Electronic Colloquium on Computational Complexity, Report No. 297 (2026)

An Exponential Top Fan-In Hierarchy for Border Depth-4 Circuits

Anakin Dey * Pranjal Dutta † Zeyu Guo ‡

Abstract

We study depth-4 algebraic circuits with bounded top and bottom fan-in over fields of characteristic 0. These circuits compute sums of at most k products of polynomials of degree at most δ. We denote this class of circuits by $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$. Our hard polynomial for this class is $P_{k+1,n}$, the sum of $k+1$ monomials of degree n on disjoint sets of variables. It can be computed by a depth-3 circuit of top fan-in $k+1$ and size $O(kn)$.

We give an impossibility result showing that for all constants k and δ and all sufficiently large n, no $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit computes $P_{k+1,n}$. For $k \geq 2$ and $\delta = 2$, this follows implicitly from the rank bound for $\Sigma^{[k]}\Pi\Sigma\Pi^{[2]}$ identities of Garg, Oliveira, Sengupta, Shalmon, and Shpilka (CCC 2026). To the best of our knowledge, this impossibility result was not known even for $k \geq 2$ and $\delta = 3$.

We also turn to border complexity, a notion central to geometric complexity theory, in which a polynomial only needs to be approximated arbitrarily well. Here impossibility fails over algebraically closed fields: Kumar (TOCT 2020) showed that $\overline{\Sigma^{[2]}\Pi\Sigma}$ circuits already approximate every polynomial, albeit with exponential size. In turn, we show that every $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ circuit computing $P_{k+1,n}$ must have size $2^{\Omega(n)}$ when k and δ are constants, in stark contrast to the exact depth-3 circuit of top fan-in $k+1$ and size $O(kn)$. This lower bound gives an exponential top fan-in hierarchy for border depth-4 circuits of bounded bottom fan-in and answers the main open question of Dutta and Saxena (FOCS 2022, JACM 2026), who gave an analogous hierarchy for the case $\delta = 1$, i.e., depth 3.

Along the way, we develop tools for border depth-4 circuits that apply to arbitrary polynomials and may have further applications.

*Department of Mathematics, The Ohio State University dey.92@osu.edu; partially supported by NSF award DMS-2231565.

†College of Computing and Data Science, Nanyang Technological University Singapore pranjal.dutta@ntu.edu.sg; supported by the Nanyang Assistant Professorship (NAP) and an SUG grant (#025774-00001), both funded by NTU.

‡Department of Computer Science and Engineering, The Ohio State University zguotcs@gmail.com; supported by NSF CAREER award CCF-2440926.

ISSN 1433-8092

Contents

1 Introduction **3**
- 1.1 Our Results . . . 5
- 1.2 Top Fan-In Hierarchy for Depth 3 . . . 7
  - 1.2.1 Impossibility of Exact Computation in Depth 3 . . . 8
  - 1.2.2 Lower Bound in the Border Setting for Depth 3 . . . 8
- 1.3 Obstacles to Extending the Depth-3 Argument to Depth 4 . . . 10
  - 1.3.1 Exact Computation in Depth 4 . . . 10
  - 1.3.2 Border Computation in Depth 4 . . . 11
- 1.4 Proof Ideas . . . 12
  - 1.4.1 Homogeneous Circuits: A Warm-Up . . . 12
  - 1.4.2 Non-Homogeneous Circuits . . . 14
  - 1.4.3 Impossibility Result . . . 20
- 1.5 Multiplicative Joins and Lower Bounds in Geometric Complexity Theory . . . 22
- 1.6 Related Work . . . 24
- 1.7 Organization . . . 25

2 Preliminaries **26**
- 2.1 Commutative Algebra . . . 26
- 2.2 Algebraic Geometry . . . 29
- 2.3 Algebraic Complexity . . . 32
- 2.4 Auxiliary Statements . . . 33

3 Impossibility Under a Constant-Freeness Assumption **37**

4 Impossibility Under a Constant-Freeness Assumption in the Border Setting **38**
- 4.1 Additional Preliminaries . . . 38
- 4.2 Proof of the Impossibility Result . . . 44

5 Low-Weight Translation **46**

6 A Border Lower Bound Without the Constant-Freeness Assumption **48**
- 6.1 Auxiliary Statements for the Border Setting . . . 49
- 6.2 Border Low-Weight Translation . . . 52
- 6.3 Applying Modified DiDIL . . . 62
  - 6.3.1 Scaling . . . 64
  - 6.3.2 Limits . . . 65
  - 6.3.3 The Bloated Model . . . 66
  - 6.3.4 Derivatives and Coefficients . . . 68
  - 6.3.5 The DiDIL Induction . . . 71
  - 6.3.6 Stitching It All Together: Proof of the Lower Bound . . . 75

7 Impossibility Without the Constant-Freeness Assumption **76**
- 7.1 Preliminaries for the Impossibility Proof . . . 77
- 7.2 Normalization and Irreducible Components at Infinity . . . 79
- 7.3 A Rank Bound for the Group of Units Modulo Constants . . . 81

7.4 Proof of the Impossibility Result 82

8 Conclusion **88**

A The Abhyankar–Jung Theorem **89**
A.1 Freeness of the Integral Closure as a Module 92
A.2 Discriminants 94
A.3 Removing Ramification by Adjoining a Root 95
A.4 The Discriminant of the Integral Closure 97
A.5 Finishing the Proof 100

1 Introduction

A basic question in complexity theory is whether allowing more of a resource yields more computational power. The Time Hierarchy Theorem [HS65] and the Space Hierarchy Theorem [SHL65] answer this question affirmatively for time and space on Turing machines. For circuits, the natural resources to study are size, depth, and fan-in. In this paper, we study a hierarchy question for algebraic circuits of constant depth, taking the fan-in of the output gate as the resource. We consider both exact computation and approximation, and the two behave very differently: An impossibility result in the exact setting becomes a question about size in the approximative setting.

Algebraic complexity theory studies the complexity of computing polynomials, such as the determinant and the permanent, with additions and multiplications. Let $\mathbb{F}$ be a field. An *algebraic circuit* over $\mathbb{F}$ is a directed acyclic graph where the input nodes are labeled by variables or by constants from $\mathbb{F}$ and the internal gates are labeled by $+$ or $\times$. Each edge is labeled with a constant from $\mathbb{F}$ for scalar multiplication. The output gate computes a polynomial in $\mathbb{F}[\mathbf{x}] = \mathbb{F}[x_1, \ldots, x_N]$. The *size* of a circuit is the number of its gates and wires. The *depth* of a circuit is the length of a longest path from an input to the output gate. The *fan-in* of a gate is its number of inputs, which is unbounded unless a restriction is specified. Valiant's conjecture $\mathsf{VP} \neq \mathsf{VNP}$ [Val79] says that the permanent has no algebraic circuits of polynomial size over fields of characteristic other than 2; this is the algebraic analogue of the P versus NP question. Since an algebraic circuit computes a formal polynomial, lower bounds can be approached with tools from algebra and geometry; see the survey of Shpilka and Yehudayoff [SY10] and the survey of Saptharishi [Sap26].

A central goal is to understand the power of *constant-depth* circuits. Throughout, $\mathbb{F}$ is a field of characteristic 0 unless stated otherwise. Suppose that a polynomial of degree d has an algebraic circuit of size s, with $d \leq s^{O(1)}$. The depth reduction of Agrawal and Vinay [AV08], refined by Koiran [Koi12] and Tavenas [Tav15], gives a depth-4 circuit of size $s^{O(\sqrt{d})}$ with bottom fan-in $O\left(\sqrt{d}\right)$. Thus, strong enough lower bounds for depth-4 circuits would separate VP from VNP. After a long line of lower bounds for restricted constant-depth circuits [Fou+15; Gup+14; Kay+14; KS17b], Limaye, Srinivasan, and Tavenas [LST25] proved the first superpolynomial lower bounds for algebraic circuits of every constant depth. They also proved a *depth hierarchy theorem*: For every fixed $r \geq 2$, there are explicit polynomials computable by depth-r circuits of size s such that every depth-$(r-1)$ circuit computing them has size $s^{\omega(1)}$. Thus, increasing the depth by one can yield a superpolynomial saving in circuit size.

We ask whether a finer hierarchy holds when the depth is fixed: *Can increasing the fan-in of*

the output addition gate from k to $k+1$ yield a superpolynomial, or even exponential, saving in size? For instance, deterministic white-box polynomial identity testing is possible for depth-3 circuits with top fan-in k in time polynomial in the size, with an exponent that grows with k [KS07]. Thus, it is natural to ask whether the computational power of these circuits grows with k as well.

Depth-4 Circuits with Bounded Top and Bottom Fan-In. A $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit over $\mathbb{F}$ computes a polynomial of the form

$$f = \sum_{i=1}^{k} T_i, \qquad T_i = \prod_{j=1}^{m_i} q_{i,j}, \qquad \deg q_{i,j} \leq \delta,$$

where each $q_{i,j} \in \mathbb{F}[\mathbf{x}]$ is given as a sum of monomials of degree at most δ (Definition 2.11). We call the T_i the *product gates* and the $q_{i,j}$ the *bottom polynomials* of the circuit. The top fan-in is at most k, and the bottom fan-in is at most δ. The numbers m_i of factors in the product gates, and hence the degrees of the product gates, are unrestricted. For $\delta = 1$, the bottom polynomials are affine linear, and the model is that of depth-3 $\Sigma^{[k]}\Pi\Sigma$ circuits. When k and δ are constants, questions about deterministic identity testing for $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits are a frontier, as we discuss in Section 1.6. We ask whether there is a top fan-in hierarchy for this model.

Question 1.1 (The Top Fan-In Hierarchy Question): For every fixed $k, \delta \geq 1$, is there an explicit family $\{ P_n \}_{n \geq 1}$ with $P_n \in \mathbb{F}[x_1, \ldots, x_n]$ of degree $\mathrm{poly}(n)$ such that each P_n is computable by $\Sigma^{[k+1]}\Pi\Sigma\Pi^{[\delta]}$ circuits of size $\mathrm{poly}(n)$, but every $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit computing P_n has size $n^{\omega(1)}$?

A natural candidate is the following polynomial, which is also the candidate polynomial for border depth-3 circuits in [DS26].

Definition 1.2: For $k \geq 0$ and $n \geq 1$, let $N := (k+1)n$. The *sum-product* polynomial [Lan17, §8.12.2] is defined to be

$$P_{k+1,n} := \sum_{j=1}^{k+1} \prod_{t=(j-1)n+1}^{jn} x_t \in \mathbb{F}[x_1, \ldots, x_N].$$

This is the sum of $k+1$ monomials of degree n in disjoint blocks of variables. For each $j \in [k+1]$, let

$$B_j := \{ (j-1)n+1, \ldots, jn \}.$$

We call B_j the j-th *block*. We also write $\mathbf{x}_j := \mathbf{x}_{B_j} = (x_{(j-1)n+1}, \ldots, x_{jn})$ and $\mathbf{x}^{B_j} := \prod_{t=(j-1)n+1}^{jn} x_t$. By a slight abuse of terminology, we also refer to the tuple $\mathbf{x}_j$ or the monomial $\mathbf{x}^{B_j}$ as the j-th block when the meaning is clear. We use this notation throughout the paper.

For $n = 2$, the polynomial $P_{k+1,2}$ is the *inner product polynomial* $\mathrm{IP}_{k+1}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{k+1} x_i y_i$, up to renaming the variables. For every $\delta \geq 1$, the polynomial $P_{k+1,n}$ has a $\Sigma^{[k+1]}\Pi\Sigma$ circuit, and hence a $\Sigma^{[k+1]}\Pi\Sigma\Pi^{[\delta]}$ circuit, of size $O(kn)$. The question is whether reducing the top fan-in to k, with δ fixed, requires superpolynomial or even exponential size in n. Since $P_{k+1,n}$ itself has a $\Sigma^{[k+1]}\Pi\Sigma$ circuit of size $O(kn)$, a lower bound for $P_{k+1,n}$ against top fan-in k has to exploit the difference between k and $k+1$ summands.

Border Complexity. In border (or approximative) complexity, a polynomial only needs to be approximated arbitrarily well. Formally, $f \in \mathbb{F}[\mathbf{x}]$ is *border computed* by a circuit if the circuit has constants in the field $\mathbb{F}((\varepsilon))$ of formal Laurent series and computes $f + \varepsilon S$ for some $S \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ (Definition 2.12). The constants occurring in the computation may have poles at $\varepsilon = 0$, but the computed polynomial tends to f as $\varepsilon \to 0$. We write $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ for the class of polynomials that are border computed by $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits.

The idea of studying efficient border computation of polynomials dates back to classical work in algebraic geometry on Waring rank by Palatini [Pal06] and Terracini [Ter11]. Border computation was first used in theoretical computer science to design fast matrix multiplication algorithms [Bin+79; Bin80]. Border computation is also central to the program of *geometric complexity theory* of Mulmuley and Sohoni [MS01; MS08], which formulates lower bounds as non-membership in group orbit closures. Grochow, Mulmuley, and Qiao [GMQ16] showed that lower bounds proved via polynomial equations, such as ones obtained from rank-based measures, automatically hold in the border setting. Whether VP equals its border closure $\overline{\mathsf{VP}}$ is a major open problem in algebraic complexity [Bür04; GMQ16]. It is clear that $\mathsf{VP} \subseteq \overline{\mathsf{VP}}$, and determining if $\overline{\mathsf{VP}} \subseteq \mathsf{VP}$ amounts to seeing if one can efficiently turn a border computation in $\overline{\mathsf{VP}}$ into an exact one in VP. The conversion of border computations into exact ones, called *debordering*, is studied for many restricted models; see the survey of Dutta and Lysikov [DL25].

Depth-2 circuits are closed under limits, meaning $\Sigma\Pi = \overline{\Sigma\Pi}$ and $\Pi\Sigma = \overline{\Pi\Sigma}$. Thus, border depth-3 circuits are the first interesting case of debordering, and they already behave very differently from exact depth-3 circuits. Kumar [Kum20] showed that, over algebraically closed fields of characteristic 0, every polynomial can be border computed by a $\Sigma^{[2]}\Pi\Sigma$ circuit, typically of exponential size. In contrast, no $\Sigma^{[2]}\Pi\Sigma$ circuit computes even IP_3 [Che+18; Kum20]. In particular, for every $k \geq 2$, the polynomial $P_{k+1,n}$ has border $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits. We ask two versions of Question 1.1 for our candidate family: whether there exists a $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit for $P_{k+1,n}$ and whether there are polynomial-size $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ circuits for $P_{k+1,n}$.

1.1 Our Results

We establish the desired top fan-in hierarchy for all constants k and δ, with $P_{k+1,n}$ as the separating polynomial. In the exact setting, we show an *impossibility*: No $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit can compute $P_{k+1,n}$ when n is sufficiently large (Theorem 1.3). In the border setting, every $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ circuit computing $P_{k+1,n}$ must have size $2^{\Omega(n)}$ (Theorem 1.4).

Let us state our results more precisely. Our results hold over any field of characteristic 0. We state a slightly weaker threshold than the one in Theorem 7.5 for simplicity.

Theorem 1.3 (Impossibility, see Theorem 7.5): For every fixed $k, \delta \geq 1$, if $n > 2(k+\delta)\delta^{2k}$, then $P_{k+1,n}$ cannot be computed by any $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit, regardless of its size.

We prove the theorem above in Section 7. Thus, depth-4 circuits with top fan-in $k+1$ and bottom fan-in 1 are already more powerful than depth-4 circuits with top fan-in k and any constant bottom fan-in. This is because $P_{k+1,n}$ can be computed by $\Sigma^{[k+1]}\Pi\Sigma$ circuits of size $O(kn)$, but no $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit computes $P_{k+1,n}$. For fixed δ, this threshold is at most singly exponential in k.

We remark that even showing the result for homogeneous circuits is nontrivial, and we prove it as a first step in Theorem 3.1.

Relation to Rank Bounds. For $\delta = 2$, the qualitative impossibility statement in Theorem 1.3 follows from a recent rank bound for $\Sigma^{[k]}\Pi\Sigma\Pi^{[2]}$ identities by Garg et al. [Gar+26]. We now explain this connection and why it does not extend to $\delta \geq 3$. Rank bounds are a central tool in identity testing for depth-4 circuits. Here, the rank of an identity is the *linear rank* of its bottom polynomials, that is, the dimension of their $\mathbb{F}$-span. The work of Garg et al. [Gar+26] uses linear rank, whereas the work of Beecken, Mittmann, and Saxena [BMS13] uses transcendence degree. The argument below works with either notion. An identity is *simple* if its product gates have no common nonconstant factor and *minimal* if no proper nonempty sub-sum vanishes. The rank-bound approach to identity testing for $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits [BMS13; Gup14] is based on the conjecture that every simple and minimal identity of this form has rank bounded by a function $r(k, \delta)$, independent of the number of variables.

Such a bound would rule out $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits for $P_{k+1,n}$ whenever $n > r(2k+1, \delta)/2$. Indeed, suppose towards contradiction that $P_{k+1,n}$ is computed by a $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit and write $P_{k+1,n} = \sum_{i=1}^{k} T_i$. Split the identity $\sum_{j=1}^{k+1} \mathbf{x}^{B_j} - \sum_{i=1}^{k} T_i = 0$ into minimal vanishing sub-sums. Block monomials alone cannot cancel, and there are $k+1$ of them but only k product gates, so one of these sub-sums contains at least two block monomials, say $\mathbf{x}^{B_{j_1}}$ and $\mathbf{x}^{B_{j_2}}$. This sub-sum is a minimal identity of top fan-in at most $2k+1$, and it is simple since $\mathbf{x}^{B_{j_1}}$ and $\mathbf{x}^{B_{j_2}}$ have no common factor. The bottom polynomials of this sub-sum include the $2n$ variables of two blocks, so its rank is at least $2n$, and the conjectured bound gives $2n \leq r(2k+1, \delta)$. For $\delta = 2$, such a rank bound was recently proved by Garg et al. [Gar+26], which gives the qualitative statement above. For $\delta \geq 3$, Garg, Oliveira, and Sengupta [GOS25] gave rank bounds for top fan-in 3. Their bounds cover the identity above only when $k = 1$. Hence, current rank bounds do not give Theorem 1.3 in general. To the best of our knowledge, for $k \geq 2$ and $\delta \geq 3$, it was not even known whether $P_{k+1,n}$ has $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits of any size.

As mentioned before in Border Complexity, the impossibility is no longer valid if we allow approximations, by a result of Kumar [Kum20]. Kumar's proof shows that for constant $k \geq 2$ and $\delta \geq 1$, the polynomial $P_{k+1,n}$ can be computed over algebraically closed fields by a $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ circuit of size $2^{O(n)}$. The next theorem shows that this is the best one could hope for. We again state a slightly weaker threshold than the one in Theorem 6.1 for simplicity.

Theorem 1.4 (Border Lower Bound, see Theorem 6.1): Let $\mathbb{F}$ be a field of characteristic 0. There is an absolute constant $c \geq 1$ such that, for every fixed $k, \delta \geq 1$ and all $n \geq (2\delta)^{ck^2}$, every $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit that border computes $P_{k+1,n}$ has size at least $2^{\Omega(n/(k(4\delta)^k))}$.

We prove the theorem above in Section 6. For constant k and δ, the theorem gives a lower bound of $2^{\Omega(n)}$ on s and hence the following hierarchy.

Corollary 1.5 (Border Depth-4 Top Fan-In Hierarchy): Let $\mathbb{F}$ be a field of characteristic 0. For every fixed $k, \delta \geq 1$, the polynomial $P_{k+1,n}$ has a $\Sigma^{[k+1]}\Pi\Sigma$ circuit of size $O(n)$, but every $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ circuit computing $P_{k+1,n}$ must have size at least $2^{\Omega(n)}$.

For $\delta = 1$, this is the hierarchy result of Dutta and Saxena [DS26], and Corollary 1.5 answers their main open question of whether this hierarchy extends to every constant bottom fan-in. Table 1 summarizes the picture.

	$\delta = 1$ (depth 3)	$\delta \geq 2$ (depth 4)
Exact	impossible for $n \geq 2$ ([Che+18; Kum20])	impossible for $n > 2(k+\delta)\delta^{2k}$ (Theorem 1.3)
Border	size $2^{\Omega(n)}$ [DS26]	size $2^{\Omega(n)}$ (Theorem 1.4)

Table 1: Computing $P_{k+1,n}$ with top fan-in k and bottom fan-in δ, for constants $k \geq 2$ and $\delta \geq 1$.

We also remark that the impossibility does hold in the border setting for circuits in which every nonconstant product gate has a factor vanishing at the origin before specializing at $\varepsilon = 0$. One such example is the class of homogeneous circuits. Such circuits cannot compute $P_{k+1,n}$ at all once $n > k\delta^k$, even in the border setting (Theorem 4.16). Thus, as in [DS26], the difficulty of Theorem 1.4 lies in the non-homogeneous case.

Other Hard Polynomials. Our results are not specific to $P_{k+1,n}$. Replacing each variable by a variable or by a constant from $\mathbb{F}$ does not increase the top fan-in, the bottom fan-in, or the size of a $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit, and the same holds for border circuits. Thus Theorems 1.3 and 1.4 hold for every polynomial that becomes $P_{k+1,n}$ after such a substitution. A natural example, already noted by Dutta and Saxena [DS26], is the *Trace-Iterated-Matrix-Multiplication* polynomial $\text{Tr-IMM}_{k+1,n}$, the trace of the product of n symbolic $(k+1) \times (k+1)$ matrices $X^{(1)}, \ldots, X^{(n)}$ whose entries are distinct variables. This polynomial is studied in geometric complexity theory; see, for example, Gesmundo [Ges16]. Replacing each $X^{(t)}$ by the diagonal matrix with diagonal entries $x_t, x_{n+t}, \ldots, x_{kn+t}$ turns $\text{Tr-IMM}_{k+1,n}$ into $P_{k+1,n}$. When the matrix dimension and the number of factors grow, this family is complete for VBP, the class of polynomials computed by polynomial-size algebraic branching programs.

At the other extreme, $P_{k+1,n}$ is a sum of only $k+1$ monomials, so it has a depth-2 circuit, and hence a formula, of size $O(kn)$. Thus, by Theorem 1.4, a polynomial with a formula of size $O(kn)$—even a sum of $k+1$ monomials—can require $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ circuits of size $2^{\Omega(n)}$. In particular, for all constants k and δ, there are polynomials in VF that require border $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits of exponential size. Superpolynomial lower bounds against this model also follow from the lower bounds of Limaye, Srinivasan, and Tavenas [LST25] for constant-depth circuits, which extend to the border setting as noted by Andrews and Forbes [AF22]. The lower bounds obtained by this route are superpolynomial but far from exponential. Exponential lower bounds for border circuits of bounded top fan-in are much harder to obtain, and here we obtain one for a sum of only $k+1$ monomials.

1.2 Top Fan-In Hierarchy for Depth 3

For $\delta = 1$, Dutta and Saxena [DS26] answered both versions of Question 1.1. Depth-3 $\Sigma^{[k]}\Pi\Sigma$ circuits of top fan-in k cannot exactly compute $P_{k+1,n}$ for any $n \geq 2$. For constant k, border computation requires size $2^{\Omega(n)}$. Both proofs start by reducing modulo linear factors of the product gates. We sketch these proofs, since our results follow the same outline and the difficulties

for $\delta \geq 2$ appear exactly where this outline breaks down. We discuss the proof ideas in detail in Section 1.4.

1.2.1 Impossibility of Exact Computation in Depth 3

It is known that IP_{k+1} cannot be computed by any $\Sigma^{[k]}\Pi\Sigma$ circuit, regardless of its size [Che+18; Kum20]. Thus, setting all but two variables of each block to 1 shows that the same holds for $P_{k+1,n}$ for every $n \geq 2$. For $n > k$, there is a particularly simple direct argument, which we sketch because our proofs follow it. Suppose towards contradiction that $P_{k+1,n}$ is computed by a $\Sigma^{[k]}\Pi\Sigma$ circuit and write

$$P_{k+1,n} = T_1 + \cdots + T_k,$$

where each T_i is a product of affine linear polynomials.

We *greedily* select at most one factor from each product gate and reduce modulo the selected factors. Start with the whole space $V := \mathbb{A}^N$, where $\mathbb{A}^N = \mathbb{F}^N$ is the affine space of points at which the polynomials in $x_1, \ldots, x_N$ are evaluated. For each $i \in [k]$, if some factor of T_i has a zero on V, then replace V by the zero set of this factor in V; otherwise, leave V unchanged. In the end, V is a nonempty affine subspace of codimension at most k. Every gate T_i is constant on V. Indeed, if one of its factors vanishes on V, then T_i vanishes on V. Otherwise, each factor is an affine linear polynomial without zeros on V and hence a nonzero constant on V, so T_i computes a nonzero constant on V. Thus the right side is *constant* on V, and it remains to see that the left side is not.

The subspace V is the solution set of at most k linear equations, which we can solve for at most k of the variables. Some block, say B_1, contains none of these variables, and since $n > k$, each of the other blocks contains a variable that we did not solve for. Setting one such variable from each of the other blocks to zero, we obtain an affine subspace of V on which $P_{k+1,n}$ equals the nonconstant block monomial $\mathbf{x}^{B_1}$ in free variables.

1.2.2 Lower Bound in the Border Setting for Depth 3

In the border setting, the argument of Section 1.2.1 breaks down, and so does its conclusion. Kumar's construction [Kum20] shows why, as explained below.

Let f be a homogeneous polynomial of positive degree d over an algebraically closed field of characteristic 0. We may write $f = \sum_{i=1}^{M} \ell_i^d$ for linear forms ℓ_i. Let ω be a primitive d-th root of unity. Since $\prod_{j=1}^{d}(1 - \omega^j t) = 1 - t^d$, we have

$$\varepsilon^{-d} - \varepsilon^{-d} \prod_{i=1}^{M} \prod_{j=1}^{d} \left(1 - \omega^j \varepsilon \ell_i\right) \;=\; \varepsilon^{-d}\left(1 - \prod_{i=1}^{M}\left(1 - \varepsilon^d \ell_i^d\right)\right) \;=\; f + \varepsilon^d S, \tag{1}$$

for some $S \in \mathbb{F}[\varepsilon][\mathbf{x}]$. Thus the expression on the left side of (1) gives a $\overline{\Sigma^{[2]}\Pi\Sigma}$ circuit for f. Every factor $1 - \omega^j \varepsilon \ell_i$ tends to 1 as $\varepsilon \to 0$, and its zero set, the hyperplane $\ell_i = \omega^{-j}\varepsilon^{-1}$, goes to infinity. Therefore, reducing modulo such a factor tells us *nothing* in the limit! The expression in Equation (1) also plays crucial roles in illustrating the subtleties of the proof ideas in Section 1.4.

Reduction. Suppose that we have a $\overline{\Sigma^{[k]}\Pi\Sigma}$ circuit for $P_{k+1,n}$. We can therefore write

$$P_{k+1,n} + \varepsilon S = T_1 + \cdots + T_k,$$

where each $T_i \in \mathbb{F}((\varepsilon))[\mathbf{x}]$ and $S \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$. Dutta and Saxena [DS26] showed that reduction modulo linear forms still helps, provided that one uses the "right" factors. Discard zero product gates and write each nonzero factor as a power of ε times a linear polynomial that has coefficients in $\mathbb{F}[[\varepsilon]]$ and a nonzero specialization at $\varepsilon = 0$, called its *limit*. If the limit is a homogeneous linear form, then solving the factor for a suitable variable is an affine substitution with coefficients in $\mathbb{F}[[\varepsilon]]$ whose limit is linear, and this substitution commutes with taking the limit. Reducing modulo such factors, at most one from each product gate, turns the left side into $P_{k+1,n}(\ell_1, \ldots, \ell_N)$ in the limit. Here, the ℓ_i are linear forms of rank at least $N - k$. For $n > k$, this polynomial is nonzero by the argument of Section 1.2.1. In particular, for $n > k$, homogeneous $\overline{\Sigma^{[k]}\Pi\Sigma}$ circuits cannot compute $P_{k+1,n}$ at all. The argument above also applies when every product gate has at least one factor that is a homogeneous linear form.

The hardest case is when the limit of every factor of every product gate has a nonzero constant term, so none of these limits is a homogeneous linear form. In this setting, some product gates survive reduction modulo linear forms. Every factor of a surviving gate has a limit with a nonzero constant term, so the gate is a power of ε times an invertible power series in ε and $\mathbf{x}$. This *all-non-homogeneous* case, which contains Kumar's circuits, is where the difficulty lies.

DiDIL in the All-Non-Homogeneous Setting. To handle the all-non-homogeneous case, Dutta and Saxena [DS26] use the *Divide–Derive–Induct–Limit* (DiDIL) technique. Dutta, Dwivedi, and Saxena [DDS22] introduced this technique to study border depth-3 circuits. DiDIL removes the product gates one at a time. Thus, the number of summands decreases by 1 at each step, while the polynomial $f := P_{k+1,n}(\ell_1, \ldots, \ell_N)$ survives every step; the trade-off is that the remaining summands become (somewhat) more complicated.

To see how one gate is removed, write the identity as $T_1 + \cdots + T_{k'} = f + \varepsilon S$ with $k' \le k$, and write $T_{k'} = \varepsilon^a \widetilde{T}$, where $\widetilde{T}$ has a nonzero limit. Dividing by $\widetilde{T}$ gives

$$\frac{T_1}{\widetilde{T}} + \cdots + \frac{T_{k'-1}}{\widetilde{T}} + \varepsilon^a \;=\; \frac{f + \varepsilon S}{\widetilde{T}}, \tag{2}$$

and differentiating removes the last summand ε^a, since it does not depend on the variables. To make sure that f is not removed as well, Dutta and Saxena [DS26] first apply the scaling $\Phi\colon x_i \mapsto z x_i$ for a new variable z and then differentiate with respect to z. Since f is homogeneous of degree n, we have $\Phi(f) = z^n f$, whereas $\Phi(\widetilde{T})$ is a power series in z whose constant term is nonzero in the limit, because every factor of $\widetilde{T}$ has a limit with a nonzero constant term. Thus, in the limit $\varepsilon \to 0$, the right side starts with z^n times a nonzero multiple of f, and its derivative starts with z^{n-1} times a nonzero multiple of f. An *elimination step* consists of dividing by the normalization of one summand and applying a differential operator that removes that summand while keeping the limit of the whole sum nonzero.

On the left side, the i-th summand becomes $\Phi(T_i)/\Phi(\widetilde{T})$ after scaling. Up to a power of ε, this is a ratio of two products of scaled linear polynomials $\Phi(\ell)$. For nonzero h, define

the *logarithmic derivative* $\operatorname{dlog}_z(h) := (\partial_z h)/h$, so $\partial_z h = h \cdot \operatorname{dlog}_z(h)$. By the product rule, the logarithmic derivative of $\Phi(T_i)/\Phi(\widetilde{T})$ is a sum of terms $\pm \operatorname{dlog}_z \Phi(\ell)$, one for each linear factor ℓ of T_i or $\widetilde{T}$; see Differentiation for more properties of dlog.

Write $\ell = c + \tilde{\ell}$, where $c := \ell(\mathbf{0})$ and $\tilde{\ell}$ is a linear form. The constant c has a nonzero limit. Thus $\Phi(\ell) = c + z\tilde{\ell}$, and the geometric series expansion gives

$$\operatorname{dlog}_z \Phi(\ell) \;=\; \frac{\tilde{\ell}}{c + z\tilde{\ell}} \;=\; \sum_{i \geq 0} (-1)^i c^{-i-1} z^i \, \tilde{\ell}^{\,i+1}. \tag{3}$$

Thus the coefficient of every power of z in the sum of these logarithmic derivatives is a sum of at most s powers of linear forms, where s is the size of the circuit. After at most $k-1$ elimination steps, comparing the coefficients of the lowest powers of z in the limit expresses f as a *ratio* of two polynomials, each the limit of a small sum of powers of linear forms.

The last ingredient is the dimension of the space spanned by all partial derivatives of a polynomial, a classical measure of its complexity. All partial derivatives of a power ℓ^d of a linear form are multiples of the powers ℓ^e with $e \leq d$. Thus the partial derivatives of a sum of m such powers span a space of dimension at most $m(d+1)$. This bound carries over to the border setting because the dimension is lower semicontinuous: it is the rank of a matrix, and that rank can only decrease in the limit.

To handle the ratio, one looks at a single monomial [For15]. If the partial derivatives of g span a space of dimension D, then the trailing monomial $\textsc{tm}(g)$, the smallest monomial of g in a fixed monomial order, involves at most $\log D$ variables (Theorem 6.6). Indeed, differentiating g with respect to subsets of the variables of $\textsc{tm}(g)$ gives linearly independent derivatives. Since trailing monomials are multiplicative, writing the ratio as $f \cdot G_2 = G_1$ shows that the trailing monomial of f involves at most $\log(m(d+1))$ variables. However, recall that the linear forms ℓ_i have rank at least $N-k$ (see Reduction), so after a linear change of coordinates, all but at most k of them are variables. Thus every block contains at least $n-k$ of these variables, so every monomial of f involves at least $n-k$ variables. Since the intermediate expressions grow in every DiDIL step, m can be as large as $s^{\exp(k)}$, and comparing the two bounds gives $s \geq 2^{\Omega(n/\exp(k))}$ [DS26].

1.3 Obstacles to Extending the Depth-3 Argument to Depth 4

The polynomial IP_{k+1} equals $P_{k+1,2}$ and is hard for the class $\Sigma^{[k]}\Pi\Sigma$. It is no longer hard for $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ when $\delta \geq 2$. Indeed, IP_{k+1} has degree 2, so it can be computed by a single $\Sigma\Pi^{[\delta]}$ circuit. More generally, $P_{k+1,n}$ can be computed by a $\Sigma\Pi^{[\delta]}$ circuit whenever $n \leq \delta$, so any separation hierarchy with this family needs $n > \delta$. More importantly, the depth-3 ideas of Section 1.2 break down; studying how these ideas break motivates our proof (see Section 1.4).

1.3.1 Exact Computation in Depth 4

To prove that no $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit computes $P_{k+1,n}$, let us try the greedy procedure from the depth-3 argument (Section 1.2.1) and see what changes. Let δ be at least 2. Suppose towards

contradiction that $P_{k+1,n}$ is computed by a $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit and write

$$P_{k+1,n} = T_1 + \cdots + T_k,$$

where each T_i is a product of polynomials of degree at most δ.

Suppose that we proceed with the greedy procedure of Section 1.2.1. Unlike reduction modulo linear factors, reduction modulo higher-degree factors need not correspond to a substitution. For a linear factor, such as $x_1 + x_{n+1} - 1$, we can solve for one variable and substitute. Indeed, quotienting $\mathbb{F}[\mathbf{x}]$ by $x_1 + x_{n+1} - 1$ imposes the relation $x_1 + x_{n+1} - 1 = 0$, so $x_1 = 1 - x_{n+1}$. By substitution, the quotient ring is again a polynomial ring in the remaining $N - 1$ variables. For a factor of degree δ with $\delta \geq 2$, obtaining a simple substitution map is impossible in general. For example, if we quotient $\mathbb{F}[\mathbf{x}]$ by the quadratic factor $x_1 x_{n+1} - 1$, then we cannot express x_1 as a polynomial in x_{n+1} and vice versa. Furthermore, the coordinate ring of a variety defined by a higher-degree relation need not be an integral domain, much less a polynomial ring. The zero sets of the chosen factors may be reducible, so we choose an irreducible component at each step of the greedy procedure. The resulting variety V has codimension at most k and degree at most δ^k. We work in the coordinate ring $\mathbb{F}[V]$, defined by $\mathbb{F}[V] := \mathbb{F}[\mathbf{x}]/\mathbf{I}(V)$, where $\mathbf{I}(V)$ is the ideal of polynomials vanishing on V, and study the equation above in this ring.

When we work modulo $\mathbf{I}(V)$, it is less clear that $P_{k+1,n}$ is not identically zero on V. At depth 3, we could set one variable from each block other than B_1 to zero, turning $P_{k+1,n}$ into the block monomial $\mathbf{x}^{B_1}$ while leaving the variables of B_1 untouched. On a more general variety such as V, setting one coordinate to zero can force other coordinates to vanish as well, including coordinates of B_1. For example, on the quadric $x_2 = x_1 x_{n+1}$, setting $x_{n+1} = 0$ forces $x_2 = 0$, in turn forcing $\mathbf{x}^{B_1} = 0$.

Moreover, the right side need not become constant. A gate whose factors have no zeros on V restricts to a product of units of $\mathbb{F}[V]$. Unlike in the depth-3 case, such units need not be constant. For example, on the quadric $\mathbf{V}(x_1 x_{n+1} - 1)$, the coordinate x_1 is a nonconstant unit with inverse x_{n+1}. After restricting to V, we must rule out an expression for $P_{k+1,n}$ as a sum of at most k products of units in $\mathbb{F}[V]$. This difficulty arises only for non-homogeneous circuits, whereas the first two difficulties also arise in the homogeneous setting.

1.3.2 Border Computation in Depth 4

Let δ be at least 2. Suppose that we have a $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ circuit for $P_{k+1,n}$. We can therefore write

$$P_{k+1,n} + \varepsilon S = T_1 + \cdots + T_k,$$

where each $T_i \in \mathbb{F}((\varepsilon))[\mathbf{x}]$ and $S \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$. In the border setting, the difficulties of Section 1.3.1 remain, and new ones appear. By Kumar's construction (1), $P_{k+1,n}$ has $\overline{\Sigma^{[2]}\Pi\Sigma}$ circuits over algebraically closed fields, so we can only hope for a size lower bound rather than impossibility.

We can try to follow the approach of Dutta and Saxena [DS26]. First, we try to reduce modulo suitably chosen bottom polynomials, so that the surviving gates become units near the origin, up to powers of ε. Next, we choose a maximal set of coordinates that are algebraically independent on the variety (in the limit); we call these coordinates *free*. Finally, we try to write the remaining

coordinates as power series in ε and the free coordinates around some point; substituting these series into the circuit would give the identity of power series that DiDIL needs. However, one soon realizes that two things could go wrong:

1. First, the variety may *ramify* as $\varepsilon \to 0$, meaning that several solutions for a coordinate may come together in the limit. For example, near the origin, the solutions of $x_1^2 = \varepsilon(1 + x_2)$ for x_1 are $\pm\varepsilon^{1/2}(1 + x_2)^{1/2}$. These solutions are power series in $\varepsilon^{1/2}$ and x_2 but not in ε and x_2. To see ramification, let t be a new indeterminate. Writing $t = x_1/\sqrt{1 + x_2}$ gives $\varepsilon = t^2$, so, with x_2 fixed, the projection to ε is a squaring map. As $\varepsilon \to 0$, the two points corresponding to t and $-t$ merge at $t = 0$, where the projection is ramified. The fractional power of ε arises when we try to invert this map.

2. Second, writing the remaining coordinates as power series in the free ones uses the implicit function theorem (Lemma 4.14), which needs certain partial derivatives to be nonzero at the point we expand around. This nonvanishing condition may fail at the origin. On the quadric $x_N^2 = x_1$, where x_1 is free, the derivative $\partial_{x_N}(x_N^2 - x_1)$ equals $2x_N$ and vanishes at the origin. Neither of the two roots $\pm x_1^{1/2}$ is a power series in x_1. Choosing a better point than the origin to expand around turns out to be subtle; see Section 1.4.

Even once both problems are resolved and we have the power-series identity, DiDIL cannot be applied verbatim: the factors of the surviving gates are now power series, known only through the equations of the variety. We need to modify DiDIL accordingly and apply it.

1.4 Proof Ideas

It may help to keep the smallest interesting case $k = 2$ in mind. We also discuss general k and explicitly indicate when we do so. For $k = 2$, let $N := (k+1)n = 3n$ denote the number of variables. The blocks of Definition 1.2 are $B_1 = \{1, \ldots, n\}$, $B_2 = \{n+1, \ldots, 2n\}$, and $B_3 = \{2n+1, \ldots, 3n\}$. Thus,

$$P_{3,n} = \mathbf{x}^{B_1} + \mathbf{x}^{B_2} + \mathbf{x}^{B_3} = x_1 \cdots x_n + x_{n+1} \cdots x_{2n} + x_{2n+1} \cdots x_{3n}.$$

Fix a constant δ with $\delta \geq 1$. We have to show that, for sufficiently large n, the polynomial $P_{3,n}$ cannot be written as a sum of two products of polynomials of degree at most δ and that approximating it by such sums requires circuits of size exponential in n.

1.4.1 Homogeneous Circuits: A Warm-Up

We first illustrate our techniques in the simpler setting of homogeneous circuits. We saw in Section 1.2.2 that the hard work is in the all-non-homogeneous setting. In Sections 3 and 4, we treat the setting where each top-level product gate of a $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit computes either a constant or a polynomial divisible by a constant-free factor of degree at most δ. This setting includes homogeneous $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits. We give an overview of the proofs in these settings.

Homogeneous Circuits in the Exact Setting: A Warm-Up. Suppose that $P_{3,n} = T_1 + T_2$, where T_1 and T_2 are nonconstant products of homogeneous polynomials of degree at most δ. Every nonconstant homogeneous polynomial vanishes at the origin, so T_1 and T_2 have factors q_1 and q_2 that vanish at the origin. Reduction modulo q_1 and q_2 now means restricting to their common zero

set $\mathbf{V}(q_1, q_2) \subseteq \mathbb{A}^N$, which is nonempty since it contains the origin. Unlike in depth 3, this set may be reducible, so we pass to an irreducible component V of $\mathbf{V}(q_1, q_2)$. By Krull's Height Theorem (Lemma 2.7), V has codimension at most 2 in $\mathbb{A}^N$. By Bézout's Inequality (Proposition 2.9), $\deg V \leq \deg q_1 \cdot \deg q_2 \leq \delta^2$. Since q_1 divides T_1 and q_2 divides T_2, both gates vanish on V, and hence so does $P_{3,n}$. We want to show that this is *impossible*.

As in Section 1.2.1, a dimension count shows that some block is free on V, in the sense that its coordinates are algebraically independent on V (Lemma 2.19). Say this block is B_1. To get rid of B_2 and B_3, we would like to set one coordinate of each of them to zero. However, this may kill $\mathbf{x}^{B_1}$ as well. For example, B_1 is free on the quadric $x_2 = x_1 x_{n+1}$ from Section 1.3.1. Setting $x_{n+1} = 0$ forces $x_2 = 0$ and kills $\mathbf{x}^{B_1}$. Setting $x_{n+2} = 0$ instead keeps B_1 free. Thus we must choose carefully which coordinate to set to zero in each block, avoiding coordinates such as x_{n+1} in the example above.

We call coordinates that kill $\mathbf{x}^{B_1}$, as in the example above, *bad* and the other coordinates *good*. Lemma 2.18 bounds the number of bad coordinates, ensuring that, for sufficiently large n, each of B_2 and B_3 contains a good coordinate. Extend B_1 to a maximal set T of coordinates that are algebraically independent on V. By maximality, every coordinate x_a with $a \in [N] \setminus T$ is algebraically dependent on $\mathbf{x}_T$ on V; we call these coordinates *dependent*. Thus x_a satisfies an equation $Q_a(\mathbf{x}_T, x_a) = 0$ on V, and by Lemma 2.16, we may choose Q_a to have degree at most $\deg V$. Let $Q \in \mathbb{F}[\mathbf{x}_T]$ be the product of the leading coefficients of the Q_a as polynomials in x_a. Thus $Q \neq 0$ and $\deg Q \leq \operatorname{codim} V \cdot (\deg V - 1)$. Over the locus $Q \neq 0$, the coordinate projection from V to $\mathbb{A}^T$ is finite and surjective (Lemma 2.17). By Lemma 2.14, we may set coordinates in $T \setminus B_1$ to zero while keeping B_1 free on some irreducible component, provided that Q does not become identically zero.

To count the bad coordinates, fix a monomial of Q. It suffices to avoid the coordinates outside T and the variables appearing in this monomial. These restrictions exclude at most $\operatorname{codim} V + \deg Q$ coordinates. By the degree bound above, $\operatorname{codim} V + \deg Q \leq \operatorname{codim} V \cdot \deg V \leq 2\delta^2$. If we take $n > 2\delta^2$, then each of B_2 and B_3 contains a good coordinate. Setting these two coordinates to zero gives an irreducible subvariety $V_0 \subseteq V$ on which B_1 is still free and $P_{3,n}$ restricts to $\mathbf{x}^{B_1}$ (Lemma 2.20). Thus $P_{3,n}$ does not vanish on $V_0 \subseteq V$, but T_1 and T_2 do vanish, a contradiction. For every k, if $n > k\delta^k$, then the same argument shows that $P_{k+1,n}$ cannot be computed by $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits in which every nonconstant product gate has a factor vanishing at the origin (Theorem 3.1). Such circuits include homogeneous circuits.

Homogeneous Circuits in the Border Setting. Now suppose that $T_1 + T_2 = P_{3,n} + \varepsilon S$ for some $S \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$. Here T_1 and T_2 are nonconstant products of homogeneous polynomials of degree at most δ with coefficients in the field $\mathbb{F}((\varepsilon))$ of formal Laurent series (Definition 2.12). Constant summands may first be removed by subtracting the identity evaluated at $\mathbf{x} = \mathbf{0}$, which only changes S. As remarked in Section 1.1, the argument of Homogeneous Circuits in the Exact Setting: A Warm-Up still works if every nonconstant product gate has a factor vanishing at the origin (Theorem 4.16).

In particular, choose one factor vanishing at the origin from each remaining gate. The ideal generated by these factors in $\mathbb{F}((\varepsilon))[\mathbf{x}]$ has at most two generators and contains $P_{3,n} + \varepsilon S$. Choose a minimal prime $\mathfrak{p}$ of this ideal contained in $\langle x_1, \ldots, x_N \rangle$. We then set $\varepsilon = 0$ in every polynomial

of $\mathfrak{p}$ that has no negative powers of ε, not just in the chosen factors, and take the common zero set of the resulting polynomials. Since $P_{3,n} + \varepsilon S$ is such a polynomial, $P_{3,n}$ vanishes on this variety. For example, take $q_1 = x_1$ and $q_2 = x_1 + \varepsilon x_2$. Setting $\varepsilon = 0$ in the factors gives only the equation $x_1 = 0$. However, their ideal also contains $(q_2 - q_1)/\varepsilon = x_2$, which has no negative powers of ε, so we obtain the subspace $x_1 = x_2 = 0$. Lemma 4.8 shows that every irreducible component of the resulting variety has the same dimension as the variety over $\mathbb{F}((\varepsilon))$ and degree at most δ^2. The rest of Homogeneous Circuits in the Exact Setting: A Warm-Up remains unchanged, with an irreducible component of the resulting variety in place of V. Thus, when $n > 2\delta^2$, homogeneous $\overline{\Sigma^{[2]}\Pi\Sigma\Pi^{[\delta]}}$ circuits *cannot* compute $P_{3,n}$ (Theorem 4.16). More generally, the same argument rules out homogeneous $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ circuits for $P_{k+1,n}$ whenever $n > k\delta^k$. Therefore, the real difficulty lies in non-homogeneous circuits, which we turn to next.

1.4.2 Non-Homogeneous Circuits

Just as in Dutta and Saxena [DS26], the case of non-homogeneous circuits is the case that requires the most work. Dealing with cancellations that can occur in this setting requires much care. We now give an overview of Sections 5 and 6, which represent the heart of this work.

In general, a gate need not have a factor vanishing at the origin, so we select factors *greedily* for reduction. Start with $\mathbb{A}^N$ and inspect the gates in turn. If a factor has a zero on the current variety, then intersect with its zero set and pass to an irreducible component; otherwise, leave the variety unchanged (Lemma 2.21). For $k = 2$, the final variety V has codimension at most 2 and degree at most δ^2. Every gate either vanishes on V or has all its factors restricting to units of $\mathbb{F}[V]$. As noted in Section 1.3.1, units of $\mathbb{F}[V]$ need not be constants, so the surviving gates do not simply disappear. In the border setting, such gates are even more powerful: Every factor $1 - \omega^j \varepsilon \ell_i$ of Kumar's construction (Equation (1)) is invertible as a power series in ε and $\mathbf{x}$, and yet two such gates approximate every polynomial.

As in [DS26], we handle the surviving gates by DiDIL, which works with power series instead of polynomials. As in Homogeneous Circuits in the Exact Setting: A Warm-Up, let T be a maximal set of coordinates that are algebraically independent on V, and call the coordinates in T *free*. We express each dependent coordinate as a power series in the free coordinates and substitute these series into the circuit. For DiDIL, every factor of a surviving gate must then become an invertible power series, as in the all-non-homogeneous case of [DS26]. Section 5 first carries this out in the exact setting, which is already nontrivial. The border setting is considerably more delicate, but the examples and lemmas of the exact setting already contain most of its ideas, so we describe and handle the exact setting first.

Low-Weight Translation. To expand the dependent coordinates as power series around a point of V, we have to choose the point carefully. The origin need not lie in V, as $x_1 x_{n+1} = 1$ illustrates. Even when it does, the dependent coordinates need not be power series in the free coordinates around the origin. For example, on $\mathbf{V}(x_N^2 - x_1)$, where x_1 must be free since it belongs to B_1, the two solutions for x_N, namely $\pm x_1^{1/2}$, are not power series in x_1. The implicit function theorem (Lemma 4.14) explains what goes wrong. Recall from Homogeneous Circuits in the Exact Setting: A Warm-Up that each dependent coordinate x_a satisfies an equation $Q_a(\mathbf{x}_T, x_a) = 0$ on V. The implicit function theorem expresses the dependent coordinates as power series in the translated

free coordinates around every point $b \in V$ at which $(\partial_{x_a} Q_a)(b) \neq 0$ for every dependent coordinate x_a (Lemma 4.15). In the example, $Q_N = x_N^2 - x_1$, and its derivative $\partial_{x_N} Q_N$ equals $2x_N$ and vanishes at the origin.

However, b *cannot* be generic: For generic b, the translated block monomial $\prod_{t \in B_1}(x_t + b_t)$ has a nonzero constant term, whereas our lower bound requires every monomial to involve many variables. Thus b must be *low-weight*; that is, it must have few nonzero coordinates. In the example, $b = (1, 0, \ldots, 0, 1)$ works. With $y_1 := x_1 - 1$, we get $x_N = (1 + y_1)^{1/2}$, a power series in y_1, and the block monomial becomes $(1 + y_1)\, x_2 \cdots x_n$.

One could immediately ask at this point: *Why do such points exist at all?* Roughly, the condition on b says that a single nonzero polynomial of low degree does not vanish at the free coordinates of b, and low-weight points are enough to avoid the zeros of such a polynomial. This polynomial R is a product of the resultants of the polynomials Q_a and their derivatives $\partial_{x_a} Q_a$. Each resultant vanishes where the corresponding polynomial and its derivative have a common root. In the example, R is a multiple of x_1, so every point of V with $b_1 \neq 0$ works. As in Homogeneous Circuits in the Exact Setting: A Warm-Up, setting all variables outside one monomial of R to zero keeps R nonzero. Thus R is nonzero at some point with at most $\deg R$ nonzero coordinates. Lifting this point to V gives the desired b (Lemma 5.1). Applying this argument to the subvariety constructed in Homogeneous Circuits in the Exact Setting: A Warm-Up, on which $P_{3,n}$ restricts to $\mathbf{x}^{B_1}$, gives Theorem 5.3. The translating point has $O\left(\delta^4\right)$ nonzero coordinates (or $O\left(k\delta^{2k}\right)$ for general k). For $k = 2$, the translated block monomial $\prod_{t \in B_1}(x_t + b_t)$ becomes a sum of at most two products of invertible power series. Each of these power series is the restriction of a polynomial of degree at most δ.

The resulting identity has the *all-non-homogeneous* form of [DS26]. Every monomial of the translated block monomial still involves at least $n - O\left(\delta^4\right)$ variables for the $k = 2$ case, or $n - O\left(k\delta^{2k}\right)$ for general k.

Border Low-Weight Translation. In the border setting, the steps of Section 1.4.2 meet new obstacles, which Theorem 6.7 overcomes as follows; we illustrate them with examples. Let $\Delta := \delta^2$; for general k we define $\Delta := \delta^k$.

First, having a zero is not the right criterion for reducing modulo a factor. For every $\varepsilon \neq 0$, the factor $1 - \varepsilon x_1$ has a zero at $x_1 = \varepsilon^{-1}$. This zero escapes to infinity as $\varepsilon \to 0$, while the factor itself tends to 1. Thus reducing modulo $1 - \varepsilon x_1$ tells us *nothing* in the limit! The factors $1 - \omega^j \varepsilon \ell_i$ of Kumar's construction (Equation (1)) behave in the same way: Their zero sets, the hyperplanes $\ell_i = \omega^{-j}\varepsilon^{-1}$, escape to infinity. One may wonder whether, as in the exact setting, it is enough that the factors of the surviving gates are nonzero at the origin. Unfortunately, nonvanishing at the origin is *not* enough. The factor $\varepsilon + x_2$ takes the nonzero value ε at the origin, but its inverse $\sum_{d \geq 0}(-x_2)^d \varepsilon^{-d-1}$ has unbounded negative powers of ε. DiDIL cannot handle such inverses because it divides by these factors and then takes the limit $\varepsilon \to 0$.

It turns out that what DiDIL needs is that every factor of a surviving gate becomes an ε-*unit*, that is, a power of ε times an invertible power series (Definition 4.2), once we substitute power series for the dependent coordinates. For example, $1 - \omega^j \varepsilon \ell_i$ and $\varepsilon(1 + x_2)$ are ε-units, but $\varepsilon + x_2$ is not. Accordingly, we reduce modulo a factor *only if* it is **not invertible** near the origin, up

to a power of ε, on the current variety (Lemma 6.9). For example, we would reduce modulo $\varepsilon + x_2$, but never modulo the factors $1 - \omega^j \varepsilon \ell_i$ of Equation (1). Setting $\varepsilon = 0$ as in Homogeneous Circuits in the Border Setting, we then choose an irreducible component V through the origin. This component has codimension at most 2 (at most k in the general setting) and degree at most Δ. Write $\mathbf{u} := \mathbf{x}_T$ for its free coordinates.

Second, the point b around which we expand must satisfy more conditions than in the exact setting. For each dependent coordinate x_a, let $\widetilde{Q}_a$ be a polynomial of degree at most Δ satisfying $\widetilde{Q}_a(\mathbf{u}, x_a) = 0$ on the variety. Scale $\widetilde{Q}_a$ by an appropriate power of ε so that its specialization at $\varepsilon = 0$ is nonzero.

As in Low-Weight Translation, b must avoid the zeros of a single polynomial R of degree at most $k\Delta^2$. Here R is formed from the leading coefficients of the polynomials $\widetilde{Q}_a(0, \mathbf{u}, x_a)$ and from the resultants of the $\widetilde{Q}_a$ and their derivatives. For each resultant, we divide out the maximal power of ε and take the limit $\varepsilon \to 0$. For example, let $\widetilde{Q}_1 := x_1^2 - \varepsilon(1 + x_2)$. Its derivative is $2x_1$, and the resultant of these two polynomials is $-4\varepsilon(1 + x_2)$. Thus this resultant contributes the factor $-4(1 + x_2)$ to R, and the translating point b must satisfy $b_2 \neq -1$.

