Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 157 (2022)

De-bordering and Geometric Complexity Theory for Waring Rank and Related Models

Pranjal Dutta* Fulvio Gesmundo † Christian Ikenmeyer‡ Gorav Jindal§ Vladimir Lysikov¶

Abstract

De-bordering is the task of proving that a border complexity measure is bounded from below, by a non-border complexity measure. This task is at the heart of understanding the difference between Valiant's determinant vs permanent conjecture, and Mulmuley and Sohoni's Geometric Complexity Theory (GCT) approach to settle the $\mathsf{P} \neq \mathsf{NP}$ conjecture. Currently, very few de-bordering results are known.

In this work, we study the question of de-bordering the border Waring rank of polynomials. Waring and border Waring rank are very well studied measures, in the context of invariant theory, algebraic geometry and matrix multiplication algorithms. For the first time, we obtain a Waring rank upper bound that is exponential in the border Waring rank and only *linear* in the degree. All previous results were known to be exponential in the degree.

According to Kumar's recent surprising result (ToCT'20), a small border Waring rank implies that the polynomial can be approximated as a sum of a constant and a small product of linear polynomials. We prove the converse of Kumar's result, and in this way we de-border Kumar's complexity, and obtain a new formulation of border Waring rank, up to a factor of the degree. We phrase this new formulation as the orbit closure problem of the product-plus-power polynomial, and we successfully de-border this orbit closure. We fully implement the GCT approach against the power sum, and we generalize the ideas of Ikenmeyer-Kandasamy (STOC'20) to this new orbit closure. In this way, we obtain new multiplicity obstructions that are constructed from just the symmetries of the points and representation theoretic branching rules, rather than explicit multilinear computations.

Furthermore, we realize that the generalization of our converse of Kumar's theorem to square matrices gives a homogeneous formulation of Ben-Or and Cleve (SICOMP'92). This results for the first time in a VF-complete family under *homogeneous projections*. We study this approach further and obtain that a homogeneous variant of the continuant polynomial, which was studied by Bringmann, Ikenmeyer, Zuiddam (JACM'18), is VQP-complete under homogeneous border qp-projections. 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).

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

*PD: School of Computing, National University of Singapore (NUS); 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

Acknowledgements

P.D. is supported by the project "Foundation of Lattice-based Cryptography", funded by NUS-NCS Joint Laboratory for Cyber Security. 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).

1 Introduction

Waring rank and border Waring rank. Given a homogeneous degree d polynomial f, the *Waring rank* of f, denoted $\mathsf{WR}(f)$, is the smallest r such that there exist homogeneous linear polynomials (also called 'homogeneous linear forms') $\ell_1, \ldots, \ell_r$, with $f = \sum_{i\in[r]} \ell_i^d$. In the case of polynomials of degree two, this notion is equivalent to the rank of the symmetric matrix associated to a quadratic form; hence Waring rank can be regarded as a generalization of the rank of a symmetric matrix. Unlike the case of matrices, when $d \geq 3$, Waring rank is not in general lower semicontinous [1]. The *border Waring rank* of f, denoted $\underline{\mathsf{WR}}(f)$, is the smallest r such that f can be written as limit of a sequence of polynomials f_ϵ with $\mathsf{WR}(f_\epsilon) = r$.

Waring rank was studied already in the eighteenth century [Cay45, Syl52, Cle61] in the context of invariant theory, with the aim to determine normal forms for homogeneous polynomials. We mention the famous Sylvester Pentahedral Theorem, stating that a generic cubic form in four variables can be written uniquely as sum of five cubes. At the beginning of the twentieth century, the early work on secant varieties in classical algebraic geometry [Pal06, Ter11] implicitly commenced the study of border Waring rank. The notion of border rank for tensors was introduced in [BCRL79] to construct faster-than-Strassen matrix multiplication algorithms. In [Bin80], Bini proved that tensor border rank and tensor rank define the same matrix multiplication exponent. Today this theory is deeply related to the study of Gorenstein algebras [IK99, BB14], the Hilbert scheme of points [Jel20], and deformation theory [BB21, JM22].

In algebraic complexity theory, Waring rank defines a model of computation also known as the homogeneous diagonal depth 3 circuits, see e.g. [Sax08]. Moreover, Bini's Theorem is not limited to tensors. In fact, the results of [CHI$^+$18] guarantee that the matrix multiplication exponent is controlled by the border Waring rank of a *symmetrized* version of the matrix multiplication tensor. This applies more generally to asymptotic notions where the order of the tensor (or equivalently the degree of the polynomial) is constant [CGJ19].

Algebraic complexity theory. A sequence $(c_n)_{n\in\mathbb{N}}$ of natural numbers is *polynomially bounded* if there exists a polynomial q with $\forall n : c_n \leq q(n)$. A p-family is a sequence of polynomials whose degree and number of variables is polynomially bounded. 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 to be the polynomial computed by the root. The size of a formula is the number of vertices of the tree. The *algebraic formula complexity* of a polynomial f is defined as the smallest size of an algebraic formula computing f. VF is the class of p-families $(f_n)_{n\in\mathbb{N}}$ for which the sequence of algebraic formula complexities of f_n is polynomially bounded. If we allow arbitrary directed acyclic graphs instead of just trees (i.e., the out degree of a node can be ≥ 2), then this gives the notion of *algebraic circuit*, with the corresponding notion of the algebraic circuit complexity. The class VP consists of the p-families $(f_n)_{n\in\mathbb{N}}$ for which the algebraic circuit complexity is polynomially bounded.

A third important model of computation is the *algebraic branching program* (ABP) model. It is a

[1] A function f is lower semicontinuous at a if $\liminf_{x\to a} f(x) \geq f(a)$. The function is lower semicontinuous if it is semicontinuous at every point of its domain.

classical result that every 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 affine linear forms. Note that if f is homogeneous, then it is clear that all entries can be chosen to be homogeneous linear. If f can be computed in the above form, we say that f has an ABP of width n. Let $\mathsf{abpw}(f)$ be the smallest n such that f has an ABP of width n. The class VBP consists of the p-families $(f_n)_{n\in\mathbb{N}}$ for which $\mathsf{abpw}(f_n)$ is polynomially bounded. Alternatively, one can define the *determinantal complexity* of a polynomial f to be the smallest n such that f can be written as the determinant of an $n \times n$ matrix of affine linear forms. The class VBP consists of all p-families $(f_n)_{n\in\mathbb{N}}$ for which the determinantal complexity is polynomially bounded, see e.g. [Mah14].

Let $\mathrm{per}_n := \sum_{\sigma\in\mathfrak{S}_n}\prod_{i=1}^{n} x_{i,\sigma(i)}$ be the permanent polynomial (on n^2 variables), where $\mathfrak{S}_n$ is the symmetric group of n elements. The *permanental complexity* of a polynomial f is the smallest n such that f can be written as the permanent of an $n \times n$ matrix of affine linear forms. The class VNP consists of all p-families $(f_n)_{n\in\mathbb{N}}$ for which the permanental complexity is polynomially bounded.

It is known that $\mathsf{VF} \subseteq \mathsf{VBP} \subseteq \mathsf{VP} \subseteq \mathsf{VNP}$ [Val79, Tod92]. The conjectures $\mathsf{VF} \neq \mathsf{VNP}$, $\mathsf{VBP} \neq \mathsf{VNP}$, $\mathsf{VP} \neq \mathsf{VNP}$, are known as *Valiant's conjectures*. Especially $\mathsf{VBP} \neq \mathsf{VNP}$ is known as the *determinant vs permanent* problem.

A sequence $(c_n)_{n\in\mathbb{N}}$ of natural numbers is *quasipolynomially bounded* if there exists a polynomial q with $\forall n \geq 2 : c_n \leq n^{q(\log_2 n)}$. In the definitions of VF, VBP, VP, VNP, if we change the upper bound on the complexity to "quasipolynomially bounded" instead of just "polynomially bounded", then we obtain the classes VQF, VQBP, VQP, VQNP, respectively. It turns out that $\mathsf{VQF} = \mathsf{VQBP} = \mathsf{VQP}$, see [Bür00]. The conjecture $\mathsf{VNP} \not\subseteq \mathsf{VQP}$ is called Valiant's *extended* conjecture.

Border complexity. Border complexity for algebraic circuits was first discussed in 2001 in [MS01] and [Bür04]. To any algebraic complexity measure one can define the corresponding *border complexity* (i.e., border formula complexity, border abpw, etc), in the same way as border Waring rank arises from Waring rank: The border complexity of f is the smallest n such that f can be approximated arbitrarily closely by polynomials of complexity at most n. In analogy to border Waring rank, border complexity measures are usually underlined, for example $\underline{\mathsf{abpw}}$. Replacing a complexity measure by its border measure in a complexity class, we obtain the *closure* of this class, $\overline{\mathsf{VF}}$, $\overline{\mathsf{VBP}}$, and so on. The operation of going to the closure is indeed a closure operator in the sense of topology, see [IS22].

The closure can be defined concisely using the notion of degenerations as follows. Let $f \in \mathbb{C}[U]_d$, $g \in \mathbb{C}[W]_d$ for finite dimensional complex vector spaces U, W. 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 these closures can be taken, equivalently, in the Euclidean or the Zariski topology, see e.g. [Kra85, AI.7.2 Folgerung]. For $\mathcal{C} \in \{\mathsf{VF}, \mathsf{VBP}, \mathsf{VP}, \mathsf{VNP}\}$ define $\overline{\mathcal{C}}$ via $(f_n)_n \in \overline{\mathcal{C}} \iff \exists (g_n)_n \in \mathcal{C}\ \forall n : f_n \trianglelefteq_{\mathrm{aff}} g_n$.

The paper [MS01] proposes to study the $\mathsf{VNP} \not\subseteq \overline{\mathsf{VBP}}$ conjecture (see [BLMW11]), and [Bür04] does so for $\mathsf{VNP} \not\subseteq \overline{\mathsf{VP}}$. These border variants of Valiant's conjecture are often referred to as the Mulmuley-Sohoni conjectures. In analogy to [Bür00], one observes $\overline{\mathsf{VQF}} = \overline{\mathsf{VQBP}} = \overline{\mathsf{VQP}}$, which also coincides with the closure of VQP in the sense of [IS22]. We define the *extended Mulmuley-Sohoni conjecture* as $\mathsf{VNP} \not\subseteq \overline{\mathsf{VQP}}$.

It is even open whether or not $\overline{\mathsf{VF}} \subseteq \mathsf{VNP}$. In order to make progress on these and related questions, one tries to prove results of the form $\overline{\mathcal{C}} \subseteq \mathcal{D}$ for algebraic complexity classes $\mathcal{C}$ and $\mathcal{D}$, a process called *de-bordering*. This can also be done directly on the complexity measures, without the need to define complexity classes. For example, $\mathsf{abpw}(f) \leq \underline{\mathsf{WR}}(f)$, see [BDI21, Thm 4.2] and [For14, Sec. 4.5.2]. In terms of complexity classes, this means $\overline{\mathsf{VWaring}} \subseteq \mathsf{VBP}$, where $\mathsf{VWaring}$ is the class of p-families that have polynomially bounded Waring rank,.

In fact, unlike in the matrix multiplication case, most questions about the relation between complexity classes and their closures is wide open, for example, it is wide open if Valiant's conjecture is equivalent to the Mulmuley-Sohoni conjecture, even in the extended case. For width 2 algebraic branching programs, the complexity class differs from is closure [BIZ18]. The same is true for the sum of two products of affine linear forms [Kum20]:

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

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

If no such m exists, we set $\mathsf{Kc}(f) := \infty$. Let $\underline{\mathsf{Kc}}$ be the corresponding border complexity. Formally, 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 homogeneous linear forms $\ell_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]_1$, $\alpha \in \mathbb{C}(\epsilon)$ with $f_\epsilon \simeq f$ and $\forall \beta \neq 0 : \mathsf{Kc}(f_\epsilon|_{\epsilon=\beta}) \leq m$. Now, $\underline{\mathsf{Kc}}$ 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 $\ell_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]_1$. 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. Kumar [Kum20] proved that for all homogeneous polynomials f, we have $\underline{\mathsf{Kc}}(f) \leq \deg(f)\cdot\mathsf{WR}(f)$, and it is easy to see that this implies $\underline{\mathsf{Kc}}(f) \leq \deg(f)\cdot\underline{\mathsf{WR}}(f)$.

Orbit closures. We write $\mathrm{GL}_n := \mathrm{GL}(\mathbb{C}^n)$. Given a homogeneous degree d polynomial f in n variables and an invertible matrix $g \in \mathrm{GL}_n$, we define gf via $\forall x \in \mathbb{C}^n : (gf)(x) := f(g^t x)$ [2]. The set $\mathrm{GL}_n f$ is called the *orbit* of f under the group action of GL_n. The closure of an orbit is defined as the set of all limit points of sequences of orbit points. It turns out that this Euclidean closure *coincides* with the Zariski closure. For example, $\overline{\mathrm{GL}_n(x_1x_2\cdots x_n)}$ is the set of all n-variable homogeneous degree n polynomials that can be written as a product of homogeneous linear polynomials (or as a limit of such, but in this specific case, this makes no difference). As another example, consider the power sum polynomial $x_1^d + \cdots + x_n^d$. It is easy to see that for a homogeneous degree d polynomial f we have $\underline{\mathsf{WR}}(f) \leq n$ if and only if $f \in \overline{\mathrm{GL}_n(x_1^d + \cdots + x_n^d)}$, provided f is a polynomial in at most n variables. Let $\mathbb{C}^{n\times n}$ denote the set of all $n \times n$ matrices, and note that in general $\mathbb{C}^{n\times n}(x_1^d + \cdots + x_n^d) \subsetneq \overline{\mathrm{GL}_n(x_1^d + \cdots + x_n^d)}$. In fact, for a homogeneous degree d polynomial f we have $\mathsf{WR}(f) \leq n$ if and only if $f \in \mathbb{C}^{n\times n}(x_1^d + \cdots + x_n^d)$, provided f is a polynomial in at most n variables.

In analogy to the power sum, consider the homogeneous degree d iterated $n \times n$ matrix multiplication polynomial in $N := n^2(d-1) + 2n$ many variables:

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

[2] The transpose is needed to ensure that $g(h(f)) = (gh)(f)$ for all $g, h \in \mathrm{GL}_n$.

It is easy to see that for a homogeneous degree d polynomial f we have $\underline{\mathsf{abpw}}(f) \leq n$ if and only if $f \in \overline{\mathrm{GL}_N \mathsf{IMM}_n^d}$, provided f is a polynomial in at most N variables. One can see that the Mulmuley-Sohoni conjecture is *equivalent* to the statement that $\underline{\mathsf{abpw}}(\mathrm{per}_n)$ grows superpolynomially. As we have seen, this is a question about orbit closures. In fact, $f \in \overline{\mathrm{GL}_N \mathsf{IMM}_n^d}$ if and only if $\overline{\mathrm{GL}_N f} \subseteq \overline{\mathrm{GL}_N \mathsf{IMM}_n^d}$, so it is a question about orbit closure containment.

In principle, the theory also works for inhomogeneous polynomials, but either the general linear group has to be replaced by the general affine group, which is not reductive, or the polynomial has to be *padded*, as in [MS01]. The *padded permanent* is the polynomial $x_0^{n-m}\mathrm{per}_m$, homogeneous of degree n in m^2+1 variables. The Mulmuley-Sohoni conjecture is equivalent to proving that there is no polynomially bounded (in m) n such that $\overline{\mathrm{GL}_{n^2}\, x_0^{n-m}\mathrm{per}_m} \subseteq \overline{\mathrm{GL}_{n^2} \det_n}$.

Representation Theory. One main idea of geometric complexity theory is to use the symmetries of the two polynomials to prove the non-containment of one orbit closure in the other.

Fix a homogeneous degree d polynomial p in n variables, i.e., $p \in \mathbb{C}[x_1,\ldots,x_n]_d =: S^d(\mathbb{C}^n)$ for short. Let $I(\overline{\mathrm{GL}_n\, p})$ denote the vanishing ideal of its orbit closure, i.e., the set of all polynomials that vanish identically on $\overline{\mathrm{GL}_n\, p}$. These polynomials are polynomials *in the coefficients* of p, in other words, for the homogeneous degree D part of the vanishing ideal we have $I(\overline{\mathrm{GL}_n\, p})_D \subseteq S^D(S^d(\mathbb{C}^n))$. The coordinate ring is defined as the quotient $\mathbb{C}[\overline{\mathrm{GL}_n\, p}]_D := S^D(S^d(\mathbb{C}^n))/I(\overline{\mathrm{GL}_n\, p})_D$.

For fixed homogeneous degree d polynomials p and q in n variables, assume that $\overline{\mathrm{GL}_n\, p} \subseteq \overline{\mathrm{GL}_n\, q}$. Hence, $I(\overline{\mathrm{GL}_n\, q}) \subseteq I(\overline{\mathrm{GL}_n\, p})$, and therefore for every D we get an equivariant surjection $\mathbb{C}[\overline{\mathrm{GL}_n\, q}]_D \twoheadrightarrow \mathbb{C}[\overline{\mathrm{GL}_n\, p}]_D$. Note that we can define a group action on a coordinate ring via $gf(x) := f(g^{-1}x)$, similar to the usual group action on polynomials, but this time on polynomials in coefficients of other polynomials. This makes both $\mathbb{C}[\overline{\mathrm{GL}_n\, p}]_D$ and $\mathbb{C}[\overline{\mathrm{GL}_n\, q}]_D$ finite dimensional representations of GL_n, which is reductive, and hence both decompose into a direct sum of irreducible representations. The *representation theoretic multiplicity* mult_λ counts the number of irreducibles of type λ in such a decomposition (the actual number is independent of the decomposition). By Schur's lemma (see [FH91]), an equivariant surjection implies an inequality of representation theoretic multiplicities: $\mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_n\, q}]_D) \geq \mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_n\, p}]_D)$. Hence, any λ violating this inequality (i.e., any λ with $\mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_n\, q}]_D) < \mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_n\, p}]_D)$) proves that $\overline{\mathrm{GL}_n\, p} \not\subseteq \overline{\mathrm{GL}_n\, q}$. This is called a *representation theoretic multiplicity obstruction*. If additionally $\mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_n\, q}]_D) = 0$, then we call this an *occurrence obstruction*. Occurrence obstructions can be used to prove lower bounds on the border rank of the matrix multiplication tensor, see [BI11, BI13]. In order to prove the required lower bound on $\mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_n\, p}]_D)$, these papers use elaborate constructions of highest weight vectors, instead of making use of the symmetry of the point.

