

PARTIALLY SYMMETRIC VARIANTS OF COMON'S PROBLEM  
VIA SIMULTANEOUS RANK\*FULVIO GESMUNDO<sup>†</sup>, ALESSANDRO ONETO<sup>‡</sup>, AND EMANUELE VENTURA<sup>§</sup>

**Abstract.** A symmetric tensor may be regarded as a partially symmetric tensor in several different ways. These produce different notions of rank for the symmetric tensor which are related by chains of inequalities. By exploiting algebraic tools such as apolarity theory, we show how the study of the simultaneous symmetric rank of partial derivatives of the homogeneous polynomial associated to the symmetric tensor can be used to prove equalities among different partially symmetric ranks. This approach aims to understand to what extent the symmetries of a tensor affect its rank. We apply this to the special cases of binary forms, ternary and quaternary cubics, monomials, and elementary symmetric polynomials.

**Key words.** tensors, partially symmetric tensors, symmetric tensors, tensor rank, partially symmetric rank, Waring rank, apolarity, simultaneous rank, partial derivatives

**AMS subject classifications.** Primary, 15A69, 13P05; Secondary, 13A02, 14N05

**DOI.** 10.1137/18M1225422

**1. Introduction.** The problem of representing tensors in convenient ways is connected to several areas of pure and applied mathematics. A line of research concerns *additive decompositions*: given a tensor of order  $d$ , say,  $t \in V^{\otimes d}$ , a *tensor decomposition* of  $t$  is a sum of *rank-one* tensors, i.e., elements of the form  $v_1 \otimes \cdots \otimes v_d$ , adding up to  $t$ . The smallest length of such a decomposition of  $t$  is the *tensor rank* of  $t$ . Whenever a tensor satisfies certain *symmetries*, it is natural to study tensor decompositions reflecting such symmetries. Thus, several possible notions of rank arise, which are usually referred to as *(partially) symmetric* tensor ranks. The study of partially symmetric tensors has recently gained interest; see, for instance, [BBCM13, BBCG18].

The space of homogeneous polynomials, or *forms*, of degree  $d$  on a vector space  $V^*$  can be naturally identified with the space of symmetric tensors in  $V^{\otimes d}$ ; denote this space by  $S^d V$ . Symmetric tensor decompositions are classically known as *Waring decompositions*; these are sums of powers of linear forms. The corresponding rank is the *Waring rank*. This class of decompositions has been studied since the XIX century, when Sylvester completed the classification of binary forms in terms of their Waring rank [Syl52], that is, the case where  $\dim V = 2$ . A great amount of work is devoted to this topic; among others, we mention Clebsch [Cle61], Lasker [Las04], Palatini [Pal03a, Pal03b], and Terracini [Ter15, Ter16]. During the last decades of

---

\*Received by the editors November 7, 2018; accepted for publication (in revised form) by A. Bernardi August 28, 2019; published electronically November 26, 2019.

<https://doi.org/10.1137/18M1225422>

**Funding:** The research of the first author was supported by the European Research Council (ERC grant agreement 337603) and VILLUM FONDEN via the QMATH Centre of Excellence (grant 10059). The research of the second author was supported by the Spanish Ministry of Economy and Competitiveness through the María de Maeztu Programme for Units of Excellence in R&D (MDM-2014-0445). The research of the third author was supported by grant 346300 for IMPAN from the Simons Foundation and the matching 2015–2019 Polish MNiSW fund.

<sup>†</sup>QMATH, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen O., Denmark ([fulges@math.ku.dk](mailto:fulges@math.ku.dk)).

<sup>‡</sup>FMA-IAG, Otto-von-Guericke Universität Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany ([alessandro.oneto@ovgu.de](mailto:alessandro.oneto@ovgu.de), [aless.oneto@gmail.com](mailto:aless.oneto@gmail.com)).

<sup>§</sup>Department of Mathematics, Texas A&M University, College Station, TX 77843-3368 ([eventura@math.tamu.edu](mailto:eventura@math.tamu.edu), [emanueleventura.sw@gmail.com](mailto:emanueleventura.sw@gmail.com)).

the last century, Sylvester's ideas were reread in the modern algebraic language of apolarity theory; see, e.g., [Rez92, DK93, ER93, IK99]. A major breakthrough in the development of the subject was accomplished in 1995 by Alexander and Hirschowitz [AH95], who resolved the long standing problem of determining Waring ranks of generic forms in any number of variables and any degree (see also [Iar95]). Throughout the years, the Waring problem attracted the attention of a broader community, and classical and modern tools from algebraic geometry as well as from other fields have been employed for a variety of questions in this subject; see, for instance, [Kle99, BCMT10, CS11, CCG12, BBT13].

Disregarding its symmetries, a symmetric tensor can be regarded as an element of the space of partially symmetric tensors for different choices of partial symmetries, and one can ask what are the relations among the different (partially symmetric) ranks which arise in this way. This was the object of a famous question raised by Comon, who asked whether the tensor rank of a symmetric tensor equals its symmetric rank; see [Oed08, Problem 15]. This problem received a great deal of attention in the last few years. Affirmative answers were derived under certain assumptions [CGLM08, BB13, ZHQ16, Fri16, Sei18]. Recently, Shitov gave an example for  $d = 3$  and  $\dim V = 800$ , where Comon's question has negative answer [Shi18].

In this article, we approach a partially symmetric version of Comon's question investigating relations among the partially symmetric ranks of a symmetric tensor. Our results will be obtained via the study of *simultaneous Waring decompositions* of the set of  $k$ th partial derivatives of homogeneous polynomials.

**1.1. Formulation of the problem.** Let  $V$  be a vector space of dimension  $n + 1$  over an algebraically closed field  $\mathbb{k}$  of characteristic zero.

Let  $d \geq 0$  be an integer and  $\underline{d} = (d_1, \dots, d_m) \in \mathbb{N}^m$  be a sequence of integers with  $\sum_{i=1}^m d_i = d$ . In this case,  $\underline{d}$  is called a *composition* of  $d$  and denoted  $\underline{d} \vdash d$ ; write  $\underline{d} \vdash d$  when the length of the composition is relevant.

Let  $\{x_0, \dots, x_n\}$  be a basis of  $V$ .

- For any multi-index  $\alpha \in \mathbb{N}^{n+1}$ , write

$$|\alpha| := \alpha_0 + \cdots + \alpha_n \quad \text{and} \quad \alpha! := \alpha_0! \cdots \alpha_n!.$$

- The symmetric algebra  $S^\bullet V = \bigoplus_{j \geq 0} S^j V$  of  $V$  is  $\mathbb{k}[x_0, \dots, x_n]$ , that is, the ring of polynomials on  $V^*$  with its standard grading.
- For each  $d \in \mathbb{N}$ ,  $S^d V$  is identified with the space of symmetric tensors in  $V^{\otimes d}$ ; a basis of  $S^d V$  is given by

$$\{\mathbf{x}^\alpha := x_0^{\alpha_0} \cdots x_n^{\alpha_n} \mid \alpha \in \mathbb{N}^{n+1}, |\alpha| = d\}.$$

- Let  $\mathbb{k}[x_{i,j} : i = 1, \dots, m, j = 0, \dots, n]$  denote the ring of polynomials in  $m(n+1)$  variables with the multigrading defined by  $\deg(x_{i,j}) = \underline{\varepsilon}_i := (0, \dots, 1_i, \dots, 0)$ .
- For  $\underline{d} \vdash d$ , let  $S^{\underline{d}} V := S^{d_1} V \otimes \cdots \otimes S^{d_m} V \subseteq V^{\otimes d}$  denote the space of partially symmetric tensors: this is isomorphic to the space of polynomials of multi-degree  $\underline{d}$ .
- Denote by  $\mathfrak{S}_d$  the symmetric group of permutations on a set of  $d$  elements;  $\mathfrak{S}_d$  acts on  $V^{\otimes d}$  by permuting the tensor factors.

In summary, we have the following diagram:

$$\begin{array}{ccccc}
 \text{Tensors} & \supseteq & \text{Partially Symmetric Tensors} & \supseteq & \text{Symmetric Tensors} \\
 V^{\otimes d} & & \text{invariant for } \mathfrak{S}_{d_1} \times \cdots \times \mathfrak{S}_{d_m} & & \text{invariant for } \mathfrak{S}_d \\
 & \Downarrow & & \Downarrow & \Downarrow \\
 \text{Multilinear} & \supseteq & \text{Multihomogeneous} & \longrightarrow & \text{Homogeneous} \\
 \text{Forms} & & \text{Polynomials} & & \text{Polynomials} \\
 & \Downarrow & \mathbb{k} \left[ x_{ij} : \begin{smallmatrix} i=1, \dots, m \\ j=0, \dots, n \end{smallmatrix} \right]_{\underline{d}} & & \mathbb{k}[x_0, \dots, x_n]_d
 \end{array}$$

The bottom right map is the surjective homomorphism defined by  $x_{ij} \mapsto x_j$ .

We give the following definition to formally introduce tensor decompositions respecting certain symmetries of a tensor.

**DEFINITION 1.1.** Let  $t \in S^{\underline{d}}V$ . A partially symmetric tensor decomposition of  $t$  is a sum of rank-one partially symmetric tensors such that

$$(1.1) \quad t = \sum_{i=1}^r v_{i,1}^{d_1} \otimes \cdots \otimes v_{i,m}^{d_m}, \quad \text{where } v_{i,j} \in V.$$

The smallest  $r$  such that a decomposition (1.1) exists is the partially symmetric rank of  $t$ . Equivalently, given a multihomogeneous polynomial  $f$  of multi-degree  $\underline{d}$ , a multihomogeneous decomposition, or  $\underline{d}$ -decomposition, of  $f$  is a sum

$$(1.2) \quad f = \sum_{i=1}^r \ell_{i,1}^{d_1} \cdots \ell_{i,m}^{d_m}, \quad \text{where } \deg(\ell_i) = \underline{e}_i.$$

The smallest  $r$  such that a decomposition (1.2) exists is the  $\underline{d}$ -rank of  $f$ , denoted  $R_{\underline{d}}(f)$ .

We will not distinguish between a multihomogeneous polynomial and the corresponding partially symmetric tensor. In particular, we always write  $f \in S^{\underline{d}}V$  for a multihomogeneous polynomial  $f$  of multidegree  $\underline{d}$ .

The space of symmetric tensors  $S^{\underline{d}}V$  is a subspace of  $S^{\underline{d}}(V)$  for any  $\underline{d} \vdash d$ . Therefore, for any  $\underline{d}$  and for any  $f \in S^{\underline{d}}V$ , we may ask the following.

**QUESTION 1.2.** Let  $f \in S^{\underline{d}}V$ , and let  $\underline{d} \vdash d$ . Is it true that  $R_d(f) = R_{\underline{d}}(f)$ ?

The original Comon's question corresponds to the case  $\underline{d} = (1, \dots, 1)$ . Note that in the cases where the original question has an affirmative answer, so does the partially symmetric version for any  $\underline{d}$ ; see Lemma 2.2. Equivalently, an example where Question 1.2 has a negative answer for some  $\underline{d}$  provides an example where the original Comon's question has negative answer as well. In this paper, we show instances where Question 1.2 has affirmative answer for some choice of  $\underline{d}$ , whereas the answer in the classical setting is not known. These are the cases of monomials and elementary symmetric polynomials.

Our approach is based on the study of *simultaneous Waring decompositions* of a collection of homogeneous polynomials. The problem of determining simultaneous ranks dates back to Terracini, see [Ter15]; some related problems were addressed more recently in [Fon02, AGMO18, CV18]. The *simultaneous Waring rank* of a collection of homogeneous polynomials is the minimum number of linear forms needed to simultaneously write a Waring decomposition for every polynomial in the collection. In this work, given a polynomial  $f$ , we consider the simultaneous rank of the collection of its partial derivatives of a given order.

DEFINITION 1.3. Let  $f \in S^d V$ , and let  $k < d$ . Let  $\nabla^k f$  be the set of partial derivatives of order  $k$  of  $f$ , i.e.,  $\nabla^k f = \{\frac{\partial^k f}{\partial \mathbf{x}^\alpha} : |\alpha| = k\}$ . The  $k$ th gradient rank of  $f$  is the simultaneous rank of  $\nabla^k f$ , i.e.,

$$R_{\nabla^k}(f) = \min \left\{ r : \exists \ell_1, \dots, \ell_r \in V \text{ such that } \frac{\partial^k f}{\partial \mathbf{x}^\alpha} = \sum_{i=1}^r c_{\alpha,i} \ell_i^{d-k} \text{ for some } c_{\alpha,i} \in \mathbb{k} \right\}.$$

If  $k = 1$ , write  $R_\nabla(f)$  for  $R_{\nabla^1}(f)$ .

Given  $f \in S^d V$ , for any  $\underline{d} \vdash d$  with  $d_m = d - k$ , we have the following chain of inequalities, which is proven in subsection 2.1 (see Lemma 2.2 and Corollary 2.5):

$$(1.3) \quad R_d(f) \geq R_{\underline{d}}(f) \geq R_{\nabla^k}(f).$$

In view of these inequalities, we will focus on the  $k$ th gradient with the following strategy: we show that for certain families of homogeneous polynomials, the  $k$ th gradient rank coincides with the Waring rank so that (1.3) is a chain of equalities.

We prove our results employing classical *apolarity theory* which dates back to Sylvester; see subsection 2.2. Briefly, apolarity relates the rank of a symmetric tensor  $f$  (respectively, the simultaneous rank of a family of polynomials  $f_1, \dots, f_s$ ) to the minimal cardinality of a set of points whose ideal is contained in the *apolar ideal* of  $f$  (respectively, the intersection of the apolar ideals of the  $f_i$ 's); see Lemma 2.9 and Lemma 2.11.

The notions of rank that we introduced have a *cactus* analogue which will be denoted by  $cR$  with the corresponding subscripts; precise definitions will be given in subsection 2.4. In terms of apolarity, this corresponds to studying *any* 0-dimensional scheme of minimal degree rather than just reduced sets of points. This terminology was introduced in [BB14, RS11], but it coincides with the notion of *scheme length* defined in [Iar95, Definition 4D]). For cactus ranks there is a chain of inequalities analogous to (1.3), and we will use the same strategy explained above to study Question 1.2 in this setting as well.

Similarly, one can consider the notion of border rank, which is the *upper semi-continuous closure* of the notion of rank: more precisely, the tensor border rank of  $t \in V^{\otimes d}$  is the minimum  $r$  such that  $t$  can be expressed as the limit of a sequence of  $r$  tensors; partially symmetric border rank is defined similarly. In [BGL13, BL13], some instances of Question 1.2 for border rank are considered.

**1.2. State of the art: old and new.** We list some known results of the various versions of Comon's question described in the previous section, and we present our main contributions.

