Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 37 (2019)

Closure of VP under taking factors: a short and simple proof

Chi-Ning Chou* Mrinal Kumar† Noam Solomon‡

Abstract

In this note, we give a short, simple and almost completely self contained proof of a classical result of Kaltofen [Kal86, Kal87, Kal89] which shows that if an n variate degree d polynomial f can be computed by an arithmetic circuit of size s, then each of its factors can be computed by an arithmetic circuit of size at most $\mathsf{poly}(s, n, d)$.

However, unlike Kaltofen's argument, our proof does not directly give an efficient algorithm for computing the circuits for the factors of f.

1 Introduction

Polynomial factorization is a fundamental problem at the intersection of algebra and computation and has been intensively studied in algebraic complexity theory. Given a multivariate polynomial $f \in \mathbb{F}[x_1, x_2, \ldots, x_n]$, the goal is to find irreducible factors of f. The nature of these algorithms as well as their efficiency varies depending upon on how the input polynomial is given. Two natural representations which are often used in this context are the monomial representation (where the polynomial is given as sum of its monomials) and the circuit representation (where the polynomial is given as an arithmetic circuit). In this note, we focus on the latter. In this setting, we are given an arithmetic circuit computing a multivariate polynomial, and the goal is to output arithmetic circuits for all its irreducible factors. The problem has been studied in both the whitebox setting (where we have access to the internal wirings of the input circuit) and in the blackbox setting (where we only have query access to the input circuit). In a sequence of extremely influential results in the 1980's [Kal86, Kal87, Kal89, KT90], Kaltofen (and Kaltofen and Trager [KT90]) gave efficient randomized algorithms for this problem. A consequence of these results which has had extremely interesting applications in algebraic complexity theory [KI04] is that if an n-variate degree d polynomial has an arithmetic circuit of size s, then each of its factors has an arithmetic circuit of size $\mathsf{poly}(s, n, d)$. In other words, the complexity class VP of polynomials is closed under taking factors.

In addition to being natural mathematical questions on their own, these *closure results* for polynomial factorization seem crucial to our current understanding of hardness randomness tradeoffs in algebraic complexity [KI04, DSY09, CKS18a]. In this note, we give a short, simple and almost completely self contained proof of the closure of VP under taking factors. More formally, we give a new[1] proof of the following result of Kaltofen.

*School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA. Supported by Boaz Barak's NSF awards CCF 1565264 and CNS 1618026. Email: chiningchou@g.harvard.edu.

†Department of Computer Science, University of Toronto, Canada. A part of this work was done during the semester on Lower bounds in Computational Complexity at Simons Institute for the Theory of Computing, Berkeley, CA, USA. Email: mrinalkumar08@gmail.com.

‡Department of Mathematics, MIT, Cambridge, Massachusetts, USA. Email: noam.solom@gmail.com.

[1] as far as we know.

ISSN 1433-8092

Theorem 1.1 (Kaltofen)**.** *Let $f \in \mathbb{F}[x_1, x_2, \ldots, x_{n-1}, y]$ be an n-variate degree d polynomial which can be computed by an arithmetic circuit of size s. Let g be a polynomial such that g divides f. Then, g can be computed by an arithmetic circuit of size at most $\mathsf{poly}(s, n, d)$.*

The original proof of Theorem 1.1 relies on some beautiful and neat mathematical ideas like Hensel's lifting, effective Hilbert's Irreducibility Theorem, etc, which are useful and interesting on their own. For our proof, we only rely on a simple and natural multivariate version of the classical Newton Iteration technique and the fact that the Resultant of two univariates tells us exactly when they have a non-trivial greatest common divisor (GCD). We hope that this simpler proof can shed some more insight on this closure result (and hopefully some others, which are yet to be discovered), and is more accessible to readers with a less detailed background in algebra. A cost of this *simplicity* is that unlike in the work of Kaltofen, we do not get an algorithm for factoring multivariate polynomials given by arithmetic circuits.

Besides Kaltofen's original proof, there is a considerably simpler proof due to Bürgisser [Bür04] showing that VP is closed under taking factors. Bürgisser uses the classical univariate Newton Iteration to obtain a power series approximation of a root of a multivariate polynomial when it is viewed as a univariate in one of the variables. This power series approximation of the root to a sufficiently high enough accuracy is then used to obtain an appropriate irreducible factor of the input polynomial. This step requires setting up and solving an appropriate system of linear equations. A variant of this argument is also present in the works of Dvir et al. [DSY09], of Oliveira [Oli16], of Dutta et al. [DSS17] and an earlier work of the authors [CKS18a, CKS18b]. At a high level, these proofs go via an iterative step to approximate a root (or many roots), and a clean up step where a factor is recovered from this approximation.