The original Geometric Complexity Theory (GCT) papers [MS01, MS08] propose to use occurrence obstructions to separate the orbit closures of the determinant and the padded permanent polynomial. This was commonly known as *the GCT approach*, until [IP17, BIP19] proved that this is impossible (the *no-go theorem*), by making use of the fact that [MS01, MS08] use the padded formulation of the Mulmuley-Sohoni conjecture. There exists *no such counterexample* when the determinant if replaced by IMM_n^d (which implies that per_d is not padded). The first paper where multiplicity obstructions could be found *without* an explicit construction of a highest weight vector is [IK20], where the symmetries of both points are used to construct a multiplicity obstruction λ that shows that the power sum polynomial is not a product of homogeneous linear forms (there are much easier ways to prove this though).

1.a Main Results

Let $V := \mathbb{C}[x_1, \ldots, x_n]_1 := \mathrm{Span}(\{x_1, \ldots, x_n\})$ be the space of linear forms in $x_1, \ldots, x_n$, and write $S^d(V) := \mathbb{C}[x_1, \ldots, x_n]_d =: \mathbb{C}[V^*]_d$ as the space of homogeneous polynomials of degree d. Fix integers d, r, s. Throughout the paper, we will be using the following 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 $x_{11}, \ldots, x_{dr}, y_1, \ldots, y_s$. Of special interest will be the case $r = s = 1$, i.e., the *product-plus-power* polynomial. When $s = 0$, it is $\sum_{i=1}^{r} \prod_{j=1}^{d} x_{ji}$, and when $r = 0$, it is the *power sum* polynomial $\sum_{i=1}^{s} y_i^d$.

We first prove a *fixed-parameter de-bordering* theorem for border Waring rank:

1.1 Theorem. *Let $f \in \mathbb{C}[x_1, \ldots, x_n]_d$ with $\underline{\mathsf{WR}}(f) = r$. Then, $\mathsf{WR}(f) \leq 4^r \cdot d$.*

For details, see Theorem 2.12 in Section 2.b. To the best of our knowledge, previous methods only allow upper bounds of the order d^r or r^d. To get $\mathsf{WR}(f) \leq O(d^r)$, note that a polynomial with border Waring rank r can be transformed into a polynomial in only r variables using a linear change of variables (see Lemma 2.2), and then take the maximal possible Waring rank of an r-variate polynomial of degree d as an upper bound. 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 ϵ (which could be *exponentially large* in the degree of the polynomial) 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.

There are several papers that prove that $\mathsf{WR}(f) = O(\deg f)$ if the value of $\underline{\mathsf{WR}}(f)$ is small using case-by-case analysis ([LT10] for $\underline{\mathsf{WR}}(f) \leq 3$, [Bal19] for $\underline{\mathsf{WR}}(f) \leq 4$). Ballico and Bernardi [BB17] conjecture that $\mathsf{WR}(f) \leq (\underline{\mathsf{WR}}(f) - 1) \cdot \deg f$ and prove an analogue of this statement for a weak version of border rank called curvilinear rank.

We continue in our study of border Waring rank by focusing on Kumar's recent result [Kum20], where he uses a connection to elementary symmetric polynomials (originally observed by Shpilka [Shp02]) to prove that small Waring rank implies small border $\Sigma^{[2]}\Pi\Sigma$-complexity, more precisely $\underline{\mathsf{Kc}}(f) \leq \deg(f) \cdot \mathsf{WR}(f)$, which also implies $\underline{\mathsf{Kc}}(f) \leq \deg(f) \cdot \underline{\mathsf{WR}}(f)$. We develop a border version of the Newton identities and use it to prove the *converse* of Kumar's statement:

1.2 Theorem (Converse of Kumar's theorem)**.** *For all homogeneous f, $\underline{\mathsf{WR}}(f) \leq \underline{\mathsf{Kc}}(f)$ or f is a product of linear forms.*

For details, see Corollary 3.8. This also immediately gives a de-bordering result for $\underline{\mathsf{Kc}}$, because $\mathsf{abpw}(f) \leq \underline{\mathsf{WR}}(f)$. Since $\underline{\mathsf{WR}}(f) \leq \underline{\mathsf{Kc}}(f) \leq \deg(f)\underline{\mathsf{WR}}(f)$, it is reasonable to study $\underline{\mathsf{Kc}}$ on its own right. We therefore set up the corresponding GL_{n+1} orbit closure problem of the polynomials $P_{1,1}^{[d]}$ and $P_{1,2}^{[d]}$. Both of them turn out to be an orbit closure that is contained in the orbit closure of the *binomial* $\mathrm{bn}_d := P_{2,0}^{[d]}$, which was studied by Jesko Hüttenhain in [Hüt17].

Interestingly, the orbit closure of $P_{1,1}^{[d]}$ and $\underline{\mathsf{Kc}}$ are intimately connected and can be described is as follows. For a polynomial $f \in S^d V$ 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

$$\underline{\mathsf{Kc}}(f) \le m \iff x_0^{m-d} f \trianglelefteq_{x_0}^{y_1} \prod_{i\in[m]} x_i + y_1^m \underset{\not\Leftarrow}{\Longrightarrow} x_0^{m-d} f \trianglelefteq P_{1,1}^{[m]} \iff f \trianglelefteq_{\mathrm{aff}} P_{1,1}^{[m]}.$$

The above formulation shows that proving de-bordering results for the orbit closure of $P_{1,1}^{[d]}$ is relevant for the study of $\underline{\mathsf{Kc}}$ and in turn for border Waring rank. In fact, we go further and even prove de-bordering for the product plus two powers, which is still contained in the orbit closure of the binomial.

1.3 Theorem (De-bordering product-plus-powers). *Let $f \in \mathbb{C}[x_1,\ldots,x_n]_d$. The following holds.*

(i) If $f \trianglelefteq P_{1,1}^{[d]}$, then either $f \le P_{1,1}^{[d]}$, or $\mathsf{abpw}(f) \le \underline{\mathsf{WR}}(f) \le O(d^5)$.

(ii) If $f \trianglelefteq P_{1,2}^{[d]}$, then either $f \le P_{1,2}^{[d]}$, or $f \le \prod_{i\in[d]} y_i + y_0^{d-1} \cdot y_{d+1}$, or $\mathsf{abpw}(f) \le \underline{\mathsf{WR}}(f) \le O(d^8)$.

See Theorem 4.11 and Theorem 4.12 for the details. Note that, both product-plus-power and product-plus-two-powers are special binomials [3], and hence, from [DDS22], it follows that their orbit closures are contained in VBP. Our results are more *fine-grained* de-bordering than VBP, since $\mathsf{WR}(f) \le \mathsf{poly}(d) \implies f \in \mathsf{VBP}$ [For16, GKS17, BDI21], and the converse does not necessarily hold, because $\mathsf{WR}(\det_d) = \exp(d)$ [Sax08, CKW11a]. In Section 4.b, we also show some lower bound results in these computational models.

Having set up the orbit closure formulation, in §5 we use it as a new test-bed for GCT. We generalize [IK20] from the product polynomial to product-plus-power by exhibiting multiplicity obstructions that are based entirely on the symmetries of the two polynomials.

1.4 Theorem (New obstructions). *Let $d \ge 3$, and let $\lambda = (5d-1,1) + ((d+1) \times (10d))$. Then we have representation theoretic multiplicity obstructions:* $\mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_{d+1} P_{1,1}^{[d]}}]) \le 4 < 5 = \mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_{d+1} P_{0,d+1}^{[d]}}])$*, and hence* $\overline{\mathrm{GL}_{d+1} P_{0,d+1}^{[d]}} \not\subseteq \overline{\mathrm{GL}_{d+1} P_{1,1}^{[d]}}$.

These obstructions are only based on the symmetries of the two polynomials as in [IK20], see Theorem 5.11 for the details. This is still a toy case for GCT, but as seen above, it was proved by Kumar [Kum20] that $P_{1,1}^{[d]}$ is much more expressive than $P_{1,0}^{[d]}$ (which is just a monomial), which was studied in [IK20].

We get another result from studying the proof of Theorem 1.2 and generalizing it to matrices. This result starts with the observation that one can lift Kumar's result and its reverse to matrices and obtain a clean and homogeneous version of Ben-Or and Cleve's characterization of VF. We use this to find a new and very simple homogeneous polynomial that is $\overline{\mathsf{VF}}$-complete under *homogeneous* degenerations (see Definition 6.1 for the details of this definition). The parity-alternating elementary symmetric polynomial $C_{n,d}$ is defined via

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

where P is 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.

[3] Note $P_{1,2}^{[d]} = \prod_{i\in[d]} x_i + y_1^d + y_2^d = \prod_{i\in[d]} x_i + \prod_{i\in[d]}(y_1 - \zeta^{2i+1} y_2)$, where ζ is the $(2d)$-th root of unity.

1.5 Theorem (Homogeneous VF). *For a homogeneous polynomial f let $\underline{c} := \min\{r \mid f \in \overline{\mathrm{GL}_r C_{r,\deg(f)}}\}$. The extended Mulmuley-Sohoni conjecture is true if and only if $\underline{c}(\mathrm{per}_m)$ grows super-quasipolynomially.*

See Theorem 6.12 for the details. This is a *homogeneous, padding-free* formulation. The only other polynomial so far that allows for such a formulation is the homogeneous iterated matrix multiplication polynomial, which so far resisted all attempts for finding lower bounds via representation theoretic obstructions. However, it must be noted that $C_{n,d}$ is *not* characterized by its stabilizer, and maybe there exist even better polynomials that have the same simple structure as $C_{n,d}$, and are at the same time characterized by their stabilizer. In any case, its homogeneity and simplicity makes $C_{n,d}$ a good candidate for studying VF using methods from algebraic geometry in future work.

As a first toy proof of concept, we used the software from [BHIM22] on a laptop to check that the GL_5 isomorphism type $(5,5,5,3,3)$ occurs in the vanishing ideal $I(\overline{\mathrm{GL}_5 C_{5,3}})$, and the corresponding highest weight vector does not vanish on a random point in the GL_5 orbit of $p := x_1^3 + x_2^3 + x_3^3 + x_1x_2x_3 + x_3x_4x_5$, hence we have that $p \notin \overline{\mathrm{GL}_5 C_{5,3}}$. This is in fact an occurrence obstruction, because the plethysm coefficient $a_{(5,5,5,3,3)}(7,3) = 1$, i.e., the irreducible GL_5-representation of type $(5,5,5,3,3)$ is unique in $S^7(S^3(\mathbb{C}^5))$, which we directly computed using the `schur` software `http://sourceforge.net/projects/schur`. The software is randomized, in the sense that the vanishing is determined by vanishing on sampled pseudorandom points, so it remains to be symbolically verified that $(5,5,5,3,3)$ is indeed in $I(\overline{\mathrm{GL}_5 C_{5,3}})$.

1.b Proof Ideas

Proof idea of Theorem 1.1: De-bordering border 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 (see Section 2.c), but we also provide an elementary proof which *does not use* the language of algebraic geometry and is based on partial derivative techniques (see Section 2.b).

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 first decomposing each g_k with respect to a basis consisting of powers of linear forms, and then using the classical fact (see also [BBT13]) that $\mathsf{WR}(\ell_1^a \ell_2^b) \leq \max(a+1, b+1)$.

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; see Lemma 2.8. The key insight is that if the degree of polynomials involved is high enough, namely when $\deg f \geq \underline{\mathsf{WR}}(f) - 1$, 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 term of the form $\ell^{d-r+1} g$, where r is the number rank one summands in the part and ℓ is the common lowest order linear form; see Lemma 2.5.

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

$$f_d = \lim_{\varepsilon \to 0} \frac{1}{d\varepsilon} \left[(x_0 + \varepsilon y_0)^d - x_0^d + (x_1 + \varepsilon y_1)^d - x_1^d + 2(x_0 + x_1 + \varepsilon y_2)^d - 2(x_0 + x_1)^d \right] \tag{1.6}$$

and a lower bound is obtained by considering the dimension of the space of second order partial derivatives. The summands of the decomposition (1.6) are divided into three pairs. The lowest order term of the first pair is x_0, the one of the second pair is x_1 and the one of the third pair is $(x_0 + x_1)$. For each pair, the sum of the two powers individually converges to a limit as $\varepsilon \to 0$; these three limits are, respectively, $x_0^{d-1}y_0$, $x_1^{d-1}y_1$, and $2(x_0 + x_1)^{d-1}y_2$, which are the summands of a generalized additive decomposition for f_d.

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

$$f_3 = \lim_{\varepsilon\to 0} \frac{1}{9\varepsilon} \left[3(x_0 + \varepsilon y_0)^3 + 3(x_1 + \varepsilon y_1)^3 + 6(x_0 + x_1 + \varepsilon y_2)^3 - (x_0 + 2x_1)^3 - (2x_0 + x_3)^3\right]. \quad (1.7)$$

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 of Theorem 1.2: Converse of Kumar's theorem. We observe that $\underline{\mathsf{Kc}}$ expressions fall into three different cases, depending on whether the scalar α converges to 0, converges to nonzero, or diverges. We study these three case independently. For the case where α converges to zero, it is easy to see that the resulting polynomial is a product of affine linear polynomials, see Lemma 3.9. For the case where α converges to a nonzero value, we use the Newton Identities to show a lower bound given by the Waring rank, see Proposition 3.10. The case of α diverging (and hence, cancellations occurring in the limit) is the most interesting. The use of Newton Identities is *not sufficient* to resolve this case, so we develop a new tool: Border Newton Relations. With this new tool, the proof is short and elegant, see Theorem 3.11.

Proof idea of Theorem 1.3: De-bordering product-plus-power and product-plus-two-powers. The de-borderings of the product-plus-power and product-plus-two-powers models are based on their representation as restricted binomials. A power is product of equal linear forms, and a sum of two powers can be represented as a product using the identity $a^d - b^d = \prod_{i=0}^{d-1}(a - \omega^i b)$ where ω is a primitive d-th root of unity, so we consider the limits of sums of two products

$$\lim_{\epsilon\to 0}\left(\prod_{i=1}^{d} \ell_i(\epsilon) + \prod_{i=1}^{d} \ell_i'(\epsilon)\right)$$

where ℓ_i, ℓ_i' are families of linear forms depending on ϵ. One of the products in this sum is *restricted*, in the sense that it has a constant number of *essential* variables (1 in the case of product-plus-power, and 2 in the case of product-plus-two-powers); for more on essential variables, see Section 2.a. Because of this, our results do not extend to more than 2 powers.

To analyze the limits of sums of two powers, we again use the idea of "locality". If the two products in the sum do not have limits individually, then they must cancel in the lowest degree terms, and this only happens if the lowest degree terms are the same up to permutation and scaling of linear forms in a product. Using this, we show that now the lowest degree terms only have a constant number of essential variables, and by fixing the values of these variables, we obtain an 'almost' Kumar-like expression of the form

$$\lim_{\epsilon\to 0}\left(\alpha(\epsilon)\prod_{i=1}^{d}(1 + a_i(\epsilon)) - \beta(\epsilon)\prod_{i=1}^{d}(1 + b_i(\epsilon))\right)$$

where α, β are scalars and a_i, b_i are linear forms depending on ϵ. This model can be analyzed using the border Newton relations, similar to Theorem 1.2; for details see Theorem 4.8. Finally, we use polynomial interpolation to get a bound on the border Waring rank of the original polynomial from the bound in this Kumar-like model.

Proof idea of Theorem 1.4: New obstructions. At a first glance, the polynomial $P_{1,1}^{[d]} = \prod_{i\in[d]} x_i + x_0^d$ looks very similar to the well-studied product polynomial $\prod_{i\in[d]} x_i$, which was the object of several GCT papers [Kum15, BI17, DIP20, IK20]. A system of set-theoretic equations for its orbit-closure is known due to Brill and Gordon [Gor94], and their representation theoretic structure has been recently described by Guan [Gua18]. We try to transfer as much of the theory as possible from the product to $P_{1,1}^{[d]}$, in order to mimic the proof technique of [IK20].

First, we determine the stabilizer of $P_{1,1}^{[d]}$ in GL_{d+1} that is $\mathbb{Z}_d \times (\mathbb{T}^{\mathrm{SL}_d} \rtimes \mathfrak{S}_d)$, see Theorem 5.2. 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 that 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 5.b. 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. For $\lambda \vdash dD$ we obtain the following identity:

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

where $a_\mu(d,\delta)$ is the *plethysm coefficients*, i.e., the multiplicity of μ in $\mathrm{Sym}^d(\mathrm{Sym}^\delta(V))$, see Proposition 5.4. We implement this formula on a computer and indeed find an abundance of multiplicity obstructions against generic polynomials, see appendix A. We use this data, and apply the [IK20] approach to lower bounds on $\mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_{d+1} P_{0,d}^{[d]}}])$, to find a sequence of partitions where $\mathrm{mult}_\lambda(\mathbb{C}[\mathrm{GL}_{d+1} P_{1,1}^{[d]}]) < \mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_{d+1} P_{0,d}^{[d]}}])$, see Theorem 5.11. This implies $\overline{\mathrm{GL}_{d+1} P_{1,1}^{[d]}} \not\subseteq \overline{\mathrm{GL}_{d+1} P_{0,d}^{[d]}}$ for $d \geq 3$; note that for $d = 2$, the fact that $\overline{\mathrm{GL}_3 P^{[2]}} \subseteq \overline{\mathrm{GL}_3 P_{0,2}^{[2]}}$ follows from $a^2 - b^2 = (a+b)(a-b)$.

