Electronic Colloquium on Computational Complexity, Report No. 157 (2022)

Border complexity via elementary symmetric polynomials

Pranjal Dutta Fulvio Gesmundo Christian Ikenmeyer Gorav Jindal Vladimir Lysikov

November 2022

Abstract

In (ToCT'20) Kumar surprisingly proved that every polynomial can be approximated as a sum of a constant and a product of linear polynomials. In this work, we prove the converse of Kumar's result which ramifies in a surprising new formulation of Waring rank and border Waring rank. From this conclusion, we branch out into two different directions, and implement the geometric complexity theory (GCT) approach in two different settings.

In the first direction, we study the orbit closure of the product-plus-power polynomial, determine its stabilizer, and determine the properties of its boundary points. We also connect its fundamental invariant to the Alon-Tarsi conjecture on Latin squares, and prove several exponential separations between related polynomials contained in the affine closure of product-plus-product polynomials. We fully implement the GCT approach and obtain several equations for the product-plus-power polynomial from its symmetries via representation theoretic multiplicity obstructions.

In the second direction, we demonstrate that the non-commutative variant of Kumar's result is intimately connected to the constructions of Ben-Or and Cleve (SICOMP'92), and Bringmann, Ikenmeyer, Zuiddam (JACM'18), which describe algebraic formulas in terms of iterated matrix multiplication. From this we obtain that a variant of the elementary symmetric polynomial is complete for V3F, a large subclass of VF, under homogeneous border projections. In the regime of quasipolynomial complexity, our polynomial has the same power as the determinant or as arbitrary circuits, i.e., VQP. This is the first completeness result under homogeneous projections for a subclass of VBP. Such results are required to set up the GCT approach in a way that avoids the no-go theorems of Bürgisser, Ikenmeyer, Panova (JAMS'19).

Finally, using general geometric considerations, we significantly improve the relationship between the Waring rank and the border Waring rank of polynomials. In particular, if the border Waring rank of a homogeneous polynomial f is k, then, the Waring rank of f can be at most $\exp(k) \cdot d$, while previously it was known to be $O(d^k)$.

Keywords: border complexity, Waring rank, arithmetic formulas, geometric complexity theory, symmetric polynomials

PD: Visiting Fellow, Chennai Mathematical Institute, India, duttpranjal@gmail.com
FG: Saarland University, Germany, gesmundo@cs.uni-saarland.de
CI: University of Warwick, UK, christian.ikenmeyer@warwick.ac.uk
GJ: Max Planck Institute for Software Systems, Saarbrücken, Germany, gjindal@mpi-sws.org
VL: QMATH, University of Copenhagen, Denmark, vl@math.ku.dk

ISSN 1433-8092

Contents

1 Introduction

Geometric Complexity Theory (GCT) is an approach towards proving algebraic variants of the $\mathsf{P} \neq \mathsf{NP}$ conjecture using algebraic geometry and representation theory [MS01, MS08]. Let $\mathrm{per}_n := \Sigma_{\sigma\in\mathfrak{S}_n}\Pi_{i=1}^n x_{i,\sigma(i)}$ be the permanent polynomial (on n^2 variables). An algebraic version of the $\mathsf{P} \neq \mathsf{NP}$ conjecture, often called Valiant's *determinant vs. permanent* conjecture, states that the smallest size of a matrix A whose entries are affine linear polynomials such that $\det(A) = \mathrm{per}_n$, is *not polynomially* bounded (equivalently, $\mathsf{VNP} \not\subseteq \mathsf{VBP}$). Mulmuley and Sohoni strengthened the conjecture by allowing arbitrary approximations while computing the permanent polynomial, i.e. $\mathsf{VNP} \not\subseteq \overline{\mathsf{VBP}}$. This question is the central question of GCT.

The Mulmuley–Sohoni conjecture can be rephrased as $\ell^{m-n}\mathrm{per}_n \notin \overline{\mathrm{GL}_{m^2}\det_m}$, if $m = \mathsf{poly}(n)$; here $\mathrm{GL}_{m^2} := \mathrm{GL}(\mathbb{C}^{m\times m})$ acts on the space of homogeneous degree m polynomials in m^2 variables by (invertible) linear transformation of the variables, and ℓ is some linear form (one can assume $\ell := x_{1,1}$). The polynomial $\ell^{m-n}\mathrm{per}_n$ is called the 'padded permanent', and the phenomenon of multiplying with a power of a linear form is called *padding*. The representation theoretic attack on the problem is done by using so-called *obstructions*. Padding is a serious issue in GCT, whose challenges were first highlighted by Kadish and Landsberg [KL14]. This eventually lead to the breakthrough work [BIP19], where it was shown that *occurrence obstructions* are not sufficient to prove Mulmuley and Sohoni's conjecture. Remarkably, in that no-go result, one can replace the permanent by any homogeneous polynomial of degree m in m^2 variables. The proof crucially relies on the padding of the polynomial. This is the main motivation to study *padding-free* models of computation, and homogeneous models are one way to go about it. An Algebraic formula, also called arithmetic formula, is one of the most robust models of computation that are studied in Algebraic Complexity. In this work, we introduce a GCT formulation for formulas, that circumvents padding.

1.1 Definition (Algebraic Formula). *An algebraic formula, over a field $\mathbb{F}$, is a directed tree with a unique sink vertex called the root. The source vertices are labelled by either formal variables or field constants, and each internal node of the graph is labelled by either $+$ or $\times$. Nodes compute formal polynomials in the input variables in the natural way. The polynomial computed by the formula is defined as the polynomial computed by the root.*

The size of a formula is the number of vertices of the tree. VF is the class of polynomial families $(f_n)_{n\in\mathbb{N}}$, with formula-size of f_n being polynomially bounded. Replacing the tree of Definition 1.1 with a directed acyclic graph (i.e., the out degree of a node can be ≥ 2), gives the notion of *algebraic circuit*. Similarly, the size of an algebraic circuit is the number of vertices in the graph. The class VP contains polynomial families $(f_n)_{n\in\mathbb{N}}$, with both circuit-size and degree of f_n being polynomially bounded. Another important model of computation is the *algebraic branching program* (ABP) model. It is a classical result that every homogeneous degree d polynomial f can be written as a product

$$f = \begin{pmatrix} \ell_{1,1,1} & \ell_{1,2,1} & \cdots & \ell_{1,n,1} \end{pmatrix} \begin{pmatrix} \ell_{1,1,2} & \cdots & \ell_{1,n,2} \\ \vdots & \ddots & \vdots \\ \ell_{n,1,2} & \cdots & \ell_{n,n,2} \end{pmatrix} \begin{pmatrix} \ell_{1,1,3} & \cdots & \ell_{1,n,3} \\ \vdots & \ddots & \vdots \\ \ell_{n,1,3} & \cdots & \ell_{n,n,3} \end{pmatrix} \cdots \begin{pmatrix} \ell_{1,1,d-1} & \cdots & \ell_{1,n,d-1} \\ \vdots & \ddots & \vdots \\ \ell_{n,1,d-1} & \cdots & \ell_{n,n,d-1} \end{pmatrix} \begin{pmatrix} \ell_{1,1,d} \\ \vdots \\ \ell_{n,1,d} \end{pmatrix}$$

of matrices whose entries are homogeneous linear forms. If f can be computed in the above form, we say that f has a small ABP of width n. In fact, if we define $w(f)$ to be the smallest possible such n, then for a sequence of homogeneous polynomials $(f_n)_{n\in\mathbb{N}}$ we have that

$$(f_n)_{n\in\mathbb{N}} \in \mathsf{VBP} \text{ iff } w(f_n) \text{ is polynomially bounded.}^{6} \tag{1.2}$$

It is known that $\mathsf{VF} \subseteq \mathsf{VBP} \subseteq \mathsf{VP}$ [Val79]. For any class $\mathcal{C}$, we can define its (Zariski) closure $\overline{\mathcal{C}}$, which contains polynomial families, where the polynomials can be efficiently approximated arbitrarily closely; for a formal definition, see Section 2. It is not known how to homogenize arithmetic formulas with only a polynomially large blow-up in size; this makes padding-free study of formulas quite challenging. In this regard, here is a meta-question.

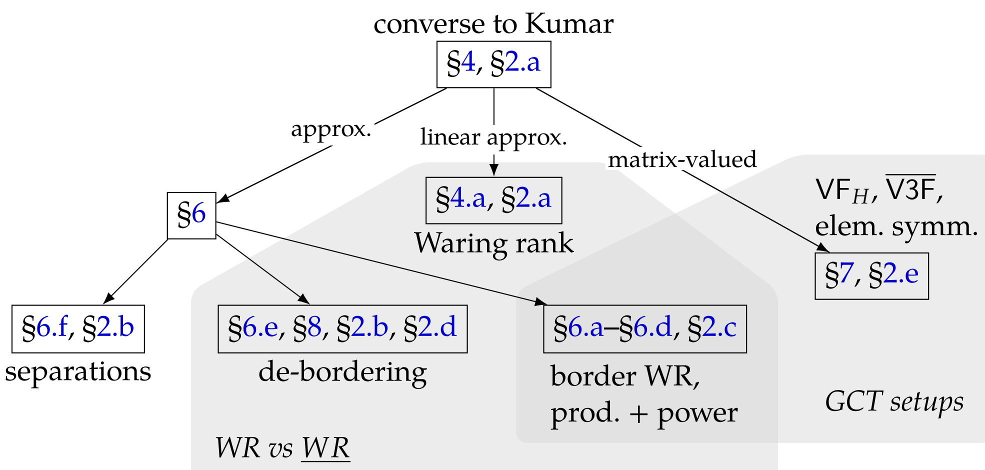
1.3 Meta-Question. Prove $(\mathrm{per}_n)_{n\in\mathbb{N}} \notin \overline{\mathsf{VF}}$, or, equivalently $\mathsf{VNP} \not\subseteq \overline{\mathsf{VF}}$, via padding-free GCT techniques.

Seemingly unrelated (to the above), the central model of our interest is the following: $\alpha \cdot \big(\prod_{i\in[m]}(1+\ell_i) - 1\big)$, for n-variate linear forms $\ell_i \in \mathbb{C}[\boldsymbol{x}]_1$, and $\alpha \in \mathbb{C}$. Note that, any polynomial f computed by the above model has degree $= m$ and is constant-free (constant term is 0). Further, f has a trivial formula size of $O(mn)$.

1.4 Problem. How powerful is the model $\alpha \cdot (\prod_{i\in[m]}(1+\ell_i) - 1)$, for linear forms ℓ_i?

Note that this is *not* a complete model: for instance, a simple monomial $x_1 \cdots x_n$ *cannot* be written as $\alpha \cdot (\prod_{i\in[m]}(1+\ell_i) - 1)$; for a quick proof see Lemma 4.3. But, what if we allow $\ell_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]_1$, $\alpha \in \mathbb{C}(\epsilon)$, and look at the limit polynomials $\lim_{\epsilon\to 0} \alpha \cdot (\prod_{i\in[m]}(1+\ell_i) - 1)$, if it exists? How much more can it compute? In this work, we answer this question in fair details by proving the converse of a recent result of Kumar [Kum20]. We also show a surprising, intimate relation to Meta-Question 1.3; for details, see Section 2.

The starting point of the paper is the study of Kumar's model in Section 4. The other sections relate to this as follows.



Acknowledgements

The work of F.G. is partially supported by the Thematic Research Programme "Tensors: geometry, complexity and quantum entanglement", University of Warsaw, Excellence Initiative – Research University and the Simons Foundation Award No. 663281 granted to the Institute of Mathematics of the Polish Academy of Sciences for the years 2021-2023. C.I. was supported by the DFG grant IK 116/2-1 and the EPSRC grant EP/W014882/1. V.L. acknowledges financial support from the European Research Council (ERC Grant Agreement No. 818761) and VILLUM FONDEN via the QMATH Centre of Excellence (Grant No. 10059).

[6] This follows from the fact that algebraic branching programs can be homogenized with only a polynomially large blow-up. This is also true for arithmetic circuits, but we do not know it for formulas.

2 Our Results

In this section, we state our results. Among many, some of the results are *de-bordering* results. A de-bordering result is an upper bound on the complexity of all polynomials which have low approximation complexity. Before stating our results, we unify some of our basic notations that will be used throughout. Let $V := \mathbb{C}[x_1, \ldots, x_n]_1 := \mathrm{Span}(\{x_1, \ldots, x_n\})$ be the space of linear forms in $x_1, \ldots, x_n$. Let $S^d V := \mathbb{C}[x_1, \ldots, x_n]_d =: \mathbb{C}[V^*]_d$, and let $f \in S^d V$ be a homogeneous polynomial of degree d. We introduce the following notions of *projection* and *degeneration*.

2.1 Definition. *Let U, W be finite dimensional complex vector spaces and let $f \in \mathbb{C}[U]_d$, $g \in \mathbb{C}[W]_d$. We say that f is a projection of g, and write $f \leq g$ if $f \in \{g \circ A \mid A\colon U \to W \text{ linear}\}$. We say that f is a degeneration of g, and write $f \trianglelefteq g$, if $f \in \overline{\{g \circ A \mid A\colon U \to W \text{ linear}\}}$. Write $f \leq_{\mathrm{aff}} g$, if $f \in \{g \circ A \mid A\colon U \to W \text{ affine linear}\} = \{v \mapsto g(Av + b) \mid A\colon U \to W \text{ linear},\ b \in W\}$. Write $f \trianglelefteq_{\mathrm{aff}} g$ if $f \in \overline{\{g \circ A \mid A\colon U \to W \text{ affine linear}\}}$.*

All closures in Definition 2.1 can be taken, equivalently, in the Euclidean or the Zariski topology, see e.g. [Kra85, AI.7.2 Folgerung].

Closure. For $\mathcal{C} \in \{\mathsf{VF}, \mathsf{VBP}, \mathsf{VP}, \mathsf{VNP}\}$ define $\overline{\mathcal{C}}$ via $(f_n)_n \in \overline{\mathcal{C}} \iff \exists (g_n)_n \in \mathcal{C}\ \forall n : f_n \trianglelefteq_{\mathrm{aff}} g_n$ [7].

2.a Kumar's complexity

For a polynomial f, let *Kumar's complexity*, denoted, $\mathsf{Kc}(f)$ be the smallest m such that there exists a constant α and homogeneous linear polynomials ℓ_i such that

$$f = \alpha\Big(\big(\prod_{i=1}^{m}(1+\ell_i)\big) - 1\Big). \tag{2.2}$$

If no such m exists, we set $\mathsf{Kc}(f) := \infty$. Note that for this definition the field does not matter, but in this paper we often restrict our attention to the complex numbers $\mathbb{C}$. Clearly, if $\mathsf{Kc}(f)$ is finite, then the highest degree homogeneous part of f is a product of homogeneous linear forms, hence not all f have a finite Kc.

For $f_\epsilon, g_\epsilon \in \mathbb{C}(\epsilon)[\boldsymbol{x}]$ we write $f_\epsilon \simeq g_\epsilon$ if both limits $\lim_{\epsilon\to 0} f_\epsilon$ and $\lim_{\epsilon \to 0} g_\epsilon$ exist, and both limits coincide. Algebraically this means that $f_\epsilon, g_\epsilon \in \mathbb{C}[[\epsilon]][\boldsymbol{x}]$ and $f_\epsilon \equiv g_\epsilon \mod \langle \epsilon \rangle$. Let $\underline{\mathsf{Kc}}(f)$ denote the smallest m such that there exists $f_\epsilon \in \mathbb{C}(\epsilon)[\boldsymbol{x}]$ and $\ell_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]_1$, $\alpha \in \mathbb{C}(\epsilon)$ with $f_\epsilon \simeq f$ and $\forall \beta \neq 0 : \mathsf{Kc}(f_\epsilon|_{\epsilon=\beta}) \leq m$. An equivalent definition for $\underline{\mathsf{Kc}}$ is the smallest m such that $f \simeq f_\epsilon$ with $f_\epsilon = \alpha\big(\big(\prod_{i=1}^m (1+\ell_i)\big) - 1\big)$, where $\alpha \in \mathbb{C}(\epsilon)$ and $\ell_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]_1$.[8] Kumar [Kum20] proved that $\underline{\mathsf{Kc}}(f)$ is indeed always finite for homogeneous polynomials (in this introductory section we limit our discussion to homogeneous polynomials for the sake of a clearer exposition, but of course using padding one sees immediately that this also works for inhomogeneous polynomials):

4.4 Proposition. For all homogeneous f we have $\underline{\mathsf{Kc}}(f) \leq \deg(f) \cdot \mathsf{WR}(f)$.

In the above, $\mathsf{WR}(f)$, the *Waring rank* of a homogeneous degree-d polynomial f, is the minimal r in an expression $f = \sum_{i \in [r]} \ell_i^d$, where ℓ_i are linear forms. We also denote $\underline{\mathsf{WR}}(f)$, the *border Waring rank* of f, as the minimum r such that $f = \lim_{\epsilon \to 0} \sum_{i\in[r]} \ell_i^d$, where $\ell_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]_1$ are linear forms.

[7] For a more general definition of the closure see [IS22]

[8] It is often more convenient to work with approximations in $\mathbb{C}[\epsilon^{-1}, \epsilon]$ instead of $\mathbb{C}(\epsilon)$. This can always be achieved by first representing rational functions by their Laurent series at 0, thus going from $\mathbb{C}(\epsilon)$ to $\mathbb{C}((\epsilon)) = \mathbb{C}[[\epsilon]][\epsilon^{-1}]$, and then truncating the Laurent series at degree high enough so that it does not affect approximations.

In this work, we prove the *converse* of the above statement. Let $\delta_f := 1$ if f is a product of homogeneous linear forms, and define $\delta_f = \infty$ otherwise. The following theorem explains the relation between border Waring rank and Kumar's complexity.

4.7 Theorem (Border Waring rank via Kumar's complexity)**.** For all homogeneous f we have

$$\min\{\deg(f)\cdot\delta_f,\ \underline{\mathsf{WR}}(f)\} \;\le\; \underline{\mathsf{Kc}}(f) \;\le\; \deg(f)\cdot\min\{\delta_f,\ \underline{\mathsf{WR}}(f)\}.$$

4.8 Corollary (De-bordering $\underline{\mathsf{Kc}}$)**.** For all homogeneous f, $\underline{\mathsf{Kc}}(f) = m \implies$ either $\underline{\mathsf{WR}}(f) \le m$, or f is a product of linear forms.

What if we only allow linear approximations? Interestingly, then the converse of Proposition 4.4 is also almost true! To this end, we define: $\underline{\mathsf{Kc}_1^-}(f)$ is the smallest m such that there $\exists \ell_i \in \mathbb{C}[\boldsymbol{x}]_1, M \ge 1, \gamma \in \mathbb{C}$ with $f \simeq \gamma\epsilon^{-M}\big(\prod_{i=1}^m(1+\epsilon\ell_i) - 1\big)$.

4.12 Proposition (Waring rank via Kumar's complexity)**.** For any homogeneous polynomial f of degree d, we have $\mathsf{WR}(f) \le \underline{\mathsf{Kc}_1^-}(f) \le d\cdot\mathsf{WR}(f)$.

Therefore, the above two theorems show that variants of Kumar's complexity exactly capture the two classical measures, namely Waring rank and border Waring rank.

2.b Orbit closures of restricted binomials

Fix integers d, r, s, with $d \ge 3$, and consider the polynomial $P_{r,s}^{[d]} = \sum_{i=1}^r \prod_{j=1}^d x_{ji} + \sum_{i=1}^s y_i^d$, in the $rd+s$ variables $x_{11}, \ldots, x_{dr}, y_1, \ldots, y_s$. We sometimes also use new x-variables for the y-variables. Hüttenhain [Hüt17] studied a polynomial family, called *generic binomials* (we use bn_d to denote it throughout the paper):

$$\mathrm{bn}_d(\boldsymbol{x},\boldsymbol{y}) \;:=\; P_{2,0}^{[d]} \;=\; \prod_{i\in[d]} x_i + \prod_{i\in[d]} y_i\,.$$

The goal in [Hüt17] was to (completely) describe the components of the *boundary* $\partial\Omega_{\mathrm{bn}_d} = \overline{\mathrm{GL}_{2d}\,\mathrm{bn}_d} \setminus \mathrm{GL}_{2d}\,\mathrm{bn}_d$. Surprisingly, in the affine closure model, Kumar's result [Kum20] showed the following:

$$\text{For every } f \in \mathbb{C}[\boldsymbol{x}]_d \text{ we have } f \trianglelefteq_{\mathrm{aff}} \mathrm{bn}_e \text{ for some } e \le \exp(\log n, d) \tag{2.3}$$

On the other hand, [DDS22] showed that if $f \trianglelefteq_{\mathrm{aff}} \mathrm{bn}_d$, then f has a small ABP. To this end, we define the *product-plus-k-powers* $P_k^{[d]}$ as:

$$P_k^{[d]} \;:=\; P_{1,k}^{[d]} \;=\; \prod_{i\in[d]} x_i + \sum_{i\in[k]} x_{d+i}^d\,.$$

In particular, when $k = 1$, we write $P^{[d]} := P_1^{[d]}$, a *product-plus-power*. For a polynomial $f \in S^dV$ that does not involve some variable x_i we write $f \trianglelefteq_{x_i}^{x_j} g$ if $f \in \overline{\{g\circ A \mid A : U \to W \text{ linear and } A(\mathbb{C}x_j) = \mathbb{C}x_i\}}$. This definition is inspired by the definition of a *parabolic subgroup* of the general linear group. If f does not involve x_0, we observe that

$$\underline{\mathsf{Kc}}(f) \le m \iff x_0^{m-d} f \trianglelefteq_{x_0}^{x_{d+1}} P^{[m]} \underset{\not\Longleftarrow}{\Longrightarrow} x_0^{m-d} f \trianglelefteq P^{[m]} \iff f \trianglelefteq_{\mathrm{aff}} P^{[m]}$$

6.20 Theorem (De-bordering product-plus-power)**.** Let $f \in S^dV$, such that $f \trianglelefteq P^{[d]}$,then either $f \le P^{[d]}$ or $\underline{\mathsf{WR}}(f) \le (d+1)^3(d^2+1)$.

Similarly, we call $P_2^{[d]}$ a *product-plus-two-powers*; it is also a *special* binomial, since,

$$\prod_{i\in[d]} x_i + x_{d+1}^d + x_{d+2}^d = \prod_{i\in[d]} x_i + \prod_i (x_{d+1} + \zeta^i x_{d+2}), \text{ where } \zeta \text{ is the } 2d-\text{th root of unity} .$$

The following theorem shows that we can de-border product-plus-two-powers.

6.25 Theorem (De-bordering product-plus-two-powers)**.** Let $f \in S^d V$, such that $f \trianglelefteq P_2^{[d]}$, then, either
1. $f \le P_2^{[d]}$, or, 2. $f \le \prod_{i\in[d]} y_i + y_0^{d-1} \cdot y_{d+1}$, or, 3. $\underline{\mathrm{WR}}(f) \le O(d^7)$.

2.4 Remark. As argued before, both product-plus-power and product-plus-two-powers are special binomials, and hence, from [DDS22] it follows that their orbit closures are contained in VBP. Our results (both Theorem 6.20 & Theorem 6.25) are more *fine-grained* de-bordering than VBP, since $\underline{\mathrm{WR}}(f) \le \mathsf{poly}(d) \implies f \in \mathsf{VBP}$ [For16, GKS17, BDI21], and the converse *does not* necessarily hold, because $\underline{\mathrm{WR}}(\det_d) = \exp(d)$ [Sax08, CKW11].

Exponential separation within the binomial.

Perhaps, the de-bordering results are must-to-understand, to understand the limitations and power of computations/approximations in different models. However, identifying explicit polynomials which are hard to compute/approximate, and proving it remains a major template in algebraic complexity. Often, proving lower bounds on the homogeneous model turns out to be *easier* than in its affine model, because of the non-trivial cancellations in the latter model.

In [DS22], it was established that if $\mathrm{trn}_d \trianglelefteq_{\mathrm{aff}} \mathrm{bn}_e$, where $\mathrm{trn}_d := \Pi_{i\in[d]} x_i + \Pi_{i\in[d]} y_i + \Pi_{i\in[d]} z_i$, is a *trinomial*, then $e \ge \exp(d)$. In fact, such an exponential hierarchy exists, because of the exponential gap between a k-monomial, to $(k+1)$-monomial; by k-monomial, we mean k-many sum of product of distinct variables (for e.g., 2-monomial is nothing but binomial). In this work, since, we are working over two restricted models of binomial, we are interested in gaps between the models (if possible). Indeed, we show that these affine models in the border, are *exponentially* separated.

6.34 Theorem. If $P_2^{[d]} \trianglelefteq_{\mathrm{aff}} P^{[e]}$, then $e \ge \exp(d)$.

6.35 Theorem. If $\mathrm{bn}_d \trianglelefteq_{\mathrm{aff}} P_2^{(e)}$, then $e \ge \exp(d)$.

2.5 Remark (Optimality)**.** Both Theorem 6.34 and Theorem 6.35 are *optimal*, since, Kumar's result [Kum20] (see Equation (2.3)), we know that there exists $e \le \exp(d)$ such that $P_2^{[d]} \trianglelefteq_{\mathrm{aff}} P^{[e]}$, and similarly $\mathrm{bn}_d \trianglelefteq_{\mathrm{aff}} P^{[e]}$.

2.c Stabilizers, fundamental invariants, and the Alon-Tarsi conjecture on Latin Squares

At a first glance the polynomial $P^{[d]} := \prod_{i\in[d]} y_i + y_0^d$ looks very similar to the well-studied product polynomial $\prod_{i\in[d]} y_i$, which was also used in several GCT papers [Kum15, BI17, DIP20, IK20]. A set of set-theoretic equations is known due to Brill and Gordon [Gor94], and their representation theoretic structure has been recently described by Guan [Gua18]. It is of interest to know how much of the theory can be transferred from the product to $P^{[d]}$, and how much of the GCT approach can be implemented. We start by determining the stabilizer of $P^{[d]}$.

6.2 Theorem (simplified version)**.** The stabilizer of $P^{[d]}$ in GL_{d+1} is $\mathbb{Z}_d \times (\mathbb{T}^{\mathrm{SL}_d} \rtimes \mathfrak{S}_d)$.

This is promisingly close to the stabilizer of the product polynomial. We then study the multiplicities in the coordinate ring of the orbit via classical representation theoretic branching rules. Recall, the irreducible representations of GL_{d+1} are indexed by partitions $\lambda = (\lambda_1, \lambda_2, \ldots)$, $\lambda_1 \geq \lambda_2 \geq \ldots$ with $\ell(\lambda) \leq d+1$, see Section 6.b. We denote by $S_\lambda(\mathbb{C}^{d+1})$ the irreducible representation of type λ. For a GL_{d+1}-representation $\mathcal{V}$ we write $\mathrm{mult}_\lambda(\mathcal{V})$ to denote the multiplicity of λ in $\mathcal{V}$, i.e., the dimension of the space of equivariant maps from $S_\lambda(\mathbb{C}^{d+1})$ to $\mathcal{V}$, or equivalently, the number of summands of isomorphism type λ in any decomposition of $\mathcal{V}$ into a direct sum of irreducible representations.

6.3 Proposition. For $\lambda \vdash dD$ we have $\mathrm{mult}_\lambda(\mathbb{C}[\mathrm{GL}_{d+1} P^{[d]}]) = \dim(S_\lambda V)^H = \sum_{\delta=0}^{D} \sum_{\mu \vdash \delta d, \mu \preceq \lambda, \ell(\mu) \leq d} a_\mu(d, \delta)$, where $a_\mu(d,\delta)$ is the *plethysm coefficients*, i.e., the multiplicity of μ in $\mathrm{Sym}^d(\mathrm{Sym}^\delta(V))$.

We implement this formula and indeed find an abundance of multiplicity obstructions, see appendix A, whose detailed study seems to be a promising future direction.

The product polynomial is *polystable*, which means that its SL-orbit is closed. The *fundamental invariant* Φ of a polystable polynomial $f \in \mathrm{Sym}^D V$ is the smallest degree $\mathrm{SL}(V)$-invariant function in $\mathbb{C}[\overline{\mathrm{GL}(V)f}]$, see Def. 3.8 in [BI17]. It describes the connection between the orbit and the orbit closure, more formally $\mathbb{C}[\overline{\mathrm{GL}(V)f}]_\Phi \simeq \mathbb{C}[\mathrm{GL}(V)f]$ is the localization at Φ, see [BI17, Pro. 3.9]. This connection can be used to exhibit multiplicity obstructions, as was done in [IK20]. For the product polynomial, if d is even, the degree of Φ is d, iff the Alon-Tarsi conjecture on Latin squares is true for d; otherwise it is of higher degree. A Latin square of side length d is a $d \times d$ matrix filled with numbers from $\{1, \ldots, d\}$ such that each row and column is a permutation. Hence, each column of a Latin square has a sign in $\{-1, 1\}$, which is the sign of the permutation. The column sign of a Latin square is the product of the signs of all its columns. If d is odd, then it is easy to see that the number of side length d Latin squares with column sign $+1$ and the number of side length d Latin squares with column sign -1 are the same. The Alon-Tarsi conjecture is that for all even d, the number of side length d Latin squares with column sign $+1$ and the number of side length d Latin squares with column sign -1 are different. This is known to be true for every side length $p+1$ and $p-1$ for odd primes p, which means that the smallest unknown case is $d = 26$, see [Dri97, Gly10].

We prove that $P^{[d]}$ is polystable, which implies that a fundamental invariant exists, see Proposition 6.4. We then prove that the situation is similar to the situation of the product polynomial:

6.6 Theorem. Let d be even. The degree of the fundamental invariant of $P^{[d]}$ is $d+1$ if and only if the Alon-Tarsi conjecture for d is true, otherwise it is in a higher degree.

2.d Fixed-parameter de-bordering of border Waring rank

In §8.b we prove a de-bordering result for border Waring rank which has the following form: there exists a function $\varphi : \mathbb{N} \to \mathbb{N}$ such that if a homogeneous polynomial f of degree d has $\underline{\mathsf{WR}}(f) \leq r$, then $\mathsf{WR}(f) \leq \varphi(r) \cdot d$. By analogy with fixed-parameter tractability in parameterized complexity, we call this *fixed-parameter de-bordering* for border Waring rank.

8.11 Theorem (simplified)**.** Let $f \in S^d V$ be such that $\underline{\mathsf{WR}}(f) = r$. Then, $\mathsf{WR}(f) \leq \exp(r) \cdot d$.

Comparison with previous bounds. To the best of our knowledge, previous methods only allow upper bounds of the order d^r or r^d. Clearly, the gap is significant even when $r = O(\log d)$, since, the multiplier function in our theorem is such that it gives $\exp(r) = \mathsf{poly}(d)$, when $r = O(\log d)$, while the previous methods could only give $\min(d^r, r^d) = d^{\log d}$, a quasipolynomial upper bound.

To get $\mathsf{WR}(f) \leq O(d^r)$, one can note that a polynomial with border Waring rank r can be transformed into a polynomial *only* in r variables using a linear change of variables (for e.g., see Theorem 5.2), and then, one can argue that the maximal possible Waring rank of an r-variate d-degree can be at most $O(d^r)$ (constant depending on r) [BT15].

Alternatively, one can use the fact that a polynomial with border Waring rank r can be computed by an noncommutative ABP of width r [BDI21]. An upper bound $\mathsf{WR}(f) \leq 2^{d-1} r^d$ can be obtained by writing an ABP as a sum of at most r^d products, one for each path.

Other known de-bordering techniques, such as the interpolation technique using the approximation degree bound of Lehmkuhl and Lickteig [LL89] or the DiDIL technique from [DDS22] can be applied in the border Waring rank setting, but do not improve over the simpler results discussed above.

We also characterize $\underline{\mathsf{WR}}(f) \leq 3$, in an elementary way (compared to [LT10]). For details, see Theorem B.1 & Theorem B.2.

2.e Homogeneous complexity for formulas

For $n, d \in \mathbb{N}$ we define the homogeneous degree d $n \times n$ iterated matrix multiplication polynomial

$$\mathrm{IMM}_n^{(d)} := \begin{pmatrix} x_{1,1,1} & x_{1,2,1} & \cdots & x_{1,n,1} \end{pmatrix} \begin{pmatrix} x_{1,1,2} & \cdots & x_{1,n,2} \\ \vdots & \ddots & \vdots \\ x_{n,1,2} & \cdots & x_{n,n,2} \end{pmatrix} \begin{pmatrix} x_{1,1,3} & \cdots & x_{1,n,3} \\ \vdots & \ddots & \vdots \\ x_{n,1,3} & \cdots & x_{n,n,3} \end{pmatrix} \cdots \begin{pmatrix} x_{1,1,d-1} & \cdots & x_{1,n,d-1} \\ \vdots & \ddots & \vdots \\ x_{n,1,d-1} & \cdots & x_{n,n,d-1} \end{pmatrix} \begin{pmatrix} x_{1,1,d} \\ \vdots \\ x_{n,1,d} \end{pmatrix},$$

which is a polynomial on $(d-2)n^2 + 2n$ variables. We define H to be the set of sequences of homogeneous polynomials. We write VBP_H instead of $\mathsf{VBP} \cap H$ for brevity. Equation (1.2) gives a characterization of VBP_H in terms of the polynomially bounded growth of w, which is a complexity measure in a homogeneous model of computation. This is desirable from the viewpoint of geometric complexity theory. It allows to set up a geometric complexity theory approach without having to rely on the padding of the hard polynomial, i.e., the border complexity of per_d can be defined as the smallest n such that $\mathrm{per}_d \in \overline{\mathrm{GL}_{(d-2)n^2+2n}\,\mathrm{IMM}_n^{(d)}}$. The statement "$\mathsf{VNP} \subseteq \overline{\mathsf{VBP}}$" is equivalent to "the border complexity of per_d is polynomially bounded" [9]. Note that the classical analogue for determinantal complexity requires padding [MS01, MS08, BLMW11]: the border determinantal complexity of per_d is the smallest n such that $x_{n,n}^{n-d} \cdot \mathrm{per}_d \in \overline{\mathrm{GL}_{n^2}\,\det_n}$. The padding leads to the problems explained in [KL14] [IP17, BIP19]. A padded setup is equivalent to a setup with general affine groups instead of general linear groups, see for example [MS21]. In classical algebraic geometry, the Waring rank, Chow rank, and related notions also do not rely on any padding [Lan15]. We would like to describe more complexity classes without padding, potentially easier ones than VBP. In this paper we provide such an example for VF in Corollary 7.5 and an even nicer example for a subclass V3F that we define in §7.c. We formalize the desired properties with the notion of a *collection* in Definition 2.6.

One approach would be to study sums of products of homogeneous linear polynomials, but note that Nisan and Wigderson [NW97] showed that every homogeneous sum of products of d homogeneous linear polynomials computing the determinant $\det_d$ must have $2^{\Omega(d)}$ many summands, hence polynomially sized (or even quasipolynomially sized) homogeneous $\Sigma\Pi\Sigma$ circuits do not capture efficient computation.

A p-family is a sequence of polynomials such that the number of variables and the degree is polynomially bounded. We write $g_{n,d}$ for the homogeneous degree d part of the n-th element of a p-family (g).

[9] This is due to the fact that the homogenization of algebraic branching programs works over every field, so also over $\mathbb{C}(\epsilon)$.

2.6 Definition. *A* collection $((f))$ *is a map* $f : \mathbb{N} \times \mathbb{N} \to \mathbb{C}[x_1, x_2, \ldots]$ *such that every* $f(n,d)$ *is homogeneous of degree* d*. Let* C *be a class of p-families (for example,* $C = \mathsf{VF}$*). We say that a collection* $((f))$ *is* C-p-hard *if for every* $(g) \in C$ *there exists a polynomially bounded function* q *such that* $\forall d > 0, n :$ $g_{n,d} \le f_{q(n),d}$*. If* q *is only quasipolynomially bounded, we say* $((f))$ *is* C-qp-hard*. We define* C-$\overline{\text{p}}$-hardness *and* C-$\overline{\text{qp}}$-hardness *analogously with* $\trianglelefteq$ *instead of* $\le$*.*

Note that in this definition it is important that the maps are homogeneous, see Definition 2.1. It is clear that homogeneous linear projections fix the constant coefficient of polynomials, hence we have $d > 0$ in the definition. Clearly, if $((f))$ is C-$\overline{\text{p}}$-hard, then $((f))$ is also $\overline{C}$-$\overline{\text{p}}$-hard.

For example, the power sum collection $x_1^d + \cdots + x_n^d$ is p-hard for VWaring. And the homogeneous iterated matrix multiplication collection $\mathrm{IMM}_n^{(d)}$ is p-hard for VBP.

A degree d monomial on $6n$ variables variables $x_{i,j,k}$, $(j,k) \in \{(1,2),(1,3),(2,1),(2,3),(3,1),(3,2)\}$, is called *suitable* if when we sort the variables increasingly by their first index, the first variable is of type $x_{*,1,*}$ and the last variable is of type $x_{*,2,*}$, and for each variable $x_{i,j,k}$ the next variable $x_{i',j',k'}$ satisfies $j' = k$. Let $D_{n,d} := \sum_{m \text{ suitable}} m$.

7.5 Corollary (simplified)**.** The collection $D_{n,d}$ is VF-p-hard.

Our main result in this direction is about the collection of parity-alternating elementary symmetric polynomials $C_{n,d} := \sum_{(i_1,i_2,\ldots,i_d)\in P} x_{i_1} x_{i_2} \cdots x_{i_d}$, where P be the set of length d increasing sequences of numbers $i_1 < i_2 < \ldots < i_d$ from $1, \ldots, n$ in which for all j the parity of i_j differs from the parity of i_{j+1}, and i_1 is odd. Let $\mathsf{V3F} \subseteq \mathsf{VF}_H$ be the class of p-families that can be computed by formulas whose inputs are homogeneous linear (no constants allowed) and whose product gates have arity exactly 3 (in order to compute polynomials f of even degree we also allow to compute all $\partial f / \partial x_i$ instead, see Section 7.c).

7.11 Theorem. The collection $C_{n,d}$ is $\overline{\mathsf{V3F}}$-$\overline{\text{p}}$-hard and $\overline{\mathsf{VQP}_H}$-$\overline{\text{qp}}$-hard.

This gives a very nice candidate for setting up a clean (i.e., without padding) geometric complexity theory by considering how large n has to be for $\mathrm{per}_d \in \overline{\mathrm{GL}_n C_{n,d}}$ to hold and asking if that n is (quasi-)polynomially bounded in d. This is polynomially bounded iff $\mathsf{VNP}_H \subseteq \overline{\mathsf{V3P}}$, and it is quasipolynomially bounded iff $\mathsf{VQNP} \subseteq \overline{\mathsf{VQP}}$.

3 Proof Ideas

In this section, we will briefly discuss the techniques and overall ideas which are used to derive the main results.

Proof idea of Theorem 6.20: De-bordering product-plus-power. If $f = \lim_{\epsilon\to 0}(T_1 + T_2)$, where $T_1 := \ell_1 \cdots \ell_d$, and $T_2 := \ell_0^d$, for $\ell_i \in \mathbb{C}[\epsilon, \epsilon^{-1}][x]_1$, then of course a trivial limit point (polynomial) is when individually $\lim_{\epsilon\to 0} T_i$ exists. This readily gives that $f \le P^{[d]}$. Otherwise, one can multiply with large power ϵ^M (where $M \ge 1$), both sides, and work over $\mathbb{C}[\epsilon]$. In particular, one can assume that $f = \lim_{\epsilon\to 0} 1/\epsilon^M \cdot (T_1 + T_2)$, where $T_1 := \ell_1 \cdots \ell_d$, and $T_2 := \ell_0^d$, for $\ell_i \in \mathbb{C}[\epsilon][x]_1$. One can further assume that $\ell_{i,0} := \ell_i|_{\epsilon=0} \ne 0$, for all i, since otherwise ϵ-power can be properly extracted and cancelled. Hypothetically, let us assume the *simplest affine* [10] setup when $\ell_i = 1 + \epsilon \cdot \widehat{\ell_i}$, for $i \in [d]$, and $\ell_0 = 1$. This is exactly the affine $\underline{\mathsf{Kc}}$ setup; an approximate Newton Identities-based proof implies that $\underline{\mathsf{Kc}}(f) = d \implies f \in \overline{\Sigma^{[O(d^2)]} \wedge \Sigma}$ (the affine border Waring rank, for definition,

[10] Note that the given form is not in the affine setup.

see Section 5); also for details see Theorem 6.14. Since, $\ell_0 \equiv 1 \bmod (\ell_0 - 1)$, we *would like* to in fact reduce the general case to the affine $\underline{\mathsf{Kc}}$ setup by working ideal that contains $(\ell_0 - 1)$.