Let  $f \in S^d V$ , let  $\underline{d} \vdash d$ , and let  $k = d - d_m$ .

- *Binary forms.* If  $\dim V = 2$ , i.e.,  $f$  is a binary form, then Question 1.2 has affirmative answer. The case  $\underline{d} = (1, \dots, 1)$  was proved in [ZH16, Corollary 3.12], and this implies an affirmative answer for any  $\underline{d}$  (see Lemma 2.2). Question 1.2 has affirmative answer also if considered for *cactus ranks*: indeed, in the case of binary forms the cactus rank coincides with the border rank, and, in this case, Question 1.2 admits a positive answer for border ranks (see [BGL13, Example 4.2.5]). In Proposition 3.1, we prove

$$\begin{aligned} R_{\nabla^k}(f) &= \min\{R_d(f), d - k + 1\}, \\ cR_{\nabla^k}(f) &= \min\{cR_d(f), d - k + 1\}. \end{aligned}$$

Consequently, if  $R_d(f) \leq d - k + 1$  (respectively,  $cR_d(f) \leq d - k + 1$ ), then we have  $R_d(f) = R_{\underline{d}}(f) = R_{\nabla^k}(f)$  (respectively,  $cR_d(f) = cR_{\underline{d}}(f) = cR_{\nabla^k}(f)$ ).

- *Ternary and quaternary cubics.* Let  $\dim V = 3$  or  $4$  and  $d = 3$ , i.e.,  $f$  defines a plane cubic curve or a cubic surface. Then, [Fri16, Theorem 7.1(4)] (for ternary cubics) and [Sei18, Theorem 1.3] (for quaternary cubics) prove that Question 1.2 has affirmative answer for  $\underline{d} = (1, \dots, 1)$ , and by Lemma 2.2 this implies an affirmative answer for any  $\underline{d}$ . [Sei18, Lemma 3.1] (for ternary cubics) and [Sei18, Theorem 1.5] (for quaternary cubics) prove that Question 1.2 has an affirmative answer when interpreted for border ranks: in the range of interest, border rank coincides with cactus rank (see [BB15, sections 3.5 and 3.6]); therefore, Question 1.2 has affirmative answer for cactus rank as well. In Corollary 2.5, we additionally prove

$$R_{1,2}(f) = R_{\nabla}(f).$$

In Proposition 3.4 (for ternary cubics) and in Proposition 3.6 (for quaternary cubics), we use a different method to prove the results described above, and we prove

$$cR_{1,2}(f) = cR_{\nabla}(f).$$

Our proof provides additional information on the relation between minimal Waring and multihomogeneous decompositions; see Remark 3.3 and Example 3.5.

- *Monomials.* Let  $f = x_0^{\alpha_0} \cdots x_n^{\alpha_n}$  be a monomial. We have the following:
  - (Theorem 3.8) if  $k \leq \min_i \{\alpha_i\}$ , then  $R_d(f) = R_{\underline{d}}(f) = R_{\nabla^k}(f)$ ;
  - (Theorem 3.12) if  $k = 1$ , then  $cR_d(f) = cR_{\underline{d}}(f) = cR_{\nabla}(f)$ .
- *Elementary symmetric forms of odd degree.* Let  $d$  be odd. In Theorem 3.16, we prove that if  $f = \sum_{i_1 < \dots < i_d} x_{i_1} \cdots x_{i_d}$  is the elementary symmetric polynomial of degree  $d$  in  $n+1$  variables, then  $R_d(f) = R_{1,d-1}(f) = R_{\nabla}(f)$ .

In addition, we provide a number of insights on different notions of rank and on their interplay. In particular, our approach suggests that classical apolarity theory is a valuable tool to study not only the symmetric rank of symmetric tensors but also their partially symmetric ranks; indeed, the multigraded versions of apolarity theory present in the literature (see, e.g., [Gal16, GRV18]), which are a priori more effective than the approach we are suggesting here, are generally more difficult to apply as they involve multigraded algebraic structures. Moreover, our approach proposes a systematic way to determine to what extent the symmetries of a tensor affect its rank. Moreover, by (1.3), an example providing a negative answer to Question 1.2 for certain  $\underline{d} \vdash d$  would be an example where the original Comon's question has a negative answer. Hence, we expect that studying the intermediate steps of the entire hierarchy of partially symmetric ranks would provide new examples where the original Comon's question has a negative answer, besides the one presented in [Shi18].

Besides a purely theoretical motivation, a better understanding of the problem posed in Question 1.2 can provide insights from a computational point of view. Tensor rank decomposition has a number of applications in pure and applied mathematics, and determining optimal decompositions is a concrete problem for many applications. Throughout the last decades several algorithms have been developed to find optimal and nearly optimal tensor decompositions for given tensors, both in the exact and in the approximate setting. We refer to [Hac12, Chapters 7 and 9] for a discussion on these algorithms and to [AR03, Com94, GSS05, Lin00] for some examples of applications. In principle, given a symmetric tensor for which Comon's question has affirmative answer, then an optimal symmetric decomposition has the same length as an optimal (not necessarily symmetric) tensor decomposition: in applications where one is not interested in the symmetries of the decomposition but only in its length, one can use interchangeably algorithms for either the symmetric or the nonsymmetric

setting to compute an optimal decomposition. Similar reasoning can be applied in the case of partially symmetric decompositions.

**1.3. Structure of the paper.** In section 2, we explain in more detail the different notions of rank we are going to consider, and we establish basic relations between them. In particular, we describe them in the framework of algebraic geometry. Moreover, we introduce algebraic tools from apolarity theory that we use in our computations. In section 3, we prove our main results.

**2. Different notions of rank and apolarity.** In this section, we introduce basic definitions of the various notions of rank that we consider, and we prove some relations among them. We also give the basics of apolarity theory which will be a fundamental tool for our approach.

Recall that  $S^d V$  is the subspace of  $V^{\otimes d}$  of symmetric tensors, namely, tensors which are invariant under the action of the symmetric group  $\mathfrak{S}_d$  that permutes the tensor factors. Similarly, given a composition  $\underline{d} = (d_1, \dots, d_m) \vdash d$ , the space  $S^{\underline{d}} V := S^{d_1} V \otimes \cdots \otimes S^{d_m} V$  is the subspace of partially symmetric tensors, namely, tensors which are invariant under the action of the subgroup  $\mathfrak{S}_{d_1} \times \cdots \times \mathfrak{S}_{d_m} \subseteq \mathfrak{S}_d$ , where  $V^{\otimes d} = \bigotimes_{j=1}^m V^{\otimes d_j}$  and  $\mathfrak{S}_{d_j}$  acts by permuting the tensor factors of  $V^{\otimes d_j}$ . In particular, a symmetric tensor may be regarded as a partially symmetric tensor, disregarding some of the additional symmetries. Hence, for any  $\underline{d} \vdash d$ , we have the inclusions

$$S^d V \subseteq S^{\underline{d}} V \subseteq V^{\otimes d}.$$

From a representation-theoretic point of view,  $S^d V$  is the Cartan component of  $S^{\underline{d}} V$  under the diagonal action of  $GL(V)$ .

Explicitly, for  $f \in S^d V$ , the *polarization* of  $f$  as a partially symmetric tensor in  $S^{\underline{d}}(V)$  (see, e.g., [Lan12, section 2.6.4]) is given by the expression

$$(2.1) \quad f = \frac{1}{d!} \sum_{\substack{\alpha_1, \dots, \alpha_{m-1} \in \mathbb{N}^{n+1} \\ |\alpha_i| = d_i}} \binom{d_1}{\alpha_1} \cdots \binom{d_{m-1}}{\alpha_{m-1}} \mathbf{x}^{\alpha_1} \otimes \cdots \otimes \mathbf{x}^{\alpha_{m-1}} \otimes \frac{\partial^{d_1 + \cdots + d_{m-1}}}{\partial \mathbf{x}^{\alpha_1} \cdots \partial \mathbf{x}^{\alpha_{m-1}}} f.$$

In the case  $\underline{d} = (1, d-1)$ , which will be particularly interesting to us, it reduces to

$$(2.2) \quad f = \frac{1}{d} \sum_{i=0}^n x_i \otimes \frac{\partial}{\partial x_i} f,$$

which can be interpreted as a tensorial version of Euler's formula.

**2.1. Ranks and projective varieties.** Now, we include the notions of rank introduced in Definition 1.1 into the geometric framework of so-called *X-ranks*. For any subset  $\mathcal{F} \subseteq \mathbb{P}^N$ , let  $\langle \mathcal{F} \rangle$  denote the linear span of  $\mathcal{F}$ , i.e., the smallest linear space containing  $\mathcal{F}$ .

**DEFINITION 2.1.** Let  $X \subseteq \mathbb{P}^N$  be a nondegenerate projective variety, and let  $p \in \mathbb{P}^N$ . The *X-rank* of  $p$ , denoted  $R_X(p)$ , is the minimal number of points of  $X$  whose linear span contains the point  $p$ , i.e., the minimal  $r$  such that  $p \in \langle q_1, \dots, q_r \rangle$  for some  $q_1, \dots, q_r \in X$ .

The notions of symmetric, tensor, and partially symmetric rank introduced in section 1 can be seen as *X-ranks* with respect to classical projective varieties such as Veronese, Segre, and Segre–Veronese varieties, respectively.

For  $\underline{d} \vdash d$ , the  *$d$ th Segre–Veronese embedding* is the map

$$\begin{aligned} \nu_{\underline{d}} : \mathbb{P} V \times \cdots \times \mathbb{P} V &\rightarrow \mathbb{P} (S^{d_1} V \otimes \cdots \otimes S^{d_m} V), \\ ([v_1], \dots, [v_m]) &\mapsto \left[ v_1^{d_1} \otimes \cdots \otimes v_m^{d_m} \right], \end{aligned}$$

where  $[v]$  denotes the class of a vector  $v \in V$  in the corresponding projective space. The image of the  $\underline{d}$ -Segre–Veronese embedding is called  $\underline{d}$ th Segre–Veronese variety. When  $\underline{d} = (d)$ ,  $\nu_d$  is the Veronese embedding of  $\mathbb{P}V$  and its image is the  $d$ th Veronese variety; when  $\underline{d} = (1, \dots, 1)$ ,  $\nu_{(1, \dots, 1)}$  is the Segre embedding of  $\mathbb{P}V^{\times d}$  and its image is the Segre variety. In particular, the  $\underline{d}$ -rank of an element  $f \in S^d V$  is the rank of  $[f]$  with respect to the  $\underline{d}$ -Segre–Veronese variety.

Fix two compositions  $\underline{d}, \underline{d}' \vdash d$ . Write  $\underline{d} \succeq \underline{d}'$  if  $\underline{d}'$  is a refinement of  $\underline{d}$ , in the sense that  $\underline{d}$  can be obtained from  $\underline{d}'$  by adding together some adjacent entries; more precisely, if  $\underline{d} = (d_1, \dots, d_m)$  and  $\underline{d}' = (d'_1, \dots, d'_{m'})$ , then  $\underline{d} \succeq \underline{d}'$  if and only if there exist  $0 = s_0 < s_1 < \dots < s_m = m'$  such that  $d_j = \sum_{i=s_{j-1}+1}^{s_j} d'_i$  for every  $j$ .

If  $\underline{d} \succeq \underline{d}'$ , then  $S^d V \subseteq S^{\underline{d}'} V$ . Moreover, directly from the definition of Segre–Veronese varieties, we have the following.

**LEMMA 2.2.** *Let  $f \in S^d V$  and  $\underline{d}, \underline{d}' \vdash d$  such that  $\underline{d} \succeq \underline{d}'$ . Then  $R_{\underline{d}}(f) \geq R_{\underline{d}'}(f)$ .*

*Proof.* If  $\underline{d} \succeq \underline{d}'$ , then  $\nu_{\underline{d}}(\mathbb{P}V^{\times m}) \subseteq \nu_{\underline{d}'}(\mathbb{P}V^{\times m'})$ . In particular, every set of points contained in  $\nu_{\underline{d}}(\mathbb{P}V^{\times m})$  whose linear span contains  $[f]$  is also a set of points contained in  $\nu_{\underline{d}'}(\mathbb{P}V^{\times m'})$  whose linear span contains  $[f]$ ; therefore, we obtain the desired inequality between the ranks.  $\square$

In fact, it is clear from its proof that Lemma 2.2 holds for every element of  $S^d V$ . Here, we only deal with elements of  $S^d V$ , namely, totally symmetric tensors. For this reason, the value  $R_{\underline{d}}(f)$  does not depend on the order of the entries of  $\underline{d}$ . Hence, one can consider an ordering similar to  $\succeq$  on the set of partitions of  $d$ , and correspondingly one has an analogue of Lemma 2.2. However, for the ease of notation, we keep working with compositions of integers rather than partitions.

The notion of simultaneous rank used in Definition 1.3 to define the gradient rank of  $f \in S^d V$  can be generalized to the setting of  $X$ -rank as well.

**DEFINITION 2.3.** *Let  $X \subseteq \mathbb{P}^N$  be a nondegenerate projective variety, and let  $\mathcal{F} \subseteq \mathbb{P}^N$  be a subset. The simultaneous  $X$ -rank of  $\mathcal{F}$ , denoted  $R_X(\mathcal{F})$ , is the minimal number of points on  $X$  whose linear span contains  $\mathcal{F}$ , i.e., the minimal  $r$  such that there exist  $q_1, \dots, q_r \in X$  with  $\mathcal{F} \subseteq \langle q_1, \dots, q_r \rangle$ , or equivalently  $\langle \mathcal{F} \rangle \subseteq \langle q_1, \dots, q_r \rangle$ .*

In this general setting, we provide several elementary facts which will give us some insight on the gradient rank. The following result shows that simultaneous  $X$ -rank of a set of points can be viewed as the rank of a unique point in a larger ambient space with respect to certain Segre variety  $\mathbb{P}^s \times X$ . The statement already appeared in [BL13, Theorem 2.5] in the setting of border rank. We include the proof to highlight that there is a one-to-one correspondence between sets of points on  $X$  providing a simultaneous decomposition for  $\mathcal{F}$  and sets of points on the Segre variety  $\mathbb{P}^s \times X$  providing a decomposition of the corresponding point. See also [Tei14, section 1.3].