In our proof, we directly recover the factors at the end of the slightly more complicated iterative step (a *multivariate* Newton iteration as opposed to a *univariate* one), and the clean up is essentially trivial.

Our proof follows immediately from the following two lemmas.

Lemma 1.2. *Let $f \in \mathbb{F}[x_1, x_2, \ldots, x_{n-1}, y]$ be an n-variate degree d polynomial which can be computed by an arithmetic circuit of size at most s. If g and h are polynomials of degree at least 1 such that $f = g \cdot h$ and $\mathsf{GCD}(g, h) = 1$, then g and h have a circuit of size at most $\mathsf{poly}(s, n, d)$.*

Lemma 1.3. *Let $f \in \mathbb{F}[x_1, x_2, \ldots, x_{n-1}, y]$ be an n-variate degree d polynomial which can be computed by an arithmetic circuit of size at most s. If there is a polynomial g and an integer e such that $f = g^e$, then g has a circuit of size at most $\mathsf{poly}(s, n, d)$.*

The proof of Lemma 1.2 and Lemma 1.3 will be provided in Section 3 and Section 4 respectively.

2 Notations and Preliminaries

We follow the following notation.

- Throughout the paper, $\mathbb{F}$ is a field of characteristic zero or sufficiently large.
- For a positive integers n, $[n]$ denotes the set $\{0, 1, 2, \ldots, n-1\}$.
- We use boldface letters to denote ordered tuples of objects. For instance, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, or $\mathbf{g} = (g_0, g_1, \ldots, g_{d_1-1})$. The length of these tuples and the precise indexing is defined before the specific notation is invoked. The sum of two such tuples of the same length is their coordinate wise sum.

- We say that a function Ψ of parameters $n_1, n_2, \ldots, n_t$ taking values in $\mathbb{Z}^+$ is $\mathsf{poly}(n_1, n_2, \ldots, n_t)$ if there is a polynomial Φ such that for all sufficiently large values of $n_1, n_2, \ldots, n_t$, $\Psi(n_1, n_2, \ldots, n_t)$ is upper bounded by $\Phi(n_1, n_2, \ldots, n_t)$.

2.1 Arithmetic Circuits

Arithmetic circuits (also historically referred to as straight line programs) provide a succinct and compact representation for multivariate polynomials. Formally, they are defined as follows.

Definition 2.1 (Arithmetic Circuit)**.** *Let $\mathbb{F}$ be a field and $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ be variables. An arithmetic circuit C over $\mathbb{F}$ and $\mathbf{x}$ is a directed acyclic graph where the vertices are called gates. Every gate with in-degree zero is an input gate and is labeled by a single variable from $\mathbf{x}$ or a field element from $\mathbb{F}$. The other gates are labeled by $+$ (sum gates) or $\times$ (product gates). The gates with out-degree zero are called output gates.*

Each gate in the circuit C computes a polynomial in $\mathbb{F}[\mathbf{x}]$ in a natural and inductive way. For an input gate g, the polynomial it computes is the corresponding variable or field element. A $+$ (resp. $\times$) gate g computes the sum (resp. products) of the polynomials computed at the gates which have a directed edge to g. The size of an arithmetic circuit C is defined as the number of edges in C. ◊

The following lemma structural lemma about arithmetic circuits will be useful for our proof.

Lemma 2.2 (Homogenization)**.** *Let C be a multi-output arithmetic circuit of size s with outputs $f_1, f_2, \ldots, f_t$. Then, for every k, there is a homogeneous circuit of size at most $O(k^2 s)$ which outputs the homogeneous components of degree at most k of $f_1, f_2, \ldots, f_t$.*

We refer the reader to any standard resource (such as the survey by Shpilka and Yehudayoff [SY+10]) for a proof for Lemma 2.2 and for a general overview of arithmetic circuit complexity.

2.2 Multivariate Taylor's Expansion

We use the following lemma which is an easy consequence of the classical multivariate Taylor expansion for polynomials.

Lemma 2.3 (Truncated Multivariate Taylor's Expansion)**.** *Let $f \in \mathbb{F}[\mathbf{x}]$ and $\mathbf{a} \in \mathbb{F}^n$, we have*

$$f(\mathbf{x}+\mathbf{a}) \equiv f(\mathbf{a}) + \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(\mathbf{a}) \cdot x_i \mod \langle \mathbf{x} \rangle^2 .$$

In the proof for Theorem 1.1, we need a variant of this lemma where the variables are from $\mathbb{F}[\mathbf{x}]$ instead of $\mathbf{x}$. We state it as a corollary of Lemma 2.3

Corollary 2.4. *Let $f_1, f_2, \ldots, f_m, p_1, p_2, \ldots, p_m \in \mathbb{F}[\mathbf{x}]$ where $\deg(p_i) \geq k$ for each $i \in [m]$ and $Q \in \mathbb{F}[z_1, z_2, \ldots, z_m]$, we have*

$$Q(\mathbf{f}+\mathbf{p}) \equiv Q(\mathbf{f}) + \sum_{i \in [m]} \frac{\partial Q}{\partial z_i}(\mathbf{f}) \cdot p_i \mod \langle \mathbf{x} \rangle^{k+1} .$$

2.3 Jacobian Matrix

The Jacobian matrix is a matrix that contains the partial derivatives of a vector of multivariate functions.