In particular, we work with $\mathcal{I}_{\epsilon,\alpha}$, where $\mathcal{I}_{\epsilon,\alpha} := \langle \ell_0 - \alpha, \epsilon \rangle$, for some $\alpha \in \mathbb{C}$. Note that, $\mathcal{I}_{\epsilon,\alpha} = \langle \ell_{0,0} - \alpha, \epsilon \rangle$. One can *reconstruct* f, from $(d+1)$-many f_α, where $f_\alpha := f \bmod \mathcal{I}_{\epsilon,\alpha} = f \bmod (\ell_{0,0} - \alpha)$, by simple interpolation. Furthermore, one can show the following *low rank property*: $\operatorname{rank}(\ell_{i,0} \mid i \in [0,d]) = 1$. This helps to prove that $\underline{\mathsf{Kc}}(f_\alpha) \le d$. As mentioned above, this implies that each f_α has small affine border Waring rank. Since, affine border Waring rank is sub-additive, this finishes the proof.

Proof idea of Theorem 6.25: De-bordering product-plus-two-powers. This case becomes a bit more complicated, although the heart of the proof lies in a similar idea as above – (i) show a *low rank property*, and (ii) collect enough information to reconstruct f. Let $f = \lim_{\epsilon\to 0}(T_1 + T_2 + T_3)$, where $T_1 := \ell_1 \cdots \ell_d$, and $T_2 := \ell_0^d$, and $T_3 = \ell_{d+1}^d$, for $\ell_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]_1$. The most interesting case is when individually $\lim_{\epsilon\to 0} T_i$ *does not* exist, and there is a nontrivial cancellation among them. Similarly as before, we clear out by multiplying ϵ^M, and hence, $f = \lim_{\epsilon\to 0} 1/\epsilon^M \cdot (T_1 + T_2 + T_3)$.

If *life were easy*, one would have worked with $\mathcal{I}_{\epsilon,\alpha,\beta} := \langle \ell_0 - \alpha, \ell_{d+1} - \beta, \epsilon \rangle$, for $\alpha, \beta \in \mathbb{C}$, and 'hope' that $\ell_i \bmod \mathcal{I}_{\epsilon,\alpha,\beta}$ becomes $1 + \epsilon \cdot \widehat{\ell}_i$, so that $(T_1 + T_2 + T_3) \bmod \mathcal{I}_{\epsilon,\alpha,\beta} \equiv \gamma \cdot \left(\prod_i (1 + \epsilon \cdot \widehat{\ell}_i) - 1\right)$, and we know how to de-border affine $\underline{\mathsf{Kc}}$ complexity (see Theorem 6.14). Note that, $f_{\alpha,\beta} = f \bmod \mathcal{I}_{\epsilon,\alpha,\beta} = f \bmod \langle \ell_{0,0} - \alpha, \ell_{d+1,0} - \beta \rangle$, where $\ell_{i,0} := \ell_i|_{\epsilon=0}$. Therefore, once one shows that $\underline{\mathsf{Kc}}(f_{\alpha,\beta})$ is small, one can interpolate to get f.

The fundamental problem with this approach is there could be cases when $1 \in \mathcal{I}_\epsilon$; in that case, working with mod $\mathcal{I}_\epsilon$ *does not* make sense. A plausible example could be $\ell_0 := x_1$, and $\ell_{d+1} := 2x_1$, clearly $1 \in \mathcal{I}_\epsilon$. However, a careful newton identities-based analysis would show that even when $g \simeq (\prod_i (1 + \epsilon a_i) - \prod_i (1 + \epsilon b_i))$, for $a_i, b_i \in \mathbb{C}[\epsilon][\boldsymbol{x}]_1$, for some g, then g has small affine border Waring rank Theorem 6.17. Hence, the idea would be to actually reduce the initial form to the above one. This again follows from the *low-rank property*: $\operatorname{rank}(\ell_{i,0} \mid i \in [0,d]) \le 2$. This low-rank property crucially depends on the structure of border Waring rank $= 2$; for details, see Theorem B.1. Hence, by fixing at most 2 linear forms, one can eventually show that the 'reduced' f (of the form $f_{\alpha,\beta}$) has small border Waring rank; One can now interpolate, and show that f has small border Waring rank. The proof is far complicated (and case-analysis dependent), however overall this is the idea.

Proof ideas for lower bounds (Theorem 6.34 & Theorem 6.35). For Theorem 6.34, let $P_2^{[d]} \trianglelefteq_{\mathrm{aff}} P^{[e]}$. If one homogenizes wrt x_0, it requires (both-side padding), and the new formulation becomes $x_0^{e-d} \cdot P_2^{[d]} \trianglelefteq P^{[e]}$. By Theorem 6.20 implies – either (i) $x_0^{e-d} \cdot P_2^{[d]} = \prod_{i\in[e]} \ell_i + \ell_0^e$, for some linear forms $\ell_i \in \mathbb{C}[\boldsymbol{x}]$, or (ii) $\underline{\mathsf{WR}}(x_0^{e-d} \cdot P_2^{[d]}) = O(e^5)$. We show that (i) is an *impossibility*, while (ii) can happen only when $e \ge \exp(d)$. Part (ii) follows because a simple extension of results in [Oed19] implies that $\underline{\mathsf{WR}}(x_0^{e-d} \cdot P_2^{[d]})$ should be exponential. On the other hand, the impossibility result, at least intuitively, follows from the fact that f is 'robust' enough such that, $f \bmod \mathcal{I} \ne 0$, where $\mathcal{I} := \langle \ell_i, \ell_0 \rangle$, for some $i \in [d]$, while trivially the RHS $\prod_{i\in[e]} \ell_i + \ell_0^e \bmod \mathcal{I} = 0$, a contradiction.

For Theorem 6.35, a similar ideal-based argument work, for the impossibility, while the lower bound on e again follows from large border Waring rank. For details, see Section 6.f.

Proof ideas for invariant-theoretic properties of product-plus-powers (Theorem 6.2, Proposition 6.3 & Proposition 6.4).

Theorem 6.2 determines the stabilizer of the polynomial $P_{r,s}^{[d]} = \sum_{i=1}^r \prod_{j=1}^d x_{ji} + \sum_{i=1}^s y_i^d$. This is done in two steps. First, the *identity component* of the stabilizer is computed via Lemma 6.1, a general result on the Lie algebra of stabilizers of sums of polynomials in disjoint sets of variables.

The *discrete part* of the stabilizer is then computed via a geometric argument, studying the action on the singular locus of the hypersurface determined by the polynomial; this is similar to the argument followed in [Lan17, Sec. 6.6].

With the stabilizer, we then determine the representation theoretic multiplicities in the orbit, which is equal to the invariant space dimension $\dim S_\lambda \mathbb{C}^{d+1}$. We split this process into two parts, first determining the $\mathbb{T}^{\mathrm{SL}_d} \times \mathfrak{S}_d$-invariants, and then determining the $\mathbb{Z}_d$-invariants within the invariant space. The proof is finished by an application of Gay's theorem. The polystability of $P^{[d]}$ is obtained using the simple criterion of [BI17], which is based on work by Hilbert, Mumford, Luna, and Kempf. The connection to the Alon-Tarsi conjecture is established by interpreting the evaluation of the invariant as a tensor contraction. We closely mimic the corresponding proof for the product polynomial and observe that the extra summand does not interfere.

Proof idea of Theorem 8.11: De-bordering Waring rank. For the Waring rank, we show a de-bordering result of the form $\mathsf{WR}(f) \leq \exp(\underline{\mathsf{WR}}(f)) \cdot \deg f$. The main ideas for this proof come from apolarity theory and the study of 0-dimensional schemes in projective space, but we provide elementary proofs which do not use the language of algebraic geometry and are based on partial derivative techniques.

To prove the de-bordering, we transform a border Waring rank decomposition for f into a *generalized additive decomposition* [Iar95, BBM14] of the form $f = \sum_{k=1}^{m} \ell_k^{d-r_k+1} g_k$ where ℓ_k are linear forms and g_k are homogeneous polynomials of degrees $r_k - 1$. We then obtain an upper bound on the Waring rank by decomposing each g_k with respect to a basis consisting of powers of linear forms.

To construct a generalized additive decomposition, we divide the summands of a border rank decomposition into several parts such that cancellations only happen between summands belonging to the same part. The key insight is that if the degree of polynomials involved is high enough ($\deg f \geq \underline{\mathsf{WR}}(f) - 1$), then all parts of the decomposition are "local" in the sense that the lowest order term in each summand is a multiple of the same linear form. Each local part gives one summand of the form $\ell^{d-r+1} g$ where r is the number of summands in the part and ℓ is the common lowest order linear form.

Consider for example the family of polynomials $f_d = x_0^{d-1} y_0 + x_1^{d-1} y_1 + 2(x_0 + x_1)^{d-1} y_2$ (This example is adapted from [BB15]). If $d > 3$, then the border Waring rank of f is 6, as evidenced by the decomposition

$$f_d = \lim_{\varepsilon \to 0} \frac{1}{d\varepsilon} \left[(x_0 + \varepsilon y_0)^d - x_0^d + (x_1 + \varepsilon y_1)^d - x_1^d + 2(x_0 + x_1 + \varepsilon y_2)^d - 2(x_0 + x_1)^d \right] \tag{3.1}$$

and a lower bound obtained by considering the dimension of the space of second order partial derivatives. The summands of the decomposition (3.1) are divided into three pairs. The first two summands have x_0 as the lowest order linear form, the second two have x_1, and the last two — $(x_0 + x_1)$. For each pair, the sum individually tends to a limit as $\varepsilon \to 0$, these limits are $x_0^{d-1} y_0$, $x_1^{d-1} y_1$, and $2(x_0 + x_1)^{d-1} y_2$, the summands of a generalized additive decomposition for f.

When $d = 3$, the polynomial f is an example of a "wild form" [BB15]. It has border Waring rank 5 given for example by the decomposition

$$f_3 = \lim_{\varepsilon \to 0} \frac{1}{9\varepsilon} \left[3(x_0 + \varepsilon y_0)^3 + 3(x_1 + \varepsilon y_1)^3 + 6(x_0 + x_1 + \varepsilon y_2)^3 - (x_0 + 2x_1)^3 - (2x_0 + x_3)^3 \right]. \tag{3.2}$$

Unlike the previous decomposition, this one cannot be divided into parts that have limits individually, and is not local — all summands have different lowest order terms. This is only possible if the degree is low.

The condition on the degree is related to algebro-geometric questions about regularity of 0-dimensional schemes [IK99, Thm. 1.69], but for the schemes arising from border rank decompositions this is ultimately a consequence of the fact that r distinct linear forms have linearly independent d-th powers when $d \geq r-1$.

Proof idea for Homogeneous complexity for VF **(Corollary 7.5, Theorem 7.11).** If one generalizes Kumar's complexity to 3×3 matrices, one observes that one gets a structure that is very similar to the proof of Ben-Or and Cleve that describes VF via affine projections of the 3×3 iterated matrix multiplication polynomial. Note that $D_{n,d}$ is the homogeneous degree d part of the $(1,2)$ entry of

$$\begin{pmatrix} 1 & x_{1,1,2} & x_{1,1,3} \\ x_{1,2,1} & 1 & x_{1,2,3} \\ x_{1,3,1} & x_{1,3,2} & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & x_{n,1,2} & x_{n,1,3} \\ x_{n,2,1} & 1 & x_{n,2,3} \\ x_{n,3,1} & x_{n,3,2} & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which makes the connection to Equation (2.2) clear. Pursuing this route, we homogenize Ben-Or & Cleve and get Corollary 7.5. Here we have to pay close attention to avoid the introduction of constants. In particular, we prove a homogeneous version of Brent's depth reduction.

Note that (for odd d) $C_{n,d}$ is the homogeneous degree d part of the $(1,2)$ entry of

$$\begin{pmatrix} 1 & x_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ x_2 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & x_n \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

For Theorem 7.11 we analyze the proof of [BIZ18], but their construction of the product gate seems to be inherently affine. We convert it into an arity 3 homogeneous product gate and then have to deal with arithmetic circuits and formulas that have only arity 3 product gates. This is surprisingly subtle, because we have to disallow constant input gates, so arity 2 product gates cannot be simulated directly. The collection $C_{n,d}$ can be seen as a homogeneous variant of the continuant in [BIZ18]. This gives the $\mathsf{V3F}$-$\overline{\mathsf{p}}$-hardness of $C_{n,d}$.

To see the VQP-$\overline{\mathsf{qp}}$-hardness, we have to show that $\mathsf{V3F}$ and VF coincide when replacing polynomial complexity by quasipolynomial complexity. This is done in two steps: We first show that $\mathsf{VF}_H \subseteq \mathsf{V3P}$, see Theorem 7.13, where we first "parity-homogenize" the formula (every gate has only even or only odd nonzero homogeneous components), and then compute $z \cdot f$ at each even-degree gate instead of f, where z is a new variable. This additional factor z is then later replaced, which is the main reason why the output of this construction is a circuit and not a formula. Since we know that $\mathsf{V3F} \subseteq \mathsf{VF}_H$, we conclude our proof by showing that $\mathsf{VQ3F} = \mathsf{VQ3P}$, for details see Theorem 7.14. It uses an arity-3 basis variant of the Valiant-Skyum-Berkowitz-Rackoff circuit depth reduction [VSBR83], which is a bit more subtle than the original result.

4 Kumar's complexity and border Waring rank

In this section, we prove the results connecting Waring and border Waring rank to Kc-complexity and its variants. First, we record an immediate observation that will be useful throughout:

4.1 Remark. It is easy to observe that

$$\prod_{i=1}^{m}(1+x_i) = \sum_{j=0}^{m} e_j(\mathbf{x})$$

where $\mathbf{x} = (x_1, \ldots, x_m)$ and e_j is the elementary symmetric polynomial of degree j; recall that by definition $e_0 = 1$. In particular, given a homogeneous polynomial $f \in \mathbb{C}[\mathbf{x}]_d$ of degree d, if

$f = \alpha(\prod_{i=1}^m (1+\ell_i) - 1)$ for homogeneous linear forms $\ell_1, \dots, \ell_m$, then

$$\begin{aligned} e_j(\ell_1, \dots, \ell_m) &= 0 \quad \text{for } j \neq d, \\ e_d(\ell_1, \dots, \ell_m) &= \tfrac{1}{\alpha} f. \end{aligned}$$

To demystify $\underline{\mathsf{Kc}}$-complexity, we will often use *Newton identities*, see Proposition 4.2. Let $e_k(x_1, \dots, x_n)$ denotes the k-th **elementary symmetric** polynomial, defined by

$$e_k(x_1, \dots, x_n) \;:=\; \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq n} x_{j_1} \cdots x_{j_k} \;;$$

let $p_k(x_1, \dots, x_n)$, the k-th **power sum** polynomial, defined as $p_k(\boldsymbol{x}) := x_1^k + \cdots + x_n^k$.

4.2 Proposition (Newton Identities, see e.g. [Mac95], Section I.2)**.** *Let n, k be integers with $n \geq k \geq 1$. Then*

$$k \cdot e_k(x_1, \dots, x_n) \;=\; \sum_{i \in [k]} (-1)^{i-1} e_{k-i}(x_1, \dots, x_n) \cdot p_i(x_1, \dots, x_n) \;.$$

As observed in Section 1, the Kc model is not complete. In fact, the only homogeneous polynomials with finite Kc-complexity are powers of linear forms, as the following lemma shows.

4.3 Lemma. *Let $f \in \mathbb{C}[\mathbf{x}]_d$ be a homogeneous polynomial such that $\mathsf{Kc}(f) < \infty$. Then $\mathsf{Kc}(f) = d$ and f is a power of a linear form.*

Proof. If f is a homogeneous polynomial of degree d, then it is immediate that $\mathsf{Kc}(f) \geq \deg(f)$.

Notice that for any linear form ℓ, we have $\ell^d = \prod_{i=1}^d (1 + \zeta^i \ell) - 1$ where ζ is a primitive d-th root of 1. This shows $\mathsf{Kc}(\ell^d) \leq d$, hence equality holds.

Assume $f \in \mathbb{C}[x]_d$ is a homogeneous polynomial with $\mathsf{Kc}(f) = m < \infty$. By definition $f = \alpha\,(\prod_{i=1}^m (1+\ell_i) - 1)$ for some homogeneous linear forms $\ell_i \in \mathbb{C}[\mathbf{x}]$. Write $\boldsymbol{\ell} = (\ell_1, \dots, \ell_m)$. By Remark 4.1, we have, $e_d(\boldsymbol{\ell}) = \frac{1}{\alpha} f$ and $e_j(\boldsymbol{\ell}) = 0$ for $j \neq d$.

First, observe $m = d$. Indeed, if $m > d$, we have $0 = e_m(\boldsymbol{\ell}) = \ell_1 \cdots \ell_m$, which implies $\ell_i = 0$ for some i, in contradiction with the minimality of m. Since $\mathsf{Kc}(f) \geq \deg(f)$, we deduce $m = d$.

Now we show that if $\boldsymbol{\ell} = (\ell_1, \dots, \ell_d)$ satisfies $e_1(\boldsymbol{\ell}) = \cdots = e_{d-1}(\boldsymbol{\ell}) = 0$ then $e_d(\boldsymbol{\ell}) = (-1)^{d-1} \cdot \ell_d^d$; in particular, by unique factorization, all ℓ_i's are equal up to scaling. Write $\widehat{\boldsymbol{\ell}} = (\ell_1, \dots, \ell_{d-1})$. We use induction on j to prove that $e_j(\widehat{\boldsymbol{\ell}}) = (-1)^j \cdot \ell_d^j$ for $j = 1, \dots, d-1$. For $j = 1$, we have

$$0 = e_1(\boldsymbol{\ell}) = (\ell_1 + \cdots + \ell_{d-1}) + \ell_d = e_1(\widehat{\boldsymbol{\ell}}) + \ell_d$$

which proves the statement. For $j = 2, \dots, d-1$, consider the recursive relation

$$e_j(\boldsymbol{\ell}) = e_j(\widehat{\boldsymbol{\ell}}) + \ell_d e_{j-1}(\widehat{\boldsymbol{\ell}}).$$

By assumption we have $e_j(\boldsymbol{\ell}) = 0$ and the induction hypothesis guarantees $e_{j-1}(\widehat{\boldsymbol{\ell}}) = (-1)^{j-1} \cdot \ell_d^{j-1}$; we deduce $e_j(\widehat{\boldsymbol{\ell}}) = -\ell_d \cdot (-1)^{j-1} \cdot \ell_d^{j-1} = (-1)^j \ell_d^j$ which proves the statement.

Finally, notice $f = e_d(\boldsymbol{\ell}) = \ell_d \cdot (-1)^{d-1} \cdot e_{d-1}(\widehat{\boldsymbol{\ell}}) = -\ell_d^d$, which concludes the proof. ☐

However, the model is complete if one allows approximations, as shown by the following result, which appears in [Kum20].

4.4 Proposition (Kumar)**.** *For all homogeneous f we have $\underline{\mathsf{Kc}}(f) \leq \deg(f) \cdot \mathsf{WR}(f)$.*

Proof. The proof is based on a construction by Shpilka [Shp02]. Let $f = \sum_{i=1}^{r} \ell_i^d$. Let ζ be a primitive d-th root of unity. Then one verifies that

$$f = -e_d(-\zeta^0\ell_1, -\zeta^1\ell_1, \ldots, -\zeta^{d-1}\ell_1, \ldots\ldots, -\zeta^0\ell_r, -\zeta^1\ell_1, \ldots, -\zeta^{d-1}\ell_r)$$

and for all $0 < i < d$ we have

$$e_i(-\zeta^0\ell_1, -\zeta^1\ell_1, \ldots, -\zeta^{d-1}\ell_1, \ldots\ldots, -\zeta^0\ell_r, -\zeta^1\ell_1, \ldots, -\zeta^{d-1}\ell_r) = 0.$$

Hence $f \simeq -\epsilon^{-d}\big(((1-\epsilon\zeta^0\ell_1)\cdots(1-\epsilon\zeta^{d-1}\ell_r)) - 1\big)$. Therefore $\underline{\mathsf{Kc}}(f) \le rd$. □

In fact, a slightly stronger statement is true:

4.5 Proposition. *For all homogeneous f we have $\underline{\mathsf{Kc}}(f) \le \deg(f) \cdot \underline{\mathsf{WR}}(f)$.*

Proof. Analogously to the proof in Proposition 4.4, $f \simeq \sum_{i=1}^{r} \ell_i^d = -e_d(-\zeta^0\ell_1, \ldots, -\zeta^{d-1}\ell_r)$. Moreover, for all $0 < i < d$ we have $e_i(-\zeta^0\ell_1, \ldots, -\zeta^{d-1}\ell_r) = 0$. Choose M large enough so that for all $d < i \le dr$ we have that $\epsilon^{-Md}e_i(-\epsilon^M\zeta^0\ell_1, \ldots, -\epsilon^M\zeta^{d-1}\ell_r) \simeq 0$. It follows that $f \simeq -\epsilon^{-Md}\big(((1-\epsilon^M\zeta^0\ell_1)\cdots(1-\epsilon^M\zeta^{d-1}\ell_r)) - 1\big)$. Therefore $\underline{\mathsf{Kc}}(f) \le rd$. □

Proposition 4.4 and Proposition 4.5 show that if $\underline{\mathsf{WR}}(f)$ is small then $\underline{\mathsf{Kc}}(f)$ is small. However, there are polynomials with large Waring (border) rank but small Kumar complexity, such as products of linear forms. Notice $\underline{\mathsf{WR}}(x_1\cdots x_n) = 2^{n-1}$ [Oed19, Corollary 1.9].

4.6 Lemma. *If $f = \ell_1 \cdots \ell_d$ is a product of homogeneous linear forms ℓ_i, then $\underline{\mathsf{Kc}}(f) = d$.*

Proof. The lower bound is immediate because $\underline{\mathsf{Kc}}(f) \ge \deg(f)$. For the upper bound, notice $f \simeq \epsilon^d\big((\prod_{i=1}^{d}(1+\epsilon^{-1}\ell_i)) - 1\big)$. □

The main result of this section is a converse of the above statements. Informally, homogeneous polynomials with small border Waring rank and product of linear forms are the only homogeneous polynomials with small border Kumar complexity. In order to state this precisely, we introduce the following notation. For $f \in \mathbb{C}[\mathbf{x}]_d$, let $\delta_f = 1$ if f is a product of homogeneous linear forms, and define $\delta_f = \infty$ otherwise. The following result explains the relation between border Waring rank and Kumar's complexity.

4.7 Theorem. *For all homogeneous f we have*

$$\min\{\deg(f)\cdot\delta_f,\ \underline{\mathsf{WR}}(f)\} \ \le\ \underline{\mathsf{Kc}}(f) \ \le\ \deg(f)\cdot\min\{\delta_f,\ \underline{\mathsf{WR}}(f)\}.$$

Proof. The right inequality follows from Proposition 4.5 and Lemma 4.6. The left inequality is a combination of Lemma 4.9, Proposition 4.10, and Theorem 4.11 below. □

4.8 Corollary (De-bordering $\underline{\mathsf{Kc}}$). *Let $f \in \mathbb{C}[\mathbf{x}]_d$ be a homogeneous polynomial. If $\underline{\mathsf{Kc}}(f) = m$ then either $\underline{\mathsf{WR}}(f) \le m$, or f is a product of linear forms.*

Proof. By Theorem 4.7, if $\underline{\mathsf{Kc}}(f) = m$ then $\min\{\deg(f)\cdot\delta_f,\ \underline{\mathsf{WR}}(f)\} \le m$. Now, if $\deg(f)\cdot\delta_f \le \underline{\mathsf{WR}}(f)$, then the minimum is $\deg(f)\cdot\delta_f$, which implies $\delta_f \ne \infty$; in this case $\delta_f = 1$, so f is a product of linear forms. Otherwise, $\underline{\mathsf{WR}}(f)$ is the minimum, which implies that $\underline{\mathsf{WR}}(f) \le m$. □

Note that in the definition of $\underline{\mathsf{Kc}}$, the factor α can be assumed to be a scalar times a power of ϵ, because only the lowest power of ϵ in α would contribute to the limit. We distinguish three cases, depending on the sign of the exponent of ϵ in α.

- $\underline{\mathsf{Kc}}^{+}(f)$ is the smallest m such that $f \simeq \gamma\epsilon^N\big(\prod_{i=1}^m(1+\ell_i)-1\big)$ for some $N \geq 1$, $\gamma \in \mathbb{C}$ and $\ell_i \in \mathbb{C}[\epsilon^{\pm 1}][\mathbf{x}]_1$; set $\underline{\mathsf{Kc}}^{+}(f) = \infty$ is such an m does not exist;
- $\underline{\mathsf{Kc}}^{-}(f)$ is the smallest m such that $f \simeq \gamma\epsilon^{-M}\big(\prod_{i=1}^m(1+\ell_i)-1\big)$ for some $M \geq 1$, $\gamma \in \mathbb{C}$ and $\ell_i \in \mathbb{C}[\epsilon^{\pm 1}][\mathbf{x}]_1$; set $\underline{\mathsf{Kc}}^{-}(f) = \infty$ is such an m does not exist;
- $\underline{\mathsf{Kc}}^{=}(f)$ is the smallest m such that $f \simeq \gamma\big(\prod_{i=1}^m(1+\ell_i)-1\big)$ for some $\gamma \in \mathbb{C}$ and $\ell_i \in \mathbb{C}[\epsilon^{\pm 1}][\mathbf{x}]_1$; set $\underline{\mathsf{Kc}}^{=}(f) = \infty$ is such an m does not exist.

We observe that $\underline{\mathsf{Kc}}(f) = \min\big\{\underline{\mathsf{Kc}}^{+}(f),\, \underline{\mathsf{Kc}}^{=}(f),\, \underline{\mathsf{Kc}}^{-}(f)\big\}$.

4.9 Lemma. *For all homogeneous f we have* $\deg(f)\cdot\delta_f \leq \underline{\mathsf{Kc}}^{+}(f)$.

Proof. Let $d := \deg(f)$. The lower bound $\deg(f) \leq \underline{\mathsf{Kc}}^{+}(f)$ is clear. Therefore, it suffices to show that if $\underline{\mathsf{Kc}}^{+}(f)$ is finite, then f is a product of linear forms. Let $f \simeq= \gamma\epsilon^N\big(\prod_{i=1}^m(1+\ell_i)-1\big)$ with $N \geq 1$. Since $\epsilon^N \simeq 0$, we have $f \simeq \gamma\epsilon^N\prod_{i=1}^m(1+\ell_i)$, namely f is limit of a product of affine linear polynomials. The property of being completely reducible is closed, therefore we deduce that f is a product of affine linear polynomials. Since f is homogeneous, its factors are homogeneous as well. This shows $\delta_f = 1$ and the statement follows. □

4.10 Proposition (Newton Identities)**.** *For all homogeneous f we have* $\mathsf{WR}(f) \leq \underline{\mathsf{Kc}}^{=}(f)$.

Proof. Let $d := \deg(f)$. Suppose $\underline{\mathsf{Kc}}^{=}(f) = m$ and write $f \simeq f_\epsilon := \gamma\big(\prod_{i=1}^m(1+\ell_i)-1\big)$. One can verify that if even one of the ℓ_i diverges, then the j-th homogeneous part of f_ϵ diverges, where j is the number of diverging ℓ_i. Hence all ℓ_i converge and we set ϵ to zero. Hence, $\underline{\mathsf{Kc}}^{=}(f) = \mathsf{Kc}(f)$. Now, since f is homogeneous, each homogeneous degree i part of f_ϵ vanishes, $i < d$. In other words, $e_i(\boldsymbol{\ell}) = 0$ for all $1 \leq i < d$, where $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_m)$. Hence $s(\boldsymbol{\ell}) = 0$ for *all* symmetric polynomials of degree $< d$. Therefore the Newton identity $p_d = (-1)^{d-1}\cdot d\cdot e_d + \sum_{i=1}^{d-1}(-1)^{d+i-1}e_{d-i}\cdot p_i$ gives that $e_d(\boldsymbol{\ell})$ and $p_d(\boldsymbol{\ell})$ are same up to multiplication by a scalar. Hence $\mathsf{WR}(f) \leq m$. □

4.11 Theorem (Approximate Newton Identities)**.** *For all homogeneous f:* $\underline{\mathsf{WR}}(f) \leq \underline{\mathsf{Kc}}^{-}(f)$.

Proof. Let $d := \deg(f)$. Let $f \simeq f_\epsilon := \gamma\epsilon^{-M}\big(\prod_{i=1}^m(1+\ell'_i)-1\big)$ with $M \geq 1$. From the convergence of f_ϵ we deduce that for each i we have $\ell'_i = \epsilon\ell_i$ with $\ell_i \in \mathbb{C}[\epsilon][x]_1$, because otherwise the homogeneous degree j part diverges, where j is the number of ℓ'_i that do not satisfy this property.

Now, let $f_{\epsilon,j}$ denote the homogeneous degree j part of f_ϵ. Since f is homogeneous of degree d, for $0 \leq j < d$ we have $f_{\epsilon,j} \simeq 0$. By expanding the product, observe that for all $0 < j < d$ we have $0 \simeq f_{\epsilon,j} = \gamma\epsilon^{-M}e_j(\epsilon\ell_1, \ldots, \epsilon\ell_m) = \gamma\epsilon^{-M+j}e_j(\ell_1, \ldots, \ell_m)$. We now show by induction that for all $1 \leq j < d$ we have $\epsilon^{-M+j}p_j(\ell_1, \ldots, \ell_m) \simeq 0$. This is clear for $j = 1$, because $p_1 = e_1$. For the step from j to $j+1$ we use Newton's identities:

$$p_{j+1} \;=\; (-1)^j\,(j+1)\,e_{j+1} + \textstyle\sum_{i=1}^{j}(-1)^{j+i}e_{j+1-i}\cdot p_i.$$

Hence $\epsilon^{-M+(j+1)}p_{j+1}(\boldsymbol{\ell})$

$$= (-1)^j\,(j+1)\,\underbrace{\epsilon^{-M+(j+1)}e_{j+1}(\boldsymbol{\ell})}_{\simeq 0} + \sum_{i=1}^{j}(-1)^{j+i}\,\underbrace{\epsilon^{-M+(j+1)-i}e_{j+1-i}(\boldsymbol{\ell})}_{\simeq 0}\cdot\underbrace{\epsilon^{M}}_{\simeq 0}\cdot\underbrace{\epsilon^{-M+i}p_i(\boldsymbol{\ell})}_{\simeq 0} \simeq 0.$$

This finishes the induction proof, but we use Newton's identities again in the same way to see that $\epsilon^{-M+d}p_d(\boldsymbol{\ell}) \simeq (-1)^{d-1} \cdot d \cdot \epsilon^{-M+d}e_d(\boldsymbol{\ell})$:

$$\epsilon^{-M+d}p_d(\boldsymbol{\ell}) = (-1)^{d-1} \cdot d \cdot \epsilon^{-M+d}e_d(\boldsymbol{\ell}) + \sum_{i=1}^{d-1}(-1)^{d-1+i}\underbrace{\epsilon^{-M+d-i}e_{d-i}(\boldsymbol{\ell})}_{\simeq 0} \cdot \underbrace{\epsilon^{M}}_{\simeq 0} \cdot \underbrace{\epsilon^{-M+i}p_i(\boldsymbol{\ell})}_{\simeq 0}.$$

We are done now, because $f \simeq f_{\epsilon,d} = \gamma\epsilon^{-M+d}e_d(\ell_1,\ldots,\ell_m) \simeq \gamma\epsilon^{-M+d} \cdot \frac{1}{d} \cdot (-1)^{d-1}p_d(\ell_1,\ldots,\ell_m)$ and hence $\underline{\mathsf{WR}}(f) \leq m$. □

4.a Linear approximations and Waring rank

We demonstrated the inequality $\underline{\mathsf{Kc}}(f) \leq \deg(f) \cdot \mathsf{WR}(f)$ in Proposition 4.4. In the proof of Proposition 4.4, only "linear approximations" have been used; we prove here a converse of Proposition 4.4 in the restricted setting of linear approximation. Given a homogeneous polynomial $f \in \mathbb{C}[\mathbf{x}]_d$, let $\underline{\mathsf{Kc}_1^-}(f)$ be the smallest m such that there exist linear forms $\ell_1,\ldots,\ell_m \in \mathbb{C}[\mathbf{x}]_1$ and $M \geq 1$ such that $f \simeq \gamma\epsilon^{-M}\big(\prod_{i=1}^m(1+\epsilon\ell_i) - 1\big)$.

4.12 Proposition. *For any homogeneous polynomial f of degree d, we have* $\mathsf{WR}(f) \leq \underline{\mathsf{Kc}_1^-}(f) \leq d \cdot \mathsf{WR}(f)$.

Proof. The inequality $\underline{\mathsf{Kc}_1^-}(f) \leq d \cdot \mathsf{WR}(f)$ is clear from the proof of Proposition 4.4, as there we obtained an expression of the form described in the definition of $\underline{\mathsf{Kc}_1^-}$. Suppose $\underline{\mathsf{Kc}_1^-}(f) = m$ and write $f \simeq f_\epsilon := \gamma\epsilon^{-M}\big(\prod_{i=1}^m(1+\epsilon\ell_i) - 1\big)$ with $M \geq 1$ and $\ell_i \in \mathbb{C}[\boldsymbol{x}]_1$. It is immediate that $m \geq M$, $f = \gamma e_M(\boldsymbol{\ell})$ and $e_j(\boldsymbol{\ell}) = 0$ for $j < M$, where $\boldsymbol{\ell} = (\ell_1,\ldots,\ell_m)$. Via the Newton identity for the power sum polynomial, we have

$$p_M(\boldsymbol{\ell}) = (-1)^{M-1}Me_M(\boldsymbol{\ell}) + \sum_{i=1}^{M-1}(-1)^{M+i-1}e_{M-i}(\boldsymbol{\ell}) \cdot p_i(\boldsymbol{\ell}).$$

Since $e_j(\boldsymbol{\ell}) = 0$ for all $1 \leq j < M$, we obtain:

$$p_M(\boldsymbol{\ell}) = (-1)^{M-1}Me_M(\boldsymbol{\ell}) = \frac{1}{\gamma}(-1)^{M-1}Mf.$$

We conclude $\mathsf{WR}(f) = \mathsf{WR}(p_M(\boldsymbol{\ell})) \leq \mathsf{WR}(p_M) = m = \underline{\mathsf{Kc}_1^-}(f)$, as desired. □

5 Preliminaries

Notation. Let $V = \langle x_1,\ldots,x_n\rangle$ be the space of linear forms in $x_1,\ldots,x_n$. Let $S^dV = \mathbb{C}[x_1,\ldots,x_n]_d$ and let $f \in S^dV$ be a homogeneous polynomial of degree d. We denote $[n] = \{1,\cdots,n\}$, and $[a,b] = \{a,a+1,\cdots,b\}$, for $a,b \in \mathbb{N}_{\geq 0}$.

$\langle\ell_1,\cdots,\ell_k\rangle$ is the ideal generated by k-many linear polynomials, i.e., $g \in \langle\ell_1,\cdots,\ell_k\rangle$, must be of the form $g = \sum_{i\in[k]} g_i\ell_i$, for some $g_i \in \mathbb{C}[\boldsymbol{x}]$.

Waring and border Waring rank. For $f \in S^dV$, denote by $\mathsf{WR}(f)$ the *Waring rank* of f, that is the *minimum* r in an expression $f = \sum_{i=1}^r \ell_i^d$, for linear forms $\ell_i \in \mathbb{C}[\boldsymbol{x}]_1$. Denote by $\underline{\mathsf{WR}}(f)$ the *border Waring rank* of f, that is the minimum r such that $f = \lim_{\epsilon\to 0}\sum_{i=1}^r \ell_i^d$, where $\ell_i \in \mathbb{C}[\epsilon^{\pm 1}][\boldsymbol{x}]_1$ are linear forms depending (rationally) on ϵ.

Given a polynomial $f \in \mathbb{C}[\mathbf{x}]$ (not necessarily homogeneous), the *affine Waring rank* of f, denoted $\mathsf{WR}_{\mathsf{aff}}(f)$, is the *minimum* r in an expression $f = \sum_{i\in[r]} \ell_i^{e_i}$, for affine linear polynomials ℓ_i, and $e_i \in \mathbb{N}$. In this case, write $f \in \Sigma^{[r]} \wedge \Sigma$, with maximum exponent $e = \max_{i\in[r]} e_i$. The affine *border Waring rank* of f, denoted $\underline{\mathsf{WR}}_{\mathsf{aff}}(f)$ is the *minimum* r such that $f = \lim_{\epsilon\to 0} \sum_{i=1}^{r} \ell_i^{e_i}$, where ℓ_i, are affine linear polynomials with coefficients in $\mathbb{C}[\epsilon^{\pm 1}]$ and $e_i \in \mathbb{N}$. In this case, write $f \in \overline{\Sigma^{[r]} \wedge \Sigma}$ with maximum exponent $e = \max_{i\in[r]} e_i$.

Essential variables. We define the concept of essential variables below, which is an important concept, both in algebraic geometry and GCT.

5.1 Definition (Essential variables)**.** *Given $f \in \mathbb{C}[\boldsymbol{x}]_d$, we call the* essential number of variables *of f the smallest integer r for which there exists a set of linear forms $\ell_1, \dots, \ell_r$ such that $f \in \mathbb{C}[\ell_1, \dots, \ell_r]$.*

It is a classical fact, which essentially already appears in [Syl52], that the first catalecticant matrix controls the number of essential variables of a homogeneous polynomial. We refer to [Car06] and [KS07, Lemma B.1] for modern proofs of this result.

5.2 Theorem. *Let $f \in \mathbb{C}[x_1, \dots, x_n]_d$, and $T = \mathbb{C}[\partial_1, \dots, \partial_n]$ denote the ring of differential operators. Then,*

$$N_{ess}(f) = \mathrm{rank}(C_f)\ ,$$

i.e. the number of essential variables of f is the rank of its first catalecticant matrix, and

$$EssVar(f) \;=\; \langle D \circ f \mid D \in T_{d-1} \rangle\ ,$$

i.e. the essential variables of f span the space of its $(d-1)$-th partial derivatives.

6 Orbit closure of restricted binomials

Fix integers d, r, s, with $d \geq 3$, and consider the polynomial

$$P_{r,s}^{[d]} = \sum_{i=1}^{r} \prod_{j=1}^{d} x_{ji} + \sum_{i=1}^{s} y_i^d,$$

in the $rd + s$ variables $x_{11}, \dots, x_{dr}, y_1, \dots, y_s$. In the special case $r = s = 1$, the polynomial $P^{[d]} = P_{1,1}^{[d]}$ coincides with the restricted binomial defining Kumar's complexity in Section 2.b.

In this section, we study computational and invariant theoretic properties of $P_{r,s}^{[d]}$. Theorem 6.2 determines the stabilizer of $P_{r,s}^{[d]}$ under the action of the group GL_{rd+s} acting on the variables. The knowledge of the stabilizer, allows us to determine the representation theoretic structure of the coordinate ring of the orbit of $P^{[d]}$, which is achieved in Proposition 6.3. In Proposition 6.4, we prove that $P_{r,s}^{[d]}$ is polystable, in the sense of invariant theory. This guarantees the existence of a *fundamental invariant*, in the sense of [BI17]: in Proposition 6.6, we show a connection between the degree of this fundamental invariant and the Alon-Tarsi conjecture on Latin squares in combinatorics.

Section 6.e focuses on complexity results for $P^{[d]}$ and $P_{1,2}^{[d]}$. These can be regarded as generalizations of Theorem 4.7. Finally, in Section 6.f we determine lower bounds for $P^{[d]}$ in certain computational models.

