A nearly optimal algorithm to decompose binary forms

Matías R Bender, Jean-Charles Faugère, Ludovic Perret, Elias Tsigaridas

A nearly optimal algorithm to decompose binary forms

Matías R. Bender

Sorbonne Université, CNRS, INRIA, *Laboratoire d'Informatique de Paris 6,* LIP6, *Équipe* POLSYS, *4 place Jussieu, F-75005, Paris, France*

Jean-Charles Faugère

Sorbonne Université, CNRS, INRIA, *Laboratoire d'Informatique de Paris 6,* LIP6, *Équipe* POLSYS, *4 place Jussieu, F-75005, Paris, France*

Ludovic Perret

Sorbonne Université, CNRS, INRIA, *Laboratoire d'Informatique de Paris 6,* LIP6, *Équipe* POLSYS, *4 place Jussieu, F-75005, Paris, France*

Elias Tsigaridas

Sorbonne Université, CNRS, INRIA, *Laboratoire d'Informatique de Paris 6,* LIP6, *Équipe* POLSYS, *4 place Jussieu, F-75005, Paris, France*

Abstract

Symmetric tensor decomposition is an important problem with applications in several areas, for example signal processing, statistics, data analysis and computational neuroscience. It is equivalent to Waring's problem for homogeneous polynomials, that is to write a homogeneous polynomial in n variables of degree D as a sum of D-th powers of linear forms, using the minimal number of summands. This minimal number is called the *rank* of the polynomial/tensor. We focus on decomposing binary forms, a problem that corresponds to the decomposition of symmetric tensors of dimension 2 and order D, that is, symmetric tensors of order D over the vector space $\mathbb{K}^2$. Under this formulation, the problem finds its roots in invariant theory where the decompositions are related to canonical forms.

We introduce a *superfast* algorithm that exploits results from *structured linear algebra*. It achieves a *softly linear* arithmetic complexity bound. To the best of our knowledge, the previously known algorithms have at least quadratic complexity bounds. Our algorithm computes a symbolic decomposition in $O(\mathtt{M}(D)\log(D))$ arithmetic operations, where $\mathtt{M}(D)$ is the complexity of multiplying two polynomials of degree D. It is deterministic when the decomposition is unique. When the decomposition is not unique, it is randomized. We also present a Monte Carlo variant as well as a modification to make it a Las Vegas one.

From the symbolic decomposition, we approximate the terms of the decomposition with an error of $2^{-\varepsilon}$, in $O\big(D\log^2(D)\big(\log^2(D)+\log(\varepsilon)\big)\big)$ arithmetic operations. We use results from Kaltofen and Yagati (1989) to bound the size of the representation of the coefficients involved in the decomposition and we bound the algebraic degree of the problem by $\min(rank, D-rank+1)$. We show that this bound can be tight. When the input polynomial has integer coefficients, our algorithm performs, up to poly-logarithmic factors, $\widetilde{O}_B(D\ell+D^4+D^3\tau)$ bit operations, where τ is the maximum bitsize of the coefficients and $2^{-\ell}$ is the relative error of the terms in the decomposition.
Keywords: Decomposition of binary forms; Tensor decomposition; Symmetric tensor; Symmetric tensor rank; Polynomial Waring's problem; Waring rank; Hankel matrix; Algebraic degree; Canonical form;

1. Introduction

The problem of decomposing a symmetric tensor consists in writing it as the sum of rank-1 symmetric tensors, using the minimal number of summands. This minimal number is known as the rank of the symmetric tensor[1]. The symmetric tensors of rank-1 correspond to, roughly speaking, the D-th outer-product of a vector. The decomposition of symmetric tensor is a common problem which appears in divers areas such as signal processing, statistics, data mining, computational neuroscience, computer vision, psychometrics, chemometrics, among others. For a modern introduction to the theory of tensor, their decompositions and applications we refer to e.g., Comon (2014); Landsberg (2012).

There is an equivalence between decomposing symmetric tensors and solving Waring's problem for homogeneous polynomials, e.g., Comon et al. (2008); Helmke (1992). Given a symmetric tensor of dimension n and order D, that is a symmetric tensor of order D over the vector space

Email addresses: matias.bender@inria.fr (Matías R. Bender), jean-charles.faugere@inria.fr (Jean-Charles Faugère), ludovic.perret@lip6.fr (Ludovic Perret), elias.tsigaridas@inria.fr (Elias Tsigaridas)

URL: http://www-polsys.lip6.fr/~bender/ (Matías R. Bender), http://www-polsys.lip6.fr/~jcf/ (Jean-Charles Faugère), http://www-polsys.lip6.fr/~perret/ (Ludovic Perret), http://www-polsys.lip6.fr/~elias/ (Elias Tsigaridas)

[1] Some authors, e.g., Comon et al. (2008), refer to this number as the symmetric rank of the tensor.

$\mathbb{K}^n$, we can construct a homogeneous polynomial in n variables of degree D. We can identify the symmetric tensors of rank-1 with the D-th power of linear forms. Hence, to decompose a symmetric tensor of order D is equivalent to write the corresponding polynomial as a sum of D-th powers of linear forms using the minimal numbers of summands. This minimal number is the rank of the polynomial/tensor.

Under this formulation, symmetric tensor decomposition dates back to the origin of modern (linear) algebra as a part of Invariant Theory. In this setting, the decomposition of generic symmetric tensors corresponds to canonical forms (Sylvester, 1904, 1851; Gundelfinger, 1887). Together with the theory of apolarity, this problem was of great importance because the decompositions provide information about the behavior of the polynomials under linear change of variables (Kung and Rota, 1984).

Binary Form Decomposition. We study the decomposition of symmetric tensors of order D and dimension 2. In terms of homogeneous polynomials, we consider a binary form

$$f(x,y) := \textstyle\sum_{i=0}^{D} \binom{D}{i} a_i x^i y^{D-i}, \tag{1}$$

where $a_i \in \mathbb{K} \subset \mathbb{C}$ and $\mathbb{K}$ is some field of characteristic zero. We want to compute a decomposition

$$f(x,y) = \textstyle\sum_{j=1}^{r} (\alpha_j x + \beta_j y)^D, \tag{2}$$

where $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_r \in \overline{\mathbb{K}}$, with $\overline{\mathbb{K}}$ being the algebraic closure of $\mathbb{K}$, and r is minimal. We say that a decomposition *unique* if, for all the decompositions, the set of points $\{(\alpha_j, \beta_j) : 1 \leq j \leq r\} \subset \mathbb{P}^1(\overline{\mathbb{K}})$ is unique, where $\mathbb{P}^1(\overline{\mathbb{K}})$ is the projective space of $\overline{\mathbb{K}}$ (Reznick, 2013a).

Previous work. The decomposition of binary forms, Equation (2), has been studied extensively for $\mathbb{K} = \mathbb{C}$. More than one century ago Sylvester (1851, 1904) described the necessary and sufficient conditions for a decomposition to exist, see Section 2.1. He related the decompositions to the kernel of Hankel matrices. For a modern approach of this topic, we refer to Kung and Rota (1984); Kung (1990); Reznick (2013a); Iarrobino and Kanev (1999). Sylvester's work was extended to a more general kind of polynomial decompositions that we do not consider in this work, e.g., Gundelfinger (1887); Reznick (1996); Iarrobino and Kanev (1999).

Sylvester's results lead to an algorithm (Algorithm 1) to decompose binary forms (see Comon and Mourrain, 1996, Sec. 3.4.3). In the case where the binary form is of odd degree, then we can compute the decompositions using Berlekamp-Massey algorithm (see Dür, 1989). When the decomposition is unique, the Catalecticant algorithm, which also works for symmetric tensors of bigger dimension (Iarrobino and Kanev, 1999; Oeding and Ottaviani, 2013), improves Sylvester's work. For an arbitrary binary form, Helmke (1992) presented a randomized algorithm based on Padé approximants and continued fractions, in which he also characterized the different possible decompositions. Unfortunately, all these algorithms have complexity at least quadratic in the degree of the binary form.

Besides the problem of computing the decomposition(s) many authors considered the subproblems of computing the rank and deciding whether there exists a unique decomposition, e.g., Sylvester (1851, 1904); Helmke (1992); Comas and Seiguer (2011); Bernardi et al. (2011). For example, Sylvester (1851, 1904) considered generic binary forms, that is binary forms with coefficients belonging to a dense algebraic open subset of $\overline{\mathbb{K}}^{D+1}$ (Comon and Mourrain, 1996, Section 3), and proved that when the degree is $2k$ or $2k+1$, for $k \in \mathbb{N}$, the rank is $k+1$ and that the minimal decomposition is unique only when the degree is odd. In the non-generic case, Helmke (1992); Comas and Seiguer (2011); Iarrobino and Kanev (1999), among others, proved that the rank is related to the kernel of a Hankel matrix and that the decomposition of a binary form of degree $2k$ or $2k-1$ and rank r, is unique if and only if $r \leq k$. With respect to the problem of computing the rank there are different variants of algorithms, e.g., Comas and Seiguer (2011); Comon et al. (2008); Bernardi et al. (2011). Even though there are not explicit complexity estimates, by exploiting recent superfast algorithms for Hankel matrices (Pan, 2001), we can deduce a nearly-optimal arithmetic complexity bound for computing the rank using the approach of Comas and Seiguer (2011).

For the general problem of symmetric tensor decomposition, Sylvester's work was successfully extended to cases in which the decomposition is unique, e.g., Brachat et al. (2010); Oeding and Ottaviani (2013). There are also homotopy techniques to solve the general problem, e.g., to decompose generic symmetric tensors (Bernardi et al., 2017) or, when there is a finite number of possible decompositions and we know at least one of them, to compute all the other decompositions (Hauenstein et al., 2016). There are no complexity estimations for these methods. Besides tensor decomposition, there are other related decompositions for binary forms and univariate polynomials that we do not consider in this work, e.g., Reznick (1996, 2013b); Giesbrecht et al. (2003); Giesbrecht and Roche (2010); García-Marco et al. (2017).

Formulation of the problem. Instead of decomposing the binary form as in Equation (2), we compute $\lambda_1 \dots \lambda_r$, $\alpha_1 \dots \alpha_r$, $\beta_1 \dots \beta_r \in \overline{\mathbb{K}}$, where r is minimal, such that,

$$f(x,y) = \sum_{j=1}^{r} \lambda_j (\alpha_j x + \beta_j y)^D. \tag{3}$$

Since every λ_j belongs to the algebraic closure of the field $\mathbb{K}$, the problems are equivalent. This approach allows us to control the algebraic degree (Bajaj, 1988; Nie et al., 2010) of the parameters λ_j, α_j, and β_j in the decomposition (Section 4.1).

Note that if the field is not algebraically closed and we force the parameters to belong to the base field, that is $\lambda_j, \alpha_j, \beta_j \in \mathbb{K}$, the decompositions induced by Equation (2) and Equation (3) are not equivalent. We do not consider the latter case and we refer to Helmke (1992); Reznick (1992); Comon et al. (2008); Boij et al. (2011); Blekherman (2015) for $\mathbb{K} = \mathbb{R}$, and to Reznick (1996, 2013a); Reznick and Tokcan (2017) for $\mathbb{K} \subset \mathbb{C}$.

Main results. We extend Sylvester's algorithm to achieve a nearly-optimal complexity bound in the degree of the binary form. By considering structural properties of the Hankel matrices, we restrict the possible values for the rank of the decompositions and we identify when the decomposition is unique. We build upon Helmke (1992) and we use the Extended Euclidean Algorithm to deduce a better complexity estimate than what was previously known. Similarly to Sylvester's algorithm, our algorithm decomposes successfully any binary form, without making any assumptions on the input.

First, we focus on *symbolic decompositions*, that is a representation of the decomposition as a sum of a rational function evaluated at the roots of a univariate polynomial (Definition 36). We introduce an algorithm to compute a symbolic decomposition of a binary form of degree D in $O(\mathtt{M}(D)\log(D))$, where $\mathtt{M}(D)$ is the arithmetic complexity of polynomial multiplication (Theorem 43). When the decomposition is unique, the algorithm is deterministic and this is a worst case bound. When the decomposition is not unique, our algorithm makes some random choices to fulfill certain genericity assumptions; thus the algorithm is a Monte Carlo one. However, we can verify if the genericity assumptions hold within the same complexity bound, that is $O(\mathtt{M}(D)\log(D))$, and hence we can also deduce a Las Vegas variant of the algorithm.

Following the standard terminology used in structured matrices (Pan, 2001), our algorithm is *superfast* as its arithmetic complexity matches the size of the input up to poly-logarithmic factors. The symbolic decomposition allow us to approximate the terms in a decomposition, with a relative error of $2^{-\varepsilon}$, in $O\big(D\log^2(D)\big(\log^2(D)+\log(\varepsilon)\big)\big)$ arithmetic operations (Pan, 2002; McNamee and Pan, 2013). Moreover, we can deduce for free the rank and the border rank of the tensor, see (Comas and Seiguer, 2011, Section 1).

Using results from Kaltofen and Yagati (1989), we bound the algebraic degree of the decompositions by $\min(\text{rank}, D-\text{rank}+1)$ (Theorem 28). Moreover, we prove lower bounds for the algebraic degree of the decomposition and we show that in certain cases the bound is tight (Section 4.1.3). For polynomials with integer coefficients, we bound the bit complexity, up to poly-logarithmic factors, by $\widetilde{O}_B(D\ell + D^4 + D^3\tau)$, where τ is the maximum bitsize of the coefficients of the input binary form and $2^{-\ell}$ is the error of the terms in the decomposition (Theorem 45). This Boolean worst case bound holds independently of whether the decomposition is unique or not.

This work is an extension of the conference paper (Bender et al., 2016). With respect to the conference version, our main algorithm (Algorithm 3) omits an initial linear change of coordinates as we now rely on fewer genericity assumptions. In contrast with our previous algorithm, we present an algorithm which is deterministic when the decomposition is unique (Theorem 43). When the decomposition is not unique, our algorithm is still randomized but we present bounds for the number of bad choices that it could make (Proposition 29). With respect to the algebraic degree of the problem, we study the tightness of the bounds that we proposed in the conference paper (Theorem 27). We introduce explicit lower bounds showing that our bounds can be tight (Section 4.1.3).