In addition, every factor of a surviving gate must become an ε-unit after the substitution. For instance, if x_3 is free, then translating by b sends a surviving factor $1 + x_3$ to $(1 + b_3) + x_3$. The resulting factor is an ε-unit only if $b_3 \neq -1$. In general, for each distinct factor of the surviving gates, Lemma 6.9 gives a polynomial h that is nonzero at the origin such that the factor becomes an ε-unit whenever $h(b) \neq 0$. For the factor $1 + x_3$ above, $h = 1 + x_3$ itself works, and the condition $h(b) \neq 0$ is exactly $b_3 \neq -1$. There may be up to s such polynomials h, and we have no bound on their degrees. If we multiplied R by all these polynomials and argued as in Low-Weight Translation, then the resulting bound on $|\mathrm{supp}(b)|$ would depend on s. We need a bound independent of s.

Instead, we use that every h is nonzero at the origin: No h vanishes identically on an irreducible variety through the origin, so a generic point of such a variety avoids the zeros of all the h at once. Lemma 6.10 finds such a variety $Z \subseteq V$ that lies in a coordinate subspace of dimension $O(\Delta^3)$ and on which R does not vanish identically either. Choosing b generically on Z gives $|\mathrm{supp}(b)| = O(\Delta^3)$, independently of s; for general k, the bound becomes $O(k^2\Delta^3)$.

Third, we write the dependent coordinates as power series in ε and the free coordinates around b, possibly after replacing ε by a power ε^e; we call such a tuple of power series a *branch* through b. Three issues arise, which we resolve one by one.

- *Non-monic equations.* The polynomial $\widetilde{Q}_a$, introduced in the second step above, need not be monic in x_a, and its degree may drop at $\varepsilon = 0$. For example, $\varepsilon x_a^2 + x_a - x_t$, with $t \in T$, has one root $x_t - \varepsilon x_t^2 + \cdots$ near the origin and another root $-\varepsilon^{-1} - x_t + \cdots$, which goes to infinity as $\varepsilon \to 0$. We resolve the problem of non-monic equations using Weierstrass preparation (Theorem 6.2), which expresses $\widetilde{Q}_a$ near b as an invertible power series times a monic polynomial W_a in x_a. The roots of W_a are exactly the roots of $\widetilde{Q}_a$ near b. In the example, with $b = \mathbf{0}$, we get $W_a = x_a - (x_t - \varepsilon x_t^2 + \cdots)$.
- *Ramification.* Even a monic equation may have roots that need fractional powers of ε. For example, the equation $x_1^2 = \varepsilon(1 + x_2)$ has the roots $\pm\varepsilon^{1/2}\sqrt{1 + x_2}$, which are not power

series in ε and x_2. We resolve this using the Abhyankar–Jung Theorem (Theorem 6.3). Since $R(b) \neq 0$, the discriminant of W_a is an ε-unit. The theorem therefore implies that the roots of W_a become power series after replacing ε by a suitable power ε^e. In the example, replacing ε by ε^2 gives the roots $\pm\varepsilon\sqrt{1+x_2}$, which are power series in ε and x_2.

- *Compatibility.* The roots chosen for different dependent coordinates must satisfy all equations of the variety, not just the equation $\widetilde{Q}_a(\mathbf{u}, x_a) = 0$ for each coordinate. For example, let $\widetilde{Q}_2 = x_2^2 - \varepsilon^2(1+x_1)$ and $\widetilde{Q}_3 = x_3^2 - \varepsilon^2(1+x_1)$. The ideal they generate contains $\widetilde{Q}_2 - \widetilde{Q}_3 = (x_2 - x_3)(x_2 + x_3)$. Thus every irreducible component of their common zero set lies in $x_3 = x_2$ or in $x_3 = -x_2$. Suppose that we work on a component with $x_3 = x_2$; its ideal contains $x_3 - x_2$. However, the choices $x_2 = \varepsilon\sqrt{1+x_1}$ and $x_3 = -\varepsilon\sqrt{1+x_1}$ satisfy $\widetilde{Q}_2 = 0$ and $\widetilde{Q}_3 = 0$ but violate $x_3 = x_2$. We resolve this in Lemma 6.11.

We translate by the point of the branch above $\mathbf{x}_T = b_T$, where $b_T := (b_i)_{i\in T}$. This point specializes to b at $\varepsilon = 0$ but may depend on ε; for the branch $x_1 = \varepsilon\sqrt{1+x_2}$ of the ramification example, it is $(\varepsilon, 0)$. After this translation, Theorem 6.7 expresses each dependent coordinate as $x_a = \varphi_a(\varepsilon, \mathbf{u})$, with $\varphi_a(\varepsilon, \mathbf{0}) = 0$. Here φ_a is a root of a polynomial Q_a of degree at most Δ. The polynomial Q_a is obtained from $\widetilde{Q}_a$ by the translation and by replacing ε by ε^e. Moreover, $(\partial_{x_a} Q_a)(\mathbf{u}, \varphi_a)$ is an ε-unit. With $(\mathbf{u}, \boldsymbol{\varphi})$ denoting the tuple with entries x_i for $i \in T$ and φ_a for $a \notin T$, the theorem, for general k, gives

$$T_1 + \cdots + T_{k'} = f_0 + \varepsilon S_0, \tag{4}$$

where $k' \leq k$, $S_0 \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$, and

$$T_i := \varepsilon^{e_i} \prod_j p_{i,j}(\mathbf{u}, \boldsymbol{\varphi}), \qquad f_0 := P_{k+1,n}(b + (\mathbf{u}, \boldsymbol{\varphi}(0, \mathbf{u}))),$$

for some integers e_i and polynomials $p_{i,j}$ of degree at most δ, each with $p_{i,j}(\mathbf{u}, \boldsymbol{\varphi})$ an ε-unit. The polynomials Q_a and $p_{i,j}$ have coefficients in $\mathbb{F}[[\varepsilon]]$, and all degree bounds refer to the polynomial variables. We henceforth assume $n > k + |\mathrm{supp}(b)|$, since otherwise the desired bound already follows from the bound on $|\mathrm{supp}(b)|$. Thus $f_0 \neq 0$, and every monomial of f_0 involves at least $n - k - |\mathrm{supp}(b)|$ variables (Lemma 6.20); in particular, $f_0(\mathbf{0}) = 0$. (Remark 6.8 explains why we substitute the power series φ_a for the dependent coordinates instead of working in a quotient ring, as in the exact setting.)

Applying Modified DiDIL. In hindsight, the identity (4) has the same form as the starting point of DiDIL in [DS26, Section 6], which we sketched in Section 1.2. There, every gate is a power of ε times a product of linear polynomials whose constant terms are nonzero at $\varepsilon = 0$. The role of f_0 is played by the polynomial f, where $f := P_{k+1,n}(\ell_1, \ldots, \ell_N)$ and the linear forms $\ell_1, \ldots, \ell_N$ have rank at least $N - k$. Table 2 lists the full correspondence.

Two main differences remain.

- First, the factors $p_{i,j}(\mathbf{u}, \boldsymbol{\varphi})$ are power series, and they are implicit in the sense that we know them only through the equations $Q_a(\mathbf{u}, \varphi_a) = 0$.
- Second, f_0 is a power series rather than a homogeneous polynomial, so the role of f is

played by the lowest homogeneous component $H_{d_0}(f_0)$ of f_0, that is, the sum of its terms of the least degree d_0. Since $f_0(\mathbf{0}) = 0$, we have $d_0 \geq 1$.

As in [DS26], we scale the free coordinates by a new variable z. Since the dependent coordinates are now the power series φ_a, we combine the substitution and the scaling into one map Ψ, as defined below.

For a polynomial p, let $\Psi(p)$ be the power series obtained from p by substituting $\varphi_a(\varepsilon, \mathbf{u})$ for every dependent coordinate x_a and then replacing $\mathbf{u}$ by $z\mathbf{u}$. Thus $\Psi(x_i) = zx_i$ for $i \in T$, and $\Psi(x_a) = \varphi_a(\varepsilon, z\mathbf{u})$ for $a \notin T$. Replacing $\mathbf{u}$ by $z\mathbf{u}$ in (4) gives

$$T_1(z\mathbf{u}) + \cdots + T_{k'}(z\mathbf{u}) = f_0(z\mathbf{u}) + \varepsilon S_0(z\mathbf{u}), \tag{5}$$

where $T_i(z\mathbf{u}) = \varepsilon^{e_i} \prod_j \Psi(p_{i,j})$. Since every $p_{i,j}(\mathbf{u}, \boldsymbol{\varphi})$ is an ε-unit, every factor $\Psi(p_{i,j})$ is a power of ε times a power series whose limit as $\varepsilon \to 0$ has a nonzero constant term in $\mathbb{F}$. This condition is the *all-non-homogeneous property* of Dutta and Saxena [DS26]. Setting $\varepsilon = 0$ on the right side of (5) gives

$$f_0(z\mathbf{u}) = H_{d_0}(f_0)\, z^{d_0} + (\text{higher powers of } z),$$

so $H_{d_0}(f_0)$ is the coefficient of the lowest power of z. The elimination steps keep track of this coefficient, as in Section 1.2.

For $k = 2$, a single elimination step suffices. Since $f_0 \neq 0$, at least one gate survives; that is, $k' \geq 1$. If $k' = 1$, then $f_0 + \varepsilon S_0$ equals T_1 and is a power of ε times an invertible power series, so its limit f_0 has a nonzero constant term, contradicting $f_0(\mathbf{0}) = 0$. Thus $k' = 2$. Write $T_2(z\mathbf{u}) = \varepsilon^a \widetilde{T}$, where $\widetilde{T}$ has a nonzero limit as $\varepsilon \to 0$, and divide (5) by $\widetilde{T}$. After this division, the summand $T_2(z\mathbf{u})/\widetilde{T}$ equals ε^a and therefore does not depend on z, so differentiating with respect to z removes it:

$$\partial_z\left(\frac{T_1(z\mathbf{u})}{\widetilde{T}}\right) \;=\; \partial_z\left(\frac{f_0(z\mathbf{u}) + \varepsilon S_0(z\mathbf{u})}{\widetilde{T}}\right). \tag{6}$$

Taking the limit $\varepsilon \to 0$ commutes with ∂_z, so we can compare the two sides of (6) in the limit.

On the right side of (6), the limit of $\widetilde{T}$ is a power series in z whose constant term c is a nonzero element of $\mathbb{F}$, by the all-non-homogeneous property. Thus the limit of the right side starts with z^{d_0-1}, with coefficient $d_0 H_{d_0}(f_0)/c$; this coefficient is nonzero, since $d_0 \geq 1$ and the characteristic is 0.

On the left side of (6), let $h := T_1(z\mathbf{u})/\widetilde{T}$ and use its logarithmic derivative, defined by $\mathrm{dlog}_z(h) := (\partial_z h)/h$; see Differentiation. Here h is a power of ε times a product of series $\Psi(p)^{\pm 1}$, where p ranges over the set $\mathcal{P}$ of the at most s distinct polynomials $p_{i,j}$. After removing this power of ε, the limit of h has a nonzero constant term in $\mathbb{F}$, so h contributes only a nonzero constant factor to the coefficient of the lowest power of z. Since dlog_z turns products into sums, $\mathrm{dlog}_z(h)$ is a linear combination, with integer coefficients, of the series $\mathrm{dlog}_z \Psi(p)$ with $p \in \mathcal{P}$, however many factors the gates have. Remove the maximal power of ε from this linear combination and take the limit $\varepsilon \to 0$. Comparing the coefficients of the lowest powers of z on the two sides of (6) shows that $H_{d_0}(f_0)$ is, up to a nonzero constant factor, the coefficient of the lowest power of z in the resulting series.

In [DS26], every factor is linear, and the next step expands each $\operatorname{dlog}_z \Phi(\ell)$ by the geometric series (3). Thus every coefficient of the linear combination is a sum of at most s powers of linear forms, and the partial derivative argument of Section 1.2 finishes the proof.

For us, $\Psi(p)$ is a power series in z *even if* p is linear. For example, suppose that a dependent coordinate x_a satisfies $x_a + x_a^2 = \ell$ for a linear form ℓ in the free coordinates. Here $Q_a = x_a + x_a^2 - \ell$, and φ_a is the root that vanishes at $\mathbf{u} = \mathbf{0}$. Since $\Psi(\ell) = z\ell$, the series $\Psi(x_a)$ satisfies $\Psi(x_a) + \Psi(x_a)^2 = z\ell$, and hence

$$\Psi(x_a) \;=\; \frac{-1+\sqrt{1+4z\ell}}{2} \;=\; z\ell - z^2\ell^2 + 2z^3\ell^3 - 5z^4\ell^4 + \cdots .$$

Now consider the linear factor $p = 1 + x_a$. The series $\Psi(1 + x_a)$ equals $1 + \Psi(x_a)$ and is *not* of the form $c + z\tilde{\ell}$, and the geometric series for $1/\Psi(1 + x_a)$ would produce powers of the power series $\Psi(x_a)$. The coefficients of the result can still be computed, but the result is no longer a product of a few simple factors, and it is unclear how to keep track of such expressions through the repeated elimination steps. Instead, we differentiate the equation $\Psi(x_a) + \Psi(x_a)^2 = z\ell$ with respect to z, which gives

$$(1 + 2\Psi(x_a)) \cdot \partial_z \Psi(x_a) = \ell.$$

Using $\ell = z^{-1}\Psi(\ell)$, we obtain

$$\operatorname{dlog}_z \Psi(1 + x_a) = \frac{\partial_z \Psi(x_a)}{1 + \Psi(x_a)} = z^{-1} \cdot \Psi(\ell) \cdot \Psi(1 + 2x_a)^{-1} \cdot \Psi(1 + x_a)^{-1}.$$

Apart from the factor z^{-1}, this is again a product of series $\Psi(q)^{\pm 1}$ with linear q, so we can differentiate it again without expanding any series.

For a general polynomial p, Lemma 6.16 applies the chain rule to p and to the equations $Q_a(\mathbf{u}, \varphi_a) = 0$. The chain rule gives

$$z\,\partial_z \Psi(p) \;=\; \Psi(Dp) - \sum_{a \notin T} \Psi(\partial_{x_a} p) \cdot \Psi(DQ_a) \cdot \Psi(\partial_{x_a} Q_a)^{-1},$$

where $D := \sum_{i \in T} x_i \partial_{x_i}$ is the *Euler derivation* in the free coordinates; see Differentiation. For $p = 1 + x_a$ and $Q_a = x_a + x_a^2 - \ell$, this is the identity above. Dividing by $\Psi(p)$ shows that, for every $p \in \mathcal{P}$, the series $z \operatorname{dlog}_z \Psi(p)$ is a sum of at most $k + 1$ products of at most four series $\Psi(q)^{\pm 1}$ with $\deg q \le \Delta$.

Replacing the geometric series by the chain rule is the *crucial difference* from [DS26]. We never expand the series $\Psi(q)^{\pm 1}$ during the elimination steps, and we extract coefficients only at the very end. This procedure works because the equations $Q_a(\mathbf{u}, \varphi_a) = 0$ determine the coefficients of the branches recursively. As a result, for every q of degree at most Δ, every coefficient of $\Psi(q)$ is a polynomial expression, with coefficients in $\mathbb{F}((\varepsilon))$, in a fixed set of polynomials of degree at most Δ (Lemma 6.17). View q and the Q_a as polynomials in the dependent coordinates. The fixed set consists of the homogeneous components of their coefficients. Its size depends only on k and Δ, not on the power of z. The same holds for $\Psi(q)^{-1}$ if $q(\mathbf{u}, \boldsymbol{\varphi})$ is an ε-unit.

In the example above, the coefficient of z^d in $\Psi(x_a)$ is a multiple of ℓ^d, so the powers of linear

forms of [DS26] reappear. In general, they are replaced by *compositions* $h(g_1, \ldots, g_r)$ of an arbitrary polynomial h with r polynomials g_j of low degree. We write $\mathcal{C}_d(m, r)$ for the class of sums of at most m such compositions over $\mathbb{F}$ in which every g_j has degree at most d. We write $\overline{\mathcal{C}}_d(m, r)$ for the corresponding border class (Definition 6.4). For $k = 2$, the linear combination of the series $\operatorname{dlog}_z \Psi(p)$ above is, up to a factor z^{-1}, a sum of at most $O(s)$ products of at most four series $\Psi(q)^{\pm 1}$, so Corollary 6.18 gives

$$H_{d_0}(f_0) \in \overline{\mathcal{C}}_{\Delta}\left(O\left(s\right), O\left(\Delta^3\right)\right).$$

For larger k, we repeat the elimination step. Throughout the process, every summand has the form $(U/V) \cdot (A/B)$. Here U and V come from the gates and are products of series $\Psi(p)$ with $p \in \mathcal{P}$. The factors A and B come from the logarithmic derivatives and are nonzero sums of terms $c \prod_t \Psi(q_t)^{\pm 1}$ with $c \in \mathbb{F}((\varepsilon))[z, z^{-1}]$ and $\deg q_t \leq \Delta$. We call this class of expressions our *bloated* model $\operatorname{Gen}(\cdot, \cdot)$ (Definition 6.15); it is the analogue of the model of Dutta and Saxena [DS26], with products of series in place of powers of linear forms. By the chain rule, every elimination step keeps the summands in this form, and the number of factors in each product stays below 4^k (Lemma 6.16 and (31)).

The process stops after at most $k - 1$ elimination steps. Comparing the coefficients of the lowest power of z in the limit, as for $k = 2$, gives Theorem 6.12:

$$H_{d_0}(f_0) \cdot G_2 = G_1$$

for nonzero $G_1, G_2 \in \overline{\mathcal{C}}_{\Delta}(m, r)$, where $\log m = O\left(k4^k(k + \log s)\right)$ and $r = O\left(k4^k(\Delta + 1)^{k+1}\right)$. What matters is that, for constant k, the parameter m is polynomial in s and that r depends neither on s nor on n.

Concluding the Lower Bound. It remains to derive the lower bound on s from the identity $H_{d_0}(f_0) \cdot G_2 = G_1$, as in Section 1.2. Forbes [For15, Proposition 4.18] proved that the trailing monomial of a nonzero element of $\mathcal{C}_d(m, r)$ involves at most $O\left(d(\log m + r \log(2r))\right)$ variables. His proof compares an upper and a lower bound on the dimension of a space of shifted partial derivatives. The upper bound is a rank condition and hence survives limits, as observed in [DDS22, Claim C.9]. Theorem 6.6 states the border version that we use, with the slightly better bound $O\left(d(\log m + r)\right)$. Since trailing monomials are multiplicative, $\textsc{tm}(G_1) = \textsc{tm}(H_{d_0}(f_0)) \cdot \textsc{tm}(G_2)$, so $\textsc{tm}(H_{d_0}(f_0))$ involves at most $O\left(\Delta(r + \log m)\right)$ variables. On the other hand, every monomial of $H_{d_0}(f_0)$ involves at least $n - k - |\operatorname{supp}(b)| = n - O\left(k^2\delta^{3k}\right)$ variables. Comparing the two bounds gives $n \leq O\left(k(4\delta)^k \log s\right) + (2\delta)^{O\left(k^2\right)}$, which proves Theorem 1.4; for $k = 2$, this is $n \leq O\left(\delta^2 \log s\right) + \delta^{O(1)}$. In particular, for constant k and δ, the size s must be $2^{\Omega(n)}$.

1.4.3 Impossibility Result

Finally, we outline the proof of Theorem 1.3, which states that, for sufficiently large n, the polynomial $P_{k+1,n}$ cannot be computed by a $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit, regardless of its size. The full proof appears in Section 7.4. We may assume that $\mathbb{F}$ is algebraically closed. Suppose towards

contradiction that $P_{k+1,n}$ is computed by a $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit and write

$$P_{k+1,n} = T_1 + \cdots + T_k,$$

where each T_i is a product of polynomials of degree at most δ.

We first apply the greedy process of selecting factors (Lemma 2.21), described in Section 1.4.2, to find an irreducible variety V with $\operatorname{codim} V \le k$ and $\deg V \le \delta^k$, on which each T_i either vanishes or has every factor invertible. As in Homogeneous Circuits in the Exact Setting: A Warm-Up, a dimension count shows that some block, say B_1, is free on V. For $n > k\delta^k$, we can set one coordinate in each of the other blocks to zero and choose an irreducible component V_0 of the resulting intersection on which B_1 remains free (Lemma 2.20). Let $I \subseteq [k]$ index the gates that do not vanish on V_0. Restricting the identity to V_0 gives

$$\mathbf{x}^{B_1} = \sum_{i \in I} T_i \tag{7}$$

on V_0, where every factor of each surviving gate T_i with $i \in I$ is a unit. Moreover, Krull's Height Theorem and Bézout's Inequality give $\operatorname{codim} V_0 \le \operatorname{codim} V + k \le 2k$ and $\deg V_0 \le \deg V \le \delta^k$.

Each surviving T_i is a unit on V_0 but need not be a constant. In the depth-3 setting, V_0 is an affine subspace, whose only units are nonzero constants. On a general affine variety, however, units can be nonconstant: for example, both x_1 and x_2 are nonconstant units on $\mathbf{V}(x_1x_2 - 1)$.

To control these units, we consider the group $\mathbb{F}[V_0]^\times/\mathbb{F}^\times$ of units modulo constants. A classical analogue in algebraic number theory is Dirichlet's Unit Theorem: for a number field K, the group $\mathcal{O}_K^\times/\mu_K$ is free abelian of finite rank, where $\mathcal{O}_K$ is the ring of integers in K and μ_K is the group of roots of unity in K. For irreducible affine varieties over an algebraically closed field, it is likewise classical that $\mathbb{F}[V_0]^\times/\mathbb{F}^\times$ is free abelian of finite rank [Ros61; Sam57; Sam66]; see also [For14; Sas11]. Combining Samuel's argument [Sam57] with a bound on the number of irreducible components at infinity (Theorem 7.3), we obtain the bound (Theorem 7.4):

$$\operatorname{rank}_{\mathbb{Z}}(\mathbb{F}[V_0]^\times/\mathbb{F}^\times) \le \deg V_0.$$

If all the coordinates $x_1, \ldots, x_n$ in B_1 were units on V_0, then this bound would give a contradiction. Indeed, since B_1 is free on V_0, their classes in $\mathbb{F}[V_0]^\times/\mathbb{F}^\times$ would be multiplicatively independent: after clearing negative exponents, any nontrivial multiplicative relation modulo constants would give a nontrivial polynomial relation among these coordinates. We would therefore have

$$n \le \operatorname{rank}_{\mathbb{Z}}(\mathbb{F}[V_0]^\times/\mathbb{F}^\times) \le \deg V_0 \le \delta^k,$$

leading to a contradiction for sufficiently large n.

Nevertheless, the coordinates $x_1, \ldots, x_n$ may have zeros on V_0. To deal with these zeros, we use the unit-group rank bound to control the span of the logarithmic differentials $\mathrm{d}q/q$ of the factors of the T_i. Combining this control with (7), we construct a polynomial H that does not vanish identically on V_0 and satisfies $\deg H \le (2k+\delta)\delta^k$ and $V_0 \cap \mathbf{V}(x_1 \cdots x_n) \subseteq \mathbf{V}(H)$ (Lemma 7.7). In particular, $x_1, \ldots, x_n$ are units on the open subset $V_0 \setminus \mathbf{V}(H)$ of V_0.

Introducing one new coordinate z, we realize $V_0 \setminus \mathbf{V}(H)$ as the irreducible affine variety W, where $W := (V_0 \times \mathbb{A}^1) \cap \mathbf{V}(zH - 1)$. The coordinates $x_1, \ldots, x_n$ remain algebraically independent on W, and Bézout's Inequality gives

$$\deg W \leq \deg V_0 \, (\deg H + 1) \leq (2k + \delta)\delta^{2k} + \delta^k.$$

Applying the unit-group rank bound to W now yields $n \leq \deg W$, a contradiction for sufficiently large n.

1.5 Multiplicative Joins and Lower Bounds in Geometric Complexity Theory

In this section, we reinterpret our results, Theorems 1.3 and 1.4, in the language of *geometric complexity theory*. Proving *robust* lower bounds in complexity theory—that is, bounds that remain valid in limiting cases—is of great interest. In the language of asymptotic geometry, this goal can be formulated as follows:

Let $\{ X_N \}$ be a sequence of spaces. For each N, let $Y_N \subseteq X_N$ be a closed subset of "easy" points, let $p_N \in X_N$ be a point whose complexity we want to study, and let G_N be a group acting on X_N and leaving Y_N invariant. Does the containment $\overline{G_N \cdot p_N} \subseteq Y_N$ fail for sufficiently large N? Equivalently, are the points p_N "hard"—that is, outside Y_N—for sufficiently large N? In the algebraic setting, this translates to proving lower bounds for the border complexity of classes of algebraic computation.

Studying the border complexity of various algebraic complexity classes is central to the geometric complexity theory program proposed by Mulmuley and Sohoni [MS01; MS08]. For this discussion, assume that $\mathbb{F}$ is algebraically closed of characteristic 0. Fix a circuit model, a number of variables, and bounds on degree and circuit size. In the resulting coefficient space, border computation corresponds to the Zariski closure of the polynomials computed by circuits satisfying these bounds. Border complexity classes are the Zariski closures of standard algebraic complexity classes. The geometric complexity theory program aims to separate these border classes using tools from algebraic geometry and representation theory.

Geometric approaches to algebraic complexity often work with spaces of *homogeneous* polynomials, identified by their coefficient vectors. However, in algebraic complexity we are interested in the cancellations that may occur in the *non-homogeneous* setting. To pass to the homogeneous setting for $P \in \mathbb{F}[x_1, \ldots, x_N]$, choose a padding degree m with $m \geq \deg P$ and introduce a new variable x_0. The *homogenization*, or *padding*, of P is the degree-m form $x_0^m P(x_1/x_0, \ldots, x_N/x_0) \in \mathbb{F}[x_0, x_1, \ldots, x_N]$. This construction allows us to apply the techniques of projective geometry to the non-homogeneous setting. We represent restricted circuit classes, such as $\overline{\Sigma^{[k+1]}\Pi\Sigma\Pi^{[\delta]}}$, by geometric spaces of homogeneous polynomials. Membership in these spaces implies the existence of the corresponding exact or border computations.

We now give notation for the spaces of depth-4 computation appearing in this work. Fix $N := (k+1)n$, and let W_N be the space of linear forms, defined by $W_N := \operatorname{span}_{\mathbb{F}} \{ x_0, x_1, \ldots, x_N \}$. The *symmetric power* $S^\delta W_N$ consists of degree-δ forms in these variables. To geometrically describe the product of m degree-δ forms, we use the *multiplicative join* [Lan17, §7.5.3]

$$\mathrm{MJ}^m\left(\mathbb{P}S^\delta W_N\right) := \{ [q_1 \cdots q_m] \mid 0 \neq q_i \in S^\delta W_N \} \subseteq \mathbb{P}S^{m\delta} W_N.$$

The set $\mathrm{MJ}^m(\mathbb{P}S^\delta W_N)$ is the image of the m-fold multiplication map $(\mathbb{P}S^\delta W_N)^m \to \mathbb{P}S^{m\delta}W_N$ and hence is a closed projective variety. If we take $\delta = 1$, then we recover the geometric perspective of Dutta and Saxena [DS26]: *Chow varieties* $\mathrm{Ch}^m(W_N) := \mathrm{MJ}^m(\mathbb{P}W_N)$, which were studied by Hermite [Her54] and Hadamard [Had96]. For $X \subseteq \mathbb{P}S^d W_N$, we use the following *secant varieties* to geometrically describe the top summation gates:

$$\sigma_k^\circ(X) := \bigcup_{p_1,\dots,p_k \in X} \langle\, p_1, \dots, p_k \,\rangle, \qquad \sigma_k(X) := \overline{\sigma_k^\circ(X)},$$

where $\langle\, p_1, \dots, p_k \,\rangle$ is the smallest projective linear subspace containing $p_1, \dots, p_k$; see the survey of Bernardi et al. [Ber+18] for an overview of secant varieties.

We now can phrase our results geometrically. Let k and δ be at least 1. For exact computation, Theorem 1.3 says that if $m\delta \geq n$ and $n > 2(k+\delta)\delta^{2k}$, then

$$\left[x_0^{m\delta-n} P_{k+1,n}\right] \in \sigma_{k+1}^\circ\left(\mathrm{Ch}^{m\delta}(W_N)\right) \setminus \sigma_k^\circ\left(\mathrm{MJ}^m\left(\mathbb{P}S^\delta W_N\right)\right).$$

For border computation, fix positive integers k and δ and suppose that $m\delta \geq n$ and

$$\left[x_0^{m\delta-n} P_{k+1,n}\right] \in \sigma_k\left(\mathrm{MJ}^m\left(\mathbb{P}S^\delta W_N\right)\right).$$

Dehomogenizing a border decomposition by setting $x_0 = 1$ gives a $\overline{\Sigma^{[k]}\Pi^{[m]}\Sigma\Pi^{[\delta]}}$ circuit for $P_{k+1,n}$ with at most m factors per upper product gate. Each bottom polynomial has degree at most δ in N variables, so it has at most $\binom{N+\delta}{\delta}$ monomials and can be computed by a $\Sigma\Pi^{[\delta]}$ circuit of size $O\left(\delta\binom{N+\delta}{\delta}\right)$. Consequently, the resulting border circuit has size

$$s = O\left(km\delta\binom{N+\delta}{\delta}\right).$$

Since $N = (k+1)n$ and k and δ are fixed, we have $s \leq mn^{O(1)}$. On the other hand, Theorem 1.4 gives $s \geq 2^{\Omega(n)}$, and therefore $m \geq 2^{\Omega(n)}$.

Fix positive integers k and δ, and choose $m = m(n)$ so that $m\delta \geq n$ and $m(n) = 2^{o(n)}$. Set $X_N = \mathbb{P}S^{m\delta}W_N$ and $Y_N = \sigma_k(\mathrm{MJ}^m(\mathbb{P}S^\delta W_N))$. The group $G_N = \mathrm{GL}(W_N)$ acts by substitution of linear forms. Our hard point is $P_N = \left[x_0^{m\delta-n} P_{k+1,n}\right]$. Thus our results give $\overline{G_N \cdot P_N} \not\subseteq Y_N$ for sufficiently large N. If we omit the restriction that $m(n) = 2^{o(n)}$, then Kumar's construction [Kum20] shows that, for $k \geq 2$, the containment $\overline{G_N \cdot P_N} \subseteq Y_N$ holds for sufficiently large padding. This padding is necessary for our discussion. Kumar's construction [Kum20] is non-homogeneous but can be made homogeneous using padding. Thus his result has the following geometric formulation: *For every degree-d homogeneous polynomial $f \in \mathbb{F}[x_0, x_1, \dots, x_N]$, we have $\left[x_0^{D-d} f\right] \in \sigma_2\left(\mathrm{Ch}^D(W_N)\right)$ for sufficiently large D.*

1.6 Related Work

For general constant-depth circuits, the first superpolynomial lower bound was obtained by Limaye, Srinivasan, and Tavenas [LST25]. Combining this lower bound with the "hardness-versus-randomness" tradeoff of Chou, Kumar, and Solomon [CKS18] yields the first deterministic black-box PIT algorithm for general constant-depth circuits. The algorithm takes the form of a hitting-set generator and runs in time $s^{O(N^\varepsilon)}$ for every $\varepsilon > 0$. Andrews and Forbes [AF22] also used the results of Limaye, Srinivasan, and Tavenas [LST25] to give a different hitting-set generator for the *border* of general constant-depth circuits.

Further work on polynomial identity testing for depth-4 circuits focuses on restricting the bottom and top fan-ins, i.e., on studying $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits. For $\delta = 1$, Kayal and Saxena [KS07] gave deterministic polynomial-time PIT for $\Sigma^{[k]}\Pi\Sigma$ circuits with constant k. The black-box algorithms rest on rank bounds for identities in this model, derived from Sylvester–Gallai-type theorems for linear forms [DS07; KS09; SS13] (see Relation to Rank Bounds). Beecken, Mittmann, and Saxena [BMS13] proposed an analogous rank-bound program for $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ identities. Gupta [Gup14] formulated Sylvester–Gallai-type conjectures for polynomials of degree δ, replacing linear spans by radicals of ideals. Carrying out this program requires substantial new ideas even for quadratic and cubic polynomials. There is also considerable work on PIT for circuits of the form $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$. For arbitrary constant δ, polynomial-time PIT is known when the top fan-in k is at most 3 [GOS25; Guo26; GW25; PS21]. For arbitrary constant k, polynomial-time PIT was given by Garg et al. [Gar+26] in the case that the bottom fan-in is at most 2. For all constant k and δ, the Divide–Derive–Induct technique of Dutta, Dwivedi, and Saxena [DDS21], the exact counterpart of Divide–Derive–Induct–Limit, gives quasipolynomial-time PIT.

There is also considerable work on obtaining lower bounds for depth-4 circuits. The depth reduction of Agrawal and Vinay [AV08], improved by Koiran [Koi12] and Tavenas [Tav15], showed that to obtain strong lower bounds for general algebraic circuits, it suffices to give sufficiently strong lower bounds for depth-4 circuits. This depth reduction motivates further study of depth-4 circuits and their restricted subclasses. A series of works [Gup+14; Kay+14; KS17b] established lower bounds for homogeneous depth-4 circuits. Kumar and Saraf [KS17a] used the notion of *algebraic rank* to generalize the homogeneous depth-4 setting and proved lower bounds for depth-4 circuits of low algebraic rank. Forbes [For15] used a measure based on trailing monomials to prove lower bounds for sums of *powers* of low-degree polynomials ($\Sigma \bigwedge \Sigma\Pi^{[\delta]}$). Kayal, Saha, and Saptharishi [KSS14] proved lower bounds for non-homogeneous depth-4 circuits with bounded fan-in at both product layers and used these bounds to prove superpolynomial lower bounds for regular formulas. Fournier et al. [Fou+15] gave lower bounds for depth-4 homogeneous formulas. The breakthrough work of Limaye, Srinivasan, and Tavenas [LST25] gave the first superpolynomial lower bounds for general depth-4 circuits, beyond the restricted classes discussed above. At the core of many of these works is a measure defined by the rank of a matrix or the dimension of a vector space. Such measures often arise from *shifted partial derivatives*, introduced by Kayal [Kay12]. Since these lower bounds are based on rank conditions and rank is lower semicontinuous, they extend to the border setting. Indeed, it was noticed by Andrews and Forbes [AF22] that the lower bounds of Limaye, Srinivasan, and Tavenas [LST25] extend to the border of constant-depth circuits. It was also noticed by Dutta, Dwivedi, and Saxena [DDS22] that the rank-based lower bounds of Forbes [For15] extend to the border class $\overline{\Sigma \bigwedge \Sigma\Pi^{[\delta]}}$.

Concurrent work. Recently, Bhat, Dutta, and Shringi [BDS26] used Wronskians to improve the size lower bound for $\overline{\Sigma^{[k]}\Pi\Sigma}$ circuits computing $P_{k+1,n}$ from $2^{n/\exp(k)}$ to $2^{n/\operatorname{poly}(k)}$. Their main idea is to replace the multi-step recursion of DiDIL by a single differential equation. This idea can also be combined with our approach. The border low-weight translation (Theorem 6.7), which is the heart of this paper, puts the circuit identity into a form to which we can apply a Wronskian argument in place of modified DiDIL. The Wronskian argument requires the same adaptations as DiDIL. The factors of the gates are power series known only through the equations of the branches. We therefore again use the chain rule with the Euler derivation and keep the series $\Psi(q)^{\pm 1}$ unexpanded (Section 6.3). This adaptation gives a lower bound of the form $s \geq 2^{n/(\operatorname{poly}(k)\cdot\delta^k)}$ for all sufficiently large n, improving the bound $s \geq 2^{n/(k4^k\cdot\delta^k)}$. The factor δ^k remains, since it comes from Bézout's Inequality in the reduction (see Theorem 6.7) and not from DiDIL. Since the combined approach does not yield a lower bound of $2^{n/\operatorname{poly}(k\delta)}$ in depth 4, we do not incorporate it here and keep the two works intentionally separate.

1.7 Organization

The rest of this paper is organized as follows.

- We collect general preliminaries and auxiliary results in Section 2.
- As a warm-up case for later sections, Sections 3 and 4 establish impossibility under a constant-freeness assumption in the exact and border settings, respectively.
- In Section 5, we develop low-weight translation in the exact setting. Strictly speaking, the rest of the paper *does not* use this section, since its results are subsumed by the border analogue in Section 6.2. We include it as a gentle introduction: the exact setting already contains most of the obstacles and ideas, and they are easier to see there than in the border setting, which adds *further* technicalities.
- Section 6 proves our general exponential border lower bound (Theorem 1.4), without the constant-freeness assumption. The heart of this section, as well as of the paper, is the border low-weight translation in Section 6.2, the border analogue of the exact translation in Section 5. The border low-weight translation puts the circuit into a form to which DiDIL can be applied.

  Even then, the lower bound does not follow directly from [DS26]: the factors of the surviving gates become power series, and we modify DiDIL accordingly in Section 6.3.
- In Section 7, we prove the exact impossibility result (Theorem 1.3) without the constant-freeness assumption, using a rank bound for the group of units modulo constants of an irreducible affine variety.
- We conclude with open problems in Section 8.

In Appendix A, we give a detailed algebraic proof of the version of the Abhyankar–Jung Theorem used in our border argument, deriving it from standard results in commutative algebra, including results on complete local rings and discrete valuation rings. We hope that this proof and the other tools developed in this work are of independent interest beyond their use in proving depth-4 circuit lower bounds.

2 Preliminaries

Write $\mathbb{N} := \mathbb{Z}_{\geq 0}$, and for a natural number $N \in \mathbb{N}$, let $[N] := \{1, \ldots, N\}$. We use bold letters for vectors indexed by a set I, writing $\mathbf{x} := (x_i)_{i \in I}$. For $\mathbf{a} \in \mathbb{N}^I$ with only finitely many nonzero coordinates, write $\mathbf{x}^{\mathbf{a}} := \prod_{i \in I} x_i^{a_i}$. Define the *support* of a vector $\mathbf{a}$ by $\operatorname{supp}(\mathbf{a}) := \{ i \in I \mid a_i \neq 0 \}$. For a monomial $\mathbf{x}^{\mathbf{a}}$, we define $\operatorname{supp}(\mathbf{x}^{\mathbf{a}}) := \operatorname{supp}(\mathbf{a})$. For finite I and $\mathbf{a} \in \mathbb{Z}^I$, define $|\mathbf{a}| := \sum_{i \in I} a_i$.

All rings are commutative with unit. Let R be a ring, let N be a positive integer, and set $\mathbf{x} := (x_1, \ldots, x_N)$. Write $R^\times$ for the *unit group* of R. For $S \subseteq [N]$, let $\mathbf{x}_S := (x_s)_{s \in S}$; the polynomial ring $R[\mathbf{x}_S]$ is a subring of $R[\mathbf{x}]$.

For $f \in R[\mathbf{x}]$, we write $\deg f$ or $\deg_{\mathbf{x}} f$ for its total degree in $\mathbf{x}$ unless otherwise stated. For $s \geq 0$, let $R[\mathbf{x}]_{\leq s}$ be the free R-module with basis the monomials in $\mathbf{x}$ of total degree at most s. For an ideal $I \subseteq R[\mathbf{x}]$, let $I_{\leq s} := I \cap R[\mathbf{x}]_{\leq s}$. A *constant* in $R[\mathbf{x}]$ is an element of R, even when R is itself a polynomial or power-series ring. Extending coefficients from a field $\mathbb{F}$ to its algebraic closure $\overline{\mathbb{F}}$ preserves an exact or border circuit computation and its size, so lower bounds over $\overline{\mathbb{F}}$ imply the corresponding bounds over $\mathbb{F}$.

2.1 Commutative Algebra

We state basic definitions and results that can also be found in standard commutative algebra texts [AM69; Eis95; Lan02; Mat86] and computational algebra texts [CLO25; vzGG13].

For $S \subseteq R$, write $\langle S \rangle$ for the ideal generated by S. The *radical* $\sqrt{I}$ of an ideal I is

$$\{ a \in R \mid a^n \in I \text{ for some } n \geq 1 \}.$$

An ideal I is *radical* if $I = \sqrt{I}$. A proper ideal $\mathfrak{p}$ is *prime* if $R/\mathfrak{p}$ is an integral domain and *maximal* if $R/\mathfrak{p}$ is a field. Maximal ideals are prime, and prime ideals are radical. A *minimal prime* of I is a prime ideal minimal among those containing I.

A ring is *Noetherian* if every ideal is finitely generated. If R is Noetherian, then so are its quotients, its localizations, finitely generated R-algebras, and formal power-series rings over R in finitely many variables [AM69, Propositions 7.1 and 7.3, Corollary 7.7, and Corollary 10.27]. In a Noetherian ring, every prime ideal containing an ideal I contains a minimal prime of I. Moreover, every proper ideal in a Noetherian ring has finitely many minimal primes, whose intersection is its radical [AM69, Proposition 4.6 and Theorem 7.13].

Unique Factorization. Elements a and b of R are *associate* if $a = bu$ for some $u \in R^\times$. A nonzero element $p \in R$ is *prime* if $\langle p \rangle$ is a prime ideal. A nonzero non-unit $x \in R$ is *irreducible* if it is not a product of two non-units. A domain R is a *unique factorization domain* (UFD) if every nonzero non-unit factors uniquely into irreducibles up to order and multiplication by units. If R is a UFD, then so is $R[\mathbf{x}]$ [Lan02, Chapter IV, Corollary 2.4]. For a UFD R, a polynomial in $R[z]$ is *primitive* if its coefficients have no common non-unit divisor. By *Gauss's Lemma*, a primitive polynomial of positive degree is irreducible in $R[z]$ if and only if it is irreducible in $\operatorname{Frac}(R)[z]$ [Lan02, Chapter IV, §2, Theorem 2.3].

Localization. A subset S of a ring R is *multiplicatively closed* if $1 \in S$ and $ab \in S$ for all $a, b \in S$. For a multiplicatively closed subset $S \subseteq R$, the *localization* $S^{-1}R$ is obtained by adjoining inverses of the elements of S. Write $S^{-1}M$ for the corresponding localization of an R-module M. If $S = \{ a^n \mid n \in \mathbb{N} \}$, then we denote $S^{-1}M$ by M_a. Every divisor of an element of S is a unit in $S^{-1}R$. To see this, if $s = bc \in S$, then $(b/1) \cdot (c/s) = 1$. In particular, every divisor of a is a unit in R_a. For a prime ideal $\mathfrak{p}$, the set $R \setminus \mathfrak{p}$ is multiplicatively closed. The *localization* $M_\mathfrak{p}$ of M at $\mathfrak{p}$ is $(R \setminus \mathfrak{p})^{-1}M$. If R is a domain and $0 \notin S$, then the natural map $R \to S^{-1}R$ is injective. A ring is *local* if it has a unique maximal ideal. The units of a local ring are precisely the elements outside its maximal ideal. For a prime ideal $\mathfrak{p}$, the ring $R_\mathfrak{p}$ is local with maximal ideal $\mathfrak{p}R_\mathfrak{p}$.

The *fraction field* (or *field of fractions*) $\operatorname{Frac}(R)$ of a domain R is $R_{\langle 0 \rangle}$. For a domain R and an indeterminate z, write $R(z) \coloneqq \operatorname{Frac}(R[z])$. The map $\mathfrak{p} \mapsto S^{-1}\mathfrak{p}$ is an inclusion-preserving bijection from the primes of R that are disjoint from S to the primes of $S^{-1}R$. Its inverse sends a prime of $S^{-1}R$ to its preimage under the natural ring homomorphism $R \to S^{-1}R$ sending r to $r/1$. Localization commutes with taking quotients [AM69, Corollary 3.4(iii)].

Algebraic Independence. Let K be a field extension of L. A subset $T \subseteq K$ is *algebraically independent* over L if no finite tuple of distinct elements of T satisfies a nonzero polynomial over L. A *transcendence basis* of K over L is an algebraically independent subset $T \subseteq K$ such that $K/L(T)$ is algebraic. All transcendence bases have the same cardinality, the *transcendence degree* $\operatorname{tr.deg}_L K$ [Stacks, Tag 030D].

Dimension Theory of Rings. The *(Krull) dimension* of R, denoted $\dim R$, is the supremum of the integers d for which there exist prime ideals $\mathfrak{p}_0, \ldots, \mathfrak{p}_d$ satisfying $\mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \cdots \subsetneq \mathfrak{p}_d \subsetneq R$. The *height* of a prime ideal $\mathfrak{p}$ is defined by $\operatorname{ht}\mathfrak{p} \coloneqq \dim R_\mathfrak{p}$.

We record some useful facts about dimension and height.

Lemma 2.1: Let R be a ring.

1. For every proper ideal $I \subsetneq R$, we have $\dim R/I = \dim R/\sqrt{I}$.
2. If R is a domain of finite dimension and $I \subsetneq R$ is a nonzero ideal, then $\dim R/I \leq \dim R - 1$.
3. If R is a finitely generated integral domain over a field L (an *affine domain*), then $\dim R = \operatorname{tr.deg}_L \operatorname{Frac}(R)$ and $\dim R/\mathfrak{p} = \dim R - \operatorname{ht}\mathfrak{p}$ for every prime ideal $\mathfrak{p}$ of R [Har77, Theorem I.1.8A].
4. If R is Noetherian, $a_1, \ldots, a_r \in R$, and $\mathfrak{p}$ is a minimal prime of $\langle a_1, \ldots, a_r \rangle$, then $\operatorname{ht}\mathfrak{p} \leq r$ [Eis95, Theorems 10.1 and 10.2]. This statement is the algebraic form of *Krull's Height Theorem* (cf. Lemma 2.7).
5. If R is a Noetherian UFD, then every prime ideal of height 1 is generated by an irreducible element [Har77, Proposition I.1.12A].

Finiteness and Integrality. For a ring map $\phi\colon A \to B$, the A-algebra B is *finite* if it is finitely generated as an A-module and *of finite type* if it is finitely generated as an A-algebra. An element $b \in B$ is *integral* over A if it satisfies a monic polynomial with coefficients in $\phi(A)$. The map ϕ is *integral* if every element of B is integral over A. The *integral closure* of A in B is the subring of

elements of B integral over A. Integrality is transitive, and an A-algebra is finite if and only if it is both integral and of finite type [AM69, Corollaries 5.3–5.4 and the intervening remark]. A domain is *normal* if it is integrally closed in its fraction field; every UFD is normal. Every nonzero localization of a normal domain is normal [Stacks, Tag 00GY].

One reason integral ring extensions are useful is that they satisfy the following *lying-over* property.

Proposition 2.2: Let $\phi\colon A \to B$ be integral, with $B \neq 0$. Every prime ideal $\mathfrak{p} \subseteq A$ containing $\ker\phi$ is $\phi^{-1}(\mathfrak{q})$ for some prime ideal $\mathfrak{q} \subseteq B$. Such an ideal $\mathfrak{q}$ is maximal if and only if $\mathfrak{p}$ is maximal [AM69, Corollary 5.8 and Theorem 5.10]. Moreover, $\dim B = \dim(A/\ker\phi)$; in particular, $\dim B = \dim A$ when ϕ is injective [Eis95, Proposition 9.2].

Monomial Orders. For a nonzero polynomial f in one distinguished variable z, its *leading coefficient* $\operatorname{LC}_z(f)$ is the coefficient of the highest power of z. We write $\operatorname{LC}(f)$ when the variable is clear. Its *leading term* is that coefficient times the corresponding power of z. Fix a monomial order $\preceq$ on $\mathbb{F}[\mathbf{x}]$ [CLO25, Chapter 2, §2]. For $0 \neq f \in \mathbb{F}[\mathbf{x}]$, its *trailing monomial* $\operatorname{TM}(f)$ is the smallest monomial occurring in f, with its coefficient omitted. For nonzero $f, g \in \mathbb{F}[\mathbf{x}]$, we have $\operatorname{TM}(fg) = \operatorname{TM}(f)\operatorname{TM}(g)$.

Discrete Valuations. A *(normalized) discrete valuation* on a field K is a surjective homomorphism $v\colon K^\times \to \mathbb{Z}$ whose extension by $v(0) := \infty$ satisfies $v(a+b) \geq \min\{v(a), v(b)\}$ for all $a, b \in K$. Its *valuation ring* R is $\{a \in K \mid v(a) \geq 0\}$ and is a *discrete valuation ring* (DVR). The ring R has fraction field K and unit group $\{a \in K \mid v(a) = 0\}$. Its unique maximal ideal $\mathfrak{m}$ is $\{a \in K \mid v(a) > 0\}$. An element $\pi \in R$ is a *uniformizer* if $v(\pi) = 1$. For a fixed uniformizer π, we have $\mathfrak{m} = \langle \pi \rangle$, and every element of $K^\times$ can be written uniquely as $\pi^n u$ with $n \in \mathbb{Z}$ and $u \in R^\times$. Equivalently, a DVR is a local principal ideal domain that is not a field [AM69, Chapter 9, pp. 94–95]. Moreover, a Noetherian normal local domain of dimension 1 is a DVR [Stacks, Tag 00PD].

Power Series.

Definition 2.3: We write $R[[\mathbf{x}]]$ for the ring of *formal power series* $\sum_{\mathbf{j}\in\mathbb{N}^N} c_{\mathbf{j}} x_1^{j_1}\cdots x_N^{j_N}$ with $c_{\mathbf{j}} \in R$. For one variable z, let $R((z)) := R[[z]][z^{-1}]$. Its elements are *formal Laurent series* $\sum_{j\geq j_0} c_j z^j$ with $c_j \in R$ and $j_0 \in \mathbb{Z}$.

If $\mathbb{F}$ is a field, then $\mathbb{F}[[\mathbf{x}]]$ is a UFD [Lan02, Chapter IV, Theorem 9.3].

Let $g_1, \ldots, g_N \in R[[\mathbf{u}]]$ satisfy $g_i(\mathbf{0}) = 0$ for all $i \in [N]$. For every $f \in R[[\mathbf{x}]]$, the composition $f(g_1(\mathbf{u}), \ldots, g_N(\mathbf{u}))$ is well defined.

Remark 2.4: For a nonempty tuple $\mathbf{x}$, we have

$$\mathbb{F}[[\varepsilon]][\mathbf{x}]_\varepsilon = \mathbb{F}((\varepsilon))[\mathbf{x}], \qquad \mathbb{F}[[\varepsilon, \mathbf{x}]]_\varepsilon \subsetneq \mathbb{F}((\varepsilon))[[\mathbf{x}]].$$

The equality follows by clearing the finitely many coefficient denominators of a polynomial. The containment is strict because $\sum_{j\geq 0} \varepsilon^{-j} x_1^j$ has no uniform lower bound on its ε-exponents. ◁

For $f \in R((z))$ and $j \in \mathbb{Z}$, write $[z^j]f$ for the coefficient of z^j in f, taking this coefficient to be zero when the term is absent. For $f \in R[[\mathbf{x}]]$ and $\mathbf{a} \in \mathbb{Z}^N$, write $[\mathbf{x}^{\mathbf{a}}]f$ for the corresponding

coefficient, with value zero if any coordinate of $\mathbf{a}$ is negative. For $f \in R[[\mathbf{x}]]$ and $d \geq 0$, let $H_d(f) \in R[\mathbf{x}]$ be its *homogeneous component* of total degree d in $\mathbf{x}$. The substitution

$$f \mapsto f(z\mathbf{x}) = \sum_{d \geq 0} H_d(f) z^d \in R[\mathbf{x}][[z]]$$

is an injective R-algebra homomorphism, and $H_d(f) = [z^d] f(z \cdot \mathbf{x})$. A coefficient parameter such as ε in $R = \mathbb{F}[[\varepsilon]]$ does not contribute to total degree in $\mathbf{x}$. Evaluation at $\mathbf{0}$ sets only the other variables to zero; evaluation at $(0, \mathbf{0})$ also sets $\varepsilon = 0$, whenever this is defined.

A power series $g \in R[[\mathbf{x}]]$ is invertible if and only if $g(\mathbf{0}) \in R^\times$ [Mat86, Chapter 1, §1, p. 4]. Consequently, $\mathbb{F}((z))$ is a field. If R is Noetherian, then $\dim R[[x_1, \ldots, x_N]] = \dim R + N$ [Mat86, Theorem 15.4].

Definition 2.5: Let R be a domain. For $0 \neq h \in R((z))$, define

$$\mathrm{val}_z(h) := \min \left\{ j \in \mathbb{Z} \;\middle|\; [z^j]h \neq 0 \right\}, \qquad \mathrm{TC}(h) := [z^{\mathrm{val}_z(h)}]h.$$

The notation $\mathrm{TC}(h)$ denotes the *trailing coefficient* of h with respect to z. Set $\mathrm{val}_z(0) := \infty$. For nonzero f and g in $R((z))$, we have

$$\mathrm{val}_z(fg) = \mathrm{val}_z(f) + \mathrm{val}_z(g), \qquad \mathrm{val}_z(f+g) \geq \min\{\mathrm{val}_z(f), \mathrm{val}_z(g)\}.$$

Also, $\mathrm{TC}(fg) = \mathrm{TC}(f)\,\mathrm{TC}(g)$. If R is a field, say K, then val_z is a normalized discrete valuation on $K((z))$ and its valuation ring $K[[z]]$ is a DVR with uniformizer z. In this case, $\mathrm{val}_z\left(\frac{f}{g}\right) = \mathrm{val}_z(f) - \mathrm{val}_z(g)$ and $\mathrm{TC}\left(\frac{f}{g}\right) = \frac{\mathrm{TC}(f)}{\mathrm{TC}(g)}$.

For a domain R and $0 \neq f \in R[[\mathbf{x}]]$, let d_0 be the least d with $H_d(f) \neq 0$. Scaling then gives

$$\mathrm{val}_z(f(z\mathbf{x})) = d_0, \qquad \mathrm{TC}(f(z\mathbf{x})) = H_{d_0}(f).$$

2.2 Algebraic Geometry

We give basic definitions and results for the algebro-geometric tools we use. Much of this can be found in standard textbooks [BCS97; CLO25; Har77; HP94; Sha13] or online sources such as the Stacks project [Stacks].

Throughout Section 2.2, let $\mathbb{F}$ be an algebraically closed field unless otherwise stated. The N-dimensional *affine space* $\mathbb{A}^N$ over $\mathbb{F}$ is $\mathbb{F}^N$. For $S \subseteq \mathbb{F}[\mathbf{x}]$, let $\mathbf{V}(S) \subseteq \mathbb{A}^N$ be its common zero set; for $T \subseteq \mathbb{A}^N$, let $\mathbf{I}(T)$ be the ideal of polynomials vanishing on T. A *hypersurface* is the zero set of a nonconstant polynomial. The sets $\mathbf{V}(S)$ are the *affine varieties* and form the closed sets of the *Zariski topology*. We allow varieties to be reducible or empty. The *Zariski closure* $\overline{T}$ of T is $\mathbf{V}(\mathbf{I}(T))$, the smallest closed subset of $\mathbb{A}^N$ containing T. A nonempty variety is *irreducible* if it is not the union of two proper closed subsets. Every affine variety V is a finite union $V_1 \cup \cdots \cup V_r$ of irreducible affine varieties with $V_i \not\subseteq V_j$ for $i \neq j$. These varieties are the *irreducible components* of V, and the decomposition is unique up to order [CLO25, Chapter 4, §6, Theorem 4].

An *affine open subset* of a variety is an open subset isomorphic to an affine variety. A property

holds *generally* or *generically* on a variety V if it holds on a dense open subset of V. In an irreducible variety, every nonempty open subset is dense and irreducible, and every finite intersection of nonempty open subsets is nonempty [Har77, Chapter I, Exercise 1.6]. In particular, if V is irreducible and $g_1, \dots, g_r \notin \mathbf{I}(V)$, then $\{\, b \in V \mid g_1(b) \cdots g_r(b) \neq 0 \,\}$ is a nonempty dense open subset of V.