**LEMMA 2.4** (see [BL13, Theorem 2.5]). *Let  $X \subseteq \mathbb{P}W$ ,  $\mathcal{F} = \{p_1, \dots, p_s\} \subseteq \mathbb{P}W$ , and fix  $w_1, \dots, w_s \in W$  such that  $p_i = [w_i] \in \mathbb{P}W$ . Let  $a_1, \dots, a_s$  be a basis of an  $s$ -dimensional vector space  $A$  and consider  $t = \sum_{i=1}^s a_i \otimes w_i \in A \otimes W$ . Then*

$$R_X(\mathcal{F}) = R_{\nu_{1,1}(\mathbb{P}A \times X)}([t]).$$

*Proof.* Suppose  $q_1, \dots, q_r$  are points of  $X$  such that  $\mathcal{F} \subseteq \langle q_1, \dots, q_r \rangle$ . Let  $z_1, \dots, z_r \in W$  such that  $[z_j] = q_j \in \mathbb{P}W$ . By definition,  $w_i = \sum_{j=1}^r \lambda_{ij} z_j$  for some scalars  $\lambda_{ij}$ .

Thus, we obtain

$$(2.3) \quad \begin{aligned} t &= \sum_{i=1}^s a_i \otimes w_i = \sum_{\substack{i=1, \dots, s \\ j=1, \dots, r}} a_i \otimes (\lambda_{ij} z_j) \\ &= \sum_{\substack{i=1, \dots, s \\ j=1, \dots, r}} (\lambda_{ij} a_i) \otimes z_j = \sum_{j=1}^r (\sum_{i=1}^s \lambda_{ij} a_i) \otimes z_j; \end{aligned}$$

hence,  $t$  is a linear combination of the  $r$  elements  $(\sum_i \lambda_{ij} a_i) \otimes z_j \in A \otimes W$  for  $j = 1, \dots, r$ . Taking the corresponding points on  $\nu_{1,1}(\mathbb{P}A \times X)$ , we get  $R_X(\mathcal{F}) \geq R_{\nu_{1,1}(\mathbb{P}A \times X)}([t])$ .

Conversely, suppose  $[t] \in \langle q_1, \dots, q_r \rangle$  for some  $q_j = [y_j \otimes z_j] \in \nu_{1,1}(\mathbb{P}A \times X)$ . Therefore, in the vector space  $A \otimes W$ , we have

$$(2.4) \quad t = \sum_{i=1}^s a_i \otimes w_i = \sum_{j=1}^r c_j y_j \otimes z_j \quad \text{for some } c_j \in \mathbb{k}.$$

Let  $b_1, \dots, b_s$  be the basis of  $A^*$  dual to  $a_1, \dots, a_s$ , i.e.,  $b_k(a_i) = \delta_{ik}$ . For every  $k = 1, \dots, s$ , apply  $b_k$  to both sides of the second equality in (2.4). Hence, we obtain  $w_k = \sum_{j=1}^r c_j b_k(y_j) z_j$ , which expresses every  $w_k$  as a linear combination of the  $r$  elements  $z_1, \dots, z_r$  with  $[z_j] \in X$ . This shows  $\mathcal{F} \subseteq \langle [z_1], \dots, [z_r] \rangle$ ; thus  $R_{\nu_{1,1}(\mathbb{P}A \times X)}(P) \geq R_X(\mathcal{F})$ . This concludes the proof.  $\square$

In the context of gradient rank, we deduce the following.

**COROLLARY 2.5.** *Let  $f \in S^d V$ . Then, for every  $\underline{d} \vdash d$  with  $d_m = d - k$ , we have*

$$(2.5) \quad R_{\underline{d}}(f) \geq R_{\nabla^k}(f).$$

In particular,

$$(2.6) \quad R_{(1,d-1)}(f) = R_{\nabla}(f).$$

*Proof.* Let  $\underline{d} = (d_1, \dots, d_{m-1})$  so that  $\underline{d} \vdash k$ . By Lemma 2.4, we have  $R_{\nabla^k}(f) = R_{\nu_{1,1}(\mathbb{P}A \times \nu_{d-k}(\mathbb{P}V))}([t])$ , where  $\dim A = \dim S^{\underline{d}} V$  and

$$t = \sum_{\substack{\underline{\alpha} = (\alpha_1, \dots, \alpha_{m-1}): \\ |\alpha_j| = d_j}} a_{\underline{\alpha}} \otimes \frac{\partial^k}{\partial \mathbf{x}^{\alpha_1 + \dots + \alpha_{m-1}}} f.$$

In particular, one can define an isomorphism  $A \simeq S^{\underline{d}} V$  by  $a_{\underline{\alpha}} \mapsto c_{\underline{\alpha}} \mathbf{x}^{\alpha_1} \otimes \dots \otimes \mathbf{x}^{\alpha_{m-1}}$ , where the  $c_{\underline{\alpha}}$ 's are suitable coefficients so that  $t$  coincides with  $f$  regarded as an element of  $S^{\underline{d}} V$  as presented in (2.1).

We have the inclusions

$$(2.7) \quad \nu_{\underline{d}}(\mathbb{P}V^{\times m}) \subseteq \nu_{1,d-k}(\mathbb{P}S^{\underline{d}} V \times \mathbb{P}V) = \nu_{1,1}(\mathbb{P}S^{\underline{d}} V \times \nu_{d-k}(\mathbb{P}V));$$

this shows  $R_{\underline{d}}([f]) \geq R_{\nu_{1,1}(\mathbb{P}S^{\underline{d}} V \times \nu_{d-k}(\mathbb{P}V))}([f])$ , and we obtain the inequality (2.5).

The second statement follows from the fact that when  $k = 1$ , we have  $\underline{d} = (1)$ , and therefore the first inclusion in (2.7) is an equality.  $\square$

The nonsymmetric analogue of the equality in (2.6) is a known characterization of tensor rank: given  $t \in V_1 \otimes \dots \otimes V_d$ , the simultaneous rank of the tensors  $\{t(\omega_1) \in V_2 \otimes \dots \otimes V_d \mid \omega_1 \in V_1^*\}$  is equal to the tensor rank of  $t$  (see [Fri13, Theorem 2.1], for  $d = 3$ , or [Lan12, Exercise 3.1.1.2]). We point out that the equality in (2.6) is a consequence of the fact that every element of  $\mathbb{P}V$  has rank one, because  $\nu_1$  is the identity map. When  $k \geq 2$ , this is no longer true, and indeed (2.5) can be a strict inequality, as shown in the following example.

*Example 2.6.* Let  $\dim V = n + 1$ . Let  $f \in S^3 V$  be any element with  $R_3(f) > n + 1$ , which exists for every  $n \geq 1$ . Then  $\langle \nabla^2 f \rangle \subseteq V$ , so  $R_{\nabla^2}(f) \leq n + 1$ , showing  $R_3(f) > R_{\nabla^2}(f)$ .

Lemma 2.2 and Corollary 2.5 establish the chain of inequalities in (1.3).

**2.2. Apolarity theory.** A classical approach to the Waring problem is based on *apolarity theory*, which is the study of the action of the ring of polynomial differential operators on the polynomial ring; see [IK99, Ger96]. In this section, we recall basic facts on classical apolarity for polynomials and its generalization to (partially symmetric) tensors and simultaneous ranks.

Given a vector space  $V$  with basis  $\{x_0, \dots, x_n\}$ , let  $\{y_0, \dots, y_n\}$  be its dual basis of  $V^*$ . The symmetric algebra  $S^\bullet V^*$  can be identified with the algebra of differential operators on  $x_0, \dots, x_n$  with constant coefficients by identifying  $y_j$  with  $\frac{\partial}{\partial x_j}$ . Hence, for every  $i, j$  with  $i \leq j$  there is a bilinear map

$$(2.8) \quad \circ : S^i V^* \times S^j V \rightarrow S^{j-i} V, \\ (\phi, f) \mapsto \phi \circ f := \phi \left( \frac{\partial}{\partial x_0}, \dots, \frac{\partial}{\partial x_n} \right) f(\mathbf{x}),$$

defined by differentiation. In particular, on the monomial basis, for any  $\alpha, \beta \in \mathbb{N}^{n+1}$  multi-indices with  $|\alpha| = j$  and  $|\beta| = i$ , we have

$$\mathbf{y}^\beta \circ \mathbf{x}^\alpha = \begin{cases} \frac{\alpha!}{(\alpha - \beta)!} \mathbf{x}^{\alpha - \beta} := \prod_{i=0}^n \frac{\alpha_i!}{(\alpha_i - \beta_i)!} x_i^{\alpha_i - \beta_i} & \text{if } \beta \leq \alpha, \text{ i.e., } \beta_i \leq \alpha_i \text{ for any } i; \\ 0 & \text{otherwise.} \end{cases}$$

Set  $S^j V = 0$  whenever  $j < 0$ , and extend this map via bilinearity to define the *apolar action* of  $S^\bullet V^*$  on  $S^\bullet V$ , which we still denote by  $\circ$ .

**DEFINITION 2.7.** Given  $f \in S^d V$ , the *apolar ideal* of  $f$  is the ideal in  $S^\bullet V^*$  of polynomial differential operators which annihilate  $f$ , i.e.,

$$\text{Ann}_d(f) := \{\phi \in S^\bullet V^* : \phi \circ f = 0\}.$$

The ideal  $\text{Ann}_d(f)$  is homogeneous and  $(\text{Ann}_d(f))_i = S^i V^*$  for  $i > d$ ; that is,  $\text{Ann}_d(f)$  is Artinian with socle degree  $d$ . The  $i$ th *catalecticant* of  $f$  is the linear map

$$(2.9) \quad \text{cat}_i(f) : S^i V^* \rightarrow S^{d-i} V, \\ \varphi \mapsto \varphi \circ f.$$

Note that  $(\text{Ann}_d(f))_i = \ker(\text{cat}_i(f))$  for every  $i$ .

**Remark 2.8.** We point out that apolar ideals of homogeneous polynomials are graded Artinian Gorenstein ideals. Moreover, *Macaulay's duality* provides a one-to-one correspondence between graded Artinian Gorenstein algebras of socle degree  $d$  and homogeneous polynomials of degree  $d$ ; see, e.g., [Ger96, Theorem 8.7] or [Eis95, section 21.2].

Together with the interpretation of  $S^\bullet V^*$  as a ring of polynomial differential operators, we have the natural structure of a ring of polynomials on  $V$ . In particular, homogeneous ideals in  $S^\bullet V^*$  define algebraic varieties and schemes in  $\mathbb{P}V$ . In this way, from the apolar ideal we may obtain Waring decompositions of  $f$  as follows.

**LEMMA 2.9** (apolarity lemma – classical version, [IK99, Lemma 1.15]). Let  $f \in S^d V$ . Let  $I_{\mathbb{X}} \subseteq S^\bullet V^*$  be the ideal defining a set of points  $\mathbb{X} = \{[\ell_1], \dots, [\ell_r]\} \subseteq \mathbb{P}V$ . Then the following are equivalent:

- (i)  $I_{\mathbb{X}} \subseteq \text{Ann}_d(f)$ ;
- (ii)  $f = \sum_{i=1}^r \lambda_i \ell_i^d$  for some  $\lambda_i \in \mathbb{k}$ .

If conditions (i) and (ii) hold, the set  $\mathbb{X}$  is said to be *apolar* to  $f$ .

Via the apolarity lemma, the problem of determining Waring ranks and Waring decompositions of a homogeneous polynomial can be approached by analyzing ideals of sets of points contained in its apolar ideal.

Note that condition (ii) of the apolarity lemma can be rephrased by saying that, in the same notation as the statement,  $[f] \in \langle \nu_d(\mathbb{X}) \rangle$ . In this form, the apolarity lemma holds more generally for possibly not reduced 0-dimensional schemes; see, e.g., [BJMR18, Lemma 1]. In particular, if  $\mathbb{X} \subseteq \mathbb{P}V$  is a 0-dimensional scheme, then  $I_{\mathbb{X}} \subseteq \text{Ann}_d(f)$  if and only if  $[f] \in \langle \nu_d(\mathbb{X}) \rangle$ , where the span of a 0-dimensional scheme is the zero set of the linear forms in its defining ideal.

Moreover, apolarity theory extends to partially symmetric tensors, and even more generally to the context of toric varieties; see, e.g., [Gał16, GRV18, Ven18]. For any  $\underline{d} \vdash d$ , the space  $S^{\underline{d}}V$  may be regarded as the multihomogeneous component of multidegree  $\underline{d}$  in the ring  $\mathbb{k}[x_{ij} : i = 1, \dots, m, j = 0, \dots, n]$ . In this setting, the apolar action is naturally multigraded and the apolar ideal of  $f \in S^{\underline{d}}V$  is multihomogeneous; denote it  $\text{Ann}_{\underline{d}}(f)$ . Recall that multihomogeneous ideals define algebraic varieties and schemes in  $\mathbb{P}V^{\times m}$ . From the toric version of the apolarity lemma, e.g., [GRV18, Lemma 1.8] or [Gał16, Proposition 3.8], the multihomogeneous analogue of Lemma 2.9 is as follows.

**LEMMA 2.10** (apolarity lemma – multigraded version). *Let  $f \in S^{\underline{d}}V$ . Let  $I_{\mathbb{X}}$  be the multihomogeneous ideal defining a set of points  $\mathbb{X} = \{q_1, \dots, q_r\} \subseteq \mathbb{P}V^{\times m}$  with  $q_j = ([\ell_{j,1}], \dots, [\ell_{j,m}])$ . Then the following are equivalent:*

- (i)  $I_{\mathbb{X}} \subseteq \text{Ann}_{\underline{d}}(f)$ ;
  - (ii)  $f = \sum_{i=1}^r \lambda_i \ell_{j,1}^{d_1} \otimes \cdots \otimes \ell_{j,m}^{d_m}$  for some  $\lambda_j \in \mathbb{k}$ .
- If conditions (i) and (ii) hold, the set  $\mathbb{X}$  is said to be  $\underline{d}$ -apolar to  $f$ .

Again, condition (ii) can be stated as  $[f] \in \langle \nu_{\underline{d}}(\mathbb{X}) \rangle$ , and Lemma 2.10 extends to the general case of 0-dimensional schemes. We use the expression  $\underline{d}$ -apolar in that setting as well.

It is easy to extend the apolarity lemma also in the case of simultaneous rank considering sets of points, or more generally 0-dimensional schemes, which are simultaneously apolar to a set of forms.

**LEMMA 2.11** (apolarity lemma – simultaneous version). *Let  $f_1, \dots, f_s \in S^{\underline{d}}V$ . Let  $\mathbb{X} = \{[\ell_1], \dots, [\ell_r]\} \subseteq \mathbb{P}V$  be a set of points defined by the ideal  $I_{\mathbb{X}}$ . Then, the following are equivalent:*

- (i)  $I_{\mathbb{X}} \subseteq \bigcap_{i=1}^s \text{Ann}_d(f_i)$ ;
- (ii)  $f_j = \sum_{i=1}^r \lambda_{j,i} \ell_i^d$  for some  $\lambda_{j,i} \in \mathbb{k}$ , for any  $j = 1, \dots, s$ .