Organization of the paper. First, we introduce the notation. In Section 2, we present the preliminaries that we need for introducing our algorithm. We present Sylvester's algorithm (Section 2.1), we recall some properties of the structure of the kernel of the Hankel matrices (Section 2.2), we analyze its relation to rational reconstructions of series/polynomials (Section 2.3), and we present the Extended Euclidean Algorithm (Section 2.4). Later, in Section 3, we present our main algorithm to decompose binary forms (Algorithm 3) and its proof of correctness (Section 3.3). This algorithm uses Algorithm 4 to compute the kernel of a family of Hankel matrices, which we consider in Section 3.1. Finally, in Section 4, we study the algebraic degree of the problem (Section 4.1), we present tight bounds for it (Section 4.1.3), and we analyze the arithmetic (Section 4.2) and bit complexity of Algorithm 3 (Section 4.3).

Notation. We denote by O, respectively O_B, the arithmetic, respectively bit, complexity and we use $\widetilde{O}$, respectively $\widetilde{O}_B$, to ignore (poly-)logarithmic factors. $\mathtt{M}(n)$ is the arithmetic complexity of multiplying two polynomial of degree n. Let $\mathbb{K}$ be a zero characteristic subfield of $\mathbb{C}$, and $\overline{\mathbb{K}}$ its algebraic closure. If $v = (v_0, \ldots, v_n)^\mathsf{T}$ then $P_v = P_{(v_0,\ldots,v_n)} := \sum_{i=0}^n v_i x^i y^{n-i}$. Given a binary form $f(x,y)$, we denote by $f(x)$ the univariate polynomial $f(x) := f(x,1)$. By $f'(x)$ we denote the derivative of $f(x)$ with respect to x. For a matrix M, $\mathtt{rk}(M)$ is its rank and $\mathtt{Ker}(M)$ its kernel.

2. Preliminaries

2.1. An algorithm based on Sylvester's theorem

Sylvester's theorem (Theorem 2) relates the minimal decomposition of a binary form to the kernel of a Hankel matrix. Moreover, it implies an (incremental) algorithm for computing the minimal decomposition. The version that we present in Algorithm 1 comes from Comon and Mourrain (1996, Section 3.2).

Definition 1. *Given a vector $a = (a_0, \ldots, a_D)^\mathsf{T}$, we denote by $\{H_a^k\}_{1 \le k \le D}$ the family of Hankel matrices indexed by k, where $H_a^k \in \mathbb{K}^{(D-k+1)\times(k+1)}$ and*

$$H_a^k := \begin{pmatrix} a_0 & a_1 & \cdots & a_{k-1} & a_k \\ a_1 & a_2 & \cdots & a_k & a_{k+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{D-k-1} & a_{D-k} & \cdots & a_{D-2} & a_{D-1} \\ a_{D-k} & a_{D-k+1} & \cdots & a_{D-1} & a_D \end{pmatrix}. \tag{4}$$

We may omit the index a in H_a^k when it is clear from the context.

Theorem 2 (Sylvester, 1851)**.** *Let $f(x,y) = \sum_{i=0}^D \binom{D}{i} a_i x^i y^{D-i}$ with $a_i \in \mathbb{K} \subseteq \mathbb{C}$. Also, consider a non-zero vector $c = (c_0, \ldots, c_r)^\mathsf{T} \in \mathbb{K}^{r+1}$, such that the polynomial*

$$P_c = \textstyle\sum_{i=0}^r c_i x^i y^{r-i} = \prod_{j=1}^r (\beta_j x - \alpha_j y)$$

is square-free and $\alpha_j, \beta_j \in \overline{\mathbb{K}}$, for all $1 \le j \le r$. Then, there are $\lambda_1, \ldots \lambda_r \in \overline{\mathbb{K}}$ such that we can write $f(x,y)$ as

$$f(x,y) = \sum_{j=1}^r \lambda_j (\alpha_j x + \beta_j y)^D,$$

if and only if $(c_0, \ldots, c_r)^\mathsf{T} \in \mathtt{Ker}(H_a^r)$.

Algorithm 1 INCRDECOMP (Comon and Mourrain, 1996, Figure 1)

1. $r := 1$
2. Get a random $c \in \mathtt{Ker}(H^r)$
3. If P_c is not square-free, $r := r+1$ and GO TO 2
4. Write P_c as $\prod_{j=1}^{r}(\beta_j x - \alpha_j y)$
5. Solve the transposed Vandermonde system:

$$\begin{pmatrix} \beta_1^D & \cdots & \beta_r^D \\ \beta_1^{D-1}\alpha_1 & \cdots & \beta_r^{D-1}\alpha_r \\ \vdots & \ddots & \vdots \\ \alpha_1^D & \cdots & \alpha_r^D \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} = \begin{pmatrix} a_0 \\ \vdots \\ a_D \end{pmatrix}. \tag{5}$$

6. Return $\sum_{j=1}^{r} \lambda_j (\alpha_j x + \beta_j y)^D$

For a proof of Theorem 2 we refer to Reznick (2013a, Theorem 2.1 & Corollary 2.2). Theorem 2 implies Algorithm 1. This algorithm will execute steps 2 and 3 as many times as the rank. At the i-th iteration it computes the kernel of H^i. The dimension of this kernel is $\leq i$ and each vector in the kernel has $i+1$ coordinates. As the rank of the binary form can be as big as the degree of the binary form, a straightforward bound for the arithmetic complexity of Algorithm 1 is at least cubic in the degree.

We can improve the complexity of Algorithm 1 by a factor of D by noticing that the rank of the binary form is either $\mathtt{rk}(H^{\lceil \frac{D}{2} \rceil})$ or $D - \mathtt{rk}(H^{\lceil \frac{D}{2} \rceil}) + 2$ (Comas and Seiguer, 2011, Section 3) (Helmke, 1992, Theorem B). Another way to compute the rank is by using minors (Bernardi et al., 2011, Algorithm 2).

The bottleneck of the previous approaches is that they have to compute the kernel of a Hankel matrix. However, even if we know that the rank of the binary form is r, the dimension of the kernel of H^r can still be as big as $O(D)$; the same bound holds for the length of the vectors in the kernel. Hence, the complexity is lower bounded by $O(D^2)$.

Our approach avoids the incremental construction. We exploit the structure of the kernel of the Hankel matrices and we prove that the rank has only two possible values (Lemma 17), see also (Comas and Seiguer, 2011, Section 3), (Helmke, 1992, Theorem B), or (Bernardi et al., 2011). Moreover, we use a compact representation of the vectors in the kernel. We describe them as a combination of two polynomials of degree $O(D)$.

2.2. *Kernel of the Hankel matrices*

The Hankel matrices are among the most studied structured matrices (Pan, 2001). They are related to polynomial multiplication. We present results about the structure of their kernel. For details, we refer to Heinig and Rost (1984, Chapter 5).

Proposition 3. *Matrix-vector multiplication of Hankel matrices is equivalent to polynomial multiplication. Given two binary forms $A := \sum_{i=0}^{D} a_i x^i y^{D-i}$ and $U := \sum_{i=0}^{k} u_i x^i y^{k-i}$, consider $R := \sum_{i=0}^{D+k} r_i x^i y^{D+k-i} = A \cdot U$. If we choose the monomial basis $\{y^{D+k}, \ldots, x^{D+k}\}$, then the equality $A \cdot U = R$ is equivalent to Equation* (6)*, where the central submatrix of the left matrix is $H^k_{(a_0,\ldots,a_D)}$ (Definition 1).*

$$
\left(\begin{array}{ccccc}
 & & & & a_0 \\
 & & & a_0 & a_1 \\
 & & \iddots & \iddots & \vdots \\
 & a_0 & \cdots & a_{k-2} & a_{k-1} \\
\hline
a_0 & a_1 & \cdots & a_{k-1} & a_k \\
a_1 & a_2 & \cdots & a_k & a_{k+1} \\
\vdots & \vdots & \iddots & \vdots & \vdots \\
a_{D-k} & a_{D-k+1} & \cdots & a_{D-1} & a_D \\
\hline
a_{D-k+1} & a_{D-k+2} & \cdots & a_D & \\
\vdots & \iddots & \iddots & & \\
a_{D-1} & a_D & & & \\
a_D & & & &
\end{array}\right)
\begin{pmatrix} u_k \\ \vdots \\ u_1 \\ u_0 \end{pmatrix}
=
\left(\begin{array}{c}
r_0 \\ r_1 \\ \vdots \\ r_{k-1} \\ \hline r_k \\ r_{k+1} \\ \vdots \\ r_D \\ \hline r_{D+1} \\ \vdots \\ r_{D+k-1} \\ r_{D+k}
\end{array}\right). \tag{6}
$$

Consider a family of Hankel matrices $\{H^k_a\}_{1 \leq k \leq D}$ as in Definition 1. There is a formula for the dimension of the kernel of each matrix in the family $\{H^k_a\}_{1 \leq k \leq D}$ that involves two numbers, N^a_1 and N^a_2. To be more specific, the following holds:

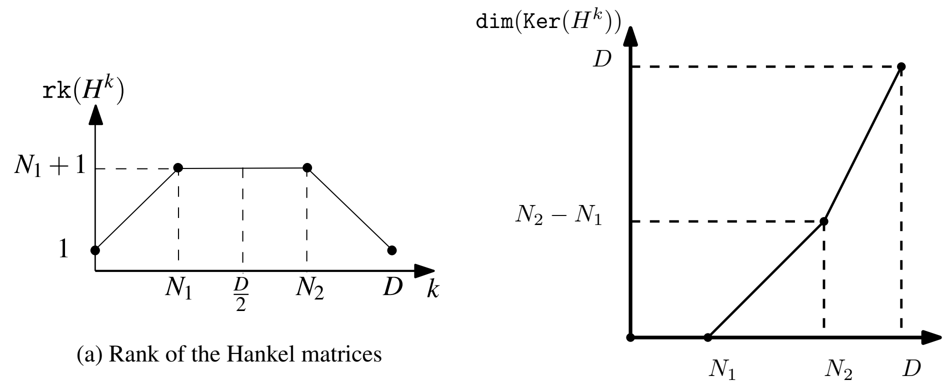
Proposition 4. *For any family of Hankel matrices $\{H_a^k\}_{1\le k\le D}$ there are two constants, N_1^a and N_2^a, such that the following hold:*

1. $0 \le N_1^a \le N_2^a \le D$.

2. *For all k,* $1 \le k \le D$, *it holds* $\mathtt{dim}(\mathtt{Ker}(H_a^k)) = \max(0; k - N_1^a) + \max(0; k - N_2^a)$.

3. $N_1^a + N_2^a = D$.

We may omit the index a in N_1^a and N_2^a when it is clear from the context.



(a) Rank of the Hankel matrices

(b) Dimension of the kernel

Figure 1: Relation between N_1, N_2, and D

An illustration of Proposition 4 appears in Figure 1. The dimension of the kernel and the rank of the matrices are piece-wise-linear functions in k, the number of columns in the matrix. The graphs of the functions consist of three line segments, as we can see in the Figures 1a and 1b. The dimension of the kernel is an increasing function of k. For k from 1 to N_1, the kernel of the matrix is trivial, so the rank increases with the number of columns. That is, the slope of the graph of the rank (Figure 1a) is 1, while the slope of the graph of the dimension of the kernel (Figure 1b) is 0. For k from N_1+1 to N_2, the rank remains constant as for each column that we add, the dimension of the kernel increases by one. Hence, the slope of the graph of the rank is 0 and the slope of the graph of the dimension of the kernel is 1. For k from N_2+1 to D, the rank decreases because the dimension of the kernel increases by 2, and so the slope of the graph of the rank is -1, while the slope of the graph of the dimension of the kernel is 2.

If $N_1 = N_2$, then the graph of both functions degenerates to two line segments. Regarding the graph of the rank, the first segment has slope 1 for k from 1 to $N_1 + 1$ and the second segment has slope -1 for k from $N_1 + 1$ to D. For the graph of the dimension of the kernel, the first segment has slope 0 from 1 to $N_1 + 1$, and the second one has slope 2 from N_1 to D.

The elements of the kernel of the matrices in $\{H^k\}$ are related. To express this relation from a linear algebra point of view, we need to introduce the **U-chains**.

Definition 5 (Heinig and Rost (1984, Definition 5.1))**.** *A **U-chain** of length k of a vector* $v = (v_0, \dots, v_n)^\mathsf{T} \in \mathbb{K}^{n+1}$ *is a set of vectors* $\{\mathbb{U}_k^0 v, \mathbb{U}_k^1 v, \dots, \mathbb{U}_k^{k-1} v\} \subset \mathbb{K}^{n+k}$*. The i-th element, for* $0 \leq i \leq k-1$*, is*

$$\mathbb{U}_k^i v = (\underbrace{0, \dots, 0}_{i}, \overbrace{v_0, \dots, v_n}^{n+1}, \underbrace{0, \dots, 0}_{k-1-i}),$$

where $\mathbb{U}_k^i$ *is an i-shifting matrix of dimension* $(n+k) \times (n+1)$ *(Heinig and Rost, 1984, page 11).*

If v is not zero, then all the elements in a U-chain of v are linearly independent. The following theorem uses U-chains to relate the vectors in the kernels of a family of Hankel matrices.

Proposition 6 (Vectors v and w)**.** *Given a family of Hankel matrices* $\{H^k\}_{1 \leq k \leq D}$*, let* N_1 *and* N_2 *be the constants of Proposition 4. There are two vectors,* $v \in \mathbb{K}^{N_1+2}$ *and* $w \in \mathbb{K}^{N_2+2}$*, such that*

- *If* $0 \leq k \leq N_1$*, then* $\mathtt{Ker}(H^k) = \{0\}$.
- *If* $N_1 < k \leq N_2$*, then the U-chain of v of length* $(k - N_1)$ *is a basis of* $\mathtt{Ker}(H^k)$*, that is*

$$\mathtt{Ker}(H^k) = \langle \mathbb{U}_{k-N_1}^0 v, \dots, \mathbb{U}_{k-N_1}^{k-N_1-1} v \rangle.$$

- *If* $N_2 < k \leq D$*, then the U-chain of v of length* $k - N_1$ *together with the U-chain of w of length* $k - N_2$ *is a basis of* $\mathtt{Ker}(H^k)$*, that is*

$$\mathtt{Ker}(H^k) = \langle \mathbb{U}_{k-N_1}^0 v, \dots, \mathbb{U}_{k-N_1}^{k-N_1-1} v, \mathbb{U}_{k-N_2}^0 w, \dots, \mathbb{U}_{k-N_2}^{k-N_2-1} w \rangle.$$

The vectors v and w of Proposition 6 are not unique. The vector v could be any vector in $\mathtt{Ker}(H^{N_1+1})$. The vector w could be any vector in $\mathtt{Ker}(H^{N_2+1})$ that does not belong to the vector space generated by the U-chain of v of length $N_2 - N_1 + 1$. From now on, given a family of Hankel matrices, we refer to v and w as the vectors of Proposition 6.