Definition 2.5 (Jacobian Matrix)**.** *Let $f_1, f_2, \ldots, f_m \in \mathbb{F}[\mathbf{x}]$, the Jacobian matrix of $\mathbf{f}$ with respect to $\mathbf{x}$ is a $m \times n$ matrix denoted as $\mathsf{Jacobian}_{\mathbf{x}}(\mathbf{f})$ where the (i,j) entry is defined as $\frac{\partial f_i}{\partial x_j}$ for each $i \in [m]$ and $j \in [n]$.* ◊

2.4 GCD and Resultant

For any two polynomials $g, h \in \mathbb{F}[\mathbf{x}]$, we can define their greatest common divisor (GCD) as follows.

Definition 2.6 (GCD)**.** *Let $\mathbb{F}$ be a field and $g, h \in \mathbb{F}[\mathbf{x}]$. The greatest common divisor (GCD) of g and h is $\mathsf{GCD}(g,h) = f$ if f divides both g and h, and for any $\tilde{f} \in \mathbb{F}[\mathbf{x}]$ that divides both g and h, we have $\tilde{f}$ dividing f. For any $g, h \in \mathbb{F}[\mathbf{x}][y]$, we define $\mathsf{GCD}_y(g,h)$ to be the greatest common divisor (GCD) of g and h with respect to y.* ◊

It turns out that there is a clean and useful mathematical condition to check whether the GCD of two polynomials is non-constant using *resultant*.

Definition 2.7 (Resultant)**.** *Let $g, h \in \mathbb{F}[\mathbf{x}][y]$ and $d_1, d_2 \in \mathbb{N}$ such that $g = \sum_{i=0}^{d_1} g_i y^i$ and $h = \sum_{j=0}^{d_2} h_j y^j$ for some $g_i, h_j \in \mathbb{F}[\mathbf{x}]$. The resultant $R_y(g,h)$ is the determinant of a following $(d_1 + d_2) \times (d_1 + d_2)$ matrix, called the Sylvester matrix $S(g,h)$.*

$$S(g,h) = \begin{pmatrix} g_0 & 0 & \cdots & 0 & h_0 & 0 & \cdots & 0 \\ g_1 & g_0 & \cdots & 0 & h_1 & h_0 & \cdots & 0 \\ g_2 & g_1 & & \vdots & \vdots & h_1 & & \vdots \\ \vdots & \vdots & \ddots & g_0 & h_{d_2} & \vdots & \ddots & 0 \\ g_{d_1} & g_{d_1-1} & & g_1 & 0 & h_{d_2} & & h_0 \\ 0 & g_{d_1} & & g_2 & 0 & 0 & & h_1 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & g_{d_1} & 0 & 0 & \cdots & h_{d_2} \end{pmatrix}.$$

Specifically, for $i \in \{1, 2, \ldots, d_2\}$, the i^{th} column of S is equal to $(0, \ldots, 0, g_{d_1}, g_{d_1-1}, \ldots, g_1, g_0, 0, \ldots, 0)$, where there are $i-1$ zeroes in the prefix. For $j \in \{d_2+1, \ldots, d_1+d_2\}$, t he j^{th} column of S equals $(0, \ldots, 0, h_{d_2}, h_{d_2-1}, \ldots, h_1, h_0, 0, \ldots, 0)$, where there are $j - d_2 - 1$ zeroes in the prefix. ◊

The following lemma shows that $R_y(g,h) = 0$ if and only if g and h have a common non-constant factor.

Lemma 2.8 (Capturing the GCD via the Resultant)**.** *Let $g, h \in \mathbb{F}[\mathbf{x}][y]$ and $d_1, d_2 \in \mathbb{N}$ such that $g = \sum_{i=0}^{d_1} g_i y^i$ and $h = \sum_{j=0}^{d_2} h_j y^j$ for some $g_i, h_j \in \mathbb{F}[\mathbf{x}]$. Then, $R_y(g,h) = 0$ if and only if $\mathsf{GCD}_y(g,h)$ has degree at least 1 in y.*

To keep this note short, we refer the reader to any standard resource (such as the lecture notes by Sudan [Sud98]) for a proof for Lemma 2.8.