Proof idea of Theorem 1.5: Homogeneous VF**.** Generalizing Kumar's complexity (1.※) to the setting of 3×3 matrices, one obtains 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 [BC92], but the version we get uses *homogeneous* projections.

Let $D_{n,d}$ be 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 (1.※) clear. We homogenize Ben-Or & Cleve and get Corollary 6.6, i.e., the VF-completeness of $D_{n,d}$. Here we have to pay close attention on how to deal with field constants, and we define the notion of input-homogeneous-linear computation, see §6.b. In particular, we prove an input-homogeneous-linear version of Brent's depth reduction, see Lemma 6.2. In order to simplify our new complete polynomial further, we then turn to 2×2 matrices. Note that (for odd d) $\mathsf{C}_{n,d}$ is the homogeneous degree d part of the $(1,2)$ entry of

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

Theorem 6.12 is based on the construction of [BIZ18], which is, however inherently *affine*. In the new homogeneous model, we convert the product gate into an arity 3 homogeneous product gate. The resulting analysis of arithmetic circuits and formulas allowing only arity 3 homogeneous product gates is surprisingly subtle. The collection $C_{n,d}$ can be seen as a homogeneous variant of the continuant in [BIZ18]. This gives the $\mathsf{V3F}$-$\overline{\mathrm{p}}$-hardness of $\mathsf{C}_{n,d}$ (i.e., $\mathsf{V3F}$-hardness under homogeneous border p-projections), where $\mathsf{V3F}$ is the class of p-families with polynomially bounded formulas over the arity 3 basis (which is a subclass of VF).

The next task in the proof of Theorem 6.12 is to translate this result to the standard basis. To see the VQP-$\overline{\mathrm{qp}}$-hardness (for definition, see Definition 6.1), 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 VF restricted to homogeneous families lies in $\mathsf{V3P}$ (the circuit analog of $\mathsf{V3F}$), see Theorem 6.15, 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 $\mathsf{V3F}$ has polynomially sized formulas, we conclude our proof by showing that $\mathsf{VQ3F} = \mathsf{VQ3P}$, for details see (6.7) and Theorem 6.16. We use an arity-3 basis variant of the Valiant-Skyum-Berkowitz-Rackoff circuit depth reduction [VSBR83], which is a bit more subtle than the original proof.

2 De-bordering border Waring rank

The goal of this section is to prove de-bordering results for border rank. In other words, given a homogeneous polynomial f, we provide upper bounds for $\mathsf{WR}(f)$ in terms of $\underline{\mathsf{WR}}(f)$ and d.

2.a Orbit closure and essential variables

The number of essential variables of a homogeneous polynomial f is the minimum integer m such that there is a linear change of coordinates after which f can be written as a polynomial in m variables. Denote the number of essential variables of f by $\mathrm{N}_{\mathrm{ess}}(f)$. It is a classical fact, which already appears in [Syl52], that the number of essential variables of f equals the dimension of the linear span of its first order partial derivatives, or equivalently the rank of the first partial derivative map. In particular $\mathrm{N}_{\mathrm{ess}}(-)$ is a lower semicontinuous function. We refer to [Car06] and [KS07, Lemma B.1] for modern proofs of this result.

We prove a structural result for orbit-closures of polynomials with non-maximal number of essential variables.

2.1 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*

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

Proof. Clearly
$$\mathrm{GL}(V)\overline{(\mathrm{GL}(W)\cdot f)} \subseteq \overline{\mathrm{GL}(V)\cdot p}.$$
To show the other inclusion, let $q \in \overline{\mathrm{GL}(V)\cdot p}$, and suppose $q = \lim_{\epsilon\to 0} g_\epsilon q$ where g_ϵ is a curve in $\mathrm{GL}(V)$. Without loss of generality, suppose $g_1 = \mathrm{id}_V$. By definition, the number of essential variables is invariant under the action of $\mathrm{GL}(V)$. In particular, since $\mathrm{N}_{\mathrm{ess}}(p) \le r$, the same holds for every element of $\mathrm{GL}(V)\cdot p$. Since $\mathrm{N}_{\mathrm{ess}}(-)$ is lower semicontinuous, we deduce $\mathrm{N}_{\mathrm{ess}}(q) \le r$ as well. This implies that there exists $h \in \mathrm{GL}(V)$ such that $h\cdot q \in S^d W$.

We are going to show $h\cdot q \in \overline{\mathrm{GL}(W)\cdot p}$. Let $\pi_W \in \mathrm{End}(V)$ be a projection onto W, that is a map $\pi_W : V \to V$ such that $\mathrm{im}(\pi_W) = W$ and $\pi_W|_W = \mathrm{id}_W$. Since the action of $\mathrm{GL}(V)$ is continuous, we have
$$hq = (\pi_W h)q = (\pi_W h)\lim_{\epsilon\to 0} g_\epsilon p = \lim_{\epsilon\to 0}(\pi_W h g_\epsilon)p.$$
Let $g'_\epsilon = \pi_W h g_\epsilon|_W$. Notice $g'_1 = \pi_W h g_1 = \mathrm{id}_W$. In particular $g'_\epsilon \in \mathrm{GL}(W)$ for generic ϵ. This shows $hq \in \overline{\mathrm{GL}(W)\cdot p}$. Hence $q = h^{-1}hq \in \mathrm{GL}(V)\cdot\overline{\mathrm{GL}(W)\cdot p}$. This concludes the proof. ☐

An immediate consequence of the semicontinuity of the number of essential variables is the following result.

2.2 Lemma. *Let $f \in S^d V$ be a homogeneous polynomial with $\underline{\mathrm{WR}}(f) \le r$. Then $\mathrm{N}_{\mathrm{ess}}(f) \le r$.*

Proof. After possibly re-embedding V is a space of larger dimension, assume $\dim V \ge r$. Then $\underline{\mathrm{WR}}(f) \le r$ implies $f \in \overline{\mathrm{GL}(V)\cdot(x_1^d + \cdots + x_r^d)}$. Since $\mathrm{N}_{\mathrm{ess}}(x_1^d + \cdots + x_r^d) = r$, we deduce $\mathrm{N}_{\mathrm{ess}}(f) \le r$. ☐

2.b Fixed-parameter de-bordering

The proof of Theorem 1.1 is based on generalized additive decompositions of polynomial, in the sense of [Iar95]. These decompositions were studied in algebraic geometry, usually in connection to 0-dimensional schemes and the notion of 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 0-dimensional schemes. 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 we see $\ell(\varepsilon) = \sum_{i=q}^{\infty} \varepsilon^i \ell_i$ as a 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.

2.3 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 $[\ell] \in \mathbb{PC}[\boldsymbol{x}]_1$ 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 $\gamma \in \mathbb{C}$.*

2.4 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 may 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 $\ell_1 = \varepsilon^{q_1} (\sum_{i=1}^n \alpha_i x_i)$ 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. ☐

2.5 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 may assume $\ell = x_1$. We prove the statement by induction on r and the difference $d - (r-1)$.

The cases $r = 1$ and $d = r - 1$ are trivial.

If $d > r - 1$, then by the previous Lemma there exists a standard local border decomposition

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

where $\ell_k = \sum_{i=1}^n \alpha_{ki} x_i$ 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

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

and

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

for $i > 1$. These decompositions involve the same linear forms ℓ_k with multiplicative coefficients and they are 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 $\ell^{d-r+1} g$.

2.6 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].

2.7 Lemma. *Let $\ell_1, \ldots, \ell_m \in \mathbb{C}[\boldsymbol{x}]_1$ be linear forms such that ℓ_i is not proportional to ℓ_j when $i \neq j$. Let $g_1, \ldots, g_m$ be homogeneous polynomials of degrees $r_1 - 1, \ldots, r_m - 1$ 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 $A \colon (y_1, y_2) \mapsto (x_1, \ldots, x_n)$ 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 $g_k = 0$. $\square$

2.8 Lemma. *Let $f \in S^d V$ be such that $\underline{\mathrm{WR}}(f) = r$. 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 2.5 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

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

From Lemma 2.7 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.

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

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

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

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 $\mathsf{WR}(f) \le \mathsf{WR}(g) \le \mathsf{maxR}(n, d_2)$. Since f is arbitrary, $\mathsf{maxR}(n, d_1) \le \mathsf{maxR}(n, d_2)$. □

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

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

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

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

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

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

Case $r \le d + 1$. By Lemma 2.8 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 $\underline{\mathsf{WR}}(\ell_k^{d-r_k+1} g_k) \le r_k$. Since $\underline{\mathsf{WR}}(\ell_k^{d-r_k+1} g_k) \le r_k$, the number of essential variables $N_{ess}(g_k) \le r_k$. If $r_k = 1$, then

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

If $r_k \ge 2$, 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[4]

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

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

Combining all parts of the decomposition together, we get

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

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

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

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

[4] it is easy to see that for $a \ge b$ 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 $a + 1$.

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 $\mathbb{C}[x_1,\dots,x_r]_{r-1}$ contains a direct sum of $x_1^q \cdot \mathbb{C}[x_1,\dots,x_p]_{p-1}$ and $x_1^{p+1} \cdot \mathbb{C}[x_{p+1},\dots,x_r]_{q-1}$. 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 $M(r) \leq \binom{2r-2}{r-1}$. $\square$

Using Blekherman-Teitler bound on the maximum rank [BT15], we can get a slightly better bound. The proof is essentially the same as for the previous theorem.

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

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

2.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 by Bernardi, Brachat and Mourrain in [BBM14] (they use a subtly different notion of generalized affine decomposition). In particular, there exists a much stronger version of Lemma 2.8, 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 the notions of cactus rank and size of a generalized additive decomposition later, for not let us state the theorem, which is based on [BBM14, Thm. 3.5].

2.14 Theorem. *If $\deg f \geq 2 \cdot \mathsf{CR}(f) - 1$, then 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.

2.15 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 explain how to measure the size of a generalized additive decomposition, review the basic notions of apolarity theory, define cactus rank and smoothable rank and explain the ideas behind the proof of Theorem 2.14 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 $D^\perp = \{f \in S_d \mid \alpha \cdot f = 0 \text{ for all } \alpha \in D\}$

Size of generalized additive decompositions. We now describe how we measure the size of a generalized additive decomposition.

2.16 Definition. *The* partial derivative space *of a polynomial $f \in S$ (not necessarily homogeneous) is the vector space $\partial^* f = T \cdot f$ spanned by f and all its partial derivatives of all orders.*

2.17 Definition. *Let ℓ be a linear form and let $\partial \in T_1$ be a partial derivative such that $\partial \ell = 1$. We define the* compression *$f_{(\partial,\ell)}$ of a homogeneous polynomial $f \in S_d$ with respect to ℓ and ∂ as follows. Write*

$$f = \sum_{i=0}^{d} \frac{\ell^i}{i!} f_i.$$

with $f_i \in \mathbb{C}[\partial^\perp]_{d-i}$. Then $f_{(\partial,\ell)} = \sum_{i=0}^{d} f_i$.

One can check that $\dim(\partial^* f_{(\partial,\ell)})$ does not depend on the choice of ∂ as long as $\partial \ell = 1$; this can be proved by hand, and it is obtained in [BJMR18] in a more intrinsic way.

2.18 Definition. *The* size *of a generalized additive decomposition*

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

is defined as $\sum_{k=1}^{m} \dim(\partial^* \overline{g}_k)$ where $\overline{g}_k = \left(\ell_k^{d-r_k+1} g_k\right)_{(\partial_k,\ell_k)}$ for some ∂_k such that $\partial_k \ell_k = 1$.

This way of measuring the size of generalized additive decompositions is compatible with the notion of cactus rank of a homogeneous polynomial, in the sense of Theorem 2.14.

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).

We will need the following property of ideals of Krull dimension 1.

2.19 Theorem ([IK99, Thm. 1.69])**.** *If I is a saturated ideal with $H_I = r$, then $h_I(p) = r$ for $p \geq r - 1$.*

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.

2.20 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 $\ell_1, \ldots, \ell_r$, consider the sequences of subspaces $E_p = \operatorname{Span}(\{\ell_1^p, \ldots, \ell_r^p\}) \subset S_p$ 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 $Z = \{[\ell_1], \ldots, [\ell_r]\}$ in the projective space $\mathbb{P}V$. Algebraically, the fact that I is a homogeneous ideal follows from the following useful proposition.

2.21 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 $\dim(T_p/I_p) = \dim T_p - \dim I_p = \dim I_p^\perp = \dim E_p$. $\square$

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

Proof. Let I be the defining ideal of Z let $E_p = I_p^\perp = \operatorname{Span}(\{\ell_1^p, \ldots, \ell_r^p\})$ 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 $I_p \subset \operatorname{Ann}(f)$. $\square$

2.23 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, 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 $E(\varepsilon) = \operatorname{Span}(\{w_1(\varepsilon), \ldots, w_r(\varepsilon)\})$ 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 $y_1, \ldots, y_q \in W^*(\varepsilon)$ 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 $E(\varepsilon) = \mathrm{Span}(\{w_1(\varepsilon), \dots, w_r(\varepsilon)\})$ 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 $E(\varepsilon) = \mathrm{Span}(\{y_1(\varepsilon), \dots, y_q(\varepsilon)\})^\perp$. 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 $\langle x_2, x_1^3 - x_1x_3^2 \rangle$ 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.

2.24 Theorem (Border apolarity, [BB21]). *$f \in S_d$ has $\underline{\mathsf{WR}}(f) \le r$ 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.

2.25 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 $\mathsf{CR}(f)$ 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 $\mathsf{SR}(f)$ corresponds to saturated limits of ideals of points. In addition, the border cactus rank $\underline{\mathsf{CR}}(f)$ is defined by considering limits of saturated ideals.

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

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

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

Proof. The inequality $\underline{\mathsf{WR}}(f) \le \mathsf{SR}(f)$ 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. $\square$

We will now prove several lemmas which connect generalized additive decompositions to apolar ideals, finishing the proof of Theorem 2.14.

2.27 Lemma ([BJMR18]). *Let ℓ be a linear form and let $f \in S_d$ be a homogeneous polynomial. Set $r = \dim \partial^* f_{[\ell]}$. There exists a homogeneous ideal I apolar to f with Hilbert polynomial $H_I = r$.*

Proof. Let $\partial \in T_1$ be such that $\partial \ell = 1$. Denote $S' = \mathbb{C}[\partial^\perp]$ and $T' = \mathbb{C}[\ell^\perp]$. The rings S' and T' are in the same dual relationship as S and T, and T is generated by T' and ∂.

We start from the ideal $J = \mathrm{Ann}(f_{(\partial,\ell)}) \subset T'$ and homogenize it using ∂. That is, define the homogenization map from $T'_{\le p} = \bigoplus_{j=0}^{p} T'_j$ to T_p sending $\alpha = \sum_{j=0}^{p} \alpha_j$ to $\sum_{j=0}^{p} \partial^{p-j}\alpha_j$. The homogeneous part I_p of the ideal I is then the image of $J_{\le p}$ under this homogenization map.

To show that the ideal I is apolar to f, write $f = \sum_{i=0}^{d} \frac{\ell^i}{i!} f_{d-i}$ with $f_i \in S'_i$. Then $f_{(\partial,\ell)} = \sum_{i=0}^{d} f_i$. If $\alpha' = \sum_{j=0}^{p} \alpha_j \in T'$ and $\alpha = \sum_{j=0}^{p} \partial^{p-j}\alpha_j$ is the α' homogenized, then the statement $\alpha' \cdot f_{(\partial,\ell)} = 0$ is equivalent to $\alpha \cdot f = 0$, since they are both equivalent to $\sum_{j=0}^{p} \alpha_j f_{j+e} = 0$ for all $e \ge 0$. Since J is apolar to $f_{(\partial,\ell)}$, I is apolar to f.

Since $f_{(\partial,\ell)}$ has degree at most d, J contains T'_p for $p > d$. Hence

$$\partial^* f_{(\partial,\ell)} \cong A(f_{(\partial,\ell)}) = T'/J \cong T'_{\leq d}/J_{\leq d}$$

as vector spaces, and for $p > d$ we have $T_p/I_p \cong T'_{\leq p}/J_{\leq p} \cong T'_{\leq d}/J_{\leq d}$. Therefore, for p large enough $\dim T_p/I_p = \dim \partial^* f_{(\partial,\ell)} = r$ and $H_I = r$. ☐

2.28 Lemma. *Suppose $f \in S_d$ is apolar to a saturated primary homogeneous ideal I with Hilbert polynomial $H_I = r$. If $d \geq 2r - 1$, then f has a one-summand generalized additive decomposition of size at most r.*

Proof. If I is an ideal of Krull dimension 1, then it defines a 0-dimensional scheme, and if it is primary, then this scheme is supported at one point $[\ell] \in \mathbb{P}V$. The ideal corresponding to this point is $J = \operatorname{Ann}(\ell) = \langle \ell^\perp \rangle$ and we have $J^m \subset I \subset J$ for some $m \leq r$. For the corresponding dual space E_p with $p \geq m$ we have $E_p \subset (J^m)^\perp_p = \{\ell^{p-m} g \mid g \in S_{\leq m}\}$. Since $f \in E_d$, it has a one-summand generalized additive decomposition $f = \ell^{d-m} g$.