6.a Stabilizer

The general linear group GL_n acts on $\mathbb{C}[x_1, \dots, x_n]$ by linear change of variables as described in Section 1. For a homogeneous polynomial $f \in \mathbb{C}[\mathbf{x}]_d$, write $\mathrm{Stab}_{\mathrm{GL}_n}(f)$ for its stabilizer under this action. It is an immediate fact that $\mathrm{Stab}_{\mathrm{GL}_n}(f)$ is a closed algebraic subgroup of GL_n. It may consists of several connected (irreducible) components: the *identity component*, denoted $\mathrm{Stab}^0_{\mathrm{GL}_n}(f)$ is the connected component containing the identity; $\mathrm{Stab}^0_{\mathrm{GL}_n}(f)$ is a closed, normal subgroup of $\mathrm{Stab}_{\mathrm{GL}_n}(f)$ [Ges16, Lemma 2.1]; the quotient $\mathrm{Stab}_{\mathrm{GL}_n}(f) / \mathrm{Stab}^0_{\mathrm{GL}_n}(f)$ is a finite group.

The Lie algebra $\mathfrak{g}$ of an algebraic group G can be geometrically identified with the tangent space to G at the identity element. Moreover, if G is a subgroup of GL_n, then $\mathfrak{g}$ is naturally a subalgebra of $\mathfrak{gl}_n = \mathrm{End}(\mathbb{C}^n)$; moreover $\mathfrak{g}$ uniquely determined the identity component of G.

It is a classical fact that the Lie algebra of $\mathrm{Stab}_{\mathrm{GL}_n}(f)$ is the annihilator of f under the Lie algebra action of $\mathfrak{gl}_n$ on $\mathbb{C}[\mathbf{x}]_d$; denote this annhilator by $\mathfrak{ann}_{\mathfrak{gl}_n}(f)$. Typically, in order to determine $\mathrm{Stab}_{\mathrm{GL}_n}(f)$, one first computes $\mathfrak{ann}_{\mathfrak{gl}_n}(f)$, which uniquely determines $\mathrm{Stab}^0_{\mathrm{GL}_n}(f)$. Then, one determines $\mathrm{Stab}_{\mathrm{GL}_n}(f)$ as a subgroup of the normalizer $N_{\mathrm{GL}_n} \mathrm{Stab}^0_{\mathrm{GL}_n}(f)$.

First, we record a general result regarding the stabilizer of sums of polynomials in disjoint sets of variables. This is the symmetric version of [CGL$^+$21, Thm. 4.1(i)].

6.1 Lemma. *Let $V = V_1 \oplus V_2$ and let $f \in \mathbb{C}[V^*]_d = S^d V$ be a homogeneous polynomial with $f = f_1 + f_2$, where $f_i \in S^d V_i$ are both concise, with $d \geq 3$. Then*

- *(i)* $\mathfrak{ann}_{\mathfrak{gl}(V)}(f_1) = \mathfrak{ann}_{\mathfrak{gl}(V_1)}(f_1) \oplus \mathrm{Hom}(V_2, V)$;
- *(ii)* $\mathfrak{ann}_{\mathfrak{gl}(V)}(f_1 + f_2) = \mathfrak{ann}_{\mathfrak{gl}(V_1)}(f_1) \oplus \mathfrak{ann}_{\mathfrak{gl}(V_2)}(f_2)$.

Proof. For both statements, the inclusion of the right-hand term into the left-hand term is clear. We prove the reverse inclusion.

For $X \in \mathfrak{gl}(V)$, write $X = \sum_{i,j=1}^2 X_{ij}$, with $X_{ij} \in \mathrm{Hom}(V_i, V_j)$.

The proof of (i) amounts to showing that if $X \in \mathfrak{ann}_{\mathfrak{gl}(V)}(f_1)$, then $X_{12} = 0$ and $X_{11} \in \mathfrak{ann}_{\mathfrak{gl}(V_1)}(f_1)$. Suppose $X.f_1 = 0$. Notice $X.f_1 = X_{11}.f_1 + X_{12}.f_1$; here $X_{11}.f_1 \in S^d V_1$ and $X_{12}.f_1 \in V_2 \otimes S^{d-1} V_1$. In particular, both terms must vanish. The term $X_{12}.f_1$ is a sum of at most $\dim V_2$ linearly independent elements, each of which is a linear combination of first order partials of f_1. Since f_1 is concise, $X_{12}.f_1 = 0$ if and only if $X_{12} = 0$. The condition $X_{11}.f_1 = 0$ is, by definition, equivalent to $X_{11} \in \mathfrak{ann}_{\mathfrak{gl}(V_1)}(f_1)$. This conclude the proof of (i).

To prove (ii), we show that if $X \in \mathfrak{ann}_{\mathfrak{gl}(V)}(f)$, then $X_{12} = 0$,$X_{21} = 0$ and $X_{ii} \in \mathfrak{ann}_{\mathfrak{gl}(V_i)}(f_i)$. Suppose $X.f = 0$. We have $X.f = (X_{11} + X_{12}).f_1 + (X_{21} + X_{22}).f_2$. Now, $(X_{11} + X_{12}).f_1 \in S^d V_1 \oplus V_2 \otimes S^{d-1} V_1$, and similarly $(X_{21} + X_{22}).f_2$. Since $d \geq 3$, the two terms are linearly independent, hence they both must vanish. This shows $(X_{11} + X_{12}) \in \mathfrak{ann}_{\mathfrak{gl}(V)}(f_1)$, therefore $X_{12} = 0$ and $X_{11} \in \mathfrak{ann}_{\mathfrak{gl}(V_1)}(f_1)$ from the previous part of the proof. The analogous condition holds for X_{21} and X_{22} and this completes the proof. ☐

We can now determine the stabilizer of $P^{[d]}_{r,s}$. Let $\mathbb{T}^{\mathrm{SL}_n}$ denote the subgroup of diagonal elements in SL_n.

6.2 Theorem. *For $d \geq 3$ and for every r, s, we have*

$$\mathrm{Stab}_{\mathrm{GL}(V)}(P^{[d]}_{r,s}) = ([\mathbb{T}^{\mathrm{SL}_d} \rtimes \mathfrak{S}_d] \wr \mathfrak{S}_r) \times (\mathbb{Z}_d \wr \mathfrak{S}_s);$$

each copy of $\mathbb{T}^{\mathrm{SL}_d} \rtimes \mathfrak{S}_d$ acts by rescaling and permuting the variables in one of the r sets $\{x_{ji} : i = 1, \dots, d\}$ for $j = 1, \dots, r$; the group $\mathfrak{S}_r$ permutes (set-wise) these sets; the group $\mathbb{Z}_d \wr \mathfrak{S}_s$ acts by rescaling (by a d-th root of 1) and permuting the variables in the set $\{y_i : i = 1, \dots, s\}$.

Proof. It is clear that the group on the right-hand side is contained in the stabilizer $\mathrm{Stab}_{\mathrm{GL}(V)}(P_{r,s}^{[d]})$. We show the reverse inclusion.

First, we determine the identity component of $\mathrm{Stab}_{\mathrm{GL}(V)}(P_{r,s}^{[d]})$. By Lemma 6.1, the annihilator of $P_{r,s}^{[d]}$ in $\mathfrak{gl}(V)$ is the direct sum of the annihilators of its summands. This guarantees that the identity component of the stabilizer of $P_{r,s}^{[d]}$ is $\mathrm{Stab}^0_{\mathrm{GL}(V)}(P_{r,s}^{[d]}) = (\mathbb{T}^{\mathrm{SL}_d})^{\times r}$, where the j-th copy of $\mathbb{T}^{\mathrm{SL}_d}$ acts by rescaling the variables $x_{1j}, \ldots, x_{dj}$; see, e.g., [Lan17, Sec. 7.1.2].

Since $\mathrm{Stab}^0_{\mathrm{GL}(V)}(P_{r,s}^{[d]})$ is a normal subgroup of $\mathrm{Stab}_{\mathrm{GL}(V)}(P_{r,s}^{[d]})$, we have

$$\mathrm{Stab}_{\mathrm{GL}(V)}(P_{r,s}^{[d]}) \subseteq N_{\mathrm{GL}(V)}(\mathbb{T}^{\mathrm{SL}_d \times r}) = ([\mathbb{T}^{\mathrm{SL}_d} \rtimes \mathfrak{S}_d] \wr \mathfrak{S}_r) \rtimes Q$$

where Q is the parabolic subgroup stabilizing the subspace spanned by the x_{ij} variables.

In order to determine the discrete component, we follow the same argument as the one used for the power sum polynomial $P_{0,s}^{[d]}$ in [Lan17, Section 8.12.1]. In particular, $\mathrm{Stab}_{\mathrm{GL}(V)}(P_{r,s}^{[d]})$ stabilizes the Hessian determinant of $P_{r,s}^{[d]}$, up to scaling. A direct calculation shows that this Hessian determinant, up to scaling, is

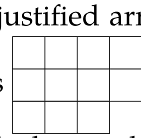
$$H = (\prod_{i,j} x_{ij} \prod_k y_k)^{d-2}.$$

Unique factorization implies that $\mathrm{Stab}_{\mathrm{GL}(V)}(P_{r,s}^{[d]}) \cap Q \subseteq \mathbb{T} \rtimes \mathfrak{S}_s$, where $\mathbb{T}$ is the torus of diagonal matrices acting on the y_j variables. Hence this subgroup commutes with $[\mathbb{T}^{\mathrm{SL}_d} \rtimes \mathfrak{S}_d] \wr \mathfrak{S}_r$ and we deduce

$$\mathrm{Stab}_{\mathrm{GL}(V)}(P_{r,s}^{[d]}) \cap Q = \mathrm{Stab}_{\mathrm{GL}_s}(y_1^d + \cdots + y_s^d) = \mathbb{Z}_d \wr \mathfrak{S}_s.$$

This concludes the proof. $\square$

6.b Multiplicities in the coordinate ring of the orbit

A *partition* $\lambda = (\lambda_1, \lambda_2, \ldots)$ is a finite nonincreasing sequence of nonnegative integers. We write $\ell(\lambda) := \max\{i \mid \lambda_i \neq 0\}$, and $\lambda \vdash D$ means $\sum_i \lambda_i = D$. To each partition λ we associate its Young diagram, which is a top-left justified array of boxes with λ_i boxes in row i. For example, the Young diagram to $\lambda = (4, 4, 3)$ is . The transpose of the Young diagram is obtained by switching rows and columns. We denote the partition corresponding to this Young diagram by λ^t, for example $(4, 4, 3)^t = (3, 3, 3, 2)$. A group homomorphism $\varrho : \mathrm{GL}_D \to \mathrm{GL}(V)$, where V is a finite dimensional complex vector space, is called a *representation* of GL_D. A representation is *polynomial* if each entry of the matrix corresponding to the linear map $\varrho(g)$ is given by a polynomial in the entries of GL_D. A linear subspace that is closed under the group operation is called a *subrepresentation*. A representation with only the two trivial subrepresentations is called *irreducible*. The irreducible polynomial representations of GL_{d+1} are indexed by partitions λ with $\ell(\lambda) \leq d+1$, see for example [Ful97, Ch. 8]. We denote by $S_\lambda(\mathbb{C}^{d+1})$ the irreducible representation of type λ. For a GL_{d+1}-representation $\mathcal{V}$ we write $\mathrm{mult}_\lambda(\mathcal{V})$ to denote the multiplicity of λ in $\mathcal{V}$, i.e., the dimension of the space of equivariant maps from $S_\lambda(\mathbb{C}^{d+1})$ to $\mathcal{V}$, or equivalently, the number of summands of isomorphism type λ in any decomposition of $\mathcal{V}$ into a direct sum of irreducible representations.

In this section we care about the special case $r = s = 1$ (which is the homogenization of Kumar's case, see Section 2.b, and we set $P^{[d]} := P_{1,1}^{[d]}$, and $G := \mathrm{GL}_{d+1}$. We now use the stabilizer to determine the multiplicities in the coordinate ring of the group orbit $\mathrm{mult}_\lambda(\mathbb{C}[\mathrm{GL}_{d+1} P^{[d]}])$. Let $H := \mathrm{Stab}_G(P^{[d]}) \simeq \mathbb{Z}_d \times (\mathbb{T}^{\mathrm{SL}_d} \rtimes \mathfrak{S}_d)$. A standard consideration in GCT is that since H

is reductive, the orbit $GP^{[d]}$ is an affine variety ([BLMW11, §4.2], [Mat60]) and a homogeneous space that is isomorphic to the quotient G/H. Its coordinate ring is determined by the Algebraic Peter-Weyl Theorem [GW09, Thm. 4.2.7]: we have $\mathbb{C}[GP^{[d]}] \simeq \mathbb{C}[G/H] \simeq \mathbb{C}[G]^H$, and therefore $\mathrm{mult}_\lambda(\mathbb{C}[GP^{[d]}]) = \dim(S_\lambda V)^H$. We show how this invariant space dimension can be determined by classical representation branching rules in Proposition 6.3.

For partitions μ and λ we define $\mu \preceq \lambda$ iff $\mu \subseteq \lambda$ (i.e. $\forall i : \mu_i \leq \lambda_i$) and the skew diagram λ/μ has at most 1 box in each column (i.e., $\lambda_i^t - \mu_i^t \leq 1$). Let $a_\mu(d, D) := \mathrm{mult}_\mu(S^d(S^D(W)))$ for any W of dimension at least d, sometimes called the plethysm coefficient.

6.3 Proposition. *For $\lambda \vdash dD$ we have*

$$\mathrm{mult}_\lambda(\mathbb{C}[\mathrm{GL}_{d+1} P^{[d]}]) = \dim(S_\lambda \mathbb{C}^{d+1})^H = \sum_{\delta=0}^{D} \sum_{\substack{\mu \vdash \delta d \\ \mu \preceq \lambda \\ \ell(\mu) \leq d}} a_\mu(d, \delta).$$

Proof.

$$(S_\lambda \mathbb{C}^{d+1})^H = (S_\lambda(\mathbb{C} \oplus \mathbb{C}^d) \downarrow^{\mathrm{GL}_{d+1}}_{\mathrm{GL}_1 \times \mathrm{GL}_d})^{\mathbb{Z}_d \times (\mathbb{T}^{\mathrm{SL}_d} \rtimes \mathfrak{S}_d)} \overset{\text{Pieri's rule}}{=} \bigoplus_{\substack{\mu \preceq \lambda \\ \ell(\mu) \leq d}} (S_{|\lambda| - |\mu|} \mathbb{C}^1)^{\mathbb{Z}_d} \otimes (S_\mu \mathbb{C}^d)^{\mathbb{T}^{\mathrm{SL}_d} \rtimes \mathfrak{S}_d},$$

where Pieri's rule is a well-known decomposition rule, see for example [FH91, p. 80, Exe. 6.12]. Now, $\dim((S_{|\lambda|-|\mu|}\mathbb{C}^1)^{\mathbb{Z}_d}) = 1$ iff $|\lambda| - |\mu|$ is a multiple of d iff $|\mu|$ is a multiple of d. Otherwise it is 0. Hence

$$\dim(S_\lambda V)^H = \sum_{\delta=0}^{d} \sum_{\substack{\mu \vdash \delta d \\ \mu \preceq \lambda \\ \ell(\mu) \leq d}} \underbrace{\dim(S_\mu \mathbb{C}^d)^{\mathbb{T}^{\mathrm{SL}_d} \rtimes \mathfrak{S}_d}}_{= a_\mu(d,\delta)}$$

The last underbrace equality is Gay's theorem [Gay76]. □

Note that the $\ell(\mu) \leq d$ requirement is not actually necessary, because if $\ell(\mu) > d$, then $a_\mu(d, \delta) = 0$.

A computer calculation (see appendix) shows that this indeed gives multiplicity obstructions. We used the HWV software [BHIM22] to directly calculate that (10, 6, 4, 4) and (8, 8, 4, 4) are the only types in the vanishing ideal for $D = 8$, $d = 3$. For $d = 3$ there are no equations in degree $1, \ldots, 7$. In particular, none of Brill's equations (which all are of degree $d + 1$) vanishes on $\mathrm{GL}_{d+1} P^{[d]} \cap S^d \mathbb{C}^d$.

6.c Polystability

A polynomial $f \in S^d V$ is called *polystable* if its $\mathrm{SL}(V)$-orbit is closed. Polystability is an important property in GCT, as it implies the existence of a *fundamental invariant* that connects the GL-orbit with the GL-orbit closure, see [BI17, Def. 3.9 and Prop. 3.10]. This connection can be used to exhibit multiplicity obstructions, as was done in [IK20].

6.4 Proposition. *Let $d \geq 2$. The polynomial $P^{[d]}_{r,s}$ is polystable, i.e., the orbit $\mathrm{SL}(V) P^{[d]}$ is closed.*

Proof. If $d = 2$, then $P^{[2]}_{r,s}$ is a polynomial of degree 2 defining a quadratic form of maximal rank. This is polystable.

Suppose $d \geq 3$. Proposition 2.8 in [BI17] gives a criterion for polystability, based on works of Hilbert, Mumford, Luna, and Kempf.

In order to apply this criterion, consider the group $R = \mathrm{Stab}_{\mathrm{GL}(V)}(P^{[d]}_{r,s}) \cap \mathbb{T}^{\mathrm{GL}(V)}$, where $\mathbb{T}^{\mathrm{GL}(V)}$ denotes the torus of diagonal matrices, in the basis defined by the variables. By Theorem 6.2, we deduce $R = (\mathbb{T}^{\mathrm{SL}_d})^{\times r} \times \mathbb{Z}_d^{\times s}$. This is a group consisting entirely of diagonal matrices and it is easy to verify that its centralizer in $\mathrm{SL}(V)$ coincides with $\mathbb{T}^{\mathrm{SL}(V)}$. This proves the first property of the criterion.

For the second property, consider the exponent vectors of the monomials appearing in $P^{[d]}_{r,s}$. For a monomial m, write $\mathrm{wt}(m)$ for its exponent vector. It is immediate to verify that

$$\sum_{i=1}^{r} \mathrm{wt}(x_{i1} \cdots x_{id}) + \frac{1}{d} \sum_{j=1}^{s} \mathrm{wt}(y_j^d) = (1, \ldots, 1);$$

this shows that the vector $(1, \ldots, 1)$ lies in the convex cone generated by the exponent vectors of the monomials of $P^{[d]}_{r,s}$. This proves the second part of the criterion and concludes the proof. □

Proposition 6.4 reduces to the following in the special case $r = s = 1$:

6.5 Corollary. *Let $d \geq 2$. The polynomial $P^{[d]}$ is polystable, i.e., the orbit $\mathrm{SL}_{d+1} P^{[d]}$ is closed.*

6.d Fundamental invariants and the Alon-Tarsi conjecture

The *fundamental invariant* Φ of a polystable polynomial $f \in S^D V$ is the smallest degree $\mathrm{SL}(V)$-invariant function in $\mathbb{C}[\overline{\mathrm{GL}(V)f}]$, see Def. 3.8 in [BI17]. It describes the connection between the orbit and the orbit-closure of f: more formally $\mathbb{C}[\overline{\mathrm{GL}(V)f}]_\Phi \simeq \mathbb{C}[\mathrm{GL}(V)f]$ is the localization at Φ, see [BI17, Pro. 3.9]. This connection can be used to exhibit multiplicity obstructions, as was done in [IK20].

It is known that for even d the orbit closure $\overline{\mathrm{GL}_d(x_1 \cdots x_d)}$ has fundamental invariant of degree d if and only if the Alon-Tarsi conjecture on Latin squares holds for d, see [BI17, Pro. 3.26]; otherwise the fundamental invariant has higher degree. In this section we show an analogous result for the orbit closure $\overline{\mathrm{GL}_{d+1}(x_1 \cdots x_d + x_{d+1}^d)}$: if d is even this orbit closure has fundamental invariant of degree $d+1$ if and only if the Alon-Tarsi conjecture on Latin squares holds for d; otherwise the fundamental invariant has higher degree.

6.6 Proposition. *Let d be even. The degree of the fundamental invariant of $P^{[d]}$ is $d+1$ if and only if the Alon-Tarsi conjecture for d is true, otherwise it is in a higher degree.*

Proof. We follow the presentation in [CIM17, BI17, BDI21]. For a partition λ we place positive integers into the boxes of the Young diagram and call it a *tableau* T of shape λ. The vector of numbers of occurrences of 1s, 2s, etc, is called the *content* of T. The content is $n \times d$ if T has exactly d many 1s, d many 2s, …, d many ns. The set of boxes of the Young diagram of λ is denoted by $\mathrm{boxes}(\lambda)$. The boxes that have the same number are said to form a *block*.

Let $m = n+1$. Fix a tableau T of shape λ with content $n \times d$ and fix a tensor $p = \sum_{i=1}^r \ell_{i,1} \otimes \cdots \otimes \ell_{i,d} \in \otimes^d \mathbb{C}^m$. A placement

$$\vartheta : \mathrm{boxes}(\lambda) \to [r] \times [d]$$

is called *proper* if the first coordinate of ϑ is constant in each block and the second coordinate of ϑ in each block is a permutation. We define the determinant of a matrix that has more rows than columns as the determinant of its largest top square submatrix.

For a tableau T with content $\Delta \times d$ we define the polynomial f_T via its evaluation on p:

$$f_T(p) := \sum_{\text{proper } \vartheta} \prod_{c=1}^{\lambda_1} \det{}_{\vartheta,c} \;\text{ with }\; \det{}_{\vartheta,c} := \det\left(\ell_{\vartheta(1,c)} \ldots \ell_{\vartheta(\mu_c,c)}\right) \tag{6.7}$$

The degree of f_T is Δ. The polynomial f_T is SL_m-invariant if and only if the shape of T is rectangular with exactly m many rows. It is easy to see that $f_T = 0$ if T has any column in which a number appears more than once. Moreover, it is easy to see that f_T is fixed (up to sign) when two entries in T are exchanged within a column. So, up to sign, there is only one T that could give an SL_m-invariant of degree $d+1$: It is the tableau with $m = d+1$ many rows and d columns that has only entries i in row i. For $n = 4$ it looks as follows.

$$T = \begin{array}{|c|c|c|c|}\hline 1 & 1 & 1 & 1 \\ \hline 2 & 2 & 2 & 2 \\ \hline 3 & 3 & 3 & 3 \\ \hline 4 & 4 & 4 & 4 \\ \hline 5 & 5 & 5 & 5 \\ \hline \end{array}$$

For this T it remains to verify that f_T does not vanish identically on to orbit closure $\overline{\mathrm{GL}_{d+1}(x_1 \cdots x_d + x_{d+1}^d)}$. Since f_T is SL_{d+1}-invariant, this is equivalent to f_T not vanishing at the point $x_1 \cdots x_d + x_{d+1}^d$. So we now evaluate $f_T(x_1 \cdots x_d + x_{d+1}^d)$. The nonzero summands in Equation (6.7) must place $(d+1, *)$ into one of the blocks. We can partition the summands according to the row in which $(d+1,*)$ is placed. Since the number of columns is even, each part of the partition contributes the same number to the overall sum. That number is the column sign of the unique Latin square that is obtained when removing the row in which $(d+1,*)$ is placed. Hence the whole sum if $d+1$ times the difference of the column-even and column-odd Latin squares, so its nonvanishing is equivalent to the Alon-Tarsi conjecture for d. $\square$

6.e De-bordering: Characterizing special binomials

In this section, we de-border two special binomials, namely product-plus-power and product-plus-two-powers, as defined in Section 2.b. Before proving them, we state and prove some basic lemmas which will be crucially used in the proof.

6.8 Proposition ([CCG12]). *For $a \le b \le c$, $\mathsf{WR}(x^a y^b) = b+1$, and $\mathsf{WR}(x^a y^b z^c) = (b+1)(c+1)$.*

6.9 Lemma (Interpolation). *Let $f(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]$ of degree d such that $f(\beta_i, x_2, \ldots, x_n) \in \overline{\Sigma^{[s]} \wedge \Sigma}$, for distinct $\beta_i \in \mathbb{C}$, $i \in [d+1]$, and let the maximum exponent in all the cases be $e (\ge d)$. Then, $f(\boldsymbol{x}) \in \overline{\Sigma^{[s(d+1)^2(e+1)]} \wedge \Sigma}$, with the maximum exponent being at most $e+d$.*

Proof. Let $f(\boldsymbol{x}) = \sum_{j=0}^{d} f_j x_1^j$. Trivially, there exists $\alpha_{ij} \in \mathbb{C}$, such that $f_j = \sum_{i \in [d+1]} \alpha_{ij} f(\beta_i, x_2, \ldots, x_n)$. By assumption, $f(\beta_i, x_2, \ldots, x_n) + \epsilon \cdot S_i(\epsilon, x_2, \ldots, x_n) = \sum_{j \in [s]} \ell_{ij}^{e_j}$, where ℓ_{ij} are linear polynomials over $\mathbb{C}(\epsilon)$, and $e_j \le e$. Hence, trivially,

$$f_j(\boldsymbol{x}) + \epsilon \cdot \Big(\sum_{i \in [d+1]} \alpha_{ij} S_i\Big) = \sum_{i \in [d+1]} \sum_{j \in [s]} \alpha_{ij} \ell_{ij}^{e_j} \implies f_j(\boldsymbol{x}) \in \overline{\Sigma^{[s(d+1)]} \wedge \Sigma}\,.$$

Note that, for any affine linear polynomial ℓ, the affine Waring rank of $(\ell^e \cdot x_1^j)$ is $e+1$ (since $j \le d \le e$), with the maximum exponent being $e+j$; this follows from Proposition 6.8. Therefore, $f_j x_1^j \in \overline{\Sigma^{[s(d+1)(e+1)]} \wedge \Sigma}$, with the maximum exponent being at most $e+j$. Finally, adding up, we get $f(\boldsymbol{x}) \in \overline{\Sigma^{[s(d+1)^2(e+1)]} \wedge \Sigma}$, with the maximum exponent being at most $e+d$. $\square$

By applying the lemma twice, we get the following.

6.10 Corollary. *Let $f(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]$ of degree d such that $f(\beta_i, \gamma_j, x_3, \ldots, x_n) \in \overline{\Sigma^{[s]} \wedge \Sigma}$, for distinct $\beta_i \in \mathbb{C}$, (and similarly γ_j), $i, j \in [d+1]$, and let the maximum exponent in all the cases be $e (\geq d)$. Then, $f(\boldsymbol{x}) \in \overline{\Sigma^{[O(sed^4]}} \wedge \Sigma$, with the maximum exponent being at most $e + 2d$.*

Next lemma is about going from $\overline{\Sigma \wedge \Sigma}$ to border waring rank in the homogeneous setting.

6.11 Lemma. *Let $f(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]_d$, such that $f(\boldsymbol{x}) \in \overline{\Sigma^{[s]} \wedge \Sigma}$. Then, $\underline{\mathrm{WR}}(f) \leq s$.*

Proof. By assumption, $f(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon) = \sum_{i \in [s]} (\alpha_i + \ell_i)^{e_i}$, where $\alpha_i \in \mathbb{C}(\epsilon)$, and $\ell_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]_1$. Since, degree of f is d, taking the degree d part of RHS gives $\sum_{i \in [s] | e_i \geq d} \binom{e_i}{d} \ell_i^d \alpha_i^{e_i - d}$. Trivially, this implies that $\underline{\mathrm{WR}}(f) \leq s$. $\square$

6.12 Lemma (Folklore)**.** *Let $f(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]$, and $A \in \mathrm{GL}_n(\mathbb{C})$. If $f(A\boldsymbol{x}) \in \overline{\Sigma^{[s]} \wedge \Sigma} \iff f(\boldsymbol{x}) \in \overline{\Sigma^{[s]} \wedge \Sigma}$.*

Proof. This follows trivially, since this is just the $(GL_n(\mathbb{C})$ group action. $\square$

The next lemma says how to extract ϵ-powers to separate out the ϵ-free part, from the rest.

6.13 Lemma. *Let $\mathcal{R}[\boldsymbol{x}]_1 \ni \ell := x_1 + \epsilon \cdot \check{\ell}$, for some linear form $\check{\ell}$ (not necessarily x_1-free), where $\mathcal{R} := \mathbb{C}[\epsilon]/\langle \epsilon^M \rangle$, for some $M \in \mathbb{N}$, then there exist another linear form $\tilde{\ell} \in \mathcal{R}[\boldsymbol{x}]_1$, such that $\ell = \tilde{\ell}$ over the ring $\mathcal{R}$, where $\tilde{\ell} = \gamma \cdot (x_1 + \epsilon \cdot \widehat{\ell})$, for some $\gamma \in \mathcal{R}$, with $\gamma \equiv 1 \bmod \langle \epsilon \rangle$, and $\widehat{\ell} \in \mathcal{R}[\boldsymbol{x}]_1$ which is x_1-free.*

Proof. $\ell = x_1 + \epsilon \check{\ell}$. Suppose, $\check{\ell} = c x_1 + \ell_0$, where ℓ_0 is x_1-free, and $c \in \mathcal{R}$. Rewrite ℓ as $\ell = (1 + c\epsilon) x_1 + \epsilon \cdot \ell_0$. Define, $\mathcal{R} \ni \gamma := (1 + c\epsilon) \bmod \langle \epsilon \rangle^M$. Note that

$$\frac{1}{1 + c\epsilon} \equiv \left(\sum_{i=0}^{M-1} (-c\epsilon)^i \right) \bmod \langle \epsilon \rangle^M =: \gamma' \in \mathcal{R} .$$

Hence, over $\mathcal{R}$, the following equality holds:

$$\ell = \gamma x_1 + \epsilon \cdot \ell_0 = \gamma (x_1 + \epsilon \cdot \gamma' \cdot \ell_0) =: \tilde{\ell} .$$

This finishes the proof. $\square$

Now we prove the non-homogeneous generalisation of Theorem 4.11. The proof technique is almost same as in the proof of Theorem 4.11.

6.14 Theorem (De-bordering affine $\underline{\mathsf{Kc}}^-$)**.** *For any degree d polynomial $f(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]$, not necessarily homogeneous, with $\underline{\mathsf{Kc}}^-(f) = m$, we have $f \in \overline{\Sigma^{[md+1]} \wedge \Sigma}$, with maximum exponent being d.*

Proof. Let $f \simeq f_\epsilon := \gamma \epsilon^{-M} \big(\prod_{i=1}^m (1 + \ell_i') - 1 \big)$ with $M \geq 1$. From the convergence of f_ϵ we deduce that for each i we have $\ell_i' = \epsilon \ell_i$ with $\ell_i \in \mathbb{C}[\epsilon][\boldsymbol{x}]_1$, because otherwise the homogeneous degree j part diverges, where j is the number of ℓ_i' that do not satisfy this property.

Let $f_{\epsilon,j}$ denote the homogeneous degree j part of f_ϵ and analogously we use f_j to denote the homogeneous degree j part of f. First we note that $f_{\epsilon,j} = \gamma \epsilon^{-M} e_j(\epsilon \ell_1, \ldots, \epsilon \ell_m) = \gamma \epsilon^{-M+j} e_j(\boldsymbol{\ell})$. We know that $f_j \simeq f_{\epsilon,j} = \gamma \epsilon^{-M+j} e_j(\boldsymbol{\ell})$. We now show by induction that for all $1 \leq j \leq d$ we have

$$\begin{aligned} p_j(\boldsymbol{\ell}) &\equiv 0 \bmod \langle \epsilon^{M-j} \rangle \\ e_j(\boldsymbol{\ell}) &\equiv 0 \bmod \langle \epsilon^{M-j} \rangle \\ p_j(\boldsymbol{\ell}) &\equiv (-1)^{j-1} j \left(e_j(\boldsymbol{\ell}) \right) \bmod \langle \epsilon^{M-j+1} \rangle \end{aligned} \tag{6.15}$$

The base case $j = 1$ is trivially true because $p_1 = e_1$. By using Equation (6.15) for $1 \le i \le j$, we know that there exist $P_i, E_i \in \mathbb{C}[\epsilon][\boldsymbol{x}]$ such that

$$\begin{aligned} p_i(\boldsymbol{\ell}) =& \epsilon^{M-i} P_i \\ e_{j+1-i}(\boldsymbol{\ell}) =& \epsilon^{M-(j+1)+i} E_i \end{aligned}$$

This implies that $e_{j+1-i}(\boldsymbol{\ell}) \cdot p_i(\boldsymbol{\ell}) = \epsilon^{M-i} P_i \epsilon^{M-(j+1)+i} E_i = \epsilon^{2M-(j+1)} P_i E_i$. Therefore, we have that $e_{j+1-i}(\boldsymbol{\ell}) \cdot p_i(\boldsymbol{\ell}) \equiv 0 \bmod \langle \epsilon^{M-j} \rangle$ as $M \ge 1$. Now for the induction step from j to $j+1$ we use Newton's identities:

$$p_{j+1}(\boldsymbol{\ell}) = (-1)^j (j+1) e_{j+1}(\boldsymbol{\ell}) + \sum_{i=1}^{j} (-1)^{j+i} e_{j+1-i}(\boldsymbol{\ell}) \cdot p_i(\boldsymbol{\ell})$$

The claim $e_{j+1-i}(\boldsymbol{\ell}) \cdot p_i(\boldsymbol{\ell}) \equiv 0 \bmod \langle \epsilon^{M-j} \rangle$ implies that $p_{j+1}(\boldsymbol{\ell}) \equiv (-1)^j (j+1) \left(e_{j+1}(\boldsymbol{\ell})\right) \bmod \langle \epsilon^{M-j} \rangle$. We also know that $f_{j+1} \simeq f_{\epsilon, j+1} = \gamma \epsilon^{-M+j+1} \left(e_j(\boldsymbol{\ell})\right)$. Since $f_{\epsilon,j+1}$ converges, $\epsilon^{M-(j+1)}$ has to divide $e_{j+1}(\boldsymbol{\ell})$. Therefore $e_{j+1}(\boldsymbol{\ell}) \equiv 0 \bmod \langle \epsilon^{M-(j+1)} \rangle$, which readily implies $p_{j+1}(\boldsymbol{\ell}) \equiv 0 \bmod \langle \epsilon^{M-(j+1)} \rangle$. This proves the induction hypothesis. Hence

$$f_j \simeq \gamma \epsilon^{-M+j} \left(e_j(\boldsymbol{\ell})\right) \simeq \gamma \epsilon^{-M+j} \cdot \frac{1}{j} \cdot (-1)^{j-1} \left(p_j(\boldsymbol{\ell})\right).$$

Which gives us that that $\underline{\mathrm{WR}}(f_j) \le m$. By the sub-additive property of the border Waring rank, we obtain that $f \in \overline{\Sigma^{[md+1]} \wedge \Sigma}$. The exponents come from the exponents from $p_j(\boldsymbol{\ell})$, which can be at most d. ☐

6.16 Corollary. *In the above model, instead of -1, if it was any arbitrary $\alpha \in \mathbb{C}[\epsilon]$, such that $\alpha \equiv 1 \bmod \langle \epsilon \rangle$ i.e., $f \simeq f_\epsilon := \gamma \epsilon^{-M} \left(\prod_{i=1}^m (1+\ell_i') - \alpha\right)$ the same proof goes through, this is essentially because α does not contribute to the higher degree terms. Note that, $\alpha \equiv 1 \bmod \langle \epsilon \rangle$, because otherwise the limit does not exist.*

6.17 Theorem. *For any degree d polynomial $f(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]$, not necessarily homogeneous, suppose we have $f \simeq \gamma \epsilon^{-M} \big(\prod_{i=1}^m (1+\epsilon a_i) - \prod_{i=1}^m (1+\epsilon b_i) \big)$ for some linear forms $a_i, b_i \in \mathbb{C}[\epsilon][\boldsymbol{x}]_1$ and $\gamma \in \mathbb{C}$ with $M \ge 1$. Then we have $f \in \overline{\Sigma^{[2md+1]} \wedge \Sigma}$, with the maximum exponent being d.*

Proof. We have $f \simeq f_\epsilon := \gamma \epsilon^{-M} \big(\prod_{i=1}^m (1+\epsilon a_i) - \prod_{i=1}^m (1+\epsilon b_i) \big)$ with $M \ge 1$. Let $f_{\epsilon,j}$ denote the homogeneous degree j part of f_ϵ and analogously we use f_j to denote the homogeneous degree j part of f. Now, note that $f_j \simeq f_{\epsilon,j} = \gamma \epsilon^{-M} \left(e_j(\epsilon a_1, \ldots, \epsilon a_m) - e_j(\epsilon b_1, \ldots, \epsilon b_m)\right)$. Therefore $f_j \simeq f_{\epsilon,j} = \gamma \epsilon^{-M+j} \left(e_j(\boldsymbol{a}) - e_j(\boldsymbol{b})\right)$, where $\boldsymbol{a} = (a_1, \ldots, a_m)$ and similarly $\boldsymbol{b} = (b_1, \ldots, b_m)$. We now show by induction that for all $1 \le j \le d$ we have

$$\begin{aligned} p_j(\boldsymbol{a}) \equiv& p_j(\boldsymbol{b}) \bmod \langle \epsilon^{M-j} \rangle \\ e_j(\boldsymbol{a}) \equiv& e_j(\boldsymbol{b}) \bmod \langle \epsilon^{M-j} \rangle \\ p_j(\boldsymbol{a}) - p_j(\boldsymbol{b}) \equiv& (-1)^{j-1} j \left(e_j(\boldsymbol{a}) - e_j(\boldsymbol{a})\right) \bmod \langle \epsilon^{M-j+1} \rangle \end{aligned} \tag{6.18}$$

The base case $j = 1$ is trivially true because $p_1 = e_1$ and $f_1 \simeq f_{1,1} = \gamma \epsilon^{-M+1} \left(e_1(\boldsymbol{a}) - e_1(\boldsymbol{b})\right)$ converges, therefore ϵ^{M-1} has to divide $e_1(\boldsymbol{a}) - e_1(\boldsymbol{b}) = p_1(\boldsymbol{a}) - p_1(\boldsymbol{b})$. For the induction step from j to $j+1$, we use Newton's identities:

$$p_{j+1} = (-1)^j (j+1) e_{j+1} + \textstyle\sum_{i=1}^{j} (-1)^{j+i} e_{j+1-i} \cdot p_i.$$

Hence

$$p_{j+1}(\boldsymbol{a}) - p_{j+1}(\boldsymbol{b}) = (-1)^j (j+1) \left(e_{j+1}(\boldsymbol{a}) - e_{j+1}(\boldsymbol{a})\right) + \sum_{i=1}^{j} (-1)^{j+i} \left(e_{j+1-i}(\boldsymbol{a}) \cdot p_i(\boldsymbol{a}) - e_{j+1-i}(\boldsymbol{b}) \cdot p_i(\boldsymbol{b})\right).$$

By using Equation (6.18) for $1 \le i \le j$, we know that

$$p_i(\boldsymbol{a}) \equiv p_i(\boldsymbol{b}) \bmod \langle \epsilon^{M-i} \rangle$$
$$e_{j+1-i}(\boldsymbol{a}) \equiv e_{j+1-i}(\boldsymbol{b}) \bmod \langle \epsilon^{M-(j+1)+i} \rangle.$$

This means that there exist $P_i, E_i \in \mathbb{C}[\epsilon][\boldsymbol{x}]$ such that

$$p_i(\boldsymbol{a}) = p_i(\boldsymbol{b}) + \epsilon^{M-i} P_i$$
$$e_{j+1-i}(\boldsymbol{a}) = e_{j+1-i}(\boldsymbol{b}) + \epsilon^{M-(j+1)+i} E_i$$

Since, $1 \le i \le j$, this implies tha

$$\begin{aligned} e_{j+1-i}(\boldsymbol{a}) \cdot p_i(\boldsymbol{a}) &= (e_{j+1-i}(\boldsymbol{b}) + \epsilon^{M-(j+1)+i} E_i)(p_i(\boldsymbol{b}) + \epsilon^{M-i} P_i) \\ &= e_{j+1-i}(\boldsymbol{b}) \cdot p_i(\boldsymbol{b}) + \epsilon^{M-j} Q_i \end{aligned}$$

for some $Q_i \in \mathbb{C}[\epsilon][\boldsymbol{x}]$. Therefore

$$p_{j+1}(\boldsymbol{a}) - p_{j+1}(\boldsymbol{b}) \equiv (-1)^j (j+1) \left(e_{j+1}(\boldsymbol{a}) - e_{j+1}(\boldsymbol{a})\right) \bmod \langle \epsilon^{M-j} \rangle$$

We also know that $f_{j+1} \simeq f_{\epsilon,j+1} = \gamma \epsilon^{-M+j+1} \left(e_j(\boldsymbol{a}) - e_j(\boldsymbol{b})\right)$. Since $f_{\epsilon,j+1}$ converges, $\epsilon^{M-(j+1)}$ has to divide $e_{j+1}(\boldsymbol{a}) - e_{j+1}(\boldsymbol{b})$. Therefore $e_{j+1}(\boldsymbol{a}) \equiv e_{j+1}(\boldsymbol{b}) \bmod \langle \epsilon^{M-(j+1)} \rangle$, which readily implies $p_{j+1}(\boldsymbol{a}) \equiv p_{j+1}(\boldsymbol{b}) \bmod \langle \epsilon^{M-(j+1)} \rangle$. This proves the induction hypothesis. Hence

$$f_j \simeq \gamma \epsilon^{-M+j} \left(e_j(\boldsymbol{a}) - e_j(\boldsymbol{b})\right) \simeq \gamma \epsilon^{-M+j} \cdot \frac{1}{j} \cdot (-1)^{j-1} \left(p_j(\boldsymbol{a}) - p_j(\boldsymbol{b})\right).$$

This easily implies that, $\underline{\mathsf{WR}}(f_j) \le 2m$. By the sub-additive property of the border Waring rank, we obtain that $f \in \overline{\Sigma^{[2md+1]} \wedge \Sigma}$. □

6.19 Corollary. *For any degree d polynomial $f(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]$, not necessarily homogeneous, suppose we have $f \simeq \gamma \epsilon^{-M} \left(\prod_{i=1}^m (1 + \epsilon a_i) - \alpha \cdot \prod_{i=1}^m (1 + \epsilon b_i) - \beta\right)$ for some linear forms $a_i, b_i \in \mathbb{C}[\epsilon][\boldsymbol{x}]$, $\alpha, \beta \in \mathbb{C}[\epsilon]$, such that $\alpha \not\equiv 0 \bmod \langle \epsilon \rangle$, and $\gamma \in \mathbb{C}$ with $M \ge 1$. Then we have $f \in \overline{\Sigma^{[2md+1]} \wedge \Sigma}$, with maximum exponent being d.*

Proof sketch. This essentially follows from the same proof as Theorem 6.17. Note that, for $j \ge 1$, β does not contribute anything, and $e_j(\boldsymbol{a}) \equiv \alpha e_j(\boldsymbol{b}) \bmod \langle \epsilon^{M-j} \rangle$; the same holds for p_j. Therefore, by induction we can show that

$$f_j \simeq \gamma \epsilon^{-M+j} \left(e_j(\boldsymbol{a}) - \alpha e_j(\boldsymbol{b})\right) \simeq \gamma \epsilon^{-M+j} \cdot \frac{1}{j} \cdot (-1)^{j-1} \left(p_j(\boldsymbol{a}) - \alpha p_j(\boldsymbol{b})\right).$$

This easily implies that, $\underline{\mathsf{WR}}(f_j) \le 2m$. By the sub-additive property of the border Waring rank, we obtain that $f \in \overline{\Sigma^{[2md+1]} \wedge \Sigma}$. □

6.20 Theorem (De-bordering product-plus-power)**.** *Let $f \in S^d V$, such that $f \trianglelefteq P^{[d]}$,then either $f \leq P^{[d]}$, or, $\underline{\mathsf{WR}}(f) \leq (d+1)^3(d^2+1)$.*

Proof. Since, $f \trianglelefteq \prod_{i\in[d]} x_i + x_0^d$, by definition, there are $\ell_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]_1$, such that $f = \lim_{\epsilon\to 0}\left(\prod_{i\in[d]} \ell_i - \ell_0^d\right)$; here, we change '$+$' to '$-$', because over $\mathbb{C}$, it is okay to do so, by an appropriate rescaling. By multiplying large powers of ϵ powers on both sides (and renaming it), one can say that there exists $\ell_i \in \mathbb{C}[\epsilon][\boldsymbol{x}]_1, S \in \mathbb{C}[\epsilon,\epsilon^{-1}][\boldsymbol{x}]_d$, and $M \in \mathbb{Z}_{\geq 0}$, such that

$$\ell_1 \cdots \ell_d - \ell_0^d = \epsilon^M \cdot f + \epsilon^{M+1} \cdot S(\boldsymbol{x},\epsilon)\ .$$

If $M = 0$, then individually the limit must exist which implies that f must be of the form $\prod \widehat{\ell}_i - \widehat{\ell}_0^d$, for some linear forms $\widehat{\ell}_i$.