For $S \subseteq [N]$, write $\mathbb{A}^S$ for the affine space with coordinates $\mathbf{x}_S$ and $\pi_S \colon \mathbb{A}^N \to \mathbb{A}^S$ for the coordinate projection $a \mapsto (a_s)_{s \in S}$. For $T \subseteq [N]$ and $S \subseteq \mathbb{F}[\mathbf{x}_T]$, the notation $\mathbf{V}(S)$ denotes the zero set of S in $\mathbb{A}^N$, not $\mathbb{A}^T$, unless otherwise noted.

For every ideal $J \subseteq \mathbb{F}[\mathbf{x}]$, the *Nullstellensatz* gives

$$\mathbf{I}(\mathbf{V}(J)) = \sqrt{J}, \qquad \mathbf{V}(J) = \emptyset \iff J = \mathbb{F}[\mathbf{x}].$$

Thus $\mathbf{V}$ and $\mathbf{I}$ give mutually inverse, inclusion-reversing bijections between radical ideals of $\mathbb{F}[\mathbf{x}]$ and affine varieties in $\mathbb{A}^N$, with prime ideals corresponding to irreducible varieties [CLO25, Chapter 4, §2, Theorems 6–7, §5, Corollary 4]. For every proper ideal J, the irreducible components of $\mathbf{V}(J)$ are precisely $\mathbf{V}(\mathfrak{p})$ as $\mathfrak{p}$ ranges over the minimal primes of J [Har77, Corollary I.1.4 and Proposition I.1.5].

For an affine variety V, define its *coordinate ring* by $\mathbb{F}[V] \coloneqq \mathbb{F}[\mathbf{x}] / \mathbf{I}(V)$. For $f \in \mathbb{F}[\mathbf{x}]$, define its restriction to V by $f|_V \coloneqq f + \mathbf{I}(V)$. If V is irreducible, then its *function field* $\mathbb{F}(V)$ is $\mathrm{Frac}(\mathbb{F}[V])$. The coordinates $\mathbf{x}_S$ are *algebraically independent on* V if the family $(x_s|_V)_{s \in S}$ is algebraically independent over $\mathbb{F}$. The coordinates $\mathbf{x}_S$ are algebraically independent on V precisely when $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_S] = \{\, 0 \,\}$. In this case we identify $\mathbb{F}[\mathbf{x}_S]$ with its image in $\mathbb{F}[V]$.

Lemma 2.6: Let $V \subseteq \mathbb{A}^N$ be an affine variety and let $q \in \mathbb{F}[\mathbf{x}]$. The restriction $q|_V$ is a unit of $\mathbb{F}[V]$ if and only if $V \cap \mathbf{V}(q) = \emptyset$.

Proof: If $q|_V$ is a unit, then evaluating an inverse at each point of V shows that q does not vanish there. Conversely, if $\mathbf{V}(\mathbf{I}(V) + \langle\, q \,\rangle) = \emptyset$, then the Nullstellensatz gives $1 = a + qh$ for some $a \in \mathbf{I}(V)$ and $h \in \mathbb{F}[\mathbf{x}]$. Restricting to V gives $(q|_V)(h|_V) = 1$. □

Morphisms. Let $f_1, \dots, f_M \in \mathbb{F}[x_1, \dots, x_N]$. These polynomials define a *polynomial map* $f \colon \mathbb{A}^N \to \mathbb{A}^M$ by $f(\mathbf{a}) \coloneqq (f_1(\mathbf{a}), \dots, f_M(\mathbf{a}))$. If $V \subseteq \mathbb{A}^N$ and $V' \subseteq \mathbb{A}^M$ are varieties with $f(V) \subseteq V'$, then the restriction $f \colon V \to V'$ is a *morphism* or *polynomial map*. Polynomial maps are continuous in the Zariski topology. A morphism is *dominant* if its image is a dense subset of the codomain. A morphism $f \colon V \to V'$ induces a map on coordinate rings $\mathbb{F}[V'] \to \mathbb{F}[V]$ by $\phi \mapsto \phi \circ f$. For nonaffine varieties, morphisms are defined locally on affine open subsets in the same way.

Dimension Theory of Varieties. The *dimension* of a nonempty irreducible affine variety V is the largest integer d for which there exist irreducible affine varieties $V_0, \dots, V_d$ satisfying $\emptyset \subsetneq V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_d = V$. The dimension $\dim V$ of a nonempty affine variety V is the maximum dimension of its irreducible components. A nonempty affine variety V is *equidimensional* of dimension d if all its irreducible components have dimension d.

In particular, $\dim \mathbb{A}^N = \dim \mathbb{F}[\mathbf{x}] = \mathrm{tr.deg}_{\mathbb{F}} \mathbb{F}(\mathbf{x}) = N$. For a nonempty variety $U \subseteq \mathbb{A}^N$, the *codimension* of U in $\mathbb{A}^N$ is the quantity $\mathrm{codim}_{\mathbb{A}^N} U \coloneqq \dim \mathbb{A}^N - \dim U = N - \dim U$. When N is

clear from context, we write $\operatorname{codim} U$ for $\operatorname{codim}_{\mathbb{A}^N} U$.

We give some useful facts on the dimension of varieties.

Lemma 2.7 (Geometric dimension facts): Let $\mathbb{F}$ be an algebraically closed field.

1. For every proper ideal $I \subsetneq \mathbb{F}[\mathbf{x}]$, we have $\dim \mathbf{V}(I) = \dim \mathbb{F}[\mathbf{x}]/I$ [Har77, Proposition I.1.7 and Theorem I.1.8A].
2. Let $X \subseteq \mathbb{A}^N$ be an irreducible affine variety and let $q_1, \ldots, q_r \in \mathbb{F}[\mathbf{x}]$. Suppose that $X \cap \mathbf{V}(q_1, \ldots, q_r)$ is nonempty, and let Z be an irreducible component of this intersection. The component Z satisfies $\dim Z \geq \dim X - r$. If, in addition, $r = 1$ and q_1 does not vanish identically on X, then $\dim Z = \dim X - 1$. In particular, every irreducible component of $\mathbf{V}(q_1, \ldots, q_r)$ has codimension at most r [Sha13, Chapter 1, § 6.2, Corollaries 1.13–1.14 and the remark following Corollary 1.13]. This bound is the geometric form of *Krull's Height Theorem*.

Degree and Hilbert Functions. We use two closely related notions of *degree*. Degree bounds are widely used in algebraic complexity theory [BCS97, Chapter 8].

Let L be a field and $I \subsetneq L[\mathbf{x}]$ a proper ideal. Its *(affine) Hilbert function*

$$\operatorname{HilbF}_I(s) := \dim_L L[\mathbf{x}]_{\leq s}/I_{\leq s}$$

agrees for sufficiently large s with the *(affine) Hilbert polynomial* HilbP_I of degree $d = \dim L[\mathbf{x}]/I$. We define the *degree* of I to be the quantity

$$\operatorname{DEG} I := d! \cdot \operatorname{LC}(\operatorname{HilbP}_I).$$

See [CLO25, Chapter 9, §3]. The dimension formula follows from homogenization and [Eis95, Corollary 13.7]. For every nonconstant polynomial $Q \in L[\mathbf{x}]$, the *hypersurface identity* gives $\operatorname{DEG}\langle Q \rangle = \deg Q$. If J is an ideal contained in I, then $\operatorname{HilbF}_I(s) \leq \operatorname{HilbF}_J(s)$ [CLO25, Chapter 9, §3, Exercise 8(b)]. We compare degrees through Hilbert functions, as in the proof of Lemma 4.8.

Lemma 2.8: Let L and L' be fields, and let $I \subseteq L[\mathbf{x}]$ and $J \subseteq L'[\mathbf{y}]$ be proper ideals, where the tuples $\mathbf{x}$ and $\mathbf{y}$ may have different lengths. If $\dim L[\mathbf{x}]/I = \dim L'[\mathbf{y}]/J$ and $\operatorname{HilbF}_J(s) \leq \operatorname{HilbF}_I(s)$ for all sufficiently large s, then $\operatorname{DEG} J \leq \operatorname{DEG} I$.

Proof: Both Hilbert polynomials have degree t, the common dimension (see Degree and Hilbert Functions). Since $\operatorname{HilbP}_J(s) \leq \operatorname{HilbP}_I(s)$ for all large s, the coefficient of s^t in HilbP_J is at most that in HilbP_I. Multiplying by $t!$ gives $\operatorname{DEG} J \leq \operatorname{DEG} I$. □

When $\mathbb{F}$ is algebraically closed, an irreducible affine variety V has *degree* $\deg V := \operatorname{DEG} \mathbf{I}(V)$. For a nonempty affine variety X with irreducible components $V_1, \ldots, V_t$, define

$$\deg X := \sum_{i=1}^{t} \deg V_i.$$

This definition agrees with the degree convention of [BCS97, Definition 8.23]. In particular,

$\deg X = \operatorname{DEG} \mathbf{I}(X)$ when X is equidimensional; in general, $\operatorname{DEG} \mathbf{I}(X)$ is the sum of the degrees of the irreducible components of largest dimension. When X is equidimensional, $\deg X$ is the number of points in its intersection with a general affine linear subspace of $\mathbb{A}^N$ of codimension $\dim X$. Every nonempty union of irreducible components of X has degree at most $\deg X$.

Proposition 2.9 (Degree Bounds): Let $\mathbb{F}$ be algebraically closed and let $X \subseteq \mathbb{A}^N$ be a nonempty affine variety.

1. For every affine linear automorphism $\phi\colon \mathbb{A}^N \to \mathbb{A}^N$, we have $\deg \phi(X) = \deg X$.
2. For every $S \subseteq [N]$, the coordinate projection π_S satisfies $\deg \overline{\pi_S(X)} \leq \deg X$. For every nonnegative integer r, we have $\deg(X \times \mathbb{A}^r) = \deg X$.
3. We have $\deg \mathbb{A}^N = 1$.
4. Let $q \in \mathbb{F}[\mathbf{x}]$ be nonconstant, and let $p_1, \ldots, p_t$ represent its distinct irreducible factors up to scalars. The irreducible components of $\mathbf{V}(q)$ are $\mathbf{V}(p_1), \ldots, \mathbf{V}(p_t)$, so $\deg \mathbf{V}(q) = \sum_{j=1}^{t} \deg p_j \leq \deg q$.
5. Let $q_1, \ldots, q_r \in \mathbb{F}[\mathbf{x}]$ be nonconstant and set $Y \coloneqq X \cap \mathbf{V}(q_1, \ldots, q_r)$. If Y is nonempty, then

$$\deg Y \leq \deg X \prod_{i=1}^{r} \deg \mathbf{V}(q_i) \leq \deg X \prod_{i=1}^{r} \deg q_i.$$

   The same upper bounds hold for the degree of any nonempty union of irreducible components of Y. These bounds are a form of *Bézout's Inequality*.

These facts follow from [BCS97, Remarks 8.20 and 8.24, Example 8.25, Corollary 8.27, Theorem 8.28, Lemma 8.32, and Remark 8.33].

We use the following consequence of the Bézout inequalities for polynomial ideals [Laz20].

Lemma 2.10: Let L be a field, and let $q_1, \ldots, q_r \in L[\mathbf{x}]$ be nonconstant polynomials of degree at most δ. If $\mathfrak{p}$ is a minimal prime of $\langle q_1, \ldots, q_r \rangle$, then

$$\operatorname{DEG} \mathfrak{p} \leq \delta^{\operatorname{ht} \mathfrak{p}} \leq \delta^r.$$

Proof: The case $r = 0$ is immediate. Otherwise, the first inequality follows from [Laz20, Corollary 19 and Lemma 8], whose notion of degree agrees with DEG by [Laz20, §5.2 and Lemma 13]. The second inequality follows from Lemma 2.1. □

2.3 Algebraic Complexity

We assume familiarity with the basics of algebraic complexity theory; see, e.g., the surveys of Dutta and Lysikov [DL25], Saptharishi [Sap26], and Shpilka and Yehudayoff [SY10] and the text of Bürgisser, Clausen, and Shokrollahi [BCS97].

We define the *size* of an algebraic circuit to be the number of gates plus the number of wires, allowing scalar multiplications along wires for free. Because we work with constant-depth circuits, we allow unbounded fan-in and fan-out.

We now define our model of depth-4 circuits with bounded top and bottom fan-in.

Definition 2.11: Let k and δ be at least 1. A $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit over R is a depth-4 circuit whose layers, from top to bottom, consist of one addition gate of fan-in at most k, product gates, addition gates, and product gates of fan-in at most δ. The gates in the middle two layers have unrestricted fan-in. Such a circuit computes a polynomial of the form $\sum_{i=1}^{k} T_i(\mathbf{x})$, where $T_i(\mathbf{x}) := \prod_j q_{i,j}(\mathbf{x})$ and each $q_{i,j} \in R[\mathbf{x}]$ has degree at most δ. A circuit of size s has at most s distinct polynomials $q_{i,j}$. The circuit is *homogeneous* if every gate computes a polynomial that is homogeneous in $\mathbf{x}$.

A polynomial $f \in R[\mathbf{x}]$ is *constant-free* if $f(\mathbf{0}) = 0$. In particular, every nonconstant homogeneous form is constant-free.

We also consider approximation by algebraic circuits in the following sense.

Definition 2.12: A polynomial $f \in \mathbb{F}[\mathbf{x}]$ is *border computed by* a circuit class $\mathcal{C}$ if some circuit $\Phi \in \mathcal{C}$ over $\mathbb{F}((\varepsilon))$ computes $f + \varepsilon S$ for some $S \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$. Equivalently, f is *approximated by* $\mathcal{C}$, *approximately computed by* $\mathcal{C}$, or *in the closure of* $\mathcal{C}$. Write $\overline{\mathcal{C}}$ for the closure of $\mathcal{C}$. The *size* of this border computation is the size of Φ.

Over an algebraically closed field $\mathbb{F}$, this definition of closure agrees with Zariski closure for the class of circuits of size at most s, for every s; see [BCS97, Lemma 20.28] and [DL25, Theorem 38].

2.4 Auxiliary Statements

This subsection collects some statements from algebra and algebraic geometry that will be useful in subsequent sections. Readers may skip this subsection on a first reading and refer back to it as needed.

Throughout Section 2.4, we assume that $\mathbb{F}$ is algebraically closed. We first state some useful facts on coordinate projections.

Lemma 2.13: Let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety, $S \subseteq [N]$, and $W_S := \overline{\pi_S(V)} \subseteq \mathbb{A}^S$.

1. $\mathbf{I}(W_S) = \mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_S]$, and $f \mapsto f|_V$ induces an isomorphism $\mathbb{F}[W_S] \cong \mathbb{F}[\mathbf{x}_S|_V]$. In particular, W_S is irreducible.
2. $\dim W_S = \operatorname{tr.deg}_{\mathbb{F}} \mathbb{F}(\mathbf{x}_S|_V) \leq \dim V$, where $\mathbb{F}(\mathbf{x}_S|_V)$ is the subfield of $\mathbb{F}(V)$ generated by the $x_s|_V$ for $s \in S$.
3. If $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_S] = \{0\}$, then there is a set T with $S \subseteq T \subseteq [N]$, $|T| = \dim V$, and $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_T] = \{0\}$. For any such T, the family $(x_t|_V)_{t \in T}$ is a transcendence basis of $\mathbb{F}(V)$ over $\mathbb{F}$.

Proof: The map $\mathbb{F}[\mathbf{x}_S] \to \mathbb{F}[V]$ has image $\mathbb{F}[\mathbf{x}_S|_V]$ and kernel $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_S]$, which is equal to $\mathbf{I}(W_S)$. This image is a domain, proving Item 1.

Item 2 follows from the dimension–transcendence degree formula in Lemma 2.1.

For Item 3, extend $(x_s|_V)_{s \in S}$ to a maximal algebraically independent subfamily $(x_t|_V)_{t \in T}$ of the coordinate functions. Every remaining coordinate is algebraic over $\mathbb{F}(\mathbf{x}_T|_V)$, so this family is a transcendence basis and $|T| = \dim V$. □

We next record two consequences of the lying-over lemma (Proposition 2.2).

Lemma 2.14: Let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety and $T \subseteq [N]$ with $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_T] = \{0\}$. Let $Q \in \mathbb{F}[\mathbf{x}_T]$ be nonzero, let $h := Q|_V$, and suppose that $\mathbb{F}[V]_h$ is integral over $\mathbb{F}[\mathbf{x}_T|_V]_h$. If $\mathfrak{p} \subseteq \mathbb{F}[\mathbf{x}_T]$ is a prime ideal with $Q \notin \mathfrak{p}$, then $V \cap \mathbf{V}(\mathfrak{p})$ has an irreducible component V_0 with $\mathbf{I}(V_0) \cap \mathbb{F}[\mathbf{x}_T] = \mathfrak{p}$.

Proof: Identify $A := \mathbb{F}[\mathbf{x}_T]$ with its image in $B := \mathbb{F}[V]$. By Proposition 2.2, there is a prime of B_h contracting to $\mathfrak{p}A_Q$. Its preimage P in $\mathbb{F}[\mathbf{x}]$ contains $\mathbf{I}(V) + \mathfrak{p}\mathbb{F}[\mathbf{x}]$ and satisfies $P \cap A = \mathfrak{p}$. Choose a minimal prime P_0 of $\mathbf{I}(V) + \mathfrak{p}\mathbb{F}[\mathbf{x}]$ contained in P. Since $\mathfrak{p} \subseteq P_0 \cap A \subseteq P \cap A = \mathfrak{p}$, we may take $V_0 := \mathbf{V}(P_0)$. □

Corollary 2.15: Let V, T, Q, and h be as in Lemma 2.14. For every $c \in \mathbb{A}^T$ with $Q(c) \neq 0$, there exists $v \in V$ such that $\pi_T(v) = c$.

Proof: The maximal ideal $\mathfrak{m}_c := \langle x_t - c_t \mid t \in T \rangle$ does not contain Q. Thus Lemma 2.14 implies that $V \cap \mathbf{V}(\mathfrak{m}_c)$ is nonempty, and every point of this intersection projects to c. □

Recall from Lemma 2.13 that for any irreducible affine variety $V \subseteq \mathbb{A}^N$ of dimension m, there exists $T \subseteq [N]$ with $|T| = m$ such that $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_T] = \{0\}$. The ring $\mathbb{F}[V]$ need not be integral over its subring $\mathbb{F}[\mathbf{x}_T|_V]$. The next two lemmas show that integrality holds after localizing both rings at a suitable nonzero polynomial $Q \in \mathbb{F}[\mathbf{x}_T]$ and give a bound on $\deg Q$.

Lemma 2.16: Let L be a field and let $\mathfrak{p} \subseteq L[\mathbf{x}]$ be a prime ideal. Let $T \subseteq [N]$ with $|T| = \dim L[\mathbf{x}]/\mathfrak{p}$ and $\mathfrak{p} \cap L[\mathbf{x}_T] = \{0\}$. For every $a \in [N] \setminus T$, the prime ideal $\mathfrak{p} \cap L[\mathbf{x}_T, x_a]$ is generated by one irreducible polynomial Q with $\deg_{x_a} Q \geq 1$ and $\deg Q \leq \mathrm{DEG}\,\mathfrak{p}$.

Proof: Fix $a \in [N] \setminus T$. Let $d := |T|$, $R := L[\mathbf{x}_T, x_a]$, and $\mathfrak{q} := \mathfrak{p} \cap R$.

The ideal $\mathfrak{q}$ is prime, and the hypotheses give inclusions of domains

$$L[\mathbf{x}_T] \hookrightarrow R/\mathfrak{q} \hookrightarrow L[\mathbf{x}]/\mathfrak{p}.$$

Indeed, the kernel of the first map is $\mathfrak{q} \cap L[\mathbf{x}_T] = \mathfrak{p} \cap L[\mathbf{x}_T] = \{0\}$ by assumption. The kernel of the second map is $(\mathfrak{p} \cap R)/\mathfrak{q} = \mathfrak{q}/\mathfrak{q} = \{0\}$. The fraction field of $L[\mathbf{x}_T]$ has transcendence degree d over L. The same holds for the fraction field of $L[\mathbf{x}]/\mathfrak{p}$ since $\dim L[\mathbf{x}]/\mathfrak{p} = |T|$. It follows that $\mathrm{Frac}(R/\mathfrak{q})$ also has transcendence degree d over L, and hence $\dim R/\mathfrak{q} = d$ by Lemma 2.1. Since R is a polynomial ring in $d+1$ variables, Lemma 2.1 gives $\mathrm{ht}\,\mathfrak{q} = 1$. Since R is a Noetherian UFD, Lemma 2.1 shows that $\mathfrak{q} = \langle Q \rangle$ for an irreducible polynomial $Q \in R$. Since Q is nonzero and $\mathfrak{p} \cap L[\mathbf{x}_T] = \{0\}$, it must involve x_a, so $\deg_{x_a} Q \geq 1$.

For every $s \geq 0$, inclusion of polynomial rings induces an injective L-linear map $R_{\leq s}/\mathfrak{q}_{\leq s} \hookrightarrow L[\mathbf{x}]_{\leq s}/\mathfrak{p}_{\leq s}$, since $\mathfrak{q} = \mathfrak{p} \cap R$. Hence $\mathrm{HilbF}_{\mathfrak{q}}(s) \leq \mathrm{HilbF}_{\mathfrak{p}}(s)$. Both $R/\mathfrak{q}$ and $L[\mathbf{x}]/\mathfrak{p}$ have dimension d, so Lemma 2.8 and the hypersurface identity $\mathrm{DEG}\,\langle Q \rangle = \deg Q$ in Degree and Hilbert Functions give

$$\deg Q = \mathrm{DEG}\,\langle Q \rangle = \mathrm{DEG}\,\mathfrak{q} \leq \mathrm{DEG}\,\mathfrak{p}.$$ □

Lemma 2.17: Let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety of dimension m and degree Δ, and

let $T \subseteq [N]$ with $|T| = m$ and $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_T] = \{0\}$. For each $a \in [N] \setminus T$, there is an irreducible polynomial $Q_a \in \mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_{T \cup \{a\}}]$ with $\deg Q_a \leq \Delta$ and $d_a := \deg_{x_a} Q_a \geq 1$. For each a, let $c_a \in \mathbb{F}[\mathbf{x}_T]$ be the leading coefficient of Q_a as a polynomial in x_a, and let

$$Q := \prod_{a \in [N] \setminus T} c_a \in \mathbb{F}[\mathbf{x}_T],$$

with $Q = 1$ if $T = [N]$. Each c_a is nonzero with $\deg c_a \leq \Delta - 1$, so $Q \neq 0$ and $\deg Q \leq (N-m)(\Delta-1)$. Moreover, $\mathbb{F}[V]_{Q|_V}$ is integral over $\mathbb{F}[\mathbf{x}_T|_V]_{Q|_V}$.

Proof: Since $\mathbf{I}(V)$ is prime with $\dim \mathbb{F}[\mathbf{x}]/\mathbf{I}(V) = m$ and $\mathrm{DEG}\,\mathbf{I}(V) = \Delta$, Lemma 2.16 gives the irreducible polynomials Q_a with $\deg Q_a \leq \Delta$ and $\deg_{x_a} Q_a \geq 1$ for each $a \in [N] \setminus T$. For each $a \in [N] \setminus T$, the leading coefficient of Q_a with respect to x_a satisfies $c_a \neq 0$ and $\deg_{\mathbf{x}_T} c_a \leq \deg Q_a - d_a \leq \Delta - 1$. Thus, Q is a product of $N - m$ nonzero polynomials of degree at most $\Delta - 1$, so $Q \neq 0$ and $\deg Q \leq (N-m)(\Delta-1)$.

Let $B := \mathbb{F}[\mathbf{x}_T|_V]$ and $h := Q|_V$. The hypothesis $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_T] = \{0\}$ ensures that $h \neq 0$. For every $a \in [N] \setminus T$, the element $c_a|_V$ divides h in B, so it is a unit in B_h by the observation in Localization. Dividing the relation $Q_a(\mathbf{x}_T|_V, x_a|_V) = 0$ by $c_a|_V$ gives a monic equation for $x_a|_V$ over B_h. These integral coordinate functions generate $\mathbb{F}[V]_h$ as a B_h-algebra, so $\mathbb{F}[V]_h$ is integral over B_h. □

We use the notation for coordinates and projections from Section 2. The following lemma gives a small exceptional set E of coordinate indices whenever the coordinates $\mathbf{x}_I$ are algebraically independent on an irreducible affine variety V. For every set J disjoint from $I \cup E$, setting $\mathbf{x}_J$ to zero preserves the algebraic independence of $\mathbf{x}_I$ on some irreducible component of $V \cap \mathbf{V}(\mathbf{x}_J)$.

Lemma 2.18: Let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety of degree Δ and codimension C, where $C = N - \dim V$. Let $I \subseteq [N]$ be such that the coordinates $\mathbf{x}_I$ are algebraically independent on V, that is, $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_I] = \{0\}$. There is a set $E \subseteq [N] \setminus I$ with $|E| \leq C\Delta$ having the following property: for every $J \subseteq [N] \setminus (I \cup E)$, the set $V \cap \mathbf{V}(\mathbf{x}_J)$ has an irreducible component V_J with $\mathbf{I}(V_J) \cap \mathbb{F}[\mathbf{x}_I] = \{0\}$.

Proof: Let $m := \dim V = N - C$. By Lemma 2.13, there is a set T with $I \subseteq T \subseteq [N]$, $|T| = m$, and $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_T] = \{0\}$. By Lemma 2.17, there is a nonzero $Q \in \mathbb{F}[\mathbf{x}_T]$ with $\deg Q \leq (N-m)(\Delta-1) = C(\Delta-1)$ such that $\mathbb{F}[V]_{Q|_V}$ is integral over $\mathbb{F}[\mathbf{x}_T|_V]_{Q|_V}$.

Fix a monomial M that occurs in Q with a nonzero coefficient, and let $S := \mathrm{supp}(M) \subseteq T$. Hence $|S| \leq \deg Q$. Let

$$E := ([N] \setminus T) \cup (S \setminus I).$$

Since $I \subseteq T$, the set E is disjoint from I, and

$$|E| \leq |[N] \setminus T| + |S| \leq C + C(\Delta - 1) = C\Delta.$$

Let $J \subseteq [N] \setminus (I \cup E) = T \setminus (I \cup S)$. The ideal $\mathfrak{p}_J := \langle x_j \mid j \in J \rangle \subseteq \mathbb{F}[\mathbf{x}_T]$ is prime, since its quotient ring is the domain $\mathbb{F}[\mathbf{x}_{T \setminus J}]$. Setting $x_j = 0$ for all $j \in J$ deletes the monomials of Q involving these variables and leaves the others unchanged. Since $J \cap S = \emptyset$, the monomial M survives with its

nonzero coefficient, so $Q \notin \mathfrak{p}_J$.

Finally, by Lemma 2.14, $V \cap \mathbf{V}(\mathbf{x}_J)$ has an irreducible component V_J with $\mathbf{I}(V_J) \cap \mathbb{F}[\mathbf{x}_T] = \mathfrak{p}_J$. Since I and J are disjoint subsets of T, it follows that $\mathbf{I}(V_J) \cap \mathbb{F}[\mathbf{x}_I] = \mathfrak{p}_J \cap \mathbb{F}[\mathbf{x}_I] = \langle\, x_j \mid j \in J \,\rangle \cap \mathbb{F}[\mathbf{x}_I] = \{\, 0 \,\}$, as desired. □

We show that if an irreducible affine variety V has sufficiently small codimension, then the coordinates in at least one block are algebraically independent on V.

Lemma 2.19: Let k and n be at least 1, and set $N := (k+1)n$. Let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety with $\operatorname{codim} V \leq k$. There is an index $j \in [k+1]$ with $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_j] = \{\, 0 \,\}$.

Proof: By Lemma 2.13, applied with $S = \emptyset$, there is a set $T \subseteq [N]$ with $|T| = \dim V$ and $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_T] = \{\, 0 \,\}$. Suppose for a contradiction that $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_j] \neq \{\, 0 \,\}$ for every $j \in [k+1]$. No block B_j is contained in T, since $\mathbb{F}[\mathbf{x}_j] \subseteq \mathbb{F}[\mathbf{x}_T]$ whenever $B_j \subseteq T$. Hence $\left|T \cap B_j\right| \leq n-1$ for every j. Therefore, $\dim V = |T| \leq (k+1)(n-1) < (k+1)n - k \leq \dim V$, a contradiction. □

The next lemma shows how to eliminate the other block monomials of $P_{k+1,n}$ while preserving the algebraic independence of $\mathbf{x}_j$ on an irreducible component.

Lemma 2.20: Let k and n be at least 1. Let $N := (k+1)n$, and let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety of degree Δ and codimension C such that $C\Delta < n$. Let $j \in [k+1]$. Suppose $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_j] = \{\, 0 \,\}$. There exist indices $\ell_1, \ldots, \ell_k \in [N] \setminus B_j$ such that $V \cap \bigcap_{i=1}^{k} \mathbf{V}(x_{\ell_i})$ has an irreducible component V_0 with $\mathbf{I}(V_0) \cap \mathbb{F}[\mathbf{x}_j] = \{\, 0 \,\}$. Moreover, $P_{k+1,n} \equiv \mathbf{x}^{B_j} \pmod{\mathbf{I}(V_0)}$.

Proof: Apply Lemma 2.18 with $I = B_j$, and let $E \subseteq [N] \setminus I$ be the resulting set of size at most $C\Delta$. Choose one index from $B_{j'} \setminus E$ for each $j' \in [k+1] \setminus \{\, j \,\}$, and let J be the set of chosen indices. Each choice is possible because $|E| \leq C\Delta < n$.

Write $J = \{\, \ell_1, \ldots, \ell_k \,\}$. By Lemma 2.18, there exists an irreducible component V_0 of $V \cap \bigcap_{i=1}^{k} \mathbf{V}(x_{\ell_i})$ such that $\mathbf{I}(V_0) \cap \mathbb{F}[\mathbf{x}_j] = \{\, 0 \,\}$. Moreover, $\{\, x_{\ell_1}, \ldots, x_{\ell_k} \,\}$ contains one variable from each term of $P_{k+1,n}$ (Definition 1.2) except the j-th term. Thus $P_{k+1,n} \equiv \mathbf{x}^{B_j} \pmod{\mathbf{I}(V_0)}$. □

We now construct a variety on which each product gate either vanishes or has every factor invertible. Each equation we impose eliminates a gate, increases the codimension by at most one, and increases the degree by a factor of at most δ.

Lemma 2.21: Let $\mathbb{F}$ be an algebraically closed field of arbitrary characteristic. Let k and δ be at least 1, and let $T_i := \prod_j q_{i,j}$ for $i \in [k]$, where each product is finite and each $q_{i,j} \in \mathbb{F}[\mathbf{x}]$ has degree at most δ. There exist an irreducible affine variety $V \subseteq \mathbb{A}^N$ and a set $K \subseteq [k]$ with the following properties:

1. $\operatorname{codim}_{\mathbb{A}^N} V \leq k - |K|$ and $\deg V \leq \delta^{k-|K|}$,
2. $T_i \in \mathbf{I}(V)$ for every $i \in [k] \setminus K$, and
3. $q_{i,j}|_V$ is a unit of $\mathbb{F}[V]$ for every $i \in K$ and every j.

In particular, $\sum_{i=1}^{k} T_i \equiv \sum_{i \in K} T_i \pmod{\mathbf{I}(V)}$.

Proof: We construct irreducible affine varieties $V_0, \dots, V_k$ satisfying $\mathbb{A}^N = V_0 \supseteq V_1 \supseteq \cdots \supseteq V_k$ and decide at each stage i whether to add i to K. We maintain the bounds $\operatorname{codim} V_i \leq i - |K|$ and $\deg V_i \leq \delta^{i-|K|}$. These bounds hold for $i = 0$ with $K = \emptyset$, since $\deg \mathbb{A}^N = 1$ by Proposition 2.9. For each $i \in [k]$, we construct V_i from V_{i-1} as follows.

- Suppose that some factor $q_{i,j}$ of T_i satisfies $V_{i-1} \cap \mathbf{V}(q_{i,j}) \neq \emptyset$. Fix such an index j and set $q_i := q_{i,j}$. Leave i out of K and choose an irreducible component V_i of $V_{i-1} \cap \mathbf{V}(q_i)$. If q_i vanishes identically on V_{i-1}, then $V_i = V_{i-1}$. Otherwise, $\dim V_i = \dim V_{i-1} - 1$ by Lemma 2.7. Furthermore, $\deg V_i \leq \delta \deg V_{i-1}$ by Proposition 2.9.
- Otherwise, add i to K and let $V_i := V_{i-1}$.

Set $V := V_k$. The invariants give Item 1. If $i \notin K$, then the chosen factor q_i of T_i vanishes on $V_i \supseteq V$, so $T_i \in \mathbf{I}(V)$. If $i \in K$, then no factor $q_{i,j}$ of T_i has a zero on $V_{i-1} \supseteq V$, so $q_{i,j}|_V$ is a unit by Lemma 2.6. □

3 Impossibility Under a Constant-Freeness Assumption

In this section, we consider a warm-up case under the following assumption: Each top-level product gate of a $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit computes either a constant or a polynomial divisible by a constant-free factor of degree at most δ. We show that, for sufficiently large n, the polynomial $P_{k+1,n}$ cannot be computed by any such circuit. In particular, this result rules out homogeneous $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits. Throughout, assume that the base field $\mathbb{F}$ is algebraically closed.

In the next section, we extend this result to show that $P_{k+1,n}$ cannot even be computed in the border setting by $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits under the same constant-freeness assumption. In Section 7, we strengthen the result in another way by removing the constant-freeness assumption in the non-border setting.[1]

Theorem 3.1: Suppose $\delta \geq 1$ and $k \geq 1$. If $n > k\delta^k$, then $P_{k+1,n}$ cannot be computed by a $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit of the form $\sum_{i=1}^{k} T_i$ in which each T_i is either a constant or divisible by a constant-free factor q_i of degree at most δ. In particular, $P_{k+1,n}$ cannot be computed by homogeneous $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits.

Proof: Let $N := (k+1)n$. Suppose towards a contradiction that $P_{k+1,n}$ can be written as $\sum_{i=1}^{k} T_i$, where each T_i is either a constant or divisible by a constant-free factor q_i of degree at most δ. After reindexing, let $T_1, \dots, T_{k'}$ be the nonconstant summands, with $k' \leq k$. For each $i \in [k']$, choose a constant-free polynomial q_i of degree at most δ dividing T_i. Since $T_i \neq 0$, each q_i is nonzero and hence nonconstant. Moreover, $T_i(\mathbf{0}) = 0$ because $q_i(\mathbf{0}) = 0$. Evaluating at $\mathbf{x} = \mathbf{0}$ and using $P_{k+1,n}(\mathbf{0}) = 0$, we see that the constant summands satisfy $\sum_{i=k'+1}^{k} T_i = 0$. Discarding these summands, we obtain

$$P_{k+1,n} = \sum_{i=1}^{k'} T_i.$$

[1] We remark that for $k \geq 2$, one cannot simultaneously remove the constant-freeness assumption and establish the impossibility result in the border setting. Indeed, if $k \geq 2$, then $P_{k+1,n}$ does admit $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits in the border setting; see [Kum20, Theorem 1.3] (Equation (1)) for the case of algebraically closed fields of characteristic zero, as well as [DS26].

The common zero set $\bigcap_{i=1}^{k'} \mathbf{V}(q_i)$ contains the origin and is therefore nonempty. Let V be an irreducible component of this set. By Lemma 2.7, the codimension of V in $\mathbb{A}^N$ is at most k' and hence at most k. Applying Bézout's Inequality (Proposition 2.9) with $X = \mathbb{A}^N$, which has degree 1, gives $\deg V \leq \delta^{k'} \leq \delta^k$.

By Lemma 2.19, there is some j such that $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_j] = \{0\}$. Fix such a j. Set $C := \operatorname{codim} V$ and $\Delta := \deg V$. We have $C\Delta \leq k\delta^k < n$. By Lemma 2.20, there exist indices $\ell_1, \ldots, \ell_k \in [N] \setminus B_j$ and an irreducible component V_0 of $V \cap \bigcap_{i=1}^{k} \mathbf{V}(x_{\ell_i})$ such that $\mathbf{I}(V_0) \cap \mathbb{F}[\mathbf{x}_j] = \{0\}$ and

$$P_{k+1,n} \equiv \mathbf{x}^{B_j} \pmod{\mathbf{I}(V_0)}. \tag{8}$$

The right side of (8) is nonzero modulo $\mathbf{I}(V_0)$ since $\mathbf{I}(V_0) \cap \mathbb{F}[\mathbf{x}_j] = \{0\}$. On the other hand, $P_{k+1,n}$ is zero on V_0 since each T_i is divisible by q_i and $q_i \in \mathbf{I}(V) \subseteq \mathbf{I}(V_0)$. Thus we obtain a contradiction. □

4 Impossibility Under a Constant-Freeness Assumption in the Border Setting

In this section, we extend Theorem 3.1 to the border setting. We show that $P_{k+1,n}$ cannot even be approximated by a $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit in which each top-level product gate computes either a constant or a polynomial divisible by a constant-free factor of degree at most δ. In particular, this result rules out border computation of $P_{k+1,n}$ by homogeneous $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits.

Throughout this section, let $\mathbb{F}$ be an algebraically closed field.

4.1 Additional Preliminaries

We give some additional preliminaries that will be used in this and later sections.

Limits and Normalization. A common operation will be to divide a nonzero Laurent series in ε by the lowest power of ε occurring in it. The resulting power series has a nonzero limit at $\varepsilon = 0$.

Definition 4.1: Let L be a field. For $g \in L[[\varepsilon]]$, write $\lim g := g|_{\varepsilon=0}$; the map $\lim\colon L[[\varepsilon]] \to L$ is a ring homomorphism. Every nonzero $g \in L((\varepsilon))$ has a unique expression

$$g = \varepsilon^a \widetilde{g}, \qquad \text{where} \quad a = \operatorname{val}_\varepsilon(g), \quad \widetilde{g} \in L[[\varepsilon]], \quad \text{and} \quad \lim \widetilde{g} \neq 0.$$

The *normalization* of g is $\widetilde{g}$.

Units up to Powers of ε. We now define the notion of ε-units. Let $\mathbf{u}$ be a tuple of variables.

Definition 4.2: A power series $g \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$ is an *ε-unit* if $g = \varepsilon^a \tilde{g}$ for some nonnegative integer a and some $\tilde{g} \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$ with $\tilde{g}(0, \mathbf{0}) \neq 0$.

For example, every nonzero element of $\mathbb{F}[[\varepsilon]]$ is an ε-unit, while $\varepsilon + u_1$ is not.

Lemma 4.3: Let $g, h \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$. If gh is an ε-unit, then so are g and h.

Proof: The ring $\mathbb{F}[[\varepsilon, \mathbf{u}]]$ is a UFD, and ε is prime, so every divisor of a power of ε times a unit is itself a power of ε times a unit. □

An element $g \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$ is an ε-unit if and only if it is invertible in $\mathbb{F}[[\varepsilon, \mathbf{u}]]_\varepsilon$. Indeed, the forward implication is immediate; conversely, an inverse gives $gh = \varepsilon^c$ for some $h \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$ and $c \geq 0$, so Lemma 4.3 applies.

Completions. A formal power series is determined by its truncations at each degree. Completion extends this idea to other rings, using reduction modulo powers of an ideal in place of truncation.

Definition 4.4 ([AM69, Chapter 10]): Let R be a ring, and let $\mathfrak{a} \subseteq R$ be an ideal. The $\mathfrak{a}$-*adic completion* of R, also called its *completion with respect to* $\mathfrak{a}$, is the ring $\widehat{R}$ defined by

$$\widehat{R} \coloneqq \varprojlim_{n \geq 1} R/\mathfrak{a}^n.$$

Concretely, an element of $\widehat{R}$ is a sequence $(r_1 + \mathfrak{a}, r_2 + \mathfrak{a}^2, r_3 + \mathfrak{a}^3, \ldots)$, where $r_n \in R$ and $r_{n+1} \equiv r_n \pmod{\mathfrak{a}^n}$ for $n \geq 1$. This compatibility condition means that reducing the $(n+1)$-st residue class modulo $\mathfrak{a}^n$ gives the n-th residue class. Addition and multiplication are performed coordinate-wise.

There is a natural ring homomorphism $R \to \widehat{R}$ sending r to $(r + \mathfrak{a}^n)_{n \geq 1}$. The ring R is *complete with respect to* $\mathfrak{a}$ if this map is an isomorphism. For a local ring, completion means completion with respect to its maximal ideal unless otherwise stated.

For example, the (t)-adic completion of $\mathbb{F}[t]$ is $\mathbb{F}[[t]]$: A formal power series determines a compatible sequence of truncations modulo t^n, and conversely, such a sequence uniquely determines every coefficient of the power series.

We record some useful facts about completions.

Proposition 4.5: Let O be a Noetherian local ring with maximal ideal $\mathfrak{n}$, and let $\widehat{O}$ be its $\mathfrak{n}$-adic completion.

1. The ring $\widehat{O}$ is a Noetherian local ring with maximal ideal $\mathfrak{n}\widehat{O}$ and dimension $\dim O$ [AM69, Proposition 10.16, Theorem 10.26, and Corollary 11.19], [Eis95, Corollary 10.12]. Moreover, $\widehat{O}$ is complete with respect to $\mathfrak{n}\widehat{O}$ [Eis95, Theorem 7.1(b)].
2. For every proper ideal $I \subseteq O$, the completion of the local ring O/I is $\widehat{O}/I\widehat{O}$. Consequently, $\dim \widehat{O}/I\widehat{O} = \dim O/I$ [AM69, Proposition 10.13 and Corollary 11.19].
3. Let $b \in \mathbb{F}^N$, let $\mathfrak{m}_b \coloneqq \langle x_1 - b_1, \ldots, x_N - b_N \rangle$, and define $\iota_b \colon \mathbb{F}[\mathbf{x}] \to \mathbb{F}[[\mathbf{y}]]$ by $\iota_b(f) = f(b + \mathbf{y})$. Translation induces an isomorphism

$$\widehat{\mathbb{F}[\mathbf{x}]_{\mathfrak{m}_b}} \cong \mathbb{F}[[\mathbf{y}]].$$

   For every ideal $J \subseteq \mathfrak{m}_b$, it also induces

$$\widehat{\mathbb{F}[\mathbf{x}]_{\mathfrak{m}_b}/J\mathbb{F}[\mathbf{x}]_{\mathfrak{m}_b}} \cong \mathbb{F}[[\mathbf{y}]] / \langle f(b + \mathbf{y}) \mid f \in J \rangle \mathbb{F}[[\mathbf{y}]].$$

The first isomorphism is the standard description of completion at a point [Eis95, §7.1, Example on p. 181], and the second follows from Item 2.

Lemma 4.6 ([Stacks, Tag 00OT]): Let $\mathbb{F}$ be algebraically closed, let $b \in \mathbb{F}^N$, and set $\mathfrak{m}_b = \langle x_i - b_i \mid i \in [N] \rangle$. Let $J \subseteq \mathfrak{m}_b$ be an ideal, and let $V_1, \ldots, V_t$ be the irreducible components of $\mathbf{V}(J)$ containing b. The following dimension formula holds:

$$\dim(\mathbb{F}[\mathbf{x}]_{\mathfrak{m}_b} / J\mathbb{F}[\mathbf{x}]_{\mathfrak{m}_b}) = \max_{i \in [t]} \dim V_i.$$

Local Rings. We will also use the following two facts about local rings.

Lemma 4.7:

1. If B is a Noetherian local domain and t is a nonzero element of its maximal ideal, then $\dim B = \dim(B/tB) + 1$.
2. For $\mathcal{O} = \mathbb{F}[[\varepsilon]][\mathbf{x}]_{\langle \varepsilon, \mathbf{x} \rangle}$, completion at the maximal ideal gives $\widehat{\mathcal{O}} \cong \mathbb{F}[[\varepsilon, \mathbf{x}]]$.

Proof: The first assertion follows from [Stacks, Tag 00KW]. For the second, the natural maps give $\mathcal{O} / \langle \varepsilon, \mathbf{x} \rangle^r \mathcal{O} \cong \mathbb{F}[\varepsilon, \mathbf{x}] / \langle \varepsilon, \mathbf{x} \rangle^r$ for $r \geq 1$. To conclude, take inverse limits. □

Modules. For an A-module M and an A-algebra B, we write $M \otimes_A B$ for the *extension of scalars* from A to B. We use the natural identifications

$$M \otimes_A (A/I) \cong M/IM \qquad \text{and} \qquad M \otimes_A S^{-1}A \cong S^{-1}M,$$

where I is an ideal and S is a multiplicatively closed subset of A.

Suppose that A is a domain with fraction field K. The *rank* of a finite A-module M is defined by

$$\operatorname{rank}_A M := \dim_K(M \otimes_A K).$$

The module M is *torsion-free* if, for every nonzero $r \in A$ and every $m \in M$, the equality $rm = 0$ implies $m = 0$ [Mat86, § 11, p. 84]. If A is a principal ideal domain (PID), then every finite torsion-free A-module is free and every submodule of A^s is free of rank at most s [Lan02, Chapter III, Theorems 7.1 and 7.3]. In particular, if A is a PID, then every finite torsion-free A-module M is isomorphic to $A^{\operatorname{rank}_A M}$.

Specialization at $\varepsilon = 0$. The next lemma controls the dimensions and degrees of the irreducible components obtained by setting $\varepsilon = 0$ in a prime ideal over $\mathbb{F}((\varepsilon))$, after contracting it to $\mathbb{F}[[\varepsilon]][\mathbf{x}]$.

Lemma 4.8: Let $R := \mathbb{F}[[\varepsilon]]$ and $K := \mathbb{F}((\varepsilon))$. Let $\pi \colon R[\mathbf{x}] \twoheadrightarrow \mathbb{F}[\mathbf{x}]$ be the $\mathbb{F}$-linear ring homomorphism sending ε to 0 and fixing $\mathbb{F}[\mathbf{x}]$. Let $\mathfrak{p} \subseteq K[\mathbf{x}]$ be a prime ideal. Let $I := \mathfrak{p} \cap R[\mathbf{x}]$, $J := \pi(I)$, and $X := \mathbf{V}(J)$. Suppose that $I \subseteq \langle \varepsilon, x_1, \ldots, x_N \rangle$, which holds in particular if $\mathfrak{p} \subseteq \langle x_1, \ldots, x_N \rangle K[\mathbf{x}]$. Under this assumption, $\mathbf{0} \in X$, and X is equidimensional of dimension m, where $m := \dim K[\mathbf{x}]/\mathfrak{p}$. Moreover, every irreducible component V of X satisfies $\deg V \leq \operatorname{DEG} \mathfrak{p}$.

Proof: The kernel of π is $\varepsilon R[\mathbf{x}]$. In fact, $\pi(R[\mathbf{x}]_{\leq s}) = \mathbb{F}[\mathbf{x}]_{\leq s}$ for every $s \geq 0$. Since π is surjective, J is an ideal.

Since $\mathfrak{p}$ is prime, its contraction I is a prime ideal of $R[\mathbf{x}]$, and $I \cap R$ is a prime ideal of R. The only prime ideals of R are $\{0\}$ and εR. Since $\mathfrak{p}$ is a proper ideal and ε is a unit of $K[\mathbf{x}]$, we have $\varepsilon \notin \mathfrak{p}$ and hence $\varepsilon \notin I$. Thus $I \cap R = \{0\}$.

Let $A := R[\mathbf{x}]/I$. Since I is prime, A is a domain. Moreover, since $I \cap R = \{0\}$, the natural map $R \to A$ is injective and embeds R as a subring of A. Clearing coefficient denominators gives $IK[\mathbf{x}] = \mathfrak{p}$. Consequently,

$$A_\varepsilon \cong K[\mathbf{x}]/\mathfrak{p} \qquad \text{and} \qquad A/\varepsilon A \cong \mathbb{F}[\mathbf{x}]/J.$$

The assumption $I \subseteq \langle \varepsilon, x_1, \ldots, x_N \rangle$ gives $J \subseteq \langle x_1, \ldots, x_N \rangle$. Thus J is proper, X contains the origin, and $A/\varepsilon A$ is nonzero.

We now apply [Stacks, Tag 00QK]. Here R is a valuation ring, A is a finitely generated R-algebra that is a domain containing R, and $A/\varepsilon A$ is nonzero. In our setting, the cited result states that every minimal prime $\mathfrak{q}$ of J satisfies

$$\dim \mathbb{F}[\mathbf{x}]/\mathfrak{q} = \dim A_\varepsilon = \dim K[\mathbf{x}]/\mathfrak{p} = m.$$

Since $\mathbf{I}(X) = \sqrt{J}$ and J and $\sqrt{J}$ have the same minimal primes, the irreducible components of X are precisely the sets $\mathbf{V}(\mathfrak{q})$ for these minimal primes $\mathfrak{q}$. Thus X is equidimensional of dimension m.

Fix an irreducible component V of X. For $s \geq 0$, let M_s be the image of $R[\mathbf{x}]_{\leq s}$ in A. The module M_s is a *finite* R-module generated over R by the images of monomials of degree at most s. Since $M_s \subseteq A$ and A is a domain containing R, the module M_s is torsion-free and hence free over the PID R.

After localizing $M_s \subseteq A$ and clearing coefficient denominators, we obtain

$$M_s \otimes_R K \cong K[\mathbf{x}]_{\leq s}/\mathfrak{p}_{\leq s}.$$

Consequently,

$$\operatorname{rank}_R M_s = \dim_K(M_s \otimes_R K) = \dim_K(K[\mathbf{x}]_{\leq s}/\mathfrak{p}_{\leq s}) = \operatorname{HilbF}_{\mathfrak{p}}(s). \tag{9}$$

Since $\pi(I) = J \subseteq \mathbf{I}(V)$, specialization induces a surjection

$$M_s/\varepsilon M_s \twoheadrightarrow \mathbb{F}[\mathbf{x}]_{\leq s}/\,\mathbf{I}(V)_{\leq s}.$$

It follows that

$$\operatorname{HilbF}_{\mathbf{I}(V)}(s) \leq \dim_{\mathbb{F}}(M_s/\varepsilon M_s) = \operatorname{rank}_R M_s = \operatorname{HilbF}_{\mathfrak{p}}(s), \tag{10}$$

where the inequality follows from the preceding surjection. The first equality follows from freeness of M_s, and the second follows from (9). Since $\dim V = \dim K[\mathbf{x}]/\mathfrak{p}$, Lemma 2.8 applied to (10) gives $\deg V \leq \operatorname{DEG} \mathfrak{p}$. □

Remark 4.9: In Section 4.2, we apply Lemma 4.8 with $\mathfrak{p} \subseteq \langle x_1, \ldots, x_N \rangle K[\mathbf{x}]$ to prove Theorem 4.16. The more relaxed condition $\mathfrak{p} \cap R[\mathbf{x}] = I \subseteq \langle \varepsilon, x_1, \ldots, x_N \rangle$ will be used in Section 6.2 to

prove Theorem 6.7. ◁

Remark 4.10: Let V be an irreducible component of the special fiber defined by specializing $\mathfrak{p} \cap R[\mathbf{x}]$ at $\varepsilon = 0$. In our applications, Lemma 2.10 bounds $\operatorname{DEG} \mathfrak{p}$ from generators of an ideal having $\mathfrak{p}$ as a minimal prime. It is tempting to reduce these generators modulo ε and apply Bézout's Inequality directly. However, powers of ε may appear in the denominators of their coefficients, so this reduction may not be defined. Even after denominators are cleared, reduction may lose equations, so V need not be an irreducible component of the resulting zero set. We instead transfer the degree bound using Lemma 4.8.

Geometrically, $\mathfrak{p}$ defines the generic fiber of a family over R, while V is an irreducible component of its special fiber at $\varepsilon = 0$. Our proof of Lemma 4.8 uses the freeness of certain R-modules to compare Hilbert functions and transfer the degree bound.[2] ◁

Resultants. We now define the resultant and record some properties that will be used later.

Definition 4.11 ([vzGG13, §6.3]): Let f and g be polynomials in $R[z]$ of degrees l and r, respectively, with $l \geq 1$ and $r \geq 0$. Write $f = a_0 z^l + a_1 z^{l-1} + \cdots + a_l$ and $g = b_0 z^r + b_1 z^{r-1} + \cdots + b_r$, where $a_0 \neq 0$ and $b_0 \neq 0$. The *Sylvester matrix* $\operatorname{Syl}_z(f, g)$ is the $(l + r) \times (l + r)$ matrix over R with columns specified as follows. For $1 \leq c \leq r$, its c-th column has the entries $a_0, \ldots, a_l$ in rows $c, \ldots, c + l$. For $1 \leq c \leq l$, its $(r + c)$-th column has the entries $b_0, \ldots, b_r$ in rows $c, \ldots, c + r$. The *resultant* of f and g with respect to z is defined by $\operatorname{Res}_z(f, g) := \det \operatorname{Syl}_z(f, g)$. For a nonzero constant $c \in R$ and a nonzero polynomial $g \in R[z]$, set $\operatorname{Res}_z(c, g) := c^{\deg_z g}$; in particular, $\operatorname{Res}_z(1, 1) = 1$. For $\deg_z f \geq 1$, set $\operatorname{Res}_z(f, 0) = 0$.

Lemma 4.12 ([Jan10, Theorems 1.10–1.11]): Let R be a domain. For nonzero $f, g, h \in R[z]$, we have

$$\operatorname{Res}_z(fg, h) = \operatorname{Res}_z(f, h) \operatorname{Res}_z(g, h) \qquad \text{and} \qquad \operatorname{Res}_z(f, gh) = \operatorname{Res}_z(f, g) \operatorname{Res}_z(f, h).$$

If f is monic of positive degree, then

$$\operatorname{Res}_z(f, g + hf) = \operatorname{Res}_z(f, g).$$

If $f(z) = \prod_{i=1}^{d} (z - r_i)$ over an extension of $\operatorname{Frac}(R)$, then

$$\operatorname{Res}_z(f, g) = \prod_{i=1}^{d} g(r_i).$$

Let R be a domain and let $f, g \in R[z]$ be as in Definition 4.11. There are $u, v \in R[z]$ with $uf + vg = \operatorname{Res}_z(f, g)$ [vzGG13, Corollary 6.21]. If R is a unique factorization domain, then $\operatorname{Res}_z(f, g) = 0$ if and only if $\gcd(f, g)$ is nonconstant in $R[z]$ [vzGG13, Corollary 6.20]. Suppose now that the characteristic of R is 0. The *discriminant* of f is defined by $\operatorname{Disc}_z(f) := \frac{(-1)^{d(d-1)/2}}{\operatorname{lc}_z(f)} \operatorname{Res}_z(f, \partial_z f)$,

[2] Since free modules are flat [Mat86, Chapter 3, §7], this argument can be viewed as an elementary manifestation of the more general role of flatness in controlling numerical invariants under specialization. A related general result is that Hilbert polynomials are locally constant in flat projective families. This result follows from local constancy of Euler characteristics [Stacks, Tag 0B9T].

where $d = \deg_z(f)$. If $f \in R[z]$, then $\mathrm{Disc}_z(f) \in R$ [Jan10, Theorem 3.5(i)].