Let u be a vector in the kernel of H^k and P_u its corresponding polynomial (see Notation). We say that P_u is a **kernel polynomial**. As $P_{\mathbb{U}_k^j v} = x^j y^{k-1-j} P_v$, we can write any kernel polynomial of a family of Hankel matrices as a combination of P_v and P_w (Heinig and Rost, 1984, Propositions 5.1 & 5.5). Moreover, P_v and P_w are relatively prime.

Proposition 7. *Consider any family of Hankel matrices $\{H^k\}_{1 \le k \le D}$. The kernel polynomials P_v and P_w are relative prime. Moreover, for each k, the set of kernel polynomials of the matrix H^k is as follows:*

- *If $0 < k \le N_1$, then it is $\{0\}$.*
- *If $N_1 < k \le N_2$, then it is $\{P_\mu \cdot P_v : \mu \in \mathbb{K}^{k-N_1}\}$.*
- *If $N_2 < k \le D$, then it is $\{P_\mu \cdot P_v + P_\rho \cdot P_w : \mu \in \mathbb{K}^{k-N_1}, \rho \in \mathbb{K}^{k-N_2}\}$.*

Corollary 8. *Let $\omega \in \mathtt{Ker}(H^{N_2+1})$ such that $P_\omega \notin \{P_\mu \cdot P_v : \mu \in \mathbb{K}^{N_2-N_1+1}\}$, then we can consider ω as the vector w from Proposition 6.*

2.3. Rational Reconstructions

A rational reconstruction for a series, respectively a polynomial, consists in approximating the series, respectively the polynomial, as the quotient of two polynomials. Rational reconstructions are the backbone of many problems e.g., Padé approximants, Cauchy Approximations, Linear Recurrent Sequences, Hermite Interpolation. They are related to Hankel matrices. For an introduction to rational reconstructions, we refer to Bostan et al. (2017, Chapter 7) and references therein.

Definition 9. *Consider $a := (a_0, \ldots, a_D)^\mathsf{T} \in \mathbb{K}^{D+1}$ and a polynomial $A := \sum_{i=0}^{D} a_i x^i \in \mathbb{K}[x]$. Given a pair of univariate polynomials (U, R), we say that they are a rational reconstruction of A modulo x^{D+1} if $A \cdot U \equiv R \mod x^{D+1}$.*

Such a reconstruction is not necessarily unique. Our interest emanates from the relation between the rational reconstructions of A modulo x^{D+1} and the kernels of the family of Hankel matrices $\{H_a^k\}_k$.

Lemma 10. *Consider* $\omega \in \mathtt{Ker}(H_a^k)$ *and* $(r_0, \ldots, r_{k-1}) \in \mathbb{K}^k$ *such that*

$$\begin{pmatrix} 0 & \cdots & 0 & a_0 \\ \vdots & \iddots & \iddots & \vdots \\ 0 & a_0 & \cdots & a_{k-1} \end{pmatrix} \begin{pmatrix} \omega_0 \\ \omega_1 \\ \vdots \\ \omega_k \end{pmatrix} = \begin{pmatrix} r_0 \\ \vdots \\ r_{k-1} \end{pmatrix}.$$

Then, $(P_\omega(1,x), \sum_{i=0}^{k-1} r_i x^i)$ *is a rational reconstruction of A modulo* x^{D+1}.

Proof. Following Equation (6), if $\omega \in \mathtt{Ker}(H_a^k)$, then

$$\left(\begin{array}{cccc} & & & a_0 \\ & & \iddots & \vdots \\ & a_0 & \cdots & a_{k-1} \\ \hline a_0 & a_1 & \cdots & a_k \\ \vdots & \vdots & \iddots & \vdots \\ a_{D-k} & a_{D-k+1} & \cdots & a_D \end{array}\right) \begin{pmatrix} \omega_0 \\ \omega_1 \\ \vdots \\ \omega_k \end{pmatrix} = \left(\begin{array}{c} r_0 \\ \vdots \\ r_{k-1} \\ \hline 0 \\ \vdots \\ 0 \end{array}\right). \tag{7}$$

Hence, $P_\omega(1,x) = \sum_{i=0}^k \omega_{k-i} x^i$ and $A \cdot P_\omega(1,x) \equiv \sum_{i=0}^{k-1} r_i x^i \mod x^{D+1}$. Therefore, $(P_\omega(1,x), \sum_{i=0}^{k-1} r_i x^i)$ is a rational reconstruction of A modulo x^{D+1}. ☐

Lemma 11. *If* (U,R) *is a rational reconstruction of A module* x^{D+1}*, then there is a vector* $\omega \in \mathtt{Ker}(H_a^{\max(\deg(U),\deg(R)+1)})$ *such that*

$$P_\omega = U\left(\frac{y}{x}\right) x^{\max(\deg(U),\deg(R)+1)}.$$

Proof. Let $k = \deg(U)$, $q = \deg(R)$, $U = \sum_i u_i x^i$ and $R = \sum_i r_i x^i$. Following Equation (6), $A\,U \equiv R \mod x^{D+1}$ is equivalent to,

$$\begin{pmatrix} & & & a_0 \\ & & \iddots & \vdots \\ & a_0 & \cdots & a_{k-1} \\ a_0 & a_1 & \cdots & a_k \\ \vdots & \vdots & \iddots & \vdots \\ a_{D-k} & a_{D-k+1} & \cdots & a_D \end{pmatrix} \begin{pmatrix} u_k \\ \vdots \\ u_1 \\ u_0 \end{pmatrix} = \begin{pmatrix} r_0 \\ r_1 \\ \vdots \\ r_q \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{8}$$

If $k > q$, Equation (8) reduces to Equation (7), and so $\omega = (u_k, \ldots, u_0) \in \mathtt{Ker}(H_a^k)$. Moreover,

$$U\left(\frac{y}{x}\right)x^k = \sum_{i=0}^{k} u_i y^i x^{k-i} \underset{(j \leftrightarrow k-i)}{=} \sum_{j=0}^{k} u_{k-j} x^j y^{j-k} = P_\omega.$$

If $q \geq k$, we extend the vector $(u_k, \ldots, u_0)$ by adding $(q+1-k)$ leading zeros. We rewrite Equation (8) as Equation (9). The concatenation of the two bottom submatrices form the matrix H_a^{q+1}, and so $\omega = (0, \ldots, 0, u_k, \ldots, u_0) \in \mathtt{Ker}(H_a^{q+1})$. Also,

$$P_\omega = \sum_{j=0}^{k} u_j x^{q+1-j} y^j + \sum_{j=k+1}^{q+1} 0 x^{q+1-j} y^j = U\left(\frac{y}{x}\right) x^{q+1}.$$

$$\left[\begin{array}{ccc|ccc} & & & & & a_0 \\ & & & & \iddots & \vdots \\ & & & a_0 & \cdots & a_k \\ & & a_0 & a_1 & \cdots & a_{k+1} \\ & \iddots & \iddots & \vdots & \iddots & \vdots \\ \hline a_0 & \cdots & a_{q-k} & a_{q+1-k} & \iddots & a_{q+1} \\ \vdots & \iddots & \vdots & \vdots & \iddots & \vdots \\ a_{D-q-1} & \cdots & a_{D-k-1} & a_{D-k} & \cdots & a_D \end{array}\right] \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \hline u_k \\ \vdots \\ u_0 \end{pmatrix} = \begin{pmatrix} r_0 \\ \vdots \\ r_k \\ \vdots \\ r_q \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{9}$$

□

Remark 12. *If (U,R) is a rational reconstruction, then the degree of the kernel polynomial $P_\omega(x,y) = U\left(\frac{y}{x}\right) x^{\max(\deg(U), \deg(R)+1)}$ is $\max(\deg(U), \deg(R)+1)$. In particular, the maximum power of x that divides the kernel polynomial P_ω is $x^{\max(0, \deg(R)+1-\deg(U))}$.*

2.4. Greatest Common Divisor and Bézout identity

The Extended Euclidean algorithm (EGCD) is a variant of the classical Euclidean algorithm that computes the Greatest Common Divisor of two univariate polynomials A and B, $\gcd(A,B)$, together with two polynomials U and V, called *cofactors*, such that $UA + VB = \gcd(A,B)$. In the process of computing these cofactors, the algorithm computes a sequence of relations between A and B that are useful to solve various problems, in particular to compute the rational reconstruction of A modulo B. For a detailed exposition of this algorithm, we refer to Bostan et al. (2017, Chapter 6) and Gathen and Gerhard (2013, Chapter 3 and 11).

Algorithm 2 Calculate the EGCD of A and B

$(U_0, V_0, R_0) \leftarrow (0, 1, B)$

$(U_1, V_1, R_1) \leftarrow (1, 0, A)$

$k \leftarrow 1$

while $R_k \neq 0$ **do**

  $k \leftarrow k+1$

  $Q_{k-1} \leftarrow R_{k-2}$ quo R_{k-1}

  $(U_k, V_k, R_k) \leftarrow (U_{k-2}, V_{k-2}, R_{k-2}) - Q_{k-1}\,(U_{k-1}, V_{k-1}, R_{k-1})$

end while

Return $\{(U_i, V_i, R_i)\}_i$

The Extended Euclidean Algorithm (Algorithm 2) computes a sequence of triples $\{(U_i, V_i, R_i)\}_i$ which form the identities

$$U_i A + V_i B = R_i, \quad \text{for all } i. \tag{10}$$

Following Gathen and Gerhard (2013), we refer to these triplets as the rows of the Extended Euclidean algorithms of A and B. Besides Equation (10), the rows are related to each other as follows.

Remark 13. *The degrees of the polynomials $\{R_i\}_i$ form a strictly decreasing sequence, that is* $\deg(R_i) > \deg(R_{i+1})$ *for every i.*

Lemma 14 (Bostan et al., 2017, Sec 7.1)**.** *For each i, $U_i V_{i+1} - U_{i+1} V_i = (-1)^i$, and so the polynomials U_i and V_i are coprime.*

Lemma 15 (Bostan et al., 2017, Lem 7.1)**.** *For each $i > 0$, the degree of U_i is the degree of B minus the degree of R_{i-1}, that is*

$$\deg(U_i) = \deg(B) - \deg(R_{i-1}), \quad \forall\, i > 0.$$

Every row of the Extended Euclidean Algorithm leads to rational reconstruction of A modulo B.

Remark 16. *For each $i \geq 0$, (U_i, R_i) is a rational reconstruction of A modulo B.*

3. The Algorithm

One of the drawbacks of Algorithm 1 and its variants is that they rely on the computation of the kernels of many Hankel matrices and they ignore the particular structure that it is present in all of them. Using Lemma 17, we can skip many calculations by computing only two vectors, v and w (Proposition 6). This is the main idea behind Algorithm 3 that leads to a softly-linear arithmetic complexity bound (Section 4.2).

Algorithm 3 performs as follows: First, step 1 computes the kernel polynomials P_v and P_w which, by Proposition 7, allow us to obtain the kernel polynomials of all the Hankel matrices (see Section 3.1). Then, step 2 computes a square-free kernel polynomial of the minimum degree r (see Section 3.2). Next, step 3 computes the coefficients $\lambda_1, \ldots, \lambda_r$ (see Section 4.1.2). Finally, step 4 recovers a decomposition for the original binary form.

Let f be a binary form as in Equation (1) and let $\{H^k\}_{1 \leq k \leq D}$ be its corresponding family of Hankel matrices (see Definition 1). The next well-known lemma establishes the rank of f.

Lemma 17. *Assume f, $\{H^k\}_k$, N_1 and N_2 of Proposition 4, and v and w of Proposition 6. If P_v (Proposition 7) is square-free then the rank of f is $N_1 + 1$, else, it is $N_2 + 1$.*

Proof. By Proposition 4, for $k < N_1 + 1$, the kernel of H^k is trivial. Hence, by Sylvester's theorem (Theorem 2), there is no decomposition with a rank smaller than $N_1 + 1$. Recall that $v \in \mathtt{Ker}(H^{N_1+1})$. So, if P_v is square-free, by Sylvester's theorem, there is a decomposition of rank $N_1 + 1$.

Assume P_v is not square-free. For $N_1 + 1 \leq k \leq N_2$, P_v divides all the kernel polynomials of the matrices H^k (Proposition 7). Therefore, none of them is square-free, and so the rank is at least $N_2 + 1$.

By Proposition 7, P_v and P_w are coprime. So, there is a polynomial P_μ of degree $N_2 - N_1$ such that $Q_\mu := P_v \cdot P_\mu + P_w$ is square-free. A formal proof of this appears in Theorem 22. By Proposition 7, Q_μ is a square-free kernel polynomial of degree $N_2 + 1$. Consequently, by Sylvester's theorem, there is a decomposition with rank $N_2 + 1$. □

For alternative proofs of Lemma 17 we refer to (Helmke, 1992, Theorem B), (Comas and Seiguer, 2011, Section 3), (Bernardi et al., 2011), or (Carlini et al., 2018, Section 4).

To relate Lemma 17 with the theory of binary form decomposition, we recall that the decompositions are identified with the square-free polynomials in the annihilator of the ideal $\langle f \rangle$ (Kung and Rota, 1984);(Iarrobino and Kanev, 1999, Chapter 1). All the kernel polynomials of $\{H^k\}_k$ belong to the annihilator of $\langle f \rangle$ and they form an ideal. If f is a binary form of degree $D = 2k$ or $2k+1$, then this ideal is generated by two binary forms of degrees $\mathtt{rk}(H^k)$ and $D+2-\mathtt{rk}(H^k)$, with no common zeros (Iarrobino and Kanev, 1999, Theorem 1.44). These are the polynomials P_v and P_w. Using this interpretation Algorithm 1 and its variants compute a (redundant) generating set of the annihilator of $\langle f \rangle$, while Algorithm 3 computes a (minimal) basis.

Algorithm 3 FASTDECOMP

Input: A binary form $f(x,y)$ of degree D

Output: A decomposition for $f(x,y)$ of rank r.

1. **Compute $\mathbf{P_v}$ and $\mathbf{P_w}$ of $\{\mathbf{H}_a^{\mathbf{k}}\}_{\mathbf{k}}$**

   We use Algorithm 4 with $(a_0, \ldots, a_D)$.