Again, condition (ii) can be stated as  $\langle f_1, \dots, f_s \rangle \subseteq \langle \nu_d(\mathbb{X}) \rangle$ , and Lemma 2.11 extends to the case of 0-dimensional schemes, possibly nonreduced. Moreover, Lemma 2.11 extends to the case of forms of different degrees.

If  $\mathcal{F} \subseteq S^{\underline{d}}V$ ,  $\text{Ann}_d(\mathcal{F}) := \bigcap_{f \in \mathcal{F}} \text{Ann}_d(f)$  is called the simultaneous apolar ideal of  $\mathcal{F}$ . Proposition 2.24 will provide a characterization of  $\text{Ann}_{d-k}(\nabla^k f)$  for every  $f \in S^{\underline{d}}V$  and every  $k$ . This will be a fundamental tool for the rest of the paper.

**2.3. Hilbert functions of 0-dimensional schemes.** Given a homogeneous ideal  $I \subseteq S^{\bullet} V^*$ , the ideal  $I$  and the quotient algebra  $A_I := S^{\bullet} V^*/I$  inherit the grading of the polynomial ring. The *Hilbert function* of the quotient algebra  $A_I$  is the function which sends an integer  $i$  to the dimension, as a  $\mathbb{k}$ -vector space, of the component of degree  $i$  of  $A_I$ , i.e.,

$$(2.10) \quad \text{HF}(A_I; i) := \dim_{\mathbb{k}} (A_I)_i = \dim_{\mathbb{k}} S^i V^* - \dim_{\mathbb{k}} I_i.$$

If  $I = I_{\mathbb{X}}$  is the defining ideal of a 0-dimensional scheme  $\mathbb{X} \subseteq \mathbb{P}V$ ,  $\text{HF}_{\mathbb{X}} := \text{HF}(A_{I_{\mathbb{X}}}; -)$  denotes the Hilbert function of the corresponding graded algebra.

The Hilbert function of a 0-dimensional scheme  $\mathbb{X}$  is strictly increasing until it reaches a constant value; the value of this constant is the *degree* of  $\mathbb{X}$ , denoted  $\deg(\mathbb{X})$ . See [IK99, Theorem 1.69]. In the case where  $\mathbb{X}$  consists of  $s$  simple points, then  $\deg(\mathbb{X}) = s$ . We refer to [EH00] for the theory of 0-dimensional schemes.

**LEMMA 2.12.** *Let  $I_{\mathbb{Y}} \subseteq S^{\bullet}V^*$  be an ideal defining a 0-dimensional scheme  $\mathbb{Y} \subseteq \mathbb{P}V$ . Let  $H = \{\ell = 0\} \subseteq \mathbb{P}V$  be a hyperplane such that  $\mathbb{Y} \cap H = \emptyset$ . Then*

$$\deg(\mathbb{Y}) = \sum_{i \geq 0} \text{HF}(A_{I_{\mathbb{Y}} + (\ell)}; i).$$

*Proof.* By assumption,  $\ell$  is a nonzero divisor in  $A_{I_{\mathbb{Y}}}$ . Therefore, for every  $i$ , multiplication by  $\ell$  induces the exact sequence

$$(2.11) \quad 0 \longrightarrow (A_{I_{\mathbb{Y}}})_{i-1} \xrightarrow{\cdot \ell} (A_{I_{\mathbb{Y}}})_i \longrightarrow (A_{I_{\mathbb{Y}} + (\ell)})_i \longrightarrow 0.$$

This gives  $\text{HF}(A_{I_{\mathbb{Y}} + (\ell)}, i) = \text{HF}_{\mathbb{Y}}(i) - \text{HF}_{\mathbb{Y}}(i-1)$ ; namely, the Hilbert function of  $A_{I_{\mathbb{Y}} + (\ell)}$  is the *first difference* of the Hilbert function of  $\mathbb{Y}$ . Hence, for  $s \gg 0$ , we have  $\deg(\mathbb{Y}) = \text{HF}_{\mathbb{Y}}(s) = \sum_{i=0}^s (\text{HF}_{\mathbb{Y}}(i) - \text{HF}_{\mathbb{Y}}(i-1))$ , where  $\text{HF}_{\mathbb{Y}}(s)$  is written as a telescopic sum, using  $(S^{\bullet}V^*/I_{\mathbb{Y}})_j = 0$  if  $j < 0$ . We conclude

$$\deg(\mathbb{Y}) = \sum_{i \geq 0} \text{HF}(A_{I_{\mathbb{Y}} + (\ell)}; i). \quad \square$$

The latter results justify the following definition.

**DEFINITION 2.13.** *Let  $\mathbb{Y} \subseteq \mathbb{P}V$  be a 0-dimensional scheme. We call the first degree where the Hilbert function of  $\mathbb{Y}$  stabilizes the regularity index of  $\mathbb{Y}$ , denoted  $\text{reg}(\mathbb{Y})$ , i.e.,*

$$\text{reg}(\mathbb{Y}) := \min\{i : \text{HF}_{\mathbb{Y}}(i) = \deg(\mathbb{Y})\}.$$

**Remark 2.14.** By [IK99, Theorem 1.69], the regularity index of a 0-dimensional scheme is one less than the *Castelnuovo–Mumford regularity* of its defining ideal; for details, we refer to [Eis95, section 20.5]. We recall that the Castelnuovo–Mumford regularity of an ideal bounds from above the maximal degree of any minimal set of generators of the ideal: in particular,

$$\max \left\{ \deg(f_i) : \begin{array}{c} \text{for any minimal set of generators} \\ I_{\mathbb{Y}} = (f_1, \dots, f_s) \end{array} \right\} \leq \text{reg}(\mathbb{Y}) + 1.$$

**Remark 2.15.** If  $I \subseteq S^{\bullet}V^* \otimes \cdots \otimes S^{\bullet}V^*$  is a multihomogeneous ideal, then again it inherits the multigrading of the ring, and we define the *multigraded Hilbert function* of the corresponding quotient algebra analogously to (2.10) by considering any multidegree. From [SV06, Proposition 1.9], one observes that: if  $\mathbb{Y}$  is a 0-dimensional scheme in  $\mathbb{P}V^{\times m}$ , the multigraded Hilbert function of  $\mathbb{Y}$  is increasing and eventually constant in each direction, that is, for any  $i \in \{1, \dots, m\}$ :

- (i)  $\text{HF}_{\mathbb{Y}}(\underline{d}) \leq \text{HF}_{\mathbb{Y}}(\underline{d} + \underline{e}_i)$  for any  $\underline{d} \in \mathbb{N}^m$ ;
- (ii)  $\text{HF}_{\mathbb{Y}}(\underline{d}) = \text{HF}_{\mathbb{Y}}(\underline{d} + \underline{e}_i)$  then  $\text{HF}_{\mathbb{Y}}(\underline{d}) = \text{HF}_{\mathbb{Y}}(\underline{d} + 2\underline{e}_i)$  for any  $\underline{d} \in \mathbb{N}^m$ .

Moreover,  $\text{HF}_{\mathbb{Y}}(\underline{d}) \leq \deg(\mathbb{Y})$  for any  $\underline{d} \in \mathbb{N}^m$ , and equality holds if  $d_i \gg 0$  for all  $i$ .

**2.4. Cactus ranks.** Considering arbitrary 0-dimensional schemes suggests the definition of a more general notion of rank: the *cactus rank*. This was introduced in [IK99] in the setting of homogeneous polynomials with the name of *scheme length*. The terminology cactus rank, which is now the one commonly used in the literature, was introduced in [RS11, BR13, BB14].

**DEFINITION 2.16.** Let  $X \subseteq \mathbb{P}^N$  be a nondegenerate projective variety, and let  $p \in \mathbb{P}^N$ . The  $X$ -cactus rank of  $p$  is

$$cR_X(p) := \min \left\{ r : \begin{array}{l} \text{there exists a 0-dimensional scheme } \mathbb{Y} \subseteq X \\ \text{with } p \in \langle \mathbb{Y} \rangle \text{ and } \deg(\mathbb{Y}) = r \end{array} \right\}.$$

By the apolarity lemma, we can characterize the cactus rank with respect to Segre–Veronese varieties as follows. Let  $f \in S^d V$ . Then the cactus rank of  $[f]$  with respect to the  $\underline{d}$ th Segre–Veronese variety is

$$cR_{\underline{d}}(f) = \min \left\{ r : \begin{array}{l} \text{there exists a 0-dimensional scheme } \mathbb{X} \subseteq \mathbb{P}^{V \times m} \\ \text{d-apolar to } f \text{ with } \deg(\mathbb{X}) = r \end{array} \right\}.$$

The analogue of Lemma 2.2, with the same proof, holds for cactus rank as well.

**LEMMA 2.17.** Let  $f \in S^d V$  and  $\underline{d} \succeq \underline{d}'$ . Then  $cR_{\underline{d}}(f) \geq cR_{\underline{d}'}(f)$ .

Also simultaneous rank has a corresponding cactus version.

**DEFINITION 2.18.** Let  $X \subseteq \mathbb{P}^N$  be a nondegenerate projective variety, and let  $\mathcal{F} \subseteq \mathbb{P}^N$ . The simultaneous  $X$ -cactus rank of  $\mathcal{F}$ , denoted  $cR_X(\mathcal{F})$ , is the minimum  $r$  such that there exist  $\mathbb{Y} \subseteq X$  with  $\deg(\mathbb{Y}) = r$  and  $\mathcal{F} \subseteq \langle \mathbb{Y} \rangle$ , or equivalently  $\langle \mathcal{F} \rangle \subseteq \langle \mathbb{Y} \rangle$ .

However, a cactus analogue of Lemma 2.4 fails, as shown in Example 2.23.

As in the case of simultaneous rank, we are interested in relations between  $cR_{\underline{d}}(f)$  and the simultaneous rank of partial derivatives of  $f$ .

**DEFINITION 2.19.** Let  $f \in S^d V$ , and let  $k < d$  be a positive integer. The  $k$ th gradient cactus rank of  $f$  is the simultaneous cactus rank of  $\nabla^k f$  with respect to the  $(d-k)$ th Veronese variety; write

$$cR_{\nabla^k}(f) := cR_{\nu_{d-k}\mathbb{P}V}(\nabla^k f).$$

The following result gives a partial analogue of Lemma 2.4 in the case of cactus rank.

**LEMMA 2.20.** Let  $X \subseteq \mathbb{P}W$ ,  $\mathcal{F} = \{p_1, \dots, p_s\} \subseteq \mathbb{P}W$ , and fix  $w_1, \dots, w_s \in W$  such that  $p_i = [w_i] \in \mathbb{P}W$ . Let  $a_1, \dots, a_s$  be a basis of an  $s$ -dimensional vector space  $A$ , and consider  $t = \sum_{i=1}^s a_i \otimes w_i \in A \otimes W$ . Then

$$cR_X(\mathcal{F}) \leq cR_{\nu_{1,1}(\mathbb{P}A \times X)}([t]).$$

*Proof.* Let  $\pi : \mathbb{P}A \times \mathbb{P}W \rightarrow \mathbb{P}W$  be the projection onto the second factor. Let  $\mathbb{X} \subseteq \mathbb{P}A \times X$  be a 0-dimensional scheme such that  $[t] \in \langle \nu_{1,1}(\mathbb{X}) \rangle \subseteq \mathbb{P}(A \otimes W)$ . Let  $\mathbb{Y} = \pi(\mathbb{X}) \subseteq X$ . Then  $\deg(\mathbb{Y}) \leq \deg(\mathbb{X}) = \deg(\nu_{1,1}(\mathbb{X}))$ . We will show  $\mathcal{F} \subseteq \langle \mathbb{Y} \rangle$ .

It suffices to show that  $t \in A \otimes E$ , where  $E$  is defined by  $\mathbb{P}E := \langle \mathbb{Y} \rangle \subseteq \mathbb{P}W$ . Indeed,  $t \in A \otimes E$  implies that the image of the linear map  $t : A^* \rightarrow W$  is contained in  $E$ , namely,  $\langle w_1, \dots, w_s \rangle \subseteq E$ . In particular  $\mathcal{F} \subseteq \mathbb{P}E$ .

We have  $\mathbb{X} \subseteq \pi^{-1}(\mathbb{Y}) \subseteq \pi^{-1}(\mathbb{P}E) = \mathbb{P}A \times \mathbb{P}E$ . Applying  $\nu_{1,1}$  and passing to the linear spans,  $\langle \nu_{1,1}(\mathbb{X}) \rangle \subseteq \langle \nu_{1,1}(\mathbb{P}A \times \mathbb{P}E) \rangle = \mathbb{P}(A \otimes E)$ . Since  $[t] \in \langle \nu_{1,1}(\mathbb{X}) \rangle$  we obtain  $t \in A \otimes E$ , and we conclude.  $\square$

A consequence of Lemma 2.20 is the cactus analogue of Corollary 2.5. In particular, Example 2.23 shows that the analogue of (2.6) does not hold for cactus rank.

**COROLLARY 2.21.** *Let  $f \in S^d V$ . Then, for every  $\underline{d} \vdash d$  with  $d_m = d - k$ , we have*

$$cR_{\underline{d}}(f) \geq cR_{\nabla^k}(f).$$

*Proof.* This follows by the same argument of Corollary 2.5, using Lemma 2.20 instead of Lemma 2.4.  $\square$

The results of Lemma 2.11, Proposition 2.24, and Corollary 2.21 provide the following chain of inequalities, which is the cactus version of (1.3). For every  $\underline{d} \vdash d$  with  $d_m = d - k$ ,

$$(2.12) \quad cR_d(f) \geq cR_{\underline{d}}(f) \geq cR_{\nabla^k}(f).$$

We conclude this section by providing some insights on the relations between the ranks of the catalecticant maps, more generalized flattening maps, and the (partially symmetric) rank of a form  $f$ . See also [BGL13, Example 4.7]

*Remark 2.22.* By the apolarity lemma (Lemma 2.9),

$$\text{HF}_{\mathbb{X}}(i) \geq \text{HF}(A_{\text{Ann}_d(f)}; i) = \text{rank}(\text{cat}_i(f)) \quad \text{for any } i \in \mathbb{N},$$

for every  $f \in S^d V$ , and every  $\mathbb{X}$  apolar to  $f$ . In particular

$$R_d(f) \geq cR_d(f) \geq \max_{i=0,\dots,d} \{\text{HF}(A_{\text{Ann}_d(f)}; i)\}.$$

The maximal value of the Hilbert function of the quotient algebra of  $\text{Ann}_d(f)$  is sometimes referred to as *catalecticant lower bound* for  $cR_d(f)$ . Similar inequalities hold for the partially symmetric rank by considering the *multigraded* Hilbert function.