Hence, without loss of generality, assume that $M \geq 1$. Let $\ell_i|_{\epsilon=0} = \ell_{i0}$. We can assume that ℓ_{i0} are nonzero for each $i \in [d]$; otherwise, say wlog $\ell_{10} = 0 \implies \ell_1 \equiv 0 \bmod \langle\epsilon\rangle \implies \prod_{i\in[d]} \ell_i \equiv 0 \bmod \langle\epsilon\rangle$. Since, $M \geq 1$, this must imply that $\ell_0^d \equiv 0 \bmod \langle\epsilon\rangle \implies \ell_0 \equiv 0 \bmod \langle\epsilon\rangle$, and hence we can further reduce M.

Since, $M \geq 1$, by comparing the coefficient of ϵ^0, we must have $\prod_{i\in[d]} \ell_{i0} = \ell_{00}^d$. This means ℓ_{i0} are multiple of each other. So, wlog $\ell_{i0} = x_1$, otherwise we can apply an invertible transformation ($\boldsymbol{x} \mapsto E\boldsymbol{x}$, for $E \in \mathrm{GL}_n(\mathbb{C})$) which sends $\ell_{00} \mapsto x_1$ and $x_i \mapsto x_i$, for $i \in [2,n]$. Here, we assumed that ℓ_{00} has x_1-variable, otherwise, we will work with a variable that is in ℓ_{00}. Also note that the desired forms in the theorem are preserved under the transformation.

So after the transformation, one can replace ℓ_i by $\tilde{\ell}_i$, using Lemma 6.13 such that $\ell = \tilde{\ell}_i$, over $\mathcal{R}$, where $\mathcal{R} := \mathbb{C}[\epsilon]/\langle\epsilon^{M+1}\rangle$, where $\tilde{\ell}_i := c_i(x_1 + \epsilon \cdot \widehat{\ell}_i)$, such that $\widehat{\ell}_i$ are x_1-free, and $c_i \equiv 1 \bmod \langle\epsilon\rangle$. Note that, by working over $\mathcal{R}$,we make sure that RHS stays unaffected $\bmod \langle\epsilon^{M+1}\rangle$, which implies the coefficient of ϵ^M, in the RHS, remains $f(E\boldsymbol{x})$.

Let us define $\alpha := \prod_{i\in[d]} c_i$. Therefore, we will get the following form:

$$\alpha \cdot \prod_{i\in[d]} (x_1 + \epsilon \cdot \widehat{\ell}_i) \ - \ c_0 \cdot (x_1 + \epsilon \cdot \widehat{\ell}_0)^d \ = \ \epsilon^M \cdot f(E\boldsymbol{x}) + \epsilon^{M+1} \cdot \widehat{S}\ .$$

Here, both α and c_0 are elements in $\mathcal{R}$, such that they are $\equiv 1 \bmod \langle\epsilon\rangle$. One can also divide both side by α. Note that, $\gamma := (c_0/\alpha \bmod \langle\epsilon\rangle^{M+1}) \in \mathcal{R}$, and further $f(E\boldsymbol{x})/\alpha \equiv f(E\boldsymbol{x}) \bmod \langle\epsilon\rangle$. This implies, even after dividing both side by α, the coefficient of ϵ^M remains unchanged. Moreover, we get the following form:

$$\prod_{i\in[d]} (x_1 + \epsilon \cdot \widehat{\ell}_i) \ - \ \gamma(x_1 + \epsilon \cdot \widehat{\ell}_0)^d \ = \ \epsilon^M \cdot f(E\boldsymbol{x}) + \epsilon^{M+1} \cdot \tilde{S}\ . \tag{6.21}$$

Let $S^d V \ni g(\boldsymbol{x}) := f(E\boldsymbol{x})$. Note that it suffices to prove that each $g(j, x_2, \cdots, x_n) \in \overline{\Sigma^{[d^2+1]} \wedge \Sigma}$, for each $j \in [1, d+1]$, with maximum exponent d, because of the following reasons:

1. If $g(j, x_2, \cdots, x_n) \in \overline{\Sigma^{[d^2+1]} \wedge \Sigma}$, for $j = 1, \ldots, d+1$, then by interpolation, we would get that $g(\boldsymbol{x}) \in \overline{\Sigma^{[(d+1)^3(d^2+1)]} \wedge \Sigma}$, see Lemma 6.9.

2. If g is homogeneous such that $g(\boldsymbol{x}) \in \overline{\Sigma^{[(d+1)^3(d^2+1)]} \wedge \Sigma}$ then $\underline{\mathsf{WR}}(g) \leq (d+1)^3(d^2+1)$ by Lemma 6.11.

3. If $\underline{\mathsf{WR}}(g) \leq (d+1)^3(d^2+1)$, then by Lemma 6.12, $\underline{\mathsf{WR}}(f) \leq (d+1)^3(d^2+1)$.

Our proof is basically the same for all j, so we will wlog work with $j = 1$, and show that $g(1, x_2, \ldots, x_n) \in \overline{\Sigma^{[d^2+1]} \wedge \Sigma}$.

6.22 Claim. $g(1, x_2, \ldots, x_n) \in \overline{\Sigma^{[d^2+1]} \wedge \Sigma}$, with the maximum exponent being d.

Proof. Substitute $x_1 = 1 - \epsilon \cdot \widehat{\ell}_0$, in Equation (6.21). LHS becomes $\prod_{i \in [d]}(1 + \epsilon \cdot L_i) - \gamma$, for some linear forms L_i in $\mathbb{C}[\epsilon, \boldsymbol{x}]$, which are x_1-free. Also note that,

$$g(1 - \epsilon \cdot \widehat{\ell}_0, x_2, \ldots, x_n) \;=\; g(1, x_2, \ldots, x_n) \,+\, \epsilon \cdot \widehat{g}(\boldsymbol{x}, \epsilon)\;,$$

for some $\widehat{g} \in \mathbb{C}[\epsilon][\boldsymbol{x}]$. Therefore, Equation (6.21) becomes

$$\prod_{i \in [d]}(1 + \epsilon \cdot L_i) \,-\, \gamma \;=\; \epsilon^M \cdot g(1, x_2, \ldots, x_n) \,+\, \epsilon^{M+1} \cdot \check{S}\;. \tag{6.23}$$

By Theorem 6.14 and its corollary (Corollary 6.16), it follows that $g(1, x_2, \ldots, x_n) \in \overline{\Sigma^{[d^2+1]} \wedge \Sigma}$, as desired. $\square$

Since, each $g(j, x_2, \ldots, x_n) \in \overline{\Sigma^{[d^2+1]} \wedge \Sigma}$, of fanin $(d^2 + 1)$ by Claim 6.22, we get that $\underline{\mathsf{WR}}(f) \le (d+1)^3(d^2+1)$, as we wanted. $\square$

6.24 Remark. One can actually simply substitute $x_1 = j$, in Equation (6.21), to get a form as of in Theorem 6.17. However, in that case, there is a 2 multiplicative factor; in particular, by doing that we would get that $g(j, x_2, \ldots, x_n) \in \overline{\Sigma^{[2d^2+1]} \wedge \Sigma}$, and hence $\underline{\mathsf{WR}}(f) \le (d+1)^3(2d^2+1)$. For this *slightly* better bound, we analyzes the proof as above.

6.25 Theorem (De-bordering product-plus-two-powers)**.** *Let $f \in S^d V$, such that $f \trianglelefteq P_2^{[d]}$, then, either*

1. $f \le P_2^{[d]}$, *or,*
2. $f \le \prod_{i \in [d]} y_i + y_0^{d-1} \cdot y_{d+1}$, *or,*
3. $\underline{\mathsf{WR}}(f) \le O(d^7)$.

6.26 Remark. The first two cases, can be compiled as $f = g + h$, where $g \le \prod_{i \in [d]} y_i$, and $\underline{\mathsf{WR}}(h) \le 2$. Note that, $\underline{\mathsf{WR}}(h) \le 2$, implies that $\mathsf{WR}(h) \le 2$, or $h \le y_0^{d-1} y_{d+1}$, form Theorem B.1. We will in fact be using this form (of $f = g + h$) in the proof.

Proof. Since, $f \trianglelefteq P_2^{[d]}$, by definition, there are linear forms $L_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]_1$, and $H \in \mathbb{C}[\epsilon][\boldsymbol{x}]$, such that

$$\prod_{i \in [d]} L_i \,+\, L_0^d \,+\, L_{d+1}^d \;=\; f \,+\, \epsilon \cdot H(\boldsymbol{x}, \epsilon)\;.$$

B multiplying large powers of ϵ both the sides, we can make sure that we are only working over $\mathbb{C}[\epsilon]$. In particular, we will have the following:

$$\epsilon^{a_1} \cdot \underbrace{\ell_1 \cdots \ell_d}_{T_1} + \epsilon^{a_2} \cdot \underbrace{\ell_0^d}_{T_2} + \epsilon^{a_3} \cdot \underbrace{\ell_{d+1}^d}_{T_3} \;=\; \epsilon^M \cdot f \,+\, \epsilon^{M+1} \cdot S(\boldsymbol{x}, \epsilon)\;.$$

In the above, $a_i, M \in \mathbb{N}_{\ge 0}$, and $\ell_i \in \mathbb{C}[\epsilon][\boldsymbol{x}]_1$, such that $\ell_i|_{\epsilon=0} = \ell_{i,0}$, is a *nonzero* linear form in $\mathbb{C}[\boldsymbol{x}]$. Further, wlog, one can assume the following (otherwise, the limit diverges, or easy de-borderings):

1. $\min a_i = 0$, otherwise, we can divide by $\epsilon^{M-\min a_i}$, both side. Note that, $M \geq \min a_i$, because otherwise, the limit f must be 0.

2. We can further assume that $a_i < M$, for all i. Otherwise, say $S := \{i \in [3] \mid a_i = M\} \subseteq [3]$, and $S' := [3] - S$. Trivially,

$$f \;=\; \sum_{i \in S} T_i|_{\epsilon=0} \;+\; \lim_{\epsilon \to 0} 1/\epsilon^M \cdot \left(\sum_{i \in S'} \epsilon^{a_i} T_i \right) \;.$$

   There are further subcases to consider.

   (a) If $S = [3]$ (and hence, S' is empty), then $f = g + h$, where g is a product of linear forms and $\mathsf{WR}(h) \leq 2$.

   (b) If $S = \{1\}$, then $f = g + h$, where g is a product of linear forms, and $\underline{\mathsf{WR}}(h) \leq 2$.

   (c) $S = \{2\}$ or $S = \{3\}$, then $f_1 := \lim_{\epsilon \to 0} 1/\epsilon^M \cdot (\sum_{i \in S'} \epsilon^{a_i} T_i)$ can be found by Theorem 6.20: f_1 is either of the form $g + h_1$, where g is product of linear forms and h_1 is a power of linear form, or $\underline{\mathsf{WR}}(f_1) \leq O(d^5)$. In this case, since $T_i|_{\epsilon=0}$, for $i \in S$, is a power of a linear form, combining them we get that either –

      i. If $f_1 = g + h_1$, then $f = g + h$, where $h := h_1 + T_i(\epsilon = 0)$, and trivially, $\mathsf{WR}(h) \leq 2$.

      ii. If $\underline{\mathsf{WR}}(f_1) \leq O(d^5)$, then again trivially, $f = f_1 + T_i|_{\epsilon=0} \implies \underline{\mathsf{WR}}(f) \leq O(d^5)$.

   (d) If S has 2 elements, then S' has exactly 1 element and thus, $\lim_{\epsilon \to 0} 1/\epsilon^M \cdot (\sum_{i \in S'} \epsilon^{a_i} T_i)$ is trivial (either a product or a power). Again, in this case, $f = g + h$, where g is a product of linear forms and $\mathsf{WR}(h) \leq 2$.

3. $\min a_i$ *does not* come from an unique element, i.e., there must exist at least two indices i and j such that $a_i = a_j = \min_{k \in [3]} a_k$. If not (i.e., if the minimum is unique), then $M = \min a_i$, and hence $f = T_i|_{\epsilon=0}$, for some i, which is either a product of linear forms, or a power of a linear form (which is trivially subsumed in the $g + h$ form).

Therefore, the remaining cases are as follows:

Case I: $a_1 = a_2 = a_3 = 0$. We can write this case as:

$$\prod_{i \in [d]} \ell_i + \ell_0^d + \ell_{d+1}^d \;=\; \epsilon^M \cdot f + \epsilon^{M+1} \cdot S(\boldsymbol{x}, \epsilon) \;.$$

Case II: $a_1 = a_2 = 0$ and $a_3 = a > 0$. We can write this case as:

$$\prod_{i \in [d]} \ell_i + \ell_0^d + \epsilon^a \cdot \ell_{d+1}^d \;=\; \epsilon^M \cdot f + \epsilon^{M+1} \cdot S(\boldsymbol{x}, \epsilon) \;.$$

Case III: $a_2 = a_3 = 0$ and $a_1 = a > 0$. We write the case as:

$$\epsilon^a \cdot \prod_{i \in [d]} \ell_i + \ell_0^d + \ell_{d+1}^d \;=\; \epsilon^M \cdot f + \epsilon^{M+1} \cdot S(\boldsymbol{x}, \epsilon) \;.$$

In the above, we have not considered the case $a_1 = a_3 = 0, a_2 > 0$, because this is symmetric as Case II. Since the minimum cannot be unique (point 3), many cases are discarded, and we have covered all the cases. Therefore, it suffices to deborder these 3 cases. We show that in all these cases $\underline{\mathrm{WR}}(f) \le O(d^7)$. All the cases, although, have a similar proof-flow, have their own subtelties.

To begin with, here is a crucial claim.

6.27 Claim. In all the 3 above cases, $\mathrm{rank}(\ell_{0,0}, \cdots, \ell_{d+1,0}) \le 2$.

Proof. Just for sanity, we will go case by case.

Case I: There are two subcases –

(i) If $\mathrm{rank}(\ell_{0,0}, \ell_{d+1,0}) = 2$, then since $\prod_{i\in[d]} \ell_{i,0} = -\prod_{j=1}^{d}(\ell_{0,0} - \zeta^j \ell_{d+1,0})$, where ζ is the $2d$-th root of unity, using the unique factorization and reducibility of linear forms, we get that $\ell_{i,0} \in \langle \ell_{0,0}, \ell_{d+1,0}\rangle$, for each $i \in [d]$, and hence, $\mathrm{rank}(\ell_{0,0}, \cdots, \ell_{d+1,0}) = 2$.

(ii) If $\mathrm{rank}(\ell_{0,0}, \ell_{d+1,0}) = 1$, then it is obvious to see that each ℓ_{i0} is also a constant multiple of $\ell_{d+1,0}$, implying $\mathrm{rank}(\ell_{0,0}, \cdots, \ell_{d+1,0}) = 1$.

Case II: Since, $M' > a > 0$, it must happen that $\prod_{i\in[d]} \ell_{i,0} = \ell_{0,0}^d$, which implies that $\mathrm{rank}(\ell_{0,0}, \cdots, \ell_{d,0}) = 1$. Therefore, $\mathrm{rank}(\ell_{0,0}, \cdots, \ell_{d+1,0}) \le 2$.

Case III: In this case, we have $\ell_0^d + \ell_{d+1}^d \equiv 0 \bmod \langle \epsilon^a \rangle$. Since, $a \ge 1$, we invoke Theorem B.1, to deduce that the coefficient of ϵ^a in $\ell_0^d + \ell_{d+1}^d$ must be of the form $\ell_{0,0}^{d-1} \cdot \ell'$, for some linear form ℓ', with trivially $\mathrm{rank}(\ell_{0,0}, \ell_{d+1,0}) = 1$; this basically follows from the proof of Theorem B.1.

Comparing the ϵ^a-th coefficient both side, we get that $\prod_{i\in[d]} \ell_{i,0} = -\ell_{0,0}^{d-1} \cdot \ell'$, it follows that each $\ell_{i,0}$ (except one) is a constant multiple of $\ell_{0,0}$, and the remaining one is constant multiple of ℓ'. This readily implies that $\mathrm{rank}(\ell_{0,0}, \cdots, \ell_{d+1,0}) \le 2$. This finishes the claim. □

Now we return to the proof of de-bordering. We divide all the three cases into two bigger cases of (a) & (b) [notation: By Case I(a), we will mean Case (a) along with the rank constraints, as per below]:

Case I-II(a): $\mathrm{rank}(\ell_{0,0}, \ell_{d+1,0}) = 2$. [We *do not consider* Case III(a) because in Case III, rank=2 is not possible from the proof of Claim 6.27.]

Case I-III(b): $\mathrm{rank}(\ell_{0,0}, \ell_{d+1,0}) = 1$.

Case I(a): Apply an invertible linear transformation ($\boldsymbol{x} \mapsto R\boldsymbol{x}$, for $R \in \mathrm{GL}_n(\mathbb{C})$), which sends $\ell_{0,0} \mapsto x_1$, $\ell_{d+1,0} \mapsto x_2$, and $x_i \mapsto x_i$, for $i \in [3, n]$. Here, we assumed that $\ell_{0,0}$ has x_1-variable, and $\ell_{d+1,0}$ has x_2, otherwise, we will work with the variables that are in them (and there must exist at least two, otherwise their rank is 1). Also note that the border waring rank is preserved under this transformation (Lemma 6.12).

So after the transformation and clubbing the x_1 and x_2-terms together, each $\widehat{\ell}_i := \ell_i(R\boldsymbol{x})$, for $i \in [d]$, becomes of the form $c_i x_1 + c_i' x_2 + \epsilon \cdot \tilde{\ell}_i$, where $\tilde{\ell}_i$ are x_1, x_2-free, $c_i, c_i' \in \mathbb{C}[\epsilon]$, such that both c_i or c_i' must be $\not\equiv 0 \bmod \langle \epsilon \rangle$. The reason being, $\prod_{i\in[d]} \left(\widehat{\ell}_i \bmod \langle \epsilon \rangle\right) = -(x_1^d + x_2^d) = -\prod_{i\in[d]}(x_1 + \zeta^i x_2)$, where ζ is the $2d$-th root of unity, implying that each $\widehat{\ell}_i \bmod \langle \epsilon \rangle$ should depend upon both x_1, x_2 by using the fact that polynmial ring is a UFD.

Similarly, $\ell_0(R\mathbf{x}) := \widehat{\ell}_0$ becomes $c_0x_1 + c'_0x_2 + \epsilon \cdot \tilde{\ell}_0$, where $\tilde{\ell}_0$ are x_1, x_2-free, and $c_0 \equiv 1 \bmod \langle\epsilon\rangle$, and $c'_0 \equiv 0 \bmod \langle\epsilon\rangle$. The reason of c_0 being $\equiv 1 \bmod \langle\epsilon\rangle$ is similar as before; see proof of Lemma 6.13. To see $c'_0 \equiv 0 \bmod \langle\epsilon\rangle$, note that the constant term of ℓ_0 is x_1, by assumption and hence, x_2 terms always come up with multiple in ϵ. Similarly, $\ell_{d+1}(R\mathbf{x}) := \widehat{\ell}_{d+1}$ becomes $c_{d+1}x_1 + c'_{d+1}x_2 + \epsilon \cdot \tilde{\ell}_{d+1}$, where $\tilde{\ell}_i$ are x_1, x_2-free, and $c_{d+1} \equiv 0 \bmod \langle\epsilon\rangle$, and $c'_{d+1} \equiv 1 \bmod \langle\epsilon\rangle$.

Since, $\widehat{\ell}_0^d + \widehat{\ell}_{d+1}^d$ can be rewritten as $\prod_{i\in[d]}(\widehat{\ell}_0 + \zeta^i \cdot \widehat{\ell}_{d+1})$, where ζ is the $2d$-th root of unity. Let $\check{\ell}_i := \widehat{\ell}_0 + \zeta^i \cdot \widehat{\ell}_{d+1} = (c_0 + \zeta^i c_{d+1})x_1 + (c'_0 + \zeta^i \cdot c'_{d+1})x_2 + \epsilon \cdot (\tilde{\ell}_0 + \zeta^i \tilde{\ell}_{d+1})$. Hence,

$$\prod_{i\in[d]} \widehat{\ell}_i + \prod_{i\in[d]} \check{\ell}_i \; = \; \epsilon^M \cdot f(R\mathbf{x}) + \epsilon^{M+1} \cdot \widehat{S} \,. \tag{6.28}$$

Let us define $c_{i,0} := c_i \bmod \langle\epsilon\rangle$, for $i \in \{0\} \cup [d+1]$, and similarly $c'_{i,0} := c'_i \bmod \langle\epsilon\rangle$. Therefore, $c_{0,0} = c'_{d+1,0} = 1$. Pick a set $\{\beta_1, \ldots, \beta_{d+1}\} =: Y \subseteq \mathbb{C}$, such that

1. $|Y| = d+1$, and
2. $c_{i,0}\beta_j + c'_{i,0}\beta_k \neq 0$, for any $\beta_j, \beta_k \in Y$, and $i \in [d]$.
3. $\beta_j + \zeta^i\beta_k \neq 0$, for $i \in [d]$, and $j, k \in [d+1]$.

Such a Y trivially exists (in fact a random Y also works). Once we have chosen Y, we now show how to deborder in this case.

The main idea is as follows: If we substitute $x_1 = \beta_1$, and $x_2 = \beta_2$ in Equation (6.28), by the property of Y, we get that

$$\prod_{i\in[d]} (\widehat{c}_i + \epsilon\tilde{\ell}_i) + \prod_{i\in[d]} (\check{c}_i + \epsilon \cdot (\tilde{\ell}_0 + \zeta^i\tilde{\ell}_{d+1})) \; = \; \epsilon^M \cdot f(R\mathbf{x})(\beta_1, \beta_2, x_3, \ldots, x_n) + \epsilon^{M+1} \cdot \tilde{S} \,. \tag{6.29}$$

In the above, $\widehat{c}_i := c_i\beta_1 + c'_i\beta_2$. Note that, $(\widehat{c}_i \bmod \langle\epsilon\rangle) = c_{i,0}\beta_1 + c'_{i,0}\beta_2 \neq 0$, by assumption. Similarly, $\check{c}_i := (c_0 + \zeta^i c_{d+1})\beta_1 + (c'_0 + \zeta^i \cdot c'_{d+1})\beta_2$. Since, $c_{0,0} = c'_{d+1,0} = 1$, by assumption (on Y): $(\check{c}_i \bmod \langle\epsilon\rangle) = \beta_1 + \zeta^i\beta_2 \neq 0$. Therefore, we can directly apply Corollary 6.19, to deduce that $\underline{\mathrm{WR}}_{\mathrm{aff}}(f(R\mathbf{x})(\beta_1, \beta_2, x_3, \ldots, x_n)) \leq O(d^2)$.

Note that the same holds if we substitute $x_1 = \beta_j$, and $x_2 = \beta_k$, for any $j, k \in [d+1]$. Therefore, we have

$$\underline{\mathrm{WR}}_{\mathrm{aff}}(f(R\mathbf{x})(\beta_j, \beta_k, x_3, \ldots, x_n)) \leq O(d^2), \quad \text{for } \; j,k \in [d+1] \overset{\text{Corollary 6.10}}{\Longrightarrow} \underline{\mathrm{WR}}(f(R\mathbf{x})) \leq O(d^7) \,.$$

In the above, we used the fact that the maximum exponent can be at most $O(d)$. Finally, by Lemma 6.12, $\underline{\mathrm{WR}}(f(R\mathbf{x})) \leq O(d^7) \implies \underline{\mathrm{WR}}(f) \leq O(d^7)$, as desired.

Case II(a): Case II(a) is almost the same as Case I(a), except one small point which we discuss now. Note that, we would not be able to directly factor $\widehat{\ell}_0^d + \epsilon^a\widehat{\ell}_{d+1}^d$, if $d \nmid a$. But, we can actually replace $\epsilon \mapsto \epsilon'^d$, for a variable ϵ'. Note that, $\epsilon \to 0 \iff \epsilon' \to 0$. By doing that, we can now factor $\widehat{\ell}_0^d + \epsilon'^{ad}\widehat{\ell}_{d+1}^d = \prod_{i\in[d]}(\widehat{\ell}_0 + \epsilon'^a\zeta^i\widehat{\ell}_{d+1})$. The rest follows similarly, by defining an appropriate Y, substituting β_j and using Corollary 6.19, and finally interpolating back f.

<u>**Case I(b):**</u> Case I(b) is as follows:

$$\prod_{i\in[d]} \ell_i + \ell_0^d + \ell_{d+1}^d \;=\; \epsilon^M \cdot f + \epsilon^{M+1} \cdot S(\boldsymbol{x}, \epsilon)\;;$$

where $\text{rank}(\ell_{0,0}, \ell_{d+1,0}) = 1$. By using Claim 6.27 (& its proof), it easily follows that $\text{rank}(\ell_{0,0}, \cdots, \ell_{d+1,0}) = 1$.

One can apply an invertible linear transformation ($\boldsymbol{x} \mapsto R\boldsymbol{x}$) that sends $\ell_{0,0} \mapsto x_1$. We can club x_1 terms together. So, $\widehat{\ell_i} := \ell_i(R\boldsymbol{x})$, for $i \in [0, d+1]$, becomes $c_i x_1 + \epsilon \cdot \tilde{\ell}_i$, where $\tilde{\ell}_i$ is x_1-free, $c_i \in \mathbb{C}[\epsilon]$, such that $c_i \not\equiv 0 \bmod \langle\epsilon\rangle$. Let us define $c_{i,0} := c_i \bmod \langle\epsilon\rangle$, which are non-zero. Fix any set (of distinct elements) $Y = \{\beta_1, \ldots, \beta_{d+1}\} \subseteq \mathbb{C}$. We would like to substitute $x_1 = (\beta_j - \epsilon \cdot \tilde{\ell}_0)/c_0$, for $j \in [d+1]$, so that $\widehat{\ell_0}$ becomes β_j. Since, we want to work over $\mathbb{C}[\epsilon]$, we can simply substitute $x_1 = (\beta_j - \epsilon \cdot \tilde{\ell}_0) \cdot \tilde{c}_0$, where $\mathbb{C}[\epsilon] \ni \tilde{c}_0 := 1/c_0 \bmod \langle\epsilon\rangle^{M+1}$. This exists since $c_{0,0} \neq 0$. Moreover, note that,

$$\lim_{\epsilon \to 0} f(R\boldsymbol{x})((\beta_j - \epsilon \cdot \tilde{\ell}_0) \cdot \tilde{c}_0, x_2, \ldots, x_n) = f(R\boldsymbol{x})(\beta_j / c_{0,0}, x_2, \ldots, x_n) := h_j(x_2, \ldots, x_n)\,.$$

We will show that $h_j(x_2, \ldots, x_n) \in \overline{\Sigma^{[O(d^2)]} \wedge \Sigma}$. By Lemma 6.9 and Lemma 6.12, it will follow that $f(\boldsymbol{x}) \in \overline{\Sigma^{[O(d^5)]} \wedge \Sigma}$.

After the substitution, we have the following:

$$\prod_{i\in[d]} (c'_{i,j} + \epsilon \cdot \ell'_i) + \beta_j^d + (c'_{d+1,j} + \epsilon \cdot \ell'_{d+1})^d \;=\; \epsilon^M \cdot h_j + \epsilon^{M+1} \cdot S'(\boldsymbol{x}, \epsilon)\;; \tag{6.30}$$

where $c'_{i,j} := c_i \beta_j \tilde{c}_0 \in \mathbb{C}[\epsilon]$, and $\ell'_i := \tilde{\ell}_i - c_i \tilde{c}_0 \tilde{\ell}_0$, for $i \in [0, d+1]$, and $j \in [d+1]$. Note that, $c'_{i,j} \bmod \langle\epsilon\rangle = c_{i,0}\beta_j / c_{0,0} \neq 0$. Moreover, $\ell'_i \in \mathbb{C}[\epsilon][x_2, \ldots, x_n]_1$, i.e., are x_1-free. We can further take out $\prod_i c'_{i,j}$, and divide both side and work over $\mathcal{R} := \mathbb{C}[\epsilon]/\langle\epsilon^{M+1}\rangle$. Clearly, the ϵ^M coefficient in RHS becomes $h_j/(c_{i,0}\beta_j/c_{0,0})$. In particular, after extracting and dividing (similar to the proof of Lemma 6.13), we have the following:

$$\prod_{i\in[d]} (1 + \epsilon \cdot \check{\ell}_i) \;+\; \alpha \cdot (1 + \epsilon \cdot \check{\ell}_{d+1})^d = \epsilon^M \cdot h_j/(c_{i,0}\beta_j/c_{0,0}) + \epsilon^{M+1} \cdot \check{S}\,.$$

By Theorem 6.17 and its Corollary 6.19, we deduce that indeed $h_j(x_2, \ldots, x_n) \in \overline{\Sigma^{[O(d^2)]} \wedge \Sigma}$, with maximum exponent being d. Therefore, as argued before, $\underline{\text{WR}}(f) \leq O(d^5)$. This finishes this case.

<u>Case II(b):</u> This is almost identical to Case I(b). In this case, we can make $\widehat{\ell}_{d+1}$ a constant after appropriate substitution (by x_1), and we will again get an expression of the form Equation (6.30).

<u>Case III(b):</u> Two possible cases can happen – (i) $\text{rank}(\ell_{0,0}, \ell_{d+1,0}) = \text{rank}(\ell_{0,0}, \ell_{1,0}, \cdots, \ell_{d+1,0}) = 1$, (ii) $\text{rank}(\ell_{0,0}, \ell_{d+1,0}) = 1$, but $\text{rank}(\ell_{0,0}, \ell_{1,0}, \cdots, \ell_{d+1,0}) = 2$.

The first case is again very similar to the above (Subcase 1), and hence we do not repeat it. For the second subcase, we will use an important structure. We rewrite this with slight change of sign (where $a > 0$):

$$\epsilon^a \cdot \prod_{i\in[d]} \ell_i + \ell_0^d - \ell_{d+1}^d \;=\; \epsilon^M \cdot f + \epsilon^{M+1} \cdot S(\boldsymbol{x}, \epsilon)\,.$$

Since, $\ell_0^d \equiv \ell_{d+1}^d \bmod \langle\epsilon\rangle^a$, by Theorem B.1, we know that, wlog, we can rewrite the above (after applying an invertible transformation $E \in \text{GL}_n(\mathbb{C})$ such that:

$$\epsilon^a \prod_{i\in[d]} \ell_i \;+\; (c_0 x_1 + \epsilon^a \check{\ell}_0)^d - (c_{d+1} x_1 + \epsilon^a \check{\ell}_{d+1})^d \;=\; \epsilon^M \cdot f(E\boldsymbol{x}) + \epsilon^{M+1} \cdot \check{S}(\boldsymbol{x}, \epsilon)\,,$$

where $c_i \in \mathbb{C}[\epsilon]$, such that $c_0 \equiv c_{d+1} \not\equiv 0 \bmod \langle\epsilon\rangle$, $\check{\ell}_i \in \mathbb{C}[\epsilon][x_2, \ldots x_n]_1$, and $\check{S} \in \mathbb{C}[\epsilon][\boldsymbol{x}]$. Let $\check{\ell}_i = \sum_j \check{\ell}_{i,j}\epsilon^j$, for $i \in \{0, d+1\}$. Since, $a < M$, $\prod_i \ell_{i,0} = c_{0,0}x_1^{d-1}(\check{\ell}_{d+1,0} - \check{\ell}_{1,0})$, where $\mathbb{C} \ni c_{0,0} \equiv c_0 \bmod \langle\epsilon\rangle$. Therefore, one can rescale the linear forms and assume that $\ell_{i,0} = x_1$, for $i \in [d-1]$, and $\ell_{d,0} = c_{0,0}(\check{\ell}_{d+1,0} - \check{\ell}_{1,0})$. By assumption, $\operatorname{rank}(\ell_{0,0}, \ldots, \ell_{d+1,0}) = 2$, it follows that x_1 and $\check{\ell}_{d+1,0} - \check{\ell}_{1,0}$ are linearly independent. Wlog, x_2 appears in the form $\check{\ell}_{d+1,0} - \check{\ell}_{1,0}$. As before, one can apply an invertible linear transformation A, which sends $x_1 \mapsto x_1, \check{\ell}_{d+1,0} - \check{\ell}_{1,0} \mapsto x_2$, and $x_i \mapsto x_i$, for $i \in [3, n]$. After this and clubbing appropriately, we get the following:

$$\epsilon^a \cdot (\tilde{c}_d x_2 + \epsilon\tilde{\ell}_d) \cdot \prod_{i\in[d-1]} (\tilde{c}_i x_1 + \epsilon\tilde{\ell}_i) + (\tilde{c}_0 x_1 + \epsilon^a\tilde{\ell}_0)^d - (\tilde{c}_{d+1}x_1 + \epsilon^a\tilde{\ell}_{d+1})^d \;=\; \epsilon^M \cdot f(B\boldsymbol{x}) + \epsilon^{M+1} \cdot \tilde{S}(\boldsymbol{x}, \epsilon)\;,$$

where $\tilde{c}_i \in \mathbb{C}[\epsilon]$, such that $\tilde{c}_i \equiv 1 \bmod \langle\epsilon\rangle$, for $i \in [d-1]$, and $\tilde{c}_0 \equiv \tilde{c}_d \equiv \tilde{c}_{d+1} \equiv c_{0,0} \bmod \langle\epsilon\rangle$; $\tilde{\ell}_i \in \mathbb{C}[\epsilon][x_2, \ldots x_n]_1$, for $i \in [d-1]$, and $\tilde{\ell}_d \in \mathbb{C}[\epsilon][x_1, x_3, \ldots, x_n]_1$, and $\tilde{S} \in \mathbb{C}[\epsilon][\boldsymbol{x}]$, and $B := A^{-1}E \in \mathrm{GL}_n(\mathbb{C})$. The proof goes similarly as before.

Fix $S = \{\beta_1, \ldots, \beta_{d+1}\} \subseteq \mathbb{C}$. Substitute $x_1 = (\beta_j - \epsilon^a\tilde{\ell}_1) \cdot \widehat{c}_0$, where $\mathbb{C}[\epsilon] \ni \widehat{c}_0 \equiv 1/\tilde{c}_0 \bmod \langle\epsilon\rangle^{M+1}$. The second term (power) becomes β_j^d, while the third term becomes $\widehat{c}_0\tilde{c}_{d+1}\beta_j + \epsilon^a x_2 + \epsilon^{a+1} \cdot \overline{\ell}_{d+1}$, where $\overline{\ell}_{d+1} \in \mathbb{C}[\epsilon][x_2, \ldots, x_n]_1$. This is particularly because of the map A so that, $\tilde{\ell}_{d+1,0} - \tilde{\ell}_{1,0} = x_2$. The coefficient of ϵ^M on RHS becomes $f(B\boldsymbol{x})(\beta_j, x_2, \ldots, x_n)$. Next, we also substitute $x_2 = (\beta_k - \epsilon\overline{\ell}_{d+1})$. Note that after this substitution, we get the following:

$$\epsilon^a \prod_{i\in[d]} (\gamma_i + \epsilon\widehat{\ell}_i) + \gamma_0 \;=\; \epsilon^M \cdot f(B\boldsymbol{x})(\beta_j, \beta_k, x_3, \ldots, x_n) + \epsilon^{M+1}\widehat{S}(x_3, \ldots, x_n, \epsilon)\;.$$

In the above, $\gamma_i \in \mathbb{C}[\epsilon]$, such that the ϵ-free terms of γ_i are nonzero, and $\widehat{\ell}_i \in \mathbb{C}[\epsilon][x_3, \ldots, x_n]_1$. This clearly implies that $\epsilon^a \mid \gamma_0$, and dividing out by ϵ^a both side, and invoking Theorem 6.14, $f(B\boldsymbol{x})(\beta_j, \beta_k, x_3, \ldots, x_n) \in \overline{\Sigma^{[O(d^2)]} \wedge \Sigma}$, with maximum exponent being at most d. Interpolating over β_i's using Corollary 6.10, we get that $f(B\boldsymbol{x}) \in \overline{\Sigma^{[O(d^7)]} \wedge \Sigma}$, which implies that $f(\boldsymbol{x}) \in \overline{\Sigma^{[O(d^7)]} \wedge \Sigma}$, by Lemma 6.12, as we wanted. This finishes the proof. □

6.f Lower Bounds

In this section, we prove several exponential separations between related polynomials contained in the affine closure of binomials.