2. **IF $\mathbf{P_v}(\mathbf{x}, \mathbf{y})$ is square-free,**

   $Q \longleftarrow P_v$

   $r \longleftarrow N_1 + 1$ {The *rank* of the decomposition is the degree of Q}

   ELSE

   Compute a square-free binary form Q

   We compute a vector μ of length $(N_2 - N_1 + 1)$,

   such that $(P_\mu \cdot P_v + P_w)$ is square-free (Section 4.1.1).

   $Q \longleftarrow P_\mu \cdot P_v + P_w$

   $r \longleftarrow N_2 + 1$ {The *rank* of the decomposition is the degree of Q}

3. **Compute the coefficients $\lambda_1, \ldots, \lambda_r$**

   Solve the system of Equation (5) where $Q(x,y) = \prod_{j=1}^{r}(\beta_j x - \alpha_j y)$.

   For details and the representation of λ_j, see Section 4.1.2.

4. **Return** $f(x,y) = \sum_{j=1}^{r} \lambda_j (\alpha_j x + \beta_j y)^D$

3.1. Computing the polynomials P_v and P_w

We use Lemma 10 and Lemma 11 to compute the polynomials P_v and P_w from Proposition 7 as a rational reconstruction of $A := \sum_{i=0}^{D} a_i x^i$ modulo x^{D+1}. Our algorithm exploits the Extended Euclidean Algorithm in a similar way as Cabay and Choi (1986) do to compute scaled Padé fractions.

In the following, let v be the vector of Proposition 6, consider $U_v := P_v(1,x)$ and $R_v \in \mathbb{K}[x]$ as the remainder of the division of $(A \cdot P_v(1,x))$ by x^{D+1}. Note that the polynomial R_v is the unique polynomial of degree strictly smaller to $N_1 + 1$ such that $A \cdot P_v(1,x) \equiv R_v \mod x^{D+1}$, see Equation (7).

Lemma 18. *If (U,R) is a rational reconstruction of A modulo x^{D+1} such that* $\max(\deg(U), \deg(R)+1) \leq N_2$*, then there is a polynomial $Q \in \mathbb{K}[x]$ such that $Q \cdot P_v(x,1) = U$ and $Q \cdot R_v = R$.*

Proof. Let $k := \deg(U)$ and $q := \deg(R)$. By Lemma 11, there is a non-zero vector $\omega \in \mathtt{Ker}(H_a^{\max(k,q+1)})$ such that the kernel polynomial P_ω is equal to $U\left(\frac{y}{x}\right) x^{\max(k,q+1)}$. Hence, $\mathtt{Ker}(H_a^{\max(k,q+1)}) \neq 0$ and so, by Proposition 6, $N_1 < \max(k, q+1)$. We assume that $\max(k, q+1) \leq N_2$, hence the degree of P_ω is smaller or equal to N_2 and, by Proposition 7, P_ω is divisible by P_v. Therefore, there is a polynomial $\bar{Q} \in \mathbb{K}[x,y]$ such that $\bar{Q} P_v = P_\omega$. Let $Q := \bar{Q}(1,x)$. By definition, $U_v = P_v(1,x)$ and $U = P_\omega(1,x)$, so $U = Q U_v$. Hence, $Q R_v \equiv R \mod x^{D+1}$, because $R_v \equiv U_v A \mod x^{D+1}$ and $Q R_v \equiv Q U_v A \equiv U A \equiv R \mod x^{D+1}$. If the degrees of $(Q R_v)$ and R are smaller than $D+1$, then $Q R_v = R$, as we want to prove. By assumption, $\deg(R) < N_2 \leq D$ and $\deg(U_v Q) = \deg(U) \leq N_2$. By definition, the degree of R_v is smaller of equal to N_1, and so $\deg(Q R_v) \leq \deg(U_v Q R_v) \leq N_2 + N_1 = D$ (Proposition 4). $\square$

We can use this lemma to recover the polynomial P_v from certain rational reconstructions.

Corollary 19. *If (U,R) is a rational reconstructions of A modulo x^{D+1} such that* $\max(\deg(U), \deg(R)+1) \leq N_2$ *and for every polynomial Q of degree strictly bigger than zero that divides U and R, $(\frac{U}{Q}, \frac{R}{Q})$ is not a rational reconstruction of A modulo x^{D+1}, then there is a non-zero constant c such that $P_v = c \cdot U\left(\frac{y}{x}\right) x^{\max(\deg(U),\deg(R)+1)}$ (Proposition 7). In particular,*

$$N_1 = \max(\deg(U) - 1, \deg(R)).$$

Proof. By Lemma 18, there is a $Q \in \mathbb{K}[x]$ such that $Q \cdot (U_v, R_v) = (U, R)$. By Lemma 10, (U_v, R_v) is a rational reconstruction, and so $\deg(Q) = 0$. Hence, $N_1 + 1 = \deg(P_v) = \max(\deg(U), \deg(R)+1)$ and $Q \cdot P_v(1, \frac{y}{x}) x^{N_1+1} = U(\frac{y}{x}) x^{N_1+1}$. $\square$

If (U,R) is a rational reconstruction of A modulo x^{D+1} such that $\deg(U)+\deg(R) \leq D$ and $U(0)=1$, then $\frac{R}{U}$ is the Padé approximant of A of type $(\deg(R),\deg(U))$ (Bostan et al., 2017, Section 7.1). When this Padé approximant exists, it is unique, meaning that for any rational reconstruction with this property the quotient $\frac{R}{U}$ is unique (we can invert $U \bmod x^{D+1}$ because $U(0)=1$). When $N_1 < N_2$, we have that $\frac{D+1}{2} \leq N_2$ (Proposition 4) and so, if the the Padé approximant of A of type $(\frac{D+1}{2}-1, \frac{D+1}{2})$ exists, by Lemma 18, we can recover P_v from it. The existence of this Padé approximant is equivalent to the condition $U_v(0)=1$, which means $v_{N_1+1}=1$. In the algorithm proposed in the conference version of this paper (Bender et al., 2016, Algorithm 3), the correctness of our algorithms relied on this condition. In that version, we ensured this property with a generic linear change of coordinates in the original polynomial f. In this paper, we skip this assumption. Following Bostan et al. (2017, Theorem 7.2), when $N_1 < N_2$, we can compute v no matter the value of v_{N_1+1}. This approach has a softly-linear arithmetic complexity and involves the computation of a row of the EGCD of A and x^{D+1}. We can compute P_w from a consecutive row.

Before going into the proof, we study the case $N_1 = N_2$. In this case, there are not rational reconstructions with the prerequisites of Lemma 18, and so we treat this case in a different way.

Lemma 20. *If $N_1 = N_2$, then there is a unique rational decomposition (U,R) such that $\deg(U) \leq \frac{D}{2}$, $\deg(R) \leq \frac{D}{2}$ and R is monic. In particular, $\deg(R) = \frac{D}{2}$ and we can consider the kernel polynomial related to $v \in \mathtt{Ker}(H^{N_1+1})$ (Proposition 6) as $P_v = U\left(\frac{y}{x}\right) x^{\frac{D}{2}+1}$.*

Proof. First note that, as $D = N_1 + N_2$ (Proposition 4), then $N_1 = \frac{D}{2}$. Following Equation (6), if we write $U = \sum_{i=0}^{N_1} u_i x^i$ and $R = \sum_{i=0}^{N_1} u_i x^i$, then we get the linear system,

$$\begin{array}{r} \\ \\ \\ H^{N_1}\left\{\begin{array}{c} \\ \\ \\ \\ \end{array}\right. \end{array}
\left(\begin{array}{cccc}
 & & & a_0 \\
 & & a_0 & a_1 \\
 & \iddots & \iddots & \vdots \\
\hline
a_0 & \cdots & a_{N_1-1} & a_{N_1} \\
a_1 & \cdots & a_{N_1} & a_{N_1+1} \\
\vdots & \iddots & \vdots & \vdots \\
a_{D-N_1} & \cdots & a_{D-1} & a_D
\end{array}\right)
\begin{pmatrix} u_{N_1} \\ \vdots \\ u_1 \\ u_0 \end{pmatrix}
=
\left(\begin{array}{c}
r_0 \\ \vdots \\ r_{N_1-1} \\
\hline
r_{N_1} \\ 0 \\ \vdots \\ 0
\end{array}\right).$$

The matrix $H^{N_1} \in \mathbb{K}^{(D-\frac{D}{2}+1)\times(\frac{D}{2}+1)}$ is square and, as $\mathtt{Ker}(H^{N_1}) = 0$ (Proposition 6), it is invertible. If $r_{N_1} = 0$, that is $\deg(R_v) < N_1$, then the polynomial U is zero and so (U,R) is not a rational reconstruction. Hence, we can consider $\deg(U) = N_1$. If R is monic, then $r_{N_1} = 1$ and so we compute the coefficients of U and R as

$$\begin{pmatrix} u_{N_1} \\ \vdots \\ u_1 \\ u_0 \end{pmatrix} = (H^{N_1+1})^{-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} r_0 \\ \vdots \\ r_{N_1-1} \\ 1 \end{pmatrix} = \begin{pmatrix} & & & a_0 \\ & & a_0 & a_1 \\ & \iddots & \iddots & \vdots \\ a_0 & \cdots & a_{N_1-1} & a_{N_1} \end{pmatrix} \begin{pmatrix} u_{N_1} \\ u_{N_1-1} \\ \vdots \\ u_0 \end{pmatrix}.$$

□

Lemma 21 (Correctness of Algorithm 4)**.** *Let $\{(U_j, V_j, R_j)\}_j$ be the set of triplets obtained from the Extended Euclidean Algorithm for the polynomials A and x^{D+1}, see Section 2.4. Let i be the index of the first row of the extended Euclidean algorithm such that $\deg(R_i) < \frac{D+1}{2}$. Then, we can compute the polynomials P_v and P_w of Proposition 7 as*

(A) $P_v = U_i(\frac{x}{y}) \cdot x^{\max(\deg(U_i), \deg(R_i)+1)}$.

(B) *If* $\deg(R_i) > \deg(U_i)$, $P_w = U_{i+1}(\frac{x}{y}) \cdot x^{\deg(U_{i+1})}$.

(C) *If* $\deg(R_i) \leq \deg(U_i)$, $P_w = U_{i-1}(\frac{x}{y}) \cdot x^{\deg(R_{i-1}+1)}$.

Proof. **(A).** First observe that if i is the first index such that the degree of R_i is strictly smaller than $\frac{D+1}{2}$, then, by Remark 13, the degree of R_{i-1} has to be bigger or equal to $\frac{D+1}{2}$. Hence, the degree of U_i is smaller or equal to $\frac{D+1}{2}$, because by Lemma 15, $\deg(U_i) = D+1-\deg(R_{i-1}) \leq D+1-\frac{D+1}{2} = \frac{D+1}{2}$. We can consider R_{i-1}, that is i is strictly bigger than 0, because the degree of $R_0 = x^{D+1}$ is strictly bigger than $\frac{D+1}{2}$.

If $N_1 = N_2$, then D is even and $N_1 = \frac{D}{2}$ (Proposition 4). As $\lfloor \frac{D+1}{2} \rfloor = \frac{D}{2}$, $\deg(R_i) \leq \frac{D}{2}$ and $\deg(U_i) \leq \frac{D}{2}$. By Lemma 20, $\max(\deg(U_i), \deg(R_i)+1) = N_1+1$ and we can consider P_v as $U_i(\frac{y}{x})x^{N_1+1}$.

If $N_1 < N_2$, assume that there is a non-zero $Q \in \mathbb{K}[x]$ such that Q divides U_i and R_i and $(\frac{U_i}{Q}, \frac{R_i}{Q})$ is a rational reconstruction of A modulo x^{D+1}. Then, $\frac{U_i}{Q}A \equiv \frac{R_i}{Q} \mod x^{D+1}$ and so there is a polynomial $\bar{V}$ such that $\bar{V}x^{D+1} + \frac{U_i}{Q}A = \frac{R_i}{Q}$. Multiplying by Q, we get the equality $Q\bar{V}x^{D+1} + U_iA = R_i$. Consider the identity $V_ix^{D+1} + U_iA = R_i$ from Equation (10). Coupling the two equalities together, we conclude that $V_i = Q\bar{V}$. As Q divides U_i and V_i, which are coprime (Lemma 14), Q is a constant, that is $\deg(Q) = 0$. If $N_1 < N_2$, then $D < 2N_2$ (Proposition 4). Hence, $\deg(U_i) \leq \frac{D+1}{2} \leq N_2$ and $\deg(R_i)+1 < \frac{D+1}{2}+1 \leq N_2+1$, that is, $\max(\deg(U_i), \deg(R_i)+1) \leq N_2$. Hence, by

Lemma 18, we can consider $U_i(\frac{y}{x})x^{\max(\deg(U_i),\deg(R_i)+1)}$ as the kernel polynomial P_v from Proposition 7, spanning $\mathtt{Ker}(H^{N_1+1})$.

(B). Assume that the degree of U_i is strictly bigger than the one of R_i, that is $\deg(U_i) > \deg(R_i)$. Then $N_1 = \deg(U_i) - 1$, as $\deg(U_i) = \deg(P_v) = N_1 + 1$ (Remark 12). Note that in this case $i > 1$ because $U_1 = 1$, $R_1 = A \neq 0$, and so $\deg(U_1) \leq \deg(R_1)$. The degree of R_{i-1} is N_2 because, by Lemma 15, $\deg(R_{i-1}) = D + 1 - \deg(U_i) = D + 1 - N_1 - 1 = N_2$ (Proposition 4). Consider the degree of U_{i-1}. By Lemma 15, $\deg(U_{i-1}) = D + 1 - \deg(R_{i-2})$. As $\deg(R_{i-2}) > \deg(R_{i-1})$ (Remark 13), then $\deg(R_{i-2}) > N_2$. Therefore, the degree of U_{i-1} is smaller or equal to the one of R_{i-1} because

$$\deg(U_{i-1}) = D + 1 - \deg(R_{i-2}) < D + 1 - N_2 = N_1 + 1, \text{ and so}$$

$$\deg(U_{i-1}) \leq N_1 \leq N_2 = \deg(R_{i-1}).$$

Hence, by Remark 16, (U_{i-1}, R_{i-1}) is a rational reconstruction of A modulo x^{D+1} such that $\deg(U_{i-1}) \leq N_1$ and $\deg(R_{i-1}) = N_2$. So, $\max(\deg(U_{i-1}), \deg(R_{i-1}) + 1) = N_2 + 1$ and, by Remark 12, there is a kernel polynomial $P_\omega = U_{i-1}(\frac{y}{x})x^{N_2+1}$ of degree $N_2 + 1$ such that $x^{N_2+1-\deg(U_{i-1})}$ divides P_ω. As $\deg(U_{i-1}) \leq N_1$, $x^{N_2+1-N_1}$ divides $x^{N_2+1-\deg(U_{i-1})}$ and so, it divides P_ω. We assumed that the degree of U_i is strictly bigger than the one of R_i, and so x does not divide P_v (Remark 12). Hence, there is no binary form Q of degree $N_2 - N_1$ such that $x^{N_2-N_1+1}$ divides QP_v. Therefore, by Corollary 8, we can consider $P_w = P_\omega$.