We observe that catalecticant lower bounds hold for  $cR_{(1^d)}(f)$ , where  $(1^d) = (1, \dots, 1)$ . More precisely, for every  $t \in V_1 \otimes \dots \otimes V_d$  and every subset  $I \subseteq \{1, \dots, d\}$ , there is an induced linear map, called a *flattening map*,

$$\text{flat}_I(t) : \bigotimes_{i \in I} V_i^* \rightarrow \bigotimes_{i \in I^c} V_i, \quad \text{where } I^c = \{1, \dots, d\} \setminus I,$$

defined by contraction; see, e.g., [Lan12, Chapter 2]. Its rank is the value of the Hilbert function of  $A_{\text{Ann}_{1^d}(t)}$  in multidegree  $e_i$ ; notice that  $A_{\text{Ann}_{(1^d)}(t)}$  is a quotient of the ring  $S^\bullet V_1 \otimes \dots \otimes S^\bullet V_d = S^\bullet(V_1 \oplus \dots \oplus V_d)$ . This rank is a lower bound for  $cR_{1^d}(t)$ . Now, if  $f \in S^d V$ , then one has that  $\text{rank}(\text{cat}_i(f)) = \text{rank}(\text{flat}_I(f))$  for every  $I$  with  $|I| = i$ . In conclusion, the catalecticant lower bound is indeed a lower bound for  $cR_{1^d}(f)$ .

More generally, some generalized flattening maps for  $f \in S^d V$ , naturally providing lower bounds for  $R_d(f)$ , give lower bounds for  $R_{1^d}(f)$  and, by [Gal17], for  $cR_{1^d}(f)$  as well. We observe this fact for Koszul flattenings [LO13]: given  $f \in S^d V$ , define  $\text{flatKos}_i^{\wedge p}(f) : S^i V^* \otimes \Lambda^p V \rightarrow S^{d-i-1} \otimes \Lambda^{p+1} V$  to be the composition

$$S^i V^* \otimes \Lambda^p V \xrightarrow{\text{cat}_i(f) \otimes \text{id}_{\Lambda^p V}} S^{d-i} V \otimes \Lambda^p V \longrightarrow S^{d-i-1} V \otimes \Lambda^{p+1} V,$$

where the second map is the Koszul differential. From [LO13, Proposition 4.1.1],

$$R_d(f) \geq \frac{\text{rank}(\text{flatKos}_i^{\wedge p}(f))}{\text{rank}(\text{flatKos}_i^{\wedge p}(x_0^{\otimes d}))} = \frac{\text{rank}(\text{flatKos}_i^{\wedge p}(f))}{\binom{n}{p-1}}$$

if  $\dim V = n + 1$ . In the nonsymmetric setting, for  $t \in V^{\otimes d}$ , one defines a Koszul flattening in a similar way, as an *augmentation* of  $\text{flat}_I$ :

$$\text{flatKos}_I^{j,\wedge p}(t) : \bigotimes_I V^* \otimes \Lambda^p V \xrightarrow{\text{flat}_I \otimes \text{id}_{\Lambda^p V}} \bigotimes_{I^c} V \otimes \Lambda^p V \longrightarrow \bigotimes_{I^c \setminus \{j\}} V \otimes \Lambda^{p+1} V.$$

This provides the lower bound

$$R_{1^d}(t) \geq \frac{\text{rank}(\text{flatKos}_I^{j,\wedge p}(t))}{\text{rank}(\text{flatKos}_I^{j,\wedge p}(x_0^{\otimes d}))} = \frac{\text{rank}(\text{flatKos}_I^{j,\wedge p}(t))}{\binom{n}{p-1}}.$$

Analogously to the standard flattening case, for  $f \in S^d V$ , one obtains that

$$\text{rank}(\text{flatKos}_I^{j,\wedge p}(f)) = \text{rank}(\text{flatKos}_i^{\wedge p}(f))$$

for every set of indices  $I$  with  $|I| = i$ ; therefore, the Koszul flattening lower bound for  $f$  holds for  $R_{1^d}(f)$ . By the results of [Ga17], these bounds hold for cactus rank as well.

**2.5. Consequences of apolarity theory.** In this section, we provide some immediate consequences of the theory introduced in subsections 2.1 and 2.3. The main result of this section is Proposition 2.24 which gives an explicit description of the simultaneous apolar ideal of the set of partial derivatives of a given order of a form  $f$ .

**2.5.1. Failure of Lemma 2.4 for cactus rank.** First, we provide an example showing that the simultaneous cactus rank of a family of forms cannot be read as the cactus rank of tensor in an bigger space, unlike what happens for the classical rank in Lemma 2.4.

*Example 2.23.* Consider  $\mathcal{F} = \{x_0^2 x_1, x_0^2 x_2\} \subseteq S^3 V$  with  $\dim V = 3$ . The apolar ideal is  $\text{Ann}_3(\mathcal{F}) = (y_0^3, y_1^2, y_2^2, y_1 y_2)$  whose Hilbert function in degree 2 is equal to 3, and, therefore, any 0-dimensional scheme apolar to  $\mathcal{F}$  has length at least 3. Indeed,  $\text{cR}_3(\mathcal{F}) = 3$  since the 2-fat point supported at  $[x_0] \in \mathbb{P}V$ , i.e., the 0-dimensional scheme of degree 3 defined by  $(y_1, y_2)^2$ , is apolar to  $\mathcal{F}$ .

Now, consider the partially symmetric tensor  $t = a_0 \otimes x_0^2 x_1 + a_1 \otimes x_0^2 x_2 \in A \otimes S^3 V$  with  $\dim A = 2$ ,  $A = \langle a_0, a_1 \rangle$ . We prove that  $\text{cR}_{(1,3)}(t) \geq 4$ . The bigraded apolar ideal of  $t$  is

$$\text{Ann}_{(1,3)}(t) = (S^2 A^*) + (b_0 y_2, b_1 y_1, b_0 y_1 - b_1 y_2, y_2^2, y_1 y_2, y_1^2, y_0^3) \subseteq S^\bullet A^* \otimes S^\bullet V^*,$$

where  $\{b_0, b_1\}$  and  $\{y_0, y_1, y_2\}$  are the bases of  $A^*$  and  $V^*$  dual to  $\{a_0, a_1\}$  and  $\{x_0, x_1, x_2\}$ , respectively. The bigraded Hilbert function of the quotient algebra  $A_{\text{Ann}_{(1,3)}(t)}$  is

$A^* \setminus V^*$	0	1	2	3	4
0	1	3	3	2	—
1	2	3	3	1	—
2	—	—	—	—	—

Therefore,  $\text{cR}_{(1,3)}(t) \geq 3$ . Assume that there exists a 0-dimensional scheme  $\mathbb{Y} \subseteq \mathbb{P}A \times \mathbb{P}V$  of length 3 apolar to  $t$ . Then,

$$(I_{\mathbb{Y}})_{(1,2)} = ((\text{Ann}_{(1,3)}(t))_{(1,2)}) = (b_0 y_2, b_1 y_1, b_0 y_1 - b_1 y_2, y_2^2, y_1 y_2, y_1^2) =: J.$$

In particular,  $I_{\mathbb{Y}} \supseteq J$ . Since  $I_{\mathbb{Y}}$  is the ideal of a 0-dimensional scheme in  $\mathbb{P}A \times \mathbb{P}V$ , it is saturated with respect to the bigraded irrelevant ideal  $\mathfrak{m} = (b_0, b_1) \cdot (y_0, y_1, y_2)$ . Let  $J^{sat} = J : \mathfrak{m}^\infty$  be the saturation of  $J$ . We deduce  $I_{\mathbb{Y}} \supseteq J^{sat} = (y_1, y_2)$ . This shows  $\mathbb{Y} \subseteq \mathbb{P}A \times (1 : 0 : 0)$ . This is a contradiction because  $\nu_{(1,3)}(\mathbb{Y}) \subseteq \nu_{(1,3)}(\mathbb{P}A \times (1 : 0 : 0))$  and the latter is a line not containing the tensor  $t$ .

**2.5.2. Structure of simultaneous apolar ideal.** The next result computes the simultaneous apolar ideal of  $\nabla^k f$  for a given  $f \in S^d V$  which, via apolarity theory, will be of key importance for our computations.

PROPOSITION 2.24. *Let  $f \in S^d V$ , and let  $k \geq 0$ . For every  $i \geq 0$ ,*

$$(\text{Ann}_{d-k}(\nabla^k f))_i = \bigcap_{|\alpha|=k} \left( \text{Ann}_{d-k}\left(\frac{\partial^k}{\partial \mathbf{x}^\alpha} f\right) \right)_i = \begin{cases} (\text{Ann}_d(f))_i & \text{if } 0 \leq i \leq d-k, \\ S^i V^* & \text{if } i \geq d-k+1. \end{cases}$$

*Proof.* For  $i \geq d-k+1$  the statement follows simply because  $k$ th partial derivatives of  $f$  have degree  $d-k$ .

For  $i \leq d-k$ , the statement is a consequence of the fact that differential operators commutes. For every  $\phi \in S^i V^*$ , we have

$$(2.13) \quad \phi \circ \frac{\partial^\alpha}{\partial \mathbf{x}^\alpha} f = \phi \circ (\mathbf{y}^\alpha \circ f) = \mathbf{y}^\alpha \circ (\phi \circ f).$$

If  $\phi \in (\text{Ann}_d(f))_i$ , the right-hand side of (2.13) is 0, showing that the left-hand side is 0 for every  $\alpha$ , and therefore  $\phi \in (\text{Ann}_{d-k}(\nabla^k f))_i$ . Conversely, if  $\phi \in (\text{Ann}_{d-k}(\nabla^k f))_i$ , then  $\phi \in \text{Ann}_{d-k}(\frac{\partial^k}{\partial \mathbf{x}^\alpha} f)$  for every  $\alpha$ ; therefore the left-hand side of (2.13) is 0, which implies that the right-hand side is 0; in this case we deduce that  $\phi \circ f$  is a homogeneous polynomial of degree  $k$  which is annihilated by all differential operators of order  $k$ . Since the apolarity pairing is nondegenerate, we conclude that  $\phi \circ f = 0$ .  $\square$

An immediate consequence of Proposition 2.24 is the following fact.

Remark 2.25. Let  $f \in S^d V$  and  $k \geq 0$  as in Proposition 2.24. Let  $\mathbb{X} \subseteq \mathbb{P}V$  be a 0-dimensional scheme such that  $I_{\mathbb{X}}$  is generated in degree at most  $d-k$ . Then  $\mathbb{X}$  is apolar to  $f$  if and only if it is apolar to  $\nabla^k f$ .

**2.5.3. Sylvester's theorem for binary forms.** As a first explicit example of application of apolarity theory to compute ranks of homogeneous polynomials, we recall Sylvester's theorem which completely describes the Waring decompositions in the case of binary forms [Syl52].

Let  $\dim V = 2$ . One can prove that if  $f \in S^d V$ , then  $\text{Ann}_d(f) = (g_1, g_2)$ , where  $\deg(g_i) = e_i$  and  $e_1 + e_2 = d+2$ ; this is a consequence of the general theory, and more precisely of the fact that Gorenstein algebras of codimension 2 are always complete intersection and that Artinian Gorenstein algebras have symmetric Hilbert function, i.e., for  $i = 0, \dots, d$ ,  $\text{HF}(A_{\text{Ann}_d(f)}; i) = \text{HF}(A_{\text{Ann}_d(f)}; d-i)$ ; see [Ger96, Proposition 8.6]. Hence, let  $e_1 \leq e_2$ .

Recall that 0-dimensional schemes in  $\mathbb{P}^1$  are defined by principal ideals. Hence, if  $g_1$  has distinct roots, we conclude that  $R_d(f) = e_1$  and a minimal set of points apolar to  $f$  is given by the roots of  $g_1$ ; moreover, if  $e_1 < e_2$ , this is the unique minimal set of points apolar to  $f$ . If  $g_1$  does not have distinct roots, then a minimal set of points apolar to  $f$  is given by the roots of  $g_1 h + g_2$  for a generic choice of  $h \in S^{e_2-e_1} V$ . For an exposition of Sylvester's theorem in modern terminology we refer to [CS11].

**THEOREM 2.26** (Sylvester's theorem). *Let  $f \in S^d V$  with  $\dim_{\mathbb{K}} V = 2$ . Let  $\text{Ann}_d(f) = (g_1, g_2)$  with  $\deg(g_1) \leq \deg(g_2)$ . Then*

$$R_d(f) = \begin{cases} \deg(g_1) & \text{if } g_1 \text{ has distinct roots;} \\ \deg(g_2) & \text{otherwise.} \end{cases}$$

As for cactus rank, with the same notation as above, one has  $cR_d(f) = e_1$ . Indeed,  $(g_1)$  always defines a 0-dimensional scheme of degree  $e_1$  apolar to  $f$ , and there are no apolar schemes of smaller degree since there are no elements of smaller degree in the apolar ideal of  $f$ . If  $e_1 < e_2$ , then the 0-dimensional scheme defined by  $g_1$  is the unique minimal 0-dimensional scheme apolar to  $f$ .

**3. Computations.** In this section, we prove our main results. We consider special families of symmetric tensors, and we study their  $k$ th gradient (cactus) ranks. As we already explained, we will focus mostly on the cases where the inequalities of (1.3) become equalities.

**3.1. Binary forms.** In this section, we obtain a complete result on the gradient ranks and gradient cactus ranks of binary forms.

**PROPOSITION 3.1.** *Let  $d \in \mathbb{N}$  and  $f \in S^d V$  with  $\dim V = 2$ . Then, for any  $k < d$ ,*

$$R_{\nabla^k}(f) = \min\{R_d(f), d - k + 1\} \quad \text{and} \quad cR_{\nabla^k}(f) = \min\{cR_d(f), d - k + 1\}.$$

Consequently, for any  $\underline{d} \vdash d$  with  $d_m = d - k$ , we have the following:

- (i) if  $R_d(f) \leq d - k + 1$ , then  $R_d(f) = R_{\underline{d}}(f) = R_{\nabla^k}(f)$ ;
- (ii) if  $cR_d(f) \leq d - k + 1$ , then  $cR_d(f) = cR_{\underline{d}}(f) = cR_{\nabla^k}(f)$ .