6.31 Lemma. *The polynomial $\prod_{i\in[d]} x_i + x_{d+1}^d + x_{d+2}^d$ cannot be written as a product of linear forms.*

Remark. One can show that the above polynomial is in fact irreducible over $\mathbb{C}$. But for our purpose, the above suffices.

Proof. Any homogeneous polynomial f of degree d which is a product of linear forms, clearly has at most d essential variables. But $\prod_{i\in[d]} x_i + x_{d+1}^d + x_{d+2}^d$ clearly has $d+2$ essential variables. □

6.32 Lemma. *The polynomial $\prod_{i=1}^d x_i + \prod_{i=d+1}^{2d} x_i$ cannot be written as a product of linear forms.*

Proof. It easily follows from a proof similar to that of Lemma 6.31. □

6.33 Lemma. *For the polynomial $f(\boldsymbol{x}) := \prod_{i\in[d]} x_i + x_{d+1}^d + x_{d+2}^d$, we have $\underline{\mathrm{WR}}(f) \geq 2^{d-1}$.*

Proof. If $\underline{\mathrm{WR}}(f) < 2^{d-1}$ then by substituting $x_{d+1} = x_{d+2} = 0$, we obtain that $\underline{\mathrm{WR}}(\prod_{i\in[d]} x_i) < 2^{d-1}$, which contradicts the fact that $\underline{\mathrm{WR}}(\prod_{i\in[d]} x_i) = 2^{d-1}$, proved in [Oed19]. □

6.34 Theorem (First exp. gap theorem)**.** *If $P_2^{[d]} \trianglelefteq_{\text{aff}} P^{[e]}$, then $e \geq \exp(d)$.*

Proof. Let $P_2^{[d]} \trianglelefteq_{\text{aff}} P^{[e]}$. If one homogenizes wrt x_0, it requires (both-side padding), and the new formulation becomes $x_0^{e-d} \cdot P_2^{[d]} \trianglelefteq P^{[e]}$. By Theorem 6.20, we know that $x_0^{e-d} \cdot P_2^{[d]} \trianglelefteq P^{[e]}$ implies – either (i) $x_0^{e-d} \cdot P_2^{[d]} = \prod_{i\in[e]} \ell_i + \ell_0^e$, for some linear forms $\ell_i \in \mathbb{C}[\boldsymbol{x}]$, or (ii) $\underline{\mathrm{WR}}(x_0^{e-d} \cdot P_2^{[d]}) = O(e^5)$. We show that (i) is an impossibility while (ii) can happen only when $e \geq \exp(d)$.

Proof of Part (ii): Fix a random $x_0 = \alpha \in \mathbb{C}$. Note that, this implies that $P_2^{[d]} + \epsilon g = \sum_{i\in[k]} \ell_i^e$ for some affine forms $\ell_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]$ and $g \in \mathbb{C}[\epsilon][\boldsymbol{x}]$ with $k \in O(e^5)$. Since $P_2^{[d]}$ is homogeneous, this also implies that $\underline{\mathrm{WR}}(P_2^{[d]}) \leq k$. But then Lemma 6.33 implies that $k \geq 2^{d-1}$, which in turn implies that $e \geq \exp(d)$.

Proof of Part (i): Let $x_0^{e-d} \cdot P_2^{[d]} = \prod_{i\in[e]} \ell_i + \ell_0^e$. Note that, by a simple derivative space argument, one can show that the number of essential variables (for definition and characterization, see Theorem 5.2) in the LHS is at least $d+2$, while the number of essential variables of the expression in RHS is at most $e+1$; since trivially $\prod_{i\in T} \ell_i$, for $T \subset [e]$, such that $|T| = e-1$, and ℓ_0^{e-1} certainly span the space of single partial derivatives. Therefore, $e \geq d+1$. This will be important since we will use the fact that $e - d \geq 1$, in the below.

Further, we can assume that $x_0 \nmid \ell_0$. Otherwise, say $\ell_0 = c \cdot x_0$, for some $c \in \mathbb{C}$, which implies that $x_0^{e-d} \mid \prod_{i\in[e]} \ell_i$. Hence, wlog we can assume that $\ell_i = x_0$, for $i \in [e-d]$ (we are assuming constants to be 1, because we can always rescale and push the constants to the other linear forms). Therefore, RHS is divisible by x_0^{e-d}. By dividing it out and renaming the linear forms appropriately, we get

$$P_2^{[d]} = \prod_{i\in[d]} \widehat{\ell}_i + c x_0^d \,,$$

where $\widehat{\ell}_i \in \mathbb{C}[\boldsymbol{x}]$. Further, we can put $x_0 = 0$. Note that, $x_0 \nmid \widehat{\ell}_i$, for any i, since otherwise x_0 divides RHS, but it doesn't divide the LHS. After substituting $x_0 = 0$, we get that

$$P_2^{[d]} = \prod_{i\in[d]} \tilde{\ell}_i \,,$$

where $\mathbb{C}[x_1, \ldots, x_{d+2}] \ni \tilde{\ell}_i = \widehat{\ell}_i|_{x_0=0} \neq 0$. From Lemma 6.31, it follows that this is not possible. A similar argument shows that $x_0 \nmid \ell_i$, for any $i \in [d]$; because otherwise that implies $x_0 \mid \ell_0$, and hence the above argument shows a contradiction.

Therefore, we assume that $x_0 \nmid \ell_i$, for $i \in [0, d]$. Now, there are two cases – (i) x_0 appears in ℓ_0, (ii) x_0 does not appear in ℓ_0.

If x_0 appears in ℓ_0, then say $\ell_0 = c_0 x_0 + \widehat{\ell}_0$, for some $c_0 \neq 0$. Note that $\widehat{\ell}_0 \in \mathbb{C}[x_1, \ldots, x_{d+2}]_1$ is non-zero, since we assume that $x_0 \nmid \ell_0$. Substitute $x_0 = -\widehat{\ell}_0/c_0$ (so that ℓ_0 vanishes). This implies:

$$(-\widehat{\ell}_0/c_0)^{e-d} \cdot P_2^{[d]} = \prod_{i\in[e]} \widehat{\ell}_i \,,$$

where $\widehat{\ell}_i = \ell_i|_{x_0 = -\widehat{\ell}_0/c_0}$. Since LHS is non-zero, so is each $\widehat{\ell}_i$. Since, everything is homogeneous, and we have unique factorization, the above implies that upto renaming, $P_2^{[d]} = c \cdot \prod_{i\in[d]} \widehat{\ell}_i$, which is a contradiction by Lemma 6.31.

If x_0 does not appear in ℓ_0, then there must exist an $i \in [e]$ such that x_0 appears in ℓ_i, otherwise RHS is x_0-free which is trivially a contradiction. We also know that x_0 cannot divide ℓ_i, by our assumption. So, say $\ell_i = c_i x_0 + \widehat{\ell}_i$, where $\widehat{\ell}_i$ is x_0-free, and $c_i \in \mathbb{C}$ is a nonzero element. Substitute $x_0 = -\widehat{\ell}_i/c_i$, so that ℓ_i vanishes. Since ℓ_0 is x_0-free, we immediately get that

$$(-\widehat{\ell}_0/c_0)^{e-d} \cdot P_2^{[d]} \;=\; \ell_0^e \,.$$

Again, by unique factorization, we get that $P_2^{[d]} = c \cdot \ell_0^d$, for some $c \in \mathbb{C}$, which is a contradiction by Lemma 6.31. This finishes the proof. □

6.35 Theorem (Second exp. gap theorem)**.** *If* $\mathrm{bn}_d \trianglelefteq_{\mathrm{aff}} P_2^{(e)}$*, then* $e \geq \exp(d)$*.*

Proof. Let $\mathrm{bn}_d \trianglelefteq_{\mathrm{aff}} P_2^{(e)}$. If one homogenizes wrt x_0, it requires (both-side padding), and the new formulation becomes $x_0^{e-d} \cdot \mathrm{bn}_d \trianglelefteq P^{[e]}$. By Theorem 6.25 (& its remark), we know that $x_0^{e-d} \cdot \mathrm{bn}_d \trianglelefteq P_2^{(e)}$ implies – either (i) $x_0^{e-d} \cdot \mathrm{bn}_d = g + h$, where $g = \prod_{i \in [e]} \ell_i$, for some linear forms $\ell_i \in \mathbb{C}[\boldsymbol{x}]_1$, and $\underline{\mathrm{WR}}(h) \leq 2$, or (ii) $\underline{\mathrm{WR}}(x_0^{e-d} \cdot \mathrm{bn}_d) = O(e^7)$. Similarly, as before, we show that (i) is an impossibility while (ii) can happen only when $e \geq \exp(d)$. Part (ii) proof is exactly to the argument in the proof of Theorem 6.34.

To prove the Part (i), there are two cases – (a) $h = \ell_0^e + \ell_{e+1}^e$, for $\ell_i \in \mathbb{C}[\boldsymbol{x}]_1$, or, (b) $h = \ell_0^{e-1} \cdot \ell_{e+1}$.

Case (a): Let $x_0^{e-d} \cdot \mathrm{bn}_d = \prod_{i \in [e]} \ell_i + \ell_0^e + \ell_{e+1}^e$. We assume that x_0 does not divide ℓ_i, for some $i \in \{0, e+1\}$, and each ℓ_i, for $i \in [e]$, otherwise, we can divide by the maximum power of x_0 on both the sides.

Note that, by a simple derivative space argument, one can show that the number of essential variables (for definition and characterization, see Theorem 5.2) in the LHS is at least $2d$ (it is $2d+1$, if $e > d$), while the number of essential variables of the expression in RHS is at most $e+2$; since trivially $\prod_{i \in T} \ell_i$, for $T \subset [e]$, such that $|T| = e-1$, and $\ell_0^{e-1}, \ell_{e-1}^e$ certainly span the space of single partial derivatives. Therefore, $e \geq 2d - 2$.

Now, we divide this into subcases –

(a1) x_0 does not appear in ℓ_i, for any $i \in [e]$,
(a2) x_0 appears in ℓ_i, for some $i \in [e]$.

Case (a1): x_0 does not appear in ℓ_i, for $i \in [e]$. In that case, say $\ell_0 = c_0 x_0 + \widehat{\ell}_0$, and $\ell_{e+1} = c_{e+1} x_0 + \widehat{\ell}_{e+1}$, where $\widehat{\ell}_0$ and $\widehat{\ell}_{e+1}$ are x_0-free, and c_0, c_{e+1} are constants (might be 0 as well, but both cannot be 0 since then RHS becomes x_0-free). Therefore, the coefficient of x_0^{e-d} (as a polynomial) in RHS is $\gamma_0 \widehat{\ell}_0^d + \gamma_{e+1} \widehat{\ell}_{e+1}^d$, where $\gamma_0 = \binom{e}{d} c_0^{e-d}$, and similarly $\gamma_{e+1} = \binom{e}{d} c_{e+1}^{e-d}$. Comparing with LHS, we get that $\mathrm{bn}_d = \gamma_0 \widehat{\ell}_0^d + \gamma_{e+1} \widehat{\ell}_{e+1}^d$. Trivially, over $\mathbb{C}$, $\gamma_0 \widehat{\ell}_0^d + \gamma_{e+1} \widehat{\ell}_{e+1}^d$ is a product of linear forms, which is a contradiction, using Lemma 6.32.

Case (a2): If x_0 appears in one of the ℓ_i, it can appear in two ways, either ℓ_i is a constant multiple of x_0, or $\ell_i = c_i x_0 + \widehat{\ell}_i$, where $\widehat{\ell}_i$ is a nonzero linear form which is x_0-free. Let $S_1 \subseteq [e]$ such that x_0 appears in $\ell_i = c_i \cdot x_0$, for $i \in S_1$, for some nonzero constant $c_i \in \mathbb{C}$, and $S_2 \subseteq [e]$, such that $\ell_i = c_i x_0 + \widehat{\ell}_i$, where $\widehat{\ell}_i$ is nonzero.

Note that if $|S_1| + |S_2| < e - d$, then x_0^{e-d} cannot be contributed from the product and hence it only gets produced from $\ell_0^e + \ell_{e+1}^e$, and we get a contradiction in the same way as above. Hence, wlog assume that $|S_1| + |S_2| \geq e - d$.

If S_2 is non-empty, say $j \in S_2$, then substitute $x_0 = -\widehat{\ell}_j/c_j$, so that ℓ_j becomes 0. This substitution gives us the following:

$$(-\widehat{\ell}_j/c_j)^{e-d} \cdot \mathrm{bn}_d \;=\; \tilde{\ell}_0^e + \tilde{\ell}_{e+1}^e \,.$$

Since, $\tilde{\ell}_0^e + \tilde{\ell}_{e+1}^e$ can be written as a product of linear forms, from the unique factorization, it follows that f must be a product of linear forms, which is a contradiction from Lemma 6.32. Hence, we are done when $|S_2|$ is non-empty.

If S_2 is empty, since $|S_1| + |S_2| \geq e - d$ by assumption, we have $|S_1| \geq e - d$. In particular, $x_0^{e-d} \mid \text{LHS} - \prod \ell_i \implies x_0^{e-d} \mid \ell_0^e + \ell_{e+1}^e = \prod_i (\ell_0 - \zeta^i \ell_{e+1})$, where ζ is $2e$-th root of unity. Since, $e - d \geq 2$ for $d \geq 4$, this simply implies that there are two indices i_1 and i_2 such that $\ell_0 - \zeta^{i_1}\ell_{e+1} = c_{i_1}x_0$, and $\ell_0 - \zeta^{i_2}\ell_{e+1} = c_{i_2}x_0$. Together, this implies that both ℓ_0 and ℓ_{e+1} are multiples of x_0, which is a contradiction, since we assumed that x_0 cannot divide each ℓ_i, for $i \in [0, e+1]$. Hence, we are done with case (a).

Case (b): Let $x_0^{e-d} \cdot \mathrm{bn}_d = \prod_{i \in [e]} \ell_i + \ell_0^{e-1} \cdot \ell_{e+1}$. We assume that x_0 does not divide both ℓ_i, for some $i \in [e]$, and one of the ℓ_0 or ℓ_{e+1}, otherwise, we can divide by the maximum power x_0 both side. Again, a similar essential variable counting argument shows that $e \geq 2d - 2$.

Similarly, as before, we divide into subcases –

(b1) x_0 does not appear in ℓ_i, for any $i \in [e]$,

(b2) x_0 appears in ℓ_i, for some $i \in [e]$.

Case (b1): If x_0 does not appear in the first product, i.e,. any of ℓ_i, for $i \in [e]$, then it must appear in ℓ_0 (because if it only appears in ℓ_{e+1}, the degree of x_0 is 1 in RHS, a contradiction). Note that, $x_0 \nmid \ell_0$ (and similarly ℓ_{e+1}), because otherwise, substituting $x_0 = 0$ makes LHS 0, while RHS remains $\prod_{i \in [e]} \ell_i$. Hence, let $\ell_0 := c_0 x_0 + \widehat{\ell_0}$, where $\widehat{\ell_0}$ is x_0-free. Substitute $x_0 = -\widehat{\ell}_0 / c_0$, so that

$$(-\widehat{\ell}_0/c_0)^{e-d} \cdot \mathrm{bn}_d \;=\; \prod_{i \in [e]} \ell_i \;.$$

This in particular implies that bn_d is a product of linear forms, which is a contradiction by Lemma 6.32.

Case (b2): In this case, wlog x_0 appears in ℓ_1. Note that, x_0 cannot divide ℓ_1, because otherwise, it must divide LHS-$\prod_{i \in [e]} = \ell_0^{e-1}\ell_{e+1}$, which implies that x_0 must divide one of the ℓ_0 or ℓ_{e+1}, contradicting the minimality of x_0-division. Therefore, $\ell_1 = c_1 x_0 + \widehat{\ell}_1$, where c_1 is a nonzero constant, and $\widehat{\ell}_1$ is a nonzero linear form which is x_0-free. Substitute $x_0 = -\widehat{\ell}_1/c_1$, both side to get that

$$(-\widehat{\ell}_1/c_1)^{e-d} \cdot \mathrm{bn}_d \;=\; \widehat{\ell}_0^{e-1}\widehat{\ell}_{e+1} \;.$$

Therefore, again by unique factorization, we get that f must a product of linear forms, which is a contradiction by Lemma 6.32.

□

7 Homogeneous complexity and the parity-alternating elementary symmetric polynomial

7.a Input-homogeneous-linear computation

We start with a technicality in the definition of arithmetic circuits. In this section every edge of an arithmetic circuit is labelled with a field constant. Instead of just forwarding the computation result of a gate to another gate, these edges rescale the polynomial along the way. For arithmetic *formulas* we do *not* allow this, as we will see that it is unnecessary.

An arithmetic formula/circuit is called *input-homogeneous-linear* (IHL) if all its leaves are labelled with *homogeneous linear* forms, in particular (contrary to ordinary arithmetic formulas/circuits) we do *not* allow any leaf to be labelled with a field constant. It now becomes clear why we needed the technicality: For any $\alpha \in \mathbb{C}$, if an IHL circuit with s gates computes a polynomial f, then using the scalars on the edges there exists an IHL circuit computing αf with also only s many gates. For formulas this rescaling can be simulated by rescaling a subset of the leaves. Indeed, we rescale the root of the formula by induction: we rescale a summation gate by rescaling both children, we rescale a product gate by rescaling an arbitrary child. Alternatively, if f is homogeneous, one can rescale the input gates by the $\sqrt[d]{\alpha}$. The latter technique works for formulas and circuits alike, but we will not use this method. It is easy to see that IHL formulas/circuits can only compute polynomials f with $f(0) = 0$. But other than that, being IHL is not a strong restriction, as the following simple lemma shows. We write $\widehat{f} := f - f(0)$.

7.1 Lemma. *Given an arithmetic circuit of size s computing a polynomial f, then there exists an IHL arithmetic circuit of size $6s$ and depth $3s$ computing $\widehat{f}$.*

There exists a polynomial q such that: Given any arithmetic formula of size s computing a polynomial f, then there exists an IHL arithmetic formula of size $q(s)$ and depth $O(\log(s))$ computing $\widehat{f}$.

Proof. We treat the case of formulas first. We first use Brent's depth reduction [Bre74] to ensure that the size is $\mathsf{poly}(s)$ and the depth is $O(\log(s))$. We now proceed in a way that is similar to the homogenization of arithmetic circuits. Let F be the formula computing f. We replace every computation gate (that computes some polynomial g) by a pair of gates (and some auxiliary gates), one computing $g(0)$ and one computing $\widehat{g}$. Clearly, $\big((g+h)(0), \widehat{g+h}\big) = \big(g(0)+h(0), \widehat{g}+\widehat{h}\big)$, hence an addition gate is just replaced by 2 addition gates. Moreover, $\big((g \cdot h)(0), \widehat{g \cdot h}\big) = \big(g(0) \cdot h(0), g(0) \cdot \widehat{h} + \widehat{g} \cdot h(0) + \widehat{g} \cdot \widehat{h}\big)$, hence a multiplication gate is replaced by 4 multiplication gates and 2 addition gates (and this gadget has depth 3). We copy the subformulas of $g(0)$, $h(0)$, $\widehat{g}$, and $\widehat{h}$, which maintains the depth, and it keeps the size $\mathsf{poly}(s)$. In this construction additions happen only between constants or between non-constants, but never between a constant and a non-constant. Therefore each maximal subformula of constant gates can be evaluated and replaced with a single constant gate, and these gates are multiplied with non-constant gates (with the one exception of the gate for $f(0)$). But in a formula, scaling a non-constant gate by a field element does not require a multiplication gate, and instead we can recursively pass this scaling operation down to the children, as explained before this lemma. At the end we remove the one remaining constant gate for $f(0)$ and are done.

For circuits we proceed similarly. We skip the depth reduction step. Let C be the formula computing f. We replace every computation gate (that computes some polynomial g) by a pair of gates (and some auxiliary gates), one computing $g(0)$ and one computing $\widehat{g}$. Clearly, $\big((\alpha g + \beta h)(0), \widehat{\alpha g + \beta h}\big) = \big(\alpha g(0) + \beta h(0), \alpha\widehat{g} + \beta\widehat{h}\big)$, hence an addition gate is just replaced by 2 addition gates. Moreover, $\big((\alpha g \cdot \beta h)(0), \widehat{\alpha g \cdot \beta h}\big) = \big(\alpha g(0) \cdot \beta h(0), \alpha g(0) \cdot \beta\widehat{h} + \alpha\widehat{g} \cdot \beta h(0) + \alpha\widehat{g} \cdot \beta\widehat{h}\big)$, hence a multiplication gate is replaced by 4 multiplication gates and 2 addition gates (and this gadget has depth 3). Here we have no need to copy subformulas, and we re-use the computation instead. In this construction additions happen only between constants or between non-constants, but never between a constant and a non-constant. Therefore each maximal subcircuit of constant gates can be evaluated and replaced with a single constant gate v, and each of these gates is multiplied with a non-constant gate w (with the one exception of the gate for $f(0)$). This rescaling of the polynomial computed at w can be simulated by just rescaling all the edge labels of the outgoing edges of w, so v can be removed. At the end we also remove the one remaining constant gate for $f(0)$ and are done. $\square$

The following corollary is obvious.

7.2 Corollary. VP *is the set of p-families* $(f_n)_{n\in\mathbb{N}}$ *for which the IHL circuit size complexity of the sequence* $(\widehat{f}_n)_{n\in\mathbb{N}}$ *is polynomially bounded.* VF *is the set of p-families* $(f_n)_{n\in\mathbb{N}}$ *for which the IHL formula size complexity of the sequence* $(\widehat{f}_n)_{n\in\mathbb{N}}$ *is polynomially bounded.*

Proof. The missing constant can be added to the IHL circuit/formula as the very last operation. $\square$

7.b IHL Ben-Or and Cleve is exactly Kumar's complexity for 3×3 matrices

Quite surprisingly, the 3×3 matrix analogue of Kumar's complexity model turns out to be the homogeneous version of Ben-Or and Cleve's construction [BC92], as the proof of the following Proposition 7.3 shows. Let $E_{i,j}$ denote the 3×3 matrix with a 1 at the entry (i,j) and zeros elsewhere. Let id_3 denote the 3×3 identity matrix.

7.3 Proposition. *Fix* $i,j \in \{1,2,3\}$, $i \neq j$. *Let* f *be a polynomial admitting an IHL formula of depth* δ. *Then there exist* 3×3 *matrices* $A_1, \dots, A_r$ *with* $r \leq 4^\delta$ *having* homogeneous linear entries *and such that*

$$f \cdot E_{i,j} = (\mathrm{id}_3 + A_1)(\mathrm{id}_3 + A_2) \cdots (\mathrm{id}_3 + A_r) - \mathrm{id}_3.$$

Proof. Consider the six positions $\{(i,j) \mid 1 \leq i,j \leq 3,\ i \neq j\}$ of the zeros in the 3×3 unit matrix. Given an IHL formula, to each input gate and to each computation gate we assign one of the 6 positions in the following way. We start at the root and assign it position (i,j). We proceed by assigning position labels recursively: For a summation gate with position (i',j'), both summands get position (i',j'). For a product gate with position (i',j'), one factor gets position (i',k) and the other gets position (k,j'), $k \neq i'$, $k \neq j'$. We now prove by induction on the depth D of the gate g (the depth of a gate it the depth of its subformula: the input have depth 0; the root has the highest depth) with position (i',j') that for each gate there is a list of at most 4^D matrices $(A_1, \dots, A_r)$ such that

$$(\mathrm{id}_3 + A_1)(\mathrm{id}_3 + A_2) \cdots (\mathrm{id}_3 + A_r) = \mathrm{id}_3 + gE_{(i',j')}$$

and the same number of matrices $B_1, \dots, B_r$ such that

$$(\mathrm{id}_3 + B_1)(\mathrm{id}_3 + B_2) \cdots (\mathrm{id}_3 + B_r) = \mathrm{id}_3 - gE_{(i',j')}.$$

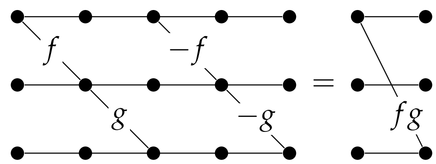
For an input gate (i.e., depth 0) with position (i',j') and input label ℓ, we set $A_1 := \ell \cdot E_{i',j'}$ and $B_1 := -\ell \cdot E_{i',j'}$. For an addition gate with position (i',j') let $(A_1, \dots, A_r)$, $(B_1, \dots, B_r)$ and $(A'_1, \dots, A'_{r'})$, $(B'_1, \dots, B'_{r'})$ be the lists coming from the induction hypothesis. We define the list for the addition gate as the concatenations $(A_1, \dots, A_r, A'_1, \dots, A'_{r'})$ and $(B_1, \dots, B_r, B'_1, \dots, B'_{r'})$. Observe that $(\mathrm{id}_3 + fE_{(i',j')}) \cdot (\mathrm{id}_3 + gE_{(i',j')}) = \mathrm{id}_3 + (f+g)E_{(i',j')}$ and that $(\mathrm{id}_3 - fE_{(i',j')}) \cdot (\mathrm{id}_3 - gE_{(i',j')}) = \mathrm{id}_3 - (f+g)E_{(i',j')}$, so this case is correct. For a product gate with position (i',j') let $(A_1, \dots, A_r)$, $(B_1, \dots, B_r)$ and $(A'_1, \dots, A'_{r'})$, $(B'_1, \dots, B'_{r'})$ be the lists coming from the induction hypothesis, i.e., $(\mathrm{id}_3 + A_1)(\mathrm{id}_3 + A_2) \cdots (\mathrm{id}_3 + A_r) = \mathrm{id}_3 + fE_{(i',k)}$, $(\mathrm{id}_3 + B_1)(\mathrm{id}_3 + B_2) \cdots (\mathrm{id}_3 + B_r) = \mathrm{id}_3 - fE_{(i',k)}$, $(\mathrm{id}_3 + A'_1)(\mathrm{id}_3 + A'_2) \cdots (\mathrm{id}_3 + A'_r) = \mathrm{id}_3 + gE_{(k,j')}$, $(\mathrm{id}_3 + B'_1)(\mathrm{id}_3 + B'_2) \cdots (\mathrm{id}_3 + B'_r) = \mathrm{id}_3 - gE_{(k',j')}$. Observe that

$$\big(\mathrm{id}_3 + fE_{(i',k)}\big)\big(\mathrm{id}_3 + gE_{(k,j')}\big)\big(\mathrm{id}_3 - fE_{(i',k)}\big)\big(\mathrm{id}_3 - gE_{(k,j')}\big) = \mathrm{id}_3 + fgE_{(i',j')}$$

and analogously

$$\big(\mathrm{id}_3 - fE_{(i',k)}\big)\big(\mathrm{id}_3 + gE_{(k,j')}\big)\big(\mathrm{id}_3 + fE_{(i',k)}\big)\big(\mathrm{id}_3 - gE_{(k,j')}\big) = \mathrm{id}_3 - fgE_{(i',j')}.$$

For illustration, in the notation of [BIZ18] the product with position (1,3) can be depicted as follows.



Since $4 \cdot 4^{D-1} = 4^D$, the size bound is satisfied. $\square$

Since the trace of a matrix can sometimes be preferrable to the (i,j)-entry, we present the result with the trace, provided approximations are allowed.

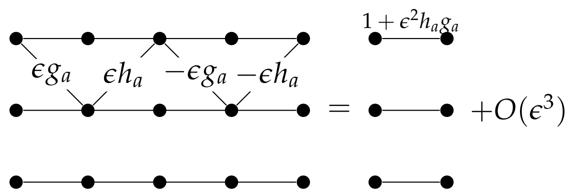
7.4 Proposition. *For every IHL formula of depth δ there exist $\leq 4^\delta$ many 3×3 matrices A_i with* homogeneous linear entries *over $\mathbb{C}[\epsilon, \epsilon^{-1}]$ and $\alpha \in \mathbb{C}[\epsilon, \epsilon^{-1}]$ such that*

$$E_{1,1} \cdot f = \lim_{\epsilon \to 0} \Big(\alpha \big((\mathrm{id}_3 + A_1)(\mathrm{id}_3 + A_2) \cdots (\mathrm{id}_3 + A_r) - \mathrm{id}_3 \big) \Big).$$

Proof. The IHL formula is a sum of products of subformulas $g_1 \cdot h_1, g_2 \cdot h_2, \ldots, g_r \cdot h_r$, and $r \leq 2^\delta$ by induction. We compute subformulas for ϵg_1, $-\epsilon g_1$, ϵh_1, $-\epsilon h_1$, ϵg_2, $-\epsilon g_2$, $\ldots$, $-\epsilon h_r$ as in the proof of Proposition 7.3 with position $(1,2)$ for each $\pm \epsilon g_i$ and position $(2,1)$ for each $\pm \epsilon h_i$. It turns out that

$$M_a := (\mathrm{id}_3 + \epsilon g_a E_{1,2})(\mathrm{id}_3 + \epsilon h_a E_{2,1})(\mathrm{id}_3 - \epsilon g_a E_{1,2})(\mathrm{id}_3 - \epsilon h_a E_{2,1}) = \mathrm{id}_3 + \epsilon^2 f_a g_a E_{1,1} + O(\epsilon^3).$$

Pictorially:



Hence $M_1 M_2 \cdots M_r = \mathrm{id}_3 + \epsilon^2 (h_1 g_1 + h_2 g_2 + \cdots + h_r g_r) + O(\epsilon^3)$. We choose $\alpha = \epsilon^{-2}$. $\square$

Let $\binom{[n]}{d}$ denote the set of cardinality d subsets of $[n]$. For a subset $S \subseteq [n]$ let $\mathrm{sort}(S)$ denote the tuple whose elements are the elements of S, sorted in ascending order. Let $\mathrm{sort}(\binom{[n]}{d}) := \{\mathrm{sort}(S) \mid S \in \binom{[n]}{d}\}$. Let $\bar{e}_d(X_1, \ldots, X_n) := \sum_{I \in \mathrm{sort}(\binom{[n]}{d})} X_{I_1} \cdots X_{I_d}$ denote the elementary symmetric polynomial over noncommuting variables $X_1, \ldots, X_n$.

7.5 Corollary. *Fix any nonzero linear form L on the space of 3×3 matrices, for example the trace. If L is supported outside the main diagonal, then the collection $L(\bar{e}_d(A_1, \ldots, A_n))$, where $A_i = \begin{pmatrix} 0 & x_{1,2,i} & x_{1,3,i} \\ x_{2,1,i} & 0 & x_{2,3,i} \\ x_{3,1,i} & x_{3,2,i} & 0 \end{pmatrix}$, is p-hard for VF_H, otherwise it is $\overline{\mathsf{p}}$-hard for VF_H.*

Proof. Given a p-family (f) of homogeneous polynomials. If f_n has polynomially bounded arithmetic formula size, then it also has IHL formulas of logarithmic depth and polynomial size (apply Brent's depth reduction and then Lemma 7.1). The first case is treated with Proposition 7.3, the second is treated completely analogously with Proposition 7.4. We only handle the slightly more difficult second case. We obtain $4^{O(\log n)} = \mathrm{poly}(n)$ many matrices A_i with

$$f_n = \lim_{\epsilon \to 0} L \Big(\alpha \big((\mathrm{id}_3 + A_1)(\mathrm{id}_3 + A_2) \cdots (\mathrm{id}_3 + A_r) - \mathrm{id}_3 \big) \Big)$$

As in §4 we can assume that $\alpha = \beta\epsilon^k$ is a scalar times a power of ϵ. Since f_n is homogeneous of some degree d, we have

$$f_n = \lim_{\epsilon\to 0} L\Big(\alpha\overline{e}_d(A_1,\ldots,A_r)\Big) = \lim_{\epsilon\to 0} L\Big(\overline{e}_d(\sqrt[d]{\beta}\epsilon^k A'_1,\ldots,\alpha\sqrt[d]{\beta}\epsilon^k A'_r)\Big)$$

where A'_i arises from A_i by replacing every ϵ by ϵ^d. □

While Corollary 7.5 gives the first collection that is hard for VF_H, the polynomials are similarly unwieldy as $\mathrm{IMM}_n^{(d)}$. In the next sections we will prove that the parity-alternating elementary symmetric polynomial is $\overline{\mathrm{p}}$-hard for a class V3F, which gives a polynomial that is just barely more complicated than the elementary symmetric polynomial.

7.c IHL computation with arity 3 products

In the light of [BIZ18] we now study the 2×2 analogues of Proposition 7.3, Proposition 7.4, Corollary 7.5. In order to do so, in this section we study IHL formulas and circuits where the additions have arity 2, but the products have *arity exactly 3*. We call this basis the *arity 3 basis*. This turns out to be rather subtle, because one would want to simulate an arity 2 product by an arity 3 product in which one of the factors is a constant 1, but that violates the IHL property. If a polynomial is computed by an IHL formula or circuit over the arity 3 basis, then all its homogeneous even-degree parts are zero. We will mostly study homogeneous polynomials that are computed over this basis. We want to also compute homogeneous even-degree polynomials f in this basis, so we define that a multi-output IHL circuit/formula over the arity 3 basis *computes* f if it computes each partial derivative $\partial f/\partial x_i$ at some output gate. Analogously to Corollary 7.2, but only for homogeneous polynomials, we define V3P and V3F to be the set of homogeneous p-families $(f_n)_{n\in\mathbb{N}}$ for which the IHL circuit/formula complexity over the arity 3 basis is polynomially bounded. For a complexity class C we write $\mathsf{C}_H := \mathsf{C}\cap H$ for brevity. We have

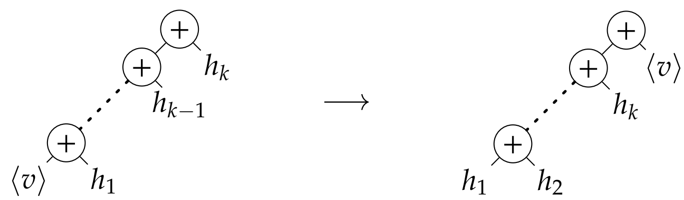
$$\begin{array}{ccccccc} \mathsf{V3F} & \subseteq & \mathsf{VF}_H & \subseteq & \mathsf{VBP}_H & \subseteq & \mathsf{VP}_H, \\ & & \rotatebox{90}{$\subseteq$} & & & & \\ & & \mathsf{V3P} & & & & \end{array} \tag{7.6}$$

where we prove the vertical inclusion in Theorem 7.13, and $\mathsf{V3F} \subseteq \mathsf{VF}_H$ follows from Euler's homogeneous function theorem that $f = \frac{1}{\deg(f)}\sum_{i=1}^m x_i \cdot \partial f/\partial x_i$, which lets us treat the even degrees (arity 3 formulas for odd degree polynomials can be directly converted gate by gate into the standard basis). Is is known that if we go to quasipolynomial complexity instead of polynomial complexity, the three classical classes coincide: $\mathsf{VQF} = \mathsf{VQBP} = \mathsf{VQP}$, which is an immediate corollary of the circuit depth reduction result of Valiant-Berkowitz-Skyum-Rackoff [VSBR83]. We prove in Theorem 7.14 that our two new classes also belong to this set: All classes in (7.6) coincide if we go to quasipolynomial complexity instead of polynomial complexity, see (7.15).

The following proposition is an adaption of Brent's depth reduction [Bre74] and it shows that instead of polynomially sized formulas we can work with formulas of logarithmic depth. Both properties, IHL and the arity 3 basis, require some changes to Brent's original argument.

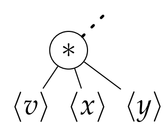
7.7 Proposition (Brent's depth reduction for IHS formulas over the arity 3 basis)**.** *Let f be a polynomial computed by an IHL formula of size s over the arity 3 basis. Then there exists an IHL formula over the arity 3 basis of size* $\mathrm{poly}(s)$ *and depth* $O(\log(s))$ *computing* f*.*

Proof. We discuss the odd-degree case, because in the even-degree case we just have one odd-degree case for each partial derivative. The construction is recursive, just as in Brent's original argument. We follow the description in [Sap21]. We start at the root and keep picking the child with the larger subformula until we reach a vertex v with $\frac{1}{3}s \leq |\langle v \rangle| \leq \frac{2}{3}s$, where $\langle v \rangle$ is the subformula at the gate v. We make a case distinction. In the first case we assume that on the path from from v to the root (excluding v) there is no product gate. We reorder the gates as follows:

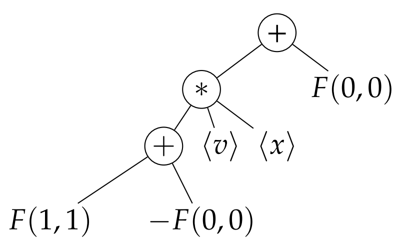


The construction applied to a size s formula gives $\text{Depth}(s) \leq \text{Depth}(\frac{2}{3}s) + 1$. The resulting size is $\text{Size}(s) \leq 2 \cdot \text{Size}(\frac{2}{3}s) + 1$.

In the second case we assume that v is the child of a product gate.



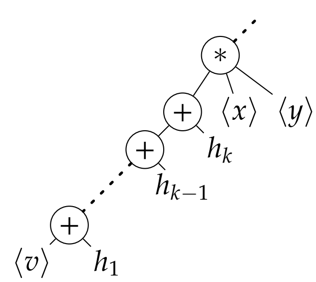
We now replace $\langle v \rangle$ by a new variable α and $\langle x \rangle$ by a new variable β. We observe that the resulting polynomial F (interpreted as a bivariate polynomial in α and β) is *linear* in the product $\alpha\beta$. Therefore $F(\alpha, \beta) = \alpha\beta(F(1,1) - F(0,0)) + F(0,0)$. Both $F(0,0)$ and $F(1,1)$ can be realized as an IHL formula over the arity 3 basis (because an arity 3 product gate with two 1s as inputs can be replaced by just the third input, and an arity 3 product gate with two 0s as input can be replaced by a constant 0, which can be simulated by removing gates), so we obtain:



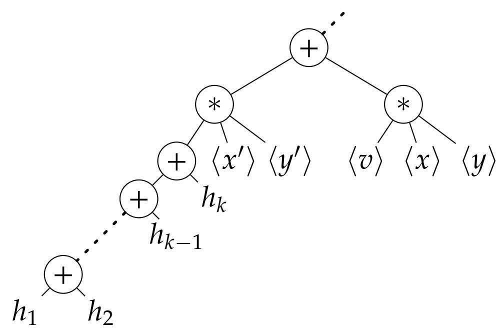
(7.8)

The construction on a size s formula gives $\text{Depth}(s) \leq \text{Depth}(\frac{2}{3}s) + 2$. The resulting size is $\text{Size}(s) \leq 5 \cdot \text{Size}(\frac{2}{3}s) + 3$.

