



# On generic and maximal $k$ -ranks of binary forms

Samuel Lundqvist, Alessandro Oneto, Bruce Reznick, Boris Shapiro

► **To cite this version:**

Samuel Lundqvist, Alessandro Oneto, Bruce Reznick, Boris Shapiro. On generic and maximal  $k$ -ranks of binary forms. 2017. <hal-01637944>

**HAL Id: hal-01637944**

**<https://hal.inria.fr/hal-01637944>**

Submitted on 18 Nov 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ON GENERIC AND MAXIMAL $k$ -RANKS OF BINARY FORMS

SAMUEL LUNDQVIST, ALESSANDRO ONETO, BRUCE REZNICK, AND BORIS SHAPIRO

*To James Joseph Sylvester, a mathematician, a polyglot, and a poet*

ABSTRACT. In what follows, we pose two general conjectures about decompositions of homogeneous polynomials as sums of powers. The first one (suggested by G. Ottaviani) deals with the generic  $k$ -rank of complex-valued forms of any degree divisible by  $k$  in any number of variables. The second one (by the fourth author) deals with the maximal  $k$ -rank of binary forms. We settle the first conjecture in the cases of two variables and the second in the first non-trivial case of the 3-rd powers of quadratic binary forms.

## 1. INTRODUCTION

A vast and currently very active area of mathematical research dealing with additive decompositions of polynomials started with the following classical result on binary forms proven in 1851 by J. J. Sylvester<sup>1</sup> [Sy51a, Sy51b].

**Theorem 1.1** (Sylvester's Theorem). (i) A general binary form  $p \in \mathbb{C}[x, y]$  of odd degree  $k = 2s - 1$  with complex coefficients can be written as

$$(1.1) \quad p(x, y) = \sum_{j=1}^s (\alpha_j x + \beta_j y)^k, \text{ for some } \alpha_j, \beta_j \in \mathbb{C}.$$

(ii) A general binary form  $p \in \mathbb{C}[x, y]$  of even degree  $k = 2s$  with complex coefficients can be written as

$$(1.2) \quad p(x, y) = \lambda x^k + \sum_{j=1}^s (\alpha_j x + \beta_j y)^k, \text{ for some } \lambda, \alpha_j, \beta_j \in \mathbb{C}.$$



FIGURE 1. J. J. Sylvester around 1890.

After Sylvester's work, decompositions of polynomials into sums of powers of linear forms have been widely studied from several perspectives starting with the geometrical point of view by the classic Italian

---

2010 *Mathematics Subject Classification*. Primary 15A21, Secondary 15A69, 14N15.

*Key words and phrases*. sums of powers, binary forms, Waring problem, generic rank, maximal rank, secant varieties.

<sup>1</sup>James Joseph (Sylvester) was born to a Jewish family in London in 1814. His remarkably original and successful mathematical career only partially helped him overcome the pervasive anti-Semitism of his era. For more on his life, see [Pa06].

school in algebraic geometry in the beginning of the 20-th century as well as current research by applied mathematicians and engineers in connection with tensor decompositions.

Such presentations are often called *Waring decompositions* and, for a given polynomial  $f$ , the smallest length of such a decomposition is called the *Waring rank*, or simply, the *rank* of  $f$ . The minimal number of linear forms required to represent a general form of degree  $k$  in  $n$  variables as a sum of their  $k$ -th powers is called the *generic rank* and denoted by  $\text{rk}^\circ(n, k)$ , while the *maximal rank*  $\text{rk}^{\max}(n, k)$  is the minimal number of linear forms required to represent *any* form of degree  $k$  in  $n$  variables. Rephrasing Theorem 1.1 in this terminology, we have that

$$\text{rk}^\circ(2, k) = \left\lceil \frac{k+1}{2} \right\rceil.$$

The explicit value of the generic Waring rank for any arbitrary  $k$  and  $n$  was obtained in the celebrated result of J. Alexander and A. Hirschowitz [AH95]. Except for the case of quadrics in all dimensions and four additional exceptions  $(n, k) = (3, 4), (4, 4), (5, 3), (5, 4)$ , the generic rank coincides with its expected value given by

$$\text{rk}^\circ(n, k) = \left\lceil \frac{1}{n} \binom{n+k-1}{k} \right\rceil.$$

In the case of quadrics, the generic rank is equal to  $n$ , while in all the exceptional cases the generic rank is by 1 bigger than the latter expected value.

Additionally, the maximal Waring rank  $\text{rk}^{\max}(2, k)$  of binary forms equals  $k$ . This was probably a classical result, but it has been recently proved in [CS11]. Also, the maximal value  $k$  is attained exactly on binary forms of the type  $l_1 l_2^{k-1}$ , where  $l_1$  and  $l_2$  are linearly independent linear forms, see [Re13b, Theorem 5.4].

Other types of additive decompositions of polynomials have been considered in the last decades. In [FOS12], the fourth author jointly with R. Fröberg and G. Ottaviani considered, for any triple of positive integers  $(k, d, n)$  with  $k, n \geq 2$ , decompositions of homogeneous polynomials of degree  $kd$  in  $n$  variables as sums of  $k$ -th powers of forms of degree  $d$ . Given a form  $f$  of degree  $kd$ , the smallest length of such a decomposition denoted by  $\text{rk}_k(f)$  is called the  $k$ -rank of  $f$ . Analogously to the classical Waring rank, we define the *generic  $k$ -rank* for forms of degree  $kd$  in  $n$  variables, denoted by  $\text{rk}_k^\circ(n, kd)$ , and the corresponding *maximal  $k$ -rank*, denoted by  $\text{rk}_k^{\max}(n, kd)$ .

The main result of [FOS12] is the following upper-bound on the generic  $k$ -rank, for any triple  $(k, d, n)$ ,

$$\text{rk}_k^\circ(n, kd) \leq k^n.$$

A remarkable property of this bound is its independence of the parameter  $d$  and also its sharpness for any fixed  $k$  and  $n$ , when  $d \gg 0$ . The following general conjecture about  $\text{rk}_k^\circ(n, kd)$  was suggested by G. Ottaviani in 2013 (private communication). For  $k = 2$ , it also coincides with [LSVB13, Conjecture 1].

**Conjecture 1.2.** *For any triple  $(k, d, n)$  of positive integers with  $k, n, d \geq 2$ ,*

$$(1.3) \quad \text{rk}_k^\circ(n, kd) = \begin{cases} \min \left\{ s \geq 1 \mid s \binom{d+n-1}{n-1} - \binom{s}{2} \geq \binom{2d+n-1}{n-1} \right\}, & \text{for } k = 2; \\ \min \left\{ s \geq 1 \mid s \binom{d+n-1}{n-1} \geq \binom{kd+n-1}{n-1} \right\}, & \text{for } k \geq 3. \end{cases}$$

Conjecture 1.2 is supported by substantial computer experiments performed by G. Ottaviani and the present authors, see e.g. [On16]. A proof of Conjecture 1.2 would extend the Alexander–Hirschowitz Theorem (which corresponds to  $d = 1$ ) to the case of arbitrary triples  $(k, d, n)$  and will complete the important project of determining generic ranks for symmetric tensors.

In this paper, we mainly consider the case of binary forms.

**Theorem 2.3.** *For  $k, d \geq 2$ , the generic  $k$ -rank of binary forms of degree  $kd$  is*

$$\mathrm{rk}_k^\circ(2, kd) = \left\lceil \frac{kd+1}{d+1} \right\rceil.$$

This result extends Theorem 1.1 to presentations of general binary forms of degree  $kd$  as sums of  $k$ -th powers of binary forms of degree  $d$  and gives a proof of Conjecture 1.2 in the case of binary forms.

An alternative proof of Theorem 2.3 using canonical forms can be found in [Re13a, Theorem 1.8]. Our approach puts this result in a much more general setting; in particular, we relate Conjecture 1.2 to Fröberg's Conjecture on Hilbert series of generic ideals. This relation extends to higher dimensions, as we explain in Appendix A, where we also settle the case of sums of squares in three variables ( $k = 2, n = 3$ ).

In Fall 2014, the fourth author formulated the following conjecture about the maximal rank of binary forms.

**Conjecture 1.4.** *For  $k \geq 2$ , the maximal  $k$ -rank of binary forms of degree  $kd$  is*

$$\mathrm{rk}_k^{\max}(2, kd) = k.$$

As we said, the case  $d = 1$  is classical and well-known, see [Re13b, Theorem 4.9]. Moreover, by [Re13b, Theorem 5.4], we know that a binary form has maximal rank if and only if it can be decomposed as  $\ell_1 \ell_2^{k-1}$ . Also, it is easy to prove that any binary polynomial of even degree can be decomposed as a sum of two squares, see [FOS12, Theorem 5]. In this paper, we prove Conjecture 1.4 in the first open case of binary sextics decomposed as sums of cubes ( $k = 3, d = 2, n = 2$ ).

**Theorem 3.1.** *Every binary sextic can be written as a sum of at most three cubes of binary quadratic forms.*

One can also suspect that  $\mathrm{rk}_k^{\max}(l_1 l_2^{kd-1}) = k$ , where  $l_1$  and  $l_2$  are non-proportional linear binary forms, similarly to what happens in the classical case. From [CO15], we know that this is an upper bound, but computing the actual  $k$ -rank is a very difficult task. In Section 4, we explain how our geometric approach can help in these computations; e.g., we can show that  $\mathrm{rk}_4^{\max}(l_1 l_2^7) = 4$ . In Appendix B, we continue the work started in [CO15] and we prove that a class of monomials has  $k$ -rank smaller than  $k$ .