Resultants commute with homomorphisms of coefficient rings that preserve the degrees of the input polynomials. More precisely, if $\varphi\colon R \to S$ is a homomorphism of domains with $\deg_z \varphi(f) = \deg_z f$ and $\deg_z \varphi(g) = \deg_z g$, then

$$\varphi(\mathrm{Res}_z(f,g)) = \mathrm{Res}_z(\varphi(f), \varphi(g)).$$

This identity follows by applying φ to the Sylvester determinant, which is an integer polynomial in the coefficients of f and g. In particular, resultants commute with the injective substitution $\varepsilon \mapsto \varepsilon^e$ on formal power-series coefficients for every integer $e \geq 1$.

Proposition 4.13: Let L be a field of characteristic 0, and let $R := L[\mathbf{x}] = L[x_1, \ldots, x_N]$. Let $f \in R[z]$ be irreducible, and set $d := \deg_z f$. Suppose $d \geq 1$ and $\deg_{\mathbf{x},z} f \leq D$.

1. $\mathrm{Res}_z(f, \partial_z f) \neq 0$ and $\deg \mathrm{Res}_z(f, \partial_z f) \leq D(D-1)$ (cf. [HP94, p. 155]).
2. If $(\mathbf{c}, \gamma) \in L^N \times L$ satisfies $f(\mathbf{c}, \gamma) = 0$ and $\mathrm{Res}_z(f, \partial_z f)(\mathbf{c}) \neq 0$, then $(\partial_z f)(\mathbf{c}, \gamma) \neq 0$.
3. Let $\ell \in R$ be the leading coefficient of f as a polynomial in z. If $\mathbf{c} \in L^N$ satisfies $\mathrm{Res}_z(f, \partial_z f)(\mathbf{c}) \neq 0$, then $\ell(\mathbf{c}) \neq 0$. In particular, $\deg_z f(\mathbf{c}, z) = \deg_z f$.

Proof: Since f is irreducible and $\partial_z f$ is a nonzero polynomial of smaller z-degree, the two polynomials are coprime, so $\mathrm{Res}_z(f, \partial_z f) \neq 0$. The coefficients of z^i in f and $\partial_z f$ have $\mathbf{x}$-degrees at most $D - i$ and $D - 1 - i$, respectively. The homogeneity properties of the resultant [Jan10, Theorem 1.4] therefore give

$$\deg \mathrm{Res}_z(f, \partial_z f) \leq (d-1)D + d(D-1) - d(d-1) = D(D-1) - (D-d)^2 \leq D(D-1).$$

Evaluating an identity $uf + v\partial_z f = \mathrm{Res}_z(f, \partial_z f)$ at $(\mathbf{c}, \gamma)$ proves Item 2.

Finally, Item 3 holds since $\mathrm{Res}_z(f, \partial_z f) = (-1)^{d(d-1)/2} \ell \, \mathrm{Disc}_z(f)$ and $\mathrm{Disc}_z(f) \in L[\mathbf{x}]$. □

Formal Implicit Function Theorem. We state a form of the formal implicit function theorem, which identifies a quotient of a power-series ring with a power-series ring in fewer variables.

Lemma 4.14: Let $\mathbb{F}$ be a field. Let m and N be integers with $0 \leq m \leq N$, and set $\mathbf{y} = (y_1, \ldots, y_N)$ and $\mathbf{y}' = (y_1, \ldots, y_m)$. Let $f_{m+1}, \ldots, f_N \in \langle\, y_1, \ldots, y_N \,\rangle\, \mathbb{F}[[\mathbf{y}]]$, and define their Jacobian matrix at the origin by

$$J := \big((\partial_{y_{a'}} f_a)(\mathbf{0})\big)_{a,a' \in [N] \setminus [m]}.$$

Suppose that J is invertible. In particular, J is invertible if each f_a lies in $\mathbb{F}[[\mathbf{y}', y_a]]$ and satisfies $(\partial_{y_a} f_a)(\mathbf{0}) \neq 0$. Under these hypotheses, the natural map

$$\mathbb{F}[[\mathbf{y}']] \to \mathbb{F}[[\mathbf{y}]] / \langle\, f_{m+1}, \ldots, f_N \,\rangle$$

is an isomorphism.

Proof: Let φ be the substitution endomorphism of $\mathbb{F}[[\mathbf{y}]]$ fixing $y_1, \ldots, y_m$ and sending y_a to f_a

for $a > m$. The Jacobian of this substitution has the following block matrix representation:

$$\begin{pmatrix} I_m & 0 \\ B & J \end{pmatrix}.$$

By the formal inverse function theorem [Eis95, Exercise 7.25], φ is an automorphism. It fixes $\mathbb{F}[[\mathbf{y}']]$ pointwise and maps $\langle\, y_{m+1}, \ldots, y_N \,\rangle$ onto $\langle\, f_{m+1}, \ldots, f_N \,\rangle$. Passing to quotients gives the asserted natural isomorphism. □

The next lemma identifies the ideal of a variety in the power-series ring around a point and shows that the quotient is itself a power-series ring.

Lemma 4.15: Let $\mathbb{F}$ be an algebraically closed field. Let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety of dimension m, and let $b \in V$. For each $a \in [N] \setminus [m]$, let $Q_a \in \mathbf{I}(V) \cap \mathbb{F}[x_1, \ldots, x_m, x_a]$ satisfy $(\partial_{x_a} Q_a)(b) \neq 0$. Let $\iota_b \colon \mathbb{F}[\mathbf{x}] \to \mathbb{F}[[\mathbf{y}]]$ be as in Proposition 4.5. Under these hypotheses,

$$\iota_b(\mathbf{I}(V))\mathbb{F}[[\mathbf{y}]] = \langle\, \iota_b(Q_{m+1}), \ldots, \iota_b(Q_N) \,\rangle\, \mathbb{F}[[\mathbf{y}]],$$

and the natural map

$$\mathbb{F}[[y_1, \ldots, y_m]] \to \mathbb{F}[[\mathbf{y}]] / \iota_b(\mathbf{I}(V))\mathbb{F}[[\mathbf{y}]]$$

is an isomorphism.

Proof: Set $S := \mathbb{F}[[\mathbf{y}]]$, $J := \langle\, \iota_b(Q_{m+1}), \ldots, \iota_b(Q_N) \,\rangle\, S$, and $\widehat{I} := \iota_b(\mathbf{I}(V))S$. Thus $J \subseteq \widehat{I}$, and Lemma 4.14 identifies S/J with $\mathbb{F}[[y_1, \ldots, y_m]]$. In particular, S/J is a domain of dimension m. On the other hand, Proposition 4.5 and Lemma 4.6 give

$$\dim(S/\widehat{I}) = \dim\left(\mathbb{F}[\mathbf{x}]_{\mathfrak{m}_b} / \mathbf{I}(V)\mathbb{F}[\mathbf{x}]_{\mathfrak{m}_b}\right) = m.$$

If $J \neq \widehat{I}$, then $\widehat{I}/J$ is a nonzero proper ideal of the m-dimensional domain S/J. In that case, Lemma 2.1 gives $\dim(S/\widehat{I}) \leq m - 1$, contradicting $\dim(S/\widehat{I}) = m$. Thus $J = \widehat{I}$, and the asserted natural isomorphism follows. □

4.2 Proof of the Impossibility Result

We now prove the main result of this section.

Theorem 4.16: Suppose $\delta \geq 1$ and $k \geq 1$. If $n > k\delta^k$, then $P_{k+1,n}$ cannot be approximately computed by circuits over $\mathbb{F}((\varepsilon))$ of the form $\sum_{i=1}^{k} T_i$, where each T_i is either constant or divisible by a constant-free polynomial $q_i \in \mathbb{F}((\varepsilon))[\mathbf{x}]$ of degree at most δ. In particular, $P_{k+1,n}$ cannot be approximately computed by homogeneous $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits over $\mathbb{F}((\varepsilon))$.

Proof: Let $R := \mathbb{F}[[\varepsilon]]$, $K := \mathbb{F}((\varepsilon))$, $N := (k+1)n$, and $\mathbf{x} := (x_1, \ldots, x_N)$. Suppose towards a contradiction that a circuit $\sum_{i=1}^{k} T_i$ of the stated form approximately computes $P_{k+1,n}$. By Definition 2.12, we have

$$\sum_{i=1}^{k} T_i = P_{k+1,n} + \varepsilon S \tag{11}$$

for some $S \in R[\mathbf{x}]$.

We first remove the constant summands. After reindexing, let $T_1, \dots, T_r$ be the terms that do not lie in K, so $T_{r+1}, \dots, T_k \in K$. For each $i \in [r]$, choose a constant-free polynomial $q_i \in K[\mathbf{x}]$ of degree at most δ dividing T_i, and write $T_i = q_i g_i$ with $g_i \in K[\mathbf{x}]$. Since $T_i \neq 0$, the polynomial q_i is nonzero and hence nonconstant. Moreover, $T_i(\mathbf{0}) = 0$ because $q_i(\mathbf{0}) = 0$. Evaluating both sides of (11) at $\mathbf{x} = \mathbf{0}$ and using $P_{k+1,n}(\mathbf{0}) = 0$ therefore gives

$$\sum_{i=r+1}^{k} T_i = \varepsilon S(\mathbf{0}). \tag{12}$$

Subtracting (12) from (11) yields

$$P_{k+1,n} + \varepsilon S' = \sum_{i=1}^{r} q_i g_i, \tag{13}$$

where $S' \coloneqq S - S(\mathbf{0}) \in R[\mathbf{x}]$ and $r \leq k$.

We next specialize a suitable minimal prime of $\langle\, q_1, \dots, q_r \,\rangle\, K[\mathbf{x}]$ at $\varepsilon = 0$. Since the q_i are constant-free, we have

$$\langle\, q_1, \dots, q_r \,\rangle\, K[\mathbf{x}] \subseteq \langle\, x_1, \dots, x_N \,\rangle\, K[\mathbf{x}].$$

The latter ideal is prime, since $K[\mathbf{x}]/\langle\, x_1, \dots, x_N \,\rangle\, K[\mathbf{x}] \cong K$. We may therefore choose a minimal prime $\mathfrak{p}$ of $\langle\, q_1, \dots, q_r \,\rangle\, K[\mathbf{x}]$ such that $\mathfrak{p} \subseteq \langle\, x_1, \dots, x_N \,\rangle\, K[\mathbf{x}]$. By Lemma 2.1, $\operatorname{ht}\mathfrak{p} \leq r$. By the same lemma, $\dim K[\mathbf{x}]/\mathfrak{p} = N - \operatorname{ht}\mathfrak{p}$. Consequently,

$$m \coloneqq \dim K[\mathbf{x}]/\mathfrak{p} = N - \operatorname{ht}\mathfrak{p} \geq N - r \geq N - k.$$

Moreover, Lemma 2.10 gives

$$\operatorname{DEG}\mathfrak{p} \leq \delta^{\operatorname{ht}\mathfrak{p}} \leq \delta^r \leq \delta^k.$$

Let $\pi\colon R[\mathbf{x}] \twoheadrightarrow \mathbb{F}[\mathbf{x}]$ be the $\mathbb{F}$-linear ring homomorphism sending ε to 0 and fixing $\mathbb{F}[\mathbf{x}]$. Let $I \coloneqq \mathfrak{p} \cap R[\mathbf{x}]$, $J \coloneqq \pi(I) \subseteq \mathbb{F}[\mathbf{x}]$, and $X \coloneqq \mathbf{V}(J) \subseteq \mathbb{A}^N$. By Lemma 4.8, X is nonempty and equidimensional of dimension m, and every irreducible component of X has degree at most $\operatorname{DEG}\mathfrak{p}$. Fix an irreducible component V of X. Thus $\operatorname{codim}_{\mathbb{A}^N} V = N - m \leq k$ and $\deg V \leq \operatorname{DEG}\mathfrak{p} \leq \delta^k$.

Moreover, $P_{k+1,n}$ vanishes on X and hence on V. Indeed, the left side of (13) lies in $R[\mathbf{x}]$, while its right side lies in the ideal $\langle\, q_1, \dots, q_r \,\rangle\, K[\mathbf{x}]$, which is contained in $\mathfrak{p}$. Thus $P_{k+1,n} + \varepsilon S' \in \mathfrak{p} \cap R[\mathbf{x}] = I$. Applying π gives

$$P_{k+1,n} = \pi(P_{k+1,n} + \varepsilon S') \in \pi(I) = J,$$

proving the claim that $P_{k+1,n}$ vanishes on X and hence on V.

By Lemma 2.19, we may choose j such that $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_j] = \{\, 0 \,\}$. Let $C \coloneqq \operatorname{codim}_{\mathbb{A}^N} V$ and $\Delta \coloneqq \deg V$. The bounds above imply $C\Delta \leq k\delta^k < n$. By Lemma 2.20, there exist indices $\ell_1, \dots, \ell_k \in [N] \setminus B_j$ and an irreducible component V_0 of $V \cap \bigcap_{i=1}^{k} \mathbf{V}(x_{\ell_i})$ such that $\mathbf{I}(V_0) \cap \mathbb{F}[\mathbf{x}_j] = \{\, 0 \,\}$. These choices give

$$P_{k+1,n} \equiv \mathbf{x}^{B_j} \pmod{\mathbf{I}(V_0)}. \tag{14}$$

The right side is nonzero modulo $\mathbf{I}(V_0)$, since $\mathbf{x}^{B_j}$ is a nonzero polynomial in $\mathbb{F}[\mathbf{x}_j]$ and $\mathbf{I}(V_0) \cap$

$\mathbb{F}[\mathbf{x}_j] = \{0\}$. On the other hand, $P_{k+1,n}$ vanishes on V and hence on its subvariety V_0, contradicting (14). □

5 Low-Weight Translation

Let k and δ be at least 1, and let n satisfy $n > k\delta^k$. Suppose that $P_{k+1,n} = \sum_{i=1}^{k} T_i$, where each T_i is a product of nonzero polynomials in $\mathbb{F}[\mathbf{x}]$ of degree at most δ. In this section, we show that, after a low-weight translation (i.e., shifting only a small number of coordinates) and passage to a suitable power-series ring, the translated polynomial can be computed by a sum of at most k product gates, each nonzero at the origin.

Throughout Section 5, let $\mathbb{F}$ be an algebraically closed field of characteristic 0. To apply Lemma 4.15, we need a point where all the chosen derivatives are nonzero. The next lemma finds such a point with few nonzero coordinates, with a bound depending only on the codimension and degree of the variety.

Lemma 5.1: Let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety of degree Δ and codimension C. Let $T \subseteq [N]$ satisfy $|T| = \dim V$ and $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_T] = \{0\}$. Under these hypotheses, there exist

- a set $E \subseteq [N]$ with $[N] \setminus T \subseteq E$ and $|E| \leq C\Delta^2$,
- a point $b \in V$ with $\operatorname{supp}(b) \subseteq E$, and
- for each $a \in [N] \setminus T$, an irreducible polynomial $Q_a \in \mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_T, x_a]$ with $\deg Q_a \leq \Delta$ and $(\partial_{x_a} Q_a)(b) \neq 0$.

Proof: By Lemma 2.17, for each $a \in [N] \setminus T$, there is an irreducible polynomial $Q_a \in \mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_{T \cup \{a\}}]$ with $\deg Q_a \leq \Delta$ and $\deg_{x_a} Q_a \geq 1$. When Q_a is viewed as a polynomial in x_a, its leading coefficient $c_a \in \mathbb{F}[\mathbf{x}_T]$ is nonzero and has degree at most $\Delta - 1$. Moreover, for $Q := \prod_a c_a$ as in Lemma 2.17 and $h := Q|_V$, the ring $\mathbb{F}[V]_h$ is integral over $\mathbb{F}[\mathbf{x}_T|_V]_h$.

For each $a \in [N] \setminus T$, let $R_a := \operatorname{Res}_{x_a}(Q_a, \partial_{x_a} Q_a) \in \mathbb{F}[\mathbf{x}_T]$. By Proposition 4.13, each R_a is nonzero and has degree at most $\Delta(\Delta - 1)$. Let

$$R := \prod_{a \in [N] \setminus T} R_a \in \mathbb{F}[\mathbf{x}_T],$$

with $R = 1$ if $T = [N]$. Thus $R \neq 0$ and $\deg R \leq |[N] \setminus T|\Delta(\Delta - 1) = C\Delta(\Delta - 1)$.

Fix a monomial M occurring in R. Let $S := \operatorname{supp}(M) \subseteq T$, and let $E := S \cup ([N] \setminus T)$. Since $|S| \leq \deg M \leq \deg R$, we have $|E| \leq C\Delta(\Delta - 1) + C \leq C\Delta^2$.

Setting all variables outside S to zero deletes some terms of R and leaves the others unchanged. The term with monomial M survives, so the resulting polynomial in $\mathbb{F}[\mathbf{x}_S]$ is nonzero. Since $\mathbb{F}$ is infinite, there is a point $c \in \mathbb{A}^T$ such that $c_t = 0$ for $t \in T \setminus S$ and $R(c) \neq 0$.

For every $a \in [N] \setminus T$, we have $R_a(c) \neq 0$ since R_a divides R. By Item 3 of Proposition 4.13, applied to Q_a as a polynomial in x_a, we have $c_a(c) \neq 0$. Hence $Q(c) \neq 0$, so Corollary 2.15 gives $b \in V$ with $\pi_T(b) = c$. Since $b_t = c_t = 0$ for $t \in T \setminus S$, we have $\operatorname{supp}(b) \subseteq S \cup ([N] \setminus T) = E$.

Finally, fix $a \in [N] \setminus T$. Since $Q_a(b) = 0$ and $R_a(c) \neq 0$, Item 2 of Proposition 4.13, applied at $(c, b_a) \in \mathbb{A}^T \times \mathbb{A}^1$, gives $(\partial_{x_a} Q_a)(b) \neq 0$. □

Remark 5.2: Characteristic 0 is used in Lemma 5.1 to ensure that the resultants R_a are nonzero. The conclusion can fail in positive characteristic. Indeed, when $\mathrm{char}(\mathbb{F}) = p > 0$, take $V = \mathbf{V}(x_2^p - x_1) \subseteq \mathbb{A}^2$ and $T = \{1\}$. For this choice, we have $\mathbf{I}(V) \cap \mathbb{F}[x_1] = \{0\}$, but $(\partial_{x_2} Q)|_V = 0$ for every $Q \in \mathbf{I}(V)$, since $\mathbf{I}(V) = \langle x_2^p - x_1 \rangle$. ◁

We now state and prove the main result of Section 5.

Theorem 5.3: Let k and δ be at least 1, let n satisfy $n > k\delta^k$, and let $N := (k+1)n$. Suppose $P_{k+1,n} = \sum_{i=1}^{k} T_i$, where each T_i is a product of nonzero polynomials in $\mathbb{F}[\mathbf{x}]$ of degree at most δ. Suppose also that $\sum_{i=1}^{k} \deg T_i \leq D$. Under these hypotheses, there exist integers m, t_0, and k', elements $b_1, \ldots, b_{t_0} \in \mathbb{F}$, and polynomials $Q_a \in \mathbb{F}[x_1, \ldots, x_m, x_a]$ for $m < a \leq N$. The integers satisfy $N - 2k \leq m \leq N$, $0 \leq t_0 \leq \min\{2k\delta^{2k}, n\}$, and $1 \leq k' \leq k$. The following properties also hold.

1. Each Q_a is irreducible of degree at most δ^k, with $Q_a(\mathbf{0}) = 0$ and $(\partial_{x_a} Q_a)(\mathbf{0}) \neq 0$. Consequently, by Lemma 4.14, the natural map $\mathbb{F}[[x_1, \ldots, x_m]] \to \mathbb{F}[[\mathbf{x}]] / \langle Q_{m+1}, \ldots, Q_N \rangle$ is an isomorphism.
2. In $\mathbb{F}[[\mathbf{x}]] / \langle Q_{m+1}, \ldots, Q_N \rangle$, we have

$$\prod_{t=1}^{t_0} (x_t + b_t) \cdot \prod_{t=t_0+1}^{n} x_t = \sum_{i=1}^{k'} T_i',$$

   where each T_i' is a product of polynomials in $\mathbb{F}[\mathbf{x}]$ of degree at most δ, none of which vanishes at $\mathbf{0}$, and $\sum_{i=1}^{k'} \deg T_i' \leq D$.

Proof: We will first find an irreducible variety V such that, modulo $\mathbf{I}(V)$, every product gate either vanishes or is a product of units. To do this, apply Lemma 2.21 to $T_1, \ldots, T_k$ to obtain V and K, let $k' := |K|$, and relabel the T_i so that $K = [k']$. Thus $\operatorname{codim} V \leq k - k' \leq k$ and $\deg V \leq \delta^k$. Every factor of T_i with $i \leq k'$ restricts to a unit on V, and

$$P_{k+1,n} \equiv \sum_{i=1}^{k'} T_i \pmod{\mathbf{I}(V)}. \tag{15}$$

Next, we restrict to a subvariety on which only one block of $P_{k+1,n}$ remains. Since $\operatorname{codim} V \leq k$, Lemma 2.19 gives a block whose coordinates are algebraically independent on V. Exchanging this block with the first one leaves $P_{k+1,n}$ unchanged, so after renaming coordinates we may assume that $\mathbf{I}(V) \cap \mathbb{F}[x_1, \ldots, x_n] = \{0\}$. Since $\operatorname{codim} V \cdot \deg V \leq k\delta^k < n$, Lemma 2.20 gives indices $\ell_1, \ldots, \ell_k \in [N] \setminus [n]$ and an irreducible component V_0 of $V \cap \mathbf{V}(x_{\ell_1}, \ldots, x_{\ell_k})$. This component satisfies $\mathbf{I}(V_0) \cap \mathbb{F}[x_1, \ldots, x_n] = \{0\}$ and $P_{k+1,n} \equiv \prod_{t=1}^{n} x_t \pmod{\mathbf{I}(V_0)}$. Since $V_0 \subseteq V$, we have $\mathbf{I}(V) \subseteq \mathbf{I}(V_0)$, so (15) gives

$$\prod_{t=1}^{n} x_t \equiv \sum_{i=1}^{k'} T_i \pmod{\mathbf{I}(V_0)}. \tag{16}$$

Since the left side is a nonzero element of $\mathbb{F}[x_1, \dots, x_n]$, it does not lie in $\mathbf{I}(V_0)$. Thus $k' \geq 1$. Let $m := \dim V_0$ and $\Delta_0 := \deg V_0$. By Lemma 2.7, applied to $X = V$ and the coordinates $x_{\ell_1}, \dots, x_{\ell_k}$, we have $m \geq \dim V - k \geq N - 2k$. Moreover, Proposition 2.9 gives $\Delta_0 \leq \deg V \leq \delta^k$.

We now choose coordinates on V_0 and a point of V_0 with few nonzero coordinates. By Lemma 2.13, applied to V_0 and $S = [n]$, there is a set U with $[n] \subseteq U \subseteq [N]$, $|U| = m$, and $\mathbf{I}(V_0) \cap \mathbb{F}[\mathbf{x}_U] = \{0\}$; in particular, $m \geq n$. Rename coordinates by a permutation that maps $[n]$ onto itself and U onto $[m]$. This permutation fixes $\prod_{t=1}^{n} x_t$, so (16) still holds and $\mathbf{I}(V_0) \cap \mathbb{F}[x_1, \dots, x_m] = \{0\}$.

Since $\operatorname{codim}_{\mathbb{A}^N} V_0 = N - m \leq 2k$, applying Lemma 5.1 to V_0 gives a point $b \in V_0$ with $|\operatorname{supp}(b)| \leq 2k\delta^{2k}$. For each $a \in [N] \setminus [m]$, the lemma also gives an irreducible polynomial $\widetilde{Q}_a \in \mathbf{I}(V_0) \cap \mathbb{F}[x_1, \dots, x_m, x_a]$ with $\deg \widetilde{Q}_a \leq \Delta_0 \leq \delta^k$ and $\partial_{x_a} \widetilde{Q}_a(b) \neq 0$. Let $t_0 := |\operatorname{supp}(b) \cap [n]|$. Thus $t_0 \leq \min\{2k\delta^{2k}, n\}$. Rename coordinates by a permutation of $[n]$ that maps $\operatorname{supp}(b) \cap [n]$ onto $[t_0]$. This permutation preserves (16) and the form of each $\widetilde{Q}_a$, so we may assume that $b_t = 0$ for $t_0 < t \leq n$.

It remains to expand (16) around b. Let $\iota_b \colon \mathbb{F}[\mathbf{x}] \to \mathbb{F}[[\mathbf{y}]]$ be the substitution $x_i \mapsto b_i + y_i$ of Proposition 4.5. By Lemma 4.15, applied to V_0, b, and the $\widetilde{Q}_a$, the map ι_b sends $\mathbf{I}(V_0)$ into the ideal of $\mathbb{F}[[\mathbf{y}]]$ generated by $\iota_b(\widetilde{Q}_{m+1}), \dots, \iota_b(\widetilde{Q}_N)$. Applying ι_b to (16) and using that $b_t = 0$ for $t_0 < t \leq n$, we obtain

$$\prod_{t=1}^{t_0} (y_t + b_t) \cdot \prod_{t=t_0+1}^{n} y_t \equiv \sum_{i=1}^{k'} T_i(b + \mathbf{y}) \left(\operatorname{mod} \left\langle \widetilde{Q}_{m+1}(b + \mathbf{y}), \dots, \widetilde{Q}_N(b + \mathbf{y}) \right\rangle \right) \tag{17}$$

in $\mathbb{F}[[\mathbf{y}]]$. Writing $\mathbf{x}$ in place of $\mathbf{y}$, let $Q_a(\mathbf{x}) := \widetilde{Q}_a(b + \mathbf{x})$ for $a > m$ and $T'_i(\mathbf{x}) := T_i(b + \mathbf{x})$ for $i \leq k'$. With this notation, (17) is the identity in the second item, and it remains to check the stated properties of the Q_a and the T'_i.

The translation $\mathbf{x} \mapsto b + \mathbf{x}$ is a degree-preserving $\mathbb{F}$-algebra automorphism of $\mathbb{F}[\mathbf{x}]$ that maps $\mathbb{F}[x_1, \dots, x_m, x_a]$ onto itself. Hence each Q_a is an irreducible element of $\mathbb{F}[x_1, \dots, x_m, x_a]$ of degree at most δ^k. Moreover, $Q_a(\mathbf{0}) = \widetilde{Q}_a(b) = 0$ because $b \in V_0$ and $\widetilde{Q}_a \in \mathbf{I}(V_0)$. Also, $\partial_{x_a} Q_a(\mathbf{0}) = \partial_{x_a} \widetilde{Q}_a(b) \neq 0$. The isomorphism in the first item then follows from Lemma 4.14, applied to $f_a = Q_a \in \mathbb{F}[x_1, \dots, x_m, x_a]$ since $Q_a(\mathbf{0}) = 0$.

Finally, each factor of T'_i has the form $q_{i,j}(b + \mathbf{x})$, where $q_{i,j}$ is a factor of T_i. Translation preserves degrees, so these factors have degree at most δ, and $\sum_{i=1}^{k'} \deg T'_i = \sum_{i=1}^{k'} \deg T_i \leq D$. Since $q_{i,j}$ restricts to a unit on V and $b \in V_0 \subseteq V$, we have $q_{i,j}(b) \neq 0$ by Lemma 2.6. Thus every factor $q_{i,j}(b + \mathbf{x})$ of T'_i is nonzero at $\mathbf{0}$. □

6 A Border Lower Bound Without the Constant-Freeness Assumption

We showed that if $n > k\delta^k$, then $P_{k+1,n}$ cannot be computed by $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ circuits in which *every* product gate is a constant or has a constant-free factor (see Theorem 4.16). Moreover, we show later that, for sufficiently large n, it cannot be computed by any $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit; this is the exact setting and not a border computation (see Theorem 7.5). However, border circuits

without the constant-freeness assumption are much more powerful: For $k \geq 2$, they can compute every polynomial over algebraically closed fields, by the completeness result of [Kum20] (see the footnote at the beginning of Section 3). In this section, we show that such circuits computing $P_{k+1,n}$ must be exponentially large when k and δ are constants.

Theorem 6.1: Let k, δ, n, and s be at least 1. If $P_{k+1,n}$ is computed by a circuit in $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ of size s, then

$$n \leq O\left(k(4\delta)^k \log s\right) + (2\delta)^{O(k^2)}.$$

In particular, for all fixed k and δ, every such circuit has size $2^{\Omega(n)}$.

Since $P_{k+1,n}$ has a $\Sigma^{[k+1]}\Pi\Sigma$ circuit of size $O(kn)$, Theorem 6.1 gives an exponential gap between top fan-in k and $k+1$ in the border setting. For $\delta = 1$, Theorem 6.1 recovers the exponential top fan-in hierarchy for border depth-3 circuits of Dutta and Saxena [DS26].

Throughout this section, $\mathbb{F}$ is algebraically closed of characteristic zero. We fix parameters k, δ, and n that are at least 1, let $N := (k+1)n$, and write $\Delta := \delta^k$. We do not restrict the number of factors of a product gate, the multiplicities of repeated factors, or the orders of the poles in ε.

6.1 Auxiliary Statements for the Border Setting

We collect a few statements used in Section 6.2, in addition to those of Section 2 and Section 2.4.

Weierstrass Preparation and Abhyankar–Jung. For the following two classical theorems, assume that $\mathbb{F}$ is algebraically closed of characteristic zero.

Theorem 6.2 (Weierstrass Preparation [Lan02, Chapter IV, §9]): Let $\mathbf{y}$ be a tuple of variables and let $g \in \mathbb{F}[[\mathbf{y}]][[Y]]$ be such that $g(\mathbf{0}, Y)$ is nonzero with lowest term of degree d. The following statements hold.

1. Every $h \in \mathbb{F}[[\mathbf{y}]][[Y]]$ has a unique expression $h = qg + r$, where $q \in \mathbb{F}[[\mathbf{y}]][[Y]]$ and $r \in \mathbb{F}[[\mathbf{y}]][Y]$ has degree less than d.
2. We have $g = UW$, where U is a unit of $\mathbb{F}[[\mathbf{y}]][[Y]]$ and $W \in \mathbb{F}[[\mathbf{y}]][Y]$ is monic of degree d with $W(\mathbf{0}, Y) = Y^d$.

In the exact setting, the implicit function theorem (Lemma 4.15) expresses the dependent coordinates as power series in the free coordinates. In the border setting, a dependent coordinate is a root of a monic polynomial whose discriminant may vanish at $\varepsilon = 0$, and then the implicit function theorem does not apply. The next theorem replaces it: if the discriminant is an ε-unit, then the roots become power series in ε and $\mathbf{u}$ once ε is replaced by a suitable power ε^e. We use it in Lemma 6.11.

Theorem 6.3 (Abhyankar–Jung [PR12, Theorem 1.3 and Remark 1.4]): Let $\mathbf{u}$ be a tuple of variables, and let $W \in \mathbb{F}[[\varepsilon, \mathbf{u}]][z]$ be a monic polynomial of positive degree d. If $\mathrm{Disc}_z(W)$ is an ε-unit, then there are a positive integer e and power series $r_1, \ldots, r_d \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$ such that

$$W(\varepsilon^e, \mathbf{u}, z) = \prod_{i=1}^{d} (z - r_i).$$

After the substitution $\varepsilon \mapsto \varepsilon^e$, the polynomial W splits into linear factors over $\mathbb{F}[[\varepsilon, \mathbf{u}]]$; equivalently, the roots of the original polynomial W are power series in $\varepsilon^{1/e}$ and $\mathbf{u}$. For example, $W = z^2 - \varepsilon(1 + u_1)$ has discriminant $4\varepsilon(1 + u_1)$, which is an ε-unit. Its roots $\pm\varepsilon^{1/2}\sqrt{1 + u_1}$ are not power series in ε and u_1, but with $e = 2$ we get

$$W(\varepsilon^2, \mathbf{u}, z) = \left(z - \varepsilon\sqrt{1 + u_1}\right)\left(z + \varepsilon\sqrt{1 + u_1}\right),$$

whose roots $\pm\varepsilon\sqrt{1 + u_1}$ lie in $\mathbb{F}[[\varepsilon, \mathbf{u}]]$.

We restate Theorem 6.3 as Theorem A.2 in the appendix, where we give a detailed algebraic proof using standard results from commutative algebra.

Polynomial Compositions. Let $\mathbf{u} = (u_1, \ldots, u_M)$. The main theorem of this subsection states that, for any fixed monomial order, every nonzero polynomial in the class $\overline{\mathcal{C}}_d(m, r)$, to be defined below, has a trailing monomial with small support.

Definition 6.4: Let L be a field, and let d, m, and r be positive integers. Let $\mathcal{C}_d(m, r; L)$ be the set of polynomials of the form

$$G = \sum_{i=1}^{m} f_i(g_{i,1}, \ldots, g_{i,r}),$$

where the *outer polynomials* $f_i \in L[y_1, \ldots, y_r]$ are arbitrary and the *inner polynomials* $g_{i,j} \in L[\mathbf{u}]$ have degree at most d. We call G a sum of m *compositions of arity* r and write $\mathcal{C}_d(m, r) := \mathcal{C}_d(m, r; \mathbb{F})$ for the class of these sums over $\mathbb{F}$. The *border class* $\overline{\mathcal{C}}_d(m, r)$ is the set of polynomials $G \in \mathbb{F}[\mathbf{u}]$ for which some $G_\varepsilon \in \mathcal{C}_d(m, r; \mathbb{F}((\varepsilon)))$ satisfies $G_\varepsilon \in \mathbb{F}[[\varepsilon]][\mathbf{u}]$ and $G_\varepsilon(0, \mathbf{u}) = G$.

Thus $G \in \overline{\mathcal{C}}_d(m, r)$ means that G is the limit, as $\varepsilon \to 0$, of a sum G_ε of m compositions of arity r over $\mathbb{F}((\varepsilon))$. It is immediate that $\mathcal{C}_d(m, r) \subseteq \overline{\mathcal{C}}_d(m, r)$.

Lemma 6.5: The classes $\mathcal{C}_d(m, r; L)$ and $\overline{\mathcal{C}}_d(m, r)$ are closed under multiplication by scalars and are nondecreasing in m and r with respect to set inclusion. Moreover, over every field L,

$$\begin{aligned} \mathcal{C}_d(m_1, r_1; L) + \mathcal{C}_d(m_2, r_2; L) &\subseteq \mathcal{C}_d(m_1 + m_2, \max\{r_1, r_2\}; L), \\ \mathcal{C}_d(m_1, r_1; L) \cdot \mathcal{C}_d(m_2, r_2; L) &\subseteq \mathcal{C}_d(m_1 m_2, r_1 + r_2; L), \end{aligned}$$

and the same inclusions hold for the border classes $\overline{\mathcal{C}}_d$.

Proof: For the product, $(\sum_i f_i(\mathbf{g}_i))(\sum_j f'_j(\mathbf{g}'_j)) = \sum_{i,j}(f_i f'_j)(\mathbf{g}_i, \mathbf{g}'_j)$, where $f_i f'_j$ is a polynomial in $r_1 + r_2$ variables. The other inclusions are clear. For the border classes, note that setting $\varepsilon = 0$ defines a ring homomorphism $\mathbb{F}[[\varepsilon]][\mathbf{u}] \to \mathbb{F}[\mathbf{u}]$ and therefore preserves addition and multiplication. Thus the same argument applies. □

Now, we state the main theorem of this subsection. We use the trailing-monomial notation of Monomial Orders.

Theorem 6.6: Suppose that $\mathbb{F}$ is a field of characteristic zero, not necessarily algebraically closed. Every nonzero $G \in \overline{\mathcal{C}}_d(m, r)$ satisfies $|\mathrm{supp}(\mathrm{TM}(G))| = O(d(\log m + r))$.

The proof of Theorem 6.6 uses shifted partial Hasse derivatives, as well as spaces of their coefficients [For15, Definitions 3.2 and 4.1]. Let L be a field. For a polynomial $h \in L[\mathbf{y}] = L[y_1, \ldots, y_n]$ and $\mathbf{a} \in \mathbb{N}^n$, the $\mathbf{a}$-th *Hasse derivative* of h is $[\mathbf{t}^{\mathbf{a}}]h(\mathbf{y}+\mathbf{t})$, where $\mathbf{t}$ is a tuple of new variables. For nonnegative integers k and ℓ, define

$$\mathbf{y}^{\leq \ell}\partial_{\mathbf{y}}^{\leq k}(h) := \{\, \mathbf{y}^{\mathbf{b}}[\mathbf{t}^{\mathbf{a}}]h(\mathbf{y}+\mathbf{t}) \mid \mathbf{a}, \mathbf{b} \in \mathbb{N}^n, |\mathbf{a}| \leq k, |\mathbf{b}| \leq \ell \,\}.$$

For $G \in \mathcal{C}_d(m,r)$, Forbes [For15, Proposition 4.18] proved the bound $O\left(d(\log m + r \log 2r)\right)$, which would also suffice for us. His proof compares an upper and a lower bound on the dimension of a space of shifted partial derivatives. The upper bound is a rank condition and therefore also holds in the border setting. This extension to the border setting was already observed by Dutta, Dwivedi, and Saxena [DDS22, Claim C.9]. We include the short proof for completeness. Our proof closely follows the proof of Forbes [For15, Proposition 4.18].

Proof of Theorem 6.6: Suppose $G \in \overline{\mathcal{C}}_d(m,r)$. Let $R = \mathbb{F}[[\varepsilon]]$ and $K = \mathbb{F}((\varepsilon))$. By the definition of $\overline{\mathcal{C}}_d(m,r)$, there exists an expression

$$G_\varepsilon = \sum_{i=1}^{m} f_i(g_{i,1}, \ldots, g_{i,r}) \in R[\mathbf{x}], \qquad f_i \in K[y_1, \ldots, y_r], \quad \deg g_{i,j} \leq d, \quad G_\varepsilon|_{\varepsilon=0} = G. \tag{18}$$

Let $\mathbf{x}^{\mathbf{a}} := \textsc{tm}(G)$, $S := \operatorname{supp}(\mathbf{x}^{\mathbf{a}})$, and $n = |S|$. If $n = 0$, then there is nothing to check. We may therefore suppose $n \geq 1$. By [For15, Lemma 4.12], there exists a morphism $\pi_S \colon K[\mathbf{x}] \to K[\mathbf{y}]$ that sends x_i to 0 for $i \notin S$ and to y_i for $i \in S$. This morphism restricts to an isomorphism $K[\mathbf{x}_S] \cong K[\mathbf{y}]$, and we equip $K[\mathbf{y}]$ with the induced monomial order. The restriction $\pi_S \colon R[\mathbf{x}] \to R[\mathbf{y}]$ commutes with specialization at $\varepsilon = 0$. Thus

$$\pi_S(G_\varepsilon) = G_\varepsilon(\pi_S(\mathbf{x})) = \sum_{i=1}^{m} f_i(g_{i,1}(\pi_S(\mathbf{x})), \ldots, g_{i,r}(\pi_S(\mathbf{x}))). \tag{19}$$

Thus $\textsc{tm}(\pi_S(G)) = \mathbf{y}^{\mathbf{a}}$ and $\pi_S(G_\varepsilon)|_{\varepsilon=0} = \pi_S(G) \neq 0$ by [For15, Lemma 4.12], where we treat $\mathbf{a} \in \mathbb{N}^N$ as $\mathbf{a}_S \in \mathbb{N}^n$.

The coefficient matrix of the shifted derivatives of $\pi_S(G_\varepsilon)$ has entries in R and specializes at $\varepsilon = 0$ to the coefficient matrix of the shifted derivatives of $\pi_S(G)$. Minors that vanish remain zero under specialization at $\varepsilon = 0$, so the rank cannot increase. By [For15, Corollary 4.5],

$$\begin{aligned}\dim_{\mathbb{F}} \operatorname{span}_{\mathbb{F}} \left(\mathbf{y}^{\leq \ell}\partial_{\mathbf{y}}^{\leq k}(\pi_S(G))\right) &\leq \dim_K \operatorname{span}_K \left(\mathbf{y}^{\leq \ell}\partial_{\mathbf{y}}^{\leq k}(\pi_S(G_\varepsilon))\right) \\ &\leq m\binom{k+r}{r}\binom{n+(d-1)k+\ell}{(d-1)k+\ell},\end{aligned}$$

where k and ℓ are to be chosen later. On the other hand, by [For15, Lemmas 4.8, 4.12, and 4.13], we have

$$\dim_{\mathbb{F}} \operatorname{span}_{\mathbb{F}} \left(\mathbf{y}^{\leq \ell}\partial_{\mathbf{y}}^{\leq k}(\pi_S(G))\right) \geq \dim_{\mathbb{F}} \operatorname{span}_{\mathbb{F}} \left(\mathbf{y}^{\leq \ell}\partial_{\mathbf{y}}^{\leq k}(\mathbf{y}^{\mathbf{a}})\right) \geq \binom{n}{k}\binom{n-k+\ell}{\ell}.$$

Thus, as in the proof of [For15, Proposition 4.18], we obtain

$$m \geq \frac{1}{\binom{k+r}{r}} \frac{\binom{n}{k}\binom{n-k+\ell}{\ell}}{\binom{n+(d-1)k+\ell}{(d-1)k+\ell}}.$$

As in [For15, Lemma A.3], choose $k = \left\lfloor \frac{n}{e^3 d} \right\rfloor$ and $\ell = (d-1)(n+(d-1)k)$. Using $\binom{k+r}{r} \leq 2^{k+r}$, we obtain

$$m \geq \frac{e^k}{2^{k+r}}, \qquad (1-\ln 2)k \leq \ln m + r \ln 2.$$

Since $k \geq \frac{n}{e^3 d} - 1$ and $r \geq 1$, we obtain

$$n \leq e^3 d \left(1 + \frac{\ln m + r \ln 2}{1 - \ln 2}\right) \leq \frac{e^3}{1-\ln 2} d(\ln m + r).$$

Thus $|\mathrm{supp}(\textsc{tm}(G))| = O\left(d(\log m + r)\right)$. □

6.2 Border Low-Weight Translation

In this subsection, we prove the border analogue of Theorem 5.3. It holds for *every* polynomial that is computed by a $\overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ circuit. We use the following notation for substitutions:

For $T \subseteq [N]$ and a tuple $g = (g_a)_{a \in [N] \setminus T}$, let $(\mathbf{x}_T, g)$ be the N-tuple whose a-th entry is x_a if $a \in T$ and g_a if $a \notin T$. Thus $p(\mathbf{x}_T, g)$ is obtained from a polynomial p by substituting g_a for x_a for every $a \notin T$. For $b \in \mathbb{F}^N$, the tuple $b + (\mathbf{x}_T, g)$ has entries $b_a + x_a$ for $a \in T$ and $b_a + g_a$ for $a \notin T$. The main theorem we want to prove is Theorem 6.7, which establishes the existence of a point b that has small support and satisfies a couple of key properties.

Theorem 6.7: Let $f \in \overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$, and write $f + \varepsilon S = \sum_{i=1}^{k} T_i$, where $S \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ and $T_i = \prod_j q_{i,j}$ for polynomials $q_{i,j} \in \mathbb{F}((\varepsilon))[\mathbf{x}]$ of degree at most δ. Let ρ satisfy $\rho \geq 1$ and $\left|\{\, q_{i,j} \mid i,j \,\}\right| \leq \rho$, meaning at most ρ of the polynomials $q_{i,j}$ are distinct as elements of $\mathbb{F}((\varepsilon))[\mathbf{x}]$. There exist a set $T \subseteq [N]$ with $N - |T| \leq k$ and a point $b \in \mathbb{F}^N$ with $|\mathrm{supp}(b)| = O\left(k^2\Delta^3\right)$, where $\mathbf{u} \coloneqq \mathbf{x}_T$, such that the following properties hold.

1. For every $a \in [N] \setminus T$, there exist a power series $\varphi_a \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$ and a polynomial $Q_a \in \mathbb{F}[[\varepsilon]][\mathbf{u}, x_a]$ of degree at most Δ such that

$$\varphi_a(\varepsilon, \mathbf{0}) = 0, \qquad Q_a(\mathbf{u}, \varphi_a) = 0, \qquad \text{and } (\partial_{x_a} Q_a)(\mathbf{u}, \varphi_a) \text{ is an } \varepsilon\text{-unit.}$$

2. Write $\boldsymbol{\varphi} \coloneqq (\varphi_a)_{a \in [N] \setminus T}$. There exist an integer k' with $0 \leq k' \leq k$, integers $e_1, \ldots, e_{k'}$, and polynomials $p_{i,j} \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ of degree at most δ with $\left|\{\, p_{i,j} \mid i,j \,\}\right| \leq \rho$ such that every $p_{i,j}(\mathbf{u}, \boldsymbol{\varphi})$ is an ε-unit and

$$\sum_{i=1}^{k'} \varepsilon^{e_i} \prod_j p_{i,j}(\mathbf{u}, \boldsymbol{\varphi}) \equiv f(b + (\mathbf{u}, \boldsymbol{\varphi})) \pmod{\varepsilon}. \tag{20}$$

The integers e_i may be negative, so the sum in (20) is computed in the ring $\mathbb{F}[[\varepsilon, \mathbf{u}]]_\varepsilon$. In particular, (20) says that the negative powers of ε cancel in the sum. Moreover, since f does not involve ε and $\varphi_a \equiv \varphi_a(0, \mathbf{u})$ modulo ε, we may replace $\boldsymbol{\varphi}$ by $\boldsymbol{\varphi}(0, \mathbf{u}) := (\varphi_a(0, \mathbf{u}))_{a \in [N] \setminus T}$ on the right side. In other words, the sum is congruent modulo ε to $f(b + (\mathbf{u}, \boldsymbol{\varphi}(0, \mathbf{u}))) \in \mathbb{F}[[\mathbf{u}]]$.

Remark 6.8: In Theorem 5.3, the dependent coordinates are power series in the free coordinates by the implicit function theorem (Lemma 4.14). In the border setting, their equations may ramify at $\varepsilon = 0$, and the Abhyankar–Jung Theorem (Theorem 6.3) takes the place of the implicit function theorem. This possible ramification is why the proof replaces ε by a power ε^e.

Over the field $\mathbb{F}((\varepsilon))$, the implicit function theorem still applies to Item 1 of Theorem 6.7. Indeed, $\varphi_a(\varepsilon, \mathbf{0}) = 0$ gives $Q_a(\mathbf{0}) = 0$. Since $(\partial_{x_a} Q_a)(\mathbf{u}, \varphi_a)$ is an ε-unit, its value $(\partial_{x_a} Q_a)(\mathbf{0})$ is a nonzero element of $\mathbb{F}[[\varepsilon]]$. Thus Q_a determines φ_a. The point of Theorem 6.7 is the *stronger* claim that φ_a has no negative powers of ε and that the series $p_{i,j}(\mathbf{u}, \boldsymbol{\varphi})$ and $(\partial_{x_a} Q_a)(\mathbf{u}, \varphi_a)$ are ε-units. These properties are essential for Section 6.3, where we divide by these series and then let $\varepsilon \to 0$: the ε-unit property guarantees that the divisions never produce unbounded negative powers of ε, so the limit makes sense.

Moreover, we *cannot* state Theorem 6.7 exactly in the form of Theorem 5.3. There, the conditions $Q_a(\mathbf{0}) = 0$ and $(\partial_{x_a} Q_a)(\mathbf{0}) \neq 0$ identify the quotient $\mathbb{F}[[\mathbf{x}]] / \langle Q_{m+1}, \dots, Q_N \rangle$ with $\mathbb{F}[[x_1, \dots, x_m]]$ (Lemma 4.14), and the identity is stated in this quotient. Applying Lemma 4.14 in the border setting with ε as an additional free variable gives $\mathbb{F}[[\varepsilon, \mathbf{x}]] / \langle Q_a \mid a \in [N] \setminus T \rangle \cong \mathbb{F}[[\varepsilon, \mathbf{u}]]$, provided that no $(\partial_{x_a} Q_a)(\mathbf{0})$ is divisible by ε. In general, however, $(\partial_{x_a} Q_a)(\mathbf{0})$ may be divisible by ε, and then this isomorphism can fail.

For example, let $\mathbf{u} = x_2$, and consider the ramified equation $x_1^2 = \varepsilon(1 + x_2)$. As in the proof, we replace ε by ε^2 and translate by the point $(\varepsilon, 0)$ of the branch $x_1 = \varepsilon\sqrt{1 + x_2}$ above $x_2 = 0$. This translation gives

$$Q := x_1^2 + 2\varepsilon x_1 - \varepsilon^2 x_2 = (x_1 - \varphi)(x_1 + 2\varepsilon + \varphi), \qquad \text{where } \varphi := \varepsilon\Big(\sqrt{1 + x_2} - 1\Big) \in \mathbb{F}[[\varepsilon, x_2]].$$

The branch $x_1 = \varphi$ is as in Item 1 of Theorem 6.7, since $(\partial_{x_1} Q)(x_2, \varphi) = 2\varepsilon\sqrt{1 + x_2}$ is an ε-unit. However, $(\partial_{x_1} Q)(\mathbf{0}) = 2\varepsilon$ is divisible by ε. In fact, both factors of Q vanish at $\varepsilon = x_1 = x_2 = 0$, so neither of them is a unit of $\mathbb{F}[[\varepsilon, x_1, x_2]]$, and $\mathbb{F}[[\varepsilon, x_1, x_2]] / \langle Q \rangle$ is not even a domain! For this reason, we substitute the power series φ_a and state (20) in $\mathbb{F}[[\varepsilon, \mathbf{u}]]$, where a congruence modulo ε makes sense. Over the field $\mathbb{F}((\varepsilon))$, Lemma 4.14 does apply, as noted above, but there ε is a unit, so a congruence modulo ε is meaningless.

Similarly, the condition of Theorem 5.3 that no factor vanishes at the origin becomes the stronger condition that every $p_{i,j}(\mathbf{u}, \boldsymbol{\varphi})$ is an ε-unit. For instance, $\varepsilon + x_2$ takes the nonzero value ε at the origin, but its reciprocal has the expansion $1/(\varepsilon + x_2) = \sum_{d \geq 0} (-x_2)^d \varepsilon^{-d-1}$, with unbounded negative powers of ε.

Finally, the left side in Theorem 5.3 is an explicit product, because its proof first passes to a subvariety on which only one block of $P_{k+1,n}$ survives (Lemma 2.20). Here we keep a general f, since the proof of Theorem 6.1 uses only the facts that $P_{k+1,n}(b + (\mathbf{u}, \boldsymbol{\varphi}(0, \mathbf{u})))$ is nonzero and that each of its monomials involves many variables. Both facts follow directly from Lemma 6.20. $\triangleleft$

The proof of Theorem 6.7 has three main steps: choosing the gates to cancel (Lemma 6.9), finding a point with few nonzero coordinates (Lemma 6.10), and constructing the required power series (Lemma 6.11). We state the three lemmas below and use them to prove the theorem. The proof of Lemma 6.11 is given afterwards.

Choosing the Gates for Cancellation. We seek a prime ideal modulo which some gates vanish and the factors of the remaining gates become invertible after inverting ε and the polynomials nonzero at the origin. For this purpose, let $\mathcal{O}_0 := (\mathbb{F}[[\varepsilon]][\mathbf{x}]_{\langle \varepsilon, \mathbf{x} \rangle})_\varepsilon$. Its elements are fractions $g/(\varepsilon^c h)$, where $g, h \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$, $c \geq 0$, and $h(0, \mathbf{0}) \neq 0$. In particular, it contains $\mathbb{F}((\varepsilon))[\mathbf{x}]$. By the correspondence for prime ideals under localization (see Localization), the map $\mathfrak{p} \mapsto \mathfrak{p} \cap \mathbb{F}[[\varepsilon]][\mathbf{x}]$ gives a bijection between the prime ideals of $\mathcal{O}_0$ and the prime ideals I of $\mathbb{F}[[\varepsilon]][\mathbf{x}]$ satisfying $I \subseteq \langle \varepsilon, \mathbf{x} \rangle$ and $\varepsilon \notin I$.

The following lemma is the border analogue of Lemma 2.21. We again inspect the gates one by one. Whenever a factor is a non-unit modulo the current ideal of $\mathcal{O}_0$, we add it to the ideal, making its gate vanish. Otherwise, every factor of that gate is already invertible.

Lemma 6.9: Let $T_1, \ldots, T_k \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$, with $T_i = \prod_j q_{i,j}$, where each $q_{i,j} \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ is nonzero and has degree at most δ in $\mathbf{x}$. There exist a set $K \subseteq [k]$, indexing *surviving gates*, and a prime ideal $\mathfrak{p} \subseteq \mathbb{F}((\varepsilon))[\mathbf{x}]$ such that $I := \mathfrak{p} \cap \mathbb{F}[[\varepsilon]][\mathbf{x}]$ has the following properties.

1. $I \subseteq \langle \varepsilon, x_1, \ldots, x_N \rangle$.
2. $\operatorname{ht} \mathfrak{p} \leq k - |K|$ and $\operatorname{DEG} \mathfrak{p} \leq \delta^{k-|K|}$.
3. $T_i \in I$ for every $i \in [k] \setminus K$.
4. For every $i \in K$ and every j, there are an integer $t \geq 0$ and polynomials $h, w \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ with $h(0, \mathbf{0}) \neq 0$ and $q_{i,j} w \equiv \varepsilon^t h \pmod{I}$.

Item 4 is the border analogue of saying that $q_{i,j}$ is a unit on the variety: it says that $q_{i,j}$ is invertible near the origin, up to a power of ε. Indeed, ε and h are invertible in $\mathcal{O}_0$, since $h(0, \mathbf{0}) \neq 0$, so dividing $q_{i,j} w \equiv \varepsilon^t h$ by $\varepsilon^t h$ shows that $w/(\varepsilon^t h)$ is an inverse of $q_{i,j}$ modulo $I\mathcal{O}_0$.

Proof: We go through the gates one by one and build an increasing chain of proper ideals $\mathfrak{a}_0 \subseteq \mathfrak{a}_1 \subseteq \cdots \subseteq \mathfrak{a}_k$ of $\mathcal{O}_0$, starting with $\mathfrak{a}_0 := \{0\}$. Simultaneously, we build a chain of subsets $K_0 \subseteq K_1 \subseteq \cdots \subseteq K_k$ of $[k]$, where $K_i \subseteq [i]$ and we start with $K_0 := \emptyset$. Let $i \geq 1$. Suppose that $\mathfrak{a}_{i-1}$ and K_{i-1} are already defined, and consider the product gate T_i.

- If some factor of T_i is not a unit modulo $\mathfrak{a}_{i-1}$, then we pick one such factor, call it the *selected factor* q_i of T_i, and let $\mathfrak{a}_i := \mathfrak{a}_{i-1} + q_i \mathcal{O}_0$. This ideal is proper precisely because q_i is not a unit modulo $\mathfrak{a}_{i-1}$. The gate T_i does not survive, and we let $K_i := K_{i-1}$.
- Otherwise, every factor of T_i is a unit modulo $\mathfrak{a}_{i-1}$. We let $\mathfrak{a}_i := \mathfrak{a}_{i-1}$, and the gate T_i survives. In this situation, let $K_i := K_{i-1} \cup \{i\}$.

In the end, let $\mathfrak{a} := \mathfrak{a}_k$ and $K := K_k$. We have that $\mathfrak{a}$ is a proper ideal of $\mathcal{O}_0$ generated by the $k - |K|$ selected factors. Let $\mathfrak{P}$ be a minimal prime of $\mathfrak{a}$, and let $\mathfrak{p} := \mathfrak{P} \cap \mathbb{F}((\varepsilon))[\mathbf{x}]$, which is a prime ideal of $\mathbb{F}((\varepsilon))[\mathbf{x}]$. For $I := \mathfrak{p} \cap \mathbb{F}[[\varepsilon]][\mathbf{x}]$ as in the statement, we have $I = \mathfrak{P} \cap \mathbb{F}[[\varepsilon]][\mathbf{x}]$.