In the third case we assume that on the path from from v to the root (excluding v) there are addition gates and then a product gate, so



As a first step we make copies of $\langle x \rangle$ and $\langle y \rangle$ and call them $\langle x' \rangle$ and $\langle y' \rangle$, respectively, and re-wire similarly as in the first case:



On the right-hand side of the tree we now proceed analogously as in the second case. We replace $\langle v \rangle$ by a new variable α and $\langle x \rangle$ by a new variable β. We observe that the resulting polynomial F (interpreted as a bivariate polynomial in α and β) is linear in the product $\alpha\beta$. Therefore $F(\alpha, \beta) = \alpha\beta(F(1,1) - F(0,0)) + F(0,0)$. Both $F(0,0)$ and $F(1,1)$ can be realized as an input-homogeneous formula over the arity 3 basis, so we obtain the same formula as in (7.8). The construction on a size s formula gives $\mathrm{Depth}(s) \le \mathrm{Depth}(\frac{2}{3}s) + 2$. The resulting size is $\mathrm{Size}(s) \le 5 \cdot \mathrm{Size}(\frac{2}{3}s) + 3$. Putting all cases together, the construction has $\mathrm{Depth}(s) \le \mathrm{Depth}(\frac{2}{3}s) + 2$ and $\mathrm{Size}(s) \le 5 \cdot \mathrm{Size}(\frac{2}{3}s) + 3$. Hence applying the construction recursively gives logarithmic depth and polynomial size. □

7.d The parity-alternating elementary symmetric polynomial

Let n be odd. For odd i let $X_i = \begin{pmatrix} 0 & x_i \\ 0 & 0 \end{pmatrix}$, and for even i let $X_i = \begin{pmatrix} 0 & 0 \\ x_i & 0 \end{pmatrix}$. Let $A := \bar{e}_d(X_1, X_2, \ldots, X_n)$. Note that in row 1 the matrix A has only one nonzero entry, and its position depends on the parity of n. Let $C_{n,d} := A_{1,1} + A_{1,2}$. A sequence a of integers numbers is called *parity-alternating* if $a_i \neq a_{i+1} \mod 2$ for all i, and a_1 is odd. Let P be the set of length d increasing parity-alternating sequences of numbers from $\{1, \ldots, n\}$. It is easy to see that

$$C_{n,d} = \sum_{(i_1, i_2, \ldots, i_d) \in P} x_{i_1} x_{i_2} \cdots x_{i_d}. \tag{7.9}$$

We usually only consider the case when the parities of d and n coincide, which is justified by the following lemma.

7.10 Lemma. *If n and d have different parity, then $C_{n,d} = C_{n-1,d}$.*

Proof. If d is odd, each parity-alternating sequence always ends with an odd parity, so if n is even we have $C_{n,d} = C_{n-1,d}$. If d is even, each parity-alternating sequence always ends with an even parity, so if n is odd we have $C_{n,d} = C_{n-1,d}$. □

Analogously to Corollary 7.5 we have the following theorem.

7.11 Theorem. *$C_{n,d}$ is $\overline{\mathsf{V3F}}$-$\overline{\mathrm{p}}$-hard and $\overline{\mathsf{VQP}_H}$-$\overline{\mathrm{qp}}$-hard.*

Proof. We start with proving $\mathsf{V3F}$-$\overline{\mathrm{p}}$-hardness (which is the same as $\overline{\mathsf{V3F}}$-$\overline{\mathrm{p}}$-hardness). Given $(f) \in \mathsf{V3F}$, then according to Proposition 7.7 we can assume that either (if f_n is of odd degree) f_n has polynomially sized formulas of logarithmic depth $\delta = O(\log n)$, or (if f_n is of even degree) its partial derivatives have polynomially sized formulas of logarithmic depth $\delta = O(\log n)$. We can assume that the gates are additions and negative cubes ($x \mapsto -x^3$), because $xyz = \frac{1}{24}\big((x+y+z)^3 - (x+y-z)^3 - (x-y+z)^3 + (x-y-z)^3\big)$, and the rescalings by $(\pm 24)^{-\frac{1}{3}}$ can be pushed to the input gates. Let d be the degree of f_n. Let $E_{\text{odd}} = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and let $E_{\text{even}} = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ and let id_2 denote the 2×2 identity matrix. We first treat the case of d being odd. We prove by induction on the depth D of a gate that there exist $\leq 3^D$ homogeneous linear forms $\ell_1, \ldots, \ell_r$ over $\mathbb{C}[\epsilon, \epsilon^{-1}, \alpha]$ such that

$$\alpha f_n \cdot E_{\text{odd}} \simeq (\mathrm{id}_2 + \ell_1 E_{\text{odd}})(\mathrm{id}_2 + \ell_2 E_{\text{even}}) \cdots (\mathrm{id}_2 + \ell_r E_{\text{odd}}) - \mathrm{id}_2$$

The induction starting at an input gate with label ℓ is done by $\ell_1 = \alpha\ell$. The addition gate is handled as follows. By induction hypothesis there exist $\ell_1, \ldots, \ell_r$ and $\ell'_1, \ldots, \ell'_{r'}$ with

$$\alpha f \cdot E_{\text{odd}} + \mathrm{id}_2 \simeq (\mathrm{id}_2 + \ell_1 E_{\text{odd}})(\mathrm{id}_2 + \ell_2 E_{\text{even}}) \cdots (\mathrm{id}_2 + \ell_r E_{\text{odd}})$$

and

$$\alpha g \cdot E_{\text{odd}} + \mathrm{id}_2 \simeq (\mathrm{id}_2 + \ell'_1 E_{\text{odd}})(\mathrm{id}_2 + \ell'_2 E_{\text{even}}) \cdots (\mathrm{id}_2 + \ell'_{r'} E_{\text{odd}})$$

Therefore $\alpha(f+g) \cdot E_{\text{odd}} + \mathrm{id}_2 = (\alpha f \cdot E_{\text{odd}} + \mathrm{id}_2)(\alpha g \cdot E_{\text{odd}} + \mathrm{id}_2) \simeq$

$$(\mathrm{id}_2 + \ell_1 E_{\text{odd}})(\mathrm{id}_2 + \ell_2 E_{\text{even}}) \cdots (\mathrm{id}_2 + \ell_r E_{\text{odd}})(\mathrm{id}_2 + \ell'_1 E_{\text{odd}})(\mathrm{id}_2 + \ell'_2 E_{\text{even}}) \cdots (\mathrm{id}_2 + \ell'_{r'} E_{\text{odd}})$$

Handling the negative cube gates is more subtle (the negative squaring gates are also the subtle cases in [BIZ18]). By induction hypothesis we have $\ell_1, \ldots, \ell_r$ such that

$$\alpha f \cdot E_{\text{odd}} \simeq (\mathrm{id}_2 + \ell_1 E_{\text{odd}})(\mathrm{id}_2 + \ell_2 E_{\text{even}}) \cdots (\mathrm{id}_2 + \ell_r E_{\text{odd}}) - \mathrm{id}_2 \tag{7.12}$$

We replace each ϵ by ϵ^k in each ℓ_i, with k so large that even when we replace α by ϵ^{-1} or $-\epsilon^{-1}$, we still have the equivalence of the LHS and RHS mod ϵ^2. We call the resulting linear forms ℓ'_i. It follows that

$$\alpha f \cdot E_{\text{odd}} \equiv \big((\mathrm{id}_2 + \ell'_1 E_{\text{odd}})(\mathrm{id}_2 + \ell'_2 E_{\text{even}}) \cdots (\mathrm{id}_2 + \ell'_r E_{\text{odd}}) - \mathrm{id}_2\big) \pmod{\epsilon^k}$$

Setting α to ϵ^{-1} we obtain

$$\epsilon^{-1} f \cdot E_{\text{odd}} \equiv \big((\mathrm{id}_2 + \ell''_1 E_{\text{odd}})(\mathrm{id}_2 + \ell''_2 E_{\text{even}}) \cdots (\mathrm{id}_2 + \ell''_r E_{\text{odd}}) - \mathrm{id}_2\big) \pmod{\epsilon^2}$$

Anaogously with $\alpha = -\epsilon^{-1}$:

$$-\epsilon^{-1} f \cdot E_{\text{odd}} \equiv \big((\mathrm{id}_2 + \tilde{\ell}''_1 E_{\text{odd}})(\mathrm{id}_2 + \tilde{\ell}''_2 E_{\text{even}}) \cdots (\mathrm{id}_2 + \tilde{\ell}''_r E_{\text{odd}}) - \mathrm{id}_2\big) \pmod{\epsilon^2}$$

The induction hypothesis (7.12) also implies (set ϵ to ϵ^3 and α to $\epsilon^2\alpha$) that

$$\epsilon^2 \alpha f \cdot E_{\text{odd}} \equiv \big((\mathrm{id}_2 + \ell'''_1 E_{\text{odd}})(\mathrm{id}_2 + \ell'''_2 E_{\text{even}}) \cdots (\mathrm{id}_2 + \ell'''_r E_{\text{odd}}) - \mathrm{id}_2\big) \pmod{\epsilon^3}$$

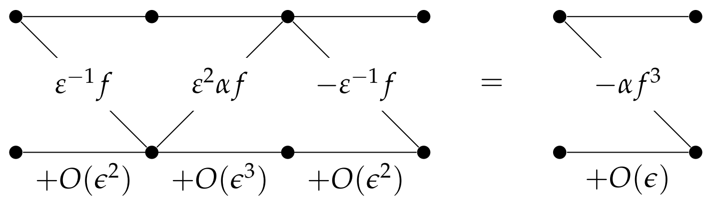
Transposing gives

$$\epsilon^2 \alpha f \cdot E_{\text{even}} \equiv \big((\mathrm{id}_2 + \ell'''_r E_{\text{even}})(\mathrm{id}_2 + \ell'''_{r-1} E_{\text{odd}}) \cdots (\mathrm{id}_2 + \ell'''_1 E_{\text{even}}) - \mathrm{id}_2\big) \pmod{\epsilon^3}$$

We now observe:

$$(\epsilon^{-1} f E_{\text{odd}} + \mathrm{id}_2 + \epsilon^2 g_1)(\epsilon^2 \alpha f E_{\text{even}} + \mathrm{id}_2 + \epsilon^3 g_2)(-\epsilon^{-1} f E_{\text{odd}} + \mathrm{id}_2 + \epsilon^2 g_3) \simeq -\alpha f^3 E_{\text{odd}} + \mathrm{id}_2.$$

Pictorially:



At the end, setting $\alpha = 1$ we obtain

$$\alpha f_n \cdot E_{\text{odd}} \simeq (\text{id}_2 + \ell_1 E_{\text{odd}})(\text{id}_2 + \ell_2 E_{\text{even}}) \cdots (\text{id}_2 + \ell_r E_{\text{odd}}) - \text{id}_2,$$

where r is only polynomially large, because we started with a formula of logarithmic depth. Since f_n is homogeneous of degree d, this implies

$$f_n \;\simeq\; \bar{e}(\ell_1 E_{\text{odd}}, \ell_2 E_{\text{even}}, \cdots, \ell_r E_{\text{odd}})_{1,2} \;=\; C_{r,d}(\ell_1, \ldots, \ell_r).$$

We now treat the case where f_n has even degree, using an argument similar to the one form Proposition 7.4. By the above construction, for each i we find

$$\alpha\left(\tfrac{1}{\deg f_n} \partial f_n / \partial x_i\right) \cdot E_{\text{odd}} \simeq (\text{id}_2 + \ell_{i,1} E_{\text{odd}})(\text{id}_2 + \ell_{i,2} E_{\text{even}}) \cdots (\text{id}_2 + \ell_{i,r_i} E_{\text{odd}}) - \text{id}_2.$$

We replace all ϵ by ϵ^3, replace all α by ϵ, and lastly add id_2:

$$\epsilon\left(\tfrac{1}{\deg f_n} \partial f_n / \partial x_i\right) \cdot E_{\text{odd}} + \text{id}_2 \equiv \left((\text{id}_2 + \ell'_{i,1} E_{\text{odd}})(\text{id}_2 + \ell'_{i,2} E_{\text{even}}) \cdots (\text{id}_2 + \ell'_{i,r_i} E_{\text{odd}})\right) \pmod{\epsilon^3}.$$

Analogously, when replacing α by $-\epsilon$ instead:

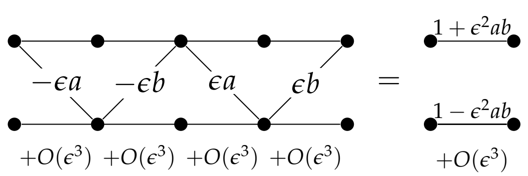
$$-\epsilon\left(\tfrac{1}{\deg f_n} \partial f_n / \partial x_i\right) \cdot E_{\text{odd}} + \text{id}_2 \equiv \left((\text{id}_2 + \ell''_{i,1} E_{\text{odd}})(\text{id}_2 + \ell''_{i,2} E_{\text{even}}) \cdots (\text{id}_2 + \ell''_{i,r_i} E_{\text{odd}})\right) \pmod{\epsilon^3}.$$

We also find corresponding linear forms for the transposes. Now observe that for any polynomials a, b we have

$$(-\epsilon a \cdot E_{\text{odd}} + \text{id}_2 + O(\epsilon^3))(-\epsilon b \cdot E_{\text{even}} + \text{id}_2 + O(\epsilon^3))(\epsilon a \cdot E_{\text{odd}} + \text{id}_2 + O(\epsilon^3))(\epsilon b \cdot E_{\text{even}} + \text{id}_2 + O(\epsilon^3))$$

$$\equiv \begin{pmatrix} 1 + \epsilon^2 a \cdot b & 0 \\ 0 & 1 - \epsilon^2 a \cdot b \end{pmatrix} \pmod{\epsilon^3}.$$

Pictorially:



Let $M(c) := \begin{pmatrix} 1 + \epsilon^2 c & 0 \\ 0 & 1 - \epsilon^2 c \end{pmatrix}$. Now note that

$$(M(a_1 b_1) + O(\epsilon^3)) \cdot (M(a_2 b_2) + O(\epsilon^3)) \cdots (M(a_n b_n) + O(\epsilon^3)) \equiv M(a_1 b_1 + a_2 b_2 + \cdots a_n b_n) \pmod{\epsilon^3}.$$

Setting $a_i = x_i$ and $b_i = \frac{1}{\deg f_n} \partial f_n / \partial x_i$, and using Euler's homogeneous function theorem, we obtain polynomially many linear forms $\ell_1, \ldots, \ell_r$ so that

$$M(f_n) \equiv \left((\text{id}_2 + \ell_1 E_{\text{odd}})(\text{id}_2 + \ell_2 E_{\text{even}}) \cdots (\text{id}_2 + \ell_r E_{\text{even}})\right) \pmod{\epsilon^3}$$

Subtracting id_2 on both sides and taking the degree d homogeneous part of the $(1,1)$ entry (note that f_n is homogeneous of degree d):

$$\epsilon^2 f_n \equiv \underbrace{\overline{e}_d(\ell_1 E_{\text{odd}}, \ell_2 E_{\text{even}}, \cdots, \ell_r E_{\text{even}})_{1,1}}_{=C_{r,d}(\ell_1,\ldots,\ell_r)} \pmod{\epsilon^3}$$

We replace all ϵ by $\epsilon^{d/2}$:

$$\epsilon^d f_n \equiv C_{r,d}(\ell'_1, \ldots, \ell'_r) \pmod{\epsilon^{3d/2}}.$$

Therefore

$$f_n \simeq C_{r,d}(\epsilon^{-1} \cdot \ell'_1, \ldots, \epsilon^{-1} \cdot \ell'_r).$$

Both cases together prove that $C_{n,d}$ is $\overline{\mathsf{V3F}}$-$\overline{\mathrm{p}}$-hard. The $\overline{\mathsf{VQP}}_H$-$\overline{\mathrm{qp}}$-hardness now follows from Theorem 7.14. $\square$

7.e Converting formulas to circuits over the arity 3 basis

In this section we prove the following theorem.

7.13 Theorem. $\mathsf{VF}_H \subseteq \mathsf{V3P}$.

Proof. Let $(f) \in \mathsf{VF}_H$. (f) has formulas of polynomial size and logarithmic depth. If f_n is of even degree, observe that if f_n has a formula of depth δ, then $\partial f_n/\partial x_i$ has a formula of depth 2δ (by induction, using the sum and product rules of derivatives, using the fact that the depth is logarithmic), which by Lemma 7.1 implies the existence of an IHL formula of depth $O(\delta)$ (note that $\partial f_n/\partial x_i$ is homogeneous of odd degree). Now we apply the odd-degree argument below for each partial derivative independently.

Let f_n be of odd degree. As a first step we convert the IHL formula into an IHL formula for which at each gate either all even homogeneous components vanish or all odd homogeneous components vanish. The construction is similar to the Lemma 7.1). It works as follows. We replace each gate v by two gates v_{odd} and v_{even}, where at v_{even} the sum of the odd degree components is computed, and at v_{odd} the sum of the odd degree components is computed. Let $f = f_{\text{even}} + _{\text{odd}}$ be the decomposition of f into the even homogeneous parts and the odd homogeneous parts. $((f+g)_{\text{even}}, (f+g)_{\text{odd}}) = (f_{\text{even}} + g_{\text{even}}, f_{\text{odd}} + g_{\text{odd}})$ so a sum gate is replaced by two sum gates. Moreover, $((f \cdot g)_{\text{even}}, (f \cdot g)_{\text{odd}}) = (f_{\text{even}} \cdot g_{\text{even}} + f_{\text{odd}} \cdot g_{\text{odd}}, f_{\text{even}} \cdot g_{\text{odd}} + f_{\text{odd}} \cdot g_{\text{even}})$, so a product gate is replaced by 4 product gates and 2 summation gates. Here we use that the depth was logarithmic.

We now convert such a formula to an IHL circuit with the same number of gates, but over the arity 3 basis. We replace each even degree gate v that computes g with a gate that computes $z \cdot g$, where z is a dummy variable. Addition gates are not changed. For product gates there are three cases.

- A product gate v of two odd-degree polynomials f and g. By induction we have an IHL circuit over the arity 3 basis for f and for g. We construct the arity 3 product $z \times f \times g$.
- A product gate v that has an odd-degree polynomial f at its child w, and that has an even-degree polynomial g at its child u. By induction we have IHL circuits C and D over the arity 3 basis for f and for zg, respectively. We take C and D, delete all instances of z in D, and feed the output of C instead. The resulting circuit computes fg.

- A product of an even-degree polynomial f and an even-degree polynomial g. By induction we have IHL circuits C and D over the arity 3 basis for zf and for zg, respectively. We take C and D, delete all instances of z in D, and feed the output of C instead. The resulting circuit computes zfg.

The size of the resulting circuit is less or equal to the size of the formula (even though the depth can increase in this construction). $\square$

A short remark: Note that the replacements of z in the second and third bullet point can only be done, because in a formula the outdegree of each gate is at most 1, i.e., we do not reuse computation results. After we replace z by f in a subcircuit that computes zg, the original subcircuit computing zg will be gone and cannot be reused.

7.f Valiant-Skyum-Berkowitz-Rackoff over the arity 3 basis

7.14 Theorem. $\mathsf{VQ3F} = \mathsf{VQ3P}$.

Proof. The entire argument is over the arity 3 basis. Given a size s circuit that computes an odd-degree polynomial, we use Theorem 7.16 to obtain a circuit of size $\mathsf{poly}(s)$ and depth $O(\log^2(s))$ that computes the same polynomial. We unfold the circuit to a formula of the same depth. The size is hence $3^{O(\log^2(s))} = s^{O(\log s)}$. If $s = n^{\mathsf{polylog}(n)}$, then $s^{O(\log s)} = n^{\mathsf{polylog}(n)}$ [11]. The even-degree case is done by treating each partial derivative independently. $\square$

Since we know that $\mathsf{VQF}_H = \mathsf{VQBP}_H = \mathsf{VQP}_H$ and $\mathsf{VQ3F} = \mathsf{VQF}_H = \mathsf{VQ3P}$, the situation of (7.6) simplifies:

$$\mathsf{VQ3F} = \mathsf{VQF}_H = \mathsf{VQBP}_H = \mathsf{VQP}_H = \mathsf{VQ3P}. \tag{7.15}$$

The following Theorem 7.16 is needed in the proof of Theorem 7.14. It lifts the classical Valiant-Skyum-Berkowitz-Rackoff [VSBR83] circuit depth reduction to the arity 3 basis. The argument is an adaption of the original argument.

7.16 Theorem (VSBR depth reduction for IHL circuits over the arity 3 basis)**.** *Let f be a polynomial computed by an IHL circuit of size s over the arity 3 basis, $\deg(f) = d$. Then there exists an IHL circuit over the arity 3 basis of size $O(\mathsf{poly}(s))$ and depth $O(\log(s) \cdot \log d)$ computing f.*

Proof. We adapt the proof from [Sap21]. We treat the odd case, because in the even degree case we can treat each partial derivative independently. We work entirely over the arity 3 basis (and hence compute a polynomial whose even degree homogeneous parts all vanish), so every circuit and subcircuit is over the arity 3 basis, and every product is of arity 3. A circuit whose root is an arity 3 product gate is denoted by $x \times y \times z$. A circuit whose root is an arity 2 addition gate is denoted by $x + y$, just as usual. Notationally, we use the same notation for gates, for their subcircuits, and for the polynomials they compute. If we want to specifically highlight that we talk about the circuit with root w, then we write $\langle w \rangle$. We write $v \le u$ is v is contained in the subcircuit with root u. We write $C \equiv C'$ to denote that the circuits C and C' compute the same polynomial.

Let z be a new dummy variable. Let the circuit $[u : v]$ be defined via $[u : v] := z$ if $u = v$, and if $u \neq v$ we have

$$[u : v] \;:=\; \begin{cases} 0 & \text{if } u \text{ is a leaf} \\ [u_1 : v] + [u_2 : v] & \text{if } u = u_1 + u_2 \\ [u_1 : v] \times u_2 \times u_3 & \text{if } u = u_1 \times u_2 \times u_3 \text{ and } u_1 \text{ has the highest} \\ & \text{degree among } \{[u_1], [u_2], [u_3]\} \end{cases}$$

[11] $(n^{\log^i(n)})^{\log^j(n^{\log^i(n)})} = n^{\log^{i+ij+j}(n)}$

It can be seen by induction that $[u : v]$ is zero or a homogeneous polynomial of degree $\deg u - \deg v + 1$, and $[u : v]$ is zero or is homogeneous linear in z. If $w \not\leq u$, then $[u : w] = 0$. For a circuit C we write $[u : v]_C := [u : v](z \leftarrow C)$, where $\leftarrow$ means that all leaves labelled z are replaced by the output of the circuit C.

We define a set of gates that is called the *m-frontier* $\mathcal{F}_m$ via

$$\mathcal{F}_m := \{u \mid u = u_1 \times u_2 \times u_3 \text{ with } \deg u_1, \deg u_2, \deg u_3 \leq m \text{ and } \deg(u) > m\}.$$

7.17 Lemma. *Fix a pair (u, m) with $\deg u > m$. Let $\mathcal{F} := \mathcal{F}_m$. Then*

$$u \equiv \sum_{w \in \mathcal{F}} [u : w]_{\langle w \rangle}.$$

Proof. For the proof we fix m and do induction on the *depth* of u, i.e., the position of u in any fixed topological ordering of the gates. Since for every gate u with $\deg(u) > m$ there exists some gate $u' \in \mathcal{F} \cap \langle u \rangle$, the induction start is the case $u \in \mathcal{F}$. In this case, since $\mathcal{F}$ is an antichain, it follows that $\sum_{w \in \mathcal{F}} [u : w] = 0 + [u : u] = z$, and hence $\sum_{w \in \mathcal{F}} [u : w]_{\langle w \rangle} = [u : u]_{\langle u \rangle} = z_{\langle u \rangle} = u$. This proves that case $u \in \mathcal{F}$. Now, let $u \notin \mathcal{F}$. If u is an addition gate:

$$\begin{aligned}
u &= u_1 + u_2 \\
&\overset{\text{I.H.}}{\equiv} \sum_{w \in \mathcal{F}} [u_1 : w]_{\langle w \rangle} + \sum_{w \in \mathcal{F}} [u_2 : w]_{\langle w \rangle} \\
&\equiv \sum_{w \in \mathcal{F}} \Big([u_1 : w]_{\langle w \rangle} + [u_2 : w]_{\langle w \rangle} \Big) \\
&= \sum_{w \in \mathcal{F}} \Big([u_1 : w] + [u_2 : w] \Big)_{\langle w \rangle} \\
&\overset{\text{Def.}}{=} \sum_{w \in \mathcal{F}} [u : w]_{\langle w \rangle}
\end{aligned}$$

If u is a multiplication gate, note that $u \notin \mathcal{F}$, so one of the children has degree $> m$ (w.l.o.g. that child is called u_1):

$$\begin{aligned}
u &= u_1 \times u_2 \times u_3 \\
&\overset{\text{I.H.}}{\equiv} \left(\sum_{w \in \mathcal{F}} [u_1 : w]_{\langle w \rangle} \right) \times u_2 \times u_3 \\
&\equiv \sum_{w \in \mathcal{F}} \Big([u_1 : w]_{\langle w \rangle} \times u_2 \times u_3 \Big) \\
&= \sum_{w \in \mathcal{F}} \Big([u_1 : w] \times u_2 \times u_3 \Big)_{\langle w \rangle} \\
&\overset{\text{Def.}}{=} \sum_{w \in \mathcal{F}} [u : w]_{\langle w \rangle}
\end{aligned}$$

$\square$

7.18 Lemma. *Fix a pair (u, m, v) with $\deg u > m \geq \deg v$. Let $\mathcal{F} := \mathcal{F}_m$.*

$$[u : v] \equiv \sum_{w \in \mathcal{F}} [u : w]_{[w:v]}.$$

Proof. For the proof we fix m and v and do induction on the *depth* of u, i.e., the position of u in any fixed topological ordering of the gates. Since for every gate u with $\deg(u) > m$ there exists some gate $u' \in \mathcal{F} \cap \langle u \rangle$, the induction start is the case $u \in \mathcal{F}$. In this case, since $\mathcal{F}$ is an antichain, it follows that $\sum_{w \in \mathcal{F}} [u : w]_{[w:v]} \equiv z_{[u:v]} = [u : v]$. This proves that case $u \in \mathcal{F}$. Now, let $u \notin \mathcal{F}$. Since $\deg u > m$ and $m \geq \deg v$ we have $u \neq v$. If u is an addition gate:

$$\begin{aligned}
[u : v] &\overset{\text{Def. } (u \neq v)}{=} [u_1 : v] + [u_2 : v] \\
&\overset{\text{I.H.}}{\equiv} \sum_{w \in \mathcal{F}} [u_1 : w]_{[w:v]} + \sum_{w \in \mathcal{F}} [u_2 : w]_{[w:v]} \\
&\equiv \sum_{w \in \mathcal{F}} \left([u_1 : w]_{[w:v]} + [u_2 : w]_{[w:v]} \right) \\
&= \sum_{w \in \mathcal{F}} \left([u_1 : w] + [u_2 : w] \right)_{[w:v]} \\
&\overset{\text{Def.}}{\equiv} \sum_{w \in \mathcal{F}} [u : w]_{[w:v]}
\end{aligned}$$

If u is a multiplication gate, note that $u \notin \mathcal{F}$, so one of the children has degree $> m$ (w.l.o.g. that child is called u_1):

$$\begin{aligned}
[u : v] &\overset{\text{Def. } (u \neq v)}{=} [u_1 : v] \times u_2 \times u_3 \\
&\overset{\text{I.H.}}{\equiv} \left(\sum_{w \in \mathcal{F}} [u_1 : w]_{[w:v]} \right) \times u_2 \times u_3 \\
&\equiv \sum_{w \in \mathcal{F}} \left([u_1 : w]_{[w:v]} \times u_2 \times u_3 \right) \\
&= \sum_{w \in \mathcal{F}} \left([u_1 : w] \times u_2 \times u_3 \right)_{[w:v]} \\
&\overset{\text{Def.}}{\equiv} \sum_{w \in \mathcal{F}} [u : w]_{[w:v]}
\end{aligned}$$

□

We now construct the shallow circuit so that the degree of each child in a multiplication gate decreases from δ to $\lceil \frac{2}{3}\delta \rceil$, so the multiplication depth (i.e., the number of multiplications on a path from leaf to root) is at most $O(\log d)$. Here we allow arity 5 multiplication gates. These can be simulated by two arity 3 multiplication gates. We construct the circuit by induction on the degree, and we construct it in a way that each u and each $[u : w]_{\langle v \rangle}$ are computed at some gate, so the size of the resulting circuit is at most $O(s^3)$. The addition gates between the multiplications can be balanced, so that we have at most $O(\log s)$ depth in each addition tree. This gives a total depth of $\log d \cdot \log s$.

The construction for u

$$\begin{aligned}
u &\overset{\text{Lem. 7.17}}{\equiv} \sum_{w \in \mathcal{F}} [u : w]_{\langle w \rangle} = \sum_{w \in \mathcal{F}} [u : w]_{\langle w_1 \rangle} \times w_2 \times w_3 = \sum_{\substack{w \in \mathcal{F} \\ \deg(u) \geq \deg(w)}} [u : w]_{\langle w_1 \rangle} \times w_2 \times w_3 \\
&\equiv \sum_{\substack{w \in \mathcal{F} \\ \deg(u) \geq \deg(w)}} [u : w]_{\langle w_3 \rangle} \times w_2 \times w_1
\end{aligned}$$

This explicit rearrangement of w_1 and w_3 is necessary and goes beyond [VSBR83]. Choose $m = \lceil \frac{2}{3} \deg u \rceil$. Recall $\deg w_i \leq m$, so we already have two of the three cases: $\deg w_1 \leq \lceil \frac{2}{3} \deg u \rceil$ and $w_2 \leq \lceil \frac{2}{3} \deg u \rceil$. But we also know $\deg(u) \geq \deg(w) = \deg(w_1) + \deg(w_2) + \deg(w_3)$, hence w.l.o.g. $\deg(w_3) \leq \lfloor \frac{1}{3} \deg(u) \rfloor$. Therefore $\deg u - \deg w + \deg w_3 \leq \lfloor \frac{4}{3} \rfloor \deg u - \underbrace{\deg w}_{>m} < \frac{2}{3} \deg u$.

The construction for $[u{:}v]$ We use fractions and "$\cdot$" multiplication signs when we do not have a circuit implementation in the intermediate equalities on polynomials. We write $w = w_1 \times w_2 \times w_3$ for $w \in \mathcal{F}$.

$$
\begin{aligned}
[u : v] &\overset{\text{Lem. 7.18}}{\equiv} \sum_{w \in \mathcal{F}} [u : w]_{[w:v]} = \sum_{\substack{w \in \mathcal{F} \\ \deg(u) \geq \deg(w)}} \frac{[u : w]}{z} \cdot [w : v] = \frac{1}{z} \sum_{\substack{w \in \mathcal{F} \\ \deg(u) \geq \deg(w)}} [u : w] \cdot [w_1 : v] \cdot w_2 \cdot w_3 \\
&\equiv \sum_{\substack{w \in \mathcal{F} \\ \deg(u) \geq \deg(w)}} [u : w]_{\langle w_3 \rangle} \times [w_1 : v] \times w_2 \\
&\overset{\text{Lem. 7.17}}{\equiv} \sum_{\substack{w \in \mathcal{F} \\ \deg(u) \geq \deg(w)}} [u : w]_{\langle w_3 \rangle} \times [w_1 : v] \times \left(\sum_{\substack{y \in \mathcal{F}' \\ \deg(w_2) \geq \deg(y)}} [w_2 : y]_{\langle y_3 \rangle} \times y_2 \times y_1 \right) \\
&\equiv \sum_{\substack{w \in \mathcal{F} \\ \deg(u) \geq \deg(w)}} \sum_{\substack{y \in \mathcal{F}' \\ \deg(w_2) \geq \deg(y)}} [u : w]_{\langle w_3 \rangle} \times [w_1 : v] \times \left([w_2 : y]_{\langle y_3 \rangle} \times y_2 \times y_1 \right)
\end{aligned}
$$

We set $m = \lceil \frac{2}{3} (\deg u + \deg v) \rceil$ and $m' = \lceil \frac{2}{3} \deg w_2 \rceil$. We calculate the degrees of the five factors:

- $\deg u - \deg w + \deg w_3 \leq (\deg u - \deg w) + \lfloor \frac{1}{3} \deg u \rfloor \leq \lfloor \frac{4}{3} \deg u \rfloor - m \leq \lceil \frac{2}{3} (\deg u - \deg v) \rceil$
- $\deg w_1 - \deg v + 1 \leq \deg w_1 \leq m \leq \lceil \frac{2}{3} (\deg u - \deg v) \rceil$
- $\deg w_2 - \deg y + \deg y_3 \leq \lfloor \frac{4}{3} \deg w_2 \rfloor - \lceil \frac{2}{3} \deg w_2 \rceil \leq \lceil \frac{2}{3} \deg w_2 \rceil \leq \lceil \frac{2}{3} (\deg u - \deg v) \rceil$
- $\deg y_2 \leq \lceil \frac{2}{3} \deg w_2 \rceil \leq \lceil \frac{2}{3} (\deg u - \deg v) \rceil$, and analogously for $\deg y_1$.

The rescaling constants on the edges can be set in the straightforward way. $\square$

8 De-bordering border Waring rank

8.a Orbit closure and essential variables

Recall from Theorem 5.2 that the number of essential variables of a polynomial $f \in S^d \mathbb{C}^n$ is the rank of its first catalecticant map. We prove a structural result for orbit-closures of polynomials with non-maximal number of essential variables.

8.1 Proposition. *Let $V = \langle x_1, \ldots, x_n \rangle$ and $W = \langle x_1, \ldots, x_r \rangle \subseteq V$. Let $f \in S^d W \subseteq S^d V$ be a homogeneous polynomial. Then*

$$\overline{\mathrm{GL}(V) \cdot f} = \mathrm{GL}(V) \overline{(\mathrm{GL}(W) \cdot f)}.$$

Proof. Clearly

$$\mathrm{GL}(V) \cdot f \subseteq \mathrm{GL}(V) \overline{(\mathrm{GL}(W) \cdot f)} \subseteq \overline{\mathrm{GL}(V) \cdot f}.$$

Therefore, it suffices to show that $\mathrm{GL}(V)\overline{(\mathrm{GL}(W)\cdot f)}$ is closed. Write $X_W = \overline{\mathrm{GL}(W)\cdot f}$. Moreover, for every $E \subseteq V$ with $\dim E = r$, let $X_E = g \cdot X_W$, where $g \in \mathrm{GL}(V)$ is any element such that $g \cdot W = E$. Notice that X_E does not depend on the choice of g.

Let $\mathrm{Gr}(r, V)$ be the Grassmannian of r-planes in V and consider the incidence correspondence

$$\begin{tikzcd} & \mathcal{I} = \{([h], E) : h \in X_E\} \quad \subseteq \mathbb{P}S^d V \times \mathrm{Gr}(r, V) \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ \mathbb{P}S^d V & & \mathrm{Gr}(r, E) \end{tikzcd}$$

where π_1, π_2 are the two projections. Then $\mathcal{I}$ is a fiber bundle over $\mathrm{Gr}(r, V)$ where the fiber over E is X_E. In particular $\mathcal{I}$ is closed in $\mathbb{P}S^d V \times \mathrm{Gr}(r, V)$. The image of a projective morphism is closed, so $\pi_1(\mathcal{I}) \subseteq \mathbb{P}S^d V$ is closed. This image is, by construction $\mathrm{GL}(V) \cdot X_W$ and this concludes the proof. $\square$

8.b Fixed-parameter de-bordering

Our proof is based on generalized additive decompositions of polynomial, in the sense of [Iar95, BBM14]. These decompositions were studied in the context of algebraic geometry, usually in connection to 0-dimensional schemes and cactus rank. We defer the discussion of connections to algebraic geometry in the next section. Here we provide elementary proofs of some statements on generalized additive decompositions based on partial derivatives techniques, without using the language of algebraic geometry. We bring from geometry a key insight: a border rank decomposition can be separated into *local* parts if the degree of the polynomial is large enough.

To define formally what it means for a border rank decomposition to be local, note that a rational family of linear forms $\ell \in \mathbb{C}(\varepsilon)[\boldsymbol{x}]_1$ always has a limit when viewed projectively. Specifically, if $\ell(\varepsilon) = \sum_{i=q}^{\infty} \varepsilon^i \ell_i$ as Laurent series, then $\lim_{\varepsilon\to 0}[\ell(\varepsilon)] = \lim_{\varepsilon\to 0}[\sum_{i=0}^{\infty} \varepsilon^i \ell_{q+i}] = [\ell_q]$. A border Waring rank decomposition is called local if for all summands in the decomposition this limit is the same. More precisely, we give the following definition.

8.2 Definition. *Let $f \in \mathbb{C}[\boldsymbol{x}]_d$ be a homogeneous polynomial. A border Waring rank decomposition*

$$f = \lim_{\varepsilon\to 0} \sum_{k=1}^{r} \ell_k^d$$

with $\ell_k \in \mathbb{C}(\varepsilon)[\boldsymbol{x}]_1$ is called a local border decomposition *if there exists a linear form $\ell \in \mathbb{C}[\boldsymbol{x}]_1$ such that $\lim_{\varepsilon\to 0}[\ell_k(\varepsilon)] = [\ell]$ for all $k \in \{1, \ldots, r\}$. We call the point $[\ell] \in \mathbb{P}\mathbb{C}[\boldsymbol{x}]_1$ the* base *of the decomposition. A local decomposition is called* standard *if $\ell_1 = \varepsilon^q \gamma \ell$ for some $q \in \mathbb{Z}$ and $\gamma \in \mathbb{C}$.*

8.3 Lemma. *If f has a local border decomposition, then it has a standard local border decomposition with the same base and the same number of summands.*

Proof. After applying a linear change of variables, we can assume that the base of the local decomposition for f is $[x_1]$. This means

$$f = \lim_{\varepsilon\to 0} \sum_{k=1}^{r} \ell_k^d$$

with $\ell_k = \varepsilon^{q_k} \cdot \gamma_k x_1 + \sum_{j=q_k+1}^{\infty} \varepsilon^j \ell_{k,j}$.

Write $\ell_1 = \varepsilon^{q_1}\left(\sum_{i=1}^n \alpha_i x_i\right)$ where $\alpha_i \in \mathbb{C}(\varepsilon)$. Let $\widehat{x}_1 = \frac{\gamma_1}{\alpha_1}x_1 - \sum_{i=2}^n \frac{\alpha_n}{\alpha_1}x_i$. Note that $\alpha_1 \simeq \gamma_1$ and $\alpha_i \simeq 0$ for $i > 1$, hence $\widehat{x}_1 \simeq x_1$ and

$$f \simeq f(\widehat{x}_1, \ldots, x_n) \simeq \ell_1(\widehat{x}_1, x_2, \ldots, x_n)^d + \sum_{k=2}^r \ell_k(\widehat{x}_1, x_2, \ldots, x_n)^d = (\varepsilon^{q_1}\gamma_1 x_1)^d + \sum_{k=2}^r \widehat{\ell}_k^d.$$

where $\widehat{\ell}_k(x_1, \ldots, x_n) = \ell_k(\widehat{x}_1, x_2, \ldots, x_n)$. This defines a new border rank decomposition of f. Moreover, notice $\lim_{\varepsilon\to 0}[\widehat{\ell}_k] = [x_1]$ for every k, so the new decomposition is again local with base $[x_1]$. Since the first summand is $\epsilon^{q_1}\gamma_1 x_1$, this is the desired standard local border decomposition. ☐

8.4 Lemma. *Suppose $f \in S^d V$ has a local border decomposition with r summands based at $[\ell]$. If $d \geq r-1$, then $f = \ell^{d-r+1}g$ for some homogeneous polynomial g of degree $r-1$.*

Proof. After applying a linear change of variables we can assume $\ell = x_1$. We prove the statement by induction on the difference $d-(r-1)$.