Choose $\partial \in T_1$ such that $\partial \ell = 1$. Write $f = \sum_{i=0}^m \frac{\ell^{d-i}}{(d-i)!} f_i$ with $f_i \in \mathbb{C}[\partial^\perp]_i$. Then $f_{(\partial,\ell)} = \sum_{i=0}^m f_i$ has degree at most m. For every $\alpha' = \sum_{j=0}^m \in \mathbb{C}[\ell^\perp]_{\leq m}$ and the corresponding homogeneous $\alpha = \sum_{j=0}^m \partial^{m-j} \alpha_j$ we have

$$\alpha' \cdot f_{(\partial,\ell)} = \sum_{j \geq i} \alpha_j \cdot f_i$$

and

$$\alpha \cdot f = \sum_{j \leq i} \frac{\ell^{d-m+j-i}}{(d-m+j-i)!} \alpha_j \cdot f_i.$$

Therefore, there is an isomorphism between $\partial^* f_{(\partial,\ell)}$ and $T^m \cdot f \subset E_{d-m}$. Note that $d - m \geq r$. By Theorem 2.19 we have $r = H_I = \dim E_{d-m} \geq \dim \partial^* f_{(\partial,\ell)}$. ☐

Proof of Theorem 2.14. 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 $I = I^{(1)} \cap \cdots \cap I^{(m)}$, and for the Hilbert polynomials it is true that $H_I = H_{I^{(1)}} + \cdots + H_{I^{(m)}}$. Defining $E_d = I_d^\perp$ and $E_d^{(k)} = (I_d^{(k)})^\perp$ we have $E_d = E_d^{(1)} + \cdots + E_d^{(m)}$. Therefore, $f = f^{(1)} + \cdots + f^{(m)}$ where $f^{(k)} \in E_d^{(k)}$. By Lemma 2.28 each $f^{(k)}$ contributes one summand to the generalized additive decomposition. The sizes of these summands are bounded by $H_{I^{(k)}}$, and the total size is bounded by r.

Conversely, if f has a generalized additive decomposition of size r, then from each summand we can construct an ideal using Lemma 2.27 and take the intersection of these ideals to get an ideal apolar to f with Hilbert function at most r. ☐

2.d Classes of the form ΣFΣ

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 $\mathsf{VWaring}$. 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.

2.29 Proposition. *Let* $\mathbf{F}$ *be a p-family in constant number of variables. Then* $\Sigma\mathbf{F}\Sigma = \mathsf{VWaring}$.

Proof. Clearly $\mathsf{VWaring} \subseteq \Sigma\mathbf{F}\Sigma$ 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

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

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 $R(m)$; since $m_1, \dots, m_r$ are polynomial functions in n, as well as $r(n)$, we conclude. □

3 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. Let $e_k(x_1, \dots, x_n)$ denotes the k-th **elementary symmetric** polynomial, defined by

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

Recall that by definition $e_0 = 1$. First, we record an immediate observation that will be useful throughout:

3.1 Remark. It is easy to observe that

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

where $\mathbf{x} = (x_1, \dots, x_m)$. 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

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

To demystify $\underline{\mathsf{Kc}}$-complexity, we will often use *Newton identities*, see Proposition 3.2. Let $p_k(x_1, \dots, x_n)$ denote the k-th **power sum** polynomial, defined as $p_k(\boldsymbol{x}) := x_1^k + \cdots + x_n^k$.

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

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

In the light of Remark 3.1, the Kc model of computation is a sum of elementary symmetric polynomials. 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. 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. In fact, the only homogeneous polynomials with finite Kc-complexity are powers of linear forms, as the following lemma shows.

3.3 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 $f = \alpha\left(\prod_{i=1}^m (1 + \ell_i) - 1\right)$ for some homogeneous linear forms $\ell_i \in \mathbb{C}[\mathbf{x}]$. Write $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_m)$. By Remark 3.1, 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 $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_d)$ 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 $\widehat{\boldsymbol{\ell}} = (\ell_1, \ldots, \ell_{d-1})$. We use induction on j to prove that $e_j(\widehat{\boldsymbol{\ell}}) = (-1)^j \cdot \ell_d^j$ for $j = 1, \ldots, d-1$. 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 $j = 2, \ldots, d-1$, 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 $f = \alpha e_d(\boldsymbol{\ell}) = \alpha \ell_d \cdot (-1)^{d-1} \cdot e_{d-1}(\widehat{\boldsymbol{\ell}}) = -\alpha \ell_d^d$, which concludes the proof. □

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

3.4 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 $\underline{\mathsf{Kc}}(f) \leq rd$. □

In fact, a slightly stronger statement is true:

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

Proof. Analogously to the proof in Proposition 3.4, $f \simeq \sum_{i=1}^r \ell_i^d = -e_d(-\zeta^0 \ell_1, \ldots, -\zeta^{d-1} \ell_r)$. 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 $d < i \leq dr$ 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 $\underline{\mathsf{Kc}}(f) \leq rd$. □

Proposition 3.4 and Proposition 3.5 show that if $\underline{\mathrm{WR}}(f)$ is small then $\underline{\mathrm{Kc}}(f)$ is small. However, there are polynomials with large Waring (border) rank but small Kumar complexity, such as products of linear forms. Notice $\underline{\mathrm{WR}}(x_1 \cdots x_n) \geq \exp(n)$, which can be easily shown by partial derivative methods, see e.g. [LT10, Sec. 11], [CKW11b, Thm. 10.4].

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

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

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.

3.7 Theorem. *For all homogeneous f we have*

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

Proof. The right inequality follows from Proposition 3.5 and Lemma 3.6. The left inequality is a combination of Lemma 3.9, Proposition 3.10, and Theorem 3.11 below. ☐

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

Proof. By Theorem 3.7, if $\underline{\mathrm{Kc}}(f) = m$ then $\min\{\deg(f)\cdot\delta_f,\ \underline{\mathrm{WR}}(f)\} \leq m$. Now, if $\deg(f)\cdot\delta_f \leq \underline{\mathrm{WR}}(f)$, then the minimum is $\deg(f)\cdot\delta_f$, which implies $\delta_f \neq \infty$; in this case $\delta_f = 1$, so f is a product of linear forms. Otherwise, $\underline{\mathrm{WR}}(f)$ is the minimum, which implies that $\underline{\mathrm{WR}}(f) \leq m$. ☐

Note that in the definition of $\underline{\mathrm{Kc}}$, 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.

- $\underline{\mathrm{Kc}}^+(f)$ 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 $\underline{\mathrm{Kc}}^+(f) = \infty$ is such an m does not exist;
- $\underline{\mathrm{Kc}}^-(f)$ 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 $\underline{\mathrm{Kc}}^-(f) = \infty$ is such an m does not exist;
- $\underline{\mathrm{Kc}}^=(f)$ 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 $\underline{\mathrm{Kc}}^=(f) = \infty$ is such an m does not exist.

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

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

Proof. Let $d := \deg(f)$. The lower bound $\deg(f) \leq \underline{\mathrm{Kc}}^+(f)$ is clear. Therefore, it suffices to show that if $\underline{\mathrm{Kc}}^+(f)$ 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. ☐

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

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 $1 \le i < d$, where $\boldsymbol{\ell} = (\ell_1, \dots, \ell_m)$. 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 $\mathsf{WR}(f) \le m$. □

3.11 Theorem (Border Newton Identities)**.** *For all homogeneous f:* $\underline{\mathsf{WR}}(f) \le \underline{\mathsf{Kc}}^{-}(f)$*.*

Proof. Let $d := \deg(f)$. Let $f \simeq f_\epsilon := \gamma\epsilon^{-M}\big(\prod_{i=1}^{m}(1+\ell_i') - 1\big)$ with $M \ge 1$. 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 $0 \le j < d$ we have $f_{\epsilon,j} \simeq 0$. By expanding the product, observe that for all $0 < j < d$ we have $0 \simeq f_{\epsilon,j} = \gamma\epsilon^{-M} e_j(\epsilon\ell_1, \dots, \epsilon\ell_m) = \gamma\epsilon^{-M+j} e_j(\ell_1, \dots, \ell_m)$. We now show by induction that for all $1 \le j < d$ we have $\epsilon^{-M+j} p_j(\ell_1, \dots, \ell_m) \simeq 0$. 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 $f \simeq f_{\epsilon,d} = \gamma\epsilon^{-M+d} e_d(\ell_1, \dots, \ell_m) \simeq \gamma\epsilon^{-M+d} \cdot \frac{1}{d} \cdot (-1)^{d-1} p_d(\ell_1, \dots, \ell_m)$ and hence $\underline{\mathsf{WR}}(f) \le m$. □

3.a Linear approximations and Waring rank

We demonstrated the inequality $\underline{\mathsf{Kc}}(f) \le \deg(f) \cdot \mathsf{WR}(f)$ in Proposition 3.4. In the proof of Proposition 3.4, only "linear approximations" have been used; we prove here a converse of Proposition 3.4 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 $\ell_1, \dots, \ell_m \in \mathbb{C}[\mathbf{x}]_1$ and $M \ge 1$ such that $f \simeq \gamma\epsilon^{-M}\big(\prod_{i=1}^{m}(1+\epsilon\ell_i) - 1\big)$.

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

Proof. The inequality $\underline{\mathsf{Kc}}_1^-(f) \leq d \cdot \mathsf{WR}(f)$ is clear from the proof of Proposition 3.4, 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

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

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

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

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

4 Restricted binomials: De-bordering and lower bounds

In this section, we study restricted binomials. A binomial bn_d is the polynomial $\mathrm{bn}_d(\boldsymbol{x}, \boldsymbol{y}) := P_{2,0}^{[d]} = x_1 \ldots x_d + y_1 \cdots y_d$. In Section 4.a, we de-border product-plus-power ($P_{1,1}^{[d]}$) and product-plus-two-powers models ($P_{1,2}^{[d]}$). In Section 4.b, we show exponential gaps between product-plus-power, product-plus-two-powers, and binomials (in the affine sense). Identifying explicit polynomials which are hard to approximate, and proving it remains a major template in algebraic and geometric complexity theory. 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. However, in the restricted setting, we are able to show *optimal* lower bounds, see Theorem 4.16 & Theorem 4.17.

4.a De-bordering: Characterizing special binomials

In this section we prove de-bordering results for product-plus-power and product-plus-two-powers models. Our method applies also for more general computational model based on restricted binomials. More specifically, we prove that polynomials obtained in the limit in our model have low border Waring rank. One can then apply a debordering result for $\underline{\mathsf{WR}}$ such as $\mathsf{abpw}(f) \leq \underline{\mathsf{WR}}(f)$ [BDI21, For14] or the results from Section 2.b to get a complete debordering.

4.1 Definition (Restricted binomial model)**.** *We say that a homogeneous polynomial $f \in S^d V$ is in the class RB_k if it can be presented as*

$$f = \prod_{i=1}^d \ell_i + \prod_{i=1}^d \ell_i'$$

for some linear forms $\ell_i, \ell_i' \in V$ such that $\mathrm{rank}(\ell_1', \ldots, \ell_d') \leq k$. We also define the corresponding approximate class $\overline{RB}_k$ in the standard way: a homogeneous polynomial $f \in S^d V$ is in $\overline{RB}_k$ if

$$f = \lim_{\varepsilon \to 0} \left(\prod_{i=1}^d \ell_i(\epsilon) + \prod_{i=1}^d \ell_i'(\epsilon) \right) \tag{4.2}$$

for some $\ell_i(\epsilon), \ell_i'(\epsilon) \in \mathbb{C}[\epsilon^{\pm 1}][\mathbf{x}]_1$ such that $\mathrm{rank}(\ell_1'(\epsilon), \ldots, \ell_d'(\epsilon)) \leq k$ for every $\epsilon \neq 0$.

4.3 Theorem (De-bordering $\overline{RB}_k$)**.** *Let $f \in S^dV$ be a polynomial in $\overline{RB}_k$. Then either $f \in RB_k$, or $\underline{\mathsf{WR}}(f) \leq O(d^{3k+2})$.*

To prove this theorem, we first need some basic lemmas which will be used in the proof. We will use non-homogeneous polynomials, so instead of Waring rank we will be working with the complexity of $\Sigma\Lambda\Sigma$-circuits. Denote by $\Sigma^{[s]}\Lambda^{[e]}\Sigma$ the class of (non-homogeneous) polynomials representable as a sum of s powers of affine linear forms with exponents not exceeding e, and by $\overline{\Sigma^{[s]}\Lambda^{[e]}\Sigma}$ the corresponding class closed under approximation. As the following lemma shows, for homogeneous polynomials this model is equal in power to border Waring rank.

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

Proof. By assumption $f \simeq \sum_{i\in[s]}(\alpha_i + \ell_i)^{e_i}$, where $\alpha_i \in \mathbb{C}(\epsilon)$, and $\ell_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]_1$. Taking the degree d part of each side, we obtain a border Waring rank decomposition $f \simeq \sum_{i\colon e_i\geq d}\binom{e_i}{d}\ell_i^d\alpha_i^{e_i-d}$ with at most s summands. □

We recall a classical result on the border Waring rank of a binary monomial.

4.5 Proposition (see e. g. [LT10])**.** *If $a \leq b$, then $\underline{\mathsf{WR}}(x^a y^b) = a + 1$.*

The next lemma bounds the $\Sigma\Lambda\Sigma$ complexity of a polynomial in terms of the complexity of polynomials obtained from it by substitution of variables.

4.6 Lemma (Interpolation)**.** *Let $f(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]$ be a polynomial of degree d such that $f(\gamma_i, x_2, \ldots, x_n) \in \overline{\Sigma^{[s]}\wedge^{[e]}\Sigma}$ for some distinct $\gamma_0, \ldots, \gamma_d \in \mathbb{C}$. Then $f(\boldsymbol{x}) \in \overline{\Sigma^{[s(d+1)^3]}\wedge^{[e+d]}\Sigma}$.*

Proof. Write $f(\boldsymbol{x}) = \sum_{j=0}^d x_1^j f_j(x_2, \ldots, x_n)$. By polynomial interpolation there exist $\alpha_{ij} \in \mathbb{C}$ such that $f_j = \sum_{i=0}^d \alpha_{ij} f(\gamma_i, x_2, \ldots, x_n)$. By assumption, $f(\gamma_i, x_2, \ldots, x_n) \simeq \sum_{j=1}^s \ell_{ij}^{e_j}$, where ℓ_{ij} are affine linear forms with coefficients in $\mathbb{C}[\epsilon^{\pm 1}]$, and $e_j \leq e$. Hence

$$f_j(\boldsymbol{x}) \simeq \sum_{i=0}^d \sum_{j=1}^s \alpha_{ij}\ell_{ij}^{e_j} \implies f_j(\boldsymbol{x}) \in \overline{\Sigma^{[s(d+1)]}\wedge^{[e]}\Sigma}\,.$$

Note that for any affine linear polynomial ℓ the polynomial $x_1^j\ell^e$ can be approximated by a $\Sigma^{[d+1]}\wedge^{[e+j]}\Sigma$-circuit using the decomposition of the monomial $x^j y^e$ with border Waring rank equal to $\min\{j+1, e+1\} \leq j+1 \leq d+1$; this follows from Proposition 4.5. Therefore $x_1^j f_j \in \overline{\Sigma^{[s(d+1)^2]}\wedge^{[e+j]}\Sigma}$, and $f(\boldsymbol{x}) = \sum_{i=0}^d x_1^j f_j \in \overline{\Sigma^{[s(d+1)^3]}\wedge^{[e+d]}\Sigma}$. □

By applying the previous Lemma several times we get the following.

4.7 Corollary. *Let $f(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]$ be a polynomial of degree d such that*

$$f(\gamma_{1i_1}, \gamma_{2i_2}, \ldots, \gamma_{ki_k}, x_{k+1}, \ldots, x_n) \in \overline{\Sigma^{[s]}\wedge^{[e]}\Sigma}$$

for some $\gamma_{ij} \in \mathbb{C}$, $1 \leq i \leq k$, $0 \leq j \leq d$, with $\gamma_{i0}, \ldots, \gamma_{id}$ distinct for each i. Then $f \in \overline{\Sigma^{[s(d+1)^{3k}]}\wedge^{[e+kd]}\Sigma}$.

Additionally, we need the following statement similar to Theorem 3.11 which considers an auxiliary Kumar-like model.

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

Proof. Let $f_\epsilon = \epsilon^{-M}\big(\prod_{i=1}^m(1+\epsilon a_i) - \prod_{i=1}^m(1+\epsilon b_i)\big)$. Denote by f_j and $f_{\epsilon,j}$ the homogeneous degree j parts of f and f_ϵ respectively. Since $f \simeq f_\epsilon$, we have

$$f_j \simeq f_{\epsilon,j} = \epsilon^{-M}\left(e_j(\epsilon a_1,\ldots,\epsilon a_m) - e_j(\epsilon b_1,\ldots,\epsilon b_m)\right) = \epsilon^{-M+j}\left(e_j(\boldsymbol{a}) - e_j(\boldsymbol{b})\right),$$

where $\boldsymbol{a} = (a_1,\ldots,a_m)$ and similarly $\boldsymbol{b} = (b_1,\ldots,b_m)$. Note that since $f_{\epsilon,j}$ converges, $e_j(\boldsymbol{a}) - e_j(\boldsymbol{b})$ is divisible by ϵ^{M-j} for all $j \geq 1$, that is,

$$e_j(\boldsymbol{a}) \equiv e_j(\boldsymbol{b}) \bmod \langle \epsilon^{M-j}\rangle$$

where we consider $e_j(\boldsymbol{a})$ and $e_j(\boldsymbol{b})$ as elements of the ring $\mathbb{C}((\epsilon))[\mathbf{x}]$.