*Acknowledgements.* We want to thank the participants of the problem-solving seminar in commutative algebra at Stockholm University and especially Ralf Fröberg for creating a nice research atmosphere. It is a pleasure to acknowledge the importance of our communication with Giorgio Ottaviani for the present study. Finally, the fourth author is sincerely grateful to the University of Illinois, Urbana-Champaign for the hospitality in summer 2015 when a part of this project was carried out.

## 2. GENERIC $k$ -RANKS

In this section we focus on *generic  $k$ -ranks*. Let  $S = \bigoplus_{d \geq 0} S_d$  be the standard graded polynomial ring in  $n$  variables with complex coefficients, where  $S_d$  denotes the vector space of degree  $d$  forms. By a count of parameters, we have a lower bound for the generic  $k$ -rank of forms of degree  $kd$ , i.e.,

$$\mathrm{rk}_k^\circ(n, kd) \geq \left\lceil \frac{\dim_{\mathbb{C}} S_d}{\dim_{\mathbb{C}} S_d} \right\rceil = \left\lceil \frac{\binom{kd+n-1}{n-1}}{\binom{d+n-1}{n-1}} \right\rceil.$$

**2.1. Secant varieties and Terracini's Lemma.** A common approach to analyse Waring-type problems is to study *secant varieties*. We describe it in our context.

Let  $\mathrm{VP}_{n, kd}^{(k)}$  be the *variety of  $k$ -th powers* in the space of polynomials of degree  $kd$ , i.e.,

$$\mathrm{VP}_{n, kd}^{(k)} = \left\{ [g^k] \mid g \in S_d \right\} \subset \mathbb{P}(S_{kd}).$$

In the case where  $d = 1$ , this is the classical *Veronese variety*  $\text{Ver}_{n,k}$ .

The  $s$ -th secant variety of  $\text{VP}_{n,kd}^{(k)}$  is the closure of the set of points that lie on the linear span of  $s$  points on the variety of powers, i.e.,

$$\sigma_s \left( \text{VP}_{n,kd}^{(k)} \right) = \overline{\bigcup_{g_1, \dots, g_s \in S_d} \langle [g_1^k], \dots, [g_s^k] \rangle} \subset \mathbb{P}(S_{kd}),$$

in other words, the  $s$ -th secant variety is the closure of the set of homogeneous polynomials of degree  $kd$  with  $k$ -th rank at most  $s$ . Observe that, since  $\text{VP}_{n,kd}^{(k)}$  is non-degenerate, secant varieties give a filtration and eventually fill the ambient space. These definitions allow us to describe the  $k$ -th generic rank as

$$\text{rk}_k^\circ(n, kd) = \min \left\{ s \mid \sigma_s \left( \text{VP}_{n,kd}^{(k)} \right) = \mathbb{P}(S_{kd}) \right\}.$$

Therefore, one way to understand generic  $k$ -ranks is to compute the dimensions of secant varieties. In order to do so, a classical tool is Terracini's Lemma which describes the generic tangent space to a *secant variety*, see [Te11]. In our setting, it is equivalent to the following statement.

**Lemma 2.1** (Terracini's Lemma). *Let  $P$  be a generic point on the linear span of  $s$  generic points on the variety of  $k$ -th powers,  $[g_1^k], \dots, [g_s^k] \in \text{VP}_{n,kd}^{(k)}$ . Then the tangent space to  $\sigma_s \left( \text{VP}_{n,kd}^{(k)} \right)$  at  $P$  coincides with the linear space of the tangent spaces to  $\text{VP}_{n,kd}^{(k)}$  at the  $s$  points, i.e.,*

$$T_P \sigma_s \left( \text{VP}_{n,kd}^{(k)} \right) = \left\langle T_{[g_1^k]} \text{VP}_{n,kd}^{(k)}, \dots, T_{[g_s^k]} \text{VP}_{n,kd}^{(k)} \right\rangle \subset \mathbb{P}(S_{kd}).$$

By Lemma 2.1, we can reduce the problem of computing generic ranks to a question in commutative algebra. Let us recall some basic definitions.

Given a homogeneous ideal  $I \subset S$ , we have an induced grading of  $I$  and of the quotient ring  $S/I$ . We define the *Hilbert function* of  $S/I$  in degree  $d$  as the dimension of the vector space formed by the degree  $d$  part of  $S/I$ , i.e.,

$$\text{HF}_{S/I}(d) = \dim_{\mathbb{C}}[S/I]_d = \dim_{\mathbb{C}} S_d - \dim_{\mathbb{C}} I_d, \text{ for } d \geq 0;$$

and the *Hilbert series* of  $S/I$  by

$$\text{HS}_{S/I}(t) = \sum_{d \geq 0} \text{HF}_{S/I}(d) t^d \in \mathbb{Z}[[t]].$$

It is easy to observe that the tangent space to the variety of powers at a point  $[g^k] \in T_{[g_1^k]} \text{VP}_{n,kd}^{(k)}$  is

$$T_{[g^k]} \text{VP}_{n,kd}^{(k)} = \{[g^{k-1}h] \mid h \in S_d\}.$$

Therefore, under the genericity assumptions of Lemma 2.1, we have that

$$T_P \sigma_s \left( \text{VP}_{n,kd}^{(k)} \right) = \mathbb{P}(I_{kd}) \subset \mathbb{P}(S_{kd}),$$

where  $I_{kd}$  is the homogeneous part of degree  $kd$  of the ideal  $I = (g_1^{k-1}, \dots, g_s^{k-1})$ , where the  $g_i$ 's are general forms of degree  $d$ . In particular, the codimension of the secant variety is

$$(2.1) \quad \text{codim } \sigma_s \left( \text{VP}_{n,kd}^{(k)} \right) = \text{HF}_{S/I}(kd) - 1,$$

and the generic  $k$ -rank  $\text{rk}_k^\circ(d, n)$  is equal to the minimal  $s$  such that  $\text{HF}_{S/I}(kd) = 0$ .

**2.2. Generic ranks of binary forms.** We consider now the case of binary forms.

In [Fr85], R. Fröberg observed that, for any ideal  $I = (h_1, \dots, h_s) \subset S$  with  $\deg(h_i) = d_i$ , for  $i = 1, \dots, s$ ,

$$(2.2) \quad \text{HS}_{S/I}(t) \succeq \left\lceil \frac{\prod_{i=1}^s (1 - t^{d_i})}{(1 - t)^n} \right\rceil,$$

where  $\lceil \cdot \rceil$  stands for the truncation of the power series at the first non-positive coefficient and the latter inequality is in the lexicographic sense. An ideal for which (2.2) is an equality is called *Hilbert generic*. Fröberg's Conjecture claims that generic ideals are Hilbert generic. We return to this in Appendix A.

A crucial observation is that the property of being Hilbert generic is a Zariski open condition on the space of ideals in  $S$  with a given number of generators of given degrees; see [FL91].

**Lemma 2.2.** *Let  $g_1, \dots, g_s$  be generic binary forms of degrees  $d_1, \dots, d_s$ , respectively. Then, for any  $s$ -tuple of positive integers  $k_1, \dots, k_s$ , the ideal  $(g_1^{k_1}, \dots, g_s^{k_s}) \subset S = \mathbb{C}[x, y]$  is Hilbert generic.*

*Proof.* As we mentioned above, it suffices to prove the statement for a specialization of the  $g_i$ 's. Assume that they are powers of generic linear forms, i.e.,  $g_i = l_i^{d_i}$ , for  $i = 1, \dots, s$ . By [GS98, Corollary 2.3], the ideal  $(l_1^{d_1 k_1}, \dots, l_s^{d_s k_s})$  is Hilbert generic and we are done.  $\square$

Using Lemma 2.2, we are able to settle Conjecture 1.2 in the case of binary forms.

**Theorem 2.3.** *For  $k, d \geq 2$ , the generic  $k$ -rank of binary forms of degree  $kd$  is*

$$\text{rk}_k^\circ(2, kd) = \left\lceil \frac{kd + 1}{d + 1} \right\rceil.$$

*Proof.* The case  $k = 2$  is covered by [FOS12, Theorem 5]. Assume that  $k \geq 3$ . By Lemma 2.1 and (2.1), in order to compute the dimension of the  $s$ -th secant variety of the variety of  $k$ -th powers, we have to compute the Hilbert function in degree  $dk$  of the ideal  $I = (g_1^{k-1}, \dots, g_s^{k-1})$ , where the  $g_i$ 's are generic binary forms of degree  $d$ . By Lemma 2.2, we know that  $I$  is Hilbert generic. Therefore,  $\text{HS}_{S/I}(t) = \lceil H(t) \rceil$ , where

$$H(t) = \frac{(1 - t^{d(k-1)})^s}{(1 - t)^2} =: \sum_{i \geq 0} H_i t^i.$$

Since  $k \geq 3$ , for any  $d(k-1) \leq i \leq dk$ , we have  $H_i = (i+1) - s(i - d(k-1) + 1)$ . Then,

$$s_i \geq s_{i+1}, \text{ for all } i = d(k-1), \dots, dk,$$

where we denote  $s_i := \min\{s \mid H_i \leq 0\}$ . From this, it follows that, for  $s \leq \left\lceil \frac{kd+1}{d+1} \right\rceil$ ,

$$\text{HF}_{S/I}(dk) = \max\{0, H_{kd}\} = \max\{0, (kd+1) - s(d+1)\}.$$

In particular, we conclude that  $\text{rk}_k^\circ(2, kd) = \left\lceil \frac{kd+1}{d+1} \right\rceil$ .  $\square$

As we mentioned in the introduction, the proof of Theorem 2.3 has been already obtained by different methods in [Re13a, Theorem 1.8]. However, the approach used here is much more general and can extend to higher dimensions. We explain this idea in Appendix A where we show that a generalized version of Fröberg's Conjecture implies Conjecture 1.2 on generic  $k$ -ranks in any number of variables. In particular, we use this to settle the case of sums of squares in three variables, see Theorem A.5.