We now check the four properties.

Item 1. By the correspondence between the primes of $\mathcal{O}_0$ and those of $\mathbb{F}[[\varepsilon]][\mathbf{x}]$, the prime $I = \mathfrak{P} \cap \mathbb{F}[[\varepsilon]][\mathbf{x}]$ is contained in $\langle \varepsilon, \mathbf{x} \rangle$.

Item 2. We show that $\mathfrak{p}$ is a minimal prime of $\langle q_i \mid i \in [k] \setminus K \rangle$; the required height and degree bounds then follow from Lemma 2.1 and Lemma 2.10. So let $\mathfrak{p}' \subseteq \mathfrak{p}$ be a prime of $\mathbb{F}((\varepsilon))[\mathbf{x}]$ that contains $\langle q_i \mid i \in [k] \setminus K \rangle$; we show that $\mathfrak{p}' = \mathfrak{p}$. Let $I' := \mathfrak{p}' \cap \mathbb{F}[[\varepsilon]][\mathbf{x}]$. Consequently, $I' \subseteq I \subseteq \langle \varepsilon, \mathbf{x} \rangle$, and $\varepsilon \notin I'$, since ε is a unit of $\mathbb{F}((\varepsilon))[\mathbf{x}]$. By the correspondence between the primes of $\mathbb{F}[[\varepsilon]][\mathbf{x}]$ and those of $\mathcal{O}_0$, it follows that $I'\,\mathcal{O}_0$ is a prime of $\mathcal{O}_0$. It contains $\mathfrak{a}$, since the selected factors lie in I', and it is contained in $\mathfrak{P}$, since $I' \subseteq I \subseteq \mathfrak{P}$. By the minimality of $\mathfrak{P}$, we get $I'\,\mathcal{O}_0 = \mathfrak{P}$, and contracting back gives $I' = \mathfrak{P} \cap \mathbb{F}[[\varepsilon]][\mathbf{x}] = I$. Since $\mathbb{F}((\varepsilon))[\mathbf{x}]$ is a localization of $\mathbb{F}[[\varepsilon]][\mathbf{x}]$, we conclude that

$$\mathfrak{p}' = I'\,\mathbb{F}((\varepsilon))[\mathbf{x}] = I\,\mathbb{F}((\varepsilon))[\mathbf{x}] = \mathfrak{p}.$$

Thus $\mathfrak{p}$ is a minimal prime of $\langle q_i \mid i \in [k] \setminus K \rangle$ generated by the $k - |K|$ selected factors, each of degree at most δ. Lemma 2.1 and Lemma 2.10 then give $\operatorname{ht}\mathfrak{p} \le k - |K|$ and $\operatorname{DEG}\mathfrak{p} \le \delta^{k-|K|}$.

Item 3. If $i \in [k] \setminus K$, then the selected factor q_i lies in $\mathfrak{a} \cap \mathbb{F}[[\varepsilon]][\mathbf{x}] \subseteq I$, and hence so does its multiple T_i.

Item 4. Let $i \in K$. Every factor $q_{i,j}$ of T_i is a unit modulo $\mathfrak{a}_{i-1}$ and hence also modulo the larger ideal $\mathfrak{a}$. Thus there is an element $w/(\varepsilon^t h) \in \mathcal{O}_0$, with $w, h \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ and $h(0, \mathbf{0}) \neq 0$, such that $q_{i,j} \cdot w/(\varepsilon^t h) - 1 \in \mathfrak{a} \subseteq \mathfrak{P}$. Multiplying by $\varepsilon^t h$ gives $q_{i,j} w - \varepsilon^t h \in \mathfrak{P} \cap \mathbb{F}[[\varepsilon]][\mathbf{x}] = I$. □

A Low-Weight Point. To apply Lemma 6.11, we need a point b on the variety at which certain polynomials do not vanish. These polynomials are of two kinds:

- In the proof of Theorem 6.7, we construct equations $\widetilde{Q}_a$ for the dependent coordinates. The first kind consists of the leading coefficients of $\widetilde{Q}_a(0, \mathbf{u}, x_a)$ in x_a and the polynomials obtained from the resultants of $\widetilde{Q}_a$ and $\partial_{x_a}\widetilde{Q}_a$ by dividing each resultant by its maximal power of ε and setting $\varepsilon = 0$. Their product is a single polynomial g of degree at most $C\Delta^2$; in the proof of Theorem 6.7, this is the polynomial R.
- The second kind consists of the polynomials $h(0, \mathbf{x})$ from Lemma 6.9, one for each distinct factor of the surviving gates.

At the same time, b should have few nonzero coordinates, since the final bound loses one variable of $P_{k+1,n}$ for every nonzero coordinate of b.

In the exact setting, Lemma 5.1 combines all the resultants into a single polynomial and finds a point at which it does not vanish, with the number of nonzero coordinates bounded in terms of its degree. In the border setting, taking the product of these resultants and all the polynomials $h(0, \mathbf{x})$ would not give a useful bound: there may be as many polynomials h as there are distinct bottom polynomials, and we have no control over their degrees.

Instead, we use that every $h(0, \mathbf{x})$ is nonzero at the origin, so none of them vanishes identically on an irreducible variety through the origin. Hence a *generic* point of such a variety avoids the zeros of all of them at once, however many there are. Lemma 6.10 finds, inside a small coordinate subspace, such a variety on which g does not vanish identically either. Thus the polynomials h come for free, and only the degree of g enters the bound.

Lemma 6.10: Let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety of codimension C that contains $\mathbf{0}$. Let $g \in \mathbb{F}[\mathbf{x}]$ be a polynomial that does not vanish identically on V, and let $h_1, \ldots, h_t \in \mathbb{F}[\mathbf{x}]$ be polynomials with $h_j(\mathbf{0}) \neq 0$ for all $j \in [t]$. There exists a point $b \in V$ such that $g(b) \cdot h_1(b) \cdots h_t(b) \neq 0$ and $|\mathrm{supp}(b)| \leq (C+1) \cdot \deg V \cdot \deg g$.

Note that the bound on $|\mathrm{supp}(b)|$ depends neither on t nor on the degrees of the h_j.

Proof: If $g(\mathbf{0}) \neq 0$, then take $b = \mathbf{0}$. Hence assume that $g(\mathbf{0}) = 0$.

For $E \subseteq [N]$, define the $|E|$-dimensional subspace $L_E := \bigcap_{i \in [N] \setminus E} \mathbf{V}(x_i)$. Choose $E \subseteq [N]$ with $|E|$ minimal such that $V \cap L_E$ has an irreducible component Z that contains the origin and on which g does not vanish identically. Such a choice exists because $E = [N]$ works. By Lemma 2.7 and Proposition 2.9, we have $\dim Z \geq |E| - C$ and $\deg Z \leq \deg V$.

Fix $i \in E$. By the minimality of $|E|$, the coordinate x_i does not vanish identically on Z: otherwise, Z would also be an irreducible component of $V \cap L_{E \setminus \{i\}}$. Let Z_i be an irreducible component of $Z \cap \mathbf{V}(x_i)$ containing the origin. By Lemma 2.7, $\dim Z_i = \dim Z - 1$. Moreover, g vanishes identically on Z_i: otherwise, an irreducible component of $V \cap L_{E \setminus \{i\}}$ containing Z_i would contradict the minimality of $|E|$. Thus $Z_i \subseteq Z \cap \mathbf{V}(g)$. Moreover, since g does not vanish identically on Z, Lemma 2.7 implies that every irreducible component of $Z \cap \mathbf{V}(g)$ has dimension $\dim Z - 1$, which equals $\dim Z_i$. Therefore, Z_i is an irreducible component of $Z \cap \mathbf{V}(g)$.

By Bézout's Inequality (Proposition 2.9), $Z \cap \mathbf{V}(g)$ has at most $\deg Z \cdot \deg g$ irreducible components. Each such component Y satisfies $\dim Y = \dim Z - 1 \geq |E| - C - 1$. If Y lies in r distinct coordinate hyperplanes $\mathbf{V}(x_i)$ with $i \in E$, then $\dim Y \leq |E| - r$, since $Y \subseteq L_E$. Hence $r \leq C + 1$. Each $i \in E$ determines an irreducible component Z_i of $Z \cap \mathbf{V}(g)$, and the same component can occur for at most $C + 1$ indices. Hence

$$|E| \leq (C+1) \cdot \deg Z \cdot \deg g \leq (C+1) \cdot \deg V \cdot \deg g.$$

Finally, no h_j vanishes identically on Z, because $\mathbf{0} \in Z$ and $h_j(\mathbf{0}) \neq 0$. Since Z is irreducible and g does not vanish identically on Z, neither does $g h_1 \cdots h_t$. Choose $b \in Z$ with $g(b) \cdot h_1(b) \cdots h_t(b) \neq 0$. The point b lies in V and satisfies $\mathrm{supp}(b) \subseteq E$, as required. □

A Power-Series Parametrization. The next lemma is the border analogue of Lemma 4.15. There, a nonzero Jacobian determinant allows us to apply the implicit function theorem (Lemma 4.14) and express the dependent coordinates as power series in the free coordinates. In the border setting, however, the equations also depend on ε, and the Jacobian determinant may vanish at $\varepsilon = 0$, so the implicit function theorem no longer applies directly.

Instead, we first make each equation monic by Weierstrass preparation (Theorem 6.2). The condition on the resultants in Item 2 of the next lemma ensures that the discriminant of the resulting monic polynomial is an ε-unit, so ramification occurs only along $\varepsilon = 0$. The Abhyankar–Jung Theorem (Theorem 6.3) then expresses the roots as power series, after replacing ε by a suitable power ε^e. The remaining task is to choose these roots compatibly: substituting them for all dependent coordinates at once must make every polynomial in the prime ideal I of the next lemma vanish, not only the individual equations $\widetilde{Q}_a$.

Lemma 6.11: Let $I \subseteq \mathbb{F}[[\varepsilon]][\mathbf{x}]$ be a prime ideal with $\varepsilon \notin I$, and let $b \in \mathbb{F}^N$ satisfy $g(0, b) = 0$ for every $g \in I$. Let $T \subseteq [N]$ satisfy $\dim(\mathbb{F}((\varepsilon))[\mathbf{x}]/I\mathbb{F}((\varepsilon))[\mathbf{x}]) = |T|$, and write $\mathbf{u} := \mathbf{x}_T$ and $b_T := (b_i)_{i \in T}$. Suppose that for every $a \in [N] \setminus T$ there is a polynomial $\widetilde{Q}_a \in I \cap \mathbb{F}[[\varepsilon]][\mathbf{u}, x_a]$ with the following properties.

1. $\widetilde{Q}_a(0, b_T, x_a) \in \mathbb{F}[x_a]$ is nonzero.
2. There are an integer $j_a \geq 0$ and a polynomial $R_a \in \mathbb{F}[[\varepsilon]][\mathbf{u}]$ such that $\operatorname{Res}_{x_a}(\widetilde{Q}_a, \partial_{x_a}\widetilde{Q}_a) = \varepsilon^{j_a} R_a$ and $R_a(0, b_T) \neq 0$.

There exist an integer $e \geq 1$ and power series $r_a \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$ with $r_a(0, \mathbf{0}) = 0$ for each $a \in [N] \setminus T$ such that the following property holds. Write $\mathbf{r} := (r_a)_{a \in [N] \setminus T}$, and define the ring homomorphism $\sigma\colon \mathbb{F}[[\varepsilon]][\mathbf{x}] \to \mathbb{F}[[\varepsilon, \mathbf{u}]]$ by $\sigma(g) = g(\varepsilon^e, b + (\mathbf{u}, \mathbf{r}))$. This homomorphism satisfies $I \subseteq \ker \sigma$. Moreover, $\sigma(\partial_{x_a}\widetilde{Q}_a)$ is an ε-unit for every $a \in [N] \setminus T$.

Here replacing ε by ε^e means applying $a(\varepsilon) \mapsto a(\varepsilon^e)$ to each coefficient series. The condition $\sigma(I) = \{0\}$ ensures that the chosen series satisfy every relation in I, not only the individual equations $\widetilde{Q}_a = 0$.

We now give the proof of Theorem 6.7. We delay the proof of Lemma 6.11 until afterwards.

Proof of Theorem 6.7: We first simplify the circuit. If every gate is zero, then $f = 0$, and the conclusion holds with $T = [N]$, $b = \mathbf{0}$, and $k' = 0$. Otherwise, after discarding the zero gates and decreasing k accordingly, we may assume that all k gates are nonzero. Next, we normalize every bottom polynomial q by setting $\widetilde{q} := \varepsilon^{-\operatorname{val}_\varepsilon(q)} q$ (see Limits and Normalization). The polynomial $\widetilde{q}$ lies in $\mathbb{F}[[\varepsilon]][\mathbf{x}]$ and does not vanish at $\varepsilon = 0$. We move the extracted powers of ε in front of the gates. There are still at most ρ distinct bottom polynomials, since $\widetilde{q}$ is determined by q. Finally, we multiply by a sufficiently large power of ε, so that no negative powers of ε remain in front of the gates. Writing again $q_{i,j}$ for the normalized bottom polynomials, we obtain

$$\varepsilon^d f = \sum_{i=1}^{k} \varepsilon^{d_i} T_i + \varepsilon^{d+1} S, \tag{21}$$

where $d, d_i \geq 0$, $S \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$, and $T_i = \prod_j q_{i,j}$ for nonzero polynomials $q_{i,j} \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ of degree at most δ in $\mathbf{x}$, with $\left|\{\, q_{i,j} \mid i, j \,\}\right| \leq \rho$.

We then choose the gates to cancel by applying Lemma 6.9 to $T_1, \ldots, T_k$. Let K, $\mathfrak{p}$, and I be its output. The gates T_i with $i \notin K$ lie in I, and those with $i \in K$ survive. Let $C := \operatorname{ht}\mathfrak{p}$. Hence $C \leq k - |K| \leq k$ and $\operatorname{DEG}\mathfrak{p} \leq \delta^{k-|K|} \leq \Delta$.

To choose the free coordinates, we set $\varepsilon = 0$. Let $J_0 := \{\, g(0, \mathbf{x}) \mid g \in I \,\}$. Since $I \subseteq \langle \varepsilon, \mathbf{x} \rangle$ by Lemma 6.9, Lemma 4.8 applies. It shows that $\mathbf{0} \in \mathbf{V}(J_0)$ and that every irreducible component of $\mathbf{V}(J_0)$ has codimension C and degree at most $\operatorname{DEG}\mathfrak{p}$, which is at most Δ. Let V be an irreducible component of $\mathbf{V}(J_0)$ that contains $\mathbf{0}$. By Lemma 2.13, applied to V and the empty set of coordinates, there is a set $T \subseteq [N]$ with $|T| = \dim V = N - C$ such that the coordinates $\mathbf{u} := \mathbf{x}_T$ satisfy $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{u}] = \{0\}$. These are our free coordinates, and the x_a with $a \notin T$ are the dependent ones.

The free coordinates remain algebraically independent modulo $\mathfrak{p}$, that is, $\mathfrak{p} \cap \mathbb{F}((\varepsilon))[\mathbf{u}] = \{0\}$.

Indeed, multiplying a nonzero element of this intersection by a suitable integer power of ε would give a polynomial $g \in \mathfrak{p} \cap \mathbb{F}[[\varepsilon]][\mathbf{u}]$ with $g(0,\mathbf{u}) \neq 0$. The polynomial g would lie in $\mathfrak{p} \cap \mathbb{F}[[\varepsilon]][\mathbf{x}] = I$. Hence $g(0,\mathbf{u})$ would be a nonzero polynomial in $\mathbf{u}$ lying in $J_0 \subseteq \mathbf{I}(V)$, contradicting the fact that $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{u}] = \{0\}$. This argument proves the claim that $\mathfrak{p} \cap \mathbb{F}((\varepsilon))[\mathbf{u}] = \{0\}$. Since $\operatorname{ht}\mathfrak{p} = C$, Lemma 2.1 also gives $\dim(\mathbb{F}((\varepsilon))[\mathbf{x}]/\mathfrak{p}) = N - C = |T|$.

For every dependent coordinate, we now find an equation of the kind needed in Lemma 6.11. Fix $a \in [N] \setminus T$. By the previous paragraph, Lemma 2.16 applies over the field $\mathbb{F}((\varepsilon))$, showing that the ideal $\mathfrak{p} \cap \mathbb{F}((\varepsilon))[\mathbf{u}, x_a]$ is generated by an irreducible polynomial $\widetilde{Q}_a$ with $\deg_{x_a} \widetilde{Q}_a \geq 1$ and $\deg \widetilde{Q}_a \leq \operatorname{DEG}\mathfrak{p} \leq \Delta$. After multiplying by a suitable integer power of ε, we may assume that $\widetilde{Q}_a \in \mathbb{F}[[\varepsilon]][\mathbf{u}, x_a]$ and $\widetilde{Q}_a(0,\mathbf{u},x_a) \neq 0$, so $\widetilde{Q}_a \in I$. The polynomial $\widetilde{Q}_a(0,\mathbf{u},x_a)$ lies in J_0 and hence vanishes on V; it must involve x_a since $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{u}] = \{0\}$. Let $c_a \in \mathbb{F}[\mathbf{u}]$ be its leading coefficient as a polynomial in x_a; thus $c_a \neq 0$ and $\deg c_a \leq \Delta - 1$. (The coefficient c_a can differ from the value at $\varepsilon = 0$ of the leading coefficient of $\widetilde{Q}_a$ in x_a, since the degree in x_a may drop at $\varepsilon = 0$, as for $\varepsilon x_a^2 + x_a - x_t$ with $t \in T$.)

By Proposition 4.13, applied over $\mathbb{F}((\varepsilon))$, the resultant $\operatorname{Res}_{x_a}(\widetilde{Q}_a, \partial_{x_a}\widetilde{Q}_a) \in \mathbb{F}[[\varepsilon]][\mathbf{u}]$ is nonzero and has degree at most $\Delta(\Delta-1)$ in $\mathbf{u}$. Extracting its maximal power of ε, we can write

$$\operatorname{Res}_{x_a}(\widetilde{Q}_a, \partial_{x_a}\widetilde{Q}_a) = \varepsilon^{j_a} R_a, \qquad \text{where } j_a \geq 0,\ R_a \in \mathbb{F}[[\varepsilon]][\mathbf{u}],\ \text{and } R_a(0,\mathbf{u}) \neq 0. \tag{22}$$

We now choose the point b. Two families of polynomials must not vanish at b. First, the hypotheses of Lemma 6.11 at $(0,b)$ will follow from $c_a(b_T) \neq 0$ and $R_a(0,b_T) \neq 0$ for all $a \notin T$. Define

$$R \coloneqq \prod_{a \in [N]\setminus T} (c_a \cdot R_a)(0,\mathbf{u}) \in \mathbb{F}[\mathbf{u}],$$

with $R = 1$ if $T = [N]$. It is a nonzero polynomial of degree at most $C(\Delta - 1 + \Delta(\Delta-1)) \leq C\Delta^2$, and it does not vanish identically on V, because $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{u}] = \{0\}$. Second, the factors of the surviving gates should become ε-units. For every $i \in K$ and every j, Lemma 6.9 provides an integer $t_{i,j} \geq 0$ and polynomials $h_{i,j}, w_{i,j} \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ with

$$\varepsilon^{t_{i,j}} h_{i,j} - q_{i,j} w_{i,j} \in I \text{ and } h_{i,j}(0,\mathbf{0}) \neq 0. \tag{23}$$

To make the factors $q_{i,j}$ into ε-units on the chosen branch, we will require $h_{i,j}(0,b) \neq 0$ for every $i \in K$ and every j. Apply Lemma 6.10 to V, to $R \in \mathbb{F}[\mathbf{u}] \subseteq \mathbb{F}[\mathbf{x}]$ in place of g, and to the finitely many polynomials $h_{i,j}(0,\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, which do not vanish at the origin. We obtain a point $b \in V$ such that $R(b) \neq 0$, $h_{i,j}(0,b) \neq 0$ for all $i \in K$ and all j, and $|\operatorname{supp}(b)| \leq (C+1)\cdot\Delta\cdot C\Delta^2 = O\left(k^2\Delta^3\right)$.

We can now apply Lemma 6.11 to I, T, b, and the polynomials $\widetilde{Q}_a$. The ideal $I = \mathfrak{p} \cap \mathbb{F}[[\varepsilon]][\mathbf{x}]$ is prime since $\mathfrak{p}$ is prime, and $\varepsilon \notin I$ since ε is a unit of $\mathbb{F}((\varepsilon))[\mathbf{x}]$. Moreover, $I\mathbb{F}((\varepsilon))[\mathbf{x}] = \mathfrak{p}$, so

$$\dim\left(\mathbb{F}((\varepsilon))[\mathbf{x}]/I\mathbb{F}((\varepsilon))[\mathbf{x}]\right) = \dim(\mathbb{F}((\varepsilon))[\mathbf{x}]/\mathfrak{p}) = N - C = |T|.$$

Every $g \in I$ satisfies $g(0,b) = 0$, since $g(0,\mathbf{x}) \in J_0$ and $b \in V \subseteq \mathbf{V}(J_0)$. Let $b_T \coloneqq (b_i)_{i\in T}$. Item 1 holds because $c_a(b_T) \neq 0$, and Item 2 is (22) with $R_a(0,b_T) \neq 0$. Both nonvanishing conditions follow from $R(b) \neq 0$. Thus the lemma provides an integer $e \geq 1$ and power series $r_a \in \mathbb{F}[[\varepsilon,\mathbf{u}]]$

with $r_a(0,\mathbf{0}) = 0$ for $a \in [N] \setminus T$. Writing $\mathbf{r} := (r_a)_{a\in[N]\setminus T}$, define the ring homomorphism $\sigma\colon \mathbb{F}[[\varepsilon]][\mathbf{x}] \to \mathbb{F}[[\varepsilon,\mathbf{u}]]$ by $\sigma(g) = g(\varepsilon^e, b + (\mathbf{u},\mathbf{r}))$. This homomorphism satisfies $\sigma(I) = \{0\}$, and $\sigma(\partial_{x_a}\widetilde{Q}_a)$ is an ε-unit for every $a \in [N] \setminus T$.

The theorem requires $\varphi_a(\varepsilon,\mathbf{0}) = 0$ as an identity in ε, whereas we currently know only $r_a(0,\mathbf{0}) = 0$. For $a \in [N] \setminus T$, let $\varphi_a := r_a - r_a(\varepsilon,\mathbf{0})$, so that $\varphi_a(\varepsilon,\mathbf{0}) = 0$. Let $b(\varepsilon) \in \mathbb{F}[[\varepsilon]]^N$ be the point with coordinates $b(\varepsilon)_i := b_i$ for $i \in T$ and $b(\varepsilon)_a := b_a + r_a(\varepsilon,\mathbf{0})$ for $a \in [N] \setminus T$. Since $r_a(0,\mathbf{0}) = 0$ for every $a \in [N] \setminus T$, we have $b(0) = b$. Define $\tau\colon \mathbb{F}[[\varepsilon]][\mathbf{x}] \to \mathbb{F}[[\varepsilon]][\mathbf{x}]$ by

$$\tau(g) = g(\varepsilon^e, b(\varepsilon) + \mathbf{x}).$$

That is, τ substitutes ε^e for ε and translates $\mathbf{x}$ by the point $b(\varepsilon)$. It is a ring homomorphism, does not increase the degree in $\mathbf{x}$, and commutes with the partial derivatives ∂_{x_a}. By the definition of $b(\varepsilon)$, we have $b(\varepsilon) + (\mathbf{u},\boldsymbol{\varphi}) = b + (\mathbf{u},\mathbf{r})$. It follows that

$$\tau(g)(\mathbf{u},\boldsymbol{\varphi}) = g(\varepsilon^e, b(\varepsilon) + (\mathbf{u},\boldsymbol{\varphi})) = g(\varepsilon^e, b + (\mathbf{u},\mathbf{r})) = \sigma(g) \tag{24}$$

for $g \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$. In other words, σ is τ followed by the substitution of φ_a for x_a for every $a \in [N] \setminus T$.

For Item 1 of the theorem, let $Q_a := \tau(\widetilde{Q}_a)$. Like $\widetilde{Q}_a$, it involves only $\mathbf{u}$ and x_a, so $Q_a \in \mathbb{F}[[\varepsilon]][\mathbf{u}, x_a]$, and it has degree at most Δ. By (24), $Q_a(\mathbf{u},\varphi_a) = \sigma(\widetilde{Q}_a) = 0$, since $\widetilde{Q}_a \in I$. Moreover, $\partial_{x_a} Q_a = \tau(\partial_{x_a}\widetilde{Q}_a)$, so (24) gives $(\partial_{x_a} Q_a)(\mathbf{u},\varphi_a) = \sigma(\partial_{x_a}\widetilde{Q}_a)$, which is an ε-unit by Lemma 6.11.

For Item 2 of the theorem, let $e_i := e(d_i - d)$ and $p_{i,j} := \tau(q_{i,j})$ for $i \in K$. These polynomials lie in $\mathbb{F}[[\varepsilon]][\mathbf{x}]$ and have degree at most δ. Since τ maps equal polynomials to equal polynomials, we have $\big|\{\, p_{i,j} \mid i \in K, j \,\}\big| \le \big|\{\, q_{i,j} \mid i, j \,\}\big| \le \rho$. By (24), $p_{i,j}(\mathbf{u},\boldsymbol{\varphi}) = \sigma(q_{i,j})$, and we show that this is an ε-unit. Since $\varepsilon^{t_{i,j}} h_{i,j} - q_{i,j} w_{i,j} \in I$ by (23) and $\sigma(I) = \{0\}$, we have $\sigma(q_{i,j})\sigma(w_{i,j}) = \varepsilon^{et_{i,j}}\sigma(h_{i,j})$. Since $r_a(0,\mathbf{0}) = 0$ for all $a \in [N] \setminus T$, the constant term of $\sigma(h_{i,j})$ is $h_{i,j}(0,b)$, which is nonzero. Thus the right side is an ε-unit. Thus $\sigma(q_{i,j})$ is an ε-unit by Lemma 4.3.

It remains to prove (20). Apply σ to (21). The gates that do not survive vanish, since $T_i \in I$ for $i \notin K$. Moreover, $\sigma(f) = f(b + (\mathbf{u},\mathbf{r}))$, since f does not involve ε. Dividing by ε^{ed} in the ring $\mathbb{F}[[\varepsilon,\mathbf{u}]]_\varepsilon$, we obtain

$$\sum_{i\in K} \varepsilon^{e_i} \prod_j p_{i,j}(\mathbf{u},\boldsymbol{\varphi}) = f(b + (\mathbf{u},\mathbf{r})) - \varepsilon^e \sigma(S),$$

where $\sigma(S) \in \mathbb{F}[[\varepsilon,\mathbf{u}]]$. Finally, since $r_a(\varepsilon,\mathbf{0}) \in \varepsilon\mathbb{F}[[\varepsilon]]$, we have $r_a = r_a(\varepsilon,\mathbf{0}) + \varphi_a \equiv \varphi_a \pmod{\varepsilon}$. Hence $f(b + (\mathbf{u},\mathbf{r})) \equiv f(b + (\mathbf{u},\boldsymbol{\varphi})) \pmod{\varepsilon}$. Since $e \ge 1$, the identity above becomes (20) after renumbering the surviving gates as $1,\dots,k'$, where $k' := |K| \le k$. □

We now give the proof of Lemma 6.11.

Proof of Lemma 6.11: If $T = [N]$, then the dimension assumption forces $I\,\mathbb{F}((\varepsilon))[\mathbf{x}] = \{0\}$, and hence $I = \{0\}$. In this case, the lemma holds trivially with $e = 1$. We may therefore assume that $T \neq [N]$.

We would like to apply the Abhyankar–Jung Theorem (Theorem 6.3) to the polynomials $\widetilde{Q}_a$. However, these polynomials are not necessarily monic. Thus, we need to first replace each $\widetilde{Q}_a$ by

a monic polynomial W_a. We make this replacement using Weierstrass preparation (Theorem 6.2). Moreover, $W_a(0, \mathbf{0}, x_a)$ will be a power of x_a, which will ensure that all its roots vanish at the origin and allow us to apply Weierstrass division with remainder degree less than $\deg_{x_a} W_a$.

First, we may assume that $b = \mathbf{0}$. Indeed, the translation $\mathbf{x} \mapsto \mathbf{x} + b$ maps $\mathbb{F}[[\varepsilon]][\mathbf{u}, x_a]$ onto itself and commutes with ∂_{x_a}. It turns the hypotheses for b into the hypotheses for $\mathbf{0}$, and composing it with a homomorphism σ for $\mathbf{0}$ gives one for b.

We will work over the power-series ring $\mathbb{F}[[\varepsilon, \mathbf{u}]]$ in ε and the free coordinates $\mathbf{u} = \mathbf{x}_T$. It is a UFD of dimension $|T| + 1$ (see Power Series). Moreover, ε is a prime element because the quotient $\mathbb{F}[[\varepsilon, \mathbf{u}]] / \langle \varepsilon \rangle$ is isomorphic to the domain $\mathbb{F}[[\mathbf{u}]]$.

Fix $a \in [N] \setminus T$. By Item 1, $\widetilde{Q}_a(0, \mathbf{0}, x_a)$ is a nonzero polynomial, and it vanishes at $x_a = 0$ because $\widetilde{Q}_a \in I$ and $g(0, \mathbf{0}) = 0$ for every $g \in I$. Thus Weierstrass preparation (Theorem 6.2) in $\mathbb{F}[[\varepsilon, \mathbf{u}, x_a]]$ gives $\widetilde{Q}_a = U_a W_a$, where U_a is a unit of $\mathbb{F}[[\varepsilon, \mathbf{u}, x_a]]$ and $W_a \in \mathbb{F}[[\varepsilon, \mathbf{u}]][x_a]$ is monic of some degree $d_a \geq 1$ with $W_a(0, \mathbf{0}, x_a) = x_a^{d_a}$. In fact, $U_a \in \mathbb{F}[[\varepsilon, \mathbf{u}]][x_a]$. To see this, note that polynomial long division of $\widetilde{Q}_a$ by the monic polynomial W_a in $\mathbb{F}[[\varepsilon, \mathbf{u}]][x_a]$ gives a quotient and a remainder, with the remainder having degree less than d_a in x_a. By the uniqueness assertion in the first item of Theorem 6.2, the quotient must be U_a and the remainder must be 0. Moreover, since $\widetilde{Q}_a \in I$ and $U_a^{-1} \in \mathbb{F}[[\varepsilon, \mathbf{u}, x_a]] \subseteq \mathbb{F}[[\varepsilon, \mathbf{x}]]$, we have $W_a = U_a^{-1} \widetilde{Q}_a \in I\, \mathbb{F}[[\varepsilon, \mathbf{x}]]$.

Next, we show that the discriminant of W_a is an ε-unit. Differentiating $\widetilde{Q}_a = U_a W_a$ gives

$$\partial_{x_a} \widetilde{Q}_a = U_a\, \partial_{x_a} W_a + W_a\, \partial_{x_a} U_a \equiv U_a\, \partial_{x_a} W_a \pmod{W_a}.$$

By Lemma 4.12 and the hypothesis on $\operatorname{Res}_{x_a}(\widetilde{Q}_a, \partial_{x_a} \widetilde{Q}_a)$, we have

$$\begin{aligned} \varepsilon^{j_a} R_a &= \operatorname{Res}_{x_a}(\widetilde{Q}_a, \partial_{x_a} \widetilde{Q}_a) \\ &= \operatorname{Res}_{x_a}(U_a, \partial_{x_a} \widetilde{Q}_a) \operatorname{Res}_{x_a}(W_a, \partial_{x_a} \widetilde{Q}_a) \\ &= \operatorname{Res}_{x_a}(U_a, \partial_{x_a} \widetilde{Q}_a) \operatorname{Res}_{x_a}(W_a, U_a) \operatorname{Res}_{x_a}(W_a, \partial_{x_a} W_a). \end{aligned} \tag{25}$$

All factors lie in $\mathbb{F}[[\varepsilon, \mathbf{u}]]$, since $\widetilde{Q}_a, U_a, W_a \in \mathbb{F}[[\varepsilon, \mathbf{u}]][x_a]$. The left side is an ε-unit because $R_a(0, \mathbf{0}) \neq 0$. By Lemma 4.3, each factor on the right is therefore an ε-unit. In particular, since W_a is monic of degree d_a, we have $\operatorname{Disc}_{x_a}(W_a) = (-1)^{d_a(d_a-1)/2} \operatorname{Res}_{x_a}(W_a, \partial_{x_a} W_a)$. Thus $\operatorname{Disc}_{x_a}(W_a)$ is an ε-unit.

Thus every W_a satisfies the hypotheses of Theorem 6.3. Apply it to each W_a, and let e_a be the resulting exponent. Let $e \geq 1$ be a common multiple of the resulting exponents e_a. Since $W_a(\varepsilon^{e_a}, \mathbf{u}, x_a)$ splits over $\mathbb{F}[[\varepsilon, \mathbf{u}]]$, the polynomial $W_a(\varepsilon^e, \mathbf{u}, x_a)$ also splits over this ring. For every $a \in [N] \setminus T$, we therefore have

$$W_a(\varepsilon^e, \mathbf{u}, x_a) = \prod_{i=1}^{d_a} (x_a - r_{a,i}) \qquad \text{with } r_{a,i} \in \mathbb{F}[[\varepsilon, \mathbf{u}]].$$

Since $W_a(0, \mathbf{0}, x_a) = x_a^{d_a}$, setting $\varepsilon = 0$ and $\mathbf{u} = \mathbf{0}$ gives $x_a^{d_a} = \prod_i (x_a - r_{a,i}(0, \mathbf{0}))$. Thus every root satisfies $r_{a,i}(0, \mathbf{0}) = 0$. Let $\iota \colon \mathbb{F}[[\varepsilon, \mathbf{u}]] \to \mathbb{F}[[\varepsilon, \mathbf{u}]]$ be the injective ring homomorphism

$g(\varepsilon, \mathbf{u}) \mapsto g(\varepsilon^e, \mathbf{u})$. The identity above says that W_a, with ι applied to its coefficients, splits over $\mathbb{F}[[\varepsilon, \mathbf{u}]]$. It remains to choose one root r_a among $r_{a,1}, \ldots, r_{a,d_a}$ for each a such that the substitution $x_a \mapsto r_a$ kills I.

To make this choice, define $\mathcal{O} := \mathbb{F}[[\varepsilon]][\mathbf{x}]_{\langle \varepsilon, \mathbf{x} \rangle}$, $B := \mathcal{O}/I\mathcal{O}$, and $D := \mathbb{F}[[\varepsilon, \mathbf{x}]]/I\mathbb{F}[[\varepsilon, \mathbf{x}]]$. We first show that $\dim D = |T| + 1$. Since I is prime and $\varepsilon \notin I$, the ideal $\mathfrak{p} := I\mathbb{F}((\varepsilon))[\mathbf{x}]$ is prime and satisfies $I = \mathfrak{p} \cap \mathbb{F}[[\varepsilon]][\mathbf{x}]$. Let $J := \{ g(0, \mathbf{x}) \mid g \in I \} \subseteq \mathbb{F}[\mathbf{x}]$. By Lemma 4.8, $\mathbf{V}(J)$ contains the origin and is equidimensional of dimension $|T|$. Moreover,

$$B/\varepsilon B \cong \mathbb{F}[\mathbf{x}]_{\langle \mathbf{x} \rangle}/J\mathbb{F}[\mathbf{x}]_{\langle \mathbf{x} \rangle}.$$

Every irreducible component of $\mathbf{V}(J)$ through the origin has dimension $|T|$, so Lemma 4.6 gives $\dim(B/\varepsilon B) = |T|$.

The ring B is a Noetherian local domain, and ε is a nonzero element of its maximal ideal. By Lemma 4.7, $\dim B = \dim(B/\varepsilon B) + 1 = |T| + 1$ and $\widehat{\mathcal{O}} \cong \mathbb{F}[[\varepsilon, \mathbf{x}]]$. Thus, Proposition 4.5 identifies D with $\widehat{B}$ and gives $\dim D = |T| + 1$.

We next show that D is a finite module over $\mathbb{F}[[\varepsilon, \mathbf{u}]]$. In $\mathbb{F}[[\varepsilon, \mathbf{x}]]$, viewed as a power-series ring in x_a over the other variables, W_a is still monic of degree d_a with $W_a(0, \mathbf{0}, x_a) = x_a^{d_a}$. Thus Weierstrass division (Theorem 6.2) lets us divide by W_a, with a remainder of degree less than d_a in x_a. We divide by the W_a one after another, each time dividing the coefficients of the previous remainder. This process shows that every element of $\mathbb{F}[[\varepsilon, \mathbf{x}]]$ is congruent modulo the ideal generated by the W_a to an $\mathbb{F}[[\varepsilon, \mathbf{u}]]$-linear combination of the monomials $\prod_{a \in [N] \setminus T} x_a^{\gamma_a}$ with $0 \le \gamma_a < d_a$. Since each W_a belongs to $I\mathbb{F}[[\varepsilon, \mathbf{x}]]$, the images of these finitely many monomials generate D as an $\mathbb{F}[[\varepsilon, \mathbf{u}]]$-module.

Next, we show that the natural map $\mathbb{F}[[\varepsilon, \mathbf{u}]] \to D$ is injective. Denote its kernel by $\mathfrak{a}$. Every element of I vanishes at the origin, so $I\mathbb{F}[[\varepsilon, \mathbf{x}]]$ is contained in the maximal ideal of $\mathbb{F}[[\varepsilon, \mathbf{x}]]$. Therefore, $D \neq 0$, so $1 \neq 0$ in D. Since the map $\mathbb{F}[[\varepsilon, \mathbf{u}]] \to D$ sends 1 to 1, its kernel $\mathfrak{a}$ is a proper ideal. We know that D is finite and hence integral over $\mathbb{F}[[\varepsilon, \mathbf{u}]]$ (see Finiteness and Integrality). If $\mathfrak{a}$ were nonzero, then D would be integral over $\mathbb{F}[[\varepsilon, \mathbf{u}]]/\mathfrak{a}$, so $\dim D = \dim \mathbb{F}[[\varepsilon, \mathbf{u}]]/\mathfrak{a} < \dim \mathbb{F}[[\varepsilon, \mathbf{u}]] = |T| + 1$ (Section 2.1), a contradiction. This argument proves that $\mathbb{F}[[\varepsilon, \mathbf{u}]] \to D$ is injective.

The ring D need not be a domain, so we pass to a quotient of D modulo a suitable prime ideal. Since D is integral over $\mathbb{F}[[\varepsilon, \mathbf{u}]]$, lying-over (Proposition 2.2) gives a prime $\mathfrak{q}$ of D with $\mathfrak{q} \cap \mathbb{F}[[\varepsilon, \mathbf{u}]] = \{ 0 \}$. The domain $D/\mathfrak{q}$ contains $\mathbb{F}[[\varepsilon, \mathbf{u}]]$ and is finite as a module over it, so its fraction field $\mathrm{Frac}(D/\mathfrak{q})$ is a finite extension of $K := \mathrm{Frac}(\mathbb{F}[[\varepsilon, \mathbf{u}]])$. Denote the algebraic closure of K by $\overline{K}$.

Recall that $\iota \colon \mathbb{F}[[\varepsilon, \mathbf{u}]] \to \mathbb{F}[[\varepsilon, \mathbf{u}]]$ is the injective ring homomorphism $g(\varepsilon, \mathbf{u}) \mapsto g(\varepsilon^e, \mathbf{u})$. It induces a field embedding $K \hookrightarrow \overline{K}$, which extends to $\iota' \colon \mathrm{Frac}(D/\mathfrak{q}) \hookrightarrow \overline{K}$, since $\overline{K}$ is algebraically closed [Lan02, Chapter V, Theorem 2.8]. Fix such an extension.

Composing $\mathbb{F}[[\varepsilon, \mathbf{x}]] \to D \to D/\mathfrak{q}$ with ι' gives a ring homomorphism $\sigma' \colon \mathbb{F}[[\varepsilon, \mathbf{x}]] \to \overline{K}$, which vanishes on I since $D = \mathbb{F}[[\varepsilon, \mathbf{x}]]/I\mathbb{F}[[\varepsilon, \mathbf{x}]]$. Moreover, σ' agrees with ι on $\mathbb{F}[[\varepsilon, \mathbf{u}]]$. For every

$a \in [N] \setminus T$, applying σ' to $W_a \in I\mathbb{F}[[\varepsilon, \mathbf{x}]]$ gives

$$0 = \sigma'(W_a) = W_a\big(\varepsilon^e, \mathbf{u}, \sigma'(x_a)\big) = \prod_{i=1}^{d_a} \big(\sigma'(x_a) - r_{a,i}\big),$$

so $\sigma'(x_a)$ is one of the roots $r_{a,i}$; we call it r_a.

Restrict σ' to $\mathbb{F}[[\varepsilon]][\mathbf{x}]$ and denote this restriction by σ. Since σ' sends every coefficient $c(\varepsilon) \in \mathbb{F}[[\varepsilon]]$ to $c(\varepsilon^e)$, fixes x_i for $i \in T$, and sends x_a to r_a for $a \notin T$, we have

$$\sigma(g) = g(\varepsilon^e, (\mathbf{u}, \mathbf{r})) \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$$

for every $g \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$, where $\mathbf{r} \coloneqq (r_a)_{a \in [N]\setminus T}$. Thus $\sigma(\mathbb{F}[[\varepsilon]][\mathbf{x}]) \subseteq \mathbb{F}[[\varepsilon, \mathbf{u}]]$, and $\sigma(I) = \{0\}$ because σ' vanishes on I.

Finally, fix $a \in [N] \setminus T$ and define

$$H_a \coloneqq \mathrm{Res}_{x_a}(W_a, \partial_{x_a} \widetilde{Q}_a) \in \mathbb{F}[[\varepsilon, \mathbf{u}]].$$

By (25) and Lemma 4.3, H_a is an ε-unit. Hence $H_a(\varepsilon^e, \mathbf{u})$ is also an ε-unit. Since $W_a(\varepsilon^e, \mathbf{u}, x_a)$ is monic with roots $r_{a,1}, \ldots, r_{a,d_a}$, the root formula for the resultant [Jan10, Theorem 1.10(i)] gives

$$H_a(\varepsilon^e, \mathbf{u}) = \prod_{i=1}^{d_a} (\partial_{x_a} \widetilde{Q}_a)(\varepsilon^e, \mathbf{u}, r_{a,i}).$$

Every factor belongs to $\mathbb{F}[[\varepsilon, \mathbf{u}]]$, so each is an ε-unit by Lemma 4.3. Since r_a is one of the $r_{a,i}$, the series $\sigma(\partial_{x_a} \widetilde{Q}_a)$ equals $(\partial_{x_a} \widetilde{Q}_a)(\varepsilon^e, \mathbf{u}, r_a)$ and is therefore an ε-unit, as desired. □

6.3 Applying Modified DiDIL

Throughout this subsection, let $f \in \overline{\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}}$ and write $f + \varepsilon S = \sum_{i=1}^{k} T_i$, where $S \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ and $T_i = \prod_j q_{i,j}$ for polynomials $q_{i,j} \in \mathbb{F}((\varepsilon))[\mathbf{x}]$ of degree at most δ. Theorem 6.7 transforms this circuit into a form to which DiDIL can be applied, as in Dutta and Saxena [DS26]. We first make this form explicit. From now on, we fix the form provided by Theorem 6.7: the set $T \subseteq [N]$ of free coordinates with $C \coloneqq N - |T| \le k$, the tuple $\mathbf{u} = \mathbf{x}_T$, the point b, and the power series φ_a with their equations Q_a of degree at most Δ. We also fix the integers e_i and the polynomials $p_{i,j}$ from Item 2; these polynomials have degree at most δ, and at most ρ of them are distinct. Let

$$T_i \coloneqq \varepsilon^{e_i} \prod_j p_{i,j}(\mathbf{u}, \boldsymbol{\varphi}) \quad (i \in [k']) \qquad \text{and} \qquad f_0 \coloneqq f(b + (\mathbf{u}, \boldsymbol{\varphi}(0, \mathbf{u}))) \in \mathbb{F}[[\mathbf{u}]]. \tag{26}$$

Thus T_i plays the role of the i-th product gate: up to the power ε^{e_i}, it is a product of polynomials of degree at most δ, restricted to the branch $x_a = \varphi_a$ for $a \notin T$. Then (20) is the same as

$$T_1 + T_2 + \cdots + T_{k'} \;=\; f_0 + \varepsilon S_0 \qquad \text{for some } S_0 \in \mathbb{F}[[\varepsilon, \mathbf{u}]], \tag{27}$$

where $k' \le k$, and every factor $p_{i,j}(\mathbf{u}, \boldsymbol{\varphi})$ of every T_i is an ε-unit, that is, a power of ε times an invertible power series. The bounds $\deg p_{i,j} \le \delta$ and $\deg Q_a \le \Delta$ are preserved by the polynomial

operations in our differentiation formulas (Lemma 6.16 and Remark 6.19). We keep the resulting substituted series and their inverses as individual factors and extract coefficients only at the end. This is why Theorem 6.12 lands in the class $\overline{\mathcal{C}}_\Delta(m,r)$, whose inner polynomials have degree at most Δ.

A Quick Comparison with [DS26]. This is the analogue of the starting point of the DiDIL in [DS26, Section 6]. There, f_0 is a polynomial of the form $P_{k+1,n}(\ell_1,\dots,\ell_N)$, where the dimension of the span of $\ell_1,\dots,\ell_N$ is "large". Each T_i is a product of linear polynomials over $\mathbb{F}[\varepsilon,\varepsilon^{-1}]$ with the *all-non-homogeneous* property: after T_i is divided by its power of ε, its constant term is a nonzero element of $\mathbb{F}$. In [DS26], this property is achieved by reducing modulo factors of the circuit whose normalized ε-limits are homogeneous linear forms [DS26, Lemma 15]. In our setting, this property follows from the border low-weight translation: after T_i is divided by its power of ε, it becomes an invertible power series in ε and $\mathbf{u}$, whose constant term is a nonzero element of $\mathbb{F}$. Two differences remain:

1. The factors $p_{i,j}(\mathbf{u},\boldsymbol{\varphi})$ are power series rather than linear polynomials, and we control them using the identities $Q_a(\mathbf{u},\varphi_a)=0$, where $\deg Q_a \le \Delta$.
2. The series f_0 need not be homogeneous, so we track its lowest nonzero homogeneous component, as specified in the theorem below.

Table 2 in Section 6.3.3 lists the full correspondence.

The main result of this subsection is the analogue of the Main Lemma of [DS26, Lemma 18]. Dutta and Saxena [DS26] express $P_{k+1,n}(\ell_1,\dots,\ell_N)$ as a ratio of two small border depth-3 diagonal circuits and conclude that this cannot happen unless the border circuit of this polynomial is exponentially large. Here the role of border depth-3 diagonal circuits is played by the border classes $\overline{\mathcal{C}}_\Delta(m,r)$ of Definition 6.4.

Theorem 6.12 (DiDIL Theorem for Bounded Depth-4): In the situation above, suppose that $f_0 \neq 0$, and let d_0 be the least d with $H_d(f_0)\neq 0$. Then there are integers $m,r\ge 1$ with

$$\log m = O\left(k4^k(k+\log\rho)\right) \qquad \text{and} \qquad r = O\left(k4^k(\Delta+1)^{k+1}\right),$$

and nonzero polynomials $G_1,G_2\in\overline{\mathcal{C}}_\Delta(m,r)$ such that $H_{d_0}(f_0)\cdot G_2 = G_1$.

Organization of this subsection. We follow the DiDIL argument of Dutta and Saxena [DS26, Section 6], with one important difference. In that work, the logarithmic derivative of a factor is expanded as a geometric series. We use the chain rule instead (Lemma 6.16) to express the logarithmic derivative through factors of the same kind, so we never have to expand anything during the elimination steps. After scaling the free coordinates by a new variable z, we *divide* by one of the summands and *derive* with respect to z, which removes that summand. We *induct*, removing the summands one at a time, and keep track of the coefficient of the lowest power of z in the *limit* $\varepsilon\to 0$ of the sum. When no summand can be removed anymore, comparing these coefficients on both sides expresses $H_{d_0}(f_0)$ as a ratio, as in Theorem 6.12. We set up the scaling in Section 6.3.1 and the limits in Section 6.3.2. Section 6.3.3 introduces the expression classes used in the DiDIL induction, and Section 6.3.4 proves their differentiation and coefficient-

extraction bounds. Section 6.3.5 carries out the DiDIL induction, and Section 6.3.6 finally deduces Theorem 6.1.

6.3.1 Scaling

As in [DS26], we do not apply DiDIL to (27) directly, but to the identity obtained from it by multiplying every free coordinate by a new variable z. We first collect the polynomials whose scaled versions we will divide by, namely those that are ε-units on the branch. Let

$$\mathcal{U} := \{\, p \in \mathbb{F}[[\varepsilon]][\mathbf{x}] \mid \deg p \le \Delta \text{ and } p(\mathbf{u}, \boldsymbol{\varphi}) \text{ is an } \varepsilon\text{-unit} \,\}.$$

By Theorem 6.7, every $p_{i,j}$ and every $\partial_a Q_a$ lies in $\mathcal{U}$, where $\partial_a := \partial_{x_a}$; so does every nonzero element of $\mathbb{F}[[\varepsilon]]$. Let $\mathcal{P} \subseteq \mathcal{U}$ be the set of distinct polynomials among the $p_{i,j}$, so that $|\mathcal{P}| \le \rho$.

Let z be a new variable. Define the $\mathbb{F}[[\varepsilon]]$-algebra homomorphism $\Psi\colon \mathbb{F}[[\varepsilon]][\mathbf{x}] \to \mathbb{F}[[\varepsilon]][\mathbf{u}][[z]]$ by

$$\Psi(x_i) := z \cdot x_i \quad (i \in T) \qquad \text{and} \qquad \Psi(x_a) := \varphi_a(\varepsilon, z \cdot \mathbf{u}) \quad (a \in [N] \setminus T).$$

Thus Ψ first substitutes $\varphi_a(\varepsilon, \mathbf{u})$ for each dependent coordinate x_a and then replaces $\mathbf{u}$ by $z\mathbf{u}$. In particular, $\Psi(Q_a) = 0$ for every $a \notin T$.

Since we will divide by powers of ε and z, let

$$\mathcal{S} := \mathbb{F}[[\varepsilon]][\mathbf{u}][[z]]_{\varepsilon z}.$$

Scaling by z extends to a homomorphism $\mathbb{F}[[\varepsilon, \mathbf{u}]]_\varepsilon \to \mathcal{S}$ by fixing ε^{-1}. Applying it to (27) gives

$$g_0 := T_1(z\mathbf{u}) + \cdots + T_{k'}(z\mathbf{u}) = f_0(z\mathbf{u}) + \varepsilon S_0(z\mathbf{u}), \tag{28}$$

where $T_i(z\mathbf{u}) = \varepsilon^{e_i} \prod_j \Psi(p_{i,j})$.

Lemma 6.13: Let $p \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$.

1. The coefficient of z^0 in $\Psi(p)$ is the constant term $p(\mathbf{0}) \in \mathbb{F}[[\varepsilon]]$ of p.
2. If $p \in \mathcal{U}$ and $p(\mathbf{u}, \boldsymbol{\varphi}) = \varepsilon^v w$ with $v \ge 0$, $w \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$, and $w(0, \mathbf{0}) \ne 0$, then $\Psi(p) = \varepsilon^v w(z\mathbf{u})$ is a unit of $\mathcal{S}$.

Proof: For the first item, the coefficient of z^0 in $g(z\mathbf{u})$ is $g(\varepsilon, \mathbf{0})$. For $g = p(\mathbf{u}, \boldsymbol{\varphi})$, this coefficient equals $p(\mathbf{0}, \boldsymbol{\varphi}(\varepsilon, \mathbf{0})) = p(\mathbf{0})$ because $\varphi_a(\varepsilon, \mathbf{0}) = 0$. For the second, the coefficient of z^0 in $w(z\mathbf{u})$ is $w(\varepsilon, \mathbf{0})$, which is a unit of $\mathbb{F}[[\varepsilon]]$, and hence is a unit of the coefficient ring $\mathbb{F}[[\varepsilon]][\mathbf{u}]$. A power series in z whose constant coefficient is a unit of the coefficient ring is a unit, so $w(z\mathbf{u})$ is a unit of $\mathbb{F}[[\varepsilon]][\mathbf{u}][[z]]$, and $\varepsilon^v w(z\mathbf{u})$ is a unit of $\mathcal{S}$. □

The second item is the *all-non-homogeneous* property after scaling. Every $\Psi(p)$ with $p \in \mathcal{U}$ is a unit of $\mathcal{S}$. After its maximal power of ε is removed, $\Psi(p)$ is a unit of the unlocalized ring $\mathbb{F}[[\varepsilon]][\mathbf{u}][[z]]$. The inverses $\Psi(p)^{-1}$ appear as soon as we divide.

6.3.2 Limits

The intermediate sums in DiDIL are ratios of elements of $\mathcal{S}$, so we work in the field

$$\mathbb{K} := L((\varepsilon)) \quad \text{with } L := \mathbb{F}(\mathbf{u})((z)).$$

The outer series variable is ε, and its coefficients lie in L. The ring $\mathcal{S}$ is a subring of $\mathbb{K}$. To see this, expand the coefficients in powers of ε:

$$\sum_d c_d z^d \mapsto \sum_{v\geq 0} \varepsilon^v \sum_d c_{d,v} z^d, \qquad c_{d,v} = [\varepsilon^v] c_d \in \mathbb{F}[\mathbf{u}].$$

This is an injective ring homomorphism $\mathbb{F}[[\varepsilon]][\mathbf{u}][[z]] \to L[[\varepsilon]]$, since in both rings the coefficient of $\varepsilon^v z^d$ in a product is the same finite sum, and it extends to $\mathcal{S}$ because ε and z are invertible in $\mathbb{K}$. In particular, every nonzero element of $\mathcal{S}$ is invertible in $\mathbb{K}$.

The following rules are used implicitly in [DS26]. We explicitly write them here for reference. We use the conventions of Limits and Normalization. For $0 \neq g \in \mathbb{K}$, put $a := \mathrm{val}_\varepsilon(g)$ and $b := \mathrm{val}_z(\lim \widetilde{g})$. DiDIL tracks the following coefficient:

$$\mathrm{TC}(\lim \widetilde{g}) = [z^b][\varepsilon^a] g.$$

Thus we first normalize in ε, then set $\varepsilon = 0$, and only then take the trailing coefficient in z.

Lemma 6.14: Let $g, h \in \mathbb{K}$ be nonzero, with normalizations $\widetilde{g}$ and $\widetilde{h}$.

1. The normalizations of $g \cdot h$ and g/h are $\widetilde{g} \cdot \widetilde{h}$ and $\widetilde{g}/\widetilde{h}$, respectively, and

$$\lim(\widetilde{g} \cdot \widetilde{h}) = (\lim \widetilde{g}) \cdot (\lim \widetilde{h}), \qquad \lim\left(\frac{\widetilde{g}}{\widetilde{h}}\right) = \frac{\lim \widetilde{g}}{\lim \widetilde{h}}.$$

   Consequently, their z-valuations and trailing coefficients obey the usual product and quotient rules.