(C). Assume that the degree of R_i is bigger or equal to the one of U_i, that is $\deg(R_i) \geq \deg(U_i)$. Hence, $\deg(R_i) + 1 = \deg(P_v) = N_1 + 1$ (Remark 12), and so $\deg(R_i) = N_1$. In particular, $R_i \neq 0$, and so the $(i+1)$-th row of the Extended Euclidean Algorithm, $(U_{i+1}, V_{i+1}, R_{i+1})$, is defined. The degree of U_{i+1} is $N_2 + 1$ because, by Lemma 15, $\deg(U_{i+1}) = D + 1 - \deg(R_i) = N_2 + 1$ (Proposition 4). The degree of R_{i+1} is strictly smaller than the one of R_i (Remark 13), which is N_1. Hence, the degree of R_{i+1} is smaller than the degree of U_{i+1} because $\deg(R_{i+1}) < N_1 \leq N_2 < \deg(U_{i+1})$. Therefore, $P_\omega = U_{i+1}(\frac{y}{x})x^{N_2+1}$ is a kernel polynomial in $\mathtt{Ker}(H^{N_2+1})$ (Lemma 11). By Remark 12, as $\deg(R_{i+1}) < \deg(U_{i+1})$, x does not divide P_ω. Also, the maximal power of x that divides P_v is $x^{\deg(R_i)+1-\deg(U_i)}$, and, as we assumed $\deg(R_i) \geq \deg(U_i)$, x divides P_v. Hence, every polynomial in $\{QP_v : \deg(Q) = N_2 - N_1\}$ is divisible by x, and so, by Corollary 8, we can consider $P_w = P_\omega$. □

Algorithm 4 COMPUTE_PV_AND_PW

Input: A sequence $(a_0, \ldots, a_D)$.

Output: Polynomials P_v and P_w as 7.

1. $i \leftarrow$ first row of EGCD$(x^{D+1}, \sum_{i=0} a_i x^i)$ such that $R_i < \frac{D+1}{2}$.
2. $P_v \leftarrow U_i(\frac{x}{y}) \cdot x^{\max(\deg(U_i), \deg(R_i)+1)}$.

   $N_1 \leftarrow \max(\deg(U_i) - 1, \deg(R_i))$
3. **IF** $\deg(R_i) > \deg(U_i)$,

   $P_w \leftarrow U_{i+1}(\frac{x}{y}) \cdot x^{\deg(U_{i+1})}$.

   $N_2 \leftarrow \deg(U_{i+1}) - 1$.

   ELSE

   $P_w \leftarrow U_{i-1}(\frac{x}{y}) \cdot x^{\deg(R_{i-1}+1)}$.

   $N_2 \leftarrow \deg(R_{i-1})$.
4. **Return** P_v and P_w

3.2. Computing a square-free polynomial Q

We can compute Q at step 2 of Algorithm 3 in different ways. If P_v is square-free, then we set Q equal to P_v. If P_v is not square-free, by Lemma 17, we need to find a vector $\mu \in \mathbb{K}^{(N_2-N_1+1)}$ such that $Q_\mu := P_\mu \cdot P_v + P_w$ is square-free. By Proposition 7, P_v and P_w are relative prime. Thus, if we take a random vector μ, generically, Q_μ would be square-free. For this to hold, we have to prove that the discriminant of Q_μ is not identically zero. To simplify notation, in the following theorem we dehomogenize the polynomials.

Theorem 22. *Given two relative prime univariate polynomials $P_v(x)$ and $P_w(x)$ of degrees N_1+1 and N_2+1 respectively, let $Q_\mu(x) := P_\mu(x,1) \cdot P_v + P_w \in \mathbb{K}[\mu_0, \ldots, \mu_{N_2-N_1}][x]$. The discriminant of $Q_\mu(x)$ with respect to x is a non-zero polynomial.*

Proof. The zeros of the discriminant of $Q_\mu(x)$ with respect to x over $\mathbb{K}$ correspond to the set $\{\mu \in \mathbb{K}^{N_2-N_1+1} : Q_\mu \text{ has double roots}\}$. We want to prove that the discriminant is not zero.

A univariate polynomial is square-free if and only if it does share any root with its derivative. Hence, $(\mu_0,\dots,\mu_{N_2-N_1})^\mathsf{T} \in \{\mu \in \mathbb{K}^{N_2-N_1+1} : Q_\mu$ has double roots$\}$ if and only if, there is $(\mu_0,\dots,\mu_{N_2-N_1},\alpha) \in \mathbb{K}^{N_2-N_1+1} \times \overline{\mathbb{K}}$ such that it is a solutions to the following system of equations

$$\begin{cases} (P_\mu \cdot P_v + P_w)(x) = 0 \\ (P_\mu \cdot P_v' + P_\mu' \cdot P_v + P_w')(x) = 0. \end{cases} \tag{11}$$

In Equation (11), μ_0 only appears in $P_\mu(x,1)$ with degree 1. We eliminate it to obtain the polynomial

$$(P_v \cdot P_\mu' + P_w')P_v - P_v' \cdot P_w.$$

This polynomial is not identically 0 as P_v' does not divide P_v and P_v and P_w are relative prime. Hence, for each $(\mu_1,\dots,\mu_{N_2-N_1})$, there is a finite number of solutions for this equation, bounded by the degree of the polynomial. Moreover, as the polynomials of Equation (11) are linear in μ_0, each solution of the deduced equation is extensible to a finite number of solutions of Equation (11). Hence, there is a $\mu \in \mathbb{K}^{N_2-N_1+1}$, such that Q_μ is square-free. Therefore, the discriminant of $Q_\mu(x)$ is not identically zero. □

Corollary 23. *For every vector $(\mu_1,\dots,\mu_{N_2-N_1}) \in \mathbb{K}^{N_2-N_1}$ such that there is a $\mu_0 \in \mathbb{K}$ such that y^2 does not divides Q_μ, where $\mu = (\mu_0,\dots,\mu_{N_2-N_1})$, there are at most $2D+2$ different values for $\mu_0 \in \mathbb{K}$ such that the polynomial $Q_\mu(x,y)$ is not square-free.*

Proof. If $Q_\mu(x,y)$ is not square-free, then it has a double root in $\mathbb{P}^1(\overline{\mathbb{K}})$. This root could be of the form $(\alpha,1) \in \mathbb{P}^1(\overline{\mathbb{K}})$ or $(1,0) \in \mathbb{P}^1(\overline{\mathbb{K}})$. We analyze separately these cases.

First, we consider the polynomial $Q_\mu(x,1) \in \mathbb{K}[\mu_0,x]$. By Theorem 22, the discriminant of $Q_\mu(x,1)$ with respect to x is not zero. As $Q_\mu(x,1)$ is a polynomial of degree N_2+1 with respect to x, and of degree 1 with respect to μ_0, the degree with respect to μ_0 of the discriminant of $Q_\mu(x,1)$ with respect to x is at most $(N_2+1)+N_2 \leq 2D+1$. Hence, there are at most $2D+1$ values for μ_0 such that $Q_\mu(x,y)$ has a root of the form $(\alpha,1) \in \mathbb{P}^1(\overline{\mathbb{K}})$ with multiplicity bigger than one.

The polynomial $Q_\mu(x,y)$ has a root of the form $(1,0) \in \mathbb{P}^1(\mathbb{K})$ with multiplicity bigger than one, if and only if y^2 divides $Q_\mu(x,y)$. If this happens, then the coefficients of the monomials $y \cdot x^{N_2-N_1-1}$ and $x^{N_2-N_1}$ in the polynomial $Q_\mu(x,y)$ are zero. By assumption, these coefficients are not identically zero as polynomials in $\mathbb{K}[\mu_0]$. As $Q_\mu(x,y)$ is a linear polynomial with respect to μ_0, there is at most one value for μ_0 such that y^2 divides $Q_\mu(x,y)$.

Therefore, there are at most $(2D+1)+1$ values such that $Q_\mu(x,y)$ is not square-free. $\square$

Remark 24. *The previous assumption is not restrictive. If y^2 divides Q_μ, where $\mu = (\mu_0,\dots,\mu_{N_2-N_1})$, then y^2 does not divide $Q_{(\mu_0,\dots,\mu_{N_2-N_1}+1)} = Q_\mu + x^{N_2+1}$ nor $Q_{(\mu_0,\dots,\mu_{N_2-N_1-1}+1,\mu_{N_2-N_1})} = Q_\mu + yx^{N_2}$. Moreover, if $N_2 - N_1 \geq 2$, y^2 divides (or not) $Q_\mu(x,y)$ regardless the value of μ_0. Conversely, if $N_2 - N_1 < 2$, there is always a μ_0 such that y^2 does not divide Q_μ.*

3.3. Correctness of Algorithm 3

For computing a decomposition for a binary form f, we need to compute the kernel of a Hankel matrix (Theorem 2). Algorithm 4 computes correctly the polynomials P_v and P_w that characterize the kernels of the family of Hankel matrices associated to f. Once we obtain these polynomials, step 2 (see Corollary 23) and step 3 computes the coefficients $\alpha_j, \beta_j, \lambda_j$ of the decomposition. Hence, we have a decomposition for f, as $f(x,y) = \sum_{j=1}^{r} \lambda_j \cdot (\alpha x + \beta y)^D$.

Example. Consider $f(x,y) = y^4 + 8xy^3 + 18x^2y^2 + 16x^3y + 5x^4$. The family of Hankel matrices associated to f are related to the vector $a := (1,2,3,4,5)^\intercal$, it is denoted by $\{H_a^k\}_k$, and it contains the following matrices:

$H_a^1 = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \\ 4 & 5 \end{pmatrix}$	$H_a^2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$	$H_a^3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \end{pmatrix}$	$H_a^4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \end{pmatrix}$

The kernel H_a^1 is trivial, so we compute the one of H_a^2. This kernel is generated by the vector $(1,-2,1)^\intercal$, so by Proposition 6 we consider $v = (1,-2,1)^\intercal$. Also, by Proposition 4, $N_1 + 1 = 2$ and $N_2 = D - N_1 = 3$. The kernel polynomial $P_v = y^2 - 2xy + x^2 = (x-y)^2$ is not square-free thus, by Lemma 17, the rank of $f(x,y)$ is $N_2 + 1 = 4$ and we have to compute the kernel polynomial P_w in the kernel of H_a^4. Following Proposition 6, the kernel of H_a^4 is generated by U-chain of v given vectors $\mathbb{U}_2^0 v = (1,-2,1,0,0)^\intercal$, $\mathbb{U}_2^1 v = (0,1,-2,1,0)^\intercal$, and $\mathbb{U}_2^2 v = (0,0,1,-2,1)^\intercal$, plus a vector w linear independent with this U-chain. We consider the vector $w = (0,0,0,5,-4)$, which fulfills that assumption. Hence, $P_v = y^2 - 2xy + x^2$ and $P_w = 5yx^3 - 4x^4$.

We proceed by computing a square-free polynomial combination of P_v and P_w. For that, we choose

$$Q := (44y^2 + 11yx + 149x^2)P_v + 36P_w = (5x - 11y)(x - 2y)(x + 2y)(x + y).$$

Finally, we solve the system given by the transposed of a Vandermonde matrix,

$$\begin{pmatrix} 5^4 & 1 & 1 & 1 \\ 11 \cdot 5^3 & 2 & -2 & -1 \\ 11^2 \cdot 5^2 & 2^2 & (-2)^2 & (-1)^2 \\ 11^3 \cdot 5 & 2^3 & (-2)^3 & (-1)^3 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}. \tag{12}$$

The unique solution of the system is $(-\frac{1}{336}, 3, \frac{1}{21}, \frac{3}{16})^{\mathsf{T}}$, and so we recover the decomposition

$$f(x,y) = -\frac{1}{336}(11x+5y)^4 + 3(2x+y)^4 + \frac{1}{21}(-2x+y)^4 - \frac{3}{16}(-x+y)^4.$$

Instead of considering incrementally the matrices in the Hankel family we can compute the polynomials P_v and P_w faster by applying Algorithm 4. For this, we consider the polynomial $A := 5x^4 + 4x^3 + 3x^2 + 2x + 1$, and the rows of the Extended Euclidean Algorithm for A and x^5.

j	V_j	U_j	R_j
0	1	0	x^5
1	0	1	$5x^4 + 4x^3 + 3x^2 + 2x + 1$
2	1	$\frac{1}{25}(5x-4)$	$\frac{1}{25}(x^3 + 2x^2 + 3x + 4)$
3	$-25(5x-6)$	$25(x-1)^2$	25
4	$\frac{1}{25}(5x^4 + 4x^3 + 3x^2 + 2x + 1)$	$-\frac{1}{25}x^5$	0

We need to consider the first j such that $\deg(R_j) < \frac{5}{2}$, which is $j = 3$. Hence, $N_1 = \max(\deg(U_3) - 1, \deg(R_3)) = 1$ and

$$P_v := U_3\Big(\frac{y}{x}\Big)\, x^{\max(\deg(U_3), \deg(R_3)+1)} = 25\left(\frac{y}{x} - 1\right)^2 x^2 = 25(y-x)^2.$$

As $\deg(R_3) \leq \deg(U_3)$, we consider $N_2 = \deg(R_2) = 3$ and

$$P_w := U_2\Big(\frac{y}{x}\Big)\, x^{\deg(R_2)+1} = \frac{1}{25}(5yx^3 - 4x^4). \qquad \Diamond$$

The real case. When we consider the decomposition of binary forms over $\mathbb{R}$, Algorithm 3 might fail. This happens either when the decomposition over $\mathbb{C}$ is not unique or when the decomposition over $\mathbb{C}$ is unique but it is not a decomposition over $\mathbb{R}$. The algorithm fails because

- The real rank of the binary form might be bigger than $N_2 + 1$, see Reznick (2013a). Lemma 17 does not hold over $\mathbb{R}$ and so we cannot find a square-free kernel polynomial Q_μ that factors over $\mathbb{R}$.
- Even when the real rank of the binary form is $N_2 + 1$, we need to perform some extra computations to compute a square-free kernel polynomial Q_μ that factors over $\mathbb{R}$. This computations are not taken into account in our algorithm, so we could never find such a decomposition.