## 3. MAXIMAL 3-RANK OF BINARY SEXTICS

In this section we settle Conjecture 1.4 in the case of sum of cubes decompositions of binary sextics.

**Theorem 3.1.** *Every binary sextic can be written as a sum of at most three cubes of binary quadratic forms.*

*Proof.* We begin with an observation about binary cubics. It has been known since the work of Sylvester that a binary cubic  $h$  can be written as sum of two cubes of linear forms unless  $h$  has a square factor, and is not a cube; see [Re13a, Theorem 5.2].

If  $h$  is a cube, it is trivially the sum of two cubes. If  $h$  does not have a square factor, then after a change of variables,  $h(x, y) = xy(\alpha x + \beta y)$  with  $\alpha\beta \neq 0$ . In this case, letting  $\omega = e^{\frac{2\pi i}{3}}$ , we have the identity

$$3\alpha\beta(\omega - \omega^2)xy(\alpha x + \beta y) = (\omega^2\alpha x - \omega\beta y)^3 - (\omega\alpha x - \omega^2\beta y)^3.$$

Otherwise, after a change of variables,  $h(x, y) = x^2y$  and

$$6x^2y = (-x + y)^3 - 2y^3 + (x + y)^3.$$

Thus, in every case, a binary cubic is a sum of at most three cubes of linear forms. Sylvester's algorithm allows one to write a cubic as a sum of cubes without factoring it, see Example 4.4.

Consider now the general binary sextic

$$p(x, y) = \sum_{k=0}^6 \binom{6}{k} a_k x^{6-k} y^k, \quad a_k \in \mathbb{C}.$$

If  $p = 0$ , then there is nothing to prove. Otherwise, we may make an invertible linear change of variables, after which  $p(1, 0) = a_0 \neq 0$ . If we set

$$q(x, y) = x^2 + \left(\frac{2a_1}{a_0}\right)xy + \left(\frac{5a_0a_2 - 4a_1^2}{a_0^2}\right)y^2,$$

then, we have the expression

$$(3.1) \quad p(x, y) = a_0q(x, y)^3 + \frac{1}{a_0^5}y^3c(x, y),$$

where  $c(x, y) = c_0x^3 + 3c_1x^2y + 3c_2xy^2 + c_3y^3$ , with

$$\begin{aligned} c_0 &= 20a_0^3(2a_1^3 - 3a_0a_1a_3 + a_0^2a_3), \\ c_1 &= 5a_0^2(4a_1^2a_2 - 5a_0a_2^2 + a_0^2a_4), \\ c_2 &= 2a_0(-16a_1^5 + 40a_0a_1^3a_2 - 25a_0^2a_1a_2^2 + a_0^4a_5), \\ c_3 &= 64a_1^6 - 240a_0a_1^4a_2 + 300a_0^2a_1^2a_2^2 - 125a_0^3a_2^3 + a_0^5a_6. \end{aligned}$$

If the discriminant  $\Delta(c)$  of  $c$  is non-vanishing, then  $c(x, y)$  has distinct factors and, by Sylvester, it is a sum of two cubes of linear forms. Thus (3.1) gives  $p$  as a sum of three cubes of quadratic forms. A computation shows that

$$\Delta(c) = -540a_0^6D(p),$$

where  $D(p)$  is a polynomial in the  $a_k$ 's of degree 18 with 128 terms. It follows that  $p$  is a sum of three cubes of linear forms unless  $D(p) = 0$ . This was known for general sextics by [Re13b, Corollary 4.2].

Suppose  $D(p) = 0$ . If  $c$  is the cube of a linear form, then (3.1) gives  $p$  as a sum of two cubes. In the remaining case,  $c$  is a cubic with a square factor. As a first step, we may rewrite  $p$  as

$$p(x, y) = (ax^2 + 2bxy + cy^2)^3 + y^3(rx + sy)^2(tx + uy),$$

where  $ru - st \neq 0$ . If  $a = 0$ , then  $p(x, y) = y^3 h(x, y)$  for a cubic  $h$ , and (as noted earlier)  $h$  must be a sum of three cubes of linear forms, so  $p$  is then a sum of three cubes. We may therefore assume that  $a \neq 0$  and distinguish two cases.

First, suppose  $r = 0$  in (3). Then  $st \neq 0$ . We may scale  $y$  so that  $s = 1$ . Then, since  $t \neq 0$ , we make a further invertible change of variables  $(tx + uy, x) \mapsto (x, y)$ . This case then reduces to

$$p^{(1)}(x, y) = (ax^2 + 2bxy + cy^2)^3 + xy^5.$$

The same argument as before shows that we may assume  $a \neq 0$  for  $p^{(1)}$ .

Otherwise,  $r \neq 0$ , so we make the change of variables  $(rx + sy, y) \mapsto (x, y)$ , so that  $y^3(rx + sy)^2(tx + uy) \mapsto y^3x^2(tx + uy)$ , where  $u \neq 0$ , and we may again scale  $y$  so that  $u = 1$ . This gives the second case

$$p^{(2)}(x, y) = (ax^2 + 2bxy + cy^2)^3 + x^2y^3(tx + y).$$

For either case  $p^{(j)}$  and  $T \in \mathbb{C}$ , let  $p_T^{(j)}(x, y) = p^{(j)}(x, Tx + y)$ , so  $p^{(j)}(x, y) = p_T^{(j)}(x, -Tx + y)$ . It will suffice to write  $p_T^{(j)}$  as a sum of three cubes. Write

$$p_T^{(j)}(x, y) = \sum_{k=0}^6 \binom{6}{k} a_k(T) x^{6-k} y^k.$$

Here,  $a_k(T)$  is a polynomial in  $T$  of degree  $6 - k$ . There are at most 6 values of  $T$  which must be avoided to ensure that  $a_0(T) \neq 0$ .

We now compute  $D(p_T^{(j)})$ , with the intent of finding a value of  $T$  for which  $D(p_T) \neq 0$ . It turns out that  $D(p_T^{(j)})$  is a massive polynomial of degree 72 in the coefficients of  $p^{(j)}$ . If this polynomial is non-zero, then except for at most  $6 + 72$  values of  $T$ ,  $p_T$  (and hence  $p$ ) is a sum of three cubes. Thus, our situation reduces to deriving a contradiction from the assumption that  $D(p_T^{(j)})$  is the zero polynomial. (For computational reasons, this is why we reduced to  $p^{(j)}$  above.) This is easiest to see by considering the lowest order term in various special cases.

A computation shows that the lowest order term of  $D(p_T^{(1)})$  is  $\frac{1}{36}a^{42}T^2$ . Since  $a \neq 0$ , it follows that this is not the zero polynomial and so for all but finitely many values of  $T$ ,  $p_T^{(1)}$  is a sum of three cubes.

The computation of  $D(p_T^{(2)})$  is trickier, and we divide it into four cases. If  $c, t \neq 0$ , the lowest order term of  $D(p_T^{(2)})$  is  $-\frac{1}{45}a^{40}c^2tT$ , and since  $a, c, t \neq 0$ , this term is not zero. If  $t = 0, c \neq 0$ , the lowest order term is  $-\frac{2}{45}a^{40}c^2T^2 \neq 0$ . If  $c = 0, t \neq 0$ , the lowest order term is  $-\frac{1}{135}a^{36}t^3T^3 \neq 0$ . If  $t = c = 0$ , the lowest order term of  $D(T)$  is  $-\frac{8}{135}a^6T^6 \neq 0$ . This completes the proof.  $\square$