2. If $g \in \mathcal{S}$ and $a := \mathrm{val}_\varepsilon(g)$, then $\lim \widetilde{g} \in \mathbb{F}[\mathbf{u}]((z))$. For every $d \in \mathbb{Z}$, we have $\varepsilon^{-a}[z^d]g \in \mathbb{F}[[\varepsilon]][\mathbf{u}]$ and

$$\left(\varepsilon^{-a}[z^d]g\right)\Big|_{\varepsilon=0} = [z^d](\lim \widetilde{g}).$$

   In particular, taking $d = \mathrm{val}_z(\lim \widetilde{g})$ gives $\mathrm{TC}(\lim \widetilde{g}) \in \mathbb{F}[\mathbf{u}]$.

3. If $p \in \mathcal{U}$ and $p(\mathbf{u}, \boldsymbol{\varphi}) = \varepsilon^v w$ as in Lemma 6.13, then $\widetilde{\Psi(p)} = w(\varepsilon, z\mathbf{u})$, and

$$\mathrm{val}_z(\lim \widetilde{\Psi(p)}) = 0, \qquad \mathrm{TC}(\lim \widetilde{\Psi(p)}) = w(0, \mathbf{0}) \in \mathbb{F}^\times.$$

Proof: The first item follows because $\widetilde{g}$ and $\widetilde{h}$ are units of $L[[\varepsilon]]$ and evaluation at $\varepsilon = 0$ is a ring homomorphism.

For the second item, write $g = \varepsilon^{-c} z^{-b} \sum_{j\geq 0} h_j z^j$, where $c, b \geq 0$ and $h_j \in \mathbb{F}[[\varepsilon]][\mathbf{u}]$. Every coefficient in its ε-expansion lies in $\mathbb{F}[\mathbf{u}]((z))$, so this holds for $\lim \widetilde{g} = [\varepsilon^a] g$. For every d, we have

$[z^d]g = \varepsilon^{-c} h_{d+b}$, taking $h_j = 0$ for $j < 0$; since $a = \mathrm{val}_\varepsilon(g)$, this contains no power of ε below a. Thus $\varepsilon^{-a}[z^d]g \in \mathbb{F}[[\varepsilon]][\mathbf{u}]$, and comparing the coefficient of $\varepsilon^a z^d$ gives the displayed identity.

For the third item, Lemma 6.13 gives $\Psi(p) = \varepsilon^v w(\varepsilon, z\mathbf{u})$. Its remaining factor has limit $w(0, z\mathbf{u})$ with nonzero constant term $w(0, \mathbf{0})$, which proves the assertions. □

Together with the first item, the third item says that the limit of the normalization of a product of series $\Psi(p)^{\pm 1}$ with $p \in \mathcal{U}$ has z-valuation 0 and a nonzero constant as its trailing coefficient. This is where the all-non-homogeneous property enters: In [DS26], the factors $\Pi\Sigma/\Pi\Sigma$ of the summands contribute only nonzero constants to the coefficient of the lowest power of z.

6.3.3 The Bloated Model

After one division and derivation, the summands in [DS26] are no longer products of linear polynomials but have the form $(U/V) \cdot (A/B)$, where $U, V \in \Pi\Sigma$ are products of scaled linear polynomials and $A, B \in \Sigma \wedge \Sigma$ are sums of powers of linear polynomials. Sums of at most k such summands form the bloated model $\Sigma^{[k]}(\Pi\Sigma/\Pi\Sigma)(\Sigma \wedge \Sigma/\Sigma \wedge \Sigma)$, denoted $\mathrm{Gen}(k, \cdot)$ in [DS26, Definition 2]. Our summands have the same form, with the following classes in place of $\Pi\Sigma$ and $\Sigma \wedge \Sigma$.

Definition 6.15 (General Bloated Model): Let $m \geq 1$ and $\ell \geq 0$ be integers.

1. For a set $\mathcal{P}$ of polynomials in $\mathbb{F}[[\varepsilon]][\mathbf{x}]$, the notation $\Pi\Psi(\mathcal{P})$ denotes the set of products of series $\Psi(p)$ with $p \in \mathcal{P}$, repetitions allowed; the empty product is 1.
2. $\Sigma\Pi\Psi(m, \ell)$ is the set of sums of at most m terms
$$c \cdot \Psi(q_1)^{\pm 1} \cdots \Psi(q_{\ell'})^{\pm 1} \qquad \text{with } \ell' \leq \ell,$$
where $c \in \mathbb{F}((\varepsilon))[z, z^{-1}]$, and each $q_t \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ has degree at most Δ and lies in $\mathcal{U}$ if its exponent is -1.
3. $\mathrm{Gen}(m, \ell)$ is the set of products $(U/V) \cdot (A/B)$ with $U, V \in \Pi\Psi(\mathcal{P})$ and nonzero $A, B \in \Sigma\Pi\Psi(m, \ell)$.

So $\Pi\Psi(\mathcal{P})$ plays the role of $\Pi\Sigma$ in [DS26], and $\Sigma\Pi\Psi(m, \ell)$ plays the role of $\Sigma \wedge \Sigma$. Note that we write Π where Dutta and Saxena [DS26] write $\wedge$: a power is a product of identical factors, whereas the products that arise in our setting may have different factors. The reason is the logarithmic derivative. In [DS26], a scaled linear factor has the form $\Phi(\ell) = c + z\tilde{\ell}$, where $c = \ell(\mathbf{0}) \neq 0$ and $\tilde{\ell} = \ell - c$, and the geometric series gives

$$\mathrm{dlog}_z \Phi(\ell) \;=\; \frac{\tilde{\ell}}{c + z\tilde{\ell}} \;=\; \sum_{i \geq 0} (-1)^i c^{-i-1} z^i \, \tilde{\ell}^{\,i+1},$$

which is a sum of powers of the single linear form $\tilde{\ell}$; this is where $\Sigma \wedge \Sigma$ comes from.

For our factors, however, the same expansion would produce powers of the power series $(\Psi(p) - p(\mathbf{0}))/z$, and it is unclear how to keep track of such expressions through the repeated elimination steps of DiDIL. Instead, the chain rule writes $\mathrm{dlog}_z \Psi(p)$, up to a factor z^{-1}, as a

sum of at most $k+1$ products, each involving at most four possibly different series $\Psi(q)^{\pm 1}$ with $\deg q \le \Delta$ (Lemma 6.16). So we keep every $\Psi(q)^{\pm 1}$ as a single factor and never expand it during the induction. We expand the series only when we extract coefficients: the coefficient of any power of z in such a product is a composition of an outer polynomial with a bounded number of inner polynomials of degree at most Δ (Corollary 6.18). These compositions take the place of the powers of linear forms in [DS26].

In [DS26], the class $\Sigma \wedge \Sigma$ plays two roles: it contains both the series that occur during the DiDIL induction and their coefficients. Here these two roles are played by two different classes: $\Sigma\Pi\Psi(m,\ell)$ for the series and $\mathcal{C}_\Delta(m,r)$ for their coefficients. Table 2 summarizes the correspondence.

	[DS26]	this paper
scaling	$\Phi\colon x_i \mapsto z x_i$	$\Psi\colon x_i \mapsto z x_i$, $x_a \mapsto \varphi_a(\varepsilon, z\mathbf{u})$
factors	$\Phi(\ell) = \ell(\mathbf{0}) + z\cdot(\cdots)$, ℓ linear	$\Psi(p) = p(\mathbf{0}) + z\cdot(\cdots)$, $p \in \mathcal{P}$
all-non-homogeneous	$\ell(\mathbf{0})\vert_{\varepsilon=0} \neq 0$	$p(\mathbf{u}, \boldsymbol{\varphi})$ is an ε-unit
DiDIL sums	$\sum_{i\le k} \frac{U_i}{V_i}\cdot\frac{A_i}{B_i}$, the class Gen	$\sum_{i\le k} \frac{U_i}{V_i}\cdot\frac{A_i}{B_i}$, each summand in $\mathrm{Gen}(\cdot,\cdot)$
with $U_i, V_i \in$	$\Pi\Sigma$	$\Pi\Psi(\mathcal{P})$
and $A_i, B_i \in$	$\Sigma \wedge \Sigma$	$\Sigma\Pi\Psi(\cdot,\cdot)$
Applying dlog_z	geometric series	chain rule (Lemma 6.16)
coefficients of A_i	$\Sigma \wedge \Sigma$ (interpolation)	$\mathcal{C}_\Delta(\cdot,\cdot)$ (Corollary 6.18)
lowest-degree polynomial	$P_{k+1,n}(\ell_1,\dots,\ell_N)$	$H_{d_0}(f_0)$
final ratio G_1/G_2	$G_1, G_2 \in \overline{\Sigma \wedge \Sigma}$	$G_1, G_2 \in \overline{\mathcal{C}}_\Delta(\cdot,\cdot)$ (Theorem 6.12)

Table 2: The DiDIL argument of [DS26, Section 6] and its counterpart here. In the scaling row, $i \in T$ and $a \in [N] \setminus T$. The summands have the form $(U/V)\cdot(A/B)$.

We keep track of both parameters of $\Sigma\Pi\Psi(m,\ell)$, since they behave differently. The number m of terms grows quickly during the induction, but the arity r in Theorem 6.12 does not depend on it; the number ℓ of factors per term stays below 4^k, and r depends only on ℓ, k, and Δ. The coefficients c may contain powers of z only because differentiation with respect to z produces a factor z^{-1} (Lemma 6.16). By Lemma 6.13, $\Sigma\Pi\Psi(m,\ell) \subseteq \mathcal{S}$, and every element of $\Pi\Psi(\mathcal{P})$ is a unit of $\mathcal{S}$. By Lemma 6.14, for every $U \in \Pi\Psi(\mathcal{P})$ with normalization $\widetilde{U}$, the limit $\lim \widetilde{U}$ has z-valuation 0, and $\mathrm{rc}(\lim \widetilde{U}) \in \mathbb{F} \setminus \{0\}$. Multiplying out shows that

$$\begin{aligned} \Sigma\Pi\Psi(m_1,\ell_1) + \Sigma\Pi\Psi(m_2,\ell_2) &\subseteq \Sigma\Pi\Psi(m_1+m_2, \max\{\ell_1,\ell_2\}), \\ \Sigma\Pi\Psi(m_1,\ell_1)\cdot\Sigma\Pi\Psi(m_2,\ell_2) &\subseteq \Sigma\Pi\Psi(m_1 m_2, \ell_1+\ell_2), \end{aligned}$$

and that the sets $\Sigma\Pi\Psi(m,\ell)$ increase with m and ℓ and are closed under multiplication by elements of $\mathbb{F}((\varepsilon))[z,z^{-1}]$. Closure under multiplication needs no separate argument here, whereas it does for $\Sigma \wedge \Sigma$ in [DS26]. As in [DS26], every summand arising in DiDIL will lie in $\mathrm{Gen}(m,\ell)$ for

suitable m and ℓ. Initially, by (28),

$$T_i(z\mathbf{u}) = \frac{U_{i,0}}{V_{i,0}} \cdot \frac{A_{i,0}}{B_{i,0}} \in \mathrm{Gen}(1,0) \qquad \text{with } U_{i,0} \coloneqq \prod_j \Psi(p_{i,j}),\ V_{i,0} \coloneqq 1,\ A_{i,0} \coloneqq \varepsilon^{e_i},\ B_{i,0} \coloneqq 1.$$

6.3.4 Derivatives and Coefficients

In [DS26], extracting a coefficient in z from a depth-3 diagonal circuit preserves this form, with a size bound obtained by interpolation. Moreover, each coefficient of the logarithmic derivative of a product of scaled linear polynomials is a sum of powers of linear forms, by the geometric-series expansion. Here, Corollary 6.18 bounds the composition complexity of coefficients. Lemma 6.16 controls differentiation and logarithmic differentiation, with the parameter growth stated there.

Differentiation. On $L = \mathbb{F}(\mathbf{u})((z))$, let ∂_z act term by term, treating $\mathbb{F}(\mathbf{u})$ as constants. Extend it to $\mathbb{K} = L((\varepsilon))$ by acting on each coefficient of the expansion in ε. This derivation preserves $\mathcal{S}$ and $L[[\varepsilon]]$, vanishes on $\mathbb{F}((\varepsilon))$, and commutes with lim on $L[[\varepsilon]]$. For $0 \neq h \in \mathbb{K}$, define $\mathrm{dlog}_z(h) \coloneqq (\partial_z h)/h$. The product rule gives

$$\mathrm{dlog}_z\left(\prod_{i=1}^{r} h_i^{e_i}\right) = \sum_{i=1}^{r} e_i \,\mathrm{dlog}_z(h_i) \qquad (h_i \neq 0,\ e_i \in \mathbb{Z}).$$

For $0 \neq h \in L$, put $b \coloneqq \mathrm{val}_z(h)$. If $b \neq 0$, characteristic zero gives

$$\mathrm{val}_z(\partial_z h) = b - 1, \qquad \mathrm{TC}(\partial_z h) = b \cdot \mathrm{TC}(h). \tag{29}$$

If $b = 0$, differentiation annihilates the trailing term; Remark 6.19 handles this case. For $p \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$, define the *Euler derivation in the free coordinates* by

$$Dp \coloneqq \sum_{i \in T} x_i \partial_{x_i} p.$$

It treats ε and the dependent coordinates as constants and does not increase degree in $\mathbf{x}$. The equations Q_a let us express $\partial_z \Psi(p)$ using these polynomial derivatives.

Lemma 6.16 (Derivatives): For every $p \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$,

$$z\,\partial_z \Psi(p) \;=\; \Psi(Dp) - \sum_{a \in [N]\setminus T} \Psi(\partial_a p) \cdot \Psi(DQ_a) \cdot \Psi(\partial_a Q_a)^{-1}. \tag{30}$$

Consequently:

1. $\partial_z\, \Sigma\Pi\Psi(m,\ell) \;\subseteq\; \Sigma\Pi\Psi((1+(k+1)\ell)m, \ell+4)$;
2. $\mathrm{dlog}_z(U/V) \;\in\; \Sigma\Pi\Psi((k+1)\rho, 4)$, for all $U, V \in \Pi\Psi(\mathcal{P})$.

Proof: Recall that $\Psi(p)$ is obtained from p by substituting zx_i for x_i ($i \in T$) and $\varphi_a(\varepsilon, z\mathbf{u})$ for x_a ($a \notin T$). By the chain rule,

$$\partial_z \Psi(p) = \sum_{i \in T} x_i\, \Psi(\partial_{x_i} p) + \sum_{a \in [N]\setminus T} \Psi(\partial_a p) \cdot \partial_z(\varphi_a(\varepsilon, z\mathbf{u})).$$

Multiplying the preceding identity by z turns its first sum into $\Psi(Dp)$ because $zx_i = \Psi(x_i)$ for $i \in T$. Since Q_a depends only on $\mathbf{u}$ and x_a, we have $\partial_b Q_a = 0$ for $b \notin T \cup \{a\}$. Applying the preceding identity to Q_a and using $\Psi(Q_a) = 0$, we get

$$0 = \Psi(DQ_a) + \Psi(\partial_a Q_a) \cdot z\partial_z(\varphi_a(\varepsilon, z\mathbf{u})).$$

Since $\partial_a Q_a \in \mathcal{U}$, the series $\Psi(\partial_a Q_a)$ is invertible, and (30) follows.

For the first item, we first differentiate a single factor. Let $\deg p \le \Delta$. Then Dp, $\partial_a p$, and DQ_a have degree at most Δ, and $\partial_a Q_a \in \mathcal{U}$. So by (30), $\partial_z \Psi(p)$ is a sum of at most $C + 1 \le k + 1$ terms, each of the form $\pm z^{-1}$ times a product of at most 3 series $\Psi(q)^{\pm 1}$ as in Definition 6.15. If moreover $p \in \mathcal{U}$, then $\partial_z(\Psi(p)^{-1}) = -\Psi(p)^{-1} \cdot \Psi(p)^{-1} \cdot \partial_z \Psi(p)$ is a sum of at most $k + 1$ such terms, each with at most 5 factors.

Now consider a term cY of an element of $\Sigma\Pi\Psi(m, \ell)$, where $c \in \mathbb{F}((\varepsilon))[z, z^{-1}]$ and Y is a product of $\ell' \le \ell$ factors $\Psi(q)^{\pm 1}$, and differentiate it by the product rule. Differentiating c gives the single term $(\partial_z c)Y$, where $\partial_z c \in \mathbb{F}((\varepsilon))[z, z^{-1}]$. Differentiating one of the ℓ' factors replaces it with at most $k + 1$ terms with at most 5 factors each, which gives at most $k + 1$ terms with at most $\ell' + 4$ factors. So $\partial_z(cY)$ is a sum of at most $1 + (k + 1)\ell$ terms with at most $\ell + 4$ factors each, and summing over the at most m terms gives the first item.

For the second item, note that dlog_z turns products into sums and powers into integer multiples. Collect the multiplicities of the factors to write

$$\operatorname{dlog}_z(U/V) = \sum_{p \in \mathcal{P}} e_p \operatorname{dlog}_z \Psi(p), \qquad e_p \in \mathbb{Z}.$$

Since $\mathcal{P} \subseteq \mathcal{U}$, each of these series is a sum of at most $k + 1$ terms with at most 4 factors by (30), and since $|\mathcal{P}| \le \rho$, we get at most $(k + 1)\rho$ terms. □

Coefficient Extraction. We seek a fixed finite family of polynomials in $\mathbf{u}$ that generates every coefficient of a given factor, so that the composition arity is independent of the degree in z. For $q \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$, write

$$q = \sum_{\mathbf{a} \in \mathbb{N}^{[N] \setminus T}} q_{\mathbf{a}} \mathbf{x}^{\mathbf{a}}_{[N] \setminus T}, \qquad q_{\mathbf{a}} \in \mathbb{F}[[\varepsilon]][\mathbf{u}],$$

and define the finite family

$$\mathcal{H}(q) := \left\{ H_i(q_{\mathbf{a}}) \;\middle|\; i \ge 0, \mathbf{a} \in \mathbb{N}^{[N] \setminus T} \right\} \setminus \{0\}.$$

Note that $H_i(q_{\mathbf{a}})$ is homogeneous of degree i in the free coordinates $\mathbf{u}$, with ε in the coefficient ring. If $\deg q \le \Delta$, a nonzero member $H_i(q_{\mathbf{a}})$ must have $i + |\mathbf{a}| \le \Delta$, so $|\mathcal{H}(q)| \le (\Delta + 1)^{C+1} \le (\Delta + 1)^{k+1}$, and every member of $\mathcal{H}(q)$ has degree at most Δ. Since Q_a involves only one dependent coordinate, $|\mathcal{H}(Q_a)| \le (\Delta + 1)^2$.

Lemma 6.17 (Coefficients): Let $q \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ have degree at most Δ. For every $d \ge 0$, the

coefficient $[z^d]\Psi(q)$ belongs to the $\mathbb{F}((\varepsilon))$-algebra generated by

$$\mathcal{H}(q) \cup \bigcup_{a \notin T} \mathcal{H}(Q_a).$$

If $q \in \mathcal{U}$, the same holds for $[z^d]\Psi(q)^{-1}$.

Proof: Fix $a \notin T$, write $Q_a = \sum_j Q_{a,j} x_a^j$, and put $\Phi_a := \varphi_a(\varepsilon, z\mathbf{u})$. By Item 1 of Theorem 6.7, we have $[z^0]\Phi_a = 0$ and $H_0(Q_{a,0}) = 0$. Evaluating the ε-unit $(\partial_a Q_a)(\mathbf{u}, \varphi_a)$ at $\mathbf{u} = \mathbf{0}$ gives $\lambda_a := H_0(Q_{a,1}) = (\partial_a Q_a)(\mathbf{0}) \in \mathbb{F}[[\varepsilon]] \setminus \{0\}$. Comparing coefficients in $0 = \Psi(Q_a) = \sum_{i,j} H_i(Q_{a,j}) z^i \Phi_a^j$ gives, for $d \geq 1$,

$$\lambda_a [z^d]\Phi_a = - \sum_{(i,j) \notin \{(0,0),(0,1)\}} H_i(Q_{a,j})[z^{d-i}]\Phi_a^j.$$

Since $[z^0]\Phi_a = 0$, the right-hand side is a polynomial in $\mathcal{H}(Q_a)$ and coefficients $[z^{d'}]\Phi_a$ with $d' < d$. Division by λ_a is allowed in the field $\mathbb{F}((\varepsilon))$, so induction on d shows that every coefficient $[z^d]\Phi_a$ belongs to the $\mathbb{F}((\varepsilon))$-algebra generated by $\mathcal{H}(Q_a)$.

Now $\Psi(q) = \sum_{\mathbf{a} \in \mathbb{N}^{[N]\setminus T}} q_{\mathbf{a}}(z\mathbf{u}) \prod_{b \notin T} \Phi_b^{a_b}$, and $[z^i] q_{\mathbf{a}}(z\mathbf{u}) = H_i(q_{\mathbf{a}})$. The assertion for $\Psi(q)$ follows by taking coefficients.

If $q \in \mathcal{U}$, then $[z^0]\Psi(q) = q(\mathbf{0}) \neq 0$ by Lemma 6.13 and the ε-unit condition. Write $a_d = [z^d]\Psi(q)$ and $b_d = [z^d]\Psi(q)^{-1}$. Comparing coefficients in the identity $\Psi(q)\Psi(q)^{-1} = 1$ gives $b_0 = a_0^{-1}$ and $b_d = -a_0^{-1} \sum_{i=1}^d a_i b_{d-i}$ for $d \geq 1$. Since $a_0^{-1} \in \mathbb{F}((\varepsilon))$, induction on d shows that every b_d belongs to the $\mathbb{F}((\varepsilon))$-algebra generated by $\mathcal{H}(q) \cup \bigcup_{a \notin T} \mathcal{H}(Q_a)$, proving the assertion for $\Psi(q)^{-1}$. □

Corollary 6.18: Let $A \in \Sigma\Pi\Psi(m, \ell)$ and set $r := (\ell + k)(\Delta + 1)^{k+1}$.

1. For every $d \in \mathbb{Z}$, we have $[z^d]A \in \mathcal{C}_\Delta(m, r; \mathbb{F}((\varepsilon)))$.
2. If $A \neq 0$ and $\widetilde{A}$ is its normalization, then $\mathrm{TC}(\lim \widetilde{A}) \in \overline{\mathcal{C}}_\Delta(m, r)$.

Proof: For Item 1, consider one term $c \prod_{t=1}^{\ell'} \Psi(q_t)^{\sigma_t}$ of A, where $\ell' \leq \ell$, $\sigma_t \in \{-1, 1\}$, and $c \in \mathbb{F}((\varepsilon))[z, z^{-1}]$. Since c is a Laurent polynomial in z and each factor is a power series in z, every coefficient of this term is a finite $\mathbb{F}((\varepsilon))$-linear combination of products of coefficients of the factors. By Lemma 6.17, all these coefficients lie in the $\mathbb{F}((\varepsilon))$-algebra generated by

$$\bigcup_{t=1}^{\ell'} \mathcal{H}(q_t) \cup \bigcup_{a \notin T} \mathcal{H}(Q_a).$$

This set consists of at most $\ell(\Delta + 1)^{k+1} + k(\Delta + 1)^2 \leq r$ polynomials, each of degree at most Δ. So every coefficient of the term is a single composition of arity at most r whose inner polynomials have degree at most Δ, and summing over the at most m terms of A proves Item 1.

For Item 2, let $a := \mathrm{val}_\varepsilon(A)$, $d' := \mathrm{val}_z(\lim \widetilde{A})$, and $G_\varepsilon := \varepsilon^{-a}[z^{d'}]A$. By Lemma 6.14, $G_\varepsilon \in \mathbb{F}[[\varepsilon]][\mathbf{u}]$ and $G_\varepsilon(0, \mathbf{u}) = \mathrm{TC}(\lim \widetilde{A})$. By Item 1 and closure of $\mathcal{C}_\Delta(m, r; \mathbb{F}((\varepsilon)))$ under multiplication by scalars, we have $G_\varepsilon \in \mathcal{C}_\Delta(m, r; \mathbb{F}((\varepsilon)))$. So $\mathrm{TC}(\lim \widetilde{A}) \in \overline{\mathcal{C}}_\Delta(m, r)$, by the definition of the border class. □

6.3.5 The DiDIL Induction

We now prove Theorem 6.12. Starting from the sum g_0 in (28), we remove the summands one at a time. Dividing the sum by the normalization of one of its summands turns that summand into a power of ε, and differentiating with respect to z then removes it. Repeating this gives sums $g_0, g_1, \ldots$ with fewer and fewer summands, so there are at most $k-1$ steps. We prove by induction on j that g_j satisfies Properties (I)–(III) below.

The quantity we track is the trailing coefficient $\operatorname{TC}(\lim g_j)$: it equals $H_{d_0}(f_0)$ for $j = 0$, and Property (III) records how it changes.

When no summand can be removed anymore, we compute this coefficient directly from the remaining summands, and comparing with Property (III) gives the theorem. Remark 6.19, after the proof, treats the degenerate case in which differentiating with respect to z would annihilate this coefficient.

Proof of Theorem 6.12: Let $m_0 := 1$ and $\ell_0 := 0$, and for $j \geq 0$ let

$$m_{j+1} := (k+1)(\rho + 4\ell_j + 2)\, m_j^4 \qquad \text{and} \qquad \ell_{j+1} := 4\ell_j + 4. \tag{31}$$

Both sequences (m_j) and (ℓ_j) are increasing, so the classes $\Sigma\Pi\Psi(m_j, \ell_j)$ and $\operatorname{Gen}(m_j, \ell_j)$ increase with j. As in [DS26, Section 6], we construct nonzero $g_0, g_1, \ldots, g_J \in \mathbb{K}$, where g_0 is the sum in (28), with the following properties for every $j \leq J$.

(I) $g_j = T_{1,j} + \cdots + T_{k_j,j}$ for some $1 \leq k_j \leq k' - j$, where $k' \leq k$ is as in (28). Moreover, $g_j \in L[[\varepsilon]]$ and $\lim g_j \neq 0$; we write $b_j := \operatorname{val}_z(\lim g_j)$.

(II) $T_{i,j} \in \operatorname{Gen}(m_j, \ell_j)$; we fix a representation $T_{i,j} = (U_{i,j}/V_{i,j}) \cdot (A_{i,j}/B_{i,j})$ as in Definition 6.15.

(III) $\operatorname{TC}(\lim g_j) = H_{d_0}(f_0) \cdot W_{j,1}/W_{j,2}$, where $W_{j,1}$ and $W_{j,2}$ are products of at most j polynomials of the form $\operatorname{TC}(\lim \widetilde{A})$, with $\widetilde{A}$ the normalization of a nonzero $A \in \Sigma\Pi\Psi(m_j, \ell_j)$.

These are Properties (I), (II), and (IV) of [DS26, Section 6]; their Property (III), that U and V are invertible with nonzero constant terms, holds automatically here, as noted after Definition 6.15. As in [DS26], we write each summand as

$$T_{i,j} = \varepsilon^{a_{i,j}}\, \widetilde{T}_{i,j}, \qquad t_{i,j} := \lim \widetilde{T}_{i,j} \in L, \qquad d'_{i,j} := \operatorname{val}_z(t_{i,j}),$$

where $\widetilde{T}_{i,j}$ is the normalization of $T_{i,j}$.

The series g_0 lies in $\mathbb{F}[[\varepsilon]][\mathbf{u}][[z]] \subseteq L[[\varepsilon]]$ and has limit $f_0(z\mathbf{u})$, so Power Series gives

$$b_0 = d_0 \qquad \text{and} \qquad \operatorname{TC}(\lim g_0) = H_{d_0}(f_0). \tag{32}$$

In particular, $g_0 \neq 0$, so $k' \geq 1$. Thus Property (I) holds for $j = 0$ with $k_0 := k'$ and $T_{i,0} := T_i(z\mathbf{u})$. Property (II) holds with $U_{i,0}$, $V_{i,0}$, $A_{i,0}$, and $B_{i,0}$ as defined after Definition 6.15. Property (III) holds with $W_{0,1} = W_{0,2} = 1$.

Divide. Suppose that g_j has been constructed. If $a_{i,j} = 0$ and $d'_{i,j} = b_j$ for every i, we stop, set $J := j$, and proceed to the *Taking the limit* step below. Otherwise, we choose one summand to remove. If $d'_{i,j} \neq b_j$ for some i, we choose such a summand; this is the main case, treated in the

Derive step below. If $d'_{i,j} = b_j$ for every i, we choose a summand with $a_{i,j} \neq 0$, which exists since we did not stop. This is the degenerate case, treated in Remark 6.19, which also gives an example.

We renumber the summands so that the chosen one is $T_{k_j,j}$ and write $\widetilde{T} := \widetilde{T}_{k_j,j}$, $t := t_{k_j,j}$, $a := a_{k_j,j}$, and $d' := d'_{k_j,j}$. Dividing the identity $g_j = T_{1,j} + \cdots + T_{k_j,j}$ by $\widetilde{T}$ turns the last summand into ε^a:

$$\frac{g_j}{\widetilde{T}} \;=\; \frac{T_{1,j}}{\widetilde{T}} + \cdots + \frac{T_{k_j-1,j}}{\widetilde{T}} + \varepsilon^a \,. \tag{33}$$

By Property (I), g_j lies in $L[[\varepsilon]]$ and has nonzero limit, so it is its own normalization. Hence, by Lemma 6.14, the left-hand side of (33) lies in $L[[\varepsilon]]$, and its limit is $\lim g_j/t$, whose z-valuation is $b_j - d'$.

Derive. Suppose that we are in the main case, that is, $d' \neq b_j$. We apply ∂_z to (33). Since $\partial_z(\varepsilon^a) = 0$, the last summand disappears:

$$g_{j+1} \;:=\; \partial_z\!\left(\frac{g_j}{\widetilde{T}}\right) \;=\; \sum_{i=1}^{k_j-1} \partial_z\!\left(\frac{T_{i,j}}{\widetilde{T}}\right) = \sum_{i=1}^{k_j-1} \frac{T_{i,j}}{\widetilde{T}} \cdot \mathrm{dlog}_z\left(\frac{T_{i,j}}{\widetilde{T}}\right) . \tag{34}$$

Since ∂_z maps $L[[\varepsilon]]$ into itself and commutes with lim, we have $g_{j+1} \in L[[\varepsilon]]$ and, as in [DS26],

$$\lim g_{j+1} = \partial_z\!\left(\frac{\lim g_j}{t}\right).$$

By (29), applied to $\lim g_j/t$, whose z-valuation $b_j - d'$ is nonzero, the limit $\lim g_{j+1}$ is nonzero (so $g_{j+1} \neq 0$), $b_{j+1} = b_j - d' - 1$, and

$$\mathrm{TC}(\lim g_{j+1}) = (b_j - d') \cdot \frac{\mathrm{TC}(\lim g_j)}{\mathrm{TC}(t)}. \tag{35}$$

So $\mathrm{TC}(\lim g_j)$ survives this step, up to the nonzero factor $(b_j - d')/\mathrm{TC}(t)$. This is why the Divide step prefers a summand with $d' \neq b_j$: if $d' = b_j$, then ∂_z would annihilate the trailing term of $\lim g_j/t$.

Induct. We verify Properties (I)–(III) for $j+1$. For Property (III), write $T_{k_j,j} = (U/V) \cdot (A/B)$ as in Property (II), and let $\widetilde{U}$, $\widetilde{V}$, $\widetilde{A}$, and $\widetilde{B}$ be the normalizations of U, V, A, and B. By Lemma 6.14, $t = (\lim \widetilde{U} \cdot \lim \widetilde{A})/(\lim \widetilde{V} \cdot \lim \widetilde{B})$, so

$$\mathrm{TC}(t) = c \cdot \frac{\mathrm{TC}(\lim \widetilde{A})}{\mathrm{TC}(\lim \widetilde{B})}, \qquad \text{where } c := \frac{\mathrm{TC}(\lim \widetilde{U})}{\mathrm{TC}(\lim \widetilde{V})} \in \mathbb{F} \setminus \{\,0\,\}\,.$$

Thus (35) and Property (III) for j give

$$\mathrm{TC}(\lim g_{j+1}) = H_{d_0}(f_0) \cdot \frac{W_{j,1} \cdot (b_j - d')\ \mathrm{TC}(\lim \widetilde{B})}{W_{j,2} \cdot c\ \mathrm{TC}(\lim \widetilde{A})}.$$

Here $(b_j - d')\ \mathrm{TC}(\lim \widetilde{B})$ and $c\ \mathrm{TC}(\lim \widetilde{A})$ are the trailing coefficients of the limits of the normal-

izations of $(b_j - d')B$ and cA, which are nonzero elements of $\Sigma\Pi\Psi(m_j, \ell_j) \subseteq \Sigma\Pi\Psi(m_{j+1}, \ell_{j+1})$. So $W_{j+1,1} \coloneqq W_{j,1} \cdot (b_j - d')\ \mathrm{TC}(\lim \widetilde{B})$ and $W_{j+1,2} \coloneqq W_{j,2} \cdot c\ \mathrm{TC}(\lim \widetilde{A})$ satisfy Property (III) for $j+1$.

For Property (II), fix $i < k_j$, and let $U_{i,j}$, $V_{i,j}$, $A_{i,j}$, and $B_{i,j}$ be as in Property (II). Since $\widetilde{T} = \varepsilon^{-a}(U/V) \cdot (A/B)$,

$$\frac{T_{i,j}}{\widetilde{T}} = \frac{U'}{V'} \cdot \frac{A'}{B'} \qquad \text{with } U' \coloneqq U_{i,j}V,\ V' \coloneqq V_{i,j}U,\ A' \coloneqq \varepsilon^a A_{i,j}B,\ B' \coloneqq B_{i,j}A,$$

where $U', V' \in \Pi\Psi(\mathcal{P})$, and A' and B' are nonzero elements of $\Sigma\Pi\Psi(m_j^2, 2\ell_j)$. Since dlog_z turns products into sums, the i-th summand of (34) is

$$\begin{aligned} &\frac{U'}{V'} \cdot \frac{A'}{B'} \cdot \left(\mathrm{dlog}_z \left(\frac{U'}{V'} \right) + \frac{\partial_z A'}{A'} - \frac{\partial_z B'}{B'} \right) \\ &= \frac{U'}{V'} \cdot \frac{\mathrm{dlog}_z(U'/V') \cdot A'B' + \partial_z(A') \cdot B' - A' \cdot \partial_z(B')}{B'^2}. \end{aligned}$$

By Lemma 6.16, we have

$$\mathrm{dlog}_z(U'/V') \in \Sigma\Pi\Psi((k+1)\rho, 4) \quad \text{and} \quad \partial_z(A'), \partial_z(B') \in \Sigma\Pi\Psi((1 + 2(k+1)\ell_j)m_j^2, 2\ell_j + 4).$$

So the numerator is a sum of at most $((k+1)\rho + 2 + 4(k+1)\ell_j)\, m_j^4$ terms with at most $4\ell_j + 4$ factors each. By (31), this number of terms is at most m_{j+1}, and $4\ell_j + 4 = \ell_{j+1}$, so the numerator lies in $\Sigma\Pi\Psi(m_{j+1}, \ell_{j+1})$. The denominator B'^2 is a nonzero element of $\Sigma\Pi\Psi(m_j^4, 4\ell_j) \subseteq \Sigma\Pi\Psi(m_{j+1}, \ell_{j+1})$. Thus every summand of (34) lies in $\mathrm{Gen}(m_{j+1}, \ell_{j+1})$ or is zero if its numerator vanishes.

For Property (I), we drop the zero summands and let $T_{1,j+1}, \ldots, T_{k_{j+1},j+1}$ be the remaining ones. Since $g_{j+1} \neq 0$, at least one remains, and $k_{j+1} \leq k_j - 1 \leq k' - (j+1)$. We have already shown that $g_{j+1} \in L[[\varepsilon]]$ and $\lim g_{j+1} \neq 0$.

Taking the limit. If $k_j = 1$, then $T_{1,j} = g_j$ lies in $L[[\varepsilon]]$ and has nonzero limit, so $a_{1,j} = 0$ and $d'_{1,j} = b_j$; hence the stopping condition holds. Since $1 \leq k_j \leq k' - j$, the DiDIL induction terminates at some $J \leq k' - 1 \leq k - 1$. At that point, every summand satisfies $a_{i,J} = 0$, so $T_{i,J} = \widetilde{T}_{i,J}$ lies in $L[[\varepsilon]]$. Since lim is additive, we have $\lim g_J = \sum_i t_{i,J}$. We compare the coefficients of z^{b_J} on both sides. Since $d'_{i,J} = b_J$ for every i, the coefficient of z^{b_J} in $t_{i,J}$ is $\mathrm{TC}(t_{i,J})$. Write $T_{i,J} = (U_i/V_i) \cdot (A_i/B_i)$ as in Property (II), and let $\widetilde{A}_i$ and $\widetilde{B}_i$ be the normalizations of A_i and B_i. As in the Induct step, $\mathrm{TC}(t_{i,J}) = c_i\ \mathrm{TC}(\lim \widetilde{A}_i) / \mathrm{TC}(\lim \widetilde{B}_i)$ for some $c_i \in \mathbb{F} \setminus \{0\}$, and so

$$\mathrm{TC}(\lim g_J) = \sum_{i=1}^{k_J} c_i \frac{\mathrm{TC}(\lim \widetilde{A}_i)}{\mathrm{TC}(\lim \widetilde{B}_i)}.$$

Combining this identity with Property (III) for J and multiplying by $W_{J,2} \prod_i \mathrm{TC}(\lim \widetilde{B}_i)$, we obtain

$H_{d_0}(f_0) \cdot G_2 = G_1$ for

$$G_2 := W_{J,1} \prod_{i=1}^{k_J} \mathrm{TC}(\lim \widetilde{B}_i) \qquad \text{and} \qquad G_1 := W_{J,2} \sum_{i=1}^{k_J} c_i \, \mathrm{TC}(\lim \widetilde{A}_i) \prod_{i' \neq i} \mathrm{TC}(\lim \widetilde{B}_{i'}).$$

By Lemma 6.14, the trailing coefficients $\mathrm{TC}(\lim \widetilde{A}_i)$ and $\mathrm{TC}(\lim \widetilde{B}_i)$ are nonzero polynomials in $\mathbf{u}$. So $G_2 \neq 0$ as a product of nonzero polynomials, and hence $G_1 = H_{d_0}(f_0) G_2 \neq 0$.

Size analysis. Every factor of G_2 and every factor of every summand of G_1 has the form $\mathrm{TC}(\lim \widetilde{A})$, where $\widetilde{A}$ is the normalization of some nonzero $A \in \Sigma\Pi\Psi(m_J, \ell_J)$. For the factor $c_i \, \mathrm{TC}(\lim \widetilde{A}_i)$, take $A = c_i A_i$. Since $J \leq k-1$, such an A lies in $\Sigma\Pi\Psi(m_{k-1}, \ell_{k-1})$, so every factor lies in $\overline{\mathcal{C}}_\Delta\big(m_{k-1}, (\ell_{k-1}+k)(\Delta+1)^{k+1}\big)$ by Corollary 6.18. The number of factors of G_2, and of each summand of G_1, is at most $J + k_J \leq k' \leq k$, and G_1 has $k_J \leq k$ summands. By Lemma 6.5,

$$G_1, G_2 \in \overline{\mathcal{C}}_\Delta(m, r) \qquad \text{with} \qquad m := k \cdot m_{k-1}^k \quad \text{and} \quad r := k(\ell_{k-1}+k)(\Delta+1)^{k+1}.$$

It remains to bound m and r. By induction, $\ell_j = (4^{j+1}-4)/3 < 4^{j+1}$, so $\ell_j < 4^k$ for $j \leq k-1$. Hence $m_{j+1} \leq (k+1)(\rho + 4^{k+1} + 2)\, m_j^4$, that is, $\log m_{j+1} \leq O(k + \log \rho) + 4 \log m_j$ with an absolute constant in $O(\cdot)$. Since $m_0 = 1$, this gives $\log m_j = O\big(4^j (k + \log \rho)\big)$ for $j \leq k-1$, and therefore $\log m = \log k + k \log m_{k-1} = O\big(k 4^k (k + \log \rho)\big)$. Finally, $r = k(\ell_{k-1}+k)(\Delta+1)^{k+1} = O\big(k 4^k (\Delta+1)^{k+1}\big)$. □

Remark 6.19 (The degenerate case): We now treat the degenerate case of the Divide step, in which $d'_{i,j} = b_j$ for every i and the chosen summand has $a \neq 0$. Then $d' = b_j$, so $\lim g_j/t$ has z-valuation 0, with trailing coefficient $\mathrm{TC}(\lim g_j)/\mathrm{TC}(t)$. Hence (29) does not apply: ∂_z would annihilate this trailing term, and we would lose track of $\mathrm{TC}(\lim g_j)$.

For example, with free coordinate u and no dependent coordinates, suppose that $g_j = T_{1,j} + T_{2,j}$ with

$$T_{1,j} = \varepsilon^{-1}(1 + \varepsilon(1+zu)), \qquad T_{2,j} = -\varepsilon^{-1}, \qquad \text{so that } g_j = 1 + zu.$$

The normalized summands $1 + \varepsilon(1+zu)$ and -1 have limits 1 and -1, so $d'_{1,j} = d'_{2,j} = 0 = b_j$, while $a_{1,j} = a_{2,j} = -1$. Dividing by the normalization $\widetilde{T} = -1$ of $T_{2,j}$, with $a = -1$, turns (33) into

$$-1 - zu = (-\varepsilon^{-1} - 1 - zu) + \varepsilon^{-1}.$$

Differentiating with respect to z removes the last summand ε^{-1}, but it also annihilates the constant term -1 on the left-hand side, leaving $-u$. So the trailing coefficient -1 is lost in this process!

To keep the trailing coefficient, we follow the remark at the end of [DS26, Section 4.3] and differentiate with respect to ε instead of z. Let ∂_ε be the derivative with respect to ε, acting term by term on the expansion in ε; thus $\varepsilon\partial_\varepsilon$ multiplies the coefficient of ε^v in an element of $\mathbb{K}$ by v.

This *Euler derivation* $\varepsilon\partial_\varepsilon$ of $\mathbb{K}$ is L-linear and maps $\mathcal{S}$ and $L[[\varepsilon]]$ into themselves. We apply to (33) the operator $h \mapsto \varepsilon\partial_\varepsilon h - ah$, which we denote by $\varepsilon\partial_\varepsilon - a$.

Note that $\varepsilon\partial_\varepsilon - a$ is L-linear, though not a derivation, so we may apply it term by term. Since

$(\varepsilon\partial_\varepsilon - a)(\varepsilon^a) = 0$, this again removes the last summand:

$$g_{j+1} \;:=\; (\varepsilon\partial_\varepsilon - a)\left(\frac{g_j}{\widetilde{T}}\right) \;=\; \sum_{i=1}^{k_j-1} \frac{T_{i,j}}{\widetilde{T}} \cdot \left(\varepsilon\,\mathrm{dlog}_\varepsilon\left(\frac{T_{i,j}}{\widetilde{T}}\right) - a\right),$$

where $\mathrm{dlog}_\varepsilon(h) := \partial_\varepsilon h/h$ is the logarithmic derivative with respect to ε, defined as for z.

In the example above, $a = -1$, and

$$(\varepsilon\partial_\varepsilon + 1)(\varepsilon^{-1}) = 0 \qquad \text{and} \qquad (\varepsilon\partial_\varepsilon + 1)(-\varepsilon^{-1} - 1 - zu) = -1 - zu.$$

So $g_{j+1} = -1 - zu$, with a single summand: the last summand has disappeared, and the trailing coefficient -1 is preserved, which is what we wanted.

In general, the Divide step gives $g_j/\widetilde{T} \in L[[\varepsilon]]$, and $\varepsilon\partial_\varepsilon$ maps $L[[\varepsilon]]$ into $\varepsilon L[[\varepsilon]]$. Consequently, $g_{j+1} \in L[[\varepsilon]]$ and $\lim g_{j+1} = -a \cdot \lim g_j/t$, which is nonzero since $a \neq 0$. So $b_{j+1} = b_j - d' = 0$, and (35) holds with $-a$ in place of $b_j - d'$. Thus the factor $b_j - d'$, which vanishes here, is replaced by $-a$, which does not; this is why the Divide step chooses a summand with $a \neq 0$. Properties (I) and (III) for $j+1$ follow as in the Induct step.

For Property (II), let $E := \varepsilon\partial_\varepsilon$ act on the coefficients of polynomials in $\mathbb{F}[[\varepsilon]][\mathbf{x}]$, treating $\mathbf{x}$ as constants. The chain rule gives

$$\varepsilon\partial_\varepsilon\Psi(p) = \Psi(Ep) - \sum_{c\in[N]\setminus T} \Psi(\partial_{x_c}p)\cdot\Psi(EQ_c)\cdot\Psi(\partial_{x_c}Q_c)^{-1}.$$

This is the analogue of (30), with $\varepsilon\partial_\varepsilon$ in place of $z\partial_z$ and E in place of D. Since E does not increase the degree in $\mathbf{x}$, the two closure bounds in Lemma 6.16 hold with $\varepsilon\partial_\varepsilon$ and $\varepsilon\,\mathrm{dlog}_\varepsilon$ in place of ∂_z and dlog_z, respectively. Compared with the Induct step, the only other change is the extra term $-aA'B'$ in the numerator of each summand. It adds at most m_j^4 terms with at most $4\ell_j$ factors, so the numerator has at most $((k+1)\rho + 3 + 4(k+1)\ell_j)\, m_j^4 \le m_{j+1}$ terms, by (31) and $k \ge 1$. Hence every summand again lies in $\mathrm{Gen}(m_{j+1}, \ell_{j+1})$ or is zero. ◁

6.3.6 Stitching It All Together: Proof of the Lower Bound

The next lemma shows that the restricted power series $f_0 = P_{k+1,n}(b + (\mathbf{u}, \boldsymbol{\varphi}(0,\mathbf{u})))$ is nonzero and that every monomial involves at least $n - C - |\mathrm{supp}(b)|$ variables when $n > C + |\mathrm{supp}(b)|$. We then apply Theorem 6.6 to the polynomial identity $H_{d_0}(f_0)G_2 = G_1$ from Theorem 6.12.

Lemma 6.20: Let $T \subseteq [N]$ with $C := N - |T| \le k$, let $\mathbf{u} := \mathbf{x}_T$, let $b \in \mathbb{F}^N$, and let $g = (g_a)_{a\in[N]\setminus T}$ with $g_a \in \mathbb{F}[[\mathbf{u}]]$. If $n > C + |\mathrm{supp}(b)|$, then the power series

$$f := P_{k+1,n}(b + (\mathbf{u}, g)) \in \mathbb{F}[[\mathbf{u}]]$$

is nonzero, and every monomial occurring in f involves at least $n - C - |\mathrm{supp}(b)|$ variables.

Proof: For $i \in [N]$, let y_i be the i-th entry of $b + (\mathbf{u}, g)$, that is, $y_i = b_i + x_i$ for $i \in T$ and $y_i = b_i + g_i$ for $i \notin T$. Let $E := ([N]\setminus T) \cup \mathrm{supp}(b)$, so that $|E| \le C + |\mathrm{supp}(b)| < n$. Fix $j \in [k+1]$ and consider the block B_j of Definition 1.2. For $i \in B_j \setminus E$, we have $i \in T$ and $b_i = 0$, so $y_i = x_i$.

Hence the image $\prod_{i\in B_j} y_i$ of the block monomial $\mathbf{x}^{B_j}$ is divisible by $M_j := \prod_{i\in B_j\setminus E} x_i$, a product of at least $n - |E| \geq 1$ distinct variables, and so is every monomial occurring in it. Every monomial occurring in f occurs in the image of some block monomial, which proves the second claim.

To see that $f \neq 0$, note that $|[N] \setminus T| \leq k$ and that the $k+1$ blocks are disjoint, so some block B_{j_0} is contained in T. The image of its block monomial is $\prod_{i\in B_{j_0}}(b_i + x_i)$, in which the monomial $\mathbf{x}^{B_{j_0}}$ has coefficient 1. For every $j \neq j_0$, every monomial occurring in the image of $\mathbf{x}^{B_j}$ is divisible by M_j, which involves a variable x_i with $i \in B_j$; as $i \notin B_{j_0}$, such a monomial differs from $\mathbf{x}^{B_{j_0}}$. Hence $\mathbf{x}^{B_{j_0}}$ occurs in f with coefficient 1. □

Proof of Theorem 6.1: Fix a $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit over $\mathbb{F}((\varepsilon))$ of size at most s that computes $P_{k+1,n} + \varepsilon S$ for some $S \in \mathbb{F}[[\varepsilon]][\mathbf{x}]$ (Definition 2.12). Let ρ be the number of distinct bottom polynomials of this circuit. Then $\rho \leq s$, and we apply Theorem 6.7. It provides a set $T \subseteq [N]$ with $C = N - |T| \leq k$, a point $b \in \mathbb{F}^N$ with $|\mathrm{supp}(b)| = O\left(k^2\delta^{3k}\right)$, and the data φ_a, Q_a, k', e_i, and $p_{i,j}$ of the identity (27). If $n \leq C + |\mathrm{supp}(b)|$, then $n \leq k + O\left(k^2\delta^{3k}\right) \leq (2\delta)^{O(k)}$, and there is nothing to prove. So assume $n > C + |\mathrm{supp}(b)|$. By Lemma 6.20, applied to $g_a := \varphi_a(0, \mathbf{u})$, the power series $f_0 = P_{k+1,n}(b + (\mathbf{u}, \boldsymbol{\varphi}(0, \mathbf{u})))$ is nonzero. Every monomial occurring in f_0, and in particular every monomial of $H_{d_0}(f_0)$, involves at least $n - C - |\mathrm{supp}(b)|$ variables.

By Theorem 6.12 and $\rho \leq s$, $H_{d_0}(f_0) \cdot G_2 = G_1$ for some nonzero polynomials $G_1, G_2 \in \overline{\mathcal{C}}_\Delta(m, r)$, where

$$\log m = O\left(k4^k(k + \log s)\right) \qquad \text{and} \qquad r = O\left(k4^k(\Delta+1)^{k+1}\right).$$

We compare trailing monomials. Since trailing monomials are multiplicative (see Monomial Orders), we have $\mathrm{TM}(G_1) = \mathrm{TM}(H_{d_0}(f_0)) \cdot \mathrm{TM}(G_2)$, so the support of $\mathrm{TM}(G_1)$ contains the support of $\mathrm{TM}(H_{d_0}(f_0))$. Combining this inclusion with Theorem 6.6, we get

$$n - C - |\mathrm{supp}(b)| \leq |\mathrm{supp}(\mathrm{TM}(H_{d_0}(f_0)))| \leq |\mathrm{supp}(\mathrm{TM}(G_1))| = O\left(\Delta(r + \log m)\right).$$

It remains to evaluate the bounds. Using $\Delta = \delta^k$, $\Delta + 1 \leq 2\delta^k$, and $k \leq 2^k \leq (2\delta)^k$, we get

$$\begin{aligned} C + |\mathrm{supp}(b)| &\leq k + O\left(k^2\delta^{3k}\right) \leq (2\delta)^{O(k)}, \\ \Delta \log m &= O\left(k(4\delta)^k(k + \log s)\right) = O\left(k(4\delta)^k \log s\right) + (2\delta)^{O(k)}, \\ \Delta r &= O\left(k4^k\delta^k \cdot (2\delta^k)^{k+1}\right) = (2\delta)^{O(k^2)}. \end{aligned}$$

Altogether,

$$n \leq O\left(k(4\delta)^k \log s\right) + (2\delta)^{O(k^2)}.$$

For fixed k and δ, this reads $n \leq O(\log s) + O(1)$, that is, $s = 2^{\Omega(n)}$. □

7 Impossibility Without the Constant-Freeness Assumption

In this section, we remove the constant-freeness assumption from Theorem 3.1, at the expense of a larger threshold on n. We prove that, for sufficiently large n, no $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit computes

$P_{k+1,n}$.

We first establish a rank bound for the group of units modulo constants of an irreducible affine variety, following the approach of Samuel [Sam57; Sam66]; see also Ford [For14] and Rosenlicht [Ros61] and the MathOverflow discussion [Sas11]. We develop this argument in Section 7.2 and Section 7.3 and then use the bound to prove the main impossibility result (Theorem 7.5) in Section 7.4.

Throughout this section, assume the base field $\mathbb{F}$ to be algebraically closed.

7.1 Preliminaries for the Impossibility Proof

We introduce some additional algebraic and geometric notions and results needed in this section.

Projective Space. The *projective space* $\mathbb{P}^N$ consists of nonzero vectors in $\mathbb{F}^{N+1}$ modulo nonzero scalar multiplication, with *homogeneous coordinates* $[x_0 : \cdots : x_N]$. Its closed subsets are the common zero sets of homogeneous polynomials, which are also called *forms*. Each standard chart $U_i := \{ x_i \neq 0 \}$ is isomorphic to $\mathbb{A}^N$ via the coordinates $(x_j/x_i)_{j\neq i}$. We identify $\mathbb{A}^N$ with U_0; its complement is the *hyperplane at infinity* $H_\infty := \mathbf{V}(x_0)$. For $V \subseteq \mathbb{A}^N$, its *projective closure* V_{cl} is its closure in $\mathbb{P}^N$ [CLO25, Chapter 8, §4, Definition 6 and Proposition 7(ii)].

The definition of dimension by chains of irreducible closed subvarieties applies to projective varieties and open subsets of varieties as well. For an irreducible variety X, one has $\dim X = \operatorname{tr.deg}_{\mathbb{F}} \mathbb{F}(X)$, and every nonempty open subset $U \subseteq X$ satisfies $\dim U = \dim X$ [Sha13, Chapter 1, § 6.1]. For an irreducible closed subvariety $E \subseteq X$, its *codimension* in X is $\dim X - \dim E$.

Rational and Regular Functions. For an irreducible variety X, its *function field* is $\mathbb{F}(X) := \operatorname{Frac}(\mathbb{F}[U])$ for any nonempty affine open subset $U \subseteq X$; this field is independent of the choice of U. Its elements are *rational functions*. A rational function is *regular* at a point if it is locally a quotient of polynomial functions with a denominator that is nonzero at that point. It is *regular* on an open subset if it is regular at every point of that subset. On an affine variety U, the everywhere regular functions are precisely the elements of $\mathbb{F}[U]$. A morphism $f\colon X \to Y$ *pulls back* a regular function h on $U \subseteq Y$ to $h \circ f$ on $f^{-1}(U)$. If X and Y are irreducible and f is dominant, then pullback also identifies $\mathbb{F}(Y)$ with a subfield of $\mathbb{F}(X)$.

Principal Open Subsets as Affine Varieties. For an affine variety V and a polynomial H, the set $V \setminus \mathbf{V}(H)$ is a *principal open subset* of V. The next lemma realizes this set as an affine variety by adding a variable for $1/H$ and bounds the degree of the resulting variety.

Lemma 7.1: Let $V \subseteq \mathbb{A}^N$ be irreducible and $H \in \mathbb{F}[\mathbf{x}] \setminus \mathbf{I}(V)$. Write $\overline{H} = H + \mathbf{I}(V)$ and set

$$W = (V \times \mathbb{A}^1) \cap \mathbf{V}(zH - 1).$$

The variety W is irreducible, $\mathbb{F}[W] \cong \mathbb{F}[V]_{\overline{H}}$, and $\deg W \leq \deg V \cdot (\deg H + 1)$.