Recently, Sylvester's algorithm was extended to the real case (Ansola et al., 2017, Algorithm 2). This algorithm performs an incremental search over r as in Algorithm 1, but it decides if there is a real decomposition of length r by checking the emptiness of a semi-algebraic set. In the step r-th of the algorithm, the semi-algebraic set is embedded in $\mathbb{R}^{\mathtt{dim}(\mathtt{Ker}(H^r))}$. Hence, the bottleneck of their algorithm is not the computation of P_v and P_w as in our case, but deciding the emptiness of the semi-algebraic set. We emphasize that when the decompositions over $\mathbb{C}$ and $\mathbb{R}$ are unique and the same, our algorithm computes such a decomposition. Moreover, given P_v, we can check if the previous condition holds by checking if P_v splits over $\mathbb{R}$.

4. Complexity

In this section we study the algebraic degree of the parameters that appear in the decomposition of a binary form as well as the arithmetic and bit complexity of Algorithm 3.

4.1. Algebraic degree of the problem

If we assume that the coefficients of the input binary form Equation (1) are rational numbers then the parameters of the decompositions, α_j, β_j, and λ_j (see Equation (3)), are algebraic numbers, that is, roots of univariate polynomials with integer coefficients. The maximum degree of these polynomials is the algebraic degree of the problem. We refer the interested reader to Bajaj (1988); Nie et al. (2010); Draisma et al. (2016) for a detailed exposition about the algebraic degree and how it address the complexity of the problem at hand at a fundamental level.

4.1.1. The complexity of computing Q

Recall that, from Lemma 17, the rank of f could be either $N_1 + 1$ or $N_2 + 1$. When the polynomial P_v is square-free, then the rank is $N_1 + 1$ and $Q = P_v$. Following the discussion of

Section 3.2, we prove that, when the rank of the binary form is N_2+1, we can compute a square-free kernel polynomial Q of this degree such that the largest degree of its irreducible factors is N_1. Moreover, we prove that for almost all the choices of (N_2-N_1+1) different points in $\mathbb{P}^1(\mathbb{K})$ (the projective space of $\mathbb{K}$) there is a square-free kernel polynomial of H^{N_2+1} which vanishes on these points. This will be our choice for Q.

Lemma 25. *Let f be a binary form of rank N_2+1. Given (N_2-N_1+1) different points $(\alpha_0,\beta_0),\dots,(\alpha_{N_2-N_1},\beta_{N_2-N_1}) \in \mathbb{P}^1(\mathbb{K})$ such that none of them is a root of P_v, then there is a unique binary form P_μ of degree N_2-N_1, such that the kernel polynomial $Q_\mu := P_\mu \cdot P_v + P_w$ vanishes on these points.*

Proof. Without loss of generality, we assume $\beta_i = 1$, for $i \in \{1,\dots,N_2-N_1\}$. By Proposition 7, for any polynomial P_μ of degree N_2-N_1, Q_μ is a kernel polynomial. Since $Q_\mu(\alpha_i,1)=0$, we can interpolate P_μ by noticing that $P_\mu(\alpha_j,1) = -\frac{P_w(\alpha_j,1)}{P_v(\alpha_j,1)}$.

The degree of P_μ is (N_2-N_1) and we interpolate it using (N_2-N_1+1) different points. Hence there is a unique interpolation polynomial P_μ. So, Q_μ is the unique kernel polynomial of H^{N_2+1} vanishing at all these points. $\square$

Example (cont.). For the example of Section 3.3, we obtained the square-free kernel polynomial Q by choosing the points $(2,1)$, $(-2,1)$ and $(-1,1) \in \mathbb{P}^1(\mathbb{K})$. If we choose other points such that Q is square-free, we will obtain a different decomposition. Hence, f does not have a unique decomposition. This holds in general. $\diamondsuit$

From Lemma 25, we deduce the following well-known result about the uniqueness of the decomposition, see also Helmke (1992); Comas and Seiguer (2011); Bernardi et al. (2011); Carlini et al. (2017).

Corollary 26. *A decomposition is unique if and only if the rank is N_1+1 and $N_1 < N_2$. A decomposition is not unique if and only if the rank is N_2+1.*

Theorem 27. *Let the rank of f be N_2+1. Then there is a square-free kernel polynomial Q such that the largest degree of its irreducible factors is at most N_1.*

Proof. If the rank of f is N_2+1, then for each set of N_2-N_1+1 different points in $\mathbb{P}^1(\overline{\mathbb{K}})$, following the assumptions of Lemma 25, there is a unique kernel polynomial. There is a rational map that realizes this relation (see the proof of Lemma 25). Let this map be $Q_{[\overline{\alpha}]}$, where

$\overline{\alpha} = ((\alpha_0,\beta_0),\ldots,(\alpha_{N_2-N_1},\beta_{N_2-N_1})) \in \mathbb{P}^1(\overline{\mathbb{K}})^{N_2-N_1+1}$. The image of the map is contained in $\{P_\mu \cdot P_v + P_w : \mu \in \overline{\mathbb{K}}^{N_2-N_1+1}\}$. This set and $\mathbb{P}^1(\overline{\mathbb{K}})^{N_2-N_1+1}$ have the same dimension, $N_2 - N_1 + 1$.

Given a kernel polynomial $\hat{Q}(x,y)$, there is a finite number of distinct points $(\alpha,\beta) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $\hat{Q}(\alpha,\beta) = 0$. Hence, the pre-image of an element in the image of $Q_{[\overline{\alpha}]}$ is a finite set. Therefore, the dimension of the image and the dimension of the domain are the same.

By Theorem 22, the non-square-free kernel polynomials form a hypersurface in the space of kernel polynomials of the shape $P_\mu \cdot P_v + P_w$. If we consider the pre-image of the intersection between this hypersurface and the image of the rational map, then its dimension is smaller than $N_2 - N_1 + 1$.

Therefore, generically, for $N_2 - N_1 + 1$ different points in $\mathbb{P}^1(\mathbb{K})$, the map $Q_{[\overline{\alpha}]}(x,y)$ results a square-free kernel polynomial. As $\overline{\mathbb{K}}$ is the algebraic closure of $\mathbb{K} \subset \mathbb{C}$, the same holds over $\mathbb{K}$. □

Theorem 28. *Given a binary form f of rank r and degree D, there is a square-free kernel polynomial of degree r such that the biggest degree of its irreducible factors is* $\min(r, D - r + 1)$.

Proof. If the rank is $r = N_2 + 1$, then $\min(r, D - r + 1) = N_1$. By Theorem 27, such a square-free kernel polynomial exists. If the rank is $r = N_1 + 1$ and $N_1 < N_2$, by Lemma 17, there is a square-free kernel polynomial of degree $\min(r, D - r + 1) = N_1 + 1$. □

The previous result is related to the decomposition of tensors of the same border rank (Comas and Seiguer, 2011, Theorem 2); (Bernardi et al., 2011, Theorem 23); (Blekherman, 2015).

We can also bound the number of possible bad choices in the proof of Theorem 27.

Proposition 29. *Let f be a binary form of rank $N_2 + 1$. For every set $S \subset \mathbb{P}^1(\mathbb{K})$ of cardinal $(N_2 - N_1)$ such that $(\forall(\alpha,\beta) \in S)$ $P_v(\alpha,\beta) \neq 0$ there are at most $D^2 + 3D + 1$ values $(\alpha_0,\beta_0) \in \mathbb{P}^1(\mathbb{K})$ such that $(\alpha_0,\beta_0) \notin S$, $P_v(\alpha_0,\beta_0) \neq 0$ and the unique kernel polynomial $Q_\mu := P_\mu \cdot P_v + P_w$ that vanish over S and (α_0,β_0) (Lemma 25) is not square-free.*

To prove this proposition we use Lagrange polynomials to construct the maps and varieties of the proof of Theorem 27.

Let $S = \{(\alpha_1,\beta_1),\ldots,(\alpha_{N_2-N_1},\beta_{N_2-N_1})\} \subset \mathbb{P}^1(\mathbb{K})$ be the set of Proposition 29. For each $(\alpha_0,\beta_0) \in \mathbb{P}^1(\mathbb{K})$ such that $(\alpha_0,\beta_0) \notin S$ and $P_v(\alpha_0,\beta_0) \neq 0$ we consider the unique kernel polynomial Q^{α_0,β_0} which vanishes at S and (α_0,β_0), see Lemma 25. Using Lagrange polynomial, we can write this polynomial as

$$Q^{\alpha_0,\beta_0}(x,y) = \left(-\frac{P_w(\alpha_0,\beta_0)}{P_v(\alpha_0,\beta_0)} \frac{M(x,y)}{M(\alpha_0,\beta_0)} + \sum_{i=1}^{N_2-N_1} \frac{\beta_0 x - \alpha_0 y}{\alpha_0\beta_i - \alpha_i\beta_0} E_i(x,y) \right) P_v(x,y) + P_w(x,y)$$

Where $E_i(x,y) := -\frac{P_w(\alpha_i,\beta_i)}{P_v(\alpha_i,\beta_i)} \prod_{j\notin\{0,i\}} \frac{\beta_j x - \alpha_j y}{\alpha_i\beta_j - \alpha_j\beta_i}$ and $M(x,y) := \prod_{j=1}^{N_2-N_1}(\beta_j x - \alpha_j y)$.[2]

For each $(\alpha_j,\beta_j) \in S$, we characterize the possible $(\alpha_0,\beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that (α_j,β_j) is a root of Q^{α_0,β_0} of multiplicity bigger than one. Then, we study the $(\alpha_0,\beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that (α_0,β_0) is a root of Q^{α_0,β_0} with multiplicity bigger than one. Finally, we reduce every case to the previous ones.

To study the multiplicities of the roots, we use the fact that (α_0,β_0) is a double root of a binary form P if and only if $P(\alpha_0,\beta_0) = \frac{\partial P}{\partial x}(\alpha_0,\beta_0) = \frac{\partial P}{\partial y}(\alpha_0,\beta_0) = 0$. Hence, for each $(\alpha_0,\beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$, we consider $\frac{\partial Q^{\alpha_0,\beta_0}}{\partial x}$ and $\frac{\partial Q^{\alpha_0,\beta_0}}{\partial y}$, where

$$\begin{aligned} \frac{\partial Q^{\alpha_0,\beta_0}}{\partial x} = & -\frac{P_w(\alpha_0,\beta_0)}{P_v(\alpha_0,\beta_0)} \frac{1}{M(\alpha_0,\beta_0)} \left(\frac{\partial M}{\partial x} P_v + M \frac{\partial P_v}{\partial x} \right)(x,y) + \\ & \sum_{i=1}^{N_2-N_1} \frac{1}{\beta_0\alpha_i - \alpha_0\beta_i} \frac{\partial((\beta_0 x - \alpha_0 y) E_i P_v)}{\partial x}(x,y) + \frac{\partial P_w}{\partial x}(x,y) \end{aligned} \tag{13}$$

Let $O_x^{\alpha_0,\beta_0}(x,y)$ be the product between the last line of Equation (13) and $M(\alpha_0,\beta_0)$, that is

$$O_x^{\alpha_0,\beta_0}(x,y) := \sum_{i=1}^{N_2-N_1} \frac{M(\alpha_0,\beta_0)}{\beta_0\alpha_i - \alpha_0\beta_i} \frac{\partial((\beta_0 x - \alpha_0 y) E_i P_v)}{\partial x}(x,y) + M(\alpha_0,\beta_0) \frac{\partial P_w}{\partial x}(x,y)$$

In what follows, instead of considering the pair (α_0,β_0) as a point in $\mathbb{P}^1(\overline{\mathbb{K}})$, we consider it as a pair of variables. Hence, for every $(\alpha_i,\beta_i) \in S$, $(\beta_0\alpha_i - \alpha_0\beta_i)$ divides $M(\alpha_0,\beta_0)$, as polynomials in $\overline{\mathbb{K}}[\alpha_0,\beta_0]$, so $O_x^{\alpha_0,\beta_0}(x,y)$ is a polynomial in $\overline{\mathbb{K}}[\alpha_0,\beta_0][x,y]$. The derivative of Q^{α_0,β_0} with respect to x is a rational function in $\overline{\mathbb{K}}(\alpha_0,\beta_0)[x,y]$, that we can write as $\frac{\partial Q^{\alpha_0,\beta_0}}{\partial x} = \frac{T^{\alpha_0,\beta_0}(x,y)}{P_v(\alpha_0,\beta_0) M(\alpha_0,\beta_0)}$ where

$$T^{\alpha_0,\beta_0}(x,y) := -P_w(\alpha_0,\beta_0) \left(\frac{\partial M}{\partial x} P_v + M \frac{\partial P_v}{\partial x} \right)(x,y) + O_x^{\alpha_0,\beta_0}(x,y) P_v(\alpha_0,\beta_0) \in \overline{\mathbb{K}}[\alpha_0,\beta_0][x,y]$$

Lemma 30. *For each $(\alpha_i,\beta_i) \in S$, there are at most N_2+1 possible $(\bar{\alpha}_0,\bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\bar{\alpha}_0,\bar{\beta}_0) \notin S$, $P_v(\bar{\alpha}_0,\bar{\beta}_0) \neq 0$ and that (α_i,β_i) is a root of multiplicity bigger than 1 in $Q^{\bar{\alpha}_0,\bar{\beta}_0}$.*

[2] For each $0 \leq i \leq N_2 - N_1$, $Q^{\alpha_0,\beta_0}(x,y)$ is a rational function of degree 0 with respect to (α_i,β_i). Hence, it is well defined the evaluation of the variables (α_i,β_i) in $Q^{\alpha_0,\beta_0}(x,y)$ at points of $\mathbb{P}^1(\mathbb{K})$.

Proof. If (α_i, β_i) is a root of multiplicity bigger than 1 in $Q^{\bar{\alpha}_0,\bar{\beta}_0}$, then $\frac{\partial Q^{\bar{\alpha}_0,\bar{\beta}_0}}{\partial x}(\alpha_i, \beta_i) = 0$. Hence, we are looking for the $(\bar{\alpha}_0, \bar{\beta}_0)$ such that $T^{\bar{\alpha}_0,\bar{\beta}_0}(\alpha_i, \beta_i) = 0$. The polynomial $T^{\alpha_0,\beta_0}(\alpha_i, \beta_i)$ belongs to $\overline{\mathbb{K}}[\alpha_0, \beta_0]$, so if it is not identically zero, then there are a finite number of points $(\bar{\alpha}_0, \bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $T^{\bar{\alpha}_0,\bar{\beta}_0}(\alpha_i, \beta_i) = 0$. Moreover, the degree of the polynomial $T^{\alpha_0,\beta_0}(\alpha_i, \beta_i)$ is at most $\max(\deg(P_w), \deg(O_x^{\alpha_0,\beta_0}(\alpha_i, \beta_i)) + \deg(P_v)) = N_2 + 1$. Hence, if the polynomial is not zero, this finite number is at most $N_2 + 1$.