**Example 3.2.** In practice, this algorithm leads to some nasty expressions. We start with a specially-cooked simple one. Suppose

$$p(x, y) = x^6 + 3x^5y - 3x^4y^2 - 11x^3y^3 + 9x^2y^4 + 21xy^5 - y^6.$$

Then,

$$p(x, y) = (x^2 + xy - 2y^2)^3 + y^3(3x^2y + 9xy^2 + 7y^3).$$

Applying Sylvester's algorithm to  $h(x, y) = 3x^2y + 9xy^2 + 7y^3 = 0 \cdot y^3 + 1 \cdot 3x^2y + 3 \cdot 3xy^2 + 7 \cdot y^3$ , we see that

$$\begin{pmatrix} 0 & 1 & 3 \\ 1 & 3 & 7 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and  $2x^2 - 3xy + y^2 = (x - y)(2x - y)$ , so there exist  $\lambda_k$  so that  $h(x, y) = \lambda_1(x + y)^3 + \lambda_2(x + 2y)^3$ . Indeed,  $\lambda_1 = -1$  and  $\lambda_2 = 1$  and

$$p(x, y) = (x^2 + xy - 2y^2)^3 + y^3(x + 2y)^3 - y^3(x + y)^3.$$



**Example 3.3.** For a less trivial example, suppose  $p(x, y) = x^6 + x^5y + x^4y^2 + x^3y^3 + x^2y^4 + xy^5 + y^6$ . Then

$$p(x, y) - \left(x^2 + \frac{1}{3}xy + \frac{2}{9}y^2\right)^3 = \frac{7}{729}y^3(54x^3 + 81x^2y + 99xy^2 + 103y^3).$$

An application of Sylvester's algorithm shows that

$$\begin{aligned} 54x^3 + 81x^2y + 99xy^2 + 103y^3 &= \\ m_1(78x + (173 - \sqrt{20153})y)^3 + m_2(78x + (173 + \sqrt{20153})y)^3, \\ m_1 &= \frac{20153 + 134\sqrt{20153}}{354209128}, \quad m_2 = \frac{20153 - 134\sqrt{20153}}{354209128} \end{aligned}$$

This gives a simple sextic  $p$  as a sum of three cubes in an ugly way and gives no hint about the existence of the formula

$$p(x, y) = \sum_{\pm} \left(\frac{9 \pm \sqrt{-3}}{18}\right) \left(x^2 + \frac{1 \pm \sqrt{-3}}{2}xy + y^2\right)^3.$$

**Example 3.4.** For another example, set

$$p(x, y) = x^6 + 3x^5y + y^6.$$

We have

$$p(x, y) - (x^2 + xy - y^2)^3 = y^3(5x^3 - 3xy^2 + 2y^3)$$

and yet another application of Sylvester's algorithm gives

$$5x^3 - 3xy^2 + 2y^3 = \frac{20-9\sqrt{5}}{20}((-5 - 2\sqrt{5})x + y)^3 + \frac{20+9\sqrt{5}}{20}((-5 + 2\sqrt{5})x + y)^3.$$

On the other hand, if  $p(x, y) = x^6 + 3xy^5 + y^6$ , then the first step of the algorithm leaves us with  $c(x, y) = y^2(3x + y)$ , and we must invoke  $p_T$ . A computation shows that  $D(T)$  equals  $T^2$  times a polynomial which is irreducible over  $\mathbb{Q}$ . We take  $T = -1$ , and

$$p_{-1}(x, y) = x^6 + 3x(-x + y)^5 + (-x + y)^6 = -x^6 + 9x^5y - 15x^4y^2 + 10x^3y^3 - 3xy^5 + y^6.$$

Following the algorithm,

$$p_1(x, y) + (x^2 - 3xy - 4y^2)^3 = y^3(55x^3 - 60x^2y - 147xy^2 - 63y^3)$$

and a further invocation of Sylvester, followed by the reversed change of variables yields

$$\begin{aligned} x^6 + 3xy^5 + y^6 &= \\ (6x^2 + 11xy + 4y^2)^3 + \lambda_+(\alpha_+x^2 + \beta xy + y^2)^3 + \lambda_-(\alpha_-x^2 + \beta xy + y^2)^3 \\ \lambda_{\pm} &= \frac{4445 \pm \sqrt{5632445}}{2282}, \quad \alpha_{\pm} = \frac{6727 \pm \sqrt{5632445}}{2282}, \quad \beta_{\pm} = \frac{9009 \pm \sqrt{5632445}}{326}. \end{aligned}$$

An alternative approach is to observe that for a sextic  $p$ , there is usually a quadratic  $q$  so that  $p - q^3$  is even, i.e., is a cubic in  $\{x^2, y^2\}$ . It is enough to look at the coefficients of  $x^5y, x^3y^3, xy^5$  and solve the equations for the coefficients of  $q$ . So, it is usually a sum of two cubes of even quadratic forms. If this doesn't work, apply it to  $p_T$ .

4. COMPUTING  $k$ -RANKS

In this section, we propose a procedure to compute  $k$ -th Waring decompositions based on a geometric description of the varieties of powers as an explicit linear projection of Veronese varieties.

The variety of powers  $\mathbb{V}\mathbb{P}_{n,kd}^{(k)}$  is the image of the regular map

$$\nu_{n,kd}^{(k)} : \mathbb{P}(S_d) \rightarrow \mathbb{P}(S_{kd}), [g] \mapsto [g^k].$$

Denote by  $B_i$  the set of multi-indices  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$  and  $|\alpha| = \alpha_1 + \dots + \alpha_n = d$ .

Let  $T$  be the coordinate ring of  $\mathbb{P}(S_d)$  whose variables are labelled by  $B_d$ , i.e.,  $T = \mathbb{C}[Y_\alpha : \alpha \in B_d]$ . In particular, if  $S = \mathbb{C}[x_1, \dots, x_n]$ , we identify  $T_1$  with  $S_d$  by setting  $Y_\alpha = x^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}$ . Given  $g \in S_d$ , we denote by  $l_g$  the corresponding linear form in  $T_1$ . Thus, we consider the usual Veronese embedding

$$\nu_{N,k} : \mathbb{P}(T_1) \rightarrow \mathbb{P}(T_k), [l] \mapsto [l^k],$$

whose image is the Veronese variety  $\text{Ver}_{N,k}$ , where  $N = \binom{d+n-1}{n-1} - 1$ .

The substitution  $Y_\alpha \mapsto x^\alpha$  gives a linear projection  $\pi_{n,kd}^{(k)} : \mathbb{P}(T_k) \rightarrow \mathbb{P}(S_{kd})$  and we get the diagram

$$\begin{array}{ccc} \mathbb{P}(T_1) \ni [l_g] & \xrightarrow{\nu_{N,k}} & [l_g^k] \in \text{Ver}_{N,k} \subset \mathbb{P}(T_k) \\ \parallel & & \downarrow \pi_{n,kd}^{(k)} \\ \mathbb{P}(S_d) \ni [g] & \xrightarrow{\nu_{n,kd}^{(k)}} & [g^k] \in \mathbb{V}\mathbb{P}_{n,d}^{(k)} \subset \mathbb{P}(S_{kd}) \end{array}$$

The center of the projection  $\pi_{n,kd}^{(k)}$  has an explicit interpretation in terms of the ideal defining a specific Veronese variety. Recall that Veronese varieties are determinantal, see [Pu98]. In particular, the Veronese variety  $\text{Ver}_{n,d} \subset \mathbb{P}(S_d)$  is generated by the  $2 \times 2$  minors of the  $i$ -th catalecticant matrix

$$\text{cat}_i(n, d) = (Y_{\beta_1 + \beta_2})_{\beta_1, \beta_2} \in \text{Mat}^{B_i \times B_{d-i}}, \text{ for any } i = 1, \dots, d-1,$$

where  $Y_\alpha$  are coordinates of the space  $\mathbb{P}(S_d)$  corresponding to the standard monomial basis of  $S_d$ .

**Lemma 4.1.** *In the same notation as above, the center of the linear projection  $\pi_{n,kd}^{(k)}$  is*

$$E = \mathbb{P}([I_{n,d}]_k),$$

where  $[I_{n,d}]_k$  is the  $k$ -th homogeneous part of the ideal  $I_{n,d} \subset T$  defining the Veronese variety  $\text{Ver}_{n,d}$ .

*Proof.* The center of the projection is given by the linear space of homogeneous polynomials  $f \in T_k$  annihilated by the substitution  $Y_\alpha \mapsto x^\alpha$ , for any  $\alpha \in B_d$ . This is, by definition, the set of forms of degree  $k$  vanishing on the Veronese variety  $\text{Ver}_{n,d}$ .  $\square$

It is easy to see that the Veronese variety  $\text{Ver}_{N,k} \subset \mathbb{P}(T_k)$  does not intersect the center of the projection. Therefore, the projection  $\pi_{n,kd}^{(k)}$  is a regular map restricted to  $\text{Ver}_{N,k}$  and it maps the Veronese variety onto the variety of powers  $\mathbb{V}\mathbb{P}_{n,d}^{(k)}$ . Hence, by [Sh94, Theorem 7, Section I.5.3], we have that the restriction  $\pi_{n,kd}^{(k)} : \text{Ver}_{N,k} \rightarrow \mathbb{V}\mathbb{P}_{n,d}^{(k)}$  is a finite surjective map.

Now, to find a  $k$ -th Waring decomposition of a given form  $f$  of degree  $kd$ , we want to project using  $\pi_{n,kd}^{(k)}$  a classical Waring decomposition in  $T_k$  of some element in the fiber over  $f$ . A similar idea has been also used in [BB11].

Given  $f \in S_{kd}$ , we denote by  $\mathcal{F}_f$  the fiber  $(\pi_{n,kd}^{(k)})^{-1}([f])$ . In particular, if  $f_0 \in T_k$  is any element such that  $\pi_{n,kd}^{(k)}([f_0]) = [f]$ , then, we have  $\mathcal{F}_f = \langle \{[f_0]\} \cup E \rangle \subset \mathbb{P}(T_k)$ .