3 Proof of Lemma 1.2

We have polynomials f, g and h such that $f = g \cdot h$ and f has an arithmetic circuit of size at most s. The goal is to show that g and h have circuits of size at most $\mathsf{poly}(s, n, d)$. Let d_1, d_2 be the degrees of g and h respectively and let $d_1 \leq d_2$.If g and h are variable disjoint then we can obtain a circuit for g by just setting the variables in h to random values such that h does not vanish and then scaling by an appropriate field constant. So, we focus on the interesting case when g and h share

at least one common variables. Let y be such a variable. By taking a random (from a large enough grid) $\mathbf{a} \in \mathbb{F}^n$ and replacing x_i by $x_i + a_i y$, we can guarantee that the coefficient of y^d in f, y^{d_1} in g and y^{d_2} in h are all non-zero field elements. Without loss of generality, we assume that these constants are all 1 (or else we scale everything by a constant). In the rest of this section, we view the identity $f = g \cdot h$ as an identity in $\mathbb{F}[\mathbf{x}][y]$. Note that at the end of the above transformation, $\mathsf{GCD}(f, g)$ continues to be 1 when viewing them as univariates in y. We know that f has a small circuit, and the goal is to show that g and h have small circuits.

Let $f_0, f_1, \ldots, f_{d-1}, g_0, g_1, \ldots, g_{d_1-1}, h_0, h_1, \ldots, h_{d_2-1} \in \mathbb{F}[\mathbf{x}]$ be polynomials such that

$$f := y^d + \sum_{i=0}^{d-1} f_i y^i \,,$$

$$g := y^{d_1} + \sum_{i=0}^{d_1-1} g_i y^i$$

and

$$h := y^{d_2} + \sum_{i=0}^{d_2-1} h_i y^i \,.$$

Now, comparing the coefficients of y^i on both sides in the equality $f = g \cdot h$ gives us a system of polynomial equations in $g_0, g_1, \ldots, g_{d_1-1}, h_0, h_1, \ldots, h_{d_2-1}$ as follows.

$$\begin{aligned}
f_0 &= g_0 \cdot h_0 \\
f_1 &= g_0 \cdot h_1 + g_1 \cdot h_0 \\
&\vdots \\
f_i &= \sum_{j=0}^{\min\{i,d_1\}} g_j \cdot h_{i-j} \\
&\vdots \\
f_{d-1} &= g_{d_1-1} + h_{d_2-1}.
\end{aligned}$$

Let $\mathbf{u} = (u_0, u_1, \ldots, u_{d_1-1})$ and $\mathbf{w} = (w_0, w_1, \ldots, w_{d_2-1})$ be new sets of variables. For $\ell \in [d]$ define the polynomials

$$Q_\ell(\mathbf{u}, \mathbf{w}) := \sum_{j=0}^{\min\{\ell, d_1-1\}} u_j \cdot w_{\ell-j} - f_\ell \,.$$

We view Q_ℓ as a polynomial in $\mathbf{u}, \mathbf{w}$ with coefficients coming from the ring $\mathbb{F}[\mathbf{x}]$. In this sense, $(\mathbf{g}, \mathbf{h}) = (g_0, g_1, \ldots, g_{d_1-1}, h_0, h_1, \ldots, h_{d_2-1})$ is a common zero of $Q_0, Q_1, \ldots, Q_{d-1}$. Our goal is to essentially *solve* the system of equations given by $\{Q_\ell(\mathbf{u}, \mathbf{v}) = 0 : \ell \in [d]\}$ to recover circuits for each g_i and h_i and prove an upper bound on their size. Note that this would not be an efficient algorithmic procedure, but we will be able to argue about the circuit complexity of the solution. To this end, we first observe some elementary properties of this system of polynomial equations.

Observation 3.1. *For every $\ell \in [d-1]$, Q_ℓ can be computed by a circuit of size at most $O(sd)$.*

Proof. Since f has a circuit of size at most s, and has degree d, each f_i can be computed by a circuit of size at $O(sd)$ by an easy application of Lemma 2.2. This immediately gives a circuit of this size for each Q_ℓ. □

Lemma 3.2. *Let $\mathcal{J}(\mathbf{u},\mathbf{v}) := \mathsf{Jacobian}_{\mathbf{u},\mathbf{w}}(Q_0, Q_1, \ldots, Q_{d-1})$ be the Jacobian of $Q_0, Q_1, \ldots, Q_{d-1}$. If $\mathsf{GCD}_y(g,h) = 1$, then $\mathcal{J}(\mathbf{g},\mathbf{h})$ is a non-singular matrix.*