The polynomial $T^{\alpha_0,\beta_0}(\alpha_i, \beta_i) \in \overline{\mathbb{K}}[\alpha_0, \beta_0]$ is not zero. Observe that as M is square-free, $M(\alpha_i, \beta_i) = 0$ and $P_v(\alpha_i, \beta_i) \neq 0$, then $\left(\frac{\partial M}{\partial x} P_v + M \frac{\partial P_v}{\partial x}\right)(\alpha_i, \beta_i) \neq 0$. Hence, as P_w and P_v are coprime, if $(\bar{\alpha}_0, \bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ and $P_v(\bar{\alpha}_0, \bar{\beta}_0) = 0$, then $T^{\bar{\alpha}_0,\bar{\beta}_0}(\alpha_i, \beta_i) \neq 0$. That is, $T^{\alpha_0,\beta_0}(\alpha_i, \beta_i)$ does not vanish in the roots of P_v. □

Lemma 31. *There are at most $2N_2 + 1$ possible $(\bar{\alpha}_0, \bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\bar{\alpha}_0, \bar{\beta}_0) \notin S$, $P_v(\bar{\alpha}_0, \bar{\beta}_0) \neq 0$ and $(\bar{\alpha}_0, \bar{\beta}_0)$ is a root of multiplicity bigger than 1 in $Q^{\bar{\alpha}_0,\bar{\beta}_0}$.*

Proof. Following the proof of Lemma 30, we study $T^{\alpha_0,\beta_0}(\alpha_0, \beta_0) \in \overline{\mathbb{K}}[\alpha_0, \beta_0]$.

$$
\begin{aligned}
T^{\alpha_0,\beta_0}(\alpha_0, \beta_0) &= -P_w(\alpha_0, \beta_0)\left(\frac{\partial M}{\partial x} P_v + M \frac{\partial P_v}{\partial x}\right)(\alpha_0, \beta_0) + O_x^{\alpha_0,\beta_0}(\alpha_0, \beta_0) P_v(\alpha_0, \beta_0) \\
&= \left(-P_w M \frac{\partial P_v}{\partial x}\right)(\alpha_0, \beta_0) + \left(O_x^{\alpha_0,\beta_0} - P_w \frac{\partial M}{\partial x}\right)(\alpha_0, \beta_0) P_v(\alpha_0, \beta_0)
\end{aligned}
$$

Note that P_w and P_v are coprime. Also, M and P_v are coprime. Hence, the polynomial $T^{\alpha_0,\beta_0}(\alpha_0, \beta_0)$ is not zero because P_v does not divide $P_w M \frac{\partial P_v}{\partial x}$. We conclude the proof by noting that the degree of $T^{\alpha_0,\beta_0}(\alpha_0, \beta_0)$ is bounded by $2N_2 + 1$. □

Lemma 32. *Let $(\bar{\alpha}_0, \bar{\beta}_0), (\hat{\alpha}_0, \hat{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\bar{\alpha}_0, \bar{\beta}_0), (\hat{\alpha}_0, \hat{\beta}_0) \notin S$, $P_v(\bar{\alpha}_0, \bar{\beta}_0) \neq 0$. Hence, $Q^{\bar{\alpha}_0,\bar{\beta}_0}(\hat{\alpha}_0, \hat{\beta}_0) = 0$ if and only if $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x, y) = Q^{\hat{\alpha}_0,\hat{\beta}_0}(x, y)$.*

Proof. Assume that $Q^{\bar{\alpha}_0,\bar{\beta}_0}(\hat{\alpha}_0, \hat{\beta}_0) = 0$. Following Lemma 25, we write $Q^{\bar{\alpha}_0,\bar{\beta}_0} = P_{\bar{\mu}} P_v + P_w$ and $Q^{\hat{\alpha}_0,\hat{\beta}_0} = P_{\hat{\mu}} P_v + P_w$. As P_v and P_w are coprime and $Q^{\bar{\alpha}_0,\bar{\beta}_0}(\hat{\alpha}_0, \hat{\beta}_0) = 0$, then $P_v(\hat{\alpha}_0, \hat{\beta}_0) \neq 0$. Consider $Q^{\bar{\alpha}_0,\bar{\beta}_0} - Q^{\hat{\alpha}_0,\hat{\beta}_0} = P_v(P_{\bar{\mu}} - P_{\hat{\mu}})$. This polynomial belongs to $\mathbb{K}[x, y]$ and it vanishes over $\mathbb{P}^1(\overline{\mathbb{K}})$ at the $N_1 + 1$ roots of P_v, at the $N_2 - N_1$ points on S, and at $(\hat{\alpha}_0, \hat{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$. Hence, $Q^{\bar{\alpha}_0,\bar{\beta}_0} - Q^{\hat{\alpha}_0,\hat{\beta}_0} = 0$ as it is a binary form in $\mathbb{K}[x, y]$ of degree at most $N_2 + 1$ with $N_2 + 2$ different roots over $\mathbb{P}^1(\overline{\mathbb{K}})$. Therefore, $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x, y) = Q^{\hat{\alpha}_0,\hat{\beta}_0}(x, y)$.

To prove the second case, note that, by definition, $Q^{\hat{\alpha}_0,\hat{\beta}_0}(\hat{\alpha}_0, \hat{\beta}_0) = 0$. Hence, if we assume that $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x, y) = Q^{\hat{\alpha}_0,\hat{\beta}_0}(x, y)$, then we have $Q^{\bar{\alpha}_0,\bar{\beta}_0}(\hat{\alpha}_0, \hat{\beta}_0) = 0$. □

Proof of Proposition 29. We want to bound the number of different points $(\bar{\alpha}_0, \bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y)$ is not a square-free binary form over $\overline{\mathbb{K}}[x,y]$. If the binary form $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y)$ is not square-free, then it has a root over $\mathbb{P}^1(\overline{\mathbb{K}})$ with multiplicity bigger than one. If such a root is $(\alpha_i, \beta_i) \in S$, we can bound the possible number of different values for $(\bar{\alpha}_0, \bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ by (N_2+1) (Lemma 31). Hence, if there is a i such that $(\alpha_i, \beta_i) \in S$ has multiplicity bigger than one as a root of $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y)$, we can bound the possible number of different values for $(\bar{\alpha}_0, \bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ by $\#S \cdot (N_2+1) = (N_2-N_1)(N_2+1)$.

If $Q^{\bar{\alpha}_0,\bar{\beta}_0}$ is not square-free and the multiplicity of every root $(\alpha_i, \beta_i) \in S$ is one, then there must be a root $(\hat{\alpha}_0, \hat{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\hat{\alpha}_0, \hat{\beta}_0) \notin S$ and its multiplicity as a root of $Q^{\bar{\alpha}_0,\bar{\beta}_0}$ is bigger strictly than one. By Lemma 32, $Q^{\hat{\alpha}_0,\hat{\beta}_0}(x,y) = Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y)$, and so $(\hat{\alpha}_0, \hat{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ has multiplicity bigger than one as a root of $Q^{\hat{\alpha}_0,\hat{\beta}_0}(x,y)$. Hence, $P_v(\hat{\alpha}_0, \hat{\beta}_0) \neq 0$ and, by Lemma 31, we can bound the possible number of different values for $(\hat{\alpha}_0, \hat{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ by $2N_2+1$. As $Q^{\hat{\alpha}_0,\hat{\beta}_0}(x,y)$ has N_1+1 roots over $\mathbb{P}^1(\overline{\mathbb{K}}) \setminus S$ then, by Lemma 32, there are N_1+1 different $(\bar{\alpha}_0, \bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $Q^{\hat{\alpha}_0,\hat{\beta}_0}(x,y) = Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y)$. Hence, for each $(\hat{\alpha}_0, \hat{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\hat{\alpha}_0, \hat{\beta}_0)$ has multiplicity bigger than one as a root of $Q^{\hat{\alpha}_0,\hat{\beta}_0}(x,y)$, there are N_1+1 points $(\bar{\alpha}_0, \bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\hat{\alpha}_0, \hat{\beta}_0)$ has multiplicity bigger than one as a root of $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y)$. Therefore, the number of different values for $(\bar{\alpha}_0, \bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y)$ has a root in $\mathbb{P}^1(\overline{\mathbb{K}}) \setminus S$ with multiplicity bigger than one is bounded by $(N_1+1)(2N_2+1)$.

Joining these bounds, we deduce that there are at most $(N_2-N_1)(N_2+1)+(N_1+1)(2N_2+1)$ different $(\bar{\alpha}_0, \bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $Q^{\bar{\alpha}_0,\bar{\beta}_0}$ is not square-free. Recalling that $N_1 = D - N_2$ and $N_2 \leq D$ (Proposition 4), we can bound $(N_2-N_1)(N_2+1)+(N_1+1)(2N_2+1)$, by D^2+3D+1. $\square$

We can also relate Proposition 29 to the *Waring locus* of the binary form $f(x,y)$, see Carlini et al. (2017). For example, this proposition improves (Carlini et al., 2017, Proposition 3.8) by a factor of two. Moreover, it shows that the uniqueness condition from (Carlini et al., 2017, Proposition 3.8) misses some assumptions to hold [3].

[3] The authors are not taking into account that the lambdas that they use in their proof are not unique, and so they give us more degrees of freedom that we can use to fix more terms in the decomposition.

4.1.2. Complexity of computing λ

We compute the coefficients λ_j of the decomposition by solving a linear system involving a transposed Vandermonde matrix (Step 3 of Algorithm 3). We follow Kaltofen and Yagati (1989) to write the solution of Equation (5) as the evaluation of a rational function over the roots of a univariate polynomial.

Definition 33. *Given a polynomial $P(x) := \sum_{i=0}^{n} a_i x^i$ and $0 < k \leq n$, let*

$$Quo(P(x), x^k) := \sum_{i=k}^{n} a_i x^{i-k}.$$

Proposition 34 (Kaltofen and Yagati, 1989, Sec. 5)**.** *If $\alpha_j \neq \alpha_i$, for all $i \neq j$, then there is a unique solution to the system of Equation* (14)*.*

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_r \\ \vdots & \vdots & & \vdots \\ \alpha_1^{r-1} & \alpha_2^{r-1} & \cdots & \alpha_r^{r-1} \end{pmatrix} \lambda = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{r-1} \end{pmatrix} \tag{14}$$

Moreover, if the solution is $\lambda = (\lambda_1, \ldots, \lambda_r)^\mathsf{T}$ then, $\lambda_j = \frac{T}{Q'}(\alpha_j)$ where $Q'(x)$ is the derivative of $Q(x) := \prod_{i=1}^{r}(x - \alpha_i)$, $R(x) := \sum_{i=1}^{r} a_{r-i} x^{i-1}$ and $T(x) := Quo\big(Q(x) \cdot R(x), x^r\big)$.

Lemma 35. *Given a binary form $f(x,y) := \sum_{i=0}^{D} \binom{D}{i} a_i x^i y^{D-i}$, let Q be a square-free kernel polynomial of degree r, obtained after step 3 of Algorithm 3. Assume that y does not divide Q. Let α_j be the j-th roots of $Q(x)$, $Q'(x)$ be the derivative of $Q(x)$ and the polynomial $T(x) := Quo\big(Q(x) \cdot R(x), x^r\big)$, with $R(x) := \sum_{i=1}^{r} a_{r-i} x^{i-1}$. Then, each λ_j from step 3 in Algorithm 4 can be written as $\lambda_j = \frac{T}{Q'}(\alpha_j)$.*

Proof. As y does not divide Q, we can write it as $Q(x,y) = \prod_i (x - \alpha_i y)$, where all the α_i are different. Hence, as the $r \times r$ leading principal submatrix of Equation (5) is invertible, we can restrict the problem to solve that $r \times r$ leading principal subsystem. This system is Equation (14). Therefore, the proof follows from Proposition 34. □

Proposition-Definition 36 (Symbolic decomposition)**.** *Let Q be a square-free kernel polynomial related to a minimal decomposition of a binary form f of degree D, such that y does not divide Q. In this case, we can write f as*

$$f(x,y) = \sum_{\{\alpha \in \overline{\mathbb{K}} | Q(\alpha) = 0\}} \frac{T}{Q'}(\alpha) \cdot (\alpha x + y)^D.$$

Remark 37. *If the square-free kernel polynomial related to a decomposition of rank r is divisible by y, we can compute* $\{\lambda_j\}_{j<r}$ *of Equation* (5) *as in Lemma 35, by taking* $\frac{Q}{y}$ *as the kernel polynomial. It is without loss of generality to consider* $Q = P_{(u_0,\ldots,-1,0)^\intercal}$*, because Q is square-free, and so* y^2 *can not divide it. Hence,* $\lambda_r = a_D - \sum_{i=0}^{r-2} u_i a_{D-r+i+1}$ *(Reznick, 2013a, Equation 2.12).*

To summarize the section, given a binary form f of rank r, there is a square-free kernel polynomial Q of the degree r, such that the largest degree of its irreducible factors is bounded by $\min(r, D-r+1)$ (Proposition-Definition 36). If $Q(x,y)$ is not divisible by y, the decomposition is

$$f(x,y) = \sum\nolimits_{\{\alpha \in \overline{\mathbb{K}} | Q(\alpha)=0\}} \frac{T}{Q'}(\alpha) \cdot (\alpha x + y)^D,$$

where T and Q' are polynomials whose coefficients belong to $\mathbb{K}$ and whose degrees are bounded by r, defined in Lemma 35. When y divides Q, the form is similar.

4.1.3. Lower bounds on the algebraic degree

In this section we analyze the tightness of the bound of Theorem 28. To do so, we construct families of examples where the bound is tight. We present two families of examples. In the first one, the decomposition is unique. In the second one, it is not.