**Lemma 4.2.** *Let  $f \in S_{kd}$ . Then,*

$$\mathrm{rk}_k(f) = \min \{ \mathrm{rk}(g) \mid [g] \in \mathcal{F}_f \setminus E \subset \mathbb{P}(T_k) \}.$$

*Proof.* Given a Waring decomposition  $g = \sum_{i=1}^r l_i^k$  of an element  $g \in \mathcal{F}_f \setminus E$ , by substituting the  $Y$ 's with the corresponding monomials in the  $x$ 's, we obtain a Waring decomposition of  $f$  as sum of  $k$ -th powers. Therefore,  $\mathrm{rk}_k(f) \leq \min \{ \mathrm{rk}(g) \mid [g] \in \mathcal{F}_f \setminus E \}$ .

Conversely, let  $f = \sum_{i=1}^r g_i^k$  be a minimal  $k$ -th Waring decomposition representing  $f$ , where  $g_i = \sum_{|\alpha|=d/k} c_{i,\alpha} x^\alpha$ , for  $i = 1, \dots, r$ . Set  $l_{g_i} := \sum_{|\alpha|=d/k} c_{i,\alpha} Y_\alpha \in T_1$ . Then,  $\sum_{i=1}^r l_{g_i}^k$  is a sum of  $k$ -th powers of linear forms of an element in  $\mathcal{F}_f \setminus E$  since its image with respect to the linear projection coincides with  $[f]$ . Therefore,  $\mathrm{rk}_k(f) \geq \min \{ \mathrm{rk}(g) \mid [g] \in \mathcal{F}_f \setminus E \}$ .  $\square$

Hence, we reduce our problem of computing a  $k$ -th Waring decomposition of  $f$  to the problem of minimizing the classical Waring rank in  $\mathcal{F}_f \setminus E$ . Several algorithms for computations of Waring ranks and Waring decompositions have been proposed in the literature. Already, J.J. Sylvester [Sy51a, Sy51b] considered the case of binary forms, see [CS11] for a modern exposition. Other solutions can be found, for example, in [BCMT10, BGI11, OO13], but they all work efficiently only under some additional constraints on the considered polynomial. However, due to Lemma 4.2, we can also use these algorithms to study  $k$ -th Waring ranks.

As an illustration, we apply Sylvester's *catalecticant method* to compute  $k$ -th ranks of binary forms.

**4.1. Sylvester's catalecticant method.** Let us recall the basic notions of Apolarity Theory.

Let  $S = \mathbb{C}[x_1, \dots, x_n]$  and  $R = \mathbb{C}[y_1, \dots, y_n]$  be standard graded polynomial rings. For  $i \leq 0$ , we set  $S_i$  and  $R_i$  to be 0. We consider the *apolar action*  $\circ$  where polynomials in  $R$  act over  $S$  as partial differentials and, for any polynomial  $f \in S_d$ , we define the *apolar ideal* as  $f^\perp = \{g \in R \mid g \circ f = 0\}$ . More explicitly, the homogeneous part in degree  $i$  of  $f^\perp$  is given by the kernel of the  *$i$ -th catalecticant matrix* of  $f$

$$\mathrm{cat}_i(f) : R_i \rightarrow S_{d-i}, \quad y^\alpha \mapsto \frac{\partial^i f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

The key lemma which relates classical Waring decompositions to ideals of reduced points is as follows.

**Lemma 4.3** (Apolarity Lemma [IK, Lemma 1.15]). *Let  $f \in S$ . Then, the following are equivalent:*

- (1)  $f$  has a decomposition of length  $s$  as sum of powers of linear forms;
- (2)  $f^\perp$  contains an ideal defining a set of reduced points of cardinality  $s$ .

**Example 4.4** (*Sylvester's algorithm*). Given a binary form  $f$ , compute the apolar ideal  $f^\perp$ . Since any apolar ideal  $f^\perp$  is artinian and Gorenstein, in the case of two variables, it is also a complete intersection. In particular,  $f^\perp = (g_1, g_2)$  where  $\deg(g_1) + \deg(g_2) = \deg(f) + 2$ .

If  $g_1$  is square-free, then the ideal  $(g_1)$  defines a set of reduced points in  $\mathbb{P}^1$  and then, by Apolarity Lemma, we conclude that the Waring rank of  $f$  is equal to  $\deg(g_1)$ . Otherwise, it is possible to find a square-free element in  $f^\perp$  of degree  $\deg(g_2)$  and we conclude that the Waring rank of  $f$  is  $\deg(g_2)$ .

Two examples below illustrate how we can use this method to study  $k$ -th Waring ranks.

**Example 4.5** (Upper bound). Here, we show how to compute a  $k$ -th Waring decomposition and obtain upper bounds on  $k$ -th ranks. Let  $f = x^6 y^2 - x^3 y^5 + x^2 y^6 - x y^7$  be a binary octic.

We pick an element in the fiber, for example  $f_0 = Y_0^3 Y_2 - Y_0 Y_1 Y_2^2 + Y_0 Y_2^3 - Y_1 Y_2^3$ . In this case, we have

$$\left( f_0^\perp \right)_2 = \langle Y_1^2, Y_0^2 - 9Y_0 Y_1 + 3Y_1 Y_2 \rangle.$$

This defines a non-reduced 0-dimensional scheme, hence, we have that the rank of  $f_0$  is at least 5. Hence, we look for another element in the fiber having rank 4.

The general element of the fiber  $\mathcal{F}_f \setminus E$  is

$$F_c = \lambda f_0 - (Y_{20}Y_{02} - Y_{11}^2)(c_0Y_{20}^2 + c_1Y_{20}Y_{11} + c_2Y_{20}Y_{02} + c_3Y_{11}^2 + c_4Y_{11}Y_{02} + c_5Y_{02}^2),$$

where  $\lambda \in \mathbb{C}$  and  $c = (c_0, \dots, c_5) \in \mathbb{C}^6$ . Since we want  $F_c \notin E$ , we may assume  $\lambda = 1$ . Now, we consider the apolar action of  $\mathbb{C}[y_{20}, y_{11}, y_{02}]$  on  $\mathbb{C}[Y_{20}, Y_{11}, Y_{02}]$ , and we get the 2-nd catalecticant matrix

$$\text{cat}_2(F_c) = \begin{pmatrix} 0 & 0 & 6c_0 + 6 & -4c_0 & 2c_1 & 4c_2 \\ 0 & -4c_0 & 2c_1 & -6c_1 & -2c_2 + 2c_3 & 2c_4 - 2 \\ 6c_0 + 6 & 2c_1 & 4c_2 & -2c_2 + 2c_3 & 2c_4 - 2 & 6c_5 + 6 \\ -4c_0 & -6c_1 & -2c_2 + 2c_3 & -24c_3 & -6c_4 & -4c_5 \\ 2c_1 & -2c_2 + 2c_3 & 2c_4 - 2 & -6c_4 & -4c_5 & -6 \\ 4c_2 & 2c_4 - 2 & 6c_5 + 6 & -4c_5 & -6 & 0 \end{pmatrix}.$$

Now, we notice that a point vanishing all the  $5 \times 5$  minors is

$$\bar{c} = (0, 0, 0, 1, 0) \in \mathbb{C}^6,$$

i.e.,  $F_{\bar{c}} = f - (Y_1^3Y_2 - Y_0Y_1Y_2^2) = Y_0^3Y_2 + Y_0Y_2^3 - Y_1^3Y_2 - Y_1Y_2^3$ . Then, the kernel of  $\text{cat}_2(F_{\bar{c}})$  is given by

$$\ker(\text{cat}_2(F_{\bar{c}})) = \langle Q_1, Q_2 \rangle = \langle Y_0Y_1, Y_0^2 + Y_1^2 - Y_2^2 \rangle.$$

The ideal  $I = (Q_1, Q_2)$  defines a set of four points; in particular,

$$\mathbb{X} = \{(0 : 1 : 1), (0 : 1 : -1), (1 : 0 : 1), (1 : 0 : -1)\}.$$

By solving an easy linear system, we obtain the following expression <sup>2</sup>

$$8f = (xy - y^2)^4 - (x^2 - y^2)^4 + (x^2 + y^2)^4 - (xy + y^2)^4.$$

**Remark 4.6.** Recall the following important properties on the Hilbert function of any homogeneous ideal  $I_{\mathbb{X}} \subset S$  defining a set of reduced points in projective space and which is contained in the apolar ideal  $f^\perp$  of any given polynomial  $f$ :

- (1)  $\text{HF}_{R/I_{\mathbb{X}}}$  is strictly increasing until it becomes constant and equal to the cardinality  $|\mathbb{X}|$ ;
- (2)  $\text{HF}_{R/I_{\mathbb{X}}}(i) \geq \text{HF}_{R/f^\perp}(i)$ , for all  $i \geq 0$ .

Hence, since the Hilbert function  $\text{HF}_{R/f^\perp}(i)$  is equal to the dimension of the image of  $\text{cat}_i(f)$  which is the rank of the matrix, we have that the rank of a catalecticant matrix  $\text{cat}_i(f)$  is always a lower bound for the Waring rank of  $f$ , for any  $i = 1, \dots, \deg(f)$ .

