - \tau_k)\right)\cdots\left(\prod_{j=1}^{d_h}(\sigma_{d_f} - \rho_j) + t\prod_{k=1}^{d_g}(\sigma_{d_f} - \tau_k)\right)\right\}\end{aligned}$$

is multi-symmetric with respect to $\left\{\sigma_1, \ldots, \sigma_{d_f}\right\} \sqcup \left\{\tau_1, \ldots, \tau_{d_g}\right\} \sqcup \left\{\rho_1, \ldots, \rho_{d_h}\right\}$ that is computable by a $\mathrm{poly}(s,d)$ size depth $\Delta + O(1)$ circuit (over $\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\rho}$). Thus, by Theorem 16, we have a circuit of size $\mathrm{poly}(s,d)$ and depth $\Delta + O(1)$ over the elementary symmetric polynomials of $\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\rho}$, which are the coefficients of f, g and h, and they are also computable by size $\mathrm{poly}(s,d)$ depth $\Delta + O(1)$ circuits (by Corollary 4). That gives us the circuit $A(\mathbf{x})$.

The circuit $B(\mathbf{x})$ is to compute $\prod_{i=1}^{d_f} h(\mathbf{x}, \sigma_i) = \mathrm{Esym}_{d_f}(h(\mathbf{x},\sigma_1), \ldots, h(\mathbf{x},\sigma_{d_f}))$ and Lemma 18 shows that $B(\mathbf{x})$ is also computable by a size $\mathrm{poly}(s,d)$ depth $\Delta + O(1)$ circuit. Since $\gcd(f,h) = 1$, they do not share any y-roots and hence the division $A(\mathbf{x})/B(\mathbf{x})$ is well-defined. ◀

References

1 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. Full version at `arXiv:2404.10839`. `doi:10.1109/FOCS61266.2024.00138`.

2 Somnath Bhattacharjee, Mrinal Kumar, Shanthanu Rai, Varun Ramanathan, Ramprasad Saptharishi, and Shubhangi Saraf. Closure under factorization from a result of furstenberg. *Electronic Colloquium on Computational Complexity (ECCC)*, TR25-084, 2025. URL: `https://eccc.weizmann.ac.il/report/2025/084`.

3 Markus Bläser and Gorav Jindal. On the Complexity of Symmetric Polynomials. In Avrim Blum, editor, *10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*, volume 124 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 47:1–47:14, Dagstuhl, Germany, 2019. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/LIPIcs.ITCS.2019.47`.

4 Prasad Chaugule, Mrinal Kumar, Nutan Limaye, Chandra Kanta Mohapatra, Adrian She, and Srikanth Srinivasan. Schur polynomials do not have small formulas if the determinant does not. *computational complexity*, 32(1):3, 2023. doi:10.1007/s00037-023-00236-x.

5 David A Cox, John Little, and Donal O'shea. *Using Algebraic Geometry*. Springer New York, NY, 2005. doi:10.1007/b138611.

6 Laszlo Csanky. Fast parallel matrix inversion algorithms. *SIAM J. Comput.*, 5(4):618–623, 1976. doi:10.1137/0205040.

7 Pranjal Dutta, Nitin Saxena, and Amit Sinhababu. Discovering the roots: Uniform closure results for algebraic classes under factoring. *J. ACM*, 69(3), June 2022. doi:10.1145/3510359.

8 Michael A. Forbes. *Polynomial identity testing of read-once oblivious algebraic branching programs*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2014. URL: https://hdl.handle.net/1721.1/89843.

9 Harry Furstenberg. Algebraic functions over finite fields. *Journal of Algebra*, 7(2):271–277, 1967. doi:10.1016/0021-8693(67)90061-0.

10 Erich Kaltofen. Factorization of Polynomials Given by Straight-Line Programs. In *Randomness and Computation*, pages 375–412. JAI Press, 1989. URL: https://users.cs.duke.edu/~elk27/bibliography/89/Ka89_slpfac.pdf.

11 Leslie G. Valiant, Sven Skyum, S. Berkowitz, and Charles Rackoff. Fast Parallel Computation of Polynomials Using Few Processors. *SIAM Journal of Computing*, 12(4):641–644, 1983. Preliminary version in the *6th Internationl Symposium on the Mathematical Foundations of Computer Science (MFCS 1981)*. doi:10.1137/0212043.

12 Joachim von zur Gathen and Jürgen Gerhard. *Modern Computer Algebra*. Cambridge University Press, 3 edition, 2013. doi:10.1017/CBO9781139856065.