We now show by induction that for all $j \geq 1$ the following additional congruences hold.

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

The case $j = 1$ is trivially true because $p_1 = e_1$. For the induction step from j to $j+1$, we use Newton's identities

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

We obtain

$$\begin{aligned} p_{j+1}(\boldsymbol{a}) - p_{j+1}(\boldsymbol{b}) =& (-1)^j(j+1)\left(e_{j+1}(\boldsymbol{a}) - e_{j+1}(\boldsymbol{b})\right) \\ &+ \sum_{i=1}^{j}(-1)^{j+i}\left(e_{j+1-i}(\boldsymbol{a})\cdot p_i(\boldsymbol{a}) - e_{j+1-i}(\boldsymbol{b})\cdot p_i(\boldsymbol{b})\right). \end{aligned} \tag{4.9}$$

By induction hypothesis we know that for $1 \leq i \leq j$

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

Since $M - j \leq M - i$ and $M - j \leq M - (j+1) + i$, this can be relaxed to

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

From (4.9) we get

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

Weakening this to an equivalence mod $\langle \epsilon^{M-(j+1)}\rangle$, we obtain

$$p_{j+1}(\boldsymbol{a}) - p_{j+1}(\boldsymbol{b}) \equiv (-1)^j(j+1)\left(e_{j+1}(\boldsymbol{a}) - e_{j+1}(\boldsymbol{b})\right) \equiv 0 \bmod \langle \epsilon^{M-(j+1)}\rangle,$$

or $p_{j+1}(\boldsymbol{a}) \equiv p_{j+1}(\boldsymbol{b}) \bmod \langle \epsilon^{M-(j+1)}\rangle$, finishing the induction.

Finally, we use the proved congruences to write an approximate decomposition of f. We have

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

which shows that $\underline{\mathrm{WR}}(f_j) \leq 2m$. Note that $f_0 = 0$, so $f = \sum_{j=1}^d f_j \in \overline{\Sigma^{[2md]}\Lambda^{[d]}\Sigma}$. $\square$

4.10 Corollary. *For any degree d polynomial $f(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]$, not necessarily homogeneous, suppose we have $f \simeq \epsilon^{-M}\big(\alpha \prod_{i=1}^{m}(1+\epsilon a_i) - \beta \prod_{i=1}^{m}(1+\epsilon b_i)\big)$ with $M \geq 1$ for some $a_i, b_i \in \mathbb{C}[[\epsilon]][\boldsymbol{x}]_1$ and $\alpha, \beta \in \mathbb{C}[\epsilon]$ such that $\alpha \simeq \beta \not\simeq 0$. Then $f \in \overline{\Sigma^{[2md+1]} \wedge^{[d]} \Sigma}$.*

Proof. Let f_j and $f_{\epsilon,j}$ be the homogeneus parts as in the proof of the preceding Theorem. Additionally, Let $\alpha_0 = \lim_{\epsilon\to 0} \alpha$ and $\gamma = \frac{\beta}{\alpha}$. From assumptions of the theorem, $\alpha_0 \neq 0$ and $\gamma \simeq 1$. We have

$$\frac{1}{\alpha_0} f \simeq \frac{1}{\alpha} f \simeq \epsilon^{-M}\big(\prod_{i=1}^{m}(1+\epsilon a_i) - \gamma \prod_{i=1}^{m}(1+\epsilon b_i)\big)$$

By taking degree 0 part we get $\frac{1}{\alpha_0} f_0 \simeq \frac{1}{\alpha_0} f_{\epsilon,0} = \epsilon^{-M}(1-\gamma)$, so

$$\frac{1}{\alpha_0} f_j \simeq \epsilon^{-M+j}(e_j(\boldsymbol{a}) - \gamma e_j(\boldsymbol{b})) = \epsilon^{-M+j}(e_j(\boldsymbol{a}) - e_j(\boldsymbol{b})) + \epsilon^j \frac{f_{\epsilon,0}}{\alpha_0} e_j(\boldsymbol{b}) \simeq \epsilon^{-M+j}(e_j(\boldsymbol{a}) - e_j(\boldsymbol{b})),$$

hence

$$f \simeq f_0 + \alpha_0 \epsilon^{-M}\big(\prod_{i=1}^{m}(1+\epsilon a_i) - \prod_{i=1}^{m}(1+\epsilon b_i)\big),$$

and we reduce to the case considered in Theorem 4.8. □

We are now ready to prove Theorem 4.3.

Proof of Theorem 4.3. Since $f \in \overline{BR}_k$, it has an approximate decomposition (4.2), which we rewrite as

$$f \simeq \epsilon^p \prod_{i=1}^{d} \ell_i - \epsilon^{p'} \prod_{i=1}^{d} \ell'_i$$

where $\ell_i, \ell'_i \in \mathbb{C}[\epsilon][\mathbf{x}]_1$ are not divisible by ϵ and $\operatorname{rank}(\ell'_1, \dots \ell'_d) \leq k$ at any $\epsilon \neq 0$. Define $\ell_{i0} \in V$ as $\ell_{i0} = \ell_i|_{\epsilon=0}$ and similarly $\ell'_{i0} = \ell'_i|_{\epsilon=0}$. ℓ_{i0} and ℓ'_{i0} are nonzero and by semicontinuity of rank we have $\operatorname{rank}(\ell'_{10}, \dots \ell'_{d0}) \leq k$.

If $p = p' = 0$, then $f = \prod_{i=0}^{d} \ell_{i0} - \prod_{i=0}^{d} \ell'_{i0}$. Similarly, if one of the exponents p and p' is positive, then the corresponding summand tends to 0 as $\epsilon \to 0$, and f is a product of linear forms, and if both p and p' are positive, then $f = 0$. In all these cases we have $f \in RB_k$.

Consider now the case when there are negative exponents. The convergence of the right hand side of the decomposition implies that $p = p'$ and the lowest degree term $\prod_{i=0}^{d} \ell_{i0} - \prod_{i=0}^{d} \ell'_{i0}$ is zero. By unique factorization the sets of linear forms ℓ_{i0} and ℓ'_{i0} are the same up to scalar multiples, and we can permute and rescale the factors in one of the products so that $\ell_{i0} = \ell'_{i0}$. Additionally we can assume that $\ell_{10}, \dots, \ell_{r0}$ are linearly independent, where $r = \operatorname{rank}(\ell_{10}, \dots \ell_{d0}) \leq k$.

Since ℓ_{i0} for $i \leq r$ are linearly independent, there exists an invertible linear map A such that $\ell_{i0}(A\mathbf{x}) = x_i$ for $i \leq r$. The linear forms ℓ_{i0} lie in the linear span of the first r of them, which means that $\ell_{i0}(A\mathbf{x}) \in \mathbb{C}[x_1, \dots, x_r]_1$ for all i.

Let $M = -p$, $L_i(\mathbf{x}) = \ell_i(A\mathbf{x})$ and $L'_i(\mathbf{x}) = \ell'_i(A\mathbf{x})$. For the polynomial $g(\mathbf{x}) = f(A\mathbf{x})$ we obtain an approximate decomposition

$$g \simeq \epsilon^{-M}\big(\prod_{i=1}^{d} L_i - \prod_{i=1}^{d} L'_i\big)$$

where $L_i, L'_i \in \mathbb{C}[\epsilon][\mathbf{x}]_1$ are such that $L_{i0} := L_i|_{\epsilon=0} = L'_i|_{\epsilon=0}$ are nonzero elements of $\mathbb{C}[x_1, \dots, x_r]$.

Choose $\gamma_{ij} \in \mathbb{C}$ for $1 \le i \le r$, $0 \le j \le d$ so that $\gamma_{i0}, \dots, \gamma_{id}$ are distinct for each i and $L_{i0}(\gamma_{1j_1}, \dots, \gamma_{rj_r}) \neq 0$ for all $i, j_1, \dots, j_r$. The choice is possible because L_{k0} are nonzero and hence the set of tuples γ not satisfying the required conditions is a nontrivial Zariski closed set. Write

$$L_i(\gamma_{1j_1}, \dots, \gamma_{rj_r}, x_{r+1}, \dots, x_r) = \alpha_i + \epsilon A_i(x_{r+1}, \dots, x_n)$$
$$L_i'(\gamma_{1j_1}, \dots, \gamma_{rj_r}, x_{r+1}, \dots, x_r) = \beta_i + \epsilon B_i(x_{r+1}, \dots, x_n)$$

with $\alpha_i, \beta_i \in \mathbb{C}[\epsilon]$ and $A_i, B_i \in \mathbb{C}[\epsilon][x_{r+1}, \dots, x_n]$, and set $\alpha = \prod_{i=1}^d \alpha_i$, $\beta = \prod_{i=1}^d \beta_i$, $a_i = \frac{A_i}{\alpha_i}$, $b_i = \frac{B_i}{\beta_i}$. Because $\alpha_i|_{\epsilon=0} = L_{i0}(\gamma_{1j_1}, \dots, \gamma_{rj_r}) \neq 0$, a_i are well defined in the ring $\mathbb{C}[[\epsilon]][\mathbf{x}]$; ditto for b_i. We obtain

$$g(\gamma_{1j_1}, \dots, \gamma_{rj_r}, x_{r+1}, \dots, x_n) = \epsilon^{-M}\big(\alpha \prod_{i=1}^d (1 + \epsilon a_i) - \beta \prod_{i=1}^d (1 + \epsilon b_i)\big).$$

By Corollary 4.10 $g(\gamma_{1j_1}, \dots, \gamma_{rj_r}, \mathbf{x}) \in \overline{\Sigma^{[2d^2+1]}\Lambda^{[d]}\Sigma}$. By Lemma 4.6 $g \in \overline{\Sigma^{[(2d^2+1)(d+1)^{3r}]}\Lambda^{[(r+1)d]}\Sigma}$, and by Lemma 4.4 $\underline{\mathrm{WR}}(g) \le (2d^2+1)(d+1)^{3r} = O(d^{3k+2})$. Since border Waring rank is invariant under invertible linear transformations, the same is true for f. $\square$

As special cases we obtain the following results for product-plus-power and product-plus-two powers. Note that RB_1 consists of polynomials of the form $\prod_{i=1}^d \ell_i + \ell_1'^d$, which are exactly the restrictions of $P_{1,1}^{[d]}$. Similarly, $f \in \overline{RB}_1$ if and only if $f \trianglelefteq P_{1,1}^{[d]}$. As a corollary of Theorem 4.3 we obtain the following statement.

4.11 Theorem (De-bordering product-plus-power)**.** *Let $f \in S^dV$ such that $f \trianglelefteq P_{1,1}^{[d]}$. Then either $f \le P_{1,1}^{[d]}$, or $\underline{\mathrm{WR}}(f) = O(d^5)$.*

The result for the product-plus-two-powers follows for the analysis of $\overline{RB}_2$, since a sum of two powers $a^d - b^d$ can be represented as a product $\prod_{i=0}^{d-1}(a - \omega^i b)$ of linear forms spanning a 2-dimensional subspace (here ω is a primitive d-th root of unity). More careful analysis gives the following theorem.

4.12 Theorem (De-bordering product-plus-two-powers)**.** *Let $f \in S^dV$ such that $f \trianglelefteq P_{1,2}^{[d]}$. One of the three alternatives is true:*

1. $f \le P_{1,2}^{[d]}$, *or*
2. $f \le \prod_{i=1}^d y_i + y_0^{d-1} \cdot y_{d+1}$, *or*
3. $\underline{\mathrm{WR}}(f) = O(d^8)$.

Proof. The proof follows the proof of Theorem 4.3. Additional step is required in the case when both summands have individual limits. In this case, the limit of the restricted summand is a polynomial of border Waring rank 2. It is known [LT10] (see also Theorem B.1) that it either has border rank two, in which case the alternative (1) holds for f, or has the form $a^{d-1}b$, which implies alternative (2). $\square$

4.b Lower Bounds

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

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

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. $\square$

4.14 Lemma. *The polynomial* $P_{2,0}^{[d]} = \prod_{i=1}^d x_i + \prod_{i=d+1}^{2d} x_i$ *cannot be written as a product of linear forms.*

Proof. It easily follows from a proof similar to that of Lemma 4.13. $\square$

4.15 Lemma. *For the polynomial* $P_{1,2}^{[d]} = \prod_{i\in[d]} x_i + x_{d+1}^d + x_{d+2}^d$*, we have* $\underline{\mathsf{WR}}(f) \geq 2^{\Omega(d)}$.

Proof. Evaluating $x_{d+1} = x_{d+2} = 0$, we obtain

$$\underline{\mathsf{WR}}(f) \geq \underline{\mathsf{WR}}(x_1 \cdots x_d) \geq \binom{d}{\lceil d/2 \rceil}$$

where the second inequality follows computing the dimension of the space of partial derivatives of order $\lfloor d/2 \rfloor$, see, e.g., [LT10, CKW11b]. $\square$

4.16 Theorem (First exp. gap theorem)**.** *If* $P_{1,2}^{[d]} \trianglelefteq_{\text{aff}} P_{1,1}^{[e]}$*, then* $e \geq \exp(d)$.

We remark that by Kumar's result [Kum20], we know that there exists $e \leq \exp(d)$, such that $P_{1,2}^{[d]} \trianglelefteq_{\text{aff}} P_{1,1}^{[e]}$. Therefore, Theorem 4.16 is *optimal*.

Proof of Theorem 4.16. Let $P_{1,2}^{[d]} \trianglelefteq_{\text{aff}} P_{1,1}^{[e]}$. That means that there are affine linear forms $L_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]$ such that $\prod_{i\in[d]} + x_{d+1}^d + x_{d+2}^d + \epsilon \cdot S(\boldsymbol{x}, \epsilon) = \prod_{i\in[e]} L_i + L_{e+1}^d$. By substituting, $x_i \mapsto x_i/x_0$, and multiplying both sides by x_0^e, we get that $x_0^{e-d} \cdot P_{1,2}^{[d]} + \epsilon \cdot \widehat{S} = \prod_{i\in[e]} \widehat{L}_i + \widehat{L}_{e+1}^e$, for homogeneous linear forms $\widehat{L}_i$, or, equivalently, $x_0^{e-d} \cdot P_{1,2}^{[d]} \trianglelefteq P_{1,1}^{[e]}$.

By Theorem 4.11, we know that $x_0^{e-d} \cdot P_{1,2}^{[d]} \trianglelefteq P_{1,1}^{[e]}$ implies either (i) $x_0^{e-d} \cdot P_{1,2}^{[d]} = \prod_{i\in[e]} \ell_i + \ell_0^e$, for some linear forms $\ell_i \in \mathbb{C}[\boldsymbol{x}]$, or (ii) $\underline{\mathsf{WR}}(x_0^{e-d} \cdot P_{1,2}^{[d]}) = O(e^5)$. 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 $P_{1,2}^{[d]} + \epsilon g = \sum_{i\in[k]} \ell_i^e$ for some affine forms $\widehat{\ell}_i \in \mathbb{C}(\epsilon)[\boldsymbol{x}]$ and $g \in \mathbb{C}[\epsilon][\boldsymbol{x}]$ with $k \in O(e^5)$. Since $P_{1,2}^{[d]}$ is homogeneous, this also implies that $\underline{\mathsf{WR}}(P_{1,2}^{[d]}) \leq k$. But then Lemma 4.15 implies that $k \geq 2^{\Omega(d)}$, which in turn implies that $e \geq 2^{\Omega(d)}$.

Proof of Part (i): Let $x_0^{e-d} \cdot P_{1,2}^{[d]} = \prod_{i\in[e]} \ell_i + \ell_0^e$. Note that, by a simple derivative space argument, one can show that the number of essential variables 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

$$P_{1,2}^{[d]} \;=\; \prod_{i\in[d]} \widehat{\ell}_i + c x_0^d \,,$$

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

$$P_{1,2}^{[d]} \;=\; \prod_{i \in [d]} \tilde{\ell}_i \;,$$

where $\mathbb{C}[x_1, \ldots, x_{d+2}] \ni \tilde{\ell}_i = \widehat{\ell}_i|_{x_0=0} \neq 0$. From Lemma 4.13, 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:

$$(-\widehat{\ell}_0/c_0)^{e-d} \cdot P_{1,2}^{[d]} \;=\; \prod_{i \in [e]} \widehat{\ell}_i \;,$$

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, $P_{1,2}^{[d]} = c \cdot \prod_{i \in [d]} \widehat{\ell}_i$, which is a contradiction by Lemma 4.13.

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

$$(-\widehat{\ell}_0/c_0)^{e-d} \cdot P_{1,2}^{[d]} \;=\; \ell_0^e \;.$$

Again, by unique factorization, we get that $P_{1,2}^{[d]} = c \cdot \ell_0^d$, for some $c \in \mathbb{C}$, which is a contradiction by Lemma 4.13. This finishes the proof. $\square$

4.17 Theorem (Second exp. gap theorem). *If $P_{2,0}^{[d]} \trianglelefteq_{\text{aff}} P_{1,2}^{[e]}$, then $e \geq \exp(d)$.*

Proof. Let $P_{2,0}^{[d]} \trianglelefteq_{\text{aff}} P_{1,2}^{[e]}$. A similar formulation as above (in the previous theorem) gives us that $x_0^{e-d} \cdot P_{2,0}^{[d]} \trianglelefteq P_{1,1}^{[e]}$. By Theorem 4.12 (& its remark), we know that $x_0^{e-d} \cdot P_{2,0}^{[d]} \trianglelefteq P_{1,2}^{[e]}$ implies – either (i) $x_0^{e-d} \cdot P_{2,0}^{[d]} = g + h$, where $g = \prod_{i \in [e]} \ell_i$, for some linear forms $\ell_i \in \mathbb{C}[\boldsymbol{x}]_1$, and $\underline{\text{WR}}(h) \leq 2$, or (ii) $\underline{\text{WR}}(x_0^{e-d} \cdot P_{2,0}^{[d]}) = O(e^8)$. 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 4.16.