**Example 4.7** (Lower bound). Using the latter remark, we can also find lower bounds on  $k$ -th ranks. We consider the binary octic  $m = xy^7 \in S_8$ . The monomial  $M_0 = Y_{11}Y_{02}^3$  satisfies  $\pi_{2,2}^{(4)}([M_0]) = [m]$  and it is well-known that its Waring rank is  $\text{rk}(M_0) = 4$ . Therefore,  $\text{rk}_4(m) \leq 4$ . We use Lemma 4.2 to prove that the equality holds. A general element in  $\mathcal{F}_m \setminus E$  is given by

$$F_c = \lambda(Y_{11}Y_{02}^3) - (Y_{20}Y_{02} - Y_{11}^2)(c_0Y_{20}^2 + c_1Y_{20}Y_{11} + c_2Y_{20}Y_{02} + c_3Y_{11}^2 + c_4Y_{11}Y_{02} + c_5Y_{02}^2),$$

where  $\lambda \in \mathbb{C}$  and  $c = (c_0, \dots, c_5) \in \mathbb{C}^6$ . Since we want  $F_c \notin E$ , we may assume  $\lambda = 1$ . We need to prove that, for *any* choice of the  $c_i$ 's, the Waring rank of  $F_c$  is at least 4. Since  $M_0$  has rank equal to 4, we can

<sup>2</sup>A *Macaulay2* script describing this example can be found in the personal webpage of the second author at <https://sites.google.com/view/alessandrooneto/research/list-of-papers>.

restrict to the case where not all the  $c_\alpha$ 's are equal to zero. Now, the 2-nd catalecticant matrix of  $F_c$  is

$$\text{cat}_2(F_c) = \begin{pmatrix} 0 & 0 & 6c_0 & -4c_0 & 2c_1 & 4c_2 \\ 0 & -4c_0 & 2c_1 & -6c_1 & -2c_2 + 2c_3 & 2c_4 \\ 6c_0 & 2c_1 & 4c_2 & -2c_2 + 2c_3 & 2c_4 & 6c_5 \\ -4c_0 & -6c_1 & -2c_2 + 2c_3 & -24c_3 & -6c_4 & -4c_5 \\ 2c_1 & -2c_2 + 2c_3 & 2c_4 & -6c_4 & -4c_5 & 6 \\ 4c_2 & 2c_4 & 6c_5 & -4c_5 & 6 & 0 \end{pmatrix}$$

Computing the radical of the ideal generated by the  $4 \times 4$  minors with the algebra software *Macaulay 2* [GS] gives the ideal  $(c_0, c_1, c_2, c_3, c_4)$ . Therefore, if  $c_i \neq 0$ , for some  $0 \leq i \leq 4$ , we have that the rank of  $\text{cat}_2(F_c)$  is at least 4 and, by Remark 4.6, the Waring rank of  $F_c$  is at least 4.

Now, if we assume that  $c_0 = \dots = c_4 = 0$  and also  $c_5 = 0$ , we have that  $F_c = Y_{11}Y_{02}^3$  which has rank 4. If  $c_5 \neq 0$ , we are left with the matrix

$$\text{cat}_2(F_c) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6c_5 \\ 0 & 0 & 0 & 0 & 0 & -4c_5 \\ 0 & 0 & 0 & 0 & -4c_5 & 6 \\ 0 & 0 & 6c_5 & -4c_5 & 6 & 0 \end{pmatrix}$$

which has rank 3 and the kernel is given by

$$\ker(\text{cat}_2(F_c)) = \langle Q_1, Q_2, Q_3 \rangle = \langle y_{20}^2, y_{20}y_{11}, 2y_{11}y_{02} + 3y_{11}^2 \rangle.$$

The ideal  $(Q_1, Q_2, Q_3)$  defines a curvilinear 0-dimensional scheme of length 3 with support at the point  $(0 : 0 : 1) \in \mathbb{P}^2$ . Therefore, as explained in [BGI11, Theorem 4], the rank of  $F_c$  is 7.

Hence, all  $F_c \in \mathcal{F}_m \setminus E$  have rank at least 4 and, by Lemma 4.2,  $xy^7$  has 4-th rank at least 4. Since, as we noticed,  $xy^7 = (xy^3)y^4$  and the monomial  $xy^3$  has Waring rank 4, we conclude that  $\text{rk}_4(xy^7) = 4$ .<sup>3</sup>

In [BCMT10, BGI11, OO13], the authors give other linear algebra methods to compute Waring ranks and Waring decompositions. These methods can be used in our procedure instead of the classical catalecticant method to find better bounds on  $k$ -th Waring ranks. However, we have to observe that these methods are not always effective and might fail. Also, if we consider polynomials with higher number of variables or higher degrees, the dimension of the fiber grows quickly and this can make the computation very difficult and heavy.

#### APPENDIX A. FRÖBERG'S CONJECTURE AND GENERIC RANKS

In [Fr85], R. Fröberg studied Hilbert series of generic ideals.

**Conjecture A.1.** [Fröberg's Conjecture, [Fr85]] *Generic ideals are Hilbert generic, i.e., for  $I = (g_1, \dots, g_s)$  where  $g_1, \dots, g_s$  are generic forms of degrees  $d_1, \dots, d_s$  in  $n$  variables, then,*

$$(A.1) \quad \text{HS}_{S/I}(t) = \left\lceil \frac{\prod_{i=1}^s (1 - t^{d_i})}{(1 - t)^n} \right\rceil \in \mathbb{Z}[[t]],$$

where  $\lceil \cdot \rceil$  is the truncation of the power series at its first non-positive coefficient.

Fröberg's Conjecture attracted the attention of many mathematicians, but, at the moment it is known to be true only in the following cases. For  $s \leq n$  (easy exercise),  $s = n + 1$  (by R. Stanley, [St78]),  $n = 2$  (by R. Fröberg, [Fr85]) and  $n = 3$  (by D. Anick, [An86]). More recently, G. Nenashev settled a large number

<sup>3</sup>A *Macaulay2* script describing this example can be found in the personal webpage of the second author at <https://sites.google.com/view/alessandrooneto/research/list-of-papers>.

of cases [Ne17]. In [Ni16], L. Nicklasson suggested the following more general version that allow us to directly relate this commutative algebra problem to our computations on generic  $k$ -ranks.

**Conjecture A.2.** [Ni16] *Let  $k \geq 2$ . Let  $I = (g_1^{k-1}, \dots, g_s^{k-1})$  be an ideal generated by  $(k-1)$ -th powers of generic homogeneous polynomials of degree  $d > 1$ . Then,  $I$  is Hilbert generic.*

By using the same approach as in our proof of Theorem 2.3, we prove the following.

**Theorem A.3.** *Conjecture A.2 implies Conjecture 1.2.*

First, we need the following lemma.

**Lemma A.4.** *Let  $d, k, n \geq 2$  be positive integers and  $H(t) = \frac{(1-t^{d(k-1)})^s}{(1-t)^n} =: \sum_{i \geq 0} H_i t^i$ . Then,*

$$s_{kd} \leq s_{kd-1} \leq \dots \leq s_{d(k-1)},$$

where  $s_i := \min\{s \mid H_i \leq 0\}$ .

*Proof.* For  $n = 2$ , the claim is part of the proof of Theorem 2.3. Then, we assume  $n \geq 3$ .

For  $k \geq 3$ , the statement directly follows from the fact that

$$H_i = \binom{i+n-1}{n-1} - s \binom{i-d(k-1)+n-1}{n-1},$$

for any  $d(k-1) \leq i \leq dk$ . For  $k = 2$ , the same formula holds for  $i < dk$ , but we have

$$H_{2d} = \binom{2d+n-1}{n-1} - s \binom{d+n-1}{n-1} + \binom{s}{2}.$$

Hence, we still have to prove that, for  $k = 2$ , we have also  $s_{2d} \leq s_{2d-1}$ .