Proof: We have $\mathbb{F}[V][z]/(z\overline{H} - 1) \cong \mathbb{F}[V]_{\overline{H}}$; see [Mag12]. Since $\mathbb{F}[V]$ is a domain and $\overline{H} \neq 0$, the localization $\mathbb{F}[V]_{\overline{H}}$ is also a domain. Hence $\langle z\overline{H} - 1 \rangle$ is a prime ideal of $\mathbb{F}[V][z]$. It follows

that W is irreducible and $\mathbb{F}[W] \cong \mathbb{F}[V]_{\overline{H}}$. The degree bound follows from $\deg(V \times \mathbb{A}^1) = \deg V$ and Proposition 2.9. □

Finite Morphisms and Degree. A morphism $f\colon X \to Y$ is *finite* if, for every affine open subset $U \subseteq Y$, its inverse image $f^{-1}(U)$ is affine and $\mathbb{F}[f^{-1}(U)]$ is a finitely generated module over $\mathbb{F}[U]$, with the module structure induced by pullback along f. Finite morphisms are closed, and their compositions are finite. For a finite morphism $f\colon X \to Y$ and every nonempty irreducible closed subvariety $Z \subseteq X$, one has $\dim f(Z) = \dim Z$ [Stacks, Tag 01WG].

Let $f\colon X \to Y$ be a finite surjective morphism of irreducible varieties. Pullback by f identifies $\mathbb{F}(Y)$ with a subfield of $\mathbb{F}(X)$, and the extension $\mathbb{F}(X)/\mathbb{F}(Y)$ is finite. The *degree* of f is defined as

$$\deg f := [\mathbb{F}(X) : \mathbb{F}(Y)].$$

This notion of degree is different from the degree of an embedded variety.

Projective Degree. For an irreducible projective variety $X \subseteq \mathbb{P}^N$, its degree is the degree of $X \cap U_i \subseteq \mathbb{A}^N$ for any standard affine chart U_i meeting X; this degree is independent of the choice of chart. In particular, projective closure preserves the degree of an irreducible affine variety [CLO25, Chapter 9, §3].

Let $X \subseteq \mathbb{P}^N$ be irreducible of dimension d. Linear forms $\ell_0, \dots, \ell_d$ with no common zero on X define a finite surjective morphism $\pi = [\ell_0 : \cdots : \ell_d]\colon X \to \mathbb{P}^d$ [Sha13, Chapter 1, §5.3, Theorem 1.16]. Moreover, $\deg \pi = \deg X$ [Stacks, Tag 0BEX].

Normalization. An irreducible variety is *normal* if its affine coordinate rings are normal; normality can be checked on an affine open cover [Stacks, Tag 033J]. In particular, $\mathbb{P}^N$ is normal because its standard charts have polynomial coordinate rings. Every irreducible variety X has a *normalization*: a finite surjective morphism $\nu\colon \widetilde{X} \to X$ with $\widetilde{X}$ normal and $\mathbb{F}(\widetilde{X}) = \mathbb{F}(X)$ under pullback [Stacks, Tag 0BXR]. For every nonempty affine open subset $U \subseteq X$, the ring $\mathbb{F}[\nu^{-1}(U)]$ is the integral closure of $\mathbb{F}[U]$ in $\mathbb{F}(X)$. Thus $\dim \widetilde{X} = \dim X$. If X is projective, then so is $\widetilde{X}$ [Stacks, Tag 0GK4].

Let $f\colon X \to Y$ be a finite surjective morphism of irreducible varieties. For $y \in Y$, the *fiber* of f over y is $f^{-1}(y) = \{\, x \in X \mid f(x) = y \,\}$. We use the following *fiber bound*: If Y is normal, then $\left|f^{-1}(y)\right| \leq \deg f$ for every $y \in Y$ [Sha13, Chapter 2, § 6.3, Theorem 2.28]. Here the cardinality counts distinct points.

Orders of Zeros and Poles. Let X be a normal irreducible variety, and let $E \subseteq X$ be an irreducible closed subvariety of codimension one. Choose an affine open subset $V \subseteq X$ meeting E, and let $\mathfrak{p} \subseteq \mathbb{F}[V]$ be the prime ideal of functions vanishing on $E \cap V$. The ring $R = \mathbb{F}[V]_{\mathfrak{p}}$ is the *local ring of X along E*. It consists of the rational functions regular on some open subset of X meeting E and is independent of the choice of V.

Since X is normal and E has codimension one, R is a discrete valuation ring (see Discrete Valuations). Let $v_E\colon \mathbb{F}(X)^\times \to \mathbb{Z}$ be its normalized valuation, called the *order of vanishing along E*;

negative values indicate poles. For $c \in \mathbb{F}^\times$, we have $v_E(c) = 0$. Moreover, $v_E(h) \geq 0$ whenever h is regular on an open subset of X meeting E.

A nonzero rational function on a normal irreducible variety is regular everywhere if its order along every codimension-one irreducible closed subvariety is nonnegative [Stacks, Tag 031T(2)]. Every regular function on an irreducible projective variety is constant [Sha13, Chapter 1, §5.2, Corollary 1.1].

Tangent Spaces, the Jacobian Matrix, and Smoothness. For $f \in \mathbb{F}[\mathbf{x}]$, the *gradient* of f is $\nabla f = (\partial_{x_1} f, \ldots, \partial_{x_N} f)$, regarded as a row vector. For $f_1, \ldots, f_m \in \mathbb{F}[\mathbf{x}]$, their *Jacobian matrix* is the $m \times N$ matrix $\mathrm{Jac}(f_1, \ldots, f_m) = (\partial_{x_j} f_i)_{i \in [m],\, j \in [N]}$. Evaluation at a point is entry-wise.

For an irreducible affine variety $V \subseteq \mathbb{A}^N$ and $p \in V$, define the *tangent space* of V at p as

$$T_p V := \{\, v \in \mathbb{F}^N \mid \nabla h(p) v = 0 \text{ for every } h \in \mathbf{I}(V) \,\} \subseteq \mathbb{F}^N,$$

where vectors $v \in \mathbb{F}^N$ are regarded as column vectors. For every $p \in V$, one has $\dim T_p V \geq \dim V$ [Har77, Chapter I, Exercise 5.10(a)]. A point $p \in V$ is *smooth* or *nonsingular* if $\dim T_p V = \dim V$.

If a polynomial $g \in \mathbb{F}[\mathbf{x}]$ does not vanish at a point $p \in \mathbb{A}^N$, then the vector $\frac{\nabla g(p)}{g(p)} \in \mathbb{F}^N$ is well-defined. In particular, this vector is well-defined whenever $g|_V$ is a unit and $p \in V$. The following lemma gives useful properties of $\frac{\nabla g(p)}{g(p)}$. In particular, it shows that the map $g \mapsto \frac{\nabla g(p)}{g(p)}$ can be viewed as a multivariate analogue of the logarithmic derivative.

Lemma 7.2: Let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety, $p \in V$, and $v \in T_p V$. Let $g, h \in \mathbb{F}[\mathbf{x}]$ be polynomials whose restrictions to V are units. The following statements hold:

1. $\frac{\nabla g(p) \cdot v}{g(p)}$ depends only on $g|_V$.
2. If $g \in \mathbb{F}^\times$, then $\frac{\nabla g(p)}{g(p)} = \mathbf{0}$.
3. $\frac{\nabla (gh)(p)}{(gh)(p)} = \frac{\nabla g(p)}{g(p)} + \frac{\nabla h(p)}{h(p)}$.
4. If $h|_V = (g|_V)^e$ for some $e \in \mathbb{Z}$, then $\frac{\nabla h(p) \cdot v}{h(p)} = e \frac{\nabla g(p) \cdot v}{g(p)}$.

Proof: If $g|_V = h|_V$, then $g(p) = h(p)$ and $\nabla(g - h)(p) \cdot v = 0$. Dividing by the common nonzero value $g(p) = h(p)$ proves the first claim. The second and third claims follow from differentiation of constants and the product rule. If $(gh)|_V = 1$, then the first three claims give $\nabla g(p) \cdot v / g(p) + \nabla h(p) \cdot v / h(p) = 0$. Repeated application of the product rule and this inverse identity proves the fourth claim for every integer exponent. □

7.2 Normalization and Irreducible Components at Infinity

To obtain the unit-group rank bound in the next subsection, we first bound the number of irreducible components lying above the hyperplane at infinity in the normalization of a projective closure.

Consider the intersection of the projective closure of an irreducible affine variety with the hyperplane at infinity. The following theorem bounds the number of irreducible components of the preimage of this intersection under the normalization map by the degree of the affine variety. We use this bound in the next subsection to establish a rank bound for the group of units modulo constants of an irreducible affine variety.

Theorem 7.3: Let $W_0 \subseteq \mathbb{A}^N$ be an irreducible affine variety of degree Δ and dimension $d \geq 1$, and let $W \subseteq \mathbb{P}^N$ be its projective closure. Let H_∞ be the hyperplane at infinity, and let $\nu\colon \widetilde{W} \to W$ be the normalization. Let $D_\infty := \nu^{-1}(W \cap H_\infty)$. The set D_∞ has at most Δ irreducible components, each of dimension $d-1$.

Proof: Use homogeneous coordinates $[x_0 : \cdots : x_N]$ with $H_\infty = \mathbf{V}(x_0)$. Choose linear forms $\ell_1, \ldots, \ell_d \in \mathbb{F}[\mathbf{x}]$ such that $W \cap \mathbf{V}(x_0, \ell_1, \ldots, \ell_d) = \emptyset$. Such a choice is possible because $\dim(W \cap H_\infty) \leq d-1$: successively choosing hyperplanes that contain none of the remaining irreducible components cuts this set down to the empty set.

Define $\pi\colon W \to \mathbb{P}^d$ by

$$\pi([x_0 : \cdots : x_N]) = [x_0 : \ell_1(\mathbf{x}) : \cdots : \ell_d(\mathbf{x})].$$

By the results in Projective Degree, π is finite and surjective.

Write $[y_0 : \cdots : y_d]$ for the homogeneous coordinates on the target $\mathbb{P}^d$, and let $H'_\infty = \mathbf{V}(y_0)$. By the definition of π, we have $\pi^{-1}(H'_\infty) = W \cap H_\infty$ and $\pi^{-1}(\mathbb{A}^d) = W_0$, where $\mathbb{A}^d = \mathbb{P}^d \setminus H'_\infty$.

Since π is a finite linear projection, the degree formula in Projective Degree gives

$$\deg \pi = \deg W = \deg W_0 = \Delta,$$

where the middle equality holds because W is the projective closure of W_0.

Let $\tau = \pi \circ \nu\colon \widetilde{W} \to \mathbb{P}^d$. Both π and ν are finite and surjective, so the same holds for τ. Since ν induces an isomorphism of function fields, $\deg \tau = \deg \pi = \Delta$. Moreover, $\tau^{-1}(H'_\infty) = \nu^{-1}(W \cap H_\infty) = D_\infty$.

Let $E_1, \ldots, E_s$ be the irreducible components of D_∞. We prove $s \leq \Delta$ by finding a point of H'_∞ with at least s preimages under τ.

First, we show that $\dim E_i = d-1$ and $\tau(E_i) = H'_\infty$ for every $i \in [s]$. Fix $i \in [s]$ and $b \in E_i$. Choose $t \in [d]$ such that $\tau(b) \notin \mathbf{V}(y_t)$. Let $U_t = \mathbb{P}^d \setminus \mathbf{V}(y_t)$ and $U = \tau^{-1}(U_t)$. The set U is affine because U_t is affine and τ is finite. It is a nonempty open subset of $\widetilde{W}$, so it is irreducible of dimension d. On U, the set $D_\infty \cap U$ is the zero set of the regular function $g = (y_0/y_t) \circ \tau$. Since $\tau(U) = U_t$, the function g does not vanish identically; also, $g(b) = 0$. The nonempty intersections $E_j \cap U$ are exactly the irreducible components of this zero set. In particular, $E_i \cap U$ is one of them. Since U is affine, the function g is represented by a polynomial. Thus Lemma 2.7 gives $\dim(E_i \cap U) = d-1$. Since $E_i \cap U$ is a nonempty open subset of the irreducible variety E_i, we have $\dim E_i = d-1$. Finally, finiteness of τ implies that $\tau(E_i)$ is an irreducible closed subset of H'_∞ of dimension $d-1$, so $\tau(E_i) = H'_\infty$.

For $i, j \in [s]$ with $i \neq j$, the intersection $E_i \cap E_j$ is a proper closed subset of E_i. Thus the

intersection is empty, or each of its irreducible components has dimension strictly less than $d-1$. Since τ is finite, $\tau(E_i \cap E_j)$ is therefore a proper closed subset of H'_∞ in either case. Since H'_∞ is irreducible, the finite union $Z := \bigcup_{1 \le i < j \le s} \tau(E_i \cap E_j)$ is also a proper closed subset of H'_∞. Choose $p \in H'_\infty \setminus Z$. For each $i \in [s]$, choose $e_i \in E_i$ with $\tau(e_i) = p$, using $\tau(E_i) = H'_\infty$. These points are distinct: Otherwise, p would belong to $\tau(E_i \cap E_j)$ for some $i \neq j$. Hence $\left|\tau^{-1}(p)\right| \ge s$.

Since τ is finite and surjective and $\mathbb{P}^d$ is normal, each fiber of τ contains at most $\deg \tau$ points [Sha13, Chapter 2, § 6.3, Theorem 2.28]. Therefore,

$$s \le \left|\tau^{-1}(p)\right| \le \deg \tau = \Delta.$$

Together with $\dim E_i = d-1$, this bound proves the theorem. □

7.3 A Rank Bound for the Group of Units Modulo Constants

We now use Theorem 7.3 to bound the rank of the group of units modulo constants of an irreducible affine variety. The idea is to record the orders of zeros and poles of each unit along the irreducible components counted in Theorem 7.3. We show that these integers determine the unit up to a nonzero constant.

Theorem 7.4: Let $W_0 \subseteq \mathbb{A}^N$ be an irreducible affine variety of degree Δ and dimension $d \ge 1$. The group $\mathbb{F}[W_0]^\times / \mathbb{F}^\times$ is free abelian of rank at most Δ.

Proof: Let W be the projective closure of W_0, and let $\nu \colon \widetilde{W} \to W$ be its normalization. Let $U = \nu^{-1}(W_0)$ and $D_\infty = \nu^{-1}(W \cap H_\infty) = \widetilde{W} \setminus U$. Let $E_1, \dots, E_s$ be the irreducible components of D_∞. By Theorem 7.3, $s \le \Delta$ and $\dim E_i = d-1$ for every $i \in [s]$. We identify $\mathbb{F}(W_0) = \mathbb{F}(W) = \mathbb{F}(\widetilde{W})$ via restriction and pullback and thus regard each unit on W_0 as a nonzero rational function on $\widetilde{W}$.

Define $\varphi \colon \mathbb{F}[W_0]^\times / \mathbb{F}^\times \to \mathbb{Z}^s$ by

$$\varphi(\bar{t}) = (v_{E_1}(t), \dots, v_{E_s}(t)),$$

where $\bar{t}$ denotes the natural image of $t \in \mathbb{F}[W_0]^\times$ in $\mathbb{F}[W_0]^\times / \mathbb{F}^\times$. The identities $v_E(ct) = v_E(t)$ for $c \in \mathbb{F}^\times$ and $v_E(t_1 t_2) = v_E(t_1) + v_E(t_2)$ show that φ is a well-defined group homomorphism.

Suppose that $t \in \mathbb{F}[W_0]^\times$ satisfies $\varphi(\bar{t}) = \mathbf{0}$. The definition of φ gives $v_{E_i}(t) = 0$ for every $i \in [s]$. Since t and t^{-1} are regular on W_0, their pullbacks are regular on U. Let $E \subseteq \widetilde{W}$ be any codimension-one irreducible closed subvariety different from the E_i. The subvariety E meets U. Otherwise, E would be contained in some E_i, and equality of their dimensions would force $E = E_i$. It follows that $v_E(t) \ge 0$ and $v_E(t^{-1}) \ge 0$. Since $v_E(t^{-1}) = -v_E(t)$, we obtain $v_E(t) = 0$.

Therefore, $v_E(t) = 0$ for every codimension-one irreducible closed subvariety $E \subseteq \widetilde{W}$. Since $\widetilde{W}$ is normal and the order of t along every codimension-one irreducible closed subvariety is nonnegative, t is regular everywhere on $\widetilde{W}$ [Stacks, Tag 031T(2)]. The variety $\widetilde{W}$ is projective and irreducible, so t is constant. Since $t \neq 0$, it belongs to $\mathbb{F}^\times$. Hence $\bar{t}$ is the identity, and φ is injective.

Since φ is injective, it follows that $\mathbb{F}[W_0]^\times / \mathbb{F}^\times$ is free abelian of rank at most $s \le \Delta$. □

7.4 Proof of the Impossibility Result

We prove the following theorem, which is the main result of this section.

Theorem 7.5: Let $\delta, k \geq 1$ and suppose that $n > (2k+\delta)\delta^{2k} + \delta^k$. The polynomial $P_{k+1,n}$ cannot be computed by $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuits.

Let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety, and let $C := \operatorname{codim} V$. The smooth locus of V is dense and open [Har77, Chapter I, Theorem 5.3]. At a smooth point $p \in V$, the space of linear forms on $\mathbb{F}^N$ vanishing on T_pV has dimension C and is spanned by the gradients of polynomials in $\mathbf{I}(V)$. We can therefore choose $f_1, \ldots, f_C \in \mathbf{I}(V)$ whose gradients at p form a basis of this space. Some $C \times C$ minor of $\operatorname{Jac}(f_1, \ldots, f_C)$ is then nonzero at p and remains nonzero on an open neighborhood of p in V. Intersecting this neighborhood with the smooth locus gives a dense smooth open subset $U \subseteq V$ such that, for every $q \in U$, the gradients $\nabla f_1(q), \ldots, \nabla f_C(q)$ form a basis of the space of linear forms vanishing on T_qV. The following lemma makes this observation *effective* by showing that the polynomials $f_1, \ldots, f_C$ can be chosen to have degree at most $\deg V$.

Lemma 7.6: Let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety of dimension d and degree Δ. Let $C := \operatorname{codim} V = N - d$. There exist $f_1, \ldots, f_C \in \mathbf{I}(V)$ of degree at most Δ and a nonempty smooth open subset $U \subseteq V$ such that, for every $p \in U$, the rows $\nabla f_1(p), \ldots, \nabla f_C(p)$ form a basis of the space of linear forms on $\mathbb{F}^N$ vanishing on T_pV.

Proof: If $C = 0$, then $V = \mathbb{A}^N$ and we may take $U = V$. Assume $C > 0$.

By Lemma 2.13, choose $T \subseteq [N]$ of size d such that $(x_i|_V)_{i \in T}$ is a transcendence basis of $\mathbb{F}(V)$ over $\mathbb{F}$. Write $[N] \setminus T = \{a_1, \ldots, a_C\}$. For each $i \in [C]$, Lemma 2.17 gives an irreducible polynomial $f_i \in \mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_{T \cup \{a_i\}}]$ with $\deg f_i \leq \Delta$ and $\deg_{x_{a_i}} f_i \geq 1$.

We first show that $\partial f_i / \partial x_{a_i}$ does not vanish identically on V. By Gauss's Lemma, f_i, viewed as a univariate polynomial in x_{a_i} over $\mathbb{F}(x_i|_V : i \in T)$, is irreducible. Since it vanishes at $x_{a_i}|_V$, it is a nonzero scalar multiple of the minimal polynomial of $x_{a_i}|_V$ over $\mathbb{F}(x_i|_V : i \in T)$. Since $\mathbb{F}$ has characteristic zero, that minimal polynomial has distinct roots. Thus $x_{a_i}|_V$ is a simple root, so the derivative of the minimal polynomial does not vanish there, i.e., $\frac{\partial f_i}{\partial x_{a_i}} \notin \mathbf{I}(V)$ for all $i \in [C]$.

Define

$$U = \left\{ p \in V \;\middle|\; \frac{\partial f_i}{\partial x_{a_i}}(p) \neq 0 \text{ for every } i \in [C] \right\}.$$

Since none of these partial derivatives vanishes identically on V and V is irreducible, U is a nonempty open subset of V.

Fix $p \in U$. We claim that $\nabla f_1(p), \ldots, \nabla f_C(p)$ are linearly independent. Each f_i involves only the variables indexed by T and the variable x_{a_i}. Hence, for $i \neq j$, the variable x_{a_j} does not occur in f_i, so $\partial f_i / \partial x_{a_j} = 0$. Therefore, the $C \times C$ submatrix of $\operatorname{Jac}(f_1, \ldots, f_C)(p)$ consisting of the columns indexed by $a_1, \ldots, a_C$ is the diagonal matrix

$$\left(\frac{\partial f_i}{\partial x_{a_j}}(p) \right)_{i,j \in [C]} = \operatorname{diag}\left(\frac{\partial f_1}{\partial x_{a_1}}(p), \ldots, \frac{\partial f_C}{\partial x_{a_C}}(p) \right).$$

All its diagonal entries are nonzero by the choice of U. Thus $\mathrm{Jac}(f_1, \ldots, f_C)(p)$ has rank C, proving the claim.

Finally, each f_i belongs to $\mathbf{I}(V)$, so each $\nabla f_i(p)$ vanishes on T_pV. Since these C gradients are linearly independent, we have

$$\dim T_pV \leq N - C = d.$$

Together with the general inequality $\dim T_pV \geq \dim V = d$, this bound gives $\dim T_pV = d$, so p is smooth. The space of linear forms on $\mathbb{F}^N$ vanishing on T_pV therefore has dimension $N - d = C$, and $\nabla f_1(p), \ldots, \nabla f_C(p)$ form a basis of this space. □

The next lemma concerns an irreducible affine variety V of small degree and codimension on which a product of n algebraically independent coordinate functions can be expressed as a sum of products of low-degree units. It gives a low-degree polynomial that does not vanish identically on V and whose zero locus contains the zero locus of this product on V.

Lemma 7.7: Let $\delta, k, n \geq 1$, and let $V \subseteq \mathbb{A}^N$ be an irreducible affine variety of degree Δ and dimension $d \geq 1$. Let $C := \operatorname{codim} V = N - d$. Suppose that $x_{i_1}, \ldots, x_{i_n}$ are distinct coordinates such that $\mathbb{F}[x_{i_1}, \ldots, x_{i_n}] \cap \mathbf{I}(V) = \{0\}$ and

$$\prod_{j=1}^{n} x_{i_j} \equiv \sum_{i=1}^{k} T_i \pmod{\mathbf{I}(V)},$$

where each T_i is a product of polynomials $q_{i,j}$ of degree at most δ that restrict to units on V. There exists a polynomial $H \in \mathbb{F}[\mathbf{x}] \setminus \mathbf{I}(V)$ of degree at most $(C + \delta)\Delta$ such that $\mathbf{V}\left(\prod_{j=1}^{n} x_{i_j}\right) \cap V \subseteq \mathbf{V}(H)$.

Proof: Choose $f_1, \ldots, f_C \in \mathbf{I}(V)$ of degree at most Δ and a nonempty smooth open subset $U \subseteq V$ such that, for every $p \in U$, the rows $\nabla f_1(p), \ldots, \nabla f_C(p)$ form a basis of the space of linear forms on $\mathbb{F}^N$ vanishing on T_pV. This choice is possible by Lemma 7.6. Write $P := \prod_{j=1}^{n} x_{i_j}$. Since $\mathbb{F}[x_{i_1}, \ldots, x_{i_n}] \cap \mathbf{I}(V) = \{0\}$, we have $P \notin \mathbf{I}(V)$. Since V is irreducible, we may shrink U to a nonempty open subset on which P is nowhere zero.

By Theorem 7.4, $\mathbb{F}[V]^\times / \mathbb{F}^\times$ is free abelian of rank $r \leq \Delta$. Choose polynomials $u_1, \ldots, u_r \in \mathbb{F}[\mathbf{x}]$ whose restrictions to V are units and whose classes in $\mathbb{F}[V]^\times / \mathbb{F}^\times$ form a basis of this group. Each factor $q_{i,j}$ therefore satisfies

$$q_{i,j}|_V = \lambda_{i,j} \prod_{a=1}^{r} (u_a|_V)^{e_{i,j,a}} \tag{36}$$

in $\mathbb{F}[V]^\times$, where $\lambda_{i,j} \in \mathbb{F}^\times$ and $e_{i,j,a} \in \mathbb{Z}$. By Lemma 7.2, for every $p \in V$ and $v \in T_pV$, this identity gives

$$\frac{\nabla q_{i,j}(p) \cdot v}{q_{i,j}(p)} = \sum_{a=1}^{r} e_{i,j,a} \frac{\nabla u_a(p) \cdot v}{u_a(p)}. \tag{37}$$

For $f \in \mathbb{F}[\mathbf{x}]$, write $\nabla f(p)|_{T_pV}$ for the linear map $v \mapsto \nabla f(p) \cdot v$ on T_pV. All denominators in (37) are nonzero. Hence the linear maps $\nabla q_{i,j}(p)|_{T_pV}$ lie in the span of $\nabla u_1(p)|_{T_pV}, \ldots, \nabla u_r(p)|_{T_pV}$ and span a space of dimension at most r.

Let J be the Jacobian matrix whose first C rows are $\nabla f_1, \dots, \nabla f_C$ and whose remaining rows are all the $\nabla q_{i,j}$, in any fixed order. Fix $p \in U$ and consider the linear map that restricts each row vector in the row space of $J(p)$ to a linear form on T_pV. Its image has dimension at most r by the preceding paragraph. Its kernel is exactly the span of $\nabla f_1(p), \dots, \nabla f_C(p)$: These rows belong to the row space of $J(p)$ and form a basis of all linear forms vanishing on T_pV. Thus the kernel has dimension C, and the rank–nullity theorem gives

$$\operatorname{rank} J(p) \le C + r.$$

In addition, the first C rows of $J(p)$ are linearly independent.

Let B be the matrix over $\mathbb{F}[\mathbf{x}]$ obtained from J by multiplying its i_j-th column by x_{i_j} for each $j \in [n]$, leaving the other columns unchanged. Since these coordinates are nonzero on U, for every $p \in U$, we have

$$\operatorname{rank} B(p) = \operatorname{rank} J(p) \le C + r,$$

and the first C rows of $B(p)$ are linearly independent.

We next exhibit a fixed nonzero vector in the row space of $B(p)$ for every $p \in U$. Since $P - \sum_{i=1}^{k} T_i \in \mathbf{I}(V)$, the row

$$\nabla P(p) - \sum_{i=1}^{k} \nabla T_i(p)$$

vanishes on T_pV and therefore belongs to the span of $\nabla f_1(p), \dots, \nabla f_C(p)$. Also, each T_i is a product of the factors $q_{i,j}$, so the product rule expresses $\nabla T_i(p)$ as a linear combination of the rows $\nabla q_{i,j}(p)$ of $J(p)$. It follows that $\nabla P(p)$ belongs to the row space of $J(p)$. Since $P(p) \neq 0$, the same holds for

$$t(p) \coloneqq \frac{\nabla P(p)}{P(p)} = \sum_{j=1}^{n} \frac{\nabla x_{i_j}(p)}{x_{i_j}(p)}, \tag{38}$$

where the equality follows from the product rule. Because $i_1, \dots, i_n$ are distinct, the i_j-th coordinate of $t(p)$ is $1/x_{i_j}(p)$, and all other coordinates are zero. Applying the same coordinate scaling used to obtain $B(p)$ from $J(p)$ sends $t(p)$ to the fixed row vector $w \in \mathbb{F}^N$ whose coordinates indexed by $i_1, \dots, i_n$ are one and whose other coordinates are zero. Thus w belongs to the row space of $B(p)$ for every $p \in U$. Note that w is defined at every point of V.

Choose $p_0 \in U$ at which $\operatorname{rank} B(p_0)$ attains its maximum, and let $m = \operatorname{rank} B(p_0)$. We have $C \le m \le C + r$, and $m \ge 1$ since $w \neq \mathbf{0}$. Extend the first C rows of $B(p_0)$ to a basis of its row space by choosing $m - C$ additional rows. Let $b_1, \dots, b_m \in \mathbb{F}[\mathbf{x}]^N$ be the corresponding rows of B. Choose a set $S \subseteq [N]$ of m columns such that the $m \times m$ matrix M formed by these rows and columns satisfies $\det M(p_0) \neq 0$. Let $H \coloneqq \det M$. By construction, $H(p_0) \neq 0$, so $H \notin \mathbf{I}(V)$.

We claim that w belongs to the span of $b_1(p), \dots, b_m(p)$ for every $p \in V \setminus \mathbf{V}(H)$. First consider $p \in U \setminus \mathbf{V}(H)$. Since $M(p)$ is invertible and $\operatorname{rank} B(p) \le m$, these m rows form a basis of the row space of $B(p)$. In particular, they span w. Let w_S be the restriction of w to the columns indexed by S. For $a \in [m]$, let A_a be the determinant of the matrix obtained from M by replacing its a-th row

with w_S. Solving for the coefficients on the columns indexed by S using Cramer's Rule, we obtain

$$w = \sum_{a=1}^{m} \frac{A_a(p)}{H(p)} b_a(p) \qquad \text{for every } p \in U \setminus \mathbf{V}(H).$$

Multiplying this identity by $H(p)$ shows that every coordinate of the polynomial row vector

$$Hw - \sum_{a=1}^{m} A_a b_a$$

vanishes on $U \setminus \mathbf{V}(H)$. The set $U \setminus \mathbf{V}(H)$ is a nonempty open subset of the irreducible variety V (it contains p_0), so it is dense in V. Consequently, all these coordinate polynomials vanish on V. For any $p \in V \setminus \mathbf{V}(H)$, we may now divide by $H(p)$ to obtain the same expression for w as a linear combination of $b_1(p), \ldots, b_m(p)$, proving the claim.

Now suppose $p \in V$ and $P(p) = 0$, i.e., $\prod_{j=1}^{n} x_{i_j}(p) = 0$. Choose $j \in [n]$ such that $x_{i_j}(p) = 0$. The i_j-th column of $B(p)$ is zero by the definition of B, whereas the i_j-th coordinate of w is one. Thus w cannot belong to the row space of $B(p)$. The claim just proved therefore forces $H(p) = 0$. We conclude that

$$\mathbf{V}(P) \cap V \subseteq \mathbf{V}(H).$$

Finally, each entry in the first C rows of B is a partial derivative of some f_i, possibly multiplied by one coordinate, and therefore has degree at most Δ. Similarly, each entry in the remaining rows is a partial derivative of some $q_{i,j}$, possibly multiplied by one coordinate, and therefore has degree at most δ. The determinant defining H uses the first C rows and $m - C$ additional rows, so

$$\deg H \leq C\Delta + (m - C)\delta \leq C\Delta + r\delta \leq (C + \delta)\Delta.$$

This degree bound completes the proof. □

The next lemma gives an upper bound on n under the assumptions of Lemma 7.7. The main idea is to pass to the open subset $V \setminus \mathbf{V}(H)$, on which $x_{i_1}, \ldots, x_{i_n}$ are units. Although this subset need not be closed in $\mathbb{A}^N$, it can be embedded as an irreducible affine variety $W \subseteq \mathbb{A}^{N+1}$ with coordinate ring isomorphic to $\mathbb{F}[V]_H$. We then apply the unit-group rank bound (Theorem 7.4) to W and use the algebraic independence of $x_{i_1}, \ldots, x_{i_n}$ to bound n.

Lemma 7.8: With the notation and assumptions of Lemma 7.7, we have $n \leq (C + \delta)\Delta^2 + \Delta$.

Proof: Let $H \in \mathbb{F}[\mathbf{x}] \setminus \mathbf{I}(V)$ be the polynomial provided by Lemma 7.7, and let $\overline{H} = H + \mathbf{I}(V) \in \mathbb{F}[V] \setminus \{0\}$. Set $W := (V \times \mathbb{A}^1) \cap \mathbf{V}(zH - 1)$. By Lemma 7.1, W is irreducible, $\mathbb{F}[W] \cong \mathbb{F}[V]_{\overline{H}}$, and

$$\deg W \leq \Delta(\deg H + 1) \leq (C + \delta)\Delta^2 + \Delta.$$

By Theorem 7.4, $\mathbb{F}[W]^\times / \mathbb{F}^\times$ is a free abelian group, and

$$\operatorname{rank}_{\mathbb{Z}}(\mathbb{F}[W]^\times / \mathbb{F}^\times) \leq \deg W \leq (C + \delta)\Delta^2 + \Delta. \tag{39}$$

By Lemma 7.7, if $\prod_{j=1}^{n} x_{i_j}$ vanishes at a point in V, then H also vanishes at that point. Thus $\prod_{j=1}^{n} x_{i_j}$ cannot vanish at any point of W, where $z\overline{H} = 1$. Consequently, each x_{i_j} is a unit in $\mathbb{F}[W]$.

Next, we show that the images of $x_{i_1}, \ldots, x_{i_n}$ in $\mathbb{F}[W]^\times/\mathbb{F}^\times$ are multiplicatively independent. Since $\mathbb{F}[V]$ is an integral domain and $\overline{H} \neq 0$ in $\mathbb{F}[V]$, the natural map $\mathbb{F}[V] \to \mathbb{F}[V]_{\overline{H}} \cong \mathbb{F}[W]$ is injective. If a polynomial $P \in \mathbb{F}[y_1, \ldots, y_n]$ satisfies $P(x_{i_1}, \ldots, x_{i_n}) = 0$ in $\mathbb{F}[W]$, then injectivity implies that the same relation holds in $\mathbb{F}[V]$. The algebraic independence of $x_{i_1}, \ldots, x_{i_n}$ on V then forces $P = 0$. Thus $x_{i_1}, \ldots, x_{i_n}$ remain algebraically independent on W. In particular, their images in $\mathbb{F}[W]^\times/\mathbb{F}^\times$ are multiplicatively independent: Any nontrivial multiplicative relation modulo constants would, after clearing negative exponents, yield a nontrivial polynomial relation among these coordinates.

Therefore, $n \leq \operatorname{rank}_{\mathbb{Z}}(\mathbb{F}[W]^\times/\mathbb{F}^\times)$. Combining this with (39) proves the lemma. □

Now we are ready to prove Theorem 7.5.

Proof of Theorem 7.5: Suppose to the contrary that $P_{k+1,n} = \sum_{i=1}^{k} T_i$ is such a circuit, and let $N := (k+1)n$. Apply Lemma 2.21 to obtain an irreducible affine variety $V \subseteq \mathbb{A}^N$ and a set $K \subseteq [k]$ with the following properties:

1. $\operatorname{codim} V \leq k - |K|$ and $\deg V \leq \delta^{k-|K|}$,
2. $T_i \in \mathbf{I}(V)$ for every $i \in [k] \setminus K$, and
3. $q_{i,j}|_V$ is a unit of $\mathbb{F}[V]$ for every $i \in K$ and every j.

Let $h := k - |K|$. After relabeling the product gates, we may assume $K = [k-h]$. In particular, $\operatorname{codim} V \leq h \leq k$ and $\deg V \leq \delta^h \leq \delta^k$. We have

$$P_{k+1,n} \equiv \sum_{i=1}^{k-h} T_i \pmod{\mathbf{I}(V)}, \tag{40}$$

where every factor of a surviving T_i is a unit on V.

By Lemma 2.19, there is $j \in [k+1]$ with $\mathbf{I}(V) \cap \mathbb{F}[\mathbf{x}_j] = \{0\}$. Our assumption implies $n > k\delta^k \geq \operatorname{codim} V \cdot \deg V$. By Lemma 2.20, there exist indices $\ell_1, \ldots, \ell_k \in [N] \setminus B_j$ and an irreducible component V_0 of $V \cap \bigcap_{i=1}^{k} \mathbf{V}(x_{\ell_i})$ such that $\mathbf{I}(V_0) \cap \mathbb{F}[\mathbf{x}_j] = \{0\}$ and

$$P_{k+1,n} \equiv \mathbf{x}^{B_j} \pmod{\mathbf{I}(V_0)}. \tag{41}$$

Combining (40) and (41), we have

$$\mathbf{x}^{B_j} \equiv \sum_{i=1}^{k-h} T_i \pmod{\mathbf{I}(V_0)}. \tag{42}$$

The left side of (42) is nonzero modulo $\mathbf{I}(V_0)$ since $\mathbf{I}(V_0) \cap \mathbb{F}[\mathbf{x}_j] = \{0\}$. Thus at least one summand T_i survives. Furthermore, $\operatorname{codim} V_0 \leq \operatorname{codim} V + k \leq 2k$ and $\deg V_0 \leq \deg V \leq \delta^k$, by Lemma 2.7 and Proposition 2.9. The factors of each surviving T_i are units on V and hence also

units on $V_0 \subseteq V$. Applying Lemma 7.8 yields

$$n \leq (\operatorname{codim} V_0 + \delta)(\deg V_0)^2 + \deg V_0 \leq (2k+\delta)\delta^{2k} + \delta^k,$$

contradicting the assumption that $n > (2k+\delta)\delta^{2k} + \delta^k$. □

Remark 7.9: Much of the work in proving Theorem 7.5 goes into constructing an auxiliary variety on which the coordinates x_{i_j} are units, while controlling its degree and dimension. This construction is designed to obtain the singly exponential threshold $n > (2k+\delta)\delta^{2k} + \delta^k$. However, a simpler argument suffices if one only seeks a threshold depending on k and δ, without requiring a singly exponential bound. We sketch the ideas as follows.

In the setting of Lemma 7.8, the unit-group rank bound shows that fixing the values of a bounded number of factors $q_{i,j}$ suffices to make all factors constant on each irreducible component of the resulting subvariety. Fix these values at a point where $\prod_{j=1}^{n} x_{i_j}$ is nonzero, and choose a component containing that point. On this component, the relation $\prod_{j=1}^{n} x_{i_j} \equiv \sum_{i=1}^{k} T_i \pmod{\mathbf{I}(V)}$ makes the left side a nonzero constant, so each x_{i_j} is a unit. A second application of the unit-group rank bound, together with a dimension count, then yields a simpler proof of impossibility for sufficiently large n. The drawback is that the number of factors fixed may already be exponential in k and the degree bound from Bézout's Inequality is exponential in this number. This degree growth leads to a doubly exponential threshold for n, rather than the singly exponential threshold in Theorem 7.5. ◁

Remark 7.10: The differential argument in the proof of Lemma 7.7 has a natural interpretation in terms of *logarithmic differentials*. For a polynomial f, its *differential* is written as

$$\mathrm{d}f = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i} \mathrm{d}x_i,$$

so the coordinates of ∇f are the coefficients of $\mathrm{d}f$ in the basis $\mathrm{d}x_1, \ldots, \mathrm{d}x_N$. For each selected coordinate x_{i_j}, we can instead write

$$\frac{\partial f}{\partial x_{i_j}} \mathrm{d}x_{i_j} = x_{i_j} \frac{\partial f}{\partial x_{i_j}} \frac{\mathrm{d}x_{i_j}}{x_{i_j}}.$$

Thus, multiplying the i_j-th column of the Jacobian matrix by x_{i_j} as in the proof of Lemma 7.7 amounts to expressing these differentials using $\mathrm{d}x_{i_j}/x_{i_j}$ in place of $\mathrm{d}x_{i_j}$. In this basis, the logarithmic differential

$$\frac{\mathrm{d}\left(\prod_{j=1}^{n} x_{i_j}\right)}{\prod_{j=1}^{n} x_{i_j}} = \sum_{j=1}^{n} \frac{\mathrm{d}x_{i_j}}{x_{i_j}}$$

is represented by the row w in the proof of Lemma 7.7.

Logarithmic differentials are standard tools in logarithmic geometry for studying *normal crossings divisors*. Informally, a normal crossings divisor is a hypersurface that locally, in suitable coordinates, looks like a union of coordinate hyperplanes, given by an equation $z_1 \cdots z_r = 0$. In

our setting, the hypersurface $\mathbf{V}\left(\prod_{j=1}^{n} x_{i_j}\right) \subseteq \mathbb{A}^N$ is precisely such a divisor. See [Del70, Chapter II, §3] and [Kat89] for background. ◁

8 Conclusion

We showed that for border depth-4 circuits with bounded bottom fan-in, increasing the top fan-in by one can exponentially reduce the required size. For fixed $k, \delta \geq 1$, the polynomial $P_{k+1,n}$ has a depth-3 circuit of top fan-in $k+1$ and size $O(kn)$. For sufficiently large n, no $\Sigma^{[k]}\Pi\Sigma\Pi^{[\delta]}$ circuit computes it exactly (Theorem 1.3). Any such circuit that border computes it must have size at least $2^{\Omega(n)}$ (Theorem 1.4). This result extends the border top fan-in hierarchy of Dutta and Saxena [DS26] from depth 3 to depth 4 with bounded bottom fan-in.

Extending the border argument of Dutta and Saxena [DS26] to $\delta \geq 2$ requires new ideas in both of its steps, and we hope that the resulting tools will be useful elsewhere. First, reducing modulo bottom polynomials of degree δ leads to varieties, which in the border setting depend on ε and may ramify as $\varepsilon \to 0$. Our border low-weight translation (Theorem 6.7), which applies to any polynomial, handles this difficulty by expanding around a point with few nonzero coordinates. The expansion uses Weierstrass preparation (Theorem 6.2) and the Abhyankar–Jung Theorem (Theorem 6.3). Second, the factors of the surviving gates become power series, so we modify DiDIL and replace the geometric series of [DS26] by the chain rule (Lemma 6.16).

We conclude with three open problems.

1. **Better dependence on k and δ.** Theorem 1.4 gives $s \geq 2^{\Omega(n/(k(4\delta)^k))}$ for all sufficiently large n. Can the denominator be improved to $\mathrm{poly}(k, \delta)$? Our approach loses the factor δ^k already in the reduction, through Bézout's Inequality, and we do not know whether this loss is necessary, even for exact computation. Combining our reduction with the Wronskian approach of Bhat, Dutta, and Shringi [BDS26] improves the dependence on k but does not improve this factor; see the discussion in Section 1.6.

2. **Unbounded bottom fan-in.** For every fixed k, is there an explicit family of polynomials that has polynomial-size $\Sigma^{[k+1]}\Pi\Sigma\Pi$ circuits but requires superpolynomial-size $\overline{\Sigma^{[k]}\Pi\Sigma\Pi}$ circuits, with unrestricted bottom fan-in in both models? Here $P_{k+1,n}$ is no longer hard, since it can itself be a bottom polynomial, and our bounds become trivial when δ is comparable to the circuit size.

3. **Beyond depth four.** Rank-based lower bounds carry over to the border setting by work of Grochow, Mulmuley, and Qiao [GMQ16]; for example, the superpolynomial lower bounds of Limaye, Srinivasan, and Tavenas [LST25] also hold for border circuits of constant depth as shown by Andrews and Forbes [AF22]. However, such rank measures are subadditive, and $P_{k+1,n}$ is a sum of only $k+1$ monomials. These measures alone are therefore unlikely to separate top fan-in k from $k+1$. Is there a superpolynomial top fan-in hierarchy for $\overline{\Sigma\Pi\Sigma\Pi\Sigma}$ circuits?

Acknowledgments. AD thanks Dave Anderson and Andrei Staicu for helpful conversations and feedback. This project started while PD was a Jane Street and Simons-Berkeley Research Fellow at the Simons Institute for the Theory of Computing during the Fall 2025 program on Complexity

and Linear Algebra. The initial ideas for exact computation were developed when PD visited ZG at The Ohio State University in October 2025. Developing the full set of results, including the extension to the border setting, took more than a year of further work. Part of this work was also carried out during ZG's visit to the Simons Institute in Fall 2026. PD and ZG thank the Simons Institute for its hospitality. PD also thanks The Ohio State University for hosting his visit.

AI Disclosure. We used AI tools extensively for proofreading, improving the exposition, finding and checking references, and developing and checking proofs. AI assistance was instrumental in identifying and correcting gaps in earlier arguments and in devising new arguments for the revised proofs, particularly for the border lower bound (Theorem 1.4) and the impossibility result without the constant-freeness assumption (Theorem 1.3). We verified all proofs, including those developed with AI assistance, and take full responsibility for the contents of this paper.

Appendix

A The Abhyankar–Jung Theorem

The Abhyankar–Jung theorem was first proved by Jung [Jun08] for convergent power series in two variables over $\mathbb{C}$ and later extended by Abhyankar [Abh55] to formal power series in several variables; see [PR12] for a constructive proof that relies in part on complex analysis. In this section, we present a detailed algebraic proof of the version used in this paper (Theorem 6.3), deriving it from standard results in commutative algebra, including those concerning complete local rings and discrete valuation rings. Our aim is to make this ingredient more accessible by providing a relatively elementary, purely algebraic treatment.

Throughout, assume $\mathbb{F}$ is an algebraically closed field of characteristic zero.[3]

First, we state a bivariate version of the Abhyankar–Jung theorem, where the coefficients are power series in two variables.

Theorem A.1 (Abhyankar–Jung, bivariate version): Let $P(z) \in \mathbb{F}[[\varepsilon, x]][z]$ be monic of degree $d \geq 1$. Suppose that $\mathrm{Disc}_z(P) = \varepsilon^\alpha U(\varepsilon, x)$ for some integer $\alpha \geq 0$ and some $U \in \mathbb{F}[[\varepsilon, x]]$ with $U(0,0) \neq 0$. Then there exist an integer $q \geq 1$ and power series $r_1, \ldots, r_d \in \mathbb{F}[[\varepsilon^{1/q}, x]]$ such that

$$P(z) = \prod_{i=1}^{d} (z - r_i).$$

We next derive the multivariate version (Theorem 6.3) from the bivariate case. For convenience, we restate Theorem 6.3 as follows. This is a special case of [PR12, Theorem 1.3 and Remark 1.4].

Theorem A.2 (Abhyankar–Jung, multivariate version): Let $\mathbf{u}$ be a tuple of variables and let $W \in \mathbb{F}[[\varepsilon, \mathbf{u}]][z]$ be monic of degree $d \geq 1$. Suppose that $\mathrm{Disc}_z(W) = \varepsilon^\alpha U(\varepsilon, \mathbf{u})$ for some integer $\alpha \geq 0$ and some $U \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$ with $U(0, \mathbf{0}) \neq 0$. Then there exist an integer $e \geq 1$ and power series

[3] The characteristic-zero assumption is used only to ensure that the relevant field extensions are tamely ramified. The same argument applies in sufficiently large positive characteristic.

$r_1, \dots, r_d \in \mathbb{F}[[\varepsilon, \mathbf{u}]]$ such that

$$W(\varepsilon^e, \mathbf{u}, z) = \prod_{i=1}^{d} (z - r_i(\varepsilon, \mathbf{u})).$$

To derive Theorem A.2 from Theorem A.1, we need the following form of Hensel's lemma to lift simple roots uniquely to power series roots.

Lemma A.3 (Hensel's Lemma [Eis95, Theorem 7.3]): Let K be a field, let $\mathbf{u} = (u_1, \dots, u_m)$, and let $P \in K[[\mathbf{u}]][z]$. Let $c \in K$ such that $P(\mathbf{0}, c) = 0$ and $(\partial_z P)(\mathbf{0}, c) \neq 0$. Then there is a unique $r \in K[[\mathbf{u}]]$ such that $r(\mathbf{0}) = c$ and $P(\mathbf{u}, r) = 0$.

Proof of Theorem A.2 assuming Theorem A.1: Write $\mathbf{u} = (u_1, \dots, u_m)$. To prove the theorem, we first factorize the polynomial at $\mathbf{u} = \mathbf{0}$ and lift its roots over $\mathbb{F}((\varepsilon))$. We then restrict $\mathbf{u}$ to lines to show that the lifted roots have no negative powers of ε.

Introduce a new variable x. View $W(\varepsilon, \mathbf{0}, z)$ as a polynomial over $\mathbb{F}[[\varepsilon, x]]$ whose coefficients are independent of x. Its discriminant is $\varepsilon^{\alpha} U(\varepsilon, \mathbf{0})$, with $U(0, \mathbf{0}) \neq 0$. By Theorem A.1, for some $e \geq 1$, the polynomial $W(\varepsilon^e, \mathbf{0}, z)$ splits over $\mathbb{F}[[\varepsilon, x]]$. Setting $x = 0$ gives $c_1, \dots, c_d \in \mathbb{F}[[\varepsilon]]$ with

$$W(\varepsilon^e, \mathbf{0}, z) = \prod_{i=1}^{d} (z - c_i(\varepsilon)).$$

Its discriminant is nonzero, so the c_i are pairwise distinct.

Let $K := \mathbb{F}((\varepsilon))$. For each $i \in [d]$, we have $(\partial_z W)(\varepsilon^e, \mathbf{0}, c_i) = \prod_{j \neq i} (c_i - c_j) \in K^{\times}$. By Hensel's lemma (Lemma A.3), applied over K, there is a unique $r_i \in K[[\mathbf{u}]]$ such that $r_i(\varepsilon, \mathbf{0}) = c_i(\varepsilon)$ and $W(\varepsilon^e, \mathbf{u}, r_i(\varepsilon, \mathbf{u})) = 0$. These roots are pairwise distinct. Since W is monic of degree d, we obtain

$$W(\varepsilon^e, \mathbf{u}, z) = \prod_{i=1}^{d} (z - r_i(\varepsilon, \mathbf{u})) \in K[[\mathbf{u}]][z]. \tag{43}$$

It remains to show that each r_i belongs to $\mathbb{F}[[\varepsilon, \mathbf{u}]]$.

Fix $\mathbf{a} \in \mathbb{F}^m$ and substitute $\mathbf{u} = x\mathbf{a}$. Since the discriminant of a monic polynomial is a polynomial in its coefficients, taking the discriminant commutes with these substitutions. So $\mathrm{Disc}_z(W(\varepsilon^e, x\mathbf{a}, z)) = \varepsilon^{e\alpha} U(\varepsilon^e, x\mathbf{a})$. As $U(0, \mathbf{0}) \neq 0$, by Theorem A.1, the polynomial $W(\varepsilon^e, x\mathbf{a}, z)$ has a factorization

$$W(\varepsilon^e, x\mathbf{a}, z) = \prod_{i=1}^{d} (z - s_i),$$

where $s_1, \dots, s_d \in \mathbb{F}[[\varepsilon^{1/h}, x]]$, and h is a positive integer that may depend on $\mathbf{a}$. Setting $x = 0$ gives

$$\prod_{i=1}^{d} (z - s_i|_{x=0}) = W(\varepsilon^e, \mathbf{0}, z) = \prod_{i=1}^{d} (z - c_i(\varepsilon)).$$

We may therefore relabel the s_i so that $s_i|_{x=0} = c_i(\varepsilon)$.

Let $L := \mathbb{F}((\varepsilon^{1/h}))$. In $L[[x]]$, both s_i and $r_i(\varepsilon, x\mathbf{a})$ are roots of $W(\varepsilon^e, x\mathbf{a}, z)$ with constant term $c_i(\varepsilon)$. Since $c_i(\varepsilon)$ is a simple root at $x = 0$, uniqueness in Lemma A.3, applied over L with

parameter x, gives $r_i(\varepsilon, x\mathbf{a}) = s_i$. Consequently, for every $\ell \geq 0$,

$$[x^\ell]\, r_i(\varepsilon, x\mathbf{a}) \in \mathbb{F}((\varepsilon)) \cap \mathbb{F}[[\varepsilon^{1/h}]] = \mathbb{F}[[\varepsilon]], \tag{44}$$

where the intersection is taken inside L. This holds for every $\mathbf{a} \in \mathbb{F}^m$.

For $i \in [d]$, write

$$r_i(\varepsilon, \mathbf{u}) = \sum_{\beta \in \mathbb{N}^m} b_{i,\beta}(\varepsilon)\mathbf{u}^\beta,$$

where $b_{i,\beta}(\varepsilon) \in \mathbb{F}((\varepsilon))$. Fix $i \in [d]$, $\ell \geq 0$, and $j < 0$. By (44), for every $\mathbf{a} \in \mathbb{F}^m$,

$$0 = [\varepsilon^j]\big([x^\ell]\, r_i(\varepsilon, x\mathbf{a})\big) = \sum_{|\beta|=\ell} ([\varepsilon^j] b_{i,\beta}(\varepsilon))\mathbf{a}^\beta.$$

The right-hand side is a polynomial in $\mathbf{a}$. Since it vanishes on $\mathbb{F}^m$ and $\mathbb{F}$ is infinite, all its coefficients are zero. Varying ℓ and $j < 0$ shows that $b_{i,\beta}(\varepsilon) \in \mathbb{F}[[\varepsilon]]$ for every $\beta \in \mathbb{N}^m$. Therefore,

$$r_i \in \mathbb{F}[[\varepsilon]][[\mathbf{u}]] = \mathbb{F}[[\varepsilon, \mathbf{u}]].$$

Combining this with (43) proves the theorem. □

Let B be an A-algebra. Recall that the *integral closure* of A in B is the set of all elements $b \in B$ that are integral over A, i.e., the set of elements of B that are roots of monic univariate polynomials over A. It is an A-subalgebra of B.

To prove Theorem A.1, we first show that it boils down to the following lemma.

Lemma A.4: Let $P(z) \in \mathbb{F}[[\varepsilon, x]][z]$ be as in Theorem A.1. There exists an integer $q > 0$ such that the following holds: Let $\eta = \varepsilon^{1/q}$. Let $A = \mathbb{F}[[\eta, x]]$, $F = \mathrm{Frac}(A)$, and $B = F[z]/(P(z))$. Let R be the integral closure of A in B. Then R is isomorphic to A^d as an A-algebra, where $d \in \mathbb{N}$.

Proof of Theorem A.1 assuming Lemma A.4: Let $\overline{z} = z + (P(z)) \in B$, which is also in R since $P(z)$ is a monic polynomial in z with coefficients in A. As B is spanned by powers of $\overline{z}$ over F, the natural map $R \otimes_A F \to B$ is an isomorphism. On the other hand, as $R \cong A^d$ by Lemma A.4, we have $R \otimes_A F \cong F^d$. So $d = \dim_F B = \deg(P)$.

Let φ be the A-algebra isomorphism from R to A^d. Suppose $\varphi(\overline{z}) = (r_1, \ldots, r_d)$, where $r_1, \ldots, r_d \in A$. As $\overline{z}$ is a root of P, the elements $r_1, \ldots, r_d$ are roots of $P(z)$ in $A \subseteq F$.

Next, we show that these roots are distinct. Every element in B is of the form $Q(\overline{z})$ for some polynomial $Q \in F[z]$, which is mapped to $(Q(r_1), \ldots, Q(r_d)) \in F^d$ under the F-algebra isomorphism $\varphi \otimes_A F \colon B \xrightarrow{\sim} F^d$. So $d = \dim_F B$ does not exceed the number of distinct roots among $r_1, \ldots, r_d$. This forces $r_1, \ldots, r_d$ to be distinct.

Finally, since $P(z)$ is monic of degree d and has at least d distinct roots in F, it must factorize completely over F as $P(z) = \prod_{i=1}^d (z - r_i)$, where $r_i \in A = \mathbb{F}[[\varepsilon^{1/q}, x]]$ for $i \in [d]$, as required. □

The rest of this section is devoted to proving Lemma A.4. To prove it, we first show that R in

Lemma A.4 is isomorphic to A^d as an A-module. We then show that the discriminant ideal of R over A is the unit ideal. Finally, we prove that R is isomorphic to A^d as an A-algebra.

A.1 Freeness of the Integral Closure as a Module

We begin with the integral closure of A in a finite separable field extension of $\mathrm{Frac}(A)$.