The case $d = r-1$ is trivial. If $d > r-1$, then by the previous Lemma there exists a standard local border decomposition

$$f = \lim_{\varepsilon\to 0}\sum_{k=1}^r l_k(\varepsilon)^d.$$

where $l_k = \sum_{i=1}^n \alpha_{ki}x_i$ for some $\alpha_{ki} \in \mathbb{C}(\varepsilon)$. Since the decomposition is standard, $\alpha_{1i} = 0$ for $i > 1$. For the derivatives of f we have the following border decompositions.

$$\frac{\partial f}{\partial x_1} = \lim_{\varepsilon\to 0}\sum_{k=1}^r d\cdot\alpha_{k1}(\varepsilon)l_k(\varepsilon)^{d-1}.$$

and

$$\frac{\partial f}{\partial x_i} = \lim_{\varepsilon\to 0}\sum_{k=2}^r d\cdot\alpha_{ki}(\varepsilon)l_k(\varepsilon)^{d-1}.$$

for $i > 1$. These decompositions involve the same linear forms ℓ_k with multiplicative coefficients, they are still local with the same base $[x_1]$. By inductive hypothesis $\frac{\partial f}{\partial x_1} = x_1^{d-r}g_1$ and $\frac{\partial f}{\partial x_i} = x_1^{d-r+1}g_i$ for some homogeneous polynomials $g_1, \ldots, g_n$ of appropriate degrees. To get an analogous expression for f, combine these expressions using Euler's formula for homogeneous polynomials as follows

$$f = \frac{1}{d}\sum_{i=0}^n x_i\frac{\partial f}{\partial x_i} = \frac{1}{d}\left(x_1\cdot x_1^{d-r}g_1 + \sum_{i=2}^n x_i x_1^{d-r+1}g_i\right) = \frac{1}{d}x_1^{d-r+1}\left(g_1 + \sum_{i=2}^n x_i g_i\right)$$ ☐

We will now extend this result to non-local border Waring rank decompositions. As long as the degree of the approximated polynomial is high enough, every border rank decomposition can be divided into local parts and transformed into a sum of terms of the form $\ell^{d-r+1}g$.

8.5 Definition. *A* generalized additive decomposition *of f is a decomposition of the form*

$$f = \sum_{k=1}^m \ell_k^{d-r_k+1}g_k$$

where ℓ_k are linear forms such that ℓ_i is not proportional to ℓ_j when $i \neq j$, and g_k are homogeneous polynomials of degrees $\deg g_k = r_k - 1$.

To show that there is no cancellations between different local parts, we need the following lemma, which in the case of 2 variables goes back to Jordan [IK99, Lem. 1.35].

8.6 Lemma. *Let $\ell_1, \dots, \ell_m \in \mathbb{C}[\boldsymbol{x}]_1$ be linear forms such that ℓ_i is not proportional to ℓ_j when $i \neq j$. Let $g_1, \dots, g_m$ be homogeneous polynomials of degrees $r_1 - 1, \dots, r_m - 1$ respectively. If*

$$\sum_{k=1}^{m} \ell_k^{d-r_k+1} g_k = 0$$

and $d \geq \sum_{k=1}^{m} r_i - 1$, then all g_k are zero.

Proof. We first prove the statement for polynomials in 2 variables y_1, y_2 by induction on the number of summands m; this proof follows [GY10, Appx.III].

The case $m = 1$ with one summand is clear. Consider the case $m > 2$. We can assume $\ell_1 = y_1$ by applying a linear change of variables if required. Note two simple facts about partial derivatives. First, for a homogeneous polynomial $f \in \mathbb{C}[y_1, y_2]_d$ we have $\partial_2^r f = 0$ if and only if $f = y_1^{d-r+1} g$ (here $\partial_2 := \frac{\partial}{\partial y_2}$). Second, differentiating r times a homogeneous polynomial of the form $\ell^{d-s+1} g$, we obtain a polynomial of the form $\ell^{d-r-s+1} h$.

Suppose

$$y_1^{d-r_1+1} g_1 + \sum_{k=2}^{m} \ell_k^{d-r_k+1} g_k = 0.$$

Differentiating r_1 times with respect to y_2, we obtain

$$\sum_{k=2}^{m} \ell_k^{d-r_1-r_k+1} h_k = 0,$$

where $\ell_k^{d-r_1-r_k+1} h_k = \partial_2^{r_1} (\ell_k^{d-r_k+1} g_k)$. The degree condition $d - r_1 \geq \sum_{k=2}^{m} r_k - 1$ holds for this new expression. Therefore, by induction hypothesis we have $h_k = 0$ and thus $\partial_2^{r_1} (\ell_k^{d-r_k+1} g_k) = 0$. It follows that $\ell_k^{d-r_k+1} g_k = y_1^{d-r_1+1} \widehat{g}_k$ for some homogeneous polynomial $\widehat{g}_k$. This implies that $y_1^{d-r_1+1}$ divides g_k, which is impossible since $d - r_1 + 1 \geq \sum_{k=2}^{m} r_k \geq r_k > \deg g_k$.

Consider now the general case where the number of variables $n \geq 2$ (the case $n = 1$ is trivial). Suppose $\sum_{k=1}^{m} \ell_k^{d-r_k+1} g_k = 0$. The set of linear maps $A \colon (y_1, y_2) \mapsto (x_1, \dots, x_n)$ such that $\ell_i \circ A$ and $\ell_j \circ A$ are not proportional to each other is a nonempty Zariski open set given by the condition $\operatorname{rank}(\ell_i \circ A, \ell_j \circ A) > 1$. Hence for a nonempty Zariski open (and therefore dense) set of linear maps A the linear forms $\ell_k \circ A$ are pairwise non-proportional. From the binary case above we have $g_k \circ A = 0$ if A lies in this open set. By continuity this implies $g_k \circ A = 0$ for all A. Since every point lies in the image of some linear map A we have $g_k = 0$. $\square$

8.7 Lemma. *Let $f \in S^d V$ be be such that $\underline{\mathsf{WR}}(f) = r$. If $d \geq r - 1$, then there exists a partition $r = r_1 + \cdots + r_m$ such that f has a generalized additive decomposition*

$$f = \sum_{k=1}^{m} \ell_k^{d-r_k+1} g_k,$$

and moreover $\underline{\mathsf{WR}}(\ell_k^{d-r_k+1} g_k) \leq r_k$.

Proof. Consider a border Waring rank decomposition

$$f = \lim_{\varepsilon \to 0} \sum_{k=1}^{r} \ell_k^d$$

Divide the summands between several local decompositions as follows. Define an equivalence relation $\sim$ on the set of indices $\{1, 2, \ldots, r\}$ as $i \sim j \Leftrightarrow \lim_{\varepsilon\to 0}[\ell_i] = \lim_{\varepsilon\to 0}[\ell_j]$ and let $I_1, \ldots, I_m$ be the equivalence classes with respect to this relation. Further, let $r_k = \sharp I_k$ and let $[L_k] = \lim_{\varepsilon\to 0}[\ell_i]$ for $i \in I_k$.

Consider the sum of all summands with indices in I_k. Let q_k be the power of ε in the lowest order term, that is,

$$\sum_{i\in I_k} \ell_i^d = \varepsilon^{q_k} f_k + \sum_{j=q_k+1}^{\infty} \varepsilon^j f_{k,j}$$

with $f_k \in \mathbb{C}[\mathbf{x}]_d$ nonzero. This expression can be transformed into a local border decomposition

$$f_k = \lim_{\varepsilon\to 0} \sum_{i\in I_k} \left(\frac{\ell_i(\varepsilon^d)}{\varepsilon^{q_k}}\right)^d.$$

based at $[L_k]$. By Lemma 8.4 we have $f_k = L_k^{d-r_k+1} g_k$ for some homogeneous polynomial g_k of degree $r_k - 1$. The decomposition also gives $\underline{\mathsf{WR}}(f_k) \leq r_k$.

Note that $q_k \leq 0$ since otherwise the summands ℓ_i with $i \in I_k$ can be removed from the original border rank decomposition of f without changing the limit. Let $q = \min\{q_1, \ldots, q_m\}$. Note that if $q < 0$, then, comparing the terms before ε^q in the left and right hand sides of the equality

$$f + O(\varepsilon) = \sum_{k=1}^{m} \sum_{i\in I_k} \ell_i^d$$

we get

$$0 = \sum_{k\colon a_k = a} f_k = \sum_{k\colon a_k=a} L_k^{d-r_k+1} g_k.$$

From Lemma 8.6 we obtain $g_k = 0$ and $f_k = 0$, in contradiction with the definition of f_k.

We conclude that $q = 0$ and

$$f = \sum_{k=1}^{m} f_k = \sum_{k=1}^{m} L_k^{d-r_k+1} g_k,$$

obtaining the required generalized additive decomposition. □

We will now take a brief detour to define a function $M(r)$ which we use to upper bound the Waring rank of generalized additive decomposition.

8.8 Definition. *We denote the* maximum Waring rank *of a degree d homogeneous polynomial in n variables by* $\mathsf{maxR}(n, d) = \max_{f\in\mathbb{C}[x_1,\ldots,x_n]_d} \mathsf{WR}(f)$. *Define the* partition-maxrank function *as*

$$M(r) = \max_{r_1+\cdots+r_m=r} \sum_{k=1}^{m} \mathsf{maxR}(r_k, r_k - 1).$$

8.9 Proposition. $\mathsf{maxR}(n, d_1) \leq \mathsf{maxR}(n, d_2)$ *when* $d_1 \leq d_2$.

Proof. Every form f of degree d_1 can be represented as a partial derivative of some form g of degree d_2. By differentiating a Waring rank decomposition of g we obtain a Waring rank decomposition of f, thus $\mathsf{WR}(f) \leq \mathsf{WR}(g) \leq \mathsf{maxR}(n, d_2)$. Since f is arbitrary, $\mathsf{maxR}(n, d_1) \leq \mathsf{maxR}(n, d_2)$. □

We are now ready to prove a de-bordering theorem for Waring rank.

8.10 Theorem. *Let $f \in S^d V$ be such that $\underline{\mathsf{WR}}(f) = r$. Then*

$$\mathsf{WR}(f) \leq M(r) \cdot d.$$

Proof. We consider two cases depending on relation of degree d and border Waring rank r.

Case $d < r - 1$. Since $\mathsf{WR}(f) = r$, the number of essential variables of f is at most r. Taking the maximum Waring rank as an upper bound, we obtain

$$\mathsf{WR}(f) \leq \mathsf{maxR}(d, r) \leq \mathsf{maxR}(r-1, r) \leq M(r) \leq M(r) \cdot d.$$

Case $r \leq d + 1$. By Lemma 8.7 f has a generalized additive decomposition

$$f = \sum_{k=1}^{m} \ell_k^{d-r_k+1} g_k$$

with $r_1 + \cdots + r_m = r$, $\deg g_k = r_k - 1$ and $\underline{\mathsf{WR}}(\ell_k^{d-r_k+1} g_k) \leq r_k$. Since $\underline{\mathsf{WR}}(\ell_k^{d-r_k+1} g_k) \leq r_k$, the number of essential variables $N_{ess}(g_k) \leq r_k$. If $r_k = 1$, then

$$\mathsf{WR}(\ell_k^{d-r_k+1} g_k) = \mathsf{WR}(\ell_k^d) = 1 \leq d.$$

If $r_k \geq 2$, then we upper bound $\mathsf{WR}(g_k)$ by $\mathsf{maxR}(N_{ess}(g_k), \deg g_k) = \mathsf{maxR}(r_k, r_k - 1)$. Taking a Waring rank decomposition $g_k = \sum_{i=1}^{\mathsf{WR}(g_k)} L_i^{r_k - 1}$ and multiplying it by $\ell_k^{d-r_k+1}$, we obtain a decomposition

$$\ell_k^{d-r_k+1} g_k = \sum_{i=1}^{\mathsf{WR}(g_k)} \ell_k^{d-r_k+1} \cdot L_i^{r_k-1}.$$

From the classical work of Sylvester (see also [BBT13]) it follows that[12]

$$\mathsf{WR}(\ell_k^{d-r_k+1} L_i^{r_k-1}) = \mathsf{WR}(y_1^{d-r_k+1} y_2^{r_k-1}) = \max\{d - r_k + 2, r_k\} \leq d.$$

Hence we have $\mathsf{WR}(\ell_k^{d-r_k+1} g_k) \leq d \cdot \mathsf{WR}(g_k) \leq d \cdot \mathsf{maxR}(r_k - 1, r_k)$.

Combining all parts of the decomposition together, we get

$$\mathsf{WR}(f) \leq d \sum_{k=1}^{m} \mathsf{maxR}(r - k - 1, r_k) \leq M(r) \cdot d. \qquad \square$$

A more explicit upper bound is provided by the following immediate corollary.

8.11 Theorem. *Let $f \in S^d \mathbb{C}^n$ and let $\underline{\mathsf{WR}}(f) = r$. Then*

$$\mathsf{WR}(f) \leq d \binom{2r-2}{r-1}.$$

Proof. The space of homogeneous polynomials of degree $r - 1$ in r variables has dimension $\binom{2r-2}{r-1}$ and is spanned by powers of linear forms. Therefore, $\mathsf{maxR}(r-1, r) \leq \binom{2r-2}{r-1}$. Note that if $r = p + q$ with $p, q \neq 0$, then the space $\mathbb{C}[x_1, \ldots, x_r]_{r-1}$ contains a direct sum of $x_1^q \cdot \mathbb{C}[x_1, \ldots, x_p]_{p-1}$ and $x_1^{p+1} \cdot \mathbb{C}[x_{p+1}, \ldots, x_r]_{q-1}$. Taking the dimensions of these spaces, we obtain $\binom{2r-2}{r-1} \geq \binom{2p-2}{p-1} + \binom{2q-2}{q-1}$. It follows that $M(r) \leq \binom{2r-2}{r-1}$. $\square$

[12] it is easy to see that for $a \geq b$ the monomial $y_1^a y_2^b$ is proportional to $\sum_{k=0}^{a} \zeta^k (\zeta^k y_1 + y_2)^{a+b}$ where ζ is a primitive root of unity of degree $a + 1$.

Using Blekherman-Teitler bound on the maximum rank [BT15], we can get a slightly better bound, but the techniques used we no longer consider elementary. The proof is essentially the same as for the previous theorem.

8.12 Corollary. *Let $f \in S^d\mathbb{C}^n$ and let $\underline{\mathsf{WR}}(f) = r$. Then*

$$\mathsf{WR}(f) \leq 2d\lceil \frac{1}{r}\binom{2r-2}{r-1} \rceil.$$

8.c Behind the scenes: generalized additive decompositions and schemes

We will now discuss how the results of the previous section can be obtained from apolarity theory and the study of 0-dimensional schemes in projective space. The connection between variations of Waring rank, apolar schemes and generalized additive decompositions is explored in detail in [BBM14]. In particular, there exists a much stronger version of Lemma 8.7, which tightly relates generalized additive decompositions of a homogeneous polynomial f to its *cactus rank* $\mathsf{CR}(f)$, a variation of Waring rank arising in apolarity theory defined in terms of 0-dimensional schemes in place of sets of linear forms. We will formally define cactus rank below, for now let us state the result.

8.13 Definition. *The* partial derivative space *of a polynomial $f \in \mathbb{C}[x_1, \dots, x_n]$ (not necessarily homogeneous) is the vector space $\partial^* f$ spanned by f and all its partial derivatives of all orders.*

8.14 Definition. *We define the* size *of a generalized additive decomposition*

$$f = \sum_{k=1}^{m} \ell_k^{d-r_k+1} g_k$$

as $\sum_{k=1}^{m} \dim \partial^* \overline{g}_k$ where $\overline{g}_k = g_k \bmod \langle \ell_k - 1 \rangle$ (note that $\overline{g}_k$ lies in $\mathbb{C}[\boldsymbol{x}]/\langle \ell_k - 1 \rangle$, which is isomorphic to a polynomial ring in $n-1$ variables).

8.15 Theorem ([BBM14])**.** *The cactus rank of a homogeneous polynomial f is equal to the minimal possible size of a generalized additive decomposition for f.*

To connect cactus rank to border rank we need and intermediate notion of *smoothable rank* $\mathsf{SR}(f)$. Smoothable rank is an upper bound on cactus rank, and it coincides with border rank for polynomials of high enough degree.

8.16 Theorem ([BB15])**.** *If $\deg f \geq \underline{\mathsf{WR}}(f) - 1$, then $\underline{\mathsf{WR}}(f) = \mathsf{SR}(f)$.*

The goal of this section is to review the basic notions of apolarity theory, define cactus rank and smoothable rank and explain the ideas behind the proofs of the theorem stated above.

Some notation. Let us fix the notation. Let $S = \mathbb{C}[x_1, \dots, x_n]$ be the algebra of polynomials and $T = \mathbb{C}[\partial_1, \dots, \partial_n]$ be the algebra of polynomial differential operators with constant coefficients (referred to as *diffoperators* in what follows), which acts on S in the standard way.

Denote by V the space of linear forms S_1. We identify T_1 with the dual space V^*. More generally, the action of T on S gives rise to a nondegenerate pairing between the homogeneous parts S_d and T_d for every d. We use orthogonality with respect to this pairing, that is, for a subset $F \subset S_d$ we denote $F^\perp = \{\alpha \in T_d \mid \alpha \cdot f = 0 \text{ for all } f \in F\}$, and vice versa, for a subset $D \subset T_d$ we let $D^\perp = \{f \in S_d \mid \alpha \cdot f = 0 \text{ for all } \alpha \in D\}$

Projective geometry. The algebra T is isomorphic to $\mathbb{C}[V]$, the algebra of polynomials in the coefficients of linear forms. The isomorphism maps a homogeneous element $\alpha \in T_d$ to $\overline{\alpha} \in \mathbb{C}[V]_d$ defined as $\overline{\alpha}(\ell) = \alpha \cdot \frac{\ell^d}{d!}$.

Recall that a homogeneous ideals in $T \cong \mathbb{C}[V]$ are in correspondence with subsets of the projective space $\mathbb{P}V$. More specifically, projective varieties are subsets of $\mathbb{P}V$ defined by vanishing of some set of polynomials. The set of all polynomials vanishing on a projective variety Z is a homogeneous ideal I, which is saturated ($\alpha T_1 \subset I \Rightarrow \alpha \in I$) and radical ($\alpha^n \in I \Rightarrow \alpha \in I$). If we consider ideals I which are saturated but not radical, we can define a projective scheme, which coincides with the variety defined by I as a topological space, but has additional structure which distinguishes it from this variety.

If $I \subset T$ is a homogeneous ideal, then the function $h_I(p) = \dim(T_p/I_p)$ is called the *Hilbert function* of I. The Hilbert function of a homogeneous ideal I always coincides with some polynomial $H_I(p)$ for p large enough. This polynomial is called the *Hilbert polynomial* of I.

Many topological and geometric properties of a projective variety or a scheme can be deduced from its Hilbert polynomial, in particular, its dimension and degree [Har77, §I.7]. We are specifically interested in ideals with constant Hilbert polynomials. These ideals corresponds to schemes of dimension 0. This means that a variety with Hilbert polynomial r is a set of r distinct points in $\mathbb{P}V$. In algebra, ideals with constant Hilbert polynomial are referred to as *ideals of Krull dimension 1* (the mismatch with the dimension of a scheme is because in algebra dimension is counted in affine space).

Apolarity theory. The connection between Waring rank and algebraic geometry is provided by the apolarity theory, which has its source in the works of Sylvester and Macaulay.

8.17 Definition. *The* apolar ideal *of a polynomial $f \in S$ is an ideal in T defined as* $\operatorname{Ann}(f) = \{\alpha \in T \mid \alpha \cdot f = 0\}$. *The* apolar algebra *of f is $A(f) = T/\operatorname{Ann}(f)$. An ideal $I \subset T$ is said to be* apolar *to f if it lies in* $\operatorname{Ann}(f)$. *A scheme $Z \subset \mathbb{P}V$ is* apolar *to f if its defining ideal is.*

Note that as a vector space, $A(f)$ is isomorphic to the space of partial derivatives $\partial^* f = T \cdot f$ via $(\alpha + \operatorname{Ann}(f)) \mapsto \alpha \cdot f$.

To relate apolarity to Waring rank, we also define an ideal associated with a set of linear forms. Given r linear forms $\ell_1, \dots, \ell_r$, consider the sequences of subspaces $E_p = \operatorname{Span}(\{\ell_1^p, \dots, \ell_r^p\}) \subset S_p$ and $I_p = E_p^\perp \subset T_p$. An important fact is that $I = \bigoplus_{p=0}^\infty I_p$ is a homogeneous ideal in T. From the geometric point of view it can be described as the vanishing ideal of the set $Z = \{[\ell_1], \dots, [\ell_r]\}$ in the projective space $\mathbb{P}V$. Algebraically, the fact that I is a homogeneous ideal follows from the following useful proposition.

8.18 Proposition. *A sequence of subspaces $E_p \subset S_p$ satisfies the property $T_1 \cdot E_{p+1} \subset E_p$ if and only if $I = \bigoplus_{p=0}^\infty E_p^\perp$ is a homogeneous ideal. If this is the case, then $h_I(p) = \dim E_p$.*

Proof. Let $I_p = E_p^\perp$. The fact that I is a homogeneous ideal can be written as $I_{p+1} \supset T_1 \cdot I_p$, which is equivalent to $T_1 \cdot E_{p+1} \subset E_p$, as both of these statements reduce to

$$(\alpha\partial) \cdot f = \alpha \cdot (\partial f) = 0 \text{ for all } \alpha \in I_p, \partial \in T_1, f \in E_{p+1}.$$

For the Hilbert function expression, note $\dim(T_p/I_p) = \dim T_p - \dim I_p = \dim I_p^\perp = \dim E_p$. □

8.19 Theorem (Apolarity lemma). *$f \in S_d$ is a linear combination of powers of linear forms $\ell_1, \dots, \ell_r$ if an only if f is apolar to $Z = \{[\ell_1], \dots, [\ell_r]\} \subset \mathbb{P}V$.*

Proof. Let I be the defining ideal of Z let $E_p = I_p^{\perp} = \operatorname{Span}(\{\ell_1^p, \dots, \ell_r^p\})$ as above.

If I is apolar to f, then $I_d \subset \operatorname{Ann}(f)_d$ and therefore $E_d \supset (\operatorname{Ann}(f)_d)^{\perp} = f^{\perp\perp} \ni f$.

For the other direction, let $f \in E_d$. Note that $\operatorname{Ann}(f)_p = T_p$ for $p > d$, so we only need to check $I_p \subset \operatorname{Ann}(f)$ for $p \leq d$.

Note that if for $\alpha \in T_p$ with $p < d$ we have $\alpha \cdot f \in S_{d-p}$ nonzero, then there exists $\partial \in T_1$ such that $\partial\alpha \cdot f = \partial \cdot (\alpha \cdot f) \neq 0$. This can be restated as $T_1\alpha \in Ann(f) \Rightarrow \alpha \in \operatorname{Ann}(f)$ for all $\alpha \in T_p$ with $p < d$

For $p \leq d$ we have $\alpha \in I_p \Rightarrow T_1^{d-p}\alpha \subset I_d = E_d^{\perp} \Rightarrow T_1^{d-p}\alpha \cdot f = 0 \Rightarrow \alpha \in \operatorname{Ann}(f)$, which proves $I_p \subset \operatorname{Ann}(f)$. ☐

8.20 Corollary. $\mathsf{WR}(f) \leq r$ *if and only if f is apolar to the vanishing ideal of r points in $\mathbb{P}V$.*

Families of subspaces and ideals and their limits. Before considering border Waring rank, we need to define limits of families of subspaces and families of ideals.

Let W be a vector space. We consider two types of families of subspaces in W. First is a family of subspaces of the form $E(\varepsilon) = \operatorname{Span}(\{w_1(\varepsilon), \dots, w_r(\varepsilon)\})$ where $w_k(\varepsilon)$ are families of vectors in W with coordinates given by rational functions of ε. We write $w_k \in W(\varepsilon)$ in this case. The second type is a family $E(\varepsilon) = \{w \mid y_1(\varepsilon; w) = \cdots = y_q(\varepsilon; w) = 0\}$ of vector spaces defined by linear forms $y_1, \dots, y_q \in W^*(\varepsilon)$ which again depend rationally on the parameter ε.

In both cases we define the limit $\widehat{E} = \lim_{\varepsilon \to 0} E(\varepsilon)$ as the subspace containing the limits of all families $w \in W(\varepsilon)$ such that $w(\varepsilon) \in E(\varepsilon)$ for $\varepsilon \neq 0$ (whenever $E(\varepsilon)$ and $w(\varepsilon)$ are defined).

For $E(\varepsilon) = \operatorname{Span}(\{w_1(\varepsilon), \dots, w_r(\varepsilon)\})$ from semicontinuity of rank we have that the maximal possible value of $\dim E(\varepsilon)$ is attained on an open set of values of ε. The situation is opposite for the family of the second type $E(\varepsilon) = \operatorname{Span}(\{y_1(\varepsilon), \dots, y_q(\varepsilon)\})^{\perp}$. In both cases the dimension of $\widehat{E}$ cannot be higher then the generic dimension. Indeed, if $\widehat{E}$ contains linearly indeoendent vectors $v_1, \dots, v_m$, then there are families $v_1(\varepsilon), \dots, v_m(\varepsilon)$ which have them as limits, and these families will be linearly independent for an open subset of values of ε. Considering two families $E(\varepsilon) \subset W$ and $E(\varepsilon)^{\perp} \subset W^*$ together, we see that $\dim \widehat{E}$ is actually equal to the generic dimension of $E(\varepsilon)$ (maximal dimension for the families of the first type, and minimal — for the families of the second type).

Alternatively, we may associate with a family of subspaces a family of points in the *Grassmannian* – the space of all k-dimensional subspaces in W. The Grassmannian can be defined as the projective variety in $\mathbb{P}\Lambda^k W$ consisting of all points of the form $[w_1 \wedge \cdots \wedge w_k]$, which represent k-dimensional subspaces spanned by $w_1, \dots, w_k$ respectively. If $E(\varepsilon)$ is a family with generic dimension k and $v_1(\varepsilon), \dots, v_k(\varepsilon) \in E(\varepsilon)$ are linearly independent for generic values of ε, then we can define a rational map $\varepsilon \mapsto [v_1(\varepsilon) \wedge \cdots \wedge v_k(\varepsilon)]$ and take the limit of this map in the Grassmannian.

Suppose $I(\varepsilon)$ is a family of homogeneous ideals in T, that is, $I(\varepsilon) = \bigoplus_{p=0}^{\infty} I_p(\varepsilon)$ for the families of subspaces $I_p(\varepsilon) \subset T_p$ such that $I_{p+1}(\varepsilon) \supset I_p(\varepsilon) \cdot T_1$. By continuity of multiplication for the limit subspaces $\widehat{I}_p = \lim_{\varepsilon \to 0} I_p(\varepsilon)$ we still have $\widehat{I}_{p+1} \supset \widehat{I}_p \cdot T_1$. Hence $\widehat{I}$ is again a homogeneous ideal in T. This notion of limit of ideals corresponds to taking limits in the *multigraded Hilbert scheme*, which is a space of ideals with given Hilbert function, see [HS04]. We refer to this limit as the *multigraded limit of a family of ideals*. The problem is that the limit in the multigraded Hilbert scheme can be non-saturated and thus not correspond to a geometric object in projective space.

For example, consider three families of points $(1 : 0 : 1), (-1 : 0 : 1), (0 : \varepsilon : 1)$ in $\mathbb{P}^2$. The family of vanishing ideals is $\langle x_1x_2, x_2(x_2 - \varepsilon x_3), \varepsilon(x_1^2 - x_3^2) + x_2x_3, x_1^3 - x_1x_3^2 \rangle$. Taking $\varepsilon \to 0$ we obtain the ideal $\langle x_1x_2, x_2^2, x_2x_3, x_1^3 - x_1x_3^2 \rangle$, which is not saturated, since it contains x_1x_2, x_2^2, x_2x_3

but not x_2. Taking the saturation, we obtain $\langle x_2, x_1^3 - x_1 x_3^2 \rangle$ which corresponds to three points $(1:0:1), (-1:0:1), (0:0:1)$ as expected.

We can take saturation after obtaining the limit ideal. This notion of limit corresponds to limits in the *Hilbert scheme*, which is the space of ideals with the fixed Hilbert polynomial. It was defined by Grothendieck [Gro61], see also [IK99, Appx.C].

Border apolarity We will now describe the basic idea of the apolarity theory for border Waring rank, which was developed by Buczyńska and Buczyński in [BB21].

Let $f = \lim_{\varepsilon \to 0} \sum_{k=1}^{r} \ell_k^d$ be a border Waring rank decomposition. Consider the families of subspaces $E_p(\varepsilon) = \mathrm{Span}(\{\ell_1(\varepsilon)^p, \dots, \ell_r(\varepsilon)^p\}) \subset S_p$ and the family of homogeneous ideals $I(\varepsilon) = \bigoplus_{p=0}^{\infty} E_p(\varepsilon)^{\perp}$ in T.

As $\varepsilon \to 0$, we obtain a sequence of subspaces $\widehat{E}_p = \lim_{\varepsilon \to 0} E_p(\varepsilon) \subset S_p$ and a homogeneous ideal $\widehat{I} = \lim_{\varepsilon \to 0} I(\varepsilon)$ (taking the limit in the multigraded Hilbert scheme). Let $\overline{f} = \sum_{k=1}^{r} \ell_k^d \in S_d(\varepsilon)$, so that $f = \lim_{\varepsilon \to 0} \overline{f}(\varepsilon)$. By the Apolarity Lemma the ideal $I(\varepsilon)$ is apolar to $\overline{f}(\varepsilon)$ for $\varepsilon \neq 0$, which means that $\alpha(\varepsilon) \cdot \overline{f}(\varepsilon) = 0$ for every $\alpha(\varepsilon) \in I(\varepsilon)$. Since the action of T on S is continuous, we obtain from this $(\lim_{\varepsilon \to 0} \alpha(\varepsilon)) \cdot f = 0$, if the limit exists. Thus $\widehat{I}$ is apolar to f.

On the other hand, suppose that $f \in S_d$ is apolar to an ideal $\widehat{I}$ which is a *limit of ideals of r points*, that is, there exists a family $I(\varepsilon)$ such that $I(\varepsilon)$ is the vanishing ideal of a set of r points in $\mathbb{P}V$. Define $E_d(\varepsilon) = I(\varepsilon)_d^{\perp} \subset S_d$. For $\varepsilon \neq 0$ the subspace $E_d(\varepsilon)$ is a span of powers of r linear forms, so it consists of polynomials with Waring rank at most r. Since f is orthogonal to $\widehat{I}_d$, it lies in the limit $\lim_{\varepsilon \to 0} E_d(\varepsilon)$ and thus has border Waring rank at most r.

8.21 Theorem (Border apolarity, [BB21])**.** *$f \in S_d$ has $\underline{\mathrm{WR}}(f) \leq r$ if and only if f is apolar to an ideal $\widehat{I}$ which is a limit of ideals of r points.*

Various ranks via apolarity. The apolarity lemma provides a template for defining different notions of rank for homogeneous polynomials by varying the class of ideals apolar to f.

8.22 Definition. *Let $\mathcal{C}$ be a class of ideals of Krull dimension 1. If $f \in S_d$ is a homogeneous polynomial, we define the $\mathcal{C}$-rank of f as the minimal r such that there exists an ideal $I \subset \mathcal{C}$ apolar to f with Hilbert polynomial $H_I = r$.*

As we have seen, Waring rank and border Waring rank are special cases of this definition corresponding to ideals of points and their limits.

We are now ready to define cactus rank and smoothable rank. The cactus rank $\mathrm{CR}(f)$ is obtained from the template definition above if we consider the class of all saturated ideals with constant Hilbert polynomial, that is, ideals of 0-dimensional schemes. The smoothable rank $\mathrm{SR}(f)$ corresponds to saturated limits of ideals of points. In addition, the border cactus rank $\underline{\mathrm{CR}}(f)$ is defined by considering limits of saturated ideals.

Class of ideals	Rank	Notation
Ideals of points (radical saturated ideals)	Waring rank	$\mathrm{WR}(f)$
Limits of ideals of points	Border Waring rank	$\underline{\mathrm{WR}}(f)$
Smoothable ideals (saturated limits of ideals of points)	Smoothable rank	$\mathrm{SR}(f)$
Saturated ideals	Cactus rank	$\mathrm{CR}(f)$
Saturable ideals (limits of saturated ideals)	Border cactus rank	$\underline{\mathrm{CR}}(f)$

The unified definition allows us to determine relations between these different ranks.

8.23 Theorem ([BBM14])**.** *The following inequalities hold:* $\underline{\mathsf{CR}}(f) \leq \mathsf{CR}(f) \leq \mathsf{SR}(f) \leq \mathsf{WR}(f)$ *and* $\underline{\mathsf{CR}}(f) \leq \underline{\mathsf{WR}}(f) \leq \mathsf{SR}(f) \leq \mathsf{WR}(f)$*.*

Proof. The inequality $\underline{\mathsf{WR}}(f) \leq \mathsf{SR}(f)$ follows from the fact that if the saturation $I^{\mathrm{sat}} \supset I$ is apolar to f, then I is also apolar to f. Other inequalities follow from the containments between corresponding classes of ideals. □

Proof idea of Theorem 8.15. If $\mathsf{CR}(f) \leq r$, then there exists a saturated homogeneous ideal I apolar to f with Hilbert polynomial r. This ideal corresponds to a 0-dimensional scheme Z, which consists of several points. Each point corresponds to a primary ideal in the primary decomposition $I = I^{(1)} \cap \cdots \cap I^{(m)}$. Defining $E_d = I_d^{\perp}$ and $E_d^{(k)} = (I_d^{(k)})^{\perp}$ we have $E_d = E_d^{(1)} + \cdots + E_d^{(m)}$. Each component will contribute one summand to the generalized additive decomposition. It remains to prove two facts. First, a primary ideal $I^{(k)}$ vanishing on $[\ell_k]$ will have polynomials of the form $\ell_k^{d-r_k+1} g_k$ in the corresponding $E^{(k)}$. This follows from the fact that $\sqrt{I^{(k)}}$ is $\langle \ell_k^{\perp} \rangle$, the vanishing ideal of $[\ell_k]$, and thus $\langle \ell_k^{\perp} \rangle^{r_k} \subset I^{(k)} \subset \langle \ell_k^{\perp} \rangle$ for some r_k. Second, given ℓ_k and g_k we can construct a primary ideal such that $E_d^{(k)}$ is isomorphic to $\partial^* \overline{g}_k$ where $\overline{g}_k = g_k \bmod \langle \ell_k - 1 \rangle$. This can be done, for example, by explicitly writing the corresponding subspaces $E_p^{(k)}$ which will contain homogenizations of all partial derivatives of $\overline{g}_k$ and proving that this sequence defines a primary ideal.

8.d Classes of the form ΣFΣ

More precisely, let $\mathbf{F} = \{F_m\}$ be a p-family and let $\Sigma\mathbf{F}\Sigma$ the class of sequences of polynomials $\{f_n\}$ such that $f_n = \sum_{i=1}^{r(n)} F_{m_i(n)}(\ell_{i1}, \ldots, \ell_{iN_{m(n)}})$ where ℓ_{ij} are linear forms in the variables of f_n and $r(n), m_i(n)$ are all polynomial functions of m; here N_m denotes the number of variables of F_m.

For instance, if $\mathbf{F} = \{x_0^m : m \in \mathbb{N}\}$, the class $\Sigma\mathbf{F}\Sigma$ coincides with the **VW**. If $\mathbf{F} = \{x_1 \cdots x_m : m \in \mathbb{N}\}$, then $\Sigma\mathbf{F}\Sigma$ is exactly $\Sigma\Pi\Sigma$. In general, it is clear that $\{f_n\}$ is a p-family.

We say that the p-family $\mathbf{F}$ has constant number of variables if the number of variables of F_m is bounded above by a constant (and in particular independently from m). In this case, we have the following immediate result.

8.24 Proposition. *Let* $\mathbf{F}$ *be a p-family in constant number of variables. Then* $\Sigma\mathbf{F}\Sigma = \boldsymbol{VW}$*.*

Proof. Clearly $\mathbf{VW} \subseteq \Sigma\mathbf{F}\Sigma$ because every polynomial restricts to powers of linear forms.

Therefore it suffices to show that if $\{f_n\}$ is a sequence of polynomials in $\Sigma\mathbf{F}\Sigma$ then $\mathsf{WR}(f_n)$ is bounded by a polynomial in n. Let N be an upper bound to the number of variables of F_m, for every m. By definition of $\Sigma\mathbf{F}\Sigma$, we have

$$f_n = F_{m_1}(\ell_{11}, \ldots, \ell_{1N}) + \cdots + F_{m_r}(\ell_{r1}, \ldots, \ell_{rN})$$

where $r = r(n)$ is a function bounded by a polynomial in n.

Since F_m is a polynomial in at most N variables, $\mathsf{WR}(F_m) \leq O(\deg(F_m)^N)$, which is a polynomial function of m. Since $r(n)$ is polynomially bounded, we conclude $\mathsf{WR}(f_n) \leq O(\deg(F_{m_1})^N) + \cdots + O(\deg(F_{m_r})^N) \leq r(n)R(m)$ for some polynomial function of m; since $m_1, \ldots, m_r$ are polynomial functions in n, as well as $r(n)$, we conclude. □

9 Remarks on related work

If we look at our definition of Kc above, it just computes a sum of elementary symmetric polynomials when the inputs are linear forms. Shpilka [Shp02] studied a similar notion of circuit complexity called s_{sym}. For a polynomial f, $s_{sym}(f)$ is defined as the smallest m such that $f = e_d(\ell_1, \ell_2, \dots, \ell_m)$ where $d = \deg(f)$ and ℓ_i are affine linear forms. Here e_d is the d^{th} elementary symmetric polynomial. It was proved in [Shp02] that $s_{sym}(f)$ is always finite, moreover several upper and lower bounds for $s_{sym}(f)$ were proven. The complexity Kc differs from $s_{sym}(f)$, as Kc can even be infinite. We also study the non-commutative generalisation of the Kc complexity, where variables are 2×2 or 3×3 matrices. In the case when the variables are 2×2 matrices, we study the special case of $\begin{pmatrix} 0 & \ell \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ \ell & 0 \end{pmatrix}$, where ℓ is a linear form. This gives rise to a product of 2×2 matrices, hence to a width two algebraic branching program. Similar models are also studied in [MS21, BIZ18]. In [BIZ18], it is shown that VF is contained in the closure of class of width two ABPs, in fact in the orbit closure of the so called continuant polynomial. [MS21] constructed polynomial sized hitting sets for affine orbits of the cyclic continuant polynomial. Out study of parity-alternating polynomials in Section 7.d can be seen as a homogeneous variant of the continuant polynomial defined in [BIZ18].

We also study the relation between Kc and Waring rank. Moreover, we study the de-bordering of various orbit closures, of product plus power, product pus two powers and that of power symmetric polynomial. The orbit closure of the power symmetric polynomials, essentially characterises the border Waring rank. In de-bordering the border Waring rank, one wants to establish that the Waring rank cannot be *too large* compared to the border Waring rank. [LT10] characterised the polynomials the polynomials which have border Waring rank at most four. The characterisation of [LT10] implies that a for homogeneous polynomial f, $\underline{\mathsf{WR}}(f) \leq 4$ implies $\mathsf{WR}(f) \in O(\deg(f))$. Ballico [Bal19] proved that if $\underline{\mathsf{WR}}(f) \leq 5$ then $\mathsf{WR}(f) \leq 4\deg(f) - 2$. One can hope to prove that if $\underline{\mathsf{WR}}(f)$ is constant then $\mathsf{WR}(f) \in O(\deg(f))$, which we prove in this paper. A similar upper bound for the Waring rank which depends linearly upon the so called Curvilinear rank was established in [BB17].

The main references on the Alon-Tarsi conjecture are [AT92, Dri97, Gly10]. [FM19] give a survey about these main results. The conjecture has been generalized in numerous directions. [SW12] prove that Drisko's proof method cannot be used without modifications to prove the Alon-Tarsi conjecture. The GCT result in [Kum15] is based on the Alon-Tarsi conjecture. The same is true for results in [BI13, BI17], some of which are based on generalizations or variants of the conjecture. The Polymath Project number 12 (`https://polymathprojects.org`) was devoted to the study of Rota's basis conjecture, which for even n is implied by the Alon-Tarsi conjecture, see [HR94]. [Alp17] proves an upper bound on the different between the even and odd Latin squares. Fundamental invariants connected to the Alon-Tarsi conjecture have recently been studied in [LZX21, AY22].

A Calculation tables

We list the partitions λ for which the plethysm coefficient $a := a_\lambda(\delta, d)$ exceeds the multiplicity $b := \mathrm{mult}_\lambda(\mathbb{C}[\mathrm{GL}_{d+1}(x_1 \cdots x_d + x_{d+1}^d)])$. We write $\lambda_{a>b}$. We list λ always with all $d+1$ parts, i.e., with all trailing zeros. λ always has $d\delta$ many boxes. If we list a case (d, δ) and not list (d, δ') with $\delta' < \delta$, then this means that (d, δ') is empty.