Proof. The key observation here is that the Jacobian matrix (see Definition 2.5) is the same as the *Sylvester matrix* (see Definition 2.7) up to a permutation of rows and columns. Concretely, let R be the resultant of g and h when they are viewed as univariates in y. Since the $\mathsf{GCD}_y(f,g)$ is equal to 1, their resultant is a non-zero polynomial of degree at most $O(d^2)$ in $\mathbb{F}[\mathbf{x}]$. Recall that R is the determinant of the following $d \times d$ matrix, called the Sylvester matrix S. For $i \in \{1, 2, \ldots, d_2\}$, the i^{th} column of S is equal to $(0, \ldots, 0, 1, g_{d_1-1}, g_{d_1-2}, \ldots, g_1, g_0, 0, \ldots, 0)$, where there are $i-1$ zeroes in the prefix. For $j \in \{d_2+1, \ldots, d\}$, t he j^{th} column of S equals $(0, \ldots, 0, 1, h_{d_2-1}, h_{d_2-2}, \ldots, h_1, h_0, 0, \ldots, 0)$, where there are $j - d_2 - 1$ zeroes in the prefix. We now write the $d \times d$ matrix $\mathcal{J}(\mathbf{u},\mathbf{w})$.

$$\mathcal{J}(\mathbf{u},\mathbf{w}) = \begin{pmatrix} \frac{\partial Q_0}{\partial u_0} & \cdots & \frac{\partial Q_0}{\partial u_{d_1-1}} & \frac{\partial Q_0}{\partial w_0} & \cdots & \frac{\partial Q_0}{\partial w_{d_2-1}} \\ \frac{\partial Q_1}{\partial u_0} & \cdots & \frac{\partial Q_1}{\partial u_{d_1-1}} & \frac{\partial Q_1}{\partial w_1} & \cdots & \frac{\partial Q_1}{\partial u_{d_2}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial Q_{d-1}}{\partial u_0} & \cdots & \frac{\partial f_{d-1}}{\partial u_{d_1-1}} & \frac{\partial Q_{d-1}}{\partial w_0} & \cdots & \frac{\partial f_{d-1}}{\partial w_{d_2-1}} \end{pmatrix}$$

Plugging in the expressions for the partial derivatives, we get that for $i \in \{1, 2, \ldots, d_1\}$, the i^{th} column of $\mathcal{J}(\mathbf{u},\mathbf{w})$ is equal to $(0, \ldots, 0, w_0, w_1, w_2, \ldots, w_{d_2-1}, 1, 0, \ldots, 0)$, where there are $i-1$ zeroes in the prefix. For $j \in \{d_1+1, \ldots, d\}$, the j^{th} column of S equals $(0, \ldots, 0, u_0, u_1, u_2, \ldots, u_{d_1-1}, 1, 0, \ldots, 0)$, where there are $j - d_2 - 1$ zeroes in the prefix. Therefore, after substitution, the columns of $\mathcal{J}(\mathbf{g},\mathbf{h})$ are precisely the same as columns of S up to a permutation of rows and columns. In other words, their ranks are equal. We know that S is non-singular, so it follows that $\mathcal{J}(\mathbf{g},\mathbf{h})$ is also non-singular. □

Remark 3.3. *For the rest of the proof, we assume without loss of generality that $\mathcal{J}(\mathbf{g}(\mathbf{0}),\mathbf{h}(\mathbf{0}))$ is non-singular. This follows from the fact that since $\mathcal{J}(\mathbf{g}(\mathbf{x}),\mathbf{h}(\mathbf{x}))$ is non-singular, there is a $\mathbf{b}$ such that $\mathcal{J}(\mathbf{g}(\mathbf{b}),\mathbf{h}(\mathbf{b}))$ is non-singular, and up to a translation of the coordinate axes, we can assume that $\mathbf{b} = \mathbf{0}$.* ◊

3.1 Newton Iteration for many variables

We now show that given the constant term for each polynomial in $\mathbf{g},\mathbf{h}$, we can recover the polynomials completely. The argument is via a natural and well known multivariate analog of the standard Newton Iteration. Clearly, the constant terms have small circuits (trivial circuits of size 1), and we show that in this iterative process, we can recover multioutput circuits for $\mathbf{g},\mathbf{h}$ of size $\mathsf{poly}(s,n,d)$.