*Proof.* By definition,  $R_{\nabla^k}(f) \leq R_d(f)$  and  $cR_{\nabla^k}(f) \leq cR_d(f)$ . For  $R_d(f) < d - k + 1$  (respectively,  $cR_d(f) < d - k + 1$ ), we conclude by Remark 2.25. Conversely, suppose  $R_d(f) \geq d - k + 1$  (respectively,  $cR_d(f) \geq d - k + 1$ ). Since  $(\text{Ann}_{d-k}(\nabla^k f))_{d-k+1} = S^{d-k+1}V^*$  by Proposition 2.24, any square-free element (respectively, any element) of  $S^{d-k+1}V^*$  defines a set of  $d - k + 1$  points (respectively, a 0-dimensional scheme of degree  $d - k + 1$ ) in  $\mathbb{P}V$  apolar to  $f$ . This implies  $R_{\nabla^k}(f) \leq d - k + 1$  (respectively,  $cR_{\nabla^k}(f) \leq d - k + 1$ ). Again, the lower bound follows by Remark 2.25. The second part of the statement follows from the first one by the chain of inequalities (1.3).  $\square$

*Remark 3.2.* In [ZH16, Corollary 3.12], the authors proved that the original Comon's question (Question 1.2 for  $\underline{d} = 1^d \vdash d$ ) has an affirmative answer in the case of binary forms. Therefore, by Lemma 2.2, it follows that Question 1.2 has an affirmative answer for any  $\underline{d} \vdash d$ . In fact, by (2.6) in Corollary 2.5, this implies the part (i) of the statement in Proposition 3.1 in the case  $k = 1$ .

*Remark 3.3.* The proof of Theorem 3.8 gives interesting insights on minimal schemes apolar to the  $k$ th gradient of a binary form and, in particular, on their relations with minimal schemes apolar to the form itself. Here, we resume some observations:

- (i) if  $R_d(f) < d - k + 1$  (respectively,  $cR_d(f) < d - k + 1$ ), the minimal reduced (respectively, not necessarily reduced) 0-dimensional schemes apolar to  $f$  are the same as the ones minimally simultaneously spanning  $\nabla^k f$ . Note that for  $R_d(f) < \frac{d+1}{2}$  (respectively,  $cR_d(f) < \frac{d+1}{2}$ ), such a reduced (respectively, not necessarily reduced) 0-dimensional scheme is unique by Sylvester's theorem (Theorem 2.26) (respectively, for the comments on cactus ranks of binary forms at the end of subsection 2.5.3);

- (ii) if  $R_d(f) = d - k + 1$ , we have that the rank  $R_d(f)$  and the gradient rank  $R_{\nabla^k}(f)$  are the same, but we can find minimal schemes apolar to  $\nabla^k f$  which are not apolar to  $f$  itself. For example,  $x_0 x_1^{d-1}$  has rank  $d$ , and any minimal apolar set of  $d$  points does not involve the point  $[x_1] \in \mathbb{P}V$ ; see [CCO17, section 3.2]. However, if we consider the first partial derivatives  $\nabla f = \{x_1^{d-1}, x_0 x_1\}$ , we have that the set of points  $\mathbb{X} = \{[x_1]\} \cup \{[x_0 + \xi x_1] : \xi^{d-1} = 1\}$  are apolar to  $\nabla f$ ; indeed,  $I_{\mathbb{X}} = (y_0(y_0^{d-1} - y_1^{d-1}))$  which is contained in  $\text{Ann}_{d-1}(\nabla f) = (y_0^2, y_1^d, y_0 y_1^{d-1})$ .
- (iii) More generally, if  $R_d(f) \geq d - k + 1$  (respectively,  $cR_d(f) \geq d - k + 1$ ), then  $R_{\nabla^k}(f) = d - k + 1$  (respectively,  $cR_d(f) = d - k + 1$ ): in such a case, any set of  $d - k + 1$  points (respectively, any 0-dimensional scheme of degree  $d - k + 1$ ) is apolar to  $\nabla^k f$ . Indeed, such a scheme is defined by a principal ideal whose generator has degree  $d - k + 1$ , and, therefore, it is contained in the apolar ideal of the  $k$ th gradient of  $f$  because, by Proposition 2.24,  $(\text{Ann}_{d-k}(\nabla^k f))_{d-k+1} = S^{d-k+1}V^*$ .

**3.2. Ternary and quaternary cubics.** Comon's question in the case of cubic forms in three or four variables, that is,  $f \in S^3V$  with  $\dim V = 3, 4$  has an affirmative answer: the proof exploits the fact that in these two cases it is possible to classify the orbits under the action of the group  $GL(V)$ : see [Fri16, Theorem 7.1(4)] for three variables and [Sei18, Theorem 1.3] for four variables. The statement for cactus rank follows from [Sei18, Lemma 3.1 and Theorem 1.5], which prove that the Question 1.2 has affirmative answer for border rank in the case of ternary and quaternary cubics, respectively, and from [BB15, sections 3.5 and 3.6], which guarantee that in these cases border rank coincides with cactus rank.

Proposition 3.4 uses the techniques developed in section 2 to recover the result in three variables and proves additionally the equality  $cR_{1,2}(f) = cR_{\nabla}(f)$ . Proposition 3.6 proves the equality  $cR_{1,2}(f) = cR_{\nabla}(f)$  in the case with four variables.

PROPOSITION 3.4. *Let  $f \in S^3V$  with  $\dim V = 3$ . Then*

$$\begin{aligned} R(f) &= R_{1,2}(f) = R_{\nabla}(f) \quad \text{and} \\ cR(f) &= cR_{1,2}(f) = cR_{\nabla}(f). \end{aligned}$$

*Proof.* If the first catalecticant of  $f$  is not full-rank, then there is a choice of coordinates such that  $f$  can be written in fewer variables; in this case  $f$  is a binary form, and the statement follows from Proposition 3.1.

Hence, assume that the first catalecticant of  $f$  is full-rank, which implies that rank and cactus rank of  $f$  are at least 3. Therefore, if  $f$  has rank 3 (cactus rank 3, respectively), then the claim directly follows.

Let  $f$  have rank 4 (cactus rank 4, respectively), and suppose that  $\nabla f$  has an apolar reduced (not necessarily reduced, respectively) 0-dimensional scheme  $\mathbb{X}$  with  $\deg(\mathbb{X}) = 3$ . By Proposition 2.24, we have that  $\text{HF}(A_{\text{Ann}_2(\nabla f)}; 1) = 3$ , which implies that  $\mathbb{X}$  is not contained in a line. Since  $\deg(\mathbb{X}) = 3$ , the ideal  $I_{\mathbb{X}}$  is generated by three quadrics, so  $I_{\mathbb{X}}$  is generated in degree 2. By Proposition 2.24, we deduce  $I_{\mathbb{X}} \subseteq \text{Ann}_3(f)$ , contradicting the assumption that  $f$  has rank (cactus rank, respectively) 4.

The cactus rank of plane cubics is at most 4 (see, e.g., [BB15, section 3.5]), so the second part of the statement is proved. The rank of plane cubics is at most 5 and there is a unique form of rank 5 up to change of coordinates, which is  $f = x_0(x_0 x_1 + x_2^2)$ ; see, for instance, [LT10]. Suppose  $R_{\nabla}(f) \leq 4$ , and let  $\mathbb{X}$  be a set of four points apolar to  $\nabla f$ . The set  $\mathbb{X} \subseteq \mathbb{P}V = \mathbb{P}^2$  may have two possible configurations: either the points in  $\mathbb{X}$  are in general linear position, or three of them lie on a line

$\ell$ . In the first case, two conics generate  $I_{\mathbb{X}}$  so  $I_{\mathbb{X}} \subseteq \text{Ann}_3(f)$  which contradicts that  $R_{\nabla}(f) = 4$ . In the second case, one can easily show that  $I_{\mathbb{X}}$  cannot be radical, which is a contradiction.  $\square$

Even though the ranks coincide, simultaneous decompositions of the gradient of a plane cubic do not always come from decompositions of the cubic itself. Indeed, as already observed in the case of binary forms (Remark 3.3(ii)), sometimes it is possible to construct a simultaneous decomposition of the gradient which contains some of the forbidden points (in the sense of [CCO17]) of the original form.

*Example 3.5.* Let  $f = x_0(x_0x_1 + x_2^2)$  be the unique plane cubic of maximal rank up to choice of coordinates; namely,  $R_3(f) = 5$ . By [CCO17, Theorem 3.18], there are no minimal Waring decompositions of  $f$  involving  $x_0^3$ , or equivalently if  $\mathbb{X}$  is a set of 5 points apolar to  $f$ , then  $[x_0] \notin \mathbb{X}$ . Consider the set of points  $\mathbb{Y}$  defined by

$$I_{\mathbb{Y}} = (y_1y_2, y_0^2y_2 + y_0y_2^2 + y_2^3, y_0^2y_1 - y_0y_1^2);$$

one can check that  $I_{\mathbb{Y}} \subseteq \text{Ann}_2(\nabla f)$  and

$$\mathbb{Y} = \{[x_0], [x_1], [x_0 - x_1], [(\omega + 1)x_0 - 2x_2], [(-\omega + 1)x_0 - 2x_2]\} \subseteq \mathbb{P}V,$$

where  $\omega^2 + 3 = 0$ . Explicitly, we have

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2x_0x_1 + x_2^2 = x_1^2 - (x_0 - x_1)^2 + 2x_0^2 + \frac{\omega+3}{24}((\omega + 1)x_0 - 2x_2)^2 + \\ &\quad + \frac{3-\omega}{24}((- \omega + 1)x_0 - 2x_2)^2; \\ \frac{\partial f}{\partial y} &= x_0^2 = x_0^2; \\ \frac{\partial f}{\partial z} &= 2x_0x_2 = x_0^2 + \frac{\omega}{12}((\omega + 1)x_0 - 2x_2)^2 - \frac{\omega}{12}((- \omega + 1)x_0 - 2x_2)^2; \end{aligned}$$

This shows that  $\mathbb{Y}$  defines a simultaneous decomposition of  $\nabla f$  containing the point  $[x_0]$  which is forbidden for  $f$ .

**PROPOSITION 3.6.** *Let  $f \in S^3V$  with  $\dim V = 4$ . Then*

$$\text{cR}_3(f) = \text{cR}_{2,1} = \text{cR}_{\nabla}(f).$$

*Proof.* Recall that  $\text{cR}_3(f) \leq 5$  (see, e.g., [BB15]). If the first catalecticant of  $f$  is not full-rank, then there is a choice of coordinates such that  $f$  can be written in fewer variables; in this case, the result follows from Proposition 3.4. Therefore suppose that the first catalecticant is full-rank, or equivalently  $\text{HF}(A_{\text{Ann}_3(f)}; 1) = 4$ .

Let  $\mathbb{X}$  be a 0-dimensional scheme apolar to  $\nabla f$  so that by apolarity  $I_{\mathbb{X}} \subseteq \text{Ann}_2(\nabla f)$ . Since  $(\text{Ann}_3(f))_1 = (\text{Ann}_3(\nabla f))_1$ , we obtain the lower bound

$$\text{cR}_{\nabla}(f) \geq \text{HF}(A_{\text{Ann}_3(f)}; 1) = 4,$$

and therefore  $\deg(\mathbb{X}) \geq 4$ , providing the result whenever  $\text{cR}_3(f) \leq 4$ .

If  $\deg(\mathbb{X}) = 4$ , by Remark 2.14 the ideal  $I_{\mathbb{X}}$  is generated by quadrics, and therefore  $\mathbb{X}$  is apolar to  $f$  because  $(\text{Ann}_3(f))_2 = (\text{Ann}_3(\nabla f))_2$  by Proposition 2.24. This shows that if  $\text{cR}_3(f) = 5$ , and  $\mathbb{X}$  is apolar to  $\nabla f$ , then  $\deg(\mathbb{X}) \geq 5$ . This concludes the proof.  $\square$

**3.3. Monomials.** We consider the case of monomials. Recall the result on Waring rank.

THEOREM 3.7 (see [CCG12]). *Let  $f = \mathbf{x}^\alpha$  with  $\alpha_0 = \min_i\{\alpha_i\}$ . Then*

$$R_d(f) = \frac{1}{\alpha_0 + 1} \prod_{i=0}^n (\alpha_i + 1).$$

Our first goal is to establish that the rank of a monomial coincides with the  $k$ th gradient rank for  $k$  at most as large as the minimal exponent appearing in the monomial.

THEOREM 3.8. *Let  $d \in \mathbb{N}$  and  $k < d$ . Let  $f = \mathbf{x}^\alpha$  be a monomial with  $k \leq \alpha_0 = \min_i\{\alpha_i\}$  and  $|\alpha| = d$ . Then, for any  $\underline{d} \vdash d$  with  $d_m = d - k$ , we have*

$$R_d(f) = R_{\underline{d}}(f) = R_{\nabla^k}(f).$$

*Proof.* By (1.3), it is enough to show  $R_{\nabla^k}(f) \geq R_d(f)$ . If  $\alpha_0 > k$ , consider  $y_0^k \circ f \in \nabla^k f$ . We have  $y_0^k \circ f = x_0^{\alpha_0-k} x_1^{\alpha_1} \cdots x_n^{\alpha_n}$  with  $\alpha_0 - k > 0$ . Therefore, by Theorem 3.7,  $R_{d-k}(y^k \circ f) = R_d(f)$ . In particular  $R_{\nabla^k}(f) \geq R_{d-k}(y_0^k \circ f) = R_d(f)$ , and we conclude.