$d = 3$, $\delta = 8$:

$(8,8,4,4)_{2>1}$, $(10,6,4,4)_{4>3}$

$d = 4$, $\delta = 6$:

$(6,6,4,4,4)_{1>0}$, $(7,7,5,5,0)_{1>0}$, $(7,7,7,3,0)_{1>0}$, $(8,5,5,3,3)_{1>0}$

$d = 4$, $\delta = 7$:

$(7,7,5,5,4)_{1>0}$, $(7,7,6,5,3)_{1>0}$, $(7,7,7,4,3)_{1>0}$, $(7,7,7,5,2)_{1>0}$, $(7,7,7,7,0)_{1>0}$, $(8,6,6,4,4)_{4>1}$, $(8,7,5,4,4)_{1>0}$, $(8,7,5,5,3)_{2>0}$, $(8,7,6,4,3)_{4>2}$, $(8,7,6,5,2)_{4>1}$, $(8,7,7,3,3)_{3>0}$, $(8,7,7,4,2)_{1>0}$, $(8,7,7,5,1)_{3>0}$, $(8,8,4,4,4)_{4>2}$, $(8,8,5,4,3)_{4>1}$, $(8,8,6,4,2)_{9>4}$, $(8,8,7,3,2)_{3>1}$, $(8,8,7,4,1)_{4>3}$, $(8,8,8,2,2)_{3>2}$, $(9,6,5,4,4)_{3>0}$, $(9,6,5,5,3)_{1>0}$, $(9,6,6,4,3)_{5>3}$, $(9,6,6,5,2)_{4>3}$, $(9,7,4,4,4)_{2>1}$, $(9,7,5,4,3)_{7>2}$, $(9,7,5,5,2)_{5>1}$, $(9,7,6,3,3)_{5>3}$, $(9,7,6,4,2)_{10>5}$, $(9,7,6,5,1)_{6>4}$, $(9,7,7,3,2)_{5>1}$, $(9,7,7,4,1)_{5>2}$, $(9,7,7,5,0)_{2>1}$, $(9,8,4,4,3)_{5>2}$, $(9,8,5,3,3)_{4>1}$, $(9,8,5,4,2)_{11>5}$, $(9,8,5,5,1)_{4>3}$, $(9,8,6,3,2)_{11>6}$, $(9,8,6,4,1)_{12>11}$, $(9,8,7,2,2)_{5>3}$, $(9,8,7,3,1)_{8>6}$, $(9,9,4,3,3)_{3>1}$, $(9,9,4,4,2)_{2>1}$, $(9,9,5,3,2)_{7>5}$, $(9,9,5,4,1)_{6>4}$, $(10,5,5,5,3)_{1>0}$, $(10,6,4,4,4)_{7>2}$, $(10,6,5,4,3)_{6>2}$, $(10,6,5,5,2)_{2>0}$, $(10,6,6,4,2)_{13>8}$, $(10,7,4,4,3)_{8>4}$, $(10,7,5,3,3)_{7>3}$, $(10,7,5,4,2)_{14>6}$, $(10,7,5,5,1)_{6>2}$, $(10,7,6,3,2)_{14>8}$, $(10,7,6,4,1)_{15>13}$, $(10,7,7,2,2)_{1>0}$, $(10,7,7,3,1)_{10>5}$, $(10,8,4,3,3)_{2>1}$, $(10,8,4,4,2)_{17>9}$, $(10,8,5,3,2)_{15>8}$, $(10,8,5,4,1)_{17>14}$, $(10,8,6,2,2)_{17>10}$, $(10,9,4,3,2)_{10>7}$, $(10,9,4,4,1)_{10>9}$, $(10,9,5,2,2)_{10>6}$, $(10,10,4,2,2)_{9>5}$, $(11,5,4,4,4)_{2>1}$, $(11,5,5,4,3)_{3>0}$, $(11,6,4,4,3)_{8>4}$, $(11,6,5,3,3)_{3>2}$, $(11,6,5,4,2)_{13>6}$, $(11,6,5,5,1)_{3>2}$, $(11,6,6,3,2)_{10>9}$, $(11,7,4,3,3)_{6>3}$, $(11,7,4,4,2)_{14>9}$, $(11,7,5,3,2)_{18>9}$, $(11,7,5,4,1)_{18>15}$, $(11,7,6,2,2)_{12>7}$, $(11,7,7,2,1)_{8>7}$, $(11,8,4,3,2)_{17>10}$, $(11,8,5,2,2)_{17>12}$, $(11,9,3,3,2)_{5>3}$, $(11,9,4,2,2)_{12>9}$, $(11,10,3,2,2)_{6>4}$, $(12,4,4,4,4)_{4>3}$, $(12,5,4,4,3)_{4>2}$, $(12,5,5,3,3)_{3>0}$, $(12,5,5,4,2)_{3>1}$, $(12,5,5,5,1)_{1>0}$, $(12,6,4,4,2)_{17>11}$, $(12,6,5,3,2)_{12>8}$, $(12,6,5,4,1)_{13>12}$, $(12,6,6,2,2)_{13>10}$, $(12,7,3,3,3)_{1>0}$, $(12,7,4,3,2)_{17>11}$, $(12,7,5,2,2)_{14>10}$, $(12,8,3,3,2)_{4>3}$, $(12,8,4,2,2)_{23>18}$, $(12,9,3,2,2)_{9>8}$, $(13,5,4,3,3)_{2>0}$, $(13,5,4,4,2)_{8>6}$, $(13,5,5,3,2)_{4>2}$, $(13,5,5,4,1)_{4>3}$, $(13,6,4,3,2)_{13>11}$, $(13,6,5,2,2)_{13>11}$, $(13,7,3,3,2)_{5>3}$, $(13,7,4,2,2)_{16>14}$, $(13,8,3,2,2)_{12>11}$, $(14,5,4,3,2)_{7>5}$, $(15,5,3,3,2)_{1>0}$

$d = 4$, $\delta = 8$:

$(7,7,7,7,4)_{1>0}$, $(8,6,6,6,6)_{2>1}$, $(8,7,6,6,5)_{1>0}$, $(8,7,7,5,5)_{3>0}$, $(8,7,7,6,4)_{1>0}$, $(8,7,7,7,3)_{2>0}$, $(8,8,6,6,4)_{7>1}$, $(8,8,7,5,4)_{3>0}$, $(8,8,7,6,3)_{5>0}$, $(8,8,8,4,4)_{8>2}$, $(8,8,8,5,3)_{2>1}$, $(8,8,8,6,2)_{7>2}$, $(9,6,6,6,5)_{2>1}$, $(9,7,6,5,5)_{3>0}$, $(9,7,6,6,4)_{5>1}$, $(9,7,7,5,4)_{7>0}$, $(9,7,7,6,3)_{6>0}$, $(9,7,7,7,2)_{3>0}$, $(9,8,5,5,5)_{1>0}$, $(9,8,6,5,4)_{14>2}$, $(9,8,6,6,3)_{12>3}$, $(9,8,7,4,4)_{10>1}$, $(9,8,7,5,3)_{18>2}$, $(9,8,7,6,2)_{13>2}$, $(9,8,7,7,1)_{3>0}$, $(9,8,8,4,3)_{11>2}$, $(9,8,8,5,2)_{12>4}$, $(9,8,8,6,1)_{7>4}$, $(9,9,5,5,4)_{6>0}$, $(9,9,6,4,4)_{5>0}$, $(9,9,6,5,3)_{15>3}$, $(9,9,6,6,2)_{5>2}$, $(9,9,7,4,3)_{14>1}$, $(9,9,7,5,2)_{17>3}$, $(9,9,7,6,1)_{7>2}$, $(9,9,7,7,0)_{2>0}$, $(9,9,8,3,3)_{8>1}$, $(9,9,8,4,2)_{8>1}$, $(9,9,8,5,1)_{9>3}$, $(9,9,9,3,2)_{3>1}$, $(9,9,9,4,1)_{3>0}$, $(10,6,6,6,4)_{9>3}$, $(10,7,5,5,5)_{3>0}$, $(10,7,6,5,4)_{15>1}$, $(10,7,6,6,3)_{13>3}$, $(10,7,7,4,4)_{5>0}$, $(10,7,7,5,3)_{19>1}$, $(10,7,7,6,2)_{8>0}$, $(10,7,7,7,1)_{4>0}$, $(10,8,5,5,4)_{7>0}$, $(10,8,6,4,4)_{31>4}$, $(10,8,6,5,3)_{32>5}$, $(10,8,6,6,2)_{29>8}$, $(10,8,7,4,3)_{35>5}$, $(10,8,7,5,2)_{34>6}$, $(10,8,7,6,1)_{18>6}$, $(10,8,8,3,3)_{4>1}$, $(10,8,8,4,2)_{33>9}$, $(10,8,8,5,1)_{15>9}$, $(10,9,5,4,4)_{15>1}$, $(10,9,5,5,3)_{16>1}$, $(10,9,6,4,3)_{39>6}$, $(10,9,6,5,2)_{38>8}$, $(10,9,6,6,1)_{16>9}$, $(10,9,7,3,3)_{21>5}$, $(10,9,7,4,2)_{43>8}$, $(10,9,7,5,1)_{28>9}$, $(10,9,8,3,2)_{24>7}$, $(10,9,8,4,1)_{24>10}$, $(10,9,9,2,2)_{2>0}$, $(10,9,9,3,1)_{8>3}$, $(10,10,4,4,4)_{12>2}$, $(10,10,5,4,3)_{18>3}$, $(10,10,5,5,2)_{7>0}$, $(10,10,6,3,3)_{8>2}$, $(10,10,6,4,2)_{42>10}$, $(10,10,6,5,1)_{18>7}$, $(10,10,6,6,0)_{11>10}$, $(10,10,7,3,2)_{23>6}$, $(10,10,7,4,1)_{26>12}$, $(10,10,8,2,2)_{17>5}$, $(10,10,8,3,1)_{13>9}$,

$(10,10,9,2,1)_{6>4}$, $(11,6,6,5,4)_{9>1}$, $(11,6,6,6,3)_{10>4}$, $(11,7,5,5,4)_{11>0}$, $(11,7,6,4,4)_{22>3}$, $(11,7,6,5,3)_{31>4}$, $(11,7,6,6,2)_{19>6}$, $(11,7,7,4,3)_{25>3}$, $(11,7,7,5,2)_{25>2}$, $(11,7,7,6,1)_{11>2}$, $(11,7,7,7,0)_{2>0}$, $(11,8,5,4,4)_{26>3}$, $(11,8,5,5,3)_{23>2}$, $(11,8,6,4,3)_{60>11}$, $(11,8,6,5,2)_{58>13}$, $(11,8,6,6,1)_{24>13}$, $(11,8,7,3,3)_{26>4}$, $(11,8,7,4,2)_{64>14}$, $(11,8,7,5,1)_{40>15}$, $(11,8,8,3,2)_{28>9}$, $(11,8,8,4,1)_{30>17}$, $(11,9,4,4,4)_{11>1}$, $(11,9,5,4,3)_{45>6}$, $(11,9,5,5,2)_{33>5}$, $(11,9,6,3,3)_{36>8}$, $(11,9,6,4,2)_{78>19}$, $(11,9,6,5,1)_{46>20}$, $(11,9,7,3,2)_{57>14}$, $(11,9,7,4,1)_{58>24}$, $(11,9,8,2,2)_{20>7}$, $(11,9,8,3,1)_{37>21}$, $(11,9,9,2,1)_{9>5}$, $(11,10,4,4,3)_{21>5}$, $(11,10,5,3,3)_{20>4}$, $(11,10,5,4,2)_{52>12}$, $(11,10,5,5,1)_{20>7}$, $(11,10,6,3,2)_{56>16}$, $(11,10,6,4,1)_{56>29}$, $(11,10,7,2,2)_{30>9}$, $(11,10,7,3,1)_{46>26}$, $(11,10,8,2,1)_{25>20}$, $(11,11,4,3,3)_{10>2}$, $(11,11,4,4,2)_{10>3}$, $(11,11,5,3,2)_{26>7}$, $(11,11,5,4,1)_{23>12}$, $(11,11,6,2,2)_{13>5}$, $(11,11,6,3,1)_{30>18}$, $(11,11,7,2,1)_{19>15}$, $(12,6,5,5,4)_{4>0}$, $(12,6,6,4,4)_{21>3}$, $(12,6,6,5,3)_{14>3}$, $(12,6,6,6,2)_{17>8}$, $(12,7,5,4,4)_{19>1}$, $(12,7,5,5,3)_{22>1}$, $(12,7,6,4,3)_{49>10}$, $(12,7,6,5,2)_{46>9}$, $(12,7,6,6,1)_{17>10}$, $(12,7,7,3,3)_{23>3}$, $(12,7,7,4,2)_{32>5}$, $(12,7,7,5,1)_{26>6}$, $(12,8,4,4,4)_{25>5}$, $(12,8,5,4,3)_{56>8}$, $(12,8,5,5,2)_{32>5}$, $(12,8,6,3,3)_{32>7}$, $(12,8,6,4,2)_{109>29}$, $(12,8,6,5,1)_{54>27}$, $(12,8,7,3,2)_{62>17}$, $(12,8,7,4,1)_{65>31}$, $(12,8,8,2,2)_{30>13}$, $(12,8,8,3,1)_{27>20}$, $(12,9,4,4,3)_{33>6}$, $(12,9,5,3,3)_{35>7}$, $(12,9,5,4,2)_{80>18}$, $(12,9,5,5,1)_{32>11}$, $(12,9,6,3,2)_{88>28}$, $(12,9,6,4,1)_{88>45}$, $(12,9,7,2,2)_{41>14}$, $(12,9,7,3,1)_{71>40}$, $(12,9,8,2,1)_{34>28}$, $(12,10,4,3,3)_{14>4}$, $(12,10,4,4,2)_{52>16}$, $(12,10,5,3,2)_{63>18}$, $(12,10,5,4,1)_{62>32}$, $(12,10,6,2,2)_{60>23}$, $(12,10,6,3,1)_{71>48}$, $(12,10,7,2,1)_{50>41}$, $(12,11,3,3,3)_{2>0}$, $(12,11,4,3,2)_{32>11}$, $(12,11,4,4,1)_{25>16}$, $(12,11,5,2,2)_{32>14}$, $(12,11,5,3,1)_{46>31}$, $(12,11,6,2,1)_{41>38}$, $(12,12,3,3,2)_{3>2}$, $(12,12,4,2,2)_{19>10}$, $(12,12,4,3,1)_{13>11}$, $(13,5,5,5,4)_{1>0}$, $(13,6,5,4,4)_{15>1}$, $(13,6,5,5,3)_{9>0}$, $(13,6,6,4,3)_{26>7}$, $(13,6,6,5,2)_{24>8}$, $(13,7,4,4,4)_{17>4}$, $(13,7,5,4,3)_{45>7}$, $(13,7,5,5,2)_{28>3}$, $(13,7,6,3,3)_{30>8}$, $(13,7,6,4,2)_{73>21}$, $(13,7,6,5,1)_{39>18}$, $(13,7,7,3,2)_{34>7}$, $(13,7,7,4,1)_{36>15}$, $(13,7,7,5,0)_{12>11}$, $(13,8,4,4,3)_{38>9}$, $(13,8,5,3,3)_{33>6}$, $(13,8,5,4,2)_{88>23}$, $(13,8,5,5,1)_{32>13}$, $(13,8,6,3,2)_{91>31}$, $(13,8,6,4,1)_{91>55}$, $(13,8,7,2,2)_{43>17}$, $(13,8,7,3,1)_{65>41}$, $(13,9,4,3,3)_{25>6}$, $(13,9,4,4,2)_{55>18}$, $(13,9,5,3,2)_{85>28}$, $(13,9,5,4,1)_{78>41}$, $(13,9,6,2,2)_{62>26}$, $(13,9,6,3,1)_{94>67}$, $(13,9,7,2,1)_{59>50}$, $(13,10,3,3,3)_{4>1}$, $(13,10,4,3,2)_{55>21}$, $(13,10,4,4,1)_{46>33}$, $(13,10,5,2,2)_{57>24}$, $(13,10,5,3,1)_{75>54}$, $(13,10,6,2,1)_{69>68}$, $(13,11,3,3,2)_{15>6}$, $(13,11,4,2,2)_{32>17}$, $(13,11,4,3,1)_{44>37}$, $(13,12,3,2,2)_{13>8}$, $(13,13,2,2,2)_{1>0}$, $(14,5,5,4,4)_{2>0}$, $(14,5,5,5,3)_{3>0}$, $(14,6,4,4,4)_{18>4}$, $(14,6,5,4,3)_{26>4}$, $(14,6,5,5,2)_{11>1}$, $(14,6,6,3,3)_{8>4}$, $(14,6,6,4,2)_{45>17}$, $(14,6,6,5,1)_{17>13}$, $(14,7,4,4,3)_{31>9}$, $(14,7,5,3,3)_{29>6}$, $(14,7,5,4,2)_{63>17}$, $(14,7,5,5,1)_{24>8}$, $(14,7,6,3,2)_{63>23}$, $(14,7,6,4,1)_{62>40}$, $(14,7,7,2,2)_{14>4}$, $(14,7,7,3,1)_{38>21}$, $(14,8,4,3,3)_{18>4}$, $(14,8,4,4,2)_{66>24}$, $(14,8,5,3,2)_{78>27}$, $(14,8,5,4,1)_{76>47}$, $(14,8,6,2,2)_{70>33}$, $(14,8,6,3,1)_{83>68}$, $(14,9,3,3,3)_{5>2}$, $(14,9,4,3,2)_{63>26}$, $(14,9,4,4,1)_{52>38}$, $(14,9,5,2,2)_{61>29}$, $(14,9,5,3,1)_{85>65}$, $(14,10,3,3,2)_{15>6}$, $(14,10,4,2,2)_{57>30}$, $(14,10,4,3,1)_{56>53}$, $(14,11,3,2,2)_{22>14}$, $(14,12,2,2,2)_{11>9}$, $(15,5,4,4,4)_{6>2}$, $(15,5,5,4,3)_{8>0}$, $(15,5,5,5,2)_{2>0}$, $(15,6,4,4,3)_{22>7}$, $(15,6,5,3,3)_{12>3}$, $(15,6,5,4,2)_{38>12}$, $(15,6,5,5,1)_{10>4}$, $(15,6,6,3,2)_{31>17}$, $(15,6,6,4,1)_{30>26}$, $(15,7,4,3,3)_{18>5}$, $(15,7,4,4,2)_{45>20}$, $(15,7,5,3,2)_{57>21}$, $(15,7,5,4,1)_{54>35}$, $(15,7,6,2,2)_{40>20}$, $(15,7,6,3,1)_{57>49}$, $(15,7,7,2,1)_{25>23}$, $(15,8,3,3,3)_{2>0}$, $(15,8,4,3,2)_{58>26}$, $(15,8,4,4,1)_{49>42}$, $(15,8,5,2,2)_{59>33}$, $(15,8,5,3,1)_{74>64}$, $(15,9,3,3,2)_{19>9}$, $(15,9,4,2,2)_{51>32}$, $(15,10,3,2,2)_{28>19}$, $(16,4,4,4,4)_{7>4}$, $(16,5,4,4,3)_{10>3}$, $(16,5,5,3,3)_{6>0}$, $(16,5,5,4,2)_{8>2}$, $(16,5,5,5,1)_{2>0}$, $(16,6,4,3,3)_{7>3}$, $(16,6,4,4,2)_{36>18}$, $(16,6,5,3,2)_{30>14}$, $(16,6,5,4,1)_{29>22}$, $(16,6,6,2,2)_{27>18}$, $(16,7,3,3,3)_{3>0}$, $(16,7,4,3,2)_{42>21}$, $(16,7,5,2,2)_{36>22}$, $(16,7,5,3,1)_{54>50}$, $(16,8,3,3,2)_{13>7}$, $(16,8,4,2,2)_{53>37}$, $(16,9,3,2,2)_{26>20}$, $(17,4,4,4,3)_{5>4}$, $(17,5,4,3,3)_{4>0}$, $(17,5,4,4,2)_{15>9}$, $(17,5,5,3,2)_{8>3}$, $(17,5,5,4,1)_{8>5}$, $(17,6,4,3,2)_{26>17}$, $(17,6,5,2,2)_{24>19}$, $(17,7,3,3,2)_{10>5}$, $(17,7,4,2,2)_{33>27}$, $(17,8,3,2,2)_{24>22}$, $(18,4,4,4,2)_{9>8}$, $(18,5,4,3,2)_{11>7}$, $(18,6,3,3,2)_{4>3}$, $(19,5,3,3,2)_{1>0}$

$d = 5$, $\delta = 7$, $\lambda_1 \le 8$:

$(8,7,7,5,5,3)_{1>0}$, $(8,7,7,6,4,3)_{1>0}$, $(8,7,7,6,5,2)_{1>0}$, $(8,7,7,7,3,3)_{1>0}$, $(8,8,7,5,4,3)_{2>1}$, $(8,8,7,6,3,3)_{1>0}$, $(8,8,7,6,4,2)_{3>2}$, $(8,8,7,6,5,1)_{2>1}$, $(8,8,7,7,4,1)_{1>0}$

B Characterizing small border Waring rank

The results on generalized additive decompositions from §8.b can be used to describe the polynomials of border rank 2 and 3, reproving the results of Landsberg and Teitler [LT10, Sec. 10].

B.1 Theorem. *A polynomial f with $\underline{\mathrm{WR}}(f) = 2$ must have the form $\ell_1^d + \ell_2^d$ or $\ell_1^{d-1}\ell_2$ where ℓ_1 and ℓ_2 are linear forms.*

In the first case, every border rank decomposition for f has the form

$$f = (\ell_1 + \varepsilon\widehat{\ell}_1)^d + (\ell_2 + \varepsilon\widehat{\ell}_2)^d$$

for some $\widehat{\ell}_1, \widehat{\ell}_2 \in \mathbb{C}[[\varepsilon]][\boldsymbol{x}]_1$.

In the second case, every border rank decomposition for f has the form

$$f = \frac{1}{\varepsilon^M}\left(a\ell_1 + \varepsilon\widehat{\ell}_1 + \varepsilon^M(\frac{1}{a^{d-1}d}\ell_2 + \ell_3)\right)^d - \frac{1}{\varepsilon^M}\left(a\ell_1 + \varepsilon\widehat{\ell}_1 + \varepsilon^M(\ell_3 + \varepsilon\widehat{\ell}_2)\right)^d$$

for some $a \in \mathbb{C}$, $\ell_3 \in \mathbb{C}[\boldsymbol{x}]_1$ and $\widehat{\ell}_1, \widehat{\ell}_2 \in \mathbb{C}[[\varepsilon]][\boldsymbol{x}]_1$.

Proof. By Lemma 8.7 f has a generalized additive decomposition

$$f = \sum_{i=1}^{m} \ell_i^{d-r_i+1} g_i$$

with $\sum_{i=1}^m r_i = \underline{\mathrm{WR}}(f) = 2$, $\deg g_i = r_i - 1$ There are only two possible partitions $\sum r_i = 2$. In the case $m = 2, r_1 = r_2 = 1$ the generalized additive decomposition is actually a Waring rank decomposition $f = \ell_1^d + \ell_2^d$. In the case $m = 1, r_1 = 2$ the polynomial g_1 is a linear form, renaming it we have $f = \ell_1^{d-1}\ell_2$.

From the proof of Lemma 8.7 it is clear that in the first case the decomposition must be a sum of two local decompositions of rank 1, and a local decomposition of rank 1 is just a power of $\ell + \varepsilon\widehat{\ell}$ for some $\widehat{\ell} \in \mathbb{C}[\varepsilon][\boldsymbol{x}]_1$.

In the second case the decomposition must be local, which means that both summands in the decomposition have the form $\varepsilon^{-M}(a\ell_1 + \varepsilon\widehat{\ell})$. To obtain $\ell_1^{d-1}\ell_2$ in the limit, the first M terms in each summand must cancel, and the terms in ε^M must differ by $\frac{1}{a^{d-1}d}\ell_2$. $\square$

B.2 Theorem. *A polynomial with $\underline{\mathrm{WR}}(f) = 3$ must have one of the three normal forms: $\ell_1^d + \ell_2^d + \ell_3^d$ or $\ell_1^d + \ell_2^{d-1}\ell_3^d$ or $\ell_1^{d-1}\ell_2 + \ell_1^{d-2}\ell_3^2$.*

Proof. By Lemma 8.7 f has a generalized additive decomposition

$$f = \sum_{i=1}^{m} \ell_i^{d-r_i+1} g_i$$

with $\sum_{i=1}^m r_i = \underline{\mathrm{WR}}(f) = 3$, $\deg g_i = r_i - 1$, and $\underline{\mathrm{WR}}(\ell_i^{d-r_i+1} g_i) \le r_i$.

In the case $m = 3$, $r_1 = r_2 = r_3 = 1$ this is a Waring rank decomposition $f = \ell_1^d + \ell_2^d + \ell_3^d$.

In the case $m = 2$, we can assume $r_1 = 1$, $r_2 = 2$. The generalized additive decomposition becomes $\ell_1^d + \ell_2^{d-1}\ell_3$, where $\ell_3 = g_2$ is a linear form.

In the case $m = 1$, $r_1 = 3$ we have $f = \ell_1^{d-2}g_1$ where g_1 is a quadratic form, and $\ell_1^{d-2}g_1$ has at most three-dimensional space of essential variables. In this case g_1 can always be presented as $\ell_1\ell_2 + \ell_3^2$ or $a\ell_1^2 + \ell_2\ell_3$ for some linear forms ℓ_2, ℓ_3. In the second case the border rank of $\ell_1^{d-2}g_1$ is at least 4 if $d > 2$, so it cannot appear. If $d = 2$ then both forms have rank 3 and are covered by the case $\ell_1^d + \ell_2^d + \ell_3^d$. □

References

[Alp17] Levent Alpoge, *Square-root cancellation for the signs of Latin squares*, Combinatorica **37** (2017), no. 2, 137–142. 60

[AT92] Noga Alon and Michael Tarsi, *Colorings and orientations of graphs*, Combinatorica **12** (1992), no. 2, 125–134. 60

[AY22] Alimzhan Amanov and Damir Yeliussizov, *Fundamental invariants of tensors, Latin hypercubes, and rectangular kronecker coefficients*, arXiv (2022). 60

[Bal19] Edoardo Ballico, *On the ranks of homogeneous polynomials of degree at least 9 and border rank 5*, Note di Matematica **38** (2019), no. 2, 55–92. 60

[BB15] W. Buczyńska and J. Buczyński, *On differences between the border rank and the smoothable rank of a polynomial*, Glasgow Math. J. **57** (2015), no. 2, 401–413. 12, 55

[BB17] Edoardo Ballico and Alessandra Bernardi, *Curvilinear schemes and maximum rank of forms*, Le Matematiche **72** (2017), no. 1, 137–144. 60

[BB21] Weronika Buczyńska and Jarosław Buczyński, *Apolarity, border rank, and multigraded Hilbert scheme*, Duke Math. J. **170** (2021), no. 16, 3659 – 3702. 58

[BBM14] A. Bernardi, J. Brachat, and B. Mourrain, *A comparison of different notions of ranks of symmetric tensors*, Linear Algebra Appl. **460** (2014), 205–230. 12, 50, 55, 59

[BBT13] Weronika Buczyńska, Jarosław Buczyński, and Zach Teitler, *Waring decompositions of monomials*, Journal of Algebra **378** (2013), 45–57. 54

[BC92] Michael Ben-Or and Richard Cleve, *Computing algebraic formulas using a constant number of registers*, SIAM J. Comput. **21** (1992), no. 1, 54–58. 38

[BDI21] Markus Bläser, Julian Dörfler, and Christian Ikenmeyer, *On the Complexity of Evaluating Highest Weight Vectors*, 36th Computational Complexity Conference (CCC 2021) **200** (2021), 29:1–29:36. 7, 9, 22

[BHIM22] Paul Breiding, Reuven Hodges, Christian Ikenmeyer, and Mateusz Michałek, *Equations for GL invariant families of polynomials*, Vietnam Journal of Mathematics **50** (2022), no. 2, 545–556. 21

[BI13] Peter Bürgisser and Christian Ikenmeyer, *Explicit lower bounds via geometric complexity theory*, Proceedings of the 45th Annual ACM symposium on Theory of Computing, 2013, pp. 141–150. 60

[BI17] Peter Bürgisser and Christian Ikenmeyer, *Fundamental invariants of orbit closures*, Journal of Algebra **477** (2017), 390–434. 7, 8, 12, 18, 21, 22, 60

[BIP19] Peter Bürgisser, Christian Ikenmeyer, and Greta Panova, *No occurrence obstructions in geometric complexity theory*, Journal of the American Mathematical Society **32** (2019), no. 1, 163–193. 3, 9

[BIZ18] Karl Bringmann, Christian Ikenmeyer, and Jeroen Zuiddam, *On algebraic branching programs of small width*, J. ACM **65** (2018), no. 5, 32:1–32:29. 13, 38, 40, 43, 60

[BLMW11] Peter Bürgisser, J. M. Landsberg, Laurent Manivel, and Jerzy Weyman, *An overview of mathematical issues arising in the geometric complexity theory approach to VP≠VNP*, SIAM J. Comput. **40** (2011), no. 4, 1179–1209. 9, 21

[Bre74] Richard P. Brent, *The parallel evaluation of general arithmetic expressions*, J. ACM **21** (1974), no. 2, 201–206. 37, 40

[BT15] G. Blekherman and Z. Teitler, *On maximum, typical and generic ranks*, Mathematische Annalen **362** (2015), no. 3-4, 1021–1031. 9, 55

[Car06] Enrico Carlini, *Reducing the number of variables of a polynomial*, Algebraic geometry and geometric modeling, Springer, 2006, pp. 237–247. 18

[CCG12] Enrico Carlini, Maria Virginia Catalisano, and Anthony V. Geramita, *The solution to the waring problem for monomials and the sum of coprime monomials*, Journal of Algebra **370** (2012), 5–14. 23

[CGL+21] A. Conner, F. Gesmundo, J. M. Landsberg, E. Ventura, and Y. Wang, *Towards a geometric approach to Strassen's asymptotic rank conjecture*, Collect. Math. **72** (2021), no. 1, 63–86. 19

[CIM17] Man-Wai Cheung, Christian Ikenmeyer, and Sevak Mkrtchyan, *Symmetrizing tableaux and the 5th case of the Foulkes conjecture*, Journal of Symbolic Computation **80** (2017), 833–843. 22

[CKW11] Xi Chen, Neeraj Kayal, and Avi Wigderson, *Partial Derivatives in Arithmetic Complexity (and beyond)*, Foundation and Trends in Theoretical Computer Science **6** (2011), no. 1–2, 1–138. 7

[DDS22] Pranjal Dutta, Prateek Dwivedi, and Nitin Saxena, *Demystifying the border of depth-3 algebraic circuits*, 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS) (2022), 92–103. 6, 7, 9

[DIP20] Julian Dörfler, Christian Ikenmeyer, and Greta Panova, *On geometric complexity theory: Multiplicity obstructions are stronger than occurrence obstructions*, SIAM Journal on Applied Algebra and Geometry **4** (2020), no. 2, 354–376. 7

[Dri97] Arthur A Drisko, *On the number of even and odd Latin squares of order p+1*, Advances in Mathematics **128** (1997), no. 1, 20–35. 8, 60

[DS22] Pranjal Dutta and Nitin Saxena, *Separated borders: Exponential-gap fanin-hierarchy theorem for approximative depth-3 circuits*, 63^{rd} Annual IEEE Symposium on the Foundations of Computer Science (FOCS 2022) (2022). 7

[FH91] William Fulton and Joe Harris, *Representation theory: a first course*, vol. 129, Springer Science & Business Media, 1991. 21

[FM19] Benjamin Friedman and Sean McGuinness, *The Alon–Tarsi conjecture: A perspective on the main results*, Discrete Math. **342** (2019), no. 8, 2234–2253. 60

[For16] Michael Forbes, *Some concrete questions on the border complexity of polynomials. presentation given at the workshop on algebraic complexity theory WACT 2016 in Tel Aviv*, 2016. 7

[Ful97] William Fulton, *Young tableaux: with applications to representation theory and geometry*, no. 35, Cambridge University Press, 1997. 20

[Gay76] David A Gay, *Characters of the Weyl group of $SU(n)$ on zero weight spaces and centralizers of permutation representations*, The Rocky Mountain Journal of Mathematics **6** (1976), no. 3, 449–455. 21

[Ges16] F. Gesmundo, *Geometric Aspects of Iterated Matrix Multiplication*, J. Algebra **461** (2016), 42–64. 19

[GKS17] Rohit Gurjar, Arpita Korwar, and Nitin Saxena, *Identity testing for constant-width, and any-order, read-once oblivious arithmetic branching programs*, Theory Comput. **13** (2017), no. 1, 1–21. 7

[Gly10] David G Glynn, *The conjectures of Alon–Tarsi and Rota in dimension prime minus one*, SIAM Journal on Discrete Mathematics **24** (2010), no. 2, 394–399. 8, 60

[Gor94] Paul Gordan, *Das Zerfallen der Curven in gerade Linien*, Mathematische Annalen **45** (1894), no. 3, 410–427. 7

[Gro61] Alexander Grothendieck, *Techniques de construction et théorèmes d'existence en géométrie algébrique IV : les schémas de Hilbert*, Séminaire Bourbaki : années 1960/61, exposés 205-222, Séminaire Bourbaki, no. 6, Société mathématique de France, 1961 (fr). MR 1611822 58

[Gua18] Yonghui Guan, *Brill's equations as a $GL(V)$-module*, Linear Algebra and its Applications **548** (2018), 273–292. 7

[GW09] R. Goodman and N. R. Wallach, *Symmetry, representations, and invariants*, Graduate Texts in Mathematics, vol. 255, Springer, Dordrecht, 2009. 21

[GY10] John Hilton Grace and Alfred Young, *The algebra of invariants*, Cambridge Library Collection - Mathematics, Cambridge University Press, 2010. 52

[Har77] Robin Hartshorne, *Algebraic geometry*, Springer-Verlag, New York, 1977, Graduate Texts in Mathematics, No. 52. MR 0463157 (57 #3116) 56

[HR94] Rosa Huang and Gian-Carlo Rota, *On the relations of various conjectures on Latin squares and straightening coefficients*, Discrete Mathematics **128** (1994), no. 1-3, 225–236. 60

[HS04] Mark Haiman and B Sturmfels, *Multigraded Hilbert schemes*, Journal of Algebraic Geometry **13** (2004), no. 4, 725–769. 57

[Hüt17] Jesko Hüttenhain, *Geometric complexity theory and orbit closures of homogeneous forms*, Ph.D. thesis, TU Berlin, 2017. 6

[Iar95] A. Iarrobino, *Inverse system of a symbolic power II. the Waring problem for forms*, Journal of Algebra **174** (1995), no. 3, 1091–1110. 12, 50

[IK99] Anthony Iarrobino and Vassil Kanev, *Power sums, Gorenstein algebras, and determinantal loci*, Lecture Notes in Mathematics, vol. 1721, Springer-Verlag, Berlin, 1999. 13, 52, 58

[IK20] Christian Ikenmeyer and Umangathan Kandasamy, *Implementing geometric complexity theory: On the separation of orbit closures via symmetries*, Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, 2020, pp. 713–726. 7, 8, 21, 22

[IP17] Christian Ikenmeyer and Greta Panova, *Rectangular Kronecker coefficients and plethysms in geometric complexity theory*, Advances in Mathematics **319** (2017), 40–66. 9

[IS22] Christian Ikenmeyer and Abhiroop Sanyal, *A note on VNP-completeness and border complexity*, Information Processing Letters **176** (2022), 106243. 5

[KL14] Harlan Kadish and Joseph M Landsberg, *Padded polynomials, their cousins, and geometric complexity theory*, Communications in Algebra **42** (2014), no. 5, 2171–2180. 3, 9

[Kra85] Hanspeter Kraft, *Geometrische Methoden in der Invariantentheorie*, Springer, 1985. 5

[KS07] Neeraj Kayal and Nitin Saxena, *Polynomial identity testing for depth 3 circuits*, computational complexity **16** (2007), no. 2, 115–138. 18

[Kum15] Shrawan Kumar, *A study of the representations supported by the orbit closure of the determinant*, Compositio Mathematica **151** (2015), no. 2, 292–312. 7, 60

[Kum20] Mrinal Kumar, *On the power of border of depth-3 arithmetic circuits*, ACM Trans. Comput. Theory **12** (2020), no. 1, 5:1–5:8. 4, 5, 6, 7, 14

[Lan15] Joseph M Landsberg, *Geometric complexity theory: an introduction for geometers*, Annali dell'università di Ferrara **61** (2015), no. 1, 65–117. 9

[Lan17] J. M. Landsberg, *Geometry and complexity theory*, Cambridge Studies in Advanced Mathematics, vol. 169, Cambridge University Press, Cambridge, 2017. 12, 20

[LL89] Thomas Lehmkuhl and Thomas Lickteig, *On the order of approximation in approximative triadic decompositions of tensors*, Theoretical Computer Science **66** (1989), no. 1, 1–14. 9

[LT10] J. M. Landsberg and Zach Teitler, *On the ranks and border ranks of symmetric tensors*, Found. Comput. Math. **10** (2010), no. 3, 339–366. 9, 60, 63

[LZX21] Xin Li, Liping Zhang, and Hanchen Xia, *Two classes of minimal generic fundamental invariants for tensors*, arXiv:2111.07343, 2021. 60

[Mac95] I. G. Macdonald, *Symmetric functions and Hall polynomials*, second ed., Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1995. 14

[Mat60] Yozô Matsushima, *Espaces homogènes de Stein des groupes de Lie complexes*, Nagoya Mathematical Journal **16** (1960), 205–218. 21

[MS01] Ketan Mulmuley and Milind A. Sohoni, *Geometric complexity theory I: an approach to the P vs. NP and related problems*, SIAM J. Comput. **31** (2001), no. 2, 496–526. 3, 9

[MS08] Ketan D Mulmuley and Milind Sohoni, *Geometric complexity theory II: towards explicit obstructions for embeddings among class varieties*, SIAM Journal on Computing **38** (2008), no. 3, 1175–1206. 3, 9

[MS21] Dori Medini and Amir Shpilka, *Hitting sets and reconstruction for dense orbits in VP_e and $\Sigma\Pi\Sigma$ circuits*, 36th Computational Complexity Conference (CCC 2021) **200** (2021), 19:1–19:27. 9, 60

[NW97] Noam Nisan and Avi Wigderson, *Lower bounds on arithmetic circuits via partial derivatives*, Computational Complexity **6** (1997), no. 3, 217–234. 9

[Oed19] Luke Oeding, *Border ranks of monomials*, arXiv:1608.02530v3, 2019. 11, 15, 33

[Sap21] Ramprasad Saptharishi, *A survey of lower bounds in arithmetic circuit complexity*, 2021, Github Survey. 41, 46

[Sax08] Nitin Saxena, *Diagonal Circuit Identity Testing and Lower Bounds*, ICALP 2008, 2008, pp. 60–71. 7

[Shp02] Amir Shpilka, *Affine projections of symmetric polynomials*, Journal of Computer and System Sciences **65** (2002), no. 4, 639–659, Special Issue on Complexity 2001. 15, 60

[SW12] Douglas S Stones and Ian M Wanless, *How not to prove the Alon-Tarsi conjecture*, Nagoya Mathematical Journal **205** (2012), 1–24. 60

[Syl52] J. J. Sylvester, *On the principles of the calculus of forms*, J. Cambridge and Dublin Math. **7** (1852), 52–97. 18

[Val79] Leslie G. Valiant, *Completeness classes in algebra*, Proceedings of the 11h Annual ACM Symposium on Theory of Computing, 1979, pp. 249–261. 4

[VSBR83] L. G. Valiant, S. Skyum, S. Berkowitz, and C. Rackoff, *Fast parallel computation of polynomials using few processors*, SIAM Journal on Computing **12** (1983), no. 4, 641–644. 13, 40, 46, 49

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