Lemma 3.4 (One step of Newton Iteration)**.** *Let $k \geq 1$ be any integer. Let C_k be a multioutput circuit of size most s_k computing polynomials $\tilde{\mathbf{g}}_k = (\tilde{g}_{0,k}, \tilde{g}_{1,k}, \ldots, \tilde{g}_{d_1-1,k})$, $\tilde{\mathbf{h}}_k = \left(\tilde{h}_{0,k}, \tilde{h}_{1,k}, \ldots, \tilde{h}_{d_2-1,k}\right)$ such that for every i and j,*

$$\tilde{g}_{i,k} \equiv g_i \mod \langle \mathbf{x} \rangle^k,$$

and

$$\tilde{h}_{j,k} \equiv h_j \mod \langle \mathbf{x} \rangle^k.$$

Then, there is a constant c independent of k such that the following is true: there is a multioutput circuit C_{k+1} of size at most $s_{k+1} = s_k + (snd)^c$ which computes the polynomials $\tilde{\mathbf{g}}_{k+1} =$

$(\tilde{g}_{0,k+1}, \tilde{g}_{1,k+1}, \ldots, \tilde{g}_{d_1-1,k+1})$, $\tilde{\mathbf{h}}_{k+1} = \left(\tilde{h}_{0,k+1}, \tilde{h}_{1,k+1}, \ldots, \tilde{h}_{d_2-1,k+1}\right)$ *such that for every* i *and* j,

$$\tilde{g}_{i,k+1} \equiv g_i \mod \langle \mathbf{x} \rangle^{k+1},$$

and

$$\tilde{h}_{j,k+1} \equiv h_j \mod \langle \mathbf{x} \rangle^{k+1}.$$

Proof. For every $i \in [d_1]$ and $j \in [d_2]$, let p_i and q_j be homogeneous polynomials of degree equal to k such that $\tilde{g}_{i,k} + p_i \equiv g_i \mod \langle \mathbf{x} \rangle^{k+1}$ and $\tilde{h}_{i,k} + q_j \equiv h_j \mod \langle \mathbf{x} \rangle^{k+1}$. Let $\mathbf{p}$ and $\mathbf{q}$ be the tuples associated to $p_i's$ and $q_j's$. Our goal is to show that these polynomials $\tilde{g}_{i,k} + p_i$ and $\tilde{h}_{j,k} + q_j$ have *small* circuits. This would complete the proof of the lemma. To this end, we set up a system of linear equations in the $p's$ and $q's$ and show that this system has a unique solution.

Let $\mathbf{g}_i$ (resp. $\mathbf{h}_i$) denote the tuple $(g_0 \mod \langle \mathbf{x} \rangle^i, g_1 \mod \langle \mathbf{x} \rangle^i, \ldots, g_{d_1-1} \mod \langle \mathbf{x} \rangle^i)$ (resp. $(h_0 \mod \langle \mathbf{x} \rangle^i, h_1 \mod \langle \mathbf{x} \rangle^i, \ldots, h_{d_2-1} \mod \langle \mathbf{x} \rangle^i)$). For each $\ell \in \{0, 1, \ldots, d-1\}$, since $Q_\ell(\mathbf{g}, \mathbf{h}) = 0$, it follows that

$$Q_\ell(\mathbf{g}_{k+1}, \mathbf{h}_{k+1}) \equiv 0 \mod \langle \mathbf{x} \rangle^{k+1}.$$

By their definition, and the hypothesis of the lemma, we know that $\mathbf{p}$ and $\mathbf{q}$ must satisfy

$$Q_\ell\left(\tilde{\mathbf{g}}_k + \mathbf{p}, \tilde{\mathbf{h}}_k + \mathbf{q}\right) \equiv 0 \mod \langle \mathbf{x} \rangle^{k+1}.$$

Since each p_i and q_j is a homogeneous polynomial of degree equal to k, via the multivariate Taylor expansion for polynomials (see Corollary 2.4) for Q_ℓ around the point $(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k)$, we get

$$Q_\ell(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) + \sum_{i \in [d_1]} \frac{\partial Q_\ell}{\partial u_i}(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \cdot p_i + \sum_{j \in [d_2]} \frac{\partial Q_\ell}{\partial w_j}(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \cdot q_j \equiv 0 \mod \langle \mathbf{x} \rangle^{k+1}.$$

Note that we used the fact that p_i and q_j are homogeneous polynomials of degree equal to k and hence their squares and higher powers vanish modulo $\langle \mathbf{x} \rangle^{k+1}$. Moreover, the only monomials of degree at most k in $\frac{\partial Q_\ell}{\partial u_i}(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \cdot p_i$ are those in $\left(\frac{\partial Q_\ell}{\partial u_i}(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \mod \langle \mathbf{x} \rangle\right) \cdot p_i$. We also know from the hypothesis that $\left(\frac{\partial Q_\ell}{\partial u_i}(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \mod \langle \mathbf{x} \rangle\right)$ is equal to $\left(\frac{\partial Q_\ell}{\partial u_i}(\mathbf{g}(\mathbf{0}), \mathbf{h}(\mathbf{0})\right)$. Applying these simplifications to the Taylor expansion for Q_ℓ, we get

$$-Q_\ell(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \equiv \sum_{i \in [d_1]} \left(\frac{\partial Q_\ell}{\partial u_i}(\mathbf{g}(\mathbf{0}), \mathbf{h}(\mathbf{0}))\right) \cdot p_i + \sum_{j \in [d_2]} \left(\frac{\partial Q_\ell}{\partial w_i}(\mathbf{g}(\mathbf{0}), \mathbf{h}(\mathbf{0}))\right) \cdot q_j \mod \langle \mathbf{x} \rangle^{k+1}.$$