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 $x_0^{e-d} \cdot P_{2,0}^{[d]} = \prod_{i \in [e]} \ell_i + \ell_0^e + \ell_{e+1}^e$. 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 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 $P_{2,0}^{[d]} = \gamma_0 \widehat{\ell}_0^d + \gamma_{e+1} \widehat{\ell}_{e+1}^d$. 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 4.14.

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 $S_1 \subseteq [e]$ be such that $\ell_i = c_i \cdot x_0$, for $i \in S_1$, for some nonzero constant $c_i \in \mathbb{C}$, and $S_2 \subseteq [e]$ be 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:

$$(-\widehat{\ell}_j / c_j)^{e-d} \cdot P_{2,0}^{[d]} \;=\; \tilde{\ell}_0^e + \tilde{\ell}_{e+1}^e \,.$$

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 4.14. 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, $x_0^{e-d} \mid \text{LHS} - \prod \ell_i \implies x_0^{e-d} \mid \ell_0^e + \ell_{e+1}^e = \prod_i (\ell_0 - \zeta^{2i+1} \ell_{e+1})$, 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 $x_0^{e-d} \cdot P_{2,0}^{[d]} = \prod_{i \in [e]} \ell_i + \ell_0^{e-1} \cdot \ell_{e+1}$. 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

$$(-\widehat{\ell}_0 / c_0)^{e-d} \cdot P_{2,0}^{[d]} \;=\; \prod_{i \in [e]} \ell_i \,.$$

This in particular implies that $P_{2,0}^{[d]}$ is a product of linear forms, which is a contradiction by Lemma 4.14.

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, $\ell_1 = c_1x_0 + \widehat{\ell}_1$, 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

$$(-\widehat{\ell}_1/c_1)^{e-d} \cdot P_{2,0}^{[d]} = \widehat{\ell}_0^{e-1}\widehat{\ell}_{e+1} \,.$$

Therefore, again by unique factorization, we get that f must a product of linear forms, which is a contradiction by Lemma 4.14.

□

5 Geometric complexity theory of product-plus-power

In this section, we study computational and invariant theoretic properties of $P_{r,s}^{[d]}$. Theorem 5.2 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 $P_{1,1}^{[d]}$, which is achieved in Proposition 5.4. In Proposition 5.5, 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 5.8, we show a connection between the degree of this fundamental invariant and the Alon-Tarsi conjecture on Latin squares in combinatorics.

5.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 §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)].

5.1 Lemma. *Let $V = V_1 \oplus V_2$ and let $f \in \mathbb{C}[V^*]_d = S^dV$ be a homogeneous polynomial with $f = f_1 + f_2$, where $f_i \in S^dV_i$ 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. $\square$

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

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

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

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 5.1, 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 $x_{1j}, \dots, x_{dj}$; 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. $\square$

In the context of geometric complexity theory it is important to know if the polynomial is characterized by its stabilizer [MS08]. While this property fails for the polynomials $P_{r,s}^{[d]}$, a slightly weaker statement is true — every polynomial stabilized by $\mathrm{Stab}(P_{r,s}^{[d]})$ is a restriction of $P_{r,s}^{[d]}$. This is similar to the properties of minrank tensors and slice rank tensors considered in [BIL+19].

5.3 Theorem. *If a polynomial $f \in \mathbb{C}[x_{11}, \ldots, x_{dr}, y_1, \ldots, y_s]_d$ is stabilized by $\mathrm{Stab}(P_{r,s}^{[d]})$, then*

$$f = \alpha \sum_{i=1}^{r} \textstyle\prod_{j=1}^{d} x_{ji} + \beta \displaystyle\sum_{i=1}^{s} y_i^d,$$

for some $\alpha, \beta \in \mathbb{C}$.

Proof. Partition the set of variables into the subsets $X_i = \{x_{1i}, \ldots, x_{di}\}$ and $Y_i = \{y_i\}$. Note that $\mathrm{Stab}(P_{r,s}^{[d]})$ contains the transformation which scales all variables in one of the subsets by a d-th root of unity, acting as identity on all other variables. It follows that each monomial of f contains variables from only one of the subsets, for otherwise the transformation described above multiplies the monomial by a coefficient different from 1. Thus we have

$$f = \sum_{i=1}^{r} f_i(x_{1i}, \ldots, x_{di}) + \sum_{i=1}^{s} \beta_i y_i^d.$$

Since f is fixed under the symmetric group $\mathfrak{S}_s$ permuting $y_1, \ldots, y_s$, the coefficients β_i are all equal. Since f is fixed under the symmetric group $\mathfrak{S}_r$ permuting the subsets $X_1, \ldots, X_r$, all the polynomials f_i also coincide.

Finally, the stabilizer group contains the transformations scaling x_{j1} by λ and x_{k1} by λ^{-1}. This transformation scales a monomial $x_{j1}^{p_j} x_{k1}^{p_k} \ldots$ by $\lambda^{p_j - p_k}$. It follows that each monomial of f_1 must have the same degree with respect to each variable, that is, $f_1 = \alpha \prod_{j=1}^{d} x_{j1}$. $\square$

5.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 $\ydiagram{4,4,3}$. The transpose of the Young diagram is obtained by switching rows and columns. 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]. 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 §1, and we set $G := \mathrm{GL}_{d+1}$. We now use the stabilizer to determine the multiplicities in the coordinate ring of the group orbit $\mathrm{mult}_\lambda(\mathbb{C}[\mathrm{GL}_{d+1} P_{1,1}^{[d]}])$. Let $H := \mathrm{Stab}_G(P_{1,1}^{[d]}) \simeq \mathbb{Z}_d \times (\mathbb{T}^{\mathrm{SL}_d} \rtimes \mathfrak{S}_d)$. A standard consideration in GCT is that since H is reductive, the orbit $GP_{1,1}^{[d]}$ 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 $\mathbb{C}[GP_{1,1}^{[d]}] \simeq \mathbb{C}[G/H] \simeq \mathbb{C}[G]^H$, and therefore $\mathrm{mult}_\lambda(\mathbb{C}[GP_{1,1}^{[d]}]) = \dim(S_\lambda V)^H$. We show how this invariant space dimension can be determined by classical representation branching rules in Proposition 5.4.

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.

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

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

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]. $\square$

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 $\mathrm{GL}_{d+1} P_{1,1}^{[d]} \cap S^d \mathbb{C}^d$.

5.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].

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

Proof. If $d = 2$, then $P_{r,s}^{[2]}$ 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 $R = \mathrm{Stab}_{\mathrm{GL}(V)}(P_{r,s}^{[d]}) \cap \mathbb{T}^{\mathrm{GL}(V)}$, where $\mathbb{T}^{\mathrm{GL}(V)}$ denotes the torus of diagonal matrices, in the basis defined by the variables. By Theorem 5.2, 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 $P_{r,s}^{[d]}$. 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 $P_{r,s}^{[d]}$. This proves the second part of the criterion and concludes the proof. $\square$

Proposition 5.5 reduces to the following in the special case $r = s = 1$:

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

5.d Fundamental invariants and the Alon-Tarsi conjecture

A *Latin square* is an $n \times n$ matrix with entries $1, \ldots, n$ such that each row and each column is a permutation. The *column sign* of a Latin square is the product of the signs of its column permutations. If n is odd, then there are exactly as many sign 1 Latin squares are sign -1 Latin squares, and a sign-reversing involution is obtained by switching the first two rows. The Alon-Tarsi conjecture states that for n even, the number of sign +1 and sign -1 Latin squares are different. The main references on the Alon-Tarsi conjecture are [AT92, Dri97, Gly10], where it is shown that the conjecture is correct for $p \pm 1$ for all odd primes p. [FM19] give a survey about these main results.

5.7 Remark. The GCT result in [Kum15] is based on the Alon-Tarsi conjecture. 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 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.

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.

5.8 Proposition. *Let d be even. The degree of the fundamental invariant of $P_{1,1}^{[d]}$ 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:

$$f_T(p) := \sum_{\text{proper } \vartheta} \prod_{c=1}^{\lambda_1} \det{}_{\vartheta,c} \quad \text{with} \quad \det{}_{\vartheta,c} := \det\left(\ell_{\vartheta(1,c)} \ldots \ell_{\vartheta(\mu_c,c)}\right) \tag{5.9}$$

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 (5.9) 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. □

5.10 Remark. Other fundamental invariants connected to the Alon-Tarsi conjecture have recently been studied in [LZX21, AY22].

5.e New obstructions

For two partitions λ and μ, their sum is defined coordinatewise, i.e., $(\lambda+\mu)_i = \lambda_i + \mu_i$. We write $a \times b$ for the partition $(b, b, \ldots, b)$ of ab. For example, if $d = 3$, then the Young diagram to $\lambda := (5d-1,1) + ((d+1)\times(10d))$ is the following:

This subsection is devoted to proving the following theorem.

5.11 Theorem. *Let $d \geq 3$ be even, and let $\lambda := (5d-1,1) + \big((d+1)\times(10d)\big)$. If d is odd, then we assume $\binom{2(d-1)}{d-1} \geq 2d$. Then we have representation theoretic multiplicity obstructions:*

$$\mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_{d+1}\, P_{1,1}^{[d]}}]) \leq 4 < 5 = \mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_{d+1}(x_1^d + \cdots + x_{d+1}^d)}]),$$

and hence $\overline{\mathrm{GL}_{d+1}(x_1^d + \cdots + x_{d+1}^d)} \not\subseteq \overline{\mathrm{GL}_{d+1}\, P_{1,1}^{[d]}}$. These obstructions are only based on the symmetries of the two polynomials as in [IK20].

The upper and the lower bound are proved independently, see Proposition 5.12 and Proposition 5.13, which proves the theorem. Let $\kappa := (5d-1,1)$, $\blacksquare := (d+1)\times(10d)$, $\square := d \times (10d)$, $\Delta := |\square|/d = 10d$, and $\lambda := \kappa + \blacksquare$.

5.12 Proposition. $\mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_{d+1}\, P_{1,1}^{[d]}}]) \leq \mathrm{mult}_\lambda(\mathbb{C}[\mathrm{GL}_{d+1}\, P_{1,1}^{[d]}]) = \mathrm{mult}_\kappa(\mathbb{C}[\mathrm{GL}_{d+1}\, P_{1,1}^{[d]}]) = 4.$

Proof. The ring $\mathbb{C}[\overline{\mathrm{GL}_{d+1}\, P_{1,1}^{[d]}}]$ is a localization of the ring $\mathbb{C}[\mathrm{GL}_{d+1}\, P_{1,1}^{[d]}]$, see [BI17], which implies $\mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_{d+1}\, P_{1,1}^{[d]}}]) \leq \mathrm{mult}_\lambda(\mathbb{C}[\mathrm{GL}_{d+1}\, P_{1,1}^{[d]}])$. We observe that $a_{\nu+\square}(d, i+\Delta) = a_\nu(d,i)$, because $\square$ has an even number of columns and exactly d rows. Then we calculate:

$$\begin{aligned}
\mathrm{mult}_\lambda(\mathbb{C}[\mathrm{GL}_{d+1}\, P_{1,1}^{[d]}]) &= \sum_{\delta=0}^{5} \sum_{\substack{\mu \vdash \delta d \\ \mu \preceq \lambda \\ \ell(\mu) \leq d}} a_\mu(d,\delta) \\
&= a_{(d)+\square}(d, 1+\Delta) + a_{(d-1,1)+\square}(d,1+\Delta) + a_{(2d)+\square}(d,2+\Delta) \\
&\quad + a_{(2d-1,1)+\square}(d,2+\Delta) + a_{(3d)+\square}(d,3+\Delta) + a_{(3d-1,1)+\square}(d,3+\Delta) \\
&\quad + a_{(4d)+\square}(d,4+\Delta) + a_{(4d-1,1)+\square}(d,4+\Delta) + a_{(5d-1,1)+\square}(d,5+\Delta) \\
&= a_{(d)}(d,1) + a_{(d-1,1)}(d,1) + a_{(2d)}(d,2) + a_{(2d-1,1)}(d,2) + a_{(3d)}(d,3) \\
&\quad + a_{(3d-1,1)}(d,3) + a_{(4d)}(d,4) + a_{(4d-1,1)}(d,4) + a_{(5d-1,1)}(d,5) \\
&= 4,
\end{aligned}$$

because $a_{(nm)}(n,m) = 1$, and $a_{(nm-1,1)}(n,m) = 0$, because $(nm-1,1)$ is of hook shape. Note that there is no summand $a_{(5d)}(d,5)$ and no summand $a_{(0)}(d,0)$, because $(5d) \not\preceq (5d-1,1)$, and $(0) \not\preceq (5d-1,1)$. $\square$

5.13 Proposition. $\mathrm{mult}_\lambda(\mathbb{C}[\overline{\mathrm{GL}_{d+1}(x_1^d + \cdots + x_{d+1}^d)}]) \geq 5.$

Proof. We use the Main Technical Theorem 4.2 from [IK20]. Consider all partitions ϱ of 5, and observe that $\sum_{i=1}^{d+1} 2\lceil \frac{\varrho_i}{2(d-2)} \rceil \leq 10$. In the notation of [IK20], we set $e_\Xi := 10$, which is exactly how many $(d+1)\times d$ blocks form $\blacksquare$.

For a partition $\varrho \vdash_m D$ the *frequency notation* $\widehat{\varrho} \in \mathbb{N}^m$ is defined via $\widehat{\varrho}_i := |\{j \mid \varrho_j = i\}|$. For example, the frequency notation of $\varrho = (3,3,2,0)$ is $\widehat{\varrho} = (0,1,2,0)$. We observe that $|\varrho| = \sum_i i\widehat{\varrho}_i$. We first use Theorem 4.1 from [IK20] (with adjusted notation):

Let $m := d+1$, $D := 5$, $\kappa = (5d-1,1) \vdash_m Dd$. Define

$$b(\kappa,\varrho,d,D) := \sum_{\substack{\mu^1,\mu^2,\ldots,\mu^D \\ \mu^i \vdash di\widehat{\varrho}_i}} c^{\kappa}_{\mu^1,\mu^2,\ldots,\mu^D} \prod_{i=1}^{D} a_{\mu^i}(\widehat{\varrho}_i, i\cdot d).$$

Then

$$\mathrm{mult}_\kappa \mathbb{C}[\mathrm{GL}_m(x_1^d + x_2^d + \cdots + x_m^d)] = \sum_{\varrho \vdash_m D} b(\kappa,\varrho,d,D).$$

For the multi-Littlewood-Richardson coefficient to be nonzero, it is necessary that all $\mu^i \subseteq (5d-1,1)$, so each μ^i is either a single row or a hook $(\widehat{\varrho}_i \cdot i \cdot d - 1, 1)$. But $a_{nm-1,1}(n,m) = 0$ and $a_{nm}(n,m) = 1$, so we can assume that the sum has only the summand with $\mu^i = (\widehat{\varrho}_i \cdot i \cdot d)$ and the product of plethysm coefficients is 1. Hence, the multi-Littlewood-Richardson coefficient counts the number of semistandard tableaux of shape $(5d-1,1)$ and content $(\mu^1,\ldots,\mu^5)$. It is instructive to look at all possible $\widehat{\varrho}$: $\widehat{(1,1,1,1,1)} = (5)$, $\widehat{(2,1,1,1)} = (3,1)$, $\widehat{(2,2,1)} = (1,2)$, $\widehat{(3,1,1)} = (2,0,1)$, $\widehat{(3,2)} = (0,1,1)$, $\widehat{(4,1)} = (1,0,0,1)$, $\widehat{(5)} = (0,0,0,0,1)$. We observe that $\widehat{\varrho}$ has exactly two nonzero entries in 5 cases, and only one nonzero entry in 2 cases. There are no semistandard tableaux of shape $(5d-1,1)$ with only one entry, and there is exactly one semistandard tableaux of shape $(5d-1,1)$ with two symbols and fixed content. Hence

$$\mathrm{mult}_\kappa \mathbb{C}[\mathrm{GL}_{d+1}(x_1^d + x_2^d + \cdots + x_{d+1}^d)] = 5.$$

Note that this argument works indeed for all $d \geq 3$, even though for $d = 3$ we do not have $\varrho = (1,1,1,1,1)$ in the sum (because it has more than $d+1 = 4$ rows, but its contribution is zero anyway).

We now apply Theorem 4.2 from [IK20], which implies

$$\mathrm{mult}_{\blacksquare+\kappa} \mathbb{C}[\overline{\mathrm{GL}_{d+1}(x_1^d + x_2^d + \cdots + x_{d+1}^d)}] \geq 5. \qquad \square$$

6 Homogeneous complexity and the parity-alternating elementary symmetric polynomial

6.a Homogeneous complexity theory

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). In the following definition we make use of the property of $\mathsf{IMM}_n^{(d)}$ to have both a complexity parameter and a degree parameter (unlike the determinant, which only has one combined parameter).

6.1 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* $\forall d > 0, n$ *:* $g_{n,d} \leq f_{q(n),d}$*. 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* $\leq$*.*

Note that in this definition it is important that the maps are homogeneous, see §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-$\overline{\mathrm{p}}$-hard, then $((f))$ is also $\overline{C}$-$\overline{\mathrm{p}}$-hard.