Since  $s_{2d-1} = \left\lceil \frac{\binom{2d+n-2}{n-1}}{\binom{n+d-2}{n-1}} \right\rceil$ , it is enough to prove that

$$H_{2d} \leq 0, \text{ for } s = \frac{\binom{2d+n-2}{n-1}}{\binom{n+d-2}{n-1}}.$$

Then, we need to show that

$$\begin{aligned} 2 \binom{n+d-2}{n-1}^2 \binom{n+2d-1}{n-1} - 2 \binom{n+2d-2}{n-1} \binom{n+d-1}{n-1} \binom{n+d-2}{n-1} + \\ + \binom{n+2d-2}{n-1}^2 - \binom{n+2d-2}{n-1} \binom{n+d-2}{n-1} \leq 0. \end{aligned}$$

It is easy to see that  $\binom{n+d-2}{n-1} \binom{n+2d-1}{n-1} - \binom{n+2d-2}{n-1} \binom{n+d-2}{n-1} = -\binom{n+d-2}{n-1} \binom{n+2d-2}{n-2}$ ; hence, we need to show that

$$(A.2) \quad \binom{n+2d-2}{n-1}^2 - \binom{n+2d-2}{n-1} \binom{n+d-2}{n-1} \leq 2 \binom{n+d-2}{n-1}^2 \binom{n+2d-2}{n-2}.$$

For  $n = 3$  and  $d = 2$ , we can directly check that this inequality holds. For  $n \geq 3$ ,  $d \geq 2$  and  $nd > 6$ , we show the following inequality, stronger than (A.2),

$$(A.3) \quad \binom{n+2d-2}{n-1}^2 \leq 2 \binom{n+d-2}{n-1}^2 \binom{n+2d-2}{n-2}$$

By expanding the binomials in (A.3), we have

$$(A.4) \quad \frac{((n+2d-2)!)^2}{((n-1)!)^2 ((2d-1)!)^2} \leq 2 \frac{((n+d-2)!)^2}{((n-1)!)^2 ((d-1)!)^2} \cdot \frac{((n+2d-2)!)^2}{((n-2)!)^2 ((2d)!)^2}.$$

Simplifying the latter inequality, we are left with

$$(A.5) \quad \prod_{i=1}^{n-2} (2d+i) \leq \frac{1}{2} \cdot \prod_{i=1}^{n-2} \frac{(d+i)^2}{i}.$$

For  $n = 3$ , we need to show that

$$2d+1 \leq \frac{1}{2} (d+1)^2 \iff 4d^2 - 8d - 4 \geq 0,$$

which holds for  $d \geq 3$ . Now, for any  $d \geq 2$ , we proceed by induction on  $n \geq 3$ . In particular, if  $d \geq 3$ , we use the case  $n = 3$  for the base of induction, while, if  $d = 2$ , we use  $n = 4$  as the base step, which can be checked directly. Assume that formula (A.5) holds for  $n$ . Since

$$2d+n-1 \leq \frac{(d+n-1)^2}{n-1} = \frac{d^2}{n-1} + 2d+n-1,$$

we conclude that (A.5) holds also for  $n+1$ .

This concludes the proof.  $\square$

*Proof of Theorem A.3.* Let  $k, d, n \geq 2$  be integers. By Lemma 2.1 and (2.1), we know that the dimension of the  $s$ -th secant variety of  $V_{n,kd}^{(k)}$  is given by the dimension of the homogeneous part in degree  $kd$  of the ideal  $I = (g_1^{k-1}, \dots, g_s^{k-1})$ , where the  $g_i$ 's are generic forms of degree  $d$ . Moreover, by Conjecture A.2, we have  $\text{HS}_{S/I}(t) = [H(t)]$ , where  $H(t) = \frac{(1-t^{d(k-1)})^s}{(1-t)^n} = \sum_{i \geq 0} H_i(t)$ . In particular,

$$H_{kd} = \begin{cases} \binom{2d+n-1}{n-1} - s \binom{d+n-1}{n-1} + \binom{s}{2}, & \text{for } k = 2; \\ \binom{kd+n-1}{n-1} - s \binom{d+n-1}{n-1}, & \text{for } k \geq 3. \end{cases}$$

By Lemma A.4, we have  $\text{HF}_{S/I}(kd) = \max\{0, H_{kd}\}$ . Thus, Conjecture 1.2 directly follows from (2.1).  $\square$

In particular, by the result of D. Anick, we obtain the following result about 2-ranks.

**Corollary A.5.** *Conjecture 1.2 holds in the case of sums of squares of ternary forms of even degree, i.e.,*

$$\text{rk}_2^\circ(3, 2d) = \left\lceil \frac{\binom{2d+2}{2}}{\binom{d+2}{2}} \right\rceil,$$

except for  $d = 1, 3, 4$ , where it is  $\left\lceil \frac{\binom{2d+2}{2}}{\binom{d+2}{2}} \right\rceil + 1$ .

**Remark A.6.** From the result of R. Stanley, we know that Fröberg's Conjecture holds for  $s \leq n+1$ . Hence, from Terracini's Lemma and the proof of Theorem A.3, we can conclude that, for  $s \leq n+1$ ,

$$\text{codim } \sigma_s V_{n,kd}^{(k)} = \left\lceil \frac{(1-t^{d(k-1)})^s}{(1-t)^n} \right\rceil.$$

**Remark A.7.** In [FOS12], the authors prove that, for any  $n, k, d \geq 2$  positive integers, we have  $\text{rk}_k^\circ(n, kd) \leq k^n$  and, for  $d \gg 0$ , this bound is sharp. Hence, fixed  $n$  and  $k$ , we have only finitely many cases left to compute the generic  $k$ -rank. In these case, if  $n$  and  $k$  are not too large, Fröberg's Conjecture can be checked by computer. For example, it is possible to conclude that,

$$\text{rk}_2^\circ(4, 2d) = \left\lceil \frac{\binom{2d+3}{3}}{\binom{d+3}{3}} \right\rceil,$$

except for  $d = 1, 2$ , where  $\text{rk}_2^\circ(4, 2d) = \left\lceil \frac{\binom{2d+3}{3}}{\binom{d+3}{3}} \right\rceil + 1$ . See [On16, Section 3.3.2] for more details.

APPENDIX B.  $k$ -RANKS OF MONOMIALS

In this section, we continue the study started in [CO15] about  $k$ -ranks of monomials.

**Proposition B.1.** *Let  $m = x_1^{a_1} \cdots x_n^{a_n}$  be a monomial of degree  $dk$  in  $S$  and suppose that  $(k-2)n \leq d$ . Then,  $\text{rk}_k(m) \leq k$ .*

*Proof.* It is enough to show that the monomial can be written as  $m = m_1 m_2^{k-1}$  with  $|m_1| = |m_2| = d$ , since it follows from a classical well-known result, e.g., see [Re13b, Theorem 5.4], that the Waring rank of the monomial  $xy^{k-1}$  is equal to  $k$ .

For each  $i$ , write  $a_i = q_i(k-1) + r_i$  with  $0 \leq r_i \leq k-2$ . We have  $a_1 + \cdots + a_n = d(k-1) + d$ , so  $a_1 + \cdots + a_n$  is equivalent to  $d$  modulo  $(k-1)$ ; that is,  $r_1 + \cdots + r_n \equiv_{k-1} d$ . From  $r_1 + \cdots + r_n \leq (k-2)n$  and the assumption that  $(k-2)n \leq d$ , we get  $r_1 + \cdots + r_n \leq d$ . It follows that  $r_1 + \cdots + r_n + b(k-1) = d$  for some non-negative integer  $b$ .

We have  $q_1 + \cdots + q_n = (dk - (r_1 + \cdots + r_n))/(k-1)$  and  $b = (d - (r_1 + \cdots + r_n))/(k-1)$ , so  $q_1 + \cdots + q_n \geq b$ . Thus, we can choose  $b_1, \dots, b_n$  such that  $b_1 + \cdots + b_n = b$  and such that  $q_i \geq b_i$  for  $i = 1, \dots, n$ .

Take  $m_1 = x_1^{r_1+b_1(k-1)} \cdots x_n^{r_n+b_n(k-1)}$  and let  $m_2 = x_1^{q_1-b_1} \cdots x_n^{q_n-b_n}$ . Then,  $m = m_1 m_2^{k-1}$  and we conclude the proof.  $\square$

**Example B.2.** Let  $n = 3, k = 4, d = 6$ , i.e., ternary monomials of degree 24 decomposed as sums of 4-th powers. Consider  $m = x_1^3 x_2^{10} x_3^{11}$ . Hence, from the construction in the proof of the previous proposition, we have that  $r_1 = 0, r_2 = 1, r_3 = 2, q_1 = 1, q_2 = 3, q_3 = 3, r_1 + r_2 + r_3 = 3$ , so  $b = 1$ . Choose  $b_2 = 1$ . Now, we write  $m_1 = x_2^4 x_3^2, m_2 = x_1 x_2^2 x_3^3$  and  $m = m_1 m_2^3$ , so the rank is at most four.

## APPENDIX C. A CANONICAL FORM FOR BINARY FORMS

The following results were obtained as a biproduct of the present study and they look similar to Sylvester's original result and other known canonical forms, see [Re13a]. Although they are not immediately relevant to our main topic, we include them here for the sake of completeness.

**Theorem C.1.** *Given positive integers  $k \geq 2$  and  $d \geq 1$ ,*

(i) *one can uniquely present a general binary form  $p := p(x, y)$  of degree  $kd$  as*

$$(C.1) \quad p = p_0^k + y^d p_1^{k-1} + y^{2d} p_2^{k-2} + \cdots + y^{(k-1)d} p_{k-1}^2 + y^{(k-2)d} p_{k-1},$$

*where every  $p_j, j = 0, 2, \dots, k-2$  is a binary form of degree  $d$  with term  $y^d$  missing and  $p_{k-1}$  is a binary form of degree  $d$  (with no additional restrictions); each  $p_i^{k-i}$  is uniquely defined which implies that  $p_i$  itself is defined up to the  $(k-i)$ -th root of unity:*

(ii) *one can (non-uniquely) present any binary form  $p := p(x, y)$  of degree  $kd$  containing the monomial  $x^{kd}$  as (C.1), where every  $p_j, j = 0, 2, \dots, k-1$  is a binary form of degree  $d$ ;*

**Remark C.2.** Observe that in case (i) the number of parameters in the right-hand side of (C.1) equals  $kd+1$  which coincides with the dimension of the linear space of binary forms of degree  $kd$ . In other words the right-hand side of (C.1) gives a canonical form of presentation of a general binary form of degree  $kd$ , comp. [Re13a].