Assume  $\alpha_0 = k$ . Let  $\mathbb{X}$  be a minimal set of points apolar to  $\nabla^k f$ , that is,  $I_{\mathbb{X}} \subseteq \text{Ann}_{d-k}(\nabla^k f)$  and  $|\mathbb{X}| = R_{\nabla^k}(f)$ . By (1.3),  $|\mathbb{X}| \leq R_d(f)$ . We will show that this inequality cannot be strict. Let  $\mathbb{X}' \subseteq \mathbb{X}$  be the set of points defined by  $I_{\mathbb{X}'} = I_{\mathbb{X}} : (y_0)$ , i.e.,  $\mathbb{X}' = \mathbb{X} \setminus \{y_0 = 0\}$ . Therefore,

$$I_{\mathbb{X}'} : (y_0) = I_{\mathbb{X}} : (y_0) + (y_0) \subseteq \text{Ann}_{d-k}(\nabla^k f) : (y_0) + (y_0).$$

By Lemma 2.12,

$$(3.1) \quad |\mathbb{X}'| = \sum_{i \geq 0} \text{HF}(A_{I_{\mathbb{X}'} : (y_0)}; i) \geq \sum_{i \geq 0} \text{HF}(A_{\text{Ann}_{d-k}(\nabla^k f) : (y_0) + (y_0)}; i).$$

Recalling  $\alpha_0 = k$ , by Proposition 2.24, we get

$$\text{Ann}_{d-k}(\nabla^k f) : (y_0) = (y_0^k, y_1^{\alpha_1+1}, \dots, y_n^{\alpha_n+1}) + (\mathbf{y}^{\beta-\epsilon_0} : |\beta| = d - k + 1, \beta \leq \alpha),$$

where  $\epsilon_0 = (1, 0, \dots, 0)$ . Hence,

$$\begin{aligned} \text{Ann}_{d-k}(\nabla^k f) : (y_0) + (y_0) &= (y_0, y_1^{\alpha_1+1}, \dots, y_n^{\alpha_n+1}) + \left( \mathbf{y}^{\beta'} : \begin{array}{l} |\beta'| = d - k, \beta' \leq \alpha, \\ \beta'_0 = 0 \end{array} \right) = \\ &= (y_0, y_1^{\alpha_1+1}, \dots, y_n^{\alpha_n+1}, y_1^{\alpha_1} \cdots y_n^{\alpha_n}). \end{aligned}$$

From (3.1) and Theorem 3.7, we derive

$$(3.2) \quad |\mathbb{X}'| \geq \frac{1}{\alpha_0 + 1} \prod_{i=1}^n (\alpha_i + 1) - 1 = R_d(f) - 1.$$

Now, if  $|\mathbb{X}'| > R_d(f) - 1$  or  $|\mathbb{X}'| = R_d(f) - 1$  and  $\mathbb{X}' \subsetneq \mathbb{X}$ , then  $|\mathbb{X}| \geq R_d(f)$  and we conclude. Thus, assume  $\mathbb{X}' = \mathbb{X}$  and  $|\mathbb{X}'| = |\mathbb{X}| = R_d(f) - 1$ . Since

$$\text{HF}(A_{I_{\mathbb{X}} : (y_0)}; i) \geq \text{HF}(A_{\text{Ann}_{d-k}(\nabla^k f) : (y_0) + (y_0)}; i) \geq 0 \quad \text{for any } i$$

with the constraint

$$\sum_{i \geq 0} \text{HF}(A_{I_{\mathbb{X}}+(y_0)}; i) = \sum_{i \geq 0} \text{HF}(A_{\text{Ann}_{d-k}(\nabla^k f):(y_0)+(y_0)}; i),$$

we obtain  $\text{HF}(A_{I_{\mathbb{X}}+(y_0)}; i) = \text{HF}(A_{\text{Ann}_{d-k}(\nabla^k f):(y_0)+(y_0)}; i)$ . Since

$$(\text{Ann}_{d-k}(\nabla^k f) : (y_0) + (y_0))_{d-k} = (y_0, y_1^{\alpha_1+1}, \dots, y_n^{\alpha_n+1}, y_1^{\alpha_1} \cdots y_n^{\alpha_n})_{d-k} = S^{d-k}V^*,$$

we deduce  $\text{HF}(A_{I_{\mathbb{X}}+(y_0)}; d-k) = 0$ . By Lemma 2.12, we have

$$\text{HF}(A_{I_{\mathbb{X}}}; d-k) = \text{HF}(A_{I_{\mathbb{X}}}; d-k-1).$$

This implies  $\text{reg}(\mathbb{X}) \leq d-k-1$ . Thus, by Remark 2.14, we have that the maximal degree of a minimal set of generators of  $I_{\mathbb{X}}$  is at most  $d-k$ . Now, by Proposition 2.24,  $\text{Ann}_{d-k}(\nabla^k f)$  coincides with  $\text{Ann}_d(f)$  up to degree  $d-k$ , so if  $I_{\mathbb{X}}$  is generated in degree at most  $d-k$ , we obtain  $I_{\mathbb{X}} \subseteq \text{Ann}_d(f)$ . This is a contradiction by the apolarity lemma.  $\square$

*Remark 3.9.* The approach adopted in the proof of Theorem 3.8 adapts the approach used in the proof of Theorem 3.7 in [CCG12] to the case of gradient rank. The same strategy is used in [CCC<sup>+</sup>18] to compute the ranks of so-called 1-computable forms; see [CCC<sup>+</sup>18, Definition 3.5]. We observe that our strategy does not necessarily compute the gradient rank of 1-computable forms in general. For example, in [CCC<sup>+</sup>18, Proposition 4.4] the authors show that  $f = x_0^a(x_1^b + \dots + x_n^b)$  has rank equal to  $(a+1)n$ . However, in this case the quotient over the ideal  $\text{Ann}_{d-1}(\nabla f) : (x_1, \dots, x_n) + (\ell)$  has in general dimension much smaller than  $(a+1)n$ . This is to stress that inequality (3.2) is peculiar to the case of monomials and, despite the structure of the proof of Theorem 3.8, it does not seem to be related to 1-computability.

We obtain a similar result about cactus gradient ranks of monomials. Recall the result on the cactus rank of monomials.

**THEOREM 3.10** (see [RS11, Corollary 2]). *Let  $f = \mathbf{x}^\alpha$  with  $\alpha_n = \max_i \{\alpha_i\}$ . Then*

$$\text{cR}_d(f) = \frac{1}{\alpha_n + 1} \prod_{i=0}^n (\alpha_i + 1).$$

This is obtained by using the following general lower bound, which is proven in [RS11, Proposition 1] in a slightly less general setting.

**LEMMA 3.11.** *Let  $A_J = S^{\bullet}V^*/J$  be a graded Artinian algebra, and let  $I_{\mathbb{X}} \subseteq J$  be an ideal defining a 0-dimensional scheme  $\mathbb{X} \subseteq \mathbb{P}V$ . Let  $\delta = \min\{i : J_i \text{ is base point free}\}$ . Then*

$$\deg(\mathbb{X}) \geq \frac{\dim_{\mathbb{k}} A_J}{\delta}.$$

*Proof.* Let  $\widehat{\mathbb{X}} \subseteq V$  be the affine cone defined by  $I_{\mathbb{X}}$ ; since  $\dim \mathbb{X} = 0$ , we have  $\dim \widehat{\mathbb{X}} = 1$ . Let  $g \in J_\delta$  be a generic form and let  $Z(g) \subseteq V$  be the affine variety defined by the form  $g$ . Since  $J_\delta$  is base point free,  $g$  does not vanish on  $\mathbb{X}$  by Bertini's theorem [Har77, Theorem 8.18]. Moreover, by the genericity assumption,  $Z(g)$  intersects  $\widehat{\mathbb{X}}$  properly, namely,  $\dim(Z(g) \cap \widehat{\mathbb{X}}) = 0$ . Let  $\text{Spec}(A_J)$  be the scheme in  $V$  defined by  $J$ , which is a 0-dimensional scheme supported at  $0 \in V$  with  $\deg(\text{Spec}(A_J)) = \dim_{\mathbb{k}}(A_J)$ . We have  $\text{Spec}(A_J) \subseteq \widehat{\mathbb{X}}$  and  $\text{Spec}(A_J) \subseteq Z(g)$ ; therefore,  $\text{Spec}(A_J) \subseteq Z(g) \cap \widehat{\mathbb{X}}$ , and since they are 0-dimensional we obtain  $\deg(\text{Spec}(A_J)) \leq \deg(Z(g) \cap \widehat{\mathbb{X}})$ . By Bézout's theorem,  $\dim_{\mathbb{k}}(A_J) \leq \deg(g) \cdot \deg(\widehat{\mathbb{X}}) = \delta \deg(\mathbb{X})$ ; that concludes the proof.  $\square$

A direct consequence of Lemma 3.11 and Proposition 2.24 is as follows. For  $f \in S^dV$  and every  $k < d$ , we have

$$(3.3) \quad cR_{\nabla^k}(f) \geq \frac{\dim_{\mathbb{K}} A_{\text{Ann}_{d-k}(\nabla^k f)}}{\delta},$$

where  $\delta := \min \{i : (\text{Ann}_{d-k}(\nabla^k f))_i \text{ is base point free}\}$ . In particular,  $\delta \leq d - k + 1$ , because  $(\text{Ann}_{d-k}(\nabla^k f))_{d-k+1} = S^{d-k+1}V^*$  by Proposition 2.24. More generally, if  $I$  is a graded Artinian ideal, the component  $I_\delta$  of degree  $\delta$  is base point free if and only if the ideal  $(I_\delta)$  that it generates is Artinian.

From this inequality, we derive the following result on cactus gradient ranks of monomials.

**THEOREM 3.12.** *Let  $d \in \mathbb{N}$  and  $n \geq 1$ . Let  $f = \mathbf{x}^\alpha$  with  $|\alpha| = d$ . Then, for any  $\underline{d} \vdash d$ ,*

$$cR_d(f) = cR_{\underline{d}}(f) = cR_{\nabla}(f).$$

*Proof.* We assume that  $\alpha_n = \max_i \{\alpha_i\}$ . By Theorem 3.10, we have that the cactus rank of the monomial is  $cR_d(f) = \frac{1}{\alpha_n + 1} \prod_{i=0}^n (\alpha_i + 1)$ , and by Corollary 2.21 we have  $cR_{\nabla}(f) \leq cR_d(f)$ ; we show the opposite inequality. Since  $\alpha_n + 1 \leq d$ , by (3.3),

$$cR_{\nabla}(f) \geq \left\lceil \frac{\prod_{i=0}^n (\alpha_i + 1) - 1}{\alpha_n + 1} \right\rceil = \frac{\prod_{i=0}^n (\alpha_i + 1)}{\alpha_n + 1}. \quad \square$$

We conclude this section with some other remarks about  $k$ th gradient ranks of monomial for  $k$  sufficiently larger than the minimal exponent.

**LEMMA 3.13.** *Let  $f \in S^dV$ , and assume that its  $k$ th catalecticant matrix (see (2.9)) is surjective. Then  $R_{\nabla^k}(f) = \binom{d-k+n}{n}$ .*

*Proof.* By assumption,  $\langle \nabla^k f \rangle = S^{d-k}V$ . Then  $R_{\nabla^k}(f) \geq \binom{d-k+n}{n}$ . On the other hand, since the Veronese variety  $\nu_{d-k}(\mathbb{P}V)$  is nondegenerate, we can find a set of points of  $\nu_{d-k}(\mathbb{P}V)$  which is a basis of the ambient space. Then  $R_{\nabla^k}(f) \leq \binom{d-k+n}{n}$ .  $\square$

**COROLLARY 3.14.** *Let  $f = \mathbf{x}^\alpha$  be a monomial with  $\alpha_0 = \min_i \{\alpha_i\}$ . Let  $k$  be an integer such that  $k \geq d - a_0$ . Then  $R_{\nabla^k}(f) = \binom{d-k+n}{n}$ .*

*Proof.* We show that  $\text{cat}_k(f) : S^kV^* \rightarrow S^{d-k}V$  is surjective. Every monomial  $f' \in S^{d-k}V$  occurs as a  $k$ th partial derivative of  $f$ . Indeed, for any  $f' = \mathbf{x}^\beta$ , where  $\beta = (\beta_0, \dots, \beta_n)$  with  $|\beta| = d - k$ , we have  $\beta_j \leq d - k \leq \alpha_0 \leq \alpha_j$  for every  $j \geq 0$ . By Lemma 3.13, we conclude the proof.  $\square$

**3.4. Elementary symmetric polynomials.** In this section, we focus on elementary symmetric polynomials. Let  $e_{n+1,d}$  denote the elementary symmetric polynomial of degree  $d$  in  $n+1$  variables, that is, the sum of all square-free monomials of degree  $d$ , i.e.,

$$e_{n+1,d} = \sum_{0 \leq i_1 < \dots < i_d \leq n} x_{i_1} \cdots x_{i_d} \in S^dV.$$

In [Lee16], Lee determined  $R_d(e_{n+1,d})$  for  $d$  odd and gave bounds when  $d$  is even.

**THEOREM 3.15** (see [Lee16, Theorem 3.4 and Corollary 4.4]). *Let  $d \in \mathbb{N}$ , and let  $n \geq 1$ .*

*If  $d$  is odd, then*

$$R_d(e_{n+1,d}) = \sum_{i=0}^{\frac{d-1}{2}} \binom{n+1}{i}.$$

If  $d$  is even, then

$$\sum_{i=0}^{\frac{d}{2}} \binom{n+1}{i} \geq R_d(e_{n+1,d}) \geq \sum_{i=0}^{\frac{d}{2}} \binom{n+1}{i} - \binom{n}{\frac{d}{2}} - 1.$$

We extend these results to the first gradient rank of  $e_{n+1,d}$ .

**THEOREM 3.16.** Let  $d \in \mathbb{N}$ , and let  $n \geq 1$ .

If  $d$  is odd, then

$$R_d(e_{n+1,d}) = R_{1,d-1}(e_{n+1,d}) = R_\nabla(e_{n+1,d}).$$

If  $d$  is even, then

$$R_d(e_{n+1,d}) \geq R_\nabla(e_{n+1,d}) \geq \sum_{i=0}^{\frac{d}{2}} \binom{n+1}{i} - \binom{n}{\frac{d}{2}} - 1.$$

*Proof.* By (1.3) and Theorem 3.15, it is enough to prove the lower bounds on  $R_\nabla(e_{n+1,d})$ .

By Proposition 2.24, we have the equality

$$\text{Ann}_{d-1}(\nabla e_{n+1,d}) = \text{Ann}_d(e_{n+1,d}) + (S^d V^*).$$

Let  $\phi = \mathbf{y}^\beta$  be any square-free monomial: notice that  $\phi \circ e_{n+1,d} \neq 0$ , and therefore

$$(3.4) \quad \text{Ann}_{d-1}(\nabla e_{n+1,d}) = \text{Ann}_d(e_{n+1,d}) + (\phi).$$

Consider a monomial  $\phi = \mathbf{y}^\beta$  divisible by  $y_0$ , so  $\phi = y_0 \tilde{\phi}$ . We are going to show that

$$(3.5) \quad \text{Ann}_{d-1}(\nabla e_{n+1,d}) : (y_0) = \text{Ann}_d(e_{n+1,d}) : (y_0) + (\tilde{\phi}).$$

The containment  $\text{Ann}_{d-1}(\nabla e_{n+1,d}) : (y_0) \supseteq \text{Ann}_d(e_{n+1,d}) : (y_0) + (\tilde{\phi})$  is clear from the definitions. For the converse, let  $\psi \in \text{Ann}_{d-1}(\nabla e_{n+1,d}) : (y_0)$  so that  $y_0 \psi \in \text{Ann}_{d-1}(\nabla e_{n+1,d})$ . By (3.4), we have  $y_0 \psi = \psi_1 + \psi_2 \cdot y_0 \tilde{\phi}$  for some  $\psi_1 \in \text{Ann}_d(e_{n+1,d})$  and  $\psi_2 \in S^d V^*$ . Hence,  $y_0$  divides  $\psi_1$ , that is,  $\psi_1 = y_0 \tilde{\psi}_1$ . We deduce  $\tilde{\psi}_1 \in \text{Ann}_d(e_{n+1,d}) : (y_0)$ . Therefore,  $\psi = \psi_1 + \psi_2 \tilde{\phi} \in \text{Ann}_d(e_{n+1,d}) : (y_0) + (\tilde{\phi})$ . This proves (3.5).