Lemma A.5: Let $A = \mathbb{F}[[\eta, x]]$, $F = \mathrm{Frac}(A)$, and let K be a finite separable[4] extension of F. Let R be the integral closure of A in K. Then R is a finite free A-module.

To prove Lemma A.5, we need the following lemma, which states that if A is an integrally closed domain, then the minimal polynomial of an integral element over A has coefficients in A.

Lemma A.6 ([Stacks, Tag 00H7]): Let $A \subseteq R$ be an inclusion of domains, with A integrally closed, and let $g \in R$ be integral over A. Then the minimal polynomial of g over $\mathrm{Frac}(A)$ has coefficients in A.

Next, we prove the following lemma.

Lemma A.7: Let A, F, K, and R be as in Lemma A.5. The module homomorphism $R/\eta R \to R/\eta R$ sending $a + \eta R$ to $xa + \eta R$ is injective.

Proof: Let $a \in R$ such that $xa \in \eta R$, i.e., $xa = \eta b$ for some $b \in R$. Let $u := a/\eta = b/x \in K$. We want to show that a is also in ηR, i.e., $u \in R$. As R is the integral closure of A in K, it suffices to show that u is integral over A.

Suppose the minimal polynomial of u over F is $f(y) = \sum_{i=0}^{n} c_i y^i \in F[y]$, where $c_n = 1$. The minimal polynomial of $a = \eta u$ over F is then $\eta^n f(y/\eta) = \sum_{i=0}^{n} \eta^{n-i} c_i y^i$. The ring A is an integrally closed domain [Stacks, Tag 0BI0]. By Lemma A.6, as $a \in R$ is integral over A, we have $\eta^{n-i} c_i \in A$ for $i = 0, 1, \ldots, n$. Applying the same argument to $b = xu \in R$ shows that $x^{n-i} c_i \in A$ for $i = 0, 1, \ldots, n$.

Fix $0 \le i < n$. Since both $\eta^{n-i} c_i$ and $x^{n-i} c_i$ belong to A, we have

$$x^{n-i}(\eta^{n-i} c_i) = \eta^{n-i}(x^{n-i} c_i) \in (\eta^{n-i}) \subseteq (\eta). \tag{45}$$

The ideal (η) is prime, since $A/(\eta) \cong \mathbb{F}[[x]]$ is a domain. As $x^{n-i} \notin (\eta)$, (45) implies $\eta^{n-i} c_i \in (\eta)$, and hence $\eta^{n-i-1} c_i \in A$. Dividing both sides of (45) by η, we obtain

$$x^{n-i}(\eta^{n-i-1} c_i) = \eta^{n-i-1}(x^{n-i} c_i).$$

If $n - i > 1$, the right-hand side still belongs to (η), so the same argument gives $\eta^{n-i-2} c_i \in A$. Continuing in this way until all $n - i$ factors of η have been removed, we conclude that $c_i \in A$.

Alternatively, we can prove $c_i \in A$ using the stronger fact that $A = \mathbb{F}[[\eta, x]]$ is a UFD [Stacks, Tag 032C, Tag 0AG0]. Indeed, write c_i as a fraction s/t with $s \in A$, $t \in A \setminus \{0\}$, and $\gcd(s, t) = 1$. As $\eta^{n-i} c_i \in A$ and $x^{n-i} c_i \in A$, the denominator t must divide both η^{n-i} and x^{n-i}. Since $\gcd(\eta^{n-i}, x^{n-i}) = 1$, we have $c_i \in A$.

[4] Separability is not necessary for this claim, but assuming it simplifies the proof that R is a finite A-module.

As $c_i \in A$ for $0 \leq i < n$ and $c_n = 1$, the minimal polynomial $f(y) = \sum_{i=0}^{n} c_i y^i$ of u over F is in $A[y]$. So u is integral over A, as desired. □

We also need the following fact as a corollary of Nakayama's lemma [Stacks, Tag 00DV]. It allows us to derive the freeness of a module from the freeness of its quotient module.

Lemma A.8 ([Stacks, Tag 00NS]): Let A be a Noetherian local ring with the maximal ideal $\mathfrak{m}$. Let M be a finite A-module. Let $\eta \in \mathfrak{m}$ such that multiplication by η on M is injective. Suppose $M/\eta M$ is free over $A/\eta A$. Then M is free over A.

We are now ready to prove Lemma A.5.

Proof of Lemma A.5: The ring $A = \mathbb{F}[[\eta, x]]$ is a Noetherian integrally closed domain [Stacks, Tag 0BI0]. By [Stacks, Tag 032L], R is a finite A-module. Consequently, $R/\eta R$ is a finite $A/\eta A$-module.

Next, we show that $R/\eta R$ is a *torsion-free* $A/\eta A$-module, meaning that if $as = 0$, where a is a non-zerodivisor of $A/\eta A$ and $s \in R/\eta R$, then $s = 0$. Since $A/\eta A \cong \mathbb{F}[[x]]$, every nonzero element of $A/\eta A$ is a power of x times a unit. By Lemma A.7, multiplication by x on $R/\eta R$ is injective. Since multiplication by a unit is invertible, multiplication by any nonzero element of $A/\eta A$ is also injective. So $R/\eta R$ is torsion-free over $A/\eta A$.

The ring $A/\eta A$ is a discrete valuation ring and hence a principal ideal domain (PID) [Stacks, Tag 00PD]. Every finite torsion-free module over a PID is finite free [Stacks, Tag 0AUW(3)]. So $R/\eta R$ is a finite free $A/\eta A$-module.

Finally, A is a Noetherian local ring with the maximal ideal $\mathfrak{m} = (\eta, x)$. As R is a domain containing A, multiplication by $\eta \in \mathfrak{m}$ on R is injective. By Lemma A.8, R is a finite free A-module. □

Lemma A.5 addresses the case of finite separable extensions. It easily extends to finite products of finite separable extensions via the following simple observation.

Lemma A.9: Let A be a ring and let $B_1, \ldots, B_n$ be A-algebras. For $i \in [n]$, let R_i be the integral closure of A in B_i. Then the integral closure of A in $B_1 \times \cdots \times B_n$ is $R_1 \times \cdots \times R_n$.

Proof: Let R be the integral closure of A in $B_1 \times \cdots \times B_n$. Suppose $a = (a_1, \ldots, a_n) \in R$. Then $f(a) = 0$ for some monic polynomial $f(z) \in A[z]$. So $f(a_i) = 0$ for $i \in [n]$, implying that $a_1, \ldots, a_n$ are integral over A. So $a \in R_1 \times \cdots \times R_n$.

Conversely, suppose $a = (a_1, \ldots, a_n) \in R_1 \times \cdots \times R_n$. Then each a_i is a root of some monic polynomial $f_i(z) \in A[z]$. Then a is a root of the monic polynomial $f(z) = \prod_{i=1}^{n} f_i(z) \in A[z]$, implying that $a \in R$. So $R = R_1 \times \cdots \times R_n$. □

The A-algebra $B = F[z]/(P(z))$ in Lemma A.4 is indeed a finite product of finite separable extensions of $\mathrm{Frac}(A)$. The following lemma records this observation.

Lemma A.10: Let $P(z)$ be as in Theorem A.1. Let $q > 0$ be an integer, and let $\eta = \varepsilon^{1/q}$. Let $A = \mathbb{F}[[\eta, x]]$, $F = \mathrm{Frac}(A)$, and $B = F[z]/(P(z))$. Then B is a finite product of finite separable

extensions of F.

Proof: As $\mathrm{Disc}_z(P) \neq 0$, the polynomial P is separable, meaning that the roots of P in the algebraic closure of F are distinct. So we may write $P = \prod_{i=1}^{k} P_i$, where the $P_i \in F[z]$ are distinct monic irreducible polynomials, each of which is separable. In particular, the P_i are pairwise coprime. By the Chinese remainder theorem [Stacks, Tag 00DT], we have an F-algebra isomorphism $B = F[z]/(P(z)) \cong \prod_{i=1}^{k} F[z]/(P_i(z))$. Each factor $F[z]/(P_i(z))$ is a finite separable field extension of F, since P_i is irreducible and separable. □

Combining Lemma A.5, Lemma A.9, and Lemma A.10, we obtain the following corollary.

Corollary A.11: Let $P(z)$ be as in Theorem A.1. Let $q > 0$ be an integer, and let $\eta = \varepsilon^{1/q}$. Let $A = \mathbb{F}[[\eta, x]]$, $F = \mathrm{Frac}(A)$, and $B = F[z]/(P(z))$. Let R be the integral closure of A in B. Then R is a finite free A-module.

A.2 Discriminants

Next, we define the discriminant of an A-algebra B that is finite free as an A-module. Geometrically, the homomorphism $A \to B$ corresponds to a finite morphism $\mathrm{Spec}(B) \to \mathrm{Spec}(A)$ from the spectrum of B to that of A [Stacks, Tag 00DY]. The discriminant vanishes precisely at the points in $\mathrm{Spec}(A)$ over which this morphism is ramified (the branch points). Our treatment, however, will be purely algebraic.

Let A be a ring, and let B be an A-algebra that is a finite free A-module of rank d. For $a \in B$, multiplication by a defines an A-linear map $\times a \colon B \to B$ sending $b \in B$ to ab. Choose an A-basis of B, and let M_a be the matrix representing $\times a$ in this basis. Define

$$\mathrm{Tr}_{B/A}(a) \coloneqq \mathrm{Tr}(M_a),$$

where the right-hand side is the usual matrix trace. This definition does not depend on the choice of the basis, since changing the basis corresponds to conjugating M_a by an invertible matrix over A, which does not change the trace.

For a d-tuple $\mathbf{c} = (c_1, \ldots, c_d)$ of elements of B, define the $d \times d$ matrix $G_{\mathbf{c}} := \left(\mathrm{Tr}_{B/A}(c_i c_j)\right)_{i,j \in [d]}$ over A, and define the *discriminant* of $\mathbf{c}$ to be

$$\mathrm{Disc}(\mathbf{c}) \coloneqq \det(G_{\mathbf{c}}) \in A.$$

Let $\mathbf{b} = (b_1, \ldots, b_d)$ be an (ordered) A-basis of B, and write $c_i = \sum_j m_{ij} b_j$, with $M = (m_{ij}) \in A^{d \times d}$. By A-linearity of the trace, we have $G_{\mathbf{c}} = M G_{\mathbf{b}} M^T$, and hence

$$\mathrm{Disc}(\mathbf{c}) = \det(M)^2 \mathrm{Disc}(\mathbf{b}). \tag{46}$$

If $\mathbf{c}$ is also an A-basis of B, then M is invertible over A, so $\det(M)$ is a unit of A. Consequently, the ideal generated by the discriminant of a basis is independent of the basis. Therefore, we can make the following definition:

Definition A.12 (Discriminant ideal): Let B be an A-algebra that is a finite free A-module. Its *discriminant ideal* (or simply *discriminant*) over A is the ideal

$$\Delta_{B/A} \coloneqq (\mathrm{Disc}(\mathbf{b})) \subseteq A$$

of A, where $\mathbf{b}$ is any A-basis of B.

We will use several properties of this definition. First, discriminants commute with base change. For a ring homomorphism $A \to A'$, an A-basis of B gives an A'-basis of $B \otimes_A A'$. The new multiplication matrices are obtained by applying $A \to A'$ to the entries of the old ones. Their traces and determinants are therefore obtained by applying the same map. So we have the following lemma.

Lemma A.13: Let B be an A-algebra that is a finite free A-module. Let $A \to A'$ be a ring homomorphism. Then $\Delta_{(B \otimes_A A')/A'} = \Delta_{B/A}\, A'$.

Second, this definition agrees with the usual polynomial discriminant when we use the power basis.

Lemma A.14 ([LQ15, Chapter III, Proposition 5.10]): Suppose A is a ring and $P(z) \in A[z]$ is monic of degree d. Let $B = A[z]/(P(z))$, and let $\overline{z} = z + (P(z)) \in B$. Then $\mathrm{Disc}(1, \overline{z}, \ldots, \overline{z}^{d-1}) = \mathrm{Disc}_z(P)$.

It is known that the discriminant of a finite separable extension is the unit ideal.

Lemma A.15 ([Stacks, Tag 0BIL]): Let K/F be a finite separable extension. Then $\Delta_{K/F} = F$.

Finally, we record the fact that the discriminant of a finite product equals the product of the discriminants of the factors.

Lemma A.16: Let A be a ring, and let $B_1, \ldots, B_n$ be A-algebras that are finite free A-modules. Let $B = \prod_{i=1}^n B_i$. Then $\Delta_{B/A} = \prod_{i=1}^n \Delta_{B_i/A}$.

Proof: Choose an ordered A-basis $\mathbf{b}_i$ of each B_i. Let $\mathbf{b}$ be the ordered A-basis of B obtained by concatenating these bases, where each $b \in \mathbf{b}_i$ is identified with the element of B whose i-th component is b and whose other components are zero.

For $a = (a_1, \ldots, a_n) \in B$, the matrix M_a with respect to $\mathbf{b}$ is block diagonal, with the i-th diagonal block equal to M_{a_i} with respect to $\mathbf{b}_i$. Consequently, $G_{\mathbf{b}}$ is also block diagonal, with diagonal blocks $G_{\mathbf{b}_1}, \ldots, G_{\mathbf{b}_n}$. The claim follows by taking determinants. □

A.3 Removing Ramification by Adjoining a Root

We start with the following lemma characterizing finite extensions of $k((\varepsilon))$, where k is a field of characteristic zero.

Lemma A.17: Let k be a field of characteristic zero, and let $F = k((\varepsilon))$. Let L/F be a finite extension. Then there exist a finite extension $\mathbb{K}/k$ contained in L, an integer $e \geq 1$, and elements $c \in \mathbb{K}^\times$ and $\tau \in L$ such that $L = \mathbb{K}((\tau))$, $\tau^e = c\varepsilon$, and $[L : \mathbb{K}((\varepsilon))] = e$.

Proof: The proof uses the classical theory of complete discrete valuation rings. See [FV02; Neu99] for a general treatment.

Let $\mathcal{O}_F = k[[\varepsilon]]$, equipped with the normalized discrete valuation $v_F\colon F \to \mathbb{Z} \cup \{\infty\}$, which sends ε to 1. For every finite extension E/F, the integral closure $\mathcal{O}_E$ of $\mathcal{O}_F$ in E is also a complete discrete valuation ring [FV02, Chapter II, Theorem 2.5 and Lemma 3.1]. Its normalized valuation v_E satisfies $v_E|_F = e_{E/F} v_F$ for an integer $e_{E/F} \geq 1$, called the *ramification index* of E/F. We say E/F is *unramified* if $e_{E/F} = 1$.

By [FV02, Chapter II, Proposition 3.4 and its corollary], there is a unique maximal unramified subextension M/F of L/F. Moreover, L/M is *totally ramified*, meaning that $e_{L/M} = [L : M]$, where the ramification index $e_{L/M}$ is the unique integer satisfying $v_L|_M = e_{L/M} v_M$. We let

$$e := e_{L/F} = e_{L/M} e_{M/F} = e_{L/M} = [L : M],$$

where we use the fact that $e_{M/F} = 1$ and the multiplicativity of ramification indices.

We now identify M as $\mathbb{K}((\varepsilon))$ for some finite extension $\mathbb{K}/k$ contained in L. By [FV02, Chapter II, Proposition 5.2 and its proof], there is a subfield $\mathbb{K} \subseteq \mathcal{O}_M$ containing k such that the natural quotient map $\mathcal{O}_M \to \kappa_M$ maps $\mathbb{K}$ isomorphically to the residue field κ_M. To ensure $k \subseteq \mathbb{K}$, apply the proof there to a maximal subfield of $\mathcal{O}_M$ containing k. Since the residue field κ_M of M is finite over k, so is $\mathbb{K}$. Furthermore, $v_M(\varepsilon) = 1$ because M/F is unramified. Thus ε is a uniformizer of $\mathcal{O}_M$. Then by [FV02, Chapter I, Proposition 5.2 and its corollary], we have $\mathcal{O}_M = \mathbb{K}[[\varepsilon]]$ and $M = \mathbb{K}((\varepsilon))$. In particular, $[L : \mathbb{K}((\varepsilon))] = [L : M] = e$.

Next, since $\mathrm{char}(\kappa_M) = \mathrm{char}(\mathbb{K}) = 0$, L/M is tamely ramified. By [FV02, Chapter II, Proposition 3.5(1)] there exists a uniformizer θ of $\mathcal{O}_L$ and a uniformizer π of $\mathcal{O}_M$ such that $L = M(\theta)$ and $\theta^e = \pi$. As $\mathcal{O}_M = \mathbb{K}[[\varepsilon]]$, we know π has the form $\varepsilon \cdot u(\varepsilon)$, where $u(\varepsilon) \in \mathbb{K}[[\varepsilon]]^\times$.

Let $c = u(0) \in \mathbb{K}^\times$. By [FV02, Chapter I, Corollary 5.5], there exists $t(\varepsilon) \in 1 + \varepsilon\mathbb{K}[[\varepsilon]]$ such that $t(\varepsilon)^e = u(\varepsilon)/c$. Define $\tau = \theta/t(\varepsilon)$. Then $\tau^e = \theta^e/t(\varepsilon)^e = c\pi/u(\varepsilon) = c\varepsilon$ and $M(\tau) = M(\theta) = L$.

Since $t(\varepsilon)$ is a unit of $\mathcal{O}_L$, the element $\tau = \theta/t(\varepsilon)$ is a uniformizer of $\mathcal{O}_L$. By completeness, we may regard $\mathbb{K}((\tau))$ as a subfield of L. Since $\varepsilon = \tau^e/c$, grouping Laurent series by their exponents modulo e gives

$$\mathbb{K}((\tau)) = \bigoplus_{j=0}^{e-1} \tau^j \mathbb{K}((\varepsilon)) = M(\tau) = L,$$

as desired. □

The finite morphism associated with the extension $\mathcal{O}_F \subseteq \mathcal{O}_L$ may be ramified. The following lemma shows that this ramification can be removed by adjoining a suitable root of ε to the base field and then taking the integral closure of the new power series ring. It will be used in the next subsection.

Lemma A.18: With the notation of Lemma A.17, let q be a positive multiple of e and let $\eta = \varepsilon^{1/q}$. Let $B = L \otimes_{k((\varepsilon))} k((\eta))$. Let R be the integral closure of $k[[\eta]]$ in B. Then R is a finite free module over $k[[\eta]]$ and $\Delta_{R/k[[\eta]]}$ is the unit ideal.

Proof: By Lemma A.17, we have $L = \mathbb{K}((\varepsilon))(\tau)$, $\tau^e = c\varepsilon$, and $[L : \mathbb{K}((\varepsilon))] = e$. So the minimal polynomial of τ over $\mathbb{K}((\varepsilon))$ is $z^e - c\varepsilon$, and the map $z \mapsto \tau$ induces a $\mathbb{K}((\varepsilon))$-algebra isomorphism

$\mathbb{K}((\varepsilon))[z]/(z^e - c\varepsilon) \cong L$. It follows that

$$\begin{aligned} B = L \otimes_{k((\varepsilon))} k((\eta)) &\cong \mathbb{K}((\varepsilon))[z]/(z^e - c\varepsilon) \otimes_{k((\varepsilon))} k((\eta)) \\ &\cong \mathbb{K}((\eta))[z]/(z^e - c\eta^q) \cong \mathbb{K}((\eta))[y]/(y^e - c) \end{aligned}$$

as $k((\eta))$-algebras, where in the last step we make the substitution $y = z/\eta^{q/e}$ and observe that $\eta^{q/e} \in \mathbb{K}((\eta))^\times$.

As $c \neq 0$ and the characteristic is zero, $y^e - c$ is a separable polynomial over $\mathbb{K}$. By the Chinese remainder theorem, $\mathbb{K}[y]/(y^e - c)$ is isomorphic to a finite product of finite separable extensions E_i of $\mathbb{K}$ as a $\mathbb{K}$-algebra:

$$\mathbb{K}[y]/(y^e - c) \cong \prod_{i=1}^{k} E_i.$$

By extending the coefficient field $\mathbb{K}$ to $\mathbb{K}((\eta))$, we see that $B \cong \mathbb{K}((\eta))[y]/(y^e - c) \cong \prod_{i=1}^{k} E_i((\eta))$ as $k((\eta))$-algebras.

Let R_i be the integral closure of $k[[\eta]]$ in $E_i((\eta))$ for $i \in [k]$. By Lemma A.9 and Lemma A.16, it suffices to show for each $i \in [k]$ that R_i is a finite free $k[[\eta]]$-module and that $\Delta_{R_i/k[[\eta]]} = k[[\eta]]$.

Fix $i \in [k]$. Since $E_i/\mathbb{K}$ and $\mathbb{K}/k$ are finite, so is E_i/k. Consequently, $E_i[[\eta]] \cong E_i \otimes_k k[[\eta]]$ is a finite free $k[[\eta]]$-module. By [Stacks, Tag 00GK], $E_i[[\eta]]$ is integral over $k[[\eta]]$, so $E_i[[\eta]] \subseteq R_i$. Conversely, every element of R_i is integral over $k[[\eta]]$ and hence also over $E_i[[\eta]]$. Since $E_i[[\eta]]$ is integrally closed in its fraction field $E_i((\eta))$ [Stacks, Tag 0BI0], it follows that $R_i \subseteq E_i[[\eta]]$. Thus $R_i = E_i[[\eta]]$, which is finite free over $k[[\eta]]$.

Finally, since the characteristic is zero, E_i/k is separable. Hence $\Delta_{E_i/k} = k$ by Lemma A.15. Using the identification $R_i = E_i[[\eta]] \cong E_i \otimes_k k[[\eta]]$ and Lemma A.13, we obtain

$$\Delta_{R_i/k[[\eta]]} = \Delta_{E_i/k}\, k[[\eta]] = k[[\eta]],$$

as desired. □

Remark A.19: Lemma A.18 can be viewed as a special case of *Abhyankar's lemma*, which shows that one can kill tame ramification by taking an extension of a base field; see [Stacks, Tag 0EXT]. In our setting, the condition $e \mid q$ ensures that every field factor of B is unramified over $k((\eta))$. The integral closure R is the product of the corresponding valuation rings, so it is finite étale over $k[[\eta]]$, or equivalently, finite free with unit discriminant [Stacks, Tag 0BJF]. ◁

A.4 The Discriminant of the Integral Closure

In this subsection, we show that for suitable $q > 0$, the discriminant ideal of the integral closure R in Lemma A.4 is the unit ideal.

First, we need the following technical lemma.

Lemma A.20: Let D be a discrete valuation ring with fraction field F and completion $\widehat{D}$, and let $\widehat{F} = \mathrm{Frac}(\widehat{D})$. Let B be a finite product of finite separable extensions of F, and let R be the integral

closure of D in B. The inclusions $R \subseteq B$ and $\widehat{D} \subseteq \widehat{F}$ induce a homomorphism $R \otimes_D \widehat{D} \to B \otimes_F \widehat{F}$. This homomorphism is injective, and its image is the integral closure of $\widehat{D}$ in $B \otimes_F \widehat{F}$.

Proof: By Lemma A.9, it suffices to prove the statement when B is a finite separable field extension of F. Let π be a uniformizer of D, and let $\widehat{R} := R \otimes_D \widehat{D}$.

We first identify $B \otimes_F \widehat{F}$ with $\widehat{R}_\pi$. Every element of B becomes integral over D after multiplication by a sufficiently large power of π, by clearing denominators in a monic polynomial equation over F. So $R_\pi = B$. Using $D_\pi = F$ and $\widehat{D}_\pi = \widehat{F}$, we have

$$\widehat{R}_\pi = (R \otimes_D \widehat{D})_\pi \cong R \otimes_D \widehat{F} \cong (R \otimes_D F) \otimes_F \widehat{F} = (R \otimes_D D_\pi) \otimes_F \widehat{F} \cong R_\pi \otimes_F \widehat{F} \cong B \otimes_F \widehat{F}.$$

We next describe $\widehat{R}$ as a product of complete DVRs. By [Stacks, Tag 09E8], R is finite over D, has finitely many maximal ideals $\mathfrak{m}_1, \ldots, \mathfrak{m}_r$, and each $R_{\mathfrak{m}_i}$ is a DVR. By [Stacks, Tag 07N9], we have

$$\widehat{R} \cong \prod_{i=1}^{r} \widehat{R_{\mathfrak{m}_i}},$$

where each $\widehat{R_{\mathfrak{m}_i}}$ is the completion of $R_{\mathfrak{m}_i}$ and is again a DVR [Stacks, Tag 0AP1]. The image of π in each $\widehat{R_{\mathfrak{m}_i}}$ is nonzero and belongs to its maximal ideal, so inverting π gives Frac $\left(\widehat{R_{\mathfrak{m}_i}}\right)$. Consequently, the following localization map is an injection:

$$\widehat{R} \cong \prod_{i=1}^{r} \widehat{R_{\mathfrak{m}_i}} \hookrightarrow \widehat{R}_\pi \cong \prod_{i=1}^{r} \text{Frac}\left(\widehat{R_{\mathfrak{m}_i}}\right).$$

Under the identifications $\widehat{R} = R \otimes_D \widehat{D}$ and $\widehat{R}_\pi \cong B \otimes_F \widehat{F}$, this is the injective map in the statement.

It remains to show that the elements of $\widehat{R}_\pi$ integral over $\widehat{D}$ are exactly those in $\widehat{R}$. Since R is finite over D, $\widehat{R} = R \otimes_D \widehat{D}$ is finite over $\widehat{D}$. For each $i \in [r]$, the projection $\widehat{R} \to \widehat{R_{\mathfrak{m}_i}}$ is surjective, so $\widehat{R_{\mathfrak{m}_i}}$ is also finite, and hence integral, over $\widehat{D}$. Moreover, $\widehat{R_{\mathfrak{m}_i}}$ is a DVR and hence is integrally closed in its fraction field. Thus $\widehat{R_{\mathfrak{m}_i}}$ is the integral closure of $\widehat{D}$ in Frac $\left(\widehat{R_{\mathfrak{m}_i}}\right)$. By Lemma A.9, $\widehat{R} \cong \prod_i \widehat{R_{\mathfrak{m}_i}}$ is therefore the integral closure of $\widehat{D}$ in $\prod_{i=1}^r$ Frac $\left(\widehat{R_{\mathfrak{m}_i}}\right) \cong B \otimes_F \widehat{F}$. □

The following lemma is the main result of this subsection.

Lemma A.21: Let $P(z) \in \mathbb{F}[[\varepsilon, x]][z]$ be as in Theorem A.1. There exists an integer $q > 0$ such that the following holds: Let $\eta = \varepsilon^{1/q}$. Let $A = \mathbb{F}[[\eta, x]]$, $F = \text{Frac}(A)$, and $B = F[z]/(P(z))$. Let R be the integral closure of A in B, which is a finite free A-module by Corollary A.11. Then $\Delta_{R/A} = A$.

Proof: Assume q is given, whose value will be determined later. The ring $A = \mathbb{F}[[\eta, x]]$ is a UFD [Stacks, Tag 032C and Tag 0AG0], and η is a prime element of A since $\mathbb{F}[[\eta, x]]/(\eta) \cong \mathbb{F}[[x]]$ is a domain. We will prove $\Delta_{R/A} = A$ by first showing that every prime divisor of $\Delta_{R/A}$ is associated with η. Then we show that for suitable q, even η does not divide $\Delta_{R/A}$, concluding the proof.

By Corollary A.11, R is a finite free A-module of rank $d = \deg(P)$, since $R \otimes_A F \cong B =$

$F[z]/(P(z))$. Choose an A-basis $\mathbf{b} = (b_1, \ldots, b_d)$ of R. Then $\Delta_{R/A} = (\mathrm{Disc}(\mathbf{b}))$. On the other hand, as $z + (P(z)) \in B$ is in R, we have $\mathrm{Disc}_z(P) = \det(M)^2 \mathrm{Disc}(\mathbf{b})$ for some $d \times d$ matrix M over A by Lemma A.14 and (46). By assumption, $\mathrm{Disc}_z(P) = \varepsilon^\alpha U(\varepsilon, x) = \eta^{q\alpha} U(\eta^q, x)$, where $U(\eta^q, x)$ is a unit of $A = \mathbb{F}[[\eta, x]]$ since $U(0,0) \neq 0$. So every prime divisor of $\mathrm{Disc}_z(P)$ is associated with η. As $\mathrm{Disc}_z(P) = \det(M)^2 \mathrm{Disc}(\mathbf{b})$, every prime divisor of $\Delta_{R/A}$ is also associated with η.

It remains to show that η does not divide $\mathrm{Disc}(\mathbf{b})$. To achieve this, we treat x as part of the coefficient field and let $k = \mathbb{F}((x))$. The DVR $A_{(\eta)}$ has fraction field F. For every $n \geq 1$, localization gives a natural isomorphism $A_{(\eta)}/\eta^n A_{(\eta)} \cong k[\eta]/(\eta^n)$. Taking inverse limits, we obtain $\widehat{A_{(\eta)}} \cong k[[\eta]]$, whose fraction field is $k((\eta))$.

Integral closure commutes with localization [Stacks, Tag 0307]. Hence $R \otimes_A A_{(\eta)}$ is the integral closure of $A_{(\eta)}$ in B. Applying Lemma A.20 with $D = A_{(\eta)}$, we conclude that the integral closure of $k[[\eta]] = \widehat{A_{(\eta)}}$ in $B \otimes_F k((\eta))$ is

$$(R \otimes_A A_{(\eta)}) \otimes_{A_{(\eta)}} \widehat{A_{(\eta)}} \cong R \otimes_A k[[\eta]].$$

We now use Lemma A.18 to show that this integral closure has unit discriminant. As $P(z)$ is a separable polynomial, by the Chinese remainder theorem [Stacks, Tag 00DT],

$$k((\varepsilon))[z]/(P(z)) \cong \prod_{i=1}^{s} L_i,$$

where each L_i is a finite separable extension of $k((\varepsilon))$. Then

$$B \otimes_F k((\eta)) \cong k((\eta))[z]/(P(z)) \cong \prod_{i=1}^{s} \left(L_i \otimes_{k((\varepsilon))} k((\eta)) \right). \tag{47}$$

For each $i \in [s]$, let e_i be the positive integer obtained by applying Lemma A.17 to the extension $L_i/k((\varepsilon))$. Choose $q = \mathrm{lcm}(e_1, \ldots, e_s)$. By Lemma A.18, the integral closure of $k[[\eta]]$ in each factor $L_i \otimes_{k((\varepsilon))} k((\eta))$ has unit discriminant. By Lemma A.9, Lemma A.16, and (47), the integral closure of $k[[\eta]]$ in $B \otimes_F k((\eta))$ also has unit discriminant. But we have shown above that this integral closure is just $R \otimes_A k[[\eta]]$. So

$$\Delta_{R/A}\, k[[\eta]] = \Delta_{(R \otimes_A k[[\eta]])/k[[\eta]]} = k[[\eta]],$$

where the first equality holds by Lemma A.13. As $\mathrm{Disc}(\mathbf{b})$ generates $\Delta_{R/A}$ in A, we see that it is a unit in $k[[\eta]] = \mathbb{F}((x))[[\eta]]$. In particular, $\mathrm{Disc}(\mathbf{b})$ cannot be divisible by η in $A = \mathbb{F}[[\eta, x]]$. Therefore, $\Delta_{R/A} = A$. □

Remark A.22: The UFD property of A is not necessary and can be weakened to integral closedness: The first part of the proof shows that $\mathrm{Disc}(\mathbf{b})^{-1} \in F$ is in $A_\mathfrak{p}$ for any height-one prime ideal $\mathfrak{p}$ of A other than (η), and the second part shows that it is also in $A_{(\eta)}$. As A is integrally closed, we have by [Stacks, Tag 031T(2)] that $\mathrm{Disc}(\mathbf{b})^{-1} \in A$. So $\mathrm{Disc}(\mathbf{b})$ is a unit of A. ◁

A.5 Finishing the Proof

To finish the proof, we use the following fact: if A is a complete Noetherian local ring, then any A-algebra that is finite as an A-module is isomorphic to a finite product of complete Noetherian local rings [Stacks, Tag 0325]. In particular, if such an algebra is a domain, then it is local, since a product of two or more nonzero rings has zero divisors.

Proof of Lemma A.4: Choose q as in Lemma A.21, so that $\Delta_{R/A} = A$. By Lemma A.10, write $B \cong \prod_{i=1}^{r} K_i$, where each K_i is a finite separable field extension of F. Let R_i be the integral closure of A in K_i. By Lemma A.9 and Lemma A.5, $R \cong \prod_{i=1}^{r} R_i$, where each R_i is a domain and a finite free A-module. By Lemma A.16, $A = \Delta_{R/A} = \prod_{i=1}^{r} \Delta_{R_i/A}$. Hence $\Delta_{R_i/A} = A$ for every i. It remains to show that each inclusion $A \hookrightarrow R_i$ is an isomorphism.

Fix i and let $S = R_i$. The standard fact above shows that S is a local ring. Let $\mathfrak{m} = (\eta, x)$ be the maximal ideal of A, and let $\overline{S} = S/\mathfrak{m}S$. Since S is integral over A, its maximal ideal $\mathfrak{m}_S$ contracts to $\mathfrak{m}$, i.e., $\mathfrak{m}_S \cap A = \mathfrak{m}$. So $\overline{S}$ is a nonzero finite-dimensional local $\mathbb{F}$-algebra. Since $\overline{S} = S/\mathfrak{m}S \cong S \otimes_A \mathbb{F}$ and $\Delta_{S/A} = A$, by Lemma A.13, we have $\Delta_{\overline{S}/\mathbb{F}} = \mathbb{F}$. Equivalently, the trace pairing $\overline{S} \times \overline{S} \to \mathbb{F}$ sending (u, v) to $\mathrm{Tr}_{\overline{S}/\mathbb{F}}(uv)$ is a nondegenerate bilinear map.

We claim that $\overline{S}$ is reduced, meaning that it has no nonzero nilpotent element. We prove this following the proof of [Stacks, Tag 0BJF]: If $u \in \overline{S}$ is nilpotent, then for every $v \in \overline{S}$, the matrix M_{uv} that represents the linear map $w \mapsto uvw$ is nilpotent and hence has trace zero. So $\mathrm{Tr}_{\overline{S}/\mathbb{F}}(uv) = 0$ for every $v \in \overline{S}$. Nondegeneracy forces $u = 0$, proving the claim.

As $\overline{S}$ is finite-dimensional over $\mathbb{F}$, it is *Artinian*, meaning that every descending chain of ideals of $\overline{S}$ stabilizes [Stacks, Tag 00J4]. The maximal ideal of an Artinian local ring is nilpotent [Stacks, Tag 00J8]. Since $\overline{S}$ is reduced, its maximal ideal must therefore be zero, so $\overline{S}$ is a field. And since $\overline{S}$ is finite-dimensional over the algebraically closed field $\mathbb{F}$, the natural map $\mathbb{F} \to \overline{S}$ is an isomorphism.

In particular, $S/\mathfrak{m}S$ is generated by 1 as an $A/\mathfrak{m}$-vector space. Nakayama's lemma [Stacks, Tag 00DV(8)] then implies that $S = A \cdot 1$. Therefore, the inclusion $A \hookrightarrow S = R_i$ is an A-algebra isomorphism. Applying this to every i gives $R \cong \prod_{i=1}^{r} R_i \cong A^r$ as A-algebras, proving Lemma A.4. □

Remark A.23: The condition $\Delta_{R/A} = A$ means that R is finite étale over A [Stacks, Tag 0BJF]. For a complete (or, more generally, henselian) local ring A with maximal ideal $\mathfrak{m}$ and residue field $\mathbb{F}$, reduction modulo $\mathfrak{m}$ gives an equivalence between the categories of finite étale A-algebras and finite étale $\mathbb{F}$-algebras; see [Stacks, Tag 04GK, Tag 04GM] or [Mil80, Chapter I, Propositions 4.4 and 4.5]. In our setting, $\mathbb{F}$ is algebraically closed, so finite étale $\mathbb{F}$-algebras are just finite products of $\mathbb{F}$. The corresponding A-algebras are thus finite products of A, giving $R \cong A^d$ as A-algebras. Therefore, the proof above amounts to working out a special case of this equivalence. ◁

References

[Abh55] Shreeram Abhyankar. "On the Ramification of Algebraic Functions". In: *American Journal of Mathematics* 77.3 (1955), pp. 575–592. DOI: 10.2307/2372643.

[AF22] Robert Andrews and Michael A. Forbes. "Ideals, Determinants, and Straightening: Proving and Using Lower Bounds for Polynomial Ideals". In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2022)*. 2022, pp. 389–402. DOI: 10.1145/3519935.3520025. arXiv: 2112.00792 [cs.CC].

[AM69] M. F. Atiyah and I. G. Macdonald. *Introduction to Commutative Algebra*. Addison-Wesley Publishing Company, 1969. DOI: 10.1201/9780429493638.

[AV08] Manindra Agrawal and V. Vinay. "Arithmetic Circuits: A Chasm at Depth Four". In: *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2008)*. 2008, pp. 67–75. DOI: 10.1109/FOCS.2008.32.

[BCS97] Peter Bürgisser, Michael Clausen, and Mohammad Amin Shokrollahi. *Algebraic Complexity Theory*. Vol. 315. Grundlehren der mathematischen Wissenschaften. Springer, 1997. DOI: 10.1007/978-3-662-03338-8.

[BDS26] Meghana Bhat, Pranjal Dutta, and Devansh Shringi. *Demystifying Border Depth-3 Circuits via Differential Equations*. Report TR26-245. Electronic Colloquium on Computational Complexity (ECCC), Oct. 11, 2026. URL: https://eccc.weizmann.ac.il/report/2026/245/.

[Ber+18] Alessandra Bernardi, Enrico Carlini, Maria Virginia Catalisano, Alessandro Gimigliano, and Alessandro Oneto. "The Hitchhiker Guide to: Secant Varieties and Tensor Decomposition". In: *Mathematics* 6.12, 314 (2018). DOI: 10.3390/math6120314. arXiv: 1812.10267 [math.AG].

[Bin+79] Dario Bini, Milvio Capovani, Francesco Romani, and Grazia Lotti. "$O(n^{2.7799})$ Complexity for $n \times n$ Approximate Matrix Multiplication". In: *Information Processing Letters* 8.5 (1979), pp. 234–235. DOI: 10.1016/0020-0190(79)90113-3.

[Bin80] D. Bini. "Relations between Exact and Approximate Bilinear Algorithms. Applications". In: *CALCOLO* 17.1 (1980), pp. 87–97. DOI: 10.1007/BF02575865.

[BMS13] M. Beecken, J. Mittmann, and N. Saxena. "Algebraic Independence and Blackbox Identity Testing". In: *Information and Computation* 222 (2013), pp. 2–19. DOI: 10.1016/j.ic.2012.10.004.

[Bür04] Peter Bürgisser. "The Complexity of Factors of Multivariate Polynomials". In: *Foundations of Computational Mathematics* 4.4 (2004), pp. 369–396. DOI: 10.1007/s10208-002-0059-5.

[Che+18] Mahdi Cheraghchi, Elena Grigorescu, Brendan Juba, Karl Wimmer, and Ning Xie. "$AC^0 \circ MOD_2$ Lower Bounds for the Boolean Inner Product". In: *Journal of Computer and System Sciences* 97 (2018), pp. 45–59. DOI: 10.1016/j.jcss.2018.04.006.

[CKS18] Chi-Ning Chou, Mrinal Kumar, and Noam Solomon. "Hardness vs Randomness for Bounded Depth Arithmetic Circuits". In: *33rd Computational Complexity Conference (CCC 2018)*. Vol. 102. Leibniz International Proceedings in Informatics (LIPIcs). 2018, 13:1–13:17. DOI: 10.4230/LIPIcs.CCC.2018.13.

[CLO25] David A. Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. 5th ed. Undergraduate Texts in Mathematics. Springer Nature Switzerland, 2025. DOI: 10.1007/978-3-031-91841-4.

[DDS21] Pranjal Dutta, Prateek Dwivedi, and Nitin Saxena. "Deterministic Identity Testing Paradigms for Bounded Top-Fanin Depth-4 Circuits". In: *36th Computational Complexity Conference (CCC 2021)*. Vol. 200. Leibniz International Proceedings in Informatics

(LIPIcs). 2021, 11:1–11:27. DOI: 10.4230/LIPIcs.CCC.2021.11. arXiv: 2304.11325 [cs.CC].

[DDS22] Pranjal Dutta, Prateek Dwivedi, and Nitin Saxena. "Demystifying the Border of Depth-3 Algebraic Circuits". In: *Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2021)*. 2022, pp. 92–103. DOI: 10.1109/FOCS52979.2021.00018.

[Del70] Pierre Deligne. *Équations différentielles à points singuliers réguliers*. Vol. 163. Lecture Notes in Mathematics. Springer, 1970. DOI: 10.1007/BFb0061194.

[DL25] Pranjal Dutta and Vladimir Lysikov. *Recent Advances in Debordering Methods*. Oct. 15, 2025. DOI: 10.48550/arXiv.2510.13049. arXiv: 2510.13049 [cs.CC]. Pre-published.

[DS07] Zeev Dvir and Amir Shpilka. "Locally Decodable Codes with Two Queries and Polynomial Identity Testing for Depth 3 Circuits". In: *SIAM Journal on Computing* 36.5 (2007), pp. 1404–1434. DOI: 10.1137/05063605X.

[DS26] Pranjal Dutta and Nitin Saxena. "Separated Borders: Exponential-Gap Fanin-Hierarchy Theorem for Approximative Depth-3 Circuits". In: *Journal of the ACM* 73.5 (2026), 31:1–31:37. DOI: 10.1145/3838180. Personal updated copy: link.

[Eis95] David Eisenbud. *Commutative Algebra: With a View Toward Algebraic Geometry*. Vol. 150. Graduate Texts in Mathematics. Springer, 1995. DOI: 10.1007/978-1-4612-5350-1.

[For14] Timothy J. Ford. "The Group of Units on an Affine Variety". In: *Journal of Algebra and Its Applications* 13.8, 1450065 (2014). DOI: 10.1142/S0219498814500650.

[For15] Michael A. Forbes. "Deterministic Divisibility Testing via Shifted Partial Derivatives". In: *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2015)*. 2015, pp. 451–465. DOI: 10.1109/FOCS.2015.35. Full version available at: https://www.ias.edu/sites/default/files/math/csdm/14-15/Forbes2015.pdf.

[Fou+15] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. "Lower Bounds for Depth-4 Formulas Computing Iterated Matrix Multiplication". In: *SIAM Journal on Computing* 44.5 (2015), pp. 1173–1201. DOI: 10.1137/140990280.

[FV02] Ivan B. Fesenko and Sergei V. Vostokov. *Local Fields and Their Extensions*. 2nd ed. Vol. 121. Translations of Mathematical Monographs. American Mathematical Society, 2002. DOI: 10.1090/mmono/121. Available at: https://ivanfesenko.org/wp-content/uploads/2021/10/vol.pdf.

[Gar+26] Abhibhav Garg, Rafael Oliveira, Akash Kumar Sengupta, Nir Shalmon, and Amir Shpilka. "Rank Bounds and Polynomial-Time PIT for $\Sigma^k\Pi\Sigma\Pi^2$ Circuits". In: *41st Computational Complexity Conference (CCC 2026)*. Vol. 383. Leibniz International Proceedings in Informatics (LIPIcs). 2026, 17:1–17:18. DOI: 10.4230/LIPIcs.CCC.2026.17.

[Ges16] Fulvio Gesmundo. "Geometric Aspects of Iterated Matrix Multiplication". In: *Journal of Algebra* 461 (2016), pp. 42–64. DOI: 10.1016/j.jalgebra.2016.04.028.

[GMQ16] Joshua A. Grochow, Ketan D. Mulmuley, and Youming Qiao. "Boundaries of VP and VNP". In: *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*. Vol. 55. Leibniz International Proceedings in Informatics (LIPIcs). 2016, 34:1–34:14. DOI: 10.4230/LIPIcs.ICALP.2016.34.

[GOS25] Abhibhav Garg, Rafael Oliveira, and Akash Kumar Sengupta. "Rank Bounds and PIT for Depth-4 Circuits with Top Fan-in 3 and Constant Bottom Fan-in via a Non-Linear Edelstein–Kelly Theorem". In: *Proceedings of the 66th Annual IEEE Symposium*

on Foundations of Computer Science (FOCS 2025). 2025, pp. 2143–2169. DOI: 10.1109/FOCS63196.2025.00114.

[Guo26] Zeyu Guo. *A Note on Deterministic PIT for $\Sigma^{[3]}\Pi\Sigma\Pi^{[\delta]}$ Circuits*. Report TR26-168. Electronic Colloquium on Computational Complexity (ECCC), Sept. 6, 2026. URL: https://eccc.weizmann.ac.il/report/2026/168/.

[Gup+14] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. "Approaching the Chasm at Depth Four". In: *Journal of the ACM* 61.6 (2014), 33:1–33:16. DOI: 10.1145/2629541.

[Gup14] Ankit Gupta. *Algebraic Geometric Techniques for Depth-4 PIT & Sylvester–Gallai Conjectures for Varieties*. Report TR14-130. Electronic Colloquium on Computational Complexity (ECCC), Oct. 17, 2014. URL: https://eccc.weizmann.ac.il/report/2014/130/.

[GW25] Zeyu Guo and Siki Wang. *Deterministic Depth-4 PIT and Normalization*. 2025. DOI: 10.48550/arXiv.2504.15143. arXiv: 2504.15143 [cs.CC].

[Had96] J. Hadamard. "Mémoire sur l'élimination". In: *Acta Mathematica* 20 (1896), pp. 201–238. DOI: 10.1007/BF02418033.

[Har77] Robin Hartshorne. *Algebraic Geometry*. Vol. 52. Graduate Texts in Mathematics. Springer, 1977. DOI: 10.1007/978-1-4757-3849-0.

[Her54] Charles Hermite. "Sur la théorie des fonctions homogènes à deux indéterminées". In: *Cambridge and Dublin Mathematical Journal* 9 (1854), pp. 172–217.

[HP94] W. V. D. Hodge and D. Pedoe. *Methods of Algebraic Geometry. Vol. I*. Cambridge Mathematical Library. Cambridge University Press, 1994. DOI: 10.1017/CBO9780511623899.

[HS65] J. Hartmanis and R. E. Stearns. "On the Computational Complexity of Algorithms". In: *Transactions of the American Mathematical Society* 117 (1965), pp. 285–306. DOI: 10.1090/S0002-9947-1965-0170805-7.

[Jan10] Svante Janson. *Resultant and Discriminant of Polynomials*. Aug. 16, 2010. URL: https://www2.math.uu.se/~svantejs/papers/sjN5.pdf.

[Jun08] Heinrich W. E. Jung. "Darstellung der Funktionen eines algebraischen Körpers zweier unabhängigen Veränderlichen x, y in der Umgebung einer Stelle $x = a, y = b$". In: *Journal für die reine und angewandte Mathematik* 133 (1908), pp. 289–314. DOI: 10.1515/crll.1908.133.289.

[Kat89] Kazuya Kato. "Logarithmic Structures of Fontaine–Illusie". In: *Algebraic Analysis, Geometry, and Number Theory*. Ed. by Jun-ichi Igusa. Johns Hopkins University Press, 1989, pp. 191–224.

[Kay+14] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. "An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas". In: *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2014)*. 2014, pp. 61–70. DOI: 10.1109/FOCS.2014.15.

[Kay12] Neeraj Kayal. *An Exponential Lower Bound for the Sum of Powers of Bounded Degree Polynomials*. Report TR12-081. Electronic Colloquium on Computational Complexity (ECCC), June 26, 2012. URL: https://eccc.weizmann.ac.il/report/2012/081/.

[Koi12] Pascal Koiran. "Arithmetic Circuits: The Chasm at Depth Four Gets Wider". In: *Theoretical Computer Science* 448 (2012), pp. 56–65. DOI: 10.1016/j.tcs.2012.03.041.

[KS07] Neeraj Kayal and Nitin Saxena. "Polynomial Identity Testing for Depth 3 Circuits". In: *Computational Complexity* 16.2 (2007), pp. 115–138. DOI: 10.1007/s00037-007-0226-9.

[KS09] Neeraj Kayal and Shubhangi Saraf. "Blackbox Polynomial Identity Testing for Depth 3 Circuits". In: *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009)*. 2009, pp. 198–207. DOI: 10.1109/FOCS.2009.67.

[KS17a] Mrinal Kumar and Shubhangi Saraf. "Arithmetic Circuits with Locally Low Algebraic Rank". In: *Theory of Computing* 13.6 (2017), pp. 1–33. DOI: 10.4086/toc.2017.v013a006.

[KS17b] Mrinal Kumar and Shubhangi Saraf. "On the Power of Homogeneous Depth 4 Arithmetic Circuits". In: *SIAM Journal on Computing* 46.1 (2017), pp. 336–387. DOI: 10.1137/140999335.

[KSS14] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. "A Super-Polynomial Lower Bound for Regular Arithmetic Formulas". In: *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC 2014)*. 2014, pp. 146–153. DOI: 10.1145/2591796.2591847.

[Kum20] Mrinal Kumar. "On the Power of Border of Depth-3 Arithmetic Circuits". In: *ACM Transactions on Computation Theory* 12.1 (2020), 5:1–5:8. DOI: 10.1145/3371506.

[Lan02] Serge Lang. *Algebra*. 3rd ed. Vol. 211. Graduate Texts in Mathematics. Springer, 2002. DOI: 10.1007/978-1-4613-0041-0.

[Lan17] J. M. Landsberg. *Geometry and Complexity Theory*. Vol. 169. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2017. DOI: 10.1017/9781108183192.

[Laz20] Daniel Lazard. *Degree of a Polynomial Ideal and Bézout Inequalities*. Preprint, HAL hal-02897419, version 1. 2020. URL: https://hal.science/hal-02897419v1.

[LQ15] Henri Lombardi and Claude Quitté. *Commutative Algebra: Constructive Methods: Finite Projective Modules*. Vol. 20. Algebra and Applications. Springer, 2015. DOI: 10.1007/978-94-017-9944-7.

[LST25] Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. "Superpolynomial Lower Bounds against Low-Depth Algebraic Circuits". In: *Journal of the ACM* 72.4 (2025), 26:1–26:35. DOI: 10.1145/3734215.

[Mag12] Arturo Magidin. *Localisation is isomorphic to a quotient of polynomial ring*. Mathematics Stack Exchange. Version: 2018-12-07. 2012. URL: https://math.stackexchange.com/a/152289.

[Mat86] Hideyuki Matsumura. *Commutative Ring Theory*. Vol. 8. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1986. DOI: 10.1017/CBO9781139171762.

[Mil80] James S. Milne. *Étale Cohomology*. Vol. 33. Princeton Mathematical Series. Princeton University Press, 1980. DOI: 10.1515/9781400883981.

[MS01] Ketan D. Mulmuley and Milind Sohoni. "Geometric Complexity Theory I: An Approach to the P vs. NP and Related Problems". In: *SIAM Journal on Computing* 31.2 (2001), pp. 496–526. DOI: 10.1137/S009753970038715X.

[MS08] Ketan D. Mulmuley and Milind Sohoni. "Geometric Complexity Theory II: Towards Explicit Obstructions for Embeddings among Class Varieties". In: *SIAM Journal on Computing* 38.3 (2008), pp. 1175–1206. DOI: 10.1137/080718115.

[Neu99] Jürgen Neukirch. *Algebraic Number Theory*. Vol. 322. Grundlehren der mathematischen Wissenschaften. Translated from the German by Norbert Schappacher. Springer-Verlag, 1999. DOI: 10.1007/978-3-662-03983-0.

[Pal06] Francesco Palatini. "Sulle superficie algebriche i cui S_h $(h+1)$-seganti non riempiono lo spazio ambiente". In: *Atti della Reale Accademia delle Scienze di Torino* 41 (1906),

pp. 634–640. URL: https://archive.org/details/attidellarealeac41real/page/634.

[PR12] Adam Parusiński and Guillaume Rond. "The Abhyankar–Jung Theorem". In: *Journal of Algebra* 365 (2012), pp. 29–41. DOI: 10.1016/j.jalgebra.2012.05.003.

[PS21] Shir Peleg and Amir Shpilka. "Polynomial Time Deterministic Identity Testing Algorithm for $\Sigma^{[3]}\Pi\Sigma\Pi^{[2]}$ Circuits via Edelstein–Kelly Type Theorem for Quadratic Polynomials". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC 2021)*. 2021, pp. 259–271. DOI: 10.1145/3406325.3451013.

[Ros61] Maxwell Rosenlicht. "Toroidal Algebraic Groups". In: *Proceedings of the American Mathematical Society* 12.6 (1961), pp. 984–988. DOI: 10.1090/S0002-9939-1961-0133328-9.

[Sam57] Pierre Samuel. "La notion de place dans un anneau". In: *Bulletin de la Société Mathématique de France* 85 (1957), pp. 123–133. DOI: 10.24033/bsmf.1484.

[Sam66] Pierre Samuel. "À propos du théorème des unités". In: *Bulletin des Sciences Mathématiques*. 2nd ser. 90 (1966), pp. 89–96.

[Sap26] Ramprasad Saptharishi. *A Survey of Lower Bounds in Arithmetic Circuit Complexity*. July 2026. URL: https://github.com/dasarpmar/lowerbounds-survey.

[Sas11] Sasha. *Are units of rings of functions on algebraic varieties finitely generated (mod. constants)?* MathOverflow. version: 2011-03-05. 2011. URL: https://mathoverflow.net/q/57352.

[Sha13] Igor R. Shafarevich. *Basic Algebraic Geometry 1: Varieties in Projective Space*. 3rd ed. Springer, 2013. DOI: 10.1007/978-3-642-37956-7.

[SHL65] R. E. Stearns, J. Hartmanis, and P. M. Lewis. "Hierarchies of Memory Limited Computations". In: *Proceedings of the 6th Annual Symposium on Switching Circuit Theory and Logical Design*. 1965, pp. 179–190. DOI: 10.1109/FOCS.1965.11.

[SS13] Nitin Saxena and C. Seshadhri. "From Sylvester-Gallai Configurations to Rank Bounds: Improved Blackbox Identity Test for Depth-3 Circuits". In: *Journal of the ACM* 60.5 (2013), 33:1–33:33. DOI: 10.1145/2528403.

[Stacks] The Stacks Project Authors. *Stacks Project*. https://stacks.math.columbia.edu. 2026.

[SY10] Amir Shpilka and Amir Yehudayoff. "Arithmetic Circuits: A Survey of Recent Results and Open Questions". In: *Foundations and Trends in Theoretical Computer Science* 5.3–4 (2010), pp. 207–388. DOI: 10.1561/0400000039.

[Tav15] Sébastien Tavenas. "Improved Bounds for Reduction to Depth 4 and Depth 3". In: *Information and Computation* 240 (2015), pp. 2–11. DOI: 10.1016/j.ic.2014.09.004.

[Ter11] Alessandro Terracini. "Sulle V_k per cui la varietà degli S_h $(h+1)$-seganti ha dimensione minore dell'ordinario". In: *Rendiconti del Circolo Matematico di Palermo* 31 (1911), pp. 392–396. DOI: 10.1007/BF03018812.

[Val79] L. G. Valiant. "Completeness Classes in Algebra". In: *Proceedings of the 11th Annual ACM Symposium on Theory of Computing (STOC 1979)*. 1979, pp. 249–261. DOI: 10.1145/800135.804419.

[vzGG13] Joachim von zur Gathen and Jürgen Gerhard. *Modern Computer Algebra*. 3rd ed. Cambridge University Press, 2013. DOI: 10.1017/CBO9781139856065.

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