For example, the power sum collection $x_1^d + \cdots + x_n^d$ is p-hard for VWaring, the class of p-families with polynomially bounded Waring rank. And the homogeneous iterated matrix multiplication collection $\mathsf{IMM}_n^{(d)}$ is p-hard for VBP.

6.b 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 $\widehat{f} := f - f(0)$.

6.2 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. □

The following corollary is obvious.

6.3 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. □

6.c 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 6.4 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.

6.4 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 $A_1, \ldots, A_r$ 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 $(A_1, \ldots, A_r)$ 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 $B_1, \ldots, B_r$ such that

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

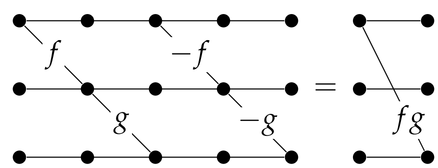
For an input gate (i.e., depth 0) with position (i', j') and input label ℓ, we set $A_1 := \ell \cdot E_{i',j'}$ and $B_1 := -\ell \cdot E_{i',j'}$. For an addition gate with position (i', j') let $(A_1, \ldots, A_r)$, $(B_1, \ldots, B_r)$ and $(A'_1, \ldots, A'_{r'})$, $(B'_1, \ldots, B'_{r'})$ be the lists coming from the induction hypothesis. We define the list for the addition gate as the concatenations $(A_1, \ldots, A_r, A'_1, \ldots, A'_{r'})$ and $(B_1, \ldots, B_r, B'_1, \ldots, B'_{r'})$. 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 $(A_1, \ldots, A_r)$, $(B_1, \ldots, B_r)$ and $(A'_1, \ldots, A'_{r'})$, $(B'_1, \ldots, B'_{r'})$ 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

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

and analogously

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

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. □

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.

6.5 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*

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

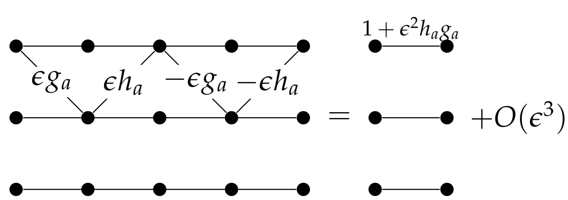
and hence

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

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, $-\epsilon g_2$, …, $-\epsilon h_r$ as in the proof of Proposition 6.4 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 $M_1 M_2 \cdots M_r = \mathrm{id}_3 + \epsilon^2(h_1 g_1 + h_2 g_2 + \cdots + h_r g_r)E_{1,1} + O(\epsilon^3)$. 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 $X_1,\ldots,X_n$.

6.6 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 6.2). The first case is treated with Proposition 6.4, the second is treated completely analogously with Proposition 6.5. We only handle the slightly more difficult second case. We obtain $4^{O(\log n)} = \mathsf{poly}(n)$ 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 §3 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

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

where A'_i arises from A_i by replacing every ϵ by ϵ^d. $\square$

While Corollary 6.6 gives the first collection that is hard for VF_H, the polynomials are similarly unwieldy as $\mathsf{IMM}_n^{(d)}$. In the next sections we will prove that the parity-alternating elementary symmetric polynomial is $\overline{\mathsf{p}}$-hard for a class V3F, which gives a polynomial that is just barely more complicated than the elementary symmetric polynomial.

6.d IHL computation with arity 3 products

In the light of [BIZ18] we now study the 2×2 analogues of Proposition 6.4, Proposition 6.5, Corollary 6.6. 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 6.3, 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

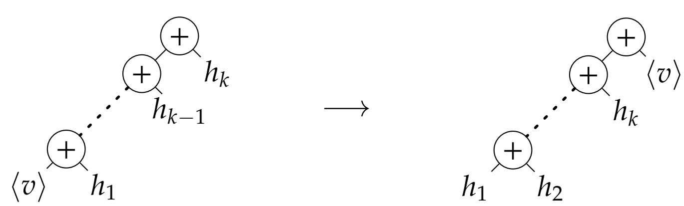
$$\begin{array}{ccccccc} \mathsf{V3F} & \subseteq & \mathsf{VF}_H & \subseteq & \mathsf{VBP}_H & \subseteq & \mathsf{VP}_H, \\ & & \rotatebox{-90}{$\subseteq$} & & & & \\ & & \mathsf{V3P} & & & & \end{array} \tag{6.7}$$

where we prove the vertical inclusion in Theorem 6.15, 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 6.16 that our two new classes also belong to this set: All classes in (6.7) coincide if we go to quasipolynomial complexity instead of polynomial complexity, see (6.17).

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.

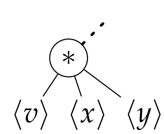
6.8 Proposition (Brent's depth reduction for IHS formulas over the arity 3 basis)**.** *Let f be a polynomial computed by an IHL formula of size s over the arity 3 basis. Then there exists an IHL formula over the arity 3 basis of size* $\mathsf{poly}(s)$ *and depth* $O(\log(s))$ *computing f.*

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:

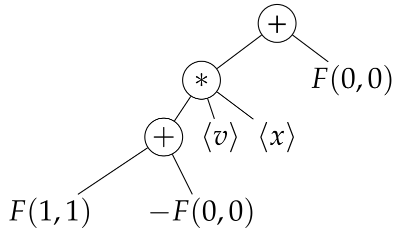


The construction applied to a size s formula gives $\mathrm{Depth}(s) \leq \mathrm{Depth}(\frac{2}{3}s) + 1$. The resulting size is $\mathrm{Size}(s) \leq 2 \cdot \mathrm{Size}(\frac{2}{3}s) + 1$.

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



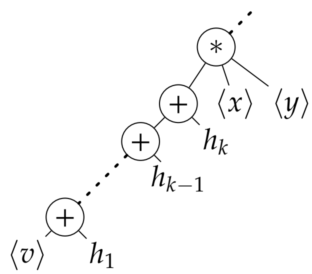
We now replace $\langle v \rangle$ by a new variable α and $\langle x \rangle$ by a new variable β. We observe that the resulting polynomial F (interpreted as a bivariate polynomial in α and β) is *linear* in the product $\alpha\beta$. Therefore $F(\alpha, \beta) = \alpha\beta(F(1,1) - F(0,0)) + F(0,0)$. 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:



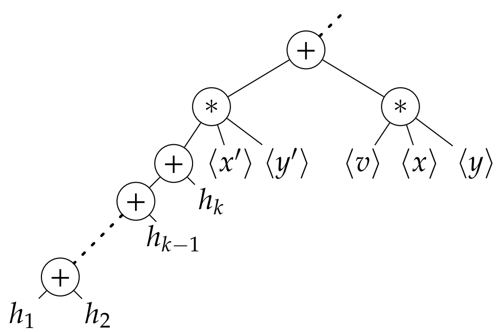
(6.9)

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

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 (6.9). The construction on a size s formula gives $\mathrm{Depth}(s) \leq \mathrm{Depth}(\frac{2}{3}s) + 2$. The resulting size is $\mathrm{Size}(s) \leq 5 \cdot \mathrm{Size}(\frac{2}{3}s) + 3$. Putting all cases together, the construction has $\mathrm{Depth}(s) \leq \mathrm{Depth}(\frac{2}{3}s) + 2$ and $\mathrm{Size}(s) \leq 5 \cdot \mathrm{Size}(\frac{2}{3}s) + 3$. Hence applying the construction recursively gives logarithmic depth and polynomial size. $\square$

6.e 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

$$C_{n,d} = \sum_{(i_1,i_2,\ldots,i_d)\in P} x_{i_1} x_{i_2} \cdots x_{i_d}. \tag{6.10}$$

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

6.11 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 6.6 we have the following theorem.

6.12 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 6.8 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

$$\alpha f \cdot E_{\text{odd}} \simeq (\mathrm{id}_2 + \ell_1 E_{\text{odd}})(\mathrm{id}_2 + \ell_2 E_{\text{even}}) \cdots (\mathrm{id}_2 + \ell_r E_{\text{odd}}) - \mathrm{id}_2 \tag{6.13}$$

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 (6.13) also implies (set ϵ to ϵ^3 and α to $\epsilon^2\alpha$) that

$$\epsilon^2 \alpha f \cdot E_{\text{odd}} \equiv \big((\mathrm{id}_2 + \ell_1''' E_{\text{odd}})(\mathrm{id}_2 + \ell_2''' E_{\text{even}}) \cdots (\mathrm{id}_2 + \ell_r''' E_{\text{odd}}) - \mathrm{id}_2\big) \pmod{\epsilon^3}$$

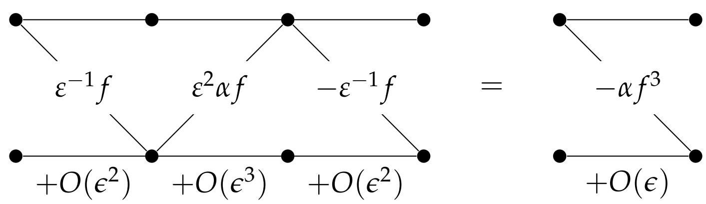
Transposing gives

$$\epsilon^2 \alpha f \cdot E_{\text{even}} \equiv \big((\mathrm{id}_2 + \ell_r''' E_{\text{even}})(\mathrm{id}_2 + \ell_{r-1}''' E_{\text{odd}}) \cdots (\mathrm{id}_2 + \ell_1''' E_{\text{even}}) - \mathrm{id}_2\big) \pmod{\epsilon^3}$$

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

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

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

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

We now treat the case where f_n has even degree, using an argument similar to the one form Proposition 6.5. By the above construction, for each i we find

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

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

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

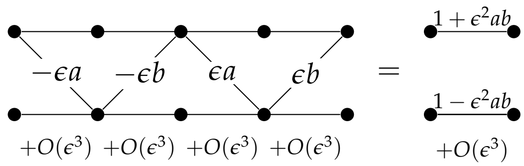
Analogously, when replacing α by $-\epsilon$ instead:

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

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

$$(-\epsilon a \cdot E_{\text{odd}} + \mathrm{id}_2 + O(\epsilon^3))(-\epsilon b \cdot E_{\text{even}} + \mathrm{id}_2 + O(\epsilon^3))(\epsilon a \cdot E_{\text{odd}} + \mathrm{id}_2 + O(\epsilon^3))(\epsilon b \cdot E_{\text{even}} + \mathrm{id}_2 + O(\epsilon^3))$$
$$\equiv \begin{pmatrix} 1 + \epsilon^2 a \cdot b & 0 \\ 0 & 1 - \epsilon^2 a \cdot b \end{pmatrix} \pmod{\epsilon^3}.$$

Pictorially:



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

$$(M(a_1b_1)+O(\epsilon^3))\cdot(M(a_2b_2)+O(\epsilon^3))\cdots(M(a_nb_n)+O(\epsilon^3)) \equiv M(a_1b_1+a_2b_2+\cdots a_nb_n) \pmod{\epsilon^3}.$$

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

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

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):

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

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 6.16. $\square$

6.14 Remark. Algebraic models of computation that are similar to using $C_{r,d}$ have also been studied in [MS21, BIZ18]. In [BIZ18], it is shown that $\overline{\mathsf{VF}}$ can be via the orbit closure of polynomially sized continuants. [MS21] constructs polynomial sized hitting sets for affine orbits of the cyclic continuant polynomial. $C_{r,d}$ is a homogeneous variant of the continuant polynomial defined in [BIZ18].

6.f Converting formulas to circuits over the arity 3 basis

In this section we prove the following theorem.

6.15 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 6.2 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 6.2). 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.

6.g Valiant-Skyum-Berkowitz-Rackoff over the arity 3 basis

6.16 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 6.18 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 $s^{O(\log s)} = n^{\mathsf{polylog}(n)}$ [5]. 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 (6.7) simplifies:

$$\mathsf{VQ3F} = \mathsf{VQF}_H = \mathsf{VQBP}_H = \mathsf{VQP}_H = \mathsf{VQ3P}. \tag{6.17}$$

The following Theorem 6.18 is needed in the proof of Theorem 6.16. 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.

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

6.18 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 $v \leq u$ 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

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

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

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

6.19 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:

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

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):

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

□

6.20 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):

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

$\square$

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*.

$$\begin{aligned}
u &\overset{\text{Lem. 6.19}}{\equiv} \sum_{w \in \mathcal{F}} [u:w]_{\langle w \rangle} = \sum_{w \in \mathcal{F}} [u:w]_{\langle w_1 \rangle} \times w_2 \times w_3 = \sum_{\substack{w \in \mathcal{F} \\ \deg(u) \geq \deg(w)}} [u:w]_{\langle w_1 \rangle} \times w_2 \times w_3 \\
&\equiv \sum_{\substack{w \in \mathcal{F} \\ \deg(u) \geq \deg(w)}} [u:w]_{\langle w_3 \rangle} \times w_2 \times w_1
\end{aligned}$$

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 $w \in \mathcal{F}$.

$$
\begin{aligned}
[u:v] &\overset{\text{Lem. 6.20}}{\equiv} \sum_{w\in\mathcal{F}} [u:w]_{[w:v]} = \sum_{\substack{w\in\mathcal{F}\\ \deg(u)\geq\deg(w)}} \frac{[u:w]}{z}\cdot[w:v] = \frac{1}{z}\sum_{\substack{w\in\mathcal{F}\\ \deg(u)\geq\deg(w)}} [u:w]\cdot[w_1:v]\cdot w_2\cdot w_3 \\
&\equiv \sum_{\substack{w\in\mathcal{F}\\ \deg(u)\geq\deg(w)}} [u:w]_{\langle w_3\rangle}\times[w_1:v]\times w_2 \\
&\overset{\text{Lem. 6.19}}{\equiv} \sum_{\substack{w\in\mathcal{F}\\ \deg(u)\geq\deg(w)}} [u:w]_{\langle w_3\rangle}\times[w_1:v]\times\left(\sum_{\substack{y\in\mathcal{F}'\\ \deg(w_2)\geq\deg(y)}} [w_2:y]_{\langle y_3\rangle}\times y_2\times y_1\right) \\
&\equiv \sum_{\substack{w\in\mathcal{F}\\ \deg(u)\geq\deg(w)}}\ \sum_{\substack{y\in\mathcal{F}'\\ \deg(w_2)\geq\deg(y)}} [u:w]_{\langle w_3\rangle}\times[w_1:v]\times\left([w_2:y]_{\langle y_3\rangle}\times y_2\times y_1\right)
\end{aligned}
$$

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. □

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.

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

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

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

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

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

$(7,7,5,5,4)_{1>0}$, $(7,7,6,5,3)_{1>0}$, $(7,7,7,4,3)_{1>0}$, $(7,7,7,5,2)_{1>0}$, $(7,7,7,7,0)_{1>0}$, $(8,6,6,4,4)_{4>1}$, $(8,7,5,4,4)_{1>0}$, $(8,7,5,5,3)_{2>0}$, $(8,7,6,4,3)_{4>2}$, $(8,7,6,5,2)_{4>1}$, $(8,7,7,3,3)_{3>0}$, $(8,7,7,4,2)_{1>0}$, $(8,7,7,5,1)_{3>0}$, $(8,8,4,4,4)_{4>2}$, $(8,8,5,4,3)_{4>1}$, $(8,8,6,4,2)_{9>4}$, $(8,8,7,3,2)_{3>1}$, $(8,8,7,4,1)_{4>3}$, $(8,8,8,2,2)_{3>2}$, $(9,6,5,4,4)_{3>0}$, $(9,6,5,5,3)_{1>0}$, $(9,6,6,4,3)_{5>3}$, $(9,6,6,5,2)_{4>3}$, $(9,7,4,4,4)_{2>1}$, $(9,7,5,4,3)_{7>2}$, $(9,7,5,5,2)_{5>1}$, $(9,7,6,3,3)_{5>3}$, $(9,7,6,4,2)_{10>5}$, $(9,7,6,5,1)_{6>4}$, $(9,7,7,3,2)_{5>1}$, $(9,7,7,4,1)_{5>2}$, $(9,7,7,5,0)_{2>1}$, $(9,8,4,4,3)_{5>2}$, $(9,8,5,3,3)_{4>1}$, $(9,8,5,4,2)_{11>5}$, $(9,8,5,5,1)_{4>3}$, $(9,8,6,3,2)_{11>6}$, $(9,8,6,4,1)_{12>11}$, $(9,8,7,2,2)_{5>3}$, $(9,8,7,3,1)_{8>6}$, $(9,9,4,3,3)_{3>1}$, $(9,9,4,4,2)_{2>1}$,