Note that  $\text{Ann}_d(e_{n+1,d}) : (y_0) + (y_0) = \text{Ann}_{d-1}(e_{n,d-1})$ , where  $e_{n,d-1}$  is the elementary symmetric polynomial in the variables  $x_1, \dots, x_n$ . Hence, from (3.5), we get

$$\begin{aligned} \text{Ann}_{d-1}(e_{n+1,d}) : (y_0) + (y_0) &= \text{Ann}_d(e_{n+1,d}) : (y_0) + (\tilde{\phi}) + (y_0) = \\ &= \text{Ann}_{d-1}(e_{n,d-1}) + (\tilde{\phi}). \end{aligned}$$

Let  $\mathbb{X}$  be a minimal set of points apolar to  $\nabla e_{n+1,d}$ , that is,  $I_{\mathbb{X}} \subseteq \text{Ann}_{d-1}(\nabla e_{n+1,d})$  with  $|\mathbb{X}| = R_\nabla(f) \leq R_d(e_{n+1,d})$ . Let  $\mathbb{X}' = \mathbb{X} \cap \{y_0 \neq 0\}$  so that  $I_{\mathbb{X}'} = I_{\mathbb{X}} : (y_0)$ . Now, we employ the same strategy as in the proof of Theorem 3.8. Using (3.4), we have

$$\begin{aligned} |\mathbb{X}'| &= \sum_{i \geq 0} \text{HF}(A_{I_{\mathbb{X}'} + (y_0)}; i) \geq \sum_{i \geq 0} \text{HF}(A_{\text{Ann}_{d-1}(\nabla e_{n+1,d}) : (y_0) + (y_0)}; i) = \\ (3.6) \quad &= \sum_{i \geq 0} \text{HF}(A_{\text{Ann}_{d-1}(e_{n,d-1}) + (\tilde{\phi})}; i). \end{aligned}$$

From the proof of [Lee16, Theorem 3.4], for  $d$  odd, we have

$$\sum_{i \geq 0} \text{HF}(A_{\text{Ann}_{d-1}(e_{n,d-1})}; i) = \sum_{i=0}^{\frac{d-1}{2}} \binom{n+1}{i} = R_d(e_{n+1,d}).$$

Now,  $\tilde{\phi}$  is a square-free monomial of degree  $d-1$  not divisible by  $y_0$ ; therefore,  $\tilde{\phi} \notin \text{Ann}_{d-1}(e_{n,d-1})$ . By (3.6), we obtain

$$|\mathbb{X}'| \geq R_d(e_{n+1,d}) - 1.$$

Applying the same argument as in the last part of the proof of Theorem 3.8, we conclude that  $|\mathbb{X}| \geq R_d(e_{n+1,d})$ , which concludes the proof for  $d$  odd.

By the proof of [Lee16, Corollary 4.4], for  $d$  even, we have

$$\sum_{i \geq 0} \text{HF}(A_{\text{Ann}_{d-1}(e_{n,d-1})}; i) = \sum_{i=0}^{\frac{d}{2}} \binom{n+1}{i} - \binom{n}{\frac{d}{2}}.$$

Again, since  $\tilde{\phi} \notin \text{Ann}_{d-1}(e_{n,d-1})$ , we obtain

$$|\mathbb{X}| \geq |\mathbb{X}'| \geq \sum_{i=0}^{\frac{d}{2}} \binom{n}{i} - \binom{n-1}{\frac{d}{2}} - 1,$$

which concludes the proof for  $d$  even.  $\square$

**Acknowledgments.** This project was realized during the Research in Pairs program at CIRM Trento in July 2018. We thank CIRM for financial support and for providing a stimulating environment for the development of the project. We also thank E. Ballico and A. Bernardi for helpful discussions and suggestions during our stay in Trento. Finally, we thank the three anonymous referees for their useful comments.

## REFERENCES

- [AGMO18] E. ANGELINI, F. GALUPPI, M. MELLA, AND G. OTTAVIANI, *On the number of Waring decompositions for a generic polynomial vector*, J. Pure Appl. Algebra, 222 (2018), pp. 950–965.
- [AH95] J. ALEXANDER AND A. HIRSCHOWITZ, *Polynomial interpolation in several variables*, J. Algebraic Geom., 4 (1995), pp. 201–222.
- [AR03] E. S. ALLMAN AND J. A. RHODES, *Phylogenetic invariants for the general Markov model of sequence mutation*, Math. Biosci., 186 (2003), pp. 113–144.
- [BB13] E. BALLICO AND A. BERNARDI, *Tensor ranks on tangent developable of Segre varieties*, Linear Multilinear Algebra, 61 (2013), pp. 881–894.
- [BB14] W. BUCZYŃSKA AND J. BUCZYŃSKI, *Secant varieties to high degree Veronese reembeddings, catalecticant matrices and smoothable Gorenstein schemes*, J. Algebraic Geom., 23 (2014), pp. 63–90.
- [BB15] W. BUCZYŃSKA AND J. BUCZYŃSKI, *On differences between the border rank and the smoothable rank of a polynomial*, Glasgow Math. J., 57 (2015), pp. 401–413.
- [BBCG18] E. BALLICO, A. BERNARDI, M. CHRISTANDL, AND F. GESMUNDO, *On the partially symmetric rank of tensor products of W-states and other symmetric tensors*, Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl., 30 (2019), pp. 93–124.
- [BBCM13] A. BERNARDI, J. BRACHAT, P. COMON, AND B. MOURRAIN, *General tensor decomposition, moment matrices and applications*, J. Symbolic Comput., 52 (2013), pp. 51–71.

- [BBT13] W. BUCZYŃSKA, J. BUCZYŃSKI, AND Z. TEITLER, *Waring decompositions of monomials*, J. Algebra, 378 (2013), pp. 45–57.
- [BCMT10] J. BRACHAT, P. COMON, B. MOURRAIN, AND E. TSIGARIDAS, *Symmetric tensor decomposition*, Linear Algebra Appl., 433 (2010), pp. 1851–1872.
- [BGL13] J. BUCZYŃSKI, A. GINENSKY, AND J. M. LANDSBERG, *Determinantal equations for secant varieties and the Eisenbud–Koh–Stillman conjecture*, J. Lond. Math. Soc., 88 (2013), pp. 1–24.
- [BJMR18] A. BERNARDI, J. JELISIEJEW, P. M. MARQUES, AND K. RANESTAD, *On polynomials with given Hilbert function and applications*, Collect. Math., 69 (2018), pp. 39–64.
- [BL13] J. BUCZYŃSKI AND J. M. LANDSBERG, *Ranks of tensors and a generalization of secant varieties*, Linear Algebra Appl., 438 (2013), pp. 668–689.
- [BR13] A. BERNARDI AND K. RANESTAD, *On the cactus rank of cubic forms*, J. Symbolic Comput., 50 (2013), pp. 291–297.
- [CCC<sup>+</sup>18] E. CARLINI, M. V. CATALISANO, L. CHIANTINI, A. V. GERAMITA, AND Y. WOO, *Symmetric tensors: Rank, Strassen’s conjecture and e-computability*, Ann. Sc. Norm. Super. Pisa Cl. Sci. 18 (2018), pp. 363–390.
- [CCG12] E. CARLINI, M. V. CATALISANO, AND A. V. GERAMITA, *The solution to the Waring problem for monomials and the sum of coprime monomials*, J. Algebra, 370 (2012), pp. 5–14.
- [CCO17] E. CARLINI, M. V. CATALISANO, AND A. ONETO, *Waring loci and the Strassen conjecture*, Adv. Math., 314 (2017), pp. 630–662.
- [CGLM08] P. COMON, G. GOLUB, L.-H. LIM, AND B. MOURRAIN, *Symmetric tensors and symmetric tensor rank*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1254–1279.
- [Cle61] A. CLEBSCH, *Über Curven vierter Ordnung*, J. Reine Angew. Math., 59 (1861), pp. 125–145.
- [Com94] P. COMON, *Independent component analysis, a new concept?*, Signal process., 36 (1994), pp. 287–314.
- [CS11] G. COMAS AND M. SEIGUER, *On the rank of a binary form*, Found. Comput. Math., 11 (2011), pp. 65–78.
- [CV18] E. CARLINI AND E. VENTURA, *A note on the simultaneous Waring rank of monomials*, Illinois J. Math., 61 (2017), pp. 517–530.
- [DK93] I. DOLGACHEV AND V. KANEV, *Polar covariants of plane cubics and quartics*, Adv. Math., 98 (1993), pp. 216–301.
- [EH00] D. EISENBUD AND J. HARRIS, *The Geometry of Schemes*, Grad. Texts in Math. 197, Springer-Verlag, New York, 2000.
- [Eis95] D. EISENBUD, *Commutative Algebra: With a View Toward Algebraic Geometry*, Grad. Texts in Math. 150, Springer-Verlag, New York, 1995.
- [ER93] R. EHRENBORG AND G.-C. ROTA, *Apolarity and canonical forms for homogeneous polynomials*, European J. Combin., 14 (1993), pp. 157–181.
- [Fon02] C. FONTANARI, *On Waring’s problem for many forms and Grassmann defective varieties*, J. Pure Appl. Algebra, 174 (2002), pp. 243–247.
- [Fri13] S. FRIEDLAND, *On tensors of border rank l in  $\mathbb{C}^{m \times n \times l}$* , Linear Algebra Appl., 438 (2013), pp. 713–737.
- [Fri16] S. FRIEDLAND, *Remarks on the symmetric rank of symmetric tensors*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 320–337.
- [Gal16] M. GALAZKA, *Multigraded Apolarity*, <https://arxiv.org/abs/1601.06211>, 2016.
- [Gal17] M. GALAZKA, *Vector bundles give equations of cactus varieties*, Linear Algebra Appl., 521 (2017), pp. 254–262.
- [Ger96] A. V. GERAMITA, *Inverse systems of fat points: Waring’s problem, secant varieties of Veronese varieties and parameter spaces for Gorenstein ideals*, in The Curves Seminar at Queen’s, A. V. Geramita, ed., Queen’s University, Kingston, ON, Canada, 1996, pp. 2–114.
- [GRV18] M. GALLET, K. RANESTAD, AND N. VILLAMIZAR, *Varieties of apolar subschemes of toric surfaces*, Ark. Mat., 56 (2018), pp. 73–99.
- [GSS05] L. D. GARCIA, M. E. STILLMAN, AND B. STURMFELS, *Algebraic geometry of Bayesian networks*, J. Symbolic Comput., 39 (2005), pp. 331–355.
- [Hac12] W. HACKBUSCH, *Tensor Spaces and Numerical Tensor Calculus*, Springer Ser. Comput. Math. 42, Springer Science & Business Media, New York, 2012.
- [Har77] R. HARTSHORNE, *Algebraic Geometry*, Grad. Texts in Math. 52, Springer-Verlag, New York, 1977.
- [Iar95] A. IARROBINO, *Inverse system of a symbolic power II. The Waring problem for forms*, J. Algebra, 174 (1995), pp. 1091–1110.

- [IK99] A. IARROBINO AND V. KANEV, *Power Sums, Gorenstein Algebras, and Determinantal Loci*, Lecture Notes in Math. 1721, Springer-Verlag, Berlin, 1999.
- [Kle99] J. KLEPPE, *Representing a Homogenous Polynomial as a Sum of Powers of Linear Forms*, Ph.D. thesis, University of Oslo, Oslo, Norway, 1999.
- [Lan12] J. M. LANDSBERG, *Tensors: Geometry and Applications*, Grad. Stud. Math. 128, American Mathematical Society, Providence, RI, 2012.
- [Las04] E. LASKER, *Zur Theorie der kanonischen Formen*, Math. Ann., 58 (1904), pp. 434–440.
- [Lee16] H. LEE, *Power sum decompositions of elementary symmetric polynomials*, Linear Algebra Appl., 492 (2016), pp. 89–97.
- [Lin00] Y. LIN, *Tensor product space ANOVA models*, Ann. Statist., 28 (2000), pp. 734–755.
- [LO13] J. M. LANDSBERG AND G. OTTAVIANI, *Equations for secant varieties of Veronese and other varieties*, Ann. Mat. Pura Appl. (4), 192 (2013), pp. 569–606.
- [LT10] J. M. LANDSBERG AND Z. TEITLER, *On the ranks and border ranks of symmetric tensors*, Found. Comput. Math., 10 (2010), pp. 339–366.
- [Oed08] L. OEDING, *Report on “Geometry and Representation Theory of Tensors for Computer Science, Statistics and Other Areas,”* <https://arxiv.org/abs/0810.3940>, 2008.
- [Pal03a] F. PALATINI, *Sulla rappresentazione delle forme ed in particolare della cubica quinaria con la somme di potenze di forme lineari*, Atti Reale Accad. Sci. Torino, 38 (1903), pp. 43–50.
- [Pal03b] F. PALATINI, *Sulla rappresentazione delle forme ternarie mediante la somma di potenze di forme lineari*, Rend. Accad. Lincei V, 12 (1903), pp. 378–384.
- [Rez92] B. REZNICK, *Sum of Even Powers of Real Linear Forms*, American Mathematical Society, Providence, RI, 1992.
- [RS11] K. RANESTAD AND F.-O. SCHREYER, *On the rank of a symmetric form*, J. Algebra, 346 (2011), pp. 340–342.
- [Sei18] A. SEIGAL, *Ranks and Symmetric Ranks of Cubic Surfaces*, <https://arxiv.org/abs/1801.05377>, 2018.
- [Shi18] Y. SHITOV, *A counterexample to Comon's conjecture*, SIAM J. Appl. Algebra Geom., 2 (2018), pp. 428–443.
- [SV06] J. SIDMAN AND A. VAN TUYL, *Multigraded regularity: syzygies and fat points*, Beitr. Algebra Geom., 47 (2006), pp. 1–22.
- [Syl52] J. J. SYLVESTER, *On the principles of the calculus of forms*, Cambridge Dublin Math. J. (1852), pp. 52–97.
- [Tei14] Z. TEITLER, *Geometric Lower Bounds for Generalized Ranks*, <https://arxiv.org/abs/1406.5145>, 2014.
- [Ter15] A. TERRACINI, *Sulla rappresentazione delle coppie di forme ternarie mediante somme di potenze di forme lineari*, Ann. Mat. Pura Appl. (1898–1922), 24 (1915), pp. 1–10.
- [Ter16] A. TERRACINI, *Sulla rappresentazione delle forme quaternarie mediante somme di potenze di forme lineari*, Atti Reale Acc. Sci. Torino, 51 (1916), pp. 107–117.
- [Ven18] E. VENTURA, *Real rank boundaries and loci of forms*, Linear Multilinear Algebra, 67 (2019), pp. 1404–1410.
- [ZHQ16] X. ZHANG, Z.-H. HUANG, AND L. QI, *Comon's conjecture, rank decomposition, and symmetric rank decomposition of symmetric tensors*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 1719–1728.