Let $\mathbf{v} = (v_0, v_1, \ldots, v_{d_1-1})$ and $\mathbf{z} = (z_0, z_1, \ldots, z_{d_2-1})$ be new sets of variables. For various values of ℓ, let us consider the affine constraint on these variables given by the equation.

$$-Q_\ell(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \mod \langle \mathbf{x} \rangle^{k+1} = \sum_{i \in [d_1]} \left(\frac{\partial Q_\ell}{\partial u_i}(\mathbf{g}(\mathbf{0}), \mathbf{h}(\mathbf{0}))\right) \cdot v_i + \sum_{j \in [d_2]} \left(\frac{\partial Q_\ell}{\partial w_i}(\mathbf{g}(\mathbf{0}), \mathbf{h}(\mathbf{0}))\right) \cdot z_j.$$

So, we get a system of non-homogeneous linear equations of the form $A \cdot (\mathbf{v}, \mathbf{z})^T - \mathbf{a} = 0$, where the matrix A equals the matrix $\mathcal{J}(\mathbf{g}(\mathbf{0}), \mathbf{h}(\mathbf{0}))$, which by Lemma 3.2 is non-singular. Thus, this system has unique solution which is given by $\mathbf{z} = A^{-1}\mathbf{a}$. From our set up above, $(\mathbf{p}, \mathbf{q})$ is a solution to this system of equations, and thus by uniqueness of solution, we get that that there are field constants $\{\beta_{i,\ell} : i \in [d_1], \ell \in [d]\}$ and $\{\gamma_{j,\ell} : j \in [d_2], \ell \in [d]\}$ in $\mathbb{F}$ such that for every $i \in [d_1]$ and $j \in [d_2]$,

$$p_i = \sum_{\ell \in [d]} \beta_{i,\ell} \left(Q_\ell(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \mod \langle \mathbf{x} \rangle^{k+1}\right),$$

$$q_j = \sum_{\ell \in [d]} \gamma_{j,\ell} \left(Q_\ell(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \mod \langle \mathbf{x} \rangle^{k+1} \right) .$$

In other words,

$$p_i = \left(\sum_{\ell \in [d]} \beta_{i,\ell} \left(Q_\ell(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \right) \right) \mod \langle \mathbf{x} \rangle^{k+1} ,$$

and

$$q_j = \left(\sum_{\ell \in [d]} \gamma_{j,\ell} \left(Q_\ell(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \right) \right) \mod \langle \mathbf{x} \rangle^{k+1} ,$$

Now, recall that for every ℓ, $\mathbb{Q}_\ell(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k)$ is a polynomial which is zero modulo $\langle \mathbf{x} \rangle^k$. Thus, the polynomial $\tilde{g}_{i,k} + \left(\sum_\ell \beta_{i,\ell} \left(Q_\ell(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \right) \right)$ is equal to g_i modulo $\langle \mathbf{x} \rangle^{k+1}$, $\tilde{g}_{i,k}$ agrees with g_i at monomials of degree less than k, and we are adding to it the *correct* homogeneous polynomial of degree equal k. Thus, we define $\tilde{g}_{i,k+1}$ and $\tilde{h}_{j,k+1}$ as

$$\tilde{g}_{i,k+1} := \tilde{g}_{i,k} + \left(\sum_{\ell \in [d]} \beta_{i,\ell} \left(Q_\ell(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \right) \right) ,$$

and

$$\tilde{h}_{j,k+1} := \tilde{h}_{j,k} + \left(\sum_{\ell \in [d]} \gamma_{j,\ell} \left(Q_\ell(\tilde{\mathbf{g}}_k, \tilde{\mathbf{h}}_k) \right) \right) .$$

All that remains now is to argue that there is a small circuit computing $\tilde{\mathbf{g}}_{k+1}$ and $\tilde{\mathbf{h}}_{k+1}$. We can obtain a circuit for computing $\tilde{\mathbf{g}}_{k+1}$ and $\tilde{\mathbf{h}}_{k+1}$ from the circuit computing $\tilde{\mathbf{g}}_k$ and $\tilde{\mathbf{h}}_k$ by adding a copy of circuits for $Q_0, Q_1, \ldots, Q_{d-1}$ at the top and a layer of addition gates above it with appropriate edge weights. The size therefore increases additively by a fixed polynomial in s, n, d in each step. □

We now complete the proof of Lemma 1.2.