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

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

$(7,7,7,7,4)_{1>0}$, $(8,6,6,6,6)_{2>1}$, $(8,7,6,6,5)_{1>0}$, $(8,7,7,5,5)_{3>0}$, $(8,7,7,6,4)_{1>0}$, $(8,7,7,7,3)_{2>0}$,
$(8,8,6,6,4)_{7>1}$, $(8,8,7,5,4)_{3>0}$, $(8,8,7,6,3)_{5>0}$, $(8,8,8,4,4)_{8>2}$, $(8,8,8,5,3)_{2>1}$, $(8,8,8,6,2)_{7>2}$,
$(9,6,6,6,5)_{2>1}$, $(9,7,6,5,5)_{3>0}$, $(9,7,6,6,4)_{5>1}$, $(9,7,7,5,4)_{7>0}$, $(9,7,7,6,3)_{6>0}$, $(9,7,7,7,2)_{3>0}$,
$(9,8,5,5,5)_{1>0}$, $(9,8,6,5,4)_{14>2}$, $(9,8,6,6,3)_{12>3}$, $(9,8,7,4,4)_{10>1}$, $(9,8,7,5,3)_{18>2}$, $(9,8,7,6,2)_{13>2}$,
$(9,8,7,7,1)_{3>0}$, $(9,8,8,4,3)_{11>2}$, $(9,8,8,5,2)_{12>4}$, $(9,8,8,6,1)_{7>4}$, $(9,9,5,5,4)_{6>0}$, $(9,9,6,4,4)_{5>0}$,
$(9,9,6,5,3)_{15>3}$, $(9,9,6,6,2)_{5>2}$, $(9,9,7,4,3)_{14>1}$, $(9,9,7,5,2)_{17>3}$, $(9,9,7,6,1)_{7>2}$, $(9,9,7,7,0)_{2>0}$,
$(9,9,8,3,3)_{8>1}$, $(9,9,8,4,2)_{8>1}$, $(9,9,8,5,1)_{9>3}$, $(9,9,9,3,2)_{3>1}$, $(9,9,9,4,1)_{3>0}$, $(10,6,6,6,4)_{9>3}$,
$(10,7,5,5,5)_{3>0}$, $(10,7,6,5,4)_{15>1}$, $(10,7,6,6,3)_{13>3}$, $(10,7,7,4,4)_{5>0}$, $(10,7,7,5,3)_{19>1}$,
$(10,7,7,6,2)_{8>0}$, $(10,7,7,7,1)_{4>0}$, $(10,8,5,5,4)_{7>0}$, $(10,8,6,4,4)_{31>4}$, $(10,8,6,5,3)_{32>5}$,
$(10,8,6,6,2)_{29>8}$, $(10,8,7,4,3)_{35>5}$, $(10,8,7,5,2)_{34>6}$, $(10,8,7,6,1)_{18>6}$, $(10,8,8,3,3)_{4>1}$,
$(10,8,8,4,2)_{33>9}$, $(10,8,8,5,1)_{15>9}$, $(10,9,5,4,4)_{15>1}$, $(10,9,5,5,3)_{16>1}$, $(10,9,6,4,3)_{39>6}$,
$(10,9,6,5,2)_{38>8}$, $(10,9,6,6,1)_{16>9}$, $(10,9,7,3,3)_{21>5}$, $(10,9,7,4,2)_{43>8}$, $(10,9,7,5,1)_{28>9}$,
$(10,9,8,3,2)_{24>7}$, $(10,9,8,4,1)_{24>10}$, $(10,9,9,2,2)_{2>0}$, $(10,9,9,3,1)_{8>3}$, $(10,10,4,4,4)_{12>2}$,
$(10,10,5,4,3)_{18>3}$, $(10,10,5,5,2)_{7>0}$, $(10,10,6,3,3)_{8>2}$, $(10,10,6,4,2)_{42>10}$, $(10,10,6,5,1)_{18>7}$,
$(10,10,6,6,0)_{11>10}$, $(10,10,7,3,2)_{23>6}$, $(10,10,7,4,1)_{26>12}$, $(10,10,8,2,2)_{17>5}$, $(10,10,8,3,1)_{13>9}$,
$(10,10,9,2,1)_{6>4}$, $(11,6,6,5,4)_{9>1}$, $(11,6,6,6,3)_{10>4}$, $(11,7,5,5,4)_{11>0}$, $(11,7,6,4,4)_{22>3}$,
$(11,7,6,5,3)_{31>4}$, $(11,7,6,6,2)_{19>6}$, $(11,7,7,4,3)_{25>3}$, $(11,7,7,5,2)_{25>2}$, $(11,7,7,6,1)_{11>2}$,
$(11,7,7,7,0)_{2>0}$, $(11,8,5,4,4)_{26>3}$, $(11,8,5,5,3)_{23>2}$, $(11,8,6,4,3)_{60>11}$, $(11,8,6,5,2)_{58>13}$,
$(11,8,6,6,1)_{24>13}$, $(11,8,7,3,3)_{26>4}$, $(11,8,7,4,2)_{64>14}$, $(11,8,7,5,1)_{40>15}$, $(11,8,8,3,2)_{28>9}$,
$(11,8,8,4,1)_{30>17}$, $(11,9,4,4,4)_{11>1}$, $(11,9,5,4,3)_{45>6}$, $(11,9,5,5,2)_{33>5}$, $(11,9,6,3,3)_{36>8}$,
$(11,9,6,4,2)_{78>19}$, $(11,9,6,5,1)_{46>20}$, $(11,9,7,3,2)_{57>14}$, $(11,9,7,4,1)_{58>24}$, $(11,9,8,2,2)_{20>7}$,
$(11,9,8,3,1)_{37>21}$, $(11,9,9,2,1)_{9>5}$, $(11,10,4,4,3)_{21>5}$, $(11,10,5,3,3)_{20>4}$, $(11,10,5,4,2)_{52>12}$,
$(11,10,5,5,1)_{20>7}$, $(11,10,6,3,2)_{56>16}$, $(11,10,6,4,1)_{56>29}$, $(11,10,7,2,2)_{30>9}$, $(11,10,7,3,1)_{46>26}$,
$(11,10,8,2,1)_{25>20}$, $(11,11,4,3,3)_{10>2}$, $(11,11,4,4,2)_{10>3}$, $(11,11,5,3,2)_{26>7}$, $(11,11,5,4,1)_{23>12}$,
$(11,11,6,2,2)_{13>5}$, $(11,11,6,3,1)_{30>18}$, $(11,11,7,2,1)_{19>15}$, $(12,6,5,5,4)_{4>0}$, $(12,6,6,4,4)_{21>3}$,
$(12,6,6,5,3)_{14>3}$, $(12,6,6,6,2)_{17>8}$, $(12,7,5,4,4)_{19>1}$, $(12,7,5,5,3)_{22>1}$, $(12,7,6,4,3)_{49>10}$,
$(12,7,6,5,2)_{46>9}$, $(12,7,6,6,1)_{17>10}$, $(12,7,7,3,3)_{23>3}$, $(12,7,7,4,2)_{32>5}$, $(12,7,7,5,1)_{26>6}$,
$(12,8,4,4,4)_{25>5}$, $(12,8,5,4,3)_{56>8}$, $(12,8,5,5,2)_{32>5}$, $(12,8,6,3,3)_{32>7}$, $(12,8,6,4,2)_{109>29}$,
$(12,8,6,5,1)_{54>27}$, $(12,8,7,3,2)_{62>17}$, $(12,8,7,4,1)_{65>31}$, $(12,8,8,2,2)_{30>13}$, $(12,8,8,3,1)_{27>20}$,
$(12,9,4,4,3)_{33>6}$, $(12,9,5,3,3)_{35>7}$, $(12,9,5,4,2)_{80>18}$, $(12,9,5,5,1)_{32>11}$, $(12,9,6,3,2)_{88>28}$,
$(12,9,6,4,1)_{88>45}$, $(12,9,7,2,2)_{41>14}$, $(12,9,7,3,1)_{71>40}$, $(12,9,8,2,1)_{34>28}$, $(12,10,4,3,3)_{14>4}$,

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

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

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

B Characterizing small border Waring rank

The results on generalized additive decompositions from §2.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 2.8 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 2.8 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 $\frac{1}{a^{d-1}d}\ell_2$. ☐

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 2.8 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 $\underline{\mathrm{WR}}(\ell_i^{d-r_i+1}g_i) \leq r_i$.

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. 39

[AT92] Noga Alon and Michael Tarsi, *Colorings and orientations of graphs*, Combinatorica **12** (1992), no. 2, 125–134. 39

[AY22] Alimzhan Amanov and Damir Yeliussizov, *Fundamental Invariants of Tensors, Latin Hypercubes, and Rectangular Kronecker Coefficients*, International Mathematics Research Notices (2022), 1–35, rnac311. 40

[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. 6

[BB14] W. Buczyńska and J. Buczyński, *Secant varieties to high degree Veronese reembeddings, catalecticant matrices and smoothable Gorenstein schemes*, J. Alg. Geom. **23** (2014), no. 1, 63–90. 2

[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. 8, 9, 17

[BB17] Edoardo Ballico and Alessandra Bernardi, *Curvilinear schemes and maximum rank of forms*, Le Matematiche **72** (2017), no. 1, 137–144. 6

[BB21] Weronika Buczyńska and Jarosław Buczyński, *Apolarity, border rank, and multigraded Hilbert scheme*, Duke Math. J. **170** (2021), no. 16, 3659 – 3702. 2, 20, 21

[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. 8, 17, 21

[BBT13] Weronika Buczyńska, Jarosław Buczyński, and Zach Teitler, *Waring decompositions of monomials*, Journal of Algebra **378** (2013), 45–57. 8, 16

[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. 10, 44

[BCRL79] Dario Bini, Milvio Capovani, Francesco Romani, and Grazia Lotti, *$O(n^{2.7799})$ complexity for n*n approximate matrix multiplication*, Inf. Process. Lett. **8** (1979), no. 5, 234–235. 2

[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. 4, 6, 7, 27, 40

[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. 8, 38

[BI11] Peter Bürgisser and Christian Ikenmeyer, *Geometric complexity theory and tensor rank*, Proceedings of the forty-third annual ACM symposium on Theory of computing, 2011, pp. 509–518. 5

[BI13] ———, *Explicit lower bounds via geometric complexity theory*, Proceedings of the 45^{th} Annual ACM symposium on Theory of Computing, 2013, pp. 141–150. 5, 39

[BI17] ———, *Fundamental invariants of orbit closures*, Journal of Algebra **477** (2017), 390–434. 10, 35, 38, 39, 40, 41

[BIL+19] Markus Bläser, Christian Ikenmeyer, Vladimir Lysikov, Anurag Pandey, and Frank-Olaf Schreyer, *Variety membership testing, algebraic natural proofs, and geometric complexity theory*, 2019. 37

[Bin80] Dario Bini, *Relations between exact and approximate bilinear algorithms. Applications*, Calcolo **17** (1980), no. 1, 87–97. 2

[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. 5

[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. 4, 11, 45, 46, 49, 51

[BJMR18] Alessandra Bernardi, Joachim Jelisiejew, Pedro Macias Marques, and Kristian Ranestad, *On polynomials with given Hilbert function and applications*, Collectanea Mathematica **69** (2018), no. 1, 39–64. 18, 21

[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. 3, 38

[Bre74] Richard P. Brent, *The parallel evaluation of general arithmetic expressions*, J. ACM **21** (1974), no. 2, 201–206. 43, 47

[BT15] G. Blekherman and Z. Teitler, *On maximum, typical and generic ranks*, Mathematische Annalen **362** (2015), no. 3-4, 1021–1031. 17

[Bür00] Peter Bürgisser, *Completeness and reduction in algebraic complexity theory*, Algorithms and Computation in Mathematics, Springer, 2000, Online version: `http://math-www.uni-paderborn.de/agpb/work/habil.ps`. 3

[Bür04] ———, *The complexity of factors of multivariate polynomials*, Found. Comput. Math. **4** (2004), no. 4, 369–396. 3

[Car06] Enrico Carlini, *Reducing the number of variables of a polynomial*, Algebraic geometry and geometric modeling, Springer, 2006, pp. 237–247. 11

[Cay45] A. Cayley, *On the theory of linear transformations*, Cambridge Math. J. **iv** (1845), 193–209. 2

[CGJ19] M. Christandl, F. Gesmundo, and A. K. Jensen, *Border rank is not multiplicative under the tensor product*, SIAM J. Appl. Alg. Geom. **3** (2019), 231–255. 2

[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. 35

[CHI+18] Luca Chiantini, Jonathan D. Hauenstein, Christian Ikenmeyer, Joseph M. Landsberg, and Giorgio Ottaviani, *Polynomials and the exponent of matrix multiplication*, Bulletin of the London Mathematical Society **50** (2018), no. 3, 369–389. 2

[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. 40

[CKW11a] Xi Chen, Neeraj Kayal, and Avi Wigderson, *Partial derivatives in arithmetic complexity and beyond*, Now Publishers Inc, 2011. 7

[CKW11b] 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. 25, 32

[Cle61] A. Clebsch, *Zur Theorie der algebraischen Flächen*, J. Reine Angew. Math. **58** (1861), 93–108. 2

[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

[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. 10

[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. 39

[FH91] William Fulton and Joe Harris, *Representation theory: a first course*, vol. 129, Springer Science & Business Media, 1991. 5, 38

[FM19] Benjamin Friedman and Sean McGuinness, *The Alon–Tarsi conjecture: A perspective on the main results*, Discrete Math. **342** (2019), no. 8, 2234–2253. 39

[For14] Michael A. Forbes, *Polynomial Identity Testing of Read-Once Oblivious Algebraic Branch-ing Programs*, PhD thesis, Massachusetts Institute of Technology (2014), 1–220. 4, 27

[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. 37

[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. 38

[Ges16] F. Gesmundo, *Geometric Aspects of Iterated Matrix Multiplication*, J. Algebra **461** (2016), 42–64. 35

[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. 39

[Gor94] Paul Gordan, *Das Zerfallen der Curven in gerade Linien*, Mathematische Annalen **45** (1894), no. 3, 410–427. 10

[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 20

[Gua18] Yonghui Guan, *Brill's equations as a $GL(V)$-module*, Linear Algebra and its Applications **548** (2018), 273–292. 10

[GW09] R. Goodman and N. R. Wallach, *Symmetry, representations, and invariants*, Graduate Texts in Mathematics, vol. 255, Springer, Dordrecht, 2009. 38

[GY10] John Hilton Grace and Alfred Young, *The algebra of invariants*, Cambridge Library Collection - Mathematics, Cambridge University Press, 2010. 14

[Har77] Robin Hartshorne, *Algebraic geometry*, Springer-Verlag, New York, 1977, Graduate Texts in Mathematics, No. 52. MR 0463157 (57 #3116) 18

[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. 39

[HS04] Mark Haiman and B Sturmfels, *Multigraded Hilbert schemes*, Journal of Algebraic Geometry **13** (2004), no. 4, 725–769. 20

[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. 8, 12

[IK99] Anthony Iarrobino and Vassil Kanev, *Power sums, Gorenstein algebras, and determinantal loci*, Lecture Notes in Mathematics, vol. 1721, Springer-Verlag, Berlin, 1999. 2, 9, 14, 18, 20

[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. 5, 7, 10, 38, 39, 41, 42

[IP17] Christian Ikenmeyer and Greta Panova, *Rectangular Kronecker coefficients and plethysms in geometric complexity theory*, Advances in Mathematics **319** (2017), 40–66. 5

[IS22] Christian Ikenmeyer and Abhiroop Sanyal, *A note on VNP-completeness and border complexity*, Information Processing Letters **176** (2022), 106243. 3

[Jel20] J. Jelisiejew, *Pathologies on the Hilbert scheme of points*, Inventiones mathematicae **220** (2020), no. 2, 581–610. 2

[JM22] J. Jelisiejew and T. Mańdziuk, *Limits of saturated ideals*, arXiv:2210.13579 (2022), 1–31. 2

[Kra85] Hanspeter Kraft, *Geometrische Methoden in der Invariantentheorie*, Springer, 1985. 3

[KS07] Neeraj Kayal and Nitin Saxena, *Polynomial identity testing for depth 3 circuits*, computational complexity **16** (2007), no. 2, 115–138. 11

[Kum15] Shrawan Kumar, *A study of the representations supported by the orbit closure of the determinant*, Compositio Mathematica **151** (2015), no. 2, 292–312. 10, 39

[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, 6, 7, 24, 32

[Lan17] J. M. Landsberg, *Geometry and complexity theory*, Cambridge Studies in Advanced Mathematics, vol. 169, Cambridge University Press, Cambridge, 2017. 36

[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. 6

[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. 6, 25, 28, 31, 32, 58

[LZX21] Xin Li, Liping Zhang, and Hanchen Xia, *Two classes of minimal generic fundamental invariants for tensors*, arXiv:2111.07343, 2021. 40

[Mac95] I. G. Macdonald, *Symmetric functions and Hall polynomials*, second ed., Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1995. 23

[Mah14] M. Mahajan, *Algebraic complexity classes*, Perspectives in Comp. Compl.: The Somenath Biswas Ann. Vol. (2014), 51–75. 3

[Mat60] Yozô Matsushima, *Espaces homogènes de Stein des groupes de Lie complexes*, Nagoya Mathematical Journal **16** (1960), 205–218. 38

[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, 5

[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. 5, 37

[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. 51

[Pal06] F. Palatini, *Sulle superficie algebriche i cui S_h $(h+1)$-seganti non riempiono lo spazio ambiente*, Atti della R. Acc. delle Scienze di Torino **41** (1906), 634–640. 2

[Sap21] Ramprasad Saptharishi, *A survey of lower bounds in arithmetic circuit complexity*, Github Survey, 2021. 47, 53

[Sax08] Nitin Saxena, *Diagonal Circuit Identity Testing and Lower Bounds*, ICALP 2008, 2008, pp. 60–71. 2, 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. 6, 23, 24

[SW12] Douglas S Stones and Ian M Wanless, *How not to prove the Alon-Tarsi conjecture*, Nagoya Mathematical Journal **205** (2012), 1–24. 39

[Syl52] J. J. Sylvester, *On the principles of the calculus of forms*, J. Cambridge and Dublin Math. **7** (1852), 52–97. 2, 11

[Ter11] A. Terracini, *Sulle v_k per cui la varietà degli $s_h(h+1)$-seganti ha dimensione minore dell'ordinario*, Rend. Circ. Mat. **31** (1911), 392–396. 2

[Tod92] Seinosuke Toda, *Classes of arithmetic circuits capturing the complexity of computing the determinant*, IEICE Transactions on Information and Systems **75** (1992), no. 1, 116–124. 3

[Val79] Leslie G. Valiant, *Completeness classes in algebra*, Proceedings of the 11h Annual ACM Symposium on Theory of Computing, 1979, pp. 249–261. 3

[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. 11, 47, 52, 55

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il