*Proof.* To prove (i), we will use induction on the power  $k \geq 1$ .

Induction base. For  $k = 1$  and any  $d \geq 1$ , the statement is trivial since one can take  $p = p_0$  in notation of Theorem C.1.

Induction step. Assume that (i) is settled for all  $d \geq 1$  and up to  $k-1 \geq 1$ . Now given a binary form  $p$  of degree  $kd$ , denote by  $\hat{p} = a_0 x^{kd} + a_1 x^{kd-1} y + \cdots + a_{d-1} x^{(k-1)d+1} y^{d-1}$  the form obtained by truncation



of  $p$  modulo all terms of the form  $x^\ell y^{kd-\ell}$  with  $\ell \geq d$ . If  $a_0 \neq 0$  (which we can assume since  $p$  is general), then one can find a form  $p_0$  of degree  $d$  such that

$$p_0^k = \widehat{p}$$

modulo the same terms  $x^\ell y^{kd-\ell}$  with  $\ell \geq d$ . Indeed, set

$$p_0 = a_0^{1/d} \left( x^d + \frac{1}{ka_0} (a_1 x^{d-1} y + a_2 x^{d-2} y^2 + \cdots + a_{d-1} x y^{d-1}) \right).$$

One can easily check that  $p_0^k$  satisfies the above relation. Observe that  $p - \widehat{p}$  is divisible by  $y^d$ . For the quotient  $(p - \widehat{p})/y^d$  we obtain the same situation with  $k$  substituted by  $k - 1$  and we can apply induction under the assumption that  $(p - \widehat{p})/y^d$  contains the term  $x^{k-1}d$ . Since we assume from the beginning that  $p$  is general and our algorithm is deterministic, we can assume that the necessary terms are non-vanishing on each step of induction decreasing  $k$  to 1. (In fact the condition of generality can be written down rather explicitly in terms of non-vanishing of  $k$  distinct discriminants.)

(ii) As we mentioned above, the only problem with the above representation under the assumption that  $p$  contains  $x^{kd}$  is that  $(p - \widehat{p})/y^d$  can miss the leading term  $x^{k-1}d$ . But if we allow to use  $p_0$  with non-vanishing term  $y^d$ , we can always obtain the latter condition. But then the choice of  $p_0$  will be non-unique. □

**Corollary C.3.** *Any univariate polynomial  $p(x)$  of degree at most  $kd$  can be presented in the form*

$$\lambda x^{kd} + p_0^k(x) + p_1^{k-1}(x) + \cdots + p_{k-2}^2(x) + p_{k-1}(x),$$

where  $\lambda$  is a complex number and  $p_j(x)$ ,  $j = 0, \dots, k - 1$  are univariate polynomials of degree at most  $d$ . (Observe the similarity of the latter formula with (1.2).)

*Proof.* If  $p(x)$  is a polynomial of degree exactly  $kd$ , then Corollary C.3 is exactly statement (ii) of Theorem C.1] in the non-homogeneous setting, i.e., for  $y = 1$ . If  $\deg p(x) < kd$ , then adding a term  $\lambda x^{kd}$ , we obtain the previous situation. □

## REFERENCES

- [AH95] J. Alexander and A. Hirschowitz, “Polynomial interpolation in several variables”, *Journal of Algebraic Geometry* **4**(2):201–222, (1995).
- [An86] D. Anick, “Thin algebras of embedding dimension three”, *Journal of Algebra* **100**(1):235–259, (1986).
- [BB11] E. Ballico, and A. Bernardi. ”On the  $X$ -rank with respect to linear projections of projective varieties.” *Mathematische Nachrichten* **284**(17-18): 2133–2140 (2011).
- [BCMT10] J. Brachat, P. Comon, B. Mourrain, and E. Tsigaridas, “Symmetric tensor decomposition” *Linear Algebra and its Applications*, 433(11-12): 1851-1872, (2010).
- [BGI11] A. Bernardi, A. Gimigliano, and M. Idà. ”Computing symmetric rank for symmetric tensors.” *Journal of Symbolic Computation* **46**(1): 34-53, (2011).
- [CO15] E. Carlini, and A. Oneto. “Monomials as sums of  $k$ -th powers of forms.” *Communications in Algebra* **43**(2): 650–658 (2015).
- [CS11] G. Comas, and M. Seiguer. ”On the rank of a binary form.” *Foundations of Computational Mathematics* **11**(1): 65-78, (2011).
- [FOS12] R. Fröberg, G. Ottaviani, and B. Shapiro, “On the Waring problem for polynomial rings”, *Proceedings of the National Academy of Sciences*, **109**(15): 5600–5602 (2012).
- [Fr85] R. Fröberg, “An inequality for Hilbert series of graded algebras”, *Mathematica Scandinavica* **56**: 117–144 (1985).
- [FL91] R. Fröberg and C. Löfwall, On Hilbert series for commutative and noncommutative graded algebras, *Journal of Pure and Applied Algebra*, **76**(1): 33–38, (1991).
- [IK] A. Iarrobino, and V. Kanev. “Power sums, Gorenstein algebras, and determinantal loci”, *Springer*, (2006).
- [GS] D. Grayson, M. Stillman, “Macaulay 2, a software system for research in algebraic geometry”, *available at* [www.math.uiuc.edu/Macaulay2](http://www.math.uiuc.edu/Macaulay2).

- [GS98] A.V. Geramita, and H. Schenck, “Fat points, inverse systems, and piecewise polynomial functions”, *Journal of Algebra* **204**(1):116–128 (1998).
- [LSVB13] T.H. Le, L. Sorber, and M. Van Barel. “The Pythagoras number of real sum of squares polynomials and sum of square magnitudes of polynomials.” *Calcolo* **50.4**: 283-303 (2013).
- [Ne17] G. Nenashev, “A note on Fröberg’s conjecture for forms of equal degrees.” *Comptes Rendus Mathématique* 355.3 (2017): 272-276.
- [Ni16] L. Nicklasson. “On the Hilbert series of ideals generated by generic forms.” *Communications in Algebra*, **45**(8): 3390–3395 (2016).
- [OO13] L. Oeding, and G. Ottaviani. ”Eigenvectors of tensors and algorithms for Waring decomposition.” *Journal of Symbolic Computation* 54 (2013): 9-35.
- [On16] A. Oneto, “Waring-type problems for polynomials: Algebra meets Geometry”, *Doctoral dissertation, Department of Mathematics, Stockholm University, available at <http://www.dissertations.se/dissertation/6c7c20155f/>* (2016).
- [Pa06] K. Parshall, “James Joseph Sylvester: Jewish Mathematician in a Victorian World”, Johns Hopkins Press, Baltimore, 2006.
- [Pu98] M. Pucci ”The Veronese variety and catalecticant matrices.” *Journal of Algebra* 202.1 (1998): 72-95.
- [Re13a] B. Reznick, “Some new canonical forms for polynomials”, *Pacific Journal of Mathematics* **266**(1):185–220 (2013).
- [Re13b] B. Reznick, “On the length of binary forms”, *Quadratic and Higher Degree Forms, (K. Alladi, M. Bhargava, D. Savitt, P. Tiep, eds.)*, *Developments in Mathematics* **31**:207–232, Springer New York (2013) .
- [Sh94] I. Shafarevich. “Basic algebraic geometry, vol. 1”, Springer-Verlag, 1994.
- [St78] R. Stanley, “Hilbert functions of graded algebras.” *Advances in Mathematics* 28.1 (1978): 57-83.
- [Sy51a] J. J. Sylvester, An Essay on Canonical Forms, Supplement to a Sketch of a Memoir on Elimination, Transformation and Canonical Forms, originally published by George Bell, Fleet Street, London, 1851; pp. 203-216 in Paper 34 in Mathematical Papers, Vol. I, Chelsea, New York, 1973. Originally published by Cambridge University Press in 1904.
- [Sy51b] J. J. Sylvester, On a remarkable discovery in the theory of canonical forms and of hyperdeterminants, originally in Philosophical Magazine, vol. I, 1851; pp. 265-283 in Paper 41 in Mathematical Papers, Vol. 1, Chelsea, New York, 1973. Originally published by Cambridge University Press in 1904.
- [Te11] A. Terracini, “Sulle  $v_k$  per cui la varietà degli  $s_h$  ( $h + 1$ )-seganti ha dimensione minore dell’ordinario”, *Rendiconti del Circolo Matematico di Palermo* (1884-1940), **31**(1):392–396, (1911).

(S. Lundqvist, B. Shapiro) DEPARTMENT OF MATHEMATICS, STOCKHOLM UNIVERSITY, SE-106 91 STOCKHOLM, SWEDEN  
*E-mail address*: [samuel@math.su.se](mailto:samuel@math.su.se), [shapiro@math.su.se](mailto:shapiro@math.su.se)

(A. Oneto) INRIA SOPHIA ANTIPOLIS MÉDITERRANÉE (TEAM AROMATH), 06902 SOPHIA ANTIPOLIS, FRANCE  
*E-mail address*: [alessandro.oneto@inria.fr](mailto:alessandro.oneto@inria.fr)

(B. Reznick) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS, URBANA, IL 61801, USA  
*E-mail address*: [reznick@illinois.edu](mailto:reznick@illinois.edu)