Proof of Lemma 1.2. Observe that modulo $\langle \mathbf{x} \rangle$, the g_i and h_j trivially have a circuit of size 1, since they are just constants. Now, using this as the base case, we use Lemma 3.4 $d+1$ for times to obtain a multioutput circuit C_{d+1} of size at most $\mathsf{poly}(s, n, d)$, which computes polynomials $\tilde{\mathbf{g}}_{d+1}$ and $\tilde{\mathbf{h}}_{d+1}$ such that for every i

$$\tilde{g}_{i,d+1} \equiv g_i \mod \langle \mathbf{x} \rangle^{d+1} .$$

But, each g_i is of degree at most d. Thus, we get

$$\tilde{g}_{i,d+1} \mod \langle \mathbf{x} \rangle^{d+1} = g_i .$$

So, we recover a circuit for each g_i and each h_j, we homogenize (see Lemma 2.2) the circuit C_{d+1} to get a homogeneous circuit C'_{d+1}, which has an output gate for each homogeneous component of degree at most d for every output of C_{d+1}. This incurs an additional multiplicative blow up of $O(d^2)$ on the size of C_{d+1}. From the circuit C'_{d+1}, we can just read off the polynomials g_i (resp. h_j) by taking an appropriate linear combinations of the outputs of C'_{d+1}, which only incurs an additive $\mathsf{poly}(s, n, d)$ blow up in the size of the circuit. This completes the proof of the lemma. □

4 Proof of Lemma 1.3

We have polynomials f and g such that $f = g^e$ where f has a circuit of size at most s. Since d be the degree of f, both e and the degree of g are upper bounded by d. The goal is showing that g has a circuit of size at most $\mathsf{poly}(s, n, d)$. Now, consider the following polynomial

$$\tilde{f} := z^e - f = z^e - g^e$$

where z is a new variable. Note that $\tilde{f}$ has a circuit of size at most $s + O(\log d)$. Next, decompose $\tilde{f}$ as follows.

$$\tilde{f} = (z - g) \cdot \left(z^{e-1} + z^{e-2}g + \cdots + g^{e-1}\right) .$$

Observe that if $g \not\equiv 0$ and the characteristic of the field is zero or large enough, then the GCD_z of the polynomial $(z - g)$ and the polynomial $\left(z^{e-1} + z^{e-2}g + \cdots + g^{e-1}\right)$ is 1. The reason is that $(z - g)$ is irreducible and does not divide $\left(z^{e-1} + z^{e-2}g + \cdots + g^{e-1}\right)$ when $g \not\equiv 0$ and the characteristic of the field is zero or large enough. Finally, by Lemma 1.2, $(z - g)$ has a circuit of size at most $\mathsf{poly}(s, n, d)$ and thus g also has a circuit of size $\mathsf{poly}(s, n, d)$.

Acknowledgements

We thank Vishwas Bhargav, Swastik Kopparty, Ramprasad Saptharishi and Srikanth Srinivasan for various insightful discussions and for encouraging us to write the proof up.

References

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

[CKS18a] Chi-Ning Chou, Mrinal Kumar, and Noam Solomon. Hardness vs Randomness for Bounded Depth Arithmetic Circuits. In *33rd Computational Complexity Conference (CCC 2018)*, volume 102, pages 13:1–13:17, 2018.

[CKS18b] Chi-Ning Chou, Mrinal Kumar, and Noam Solomon. Some Closure Results for Polynomial Factorization and Applications. *To appear in Theory of Computing. Preprint arXiv:1803.05933*, 2018.

[DSS17] Pranjal Dutta, Nitin Saxena, and Amit Sinhababu. Discovering the roots: Uniform closure results for algebraic classes under factoring. *CoRR*, abs/1710.03214, 2017.

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

[Kal86] Erich Kaltofen. Uniform Closure Properties of P-Computable Functions. In *Proceedings of the 18th Annual ACM Symposium on Theory of Computing, May 28-30, 1986, Berkeley, California, USA*, pages 330–337, 1986.

[Kal87] Erich Kaltofen. Single-Factor Hensel Lifting and its Application to the Straight-Line Complexity of Certain Polynomials. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*, pages 443–452, 1987.

[Kal89] Erich Kaltofen. Factorization of polynomials given by straight-line programs. *Randomness and Computation*, 5(375-412):2–3, 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)*.

[KT90] Erich Kaltofen and Barry M. Trager. Computing with Polynomials Given By Black Boxes for Their Evaluations: Greatest Common Divisors, Factorization, Separation of Numerators and Denominators. *J. Symb. Comput.*, 9(3):301–320, 1990.

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

[Sud98] Madhu Sudan. Algebra and computation - lecture notes, 1998.

[SY+10] Amir Shpilka, Amir Yehudayoff, et al. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends® in Theoretical Computer Science*, 5(3–4):207–388, 2010.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il