Proposition 38 (Heinig and Rost, 1984, Theorem 5.2)**.** *For every pair of relatively prime binary forms,* $\bar{P}_v$ *and* $\bar{P}_w$*, of degrees* $\bar{N}_1+1$ *and* $\bar{N}_2+1$*,* $\bar{N}_1 \leq \bar{N}_2$*, there is a sequence* $a = (a_0, \ldots, a_{\bar{N}_1+\bar{N}_2})$ *such that* $N_1^a = \bar{N}_1$*,* $N_2^a = \bar{N}_2$*, and we can consider the polynomials* $\bar{P}_v$ *and* $\bar{P}_w$ *as the kernel polynomials* P_v *and* P_w *from Proposition 7 with respect to the family of Hankel matrices* $\{H_a^k\}_k$*.*

Corollary 39. *If there is an irreducible binary form of degree* $\bar{N}_1+1$ *in* $\mathbb{K}[x,y]$*, then for every* $D > 2\bar{N}_1$*, there is a binary form f of degree D such that its decomposition is unique, its rank* $\bar{N}_1+1$*, and the degree of the biggest irreducible factor of the polynomial Q from Algorithm 3 in the decomposition is* $\min(\bar{N}_1+1, D-\bar{N}_1) = \bar{N}_1+1$*. That is, the algebraic degree of the minimal decomposition over* $\mathbb{K}$ *is* $\bar{N}_1+1$ *and the bound of Theorem 28 is tight.*

Proof. Let $\bar{P}_v$ be a irreducible binary form of degree $\bar{N}_1+1$. Let $\bar{P}_w$ be any binary form of degree $\bar{N}_2+1 := D-\bar{N}_1+1$ relatively prime with $\bar{P}_v$. Consider the sequence $a = (a_0, \ldots, a_{\bar{N}_1+\bar{N}_2})$ of Proposition 38 with respect to $\bar{P}_v$ and $\bar{P}_w$, and the binary form $f(x,y) := \sum_{i=0}^{D} \binom{D}{i} a_i x^i y^{D-i}$. As $\mathbb{K}$ is of characteristic zero, $\mathbb{K}$ is a perfect field, and so, as $\bar{P}_v$ is irreducible, it is square-free. Then, by Lemma 17, the rank of the decomposition is $N_1^a+1 = \bar{N}_1+1$, and by Corollary 26 the decomposition is unique. Following Algorithm 3, the polynomial Q is equal to $\bar{P}_v$, which is an irreducible polynomial of degree $\bar{N}_1+1$. As $D > 2\bar{N}_1$, then $\min(\bar{N}_1+1, D-\bar{N}_1) = \bar{N}_1+1$ and the bound of Theorem 28 is tight. □

Lemma 40. *Let $\mathbb{K} = \mathbb{Q}$ and $p \in \mathbb{N}$ a prime number. Then, there is a binary form f of degree $2(p-1)$ whose decomposition is not unique and the bound of Theorem 28 is tight.*

Proof. Consider the binary form $f(x,y) := \binom{2(p-1)}{p-1} x^{p-1} y^{p-1}$. Using Algorithm 4, we obtain $P_v = -y^p$ and $P_w = x^p$, $N_1 = N_2 = p-1$. The polynomial P_v is not square-free, so we have to consider a square-free kernel polynomial in $\mathtt{Ker}(H^{N_2+1})$. Moreover, the rank of the decomposition is $N_2+1 = p$. Every kernel polynomial in $\mathtt{Ker}(H^{N_2+1})$ in $\mathbb{Q}[x,y]$ can be written as $\mu_w x^p - \mu_v y^p$ for some $\mu_w, \mu_v \in \mathbb{Q}$. We are interested in the zeros of these polynomials (step 3 of Algorithm 3), thus we consider coprime $\mu_w, \mu_v \in \mathbb{Z}$, as the zeros do not change. As we want to consider square-free kernel polynomials, neither μ_w nor μ_v can be zero, and so $(1,0) \in \mathbb{P}^1(\mathbb{Q})$ is not a root of any of these polynomials. Hence, we rewrite our polynomial as $\frac{1}{\mu_v y^p}(\frac{\mu_w x^p}{\mu_v y^p} - 1)$, and so we look for the factorization over $\mathbb{Q}[z]$ of $\frac{\mu_w}{\mu_v} z^p - 1$, where $z = \frac{x}{y}$. We can use the Newton's polygon criterion, e.g., Cassels (1986, Chapter 6.3), to show that, if $\sqrt[p]{\frac{\mu_w}{\mu_v}} \notin \mathbb{Q}$, then $\frac{\mu_w}{\mu_v} z^p - 1$ is irreducible over $\mathbb{Q}[x,y]$ and so the degree of its biggest irreducible factor is $p > \min(p, 2(p-1)-p+1) = p-1$ (Theorem 28). If this is not the case, then $\sqrt[p]{|\frac{\mu_w}{\mu_v}|} \in \mathbb{Q}$, and so we can factor it as

$$\left(\sqrt[p]{|\frac{\mu_w}{\mu_v}|} \cdot z\right)^p - 1 = \left(\sqrt[p]{|\frac{\mu_w}{\mu_v}|} \cdot z - 1\right) \left(\sum_{i=0}^{p-1} \left(\sqrt[p]{|\frac{\mu_w}{\mu_v}|} \cdot z\right)^i\right).$$

The second factor is irreducible because there is an automorphism in $\mathbb{Q}[x]$ (given by $z \mapsto \sqrt[p]{|\frac{\mu_v}{\mu_w}|} z$) that transforms it into the p-th cyclotomic polynomial, which is irreducible as p is prime. Hence, the biggest irreducible factor of this polynomial has degree $p-1 = \min(p, 2(p-1)-p+1)$ and the bound of Theorem 28 is tight. □

4.2. Arithmetic complexity

Lemma 41 (Complexity of Algorithm 4)**.** *Given a binary form $f = \sum_{i=0}^{D} \binom{D}{i} a_i x^i y^{D-i}$ of degree D, Algorithm 4 computes P_v and P_w in $O(\mathtt{M}(D) \cdot \log(D))$ arithmetic operations.*

Proof. The complexity of the algorithm is the complexity of computing the rows $(i+1)$, i and $(i-1)$ of the Extended Euclidean algorithm between $\sum_{i=0} a_i x^i$ and x^{D+1}, where the i-th row is the first row i such that $\deg(R_i) < \frac{D+1}{2}$ (Lemma 21). This can be done using the *Half-GCD algorithm*, which computes these rows in $O(\mathtt{M}(D)\cdot\log(D))$. For a detailed reference of how this algorithm works see Bostan et al. (2017, Chapter 6.3) or Gathen and Gerhard (2013, Chapter 11). $\square$

Lemma 42 (Complexity of computing Q)**.** *Given the kernel polynomials P_v and P_w from Proposition 7, we compute a square-free polynomial $Q_\mu := P_\mu \cdot P_v + P_w$ such that the maximal degree of its irreducible factors is bounded by Theorem 28 in $O(\mathtt{M}(D)\cdot\log(D))$.*

Proof. To compute the vector μ, we choose randomly $N_2 - N_1 + 1$ linear forms in $\mathbb{K}[x,y]$ and we proceed as in Lemma 25. The complexity bound is due to multi-point evaluation and interpolation of a univariate polynomial (Gathen and Gerhard, 2013, Chapter 10). $\square$

Theorem 43. *When the decomposition is unique, that is when the rank is N_1+1, then Algorithm 3 computes deterministically a symbolic decomposition (Proposition-Definition 36) of a binary form in $O(\mathtt{M}(D)\log(D))$.*

When the decomposition is not unique, that is when the rank is N_2+1 and $N_1 < N_2$, then Algorithm 3 is a Monte Carlo algorithm that computes a symbolic decomposition of a binary form in $O(\mathtt{M}(D)\log(D))$.

Proof. The first step of the algorithm, in both cases, is to compute the kernel polynomials P_v and P_w of Proposition 7 using Algorithm 4. By Lemma 41, we compute them deterministically in $O(\mathtt{M}(D)\cdot\log(D))$.

If P_v is square-free, which means that the decomposition is unique, then $Q = P_v$. Otherwise, in step 2, we need to choose some random values to construct the square-free polynomial Q (from the kernel polynomials P_v and P_w) in $O(\mathtt{M}(D)\cdot\log(D))$ as in Lemma 42. This is the step that makes the algorithm a Monte Carlo one, as we might fail to produce a square-free polynomial Q.

In both cases, at step 3 we compute the rational function that describes the solution of the system in Equation (5) in $O(\mathtt{M}(D)\cdot\log(D))$ (Lemma 35). At step 4 of the algorithm, we return the decomposition. $\square$

We can bound the probability of error of Algorithm 3 using Proposition 29, which bounds the number of bad values that lead us to a non square-free polynomial Q. Moreover, we can introduce a Las Vegas version of Algorithm 3 by checking if the values that we choose to construct a polynomial Q result indeed a square-free polynomial. We can do this check in $O(\mathtt{M}(D)\cdot\log(D))$, by computing the GCD between the Q and its derivatives.

Remark 44. *If we want to output an approximation of the terms of the minimal decomposition, with a relative error of* $2^{-\varepsilon}$*, we can use Pan's algorithm (Pan, 2002) (McNamee and Pan, 2013, Theorem 15.1.1) to approximate the roots of Q. In this case the complexity becomes* $O\big(D\log^2(D)\big(\log^2(D)+\log(\varepsilon)\big)\big)$.

4.3. Bit complexity

Let $f \in \mathbb{Z}[x,y]$ be a binary form as in Equation (1), of degree D and let τ be the maximum bitsize of the coefficients a_i. We study the bit complexity of computing suitable approximations of the α_j's, β_j's, and λ_j's of Equation (3), say $\widetilde{\alpha}_j$, $\widetilde{\beta}_j$ and $\widetilde{\lambda}_j$ respectively, that induce an approximate decomposition correct up to ℓ bits. That is a decomposition such that $\|f-\sum_j \widetilde{\lambda}_j(\widetilde{\alpha}_j x+\widetilde{\beta}_j y)^D\|_\infty \leq 2^{-\ell}$. We need to estimate an upper bound on the number of bits that are necessary to perform all the operations of the algorithm.

The first step of the algorithm is to compute P_v and P_w, via the computation of three rows of the Extended GCD of two polynomials of degree D and $D+1$ with coefficients of maximal sized τ. This can be achieved in $\widetilde{O}_B(D^2\tau)$ bit operations (Gathen and Gerhard, 2013, Corollary 11.14.B), and the maximal bit size of P_v and P_w is $\widetilde{O}(D\tau)$. We check if P_v is a square-free polynomial in $\widetilde{O}(D^2\tau)$, via the computation of the GCD of $P_v(x,1)$ and its derivative (Gathen and Gerhard, 2013, Corollary 11.14.A), and by checking if y^2 divides it.

If P_v is square-free polynomial, then $Q = P_v$. If P_v is not square-free, then we can compute Q by assigning values to the coefficients of P_μ. We assume that y^2 does not divide P_w, if this does not hold, we replace P_w by the kernel polynomial $x^{N_2-N_1}P_v + P_w$, which is coprime to P_v, and so not divisible by y, as P_v and the original P_w are coprime (Proposition 7). We set all the coefficients of P_μ to zero, except the constant term. Then $Q = \mu_0 y^{N_2-N_1}P_v + P_w$. Now we have to choose μ_0 so that Q is square-free. As $y^{N_2-N_1}P_v$ and P_w are coprime, there are at most $2D+2$ forbidden values for μ_0 such that Q is not square-free (Corollary 23), thus at least one of the first $2D+3$ integer fits our requirements. We test them all. Each test corresponds to a GCD computation, that costs $\widetilde{O}_B(D^2\tau)$ and so the overall cost is $\widetilde{O}_B(D^3\tau)$.

Let $\sigma = \widetilde{O}(D\tau)$ be the maximal bit size of Q. By Remark 37, we can assume that y does not divide Q, consider $y=1$ and treat Q as an univariate polynomial.

Let $\{\alpha_j\}_j$ be the roots of Q. We isolate them in $\widetilde{O}_B(D^2\sigma)$ (Pan, 2002). For the (aggregate) separation bound of the roots it holds that $-\lg\prod_j \Delta(\alpha_j) = O(D\sigma + D\lg(D))$. We approximate all the roots up to accuracy $2^{-\ell_1}$ in $\widetilde{O}_B(D^2\sigma + D\ell_1)$ (Pan and Tsigaridas, 2017a). That is, we compute absolute approximations of α_j, say $\widetilde{\alpha}_j$, such that $|\alpha_j - \widetilde{\alpha}_j| \leq 2^{-\ell_1}$.

The next step consists in solving the (transposed) Vandermonde system, $V(\widetilde{\alpha})^{\mathsf{T}}\lambda = a$, where $V(\widetilde{\alpha})$ is the Vandermonde matrix we construct with the roots of Q, λ is a vector contains the coefficients of decomposition, and a is a vector containing the coefficients of F, see also Equation (5). We know the entries of $V(\widetilde{\alpha})$ up to ℓ_1 bits. Therefore, we can compute the elements of the solution vector λ with an absolute approximation correct up to $\ell_2 = \ell_1 - O(D\lg(D)\sigma - \lg\prod_j \Delta(\alpha_j)) = \ell_1 - O(D\lg(D)\sigma)$ bits (Pan and Tsigaridas, 2017b, Theorem 29). That is, we compute $\widetilde{\lambda}_j$'s such that $|\lambda_j - \widetilde{\lambda}_j| \leq 2^{-\ell_2}$. At this point we have obtained the approximate decomposition

$$\widetilde{f}(x,y) := \sum_{j=1}^{r} \widetilde{\lambda}_j (\widetilde{\alpha}_j x + y)^D.$$

To estimate the accuracy of $\widetilde{f}$ we need to expand the approximate decomposition and consider it as a polynomial in x. We do not actually perform this operation; we only estimate the accuracy as if we were. First, we expand each $(\widetilde{\alpha}_j x + y)^D$. This results polynomials with coefficients correct up $\ell_3 = \ell_2 - O(D\sigma) = \ell_1 - O(D\lg(D)\sigma) - O(D\sigma) = \ell_1 - O(D\lg(D)\sigma)$ bits (Pan and Tsigaridas, 2017b, Lemma 19). Next, we multiply each such polynomial with $\widetilde{\lambda}_i$, and we collect the coefficients for the various powers of x. Each coefficient is the sum of $r \leq D$ terms. The last two operations do not affect, asymptotically, the precision. Therefore, the polynomial $\widetilde{f} = \sum_{j=1}^{r} \widetilde{\lambda}_j(\widetilde{\alpha}_j x + (1 - \widetilde{\alpha}_j t)y)^D$ that corresponds to the approximate decomposition has an absolute approximation such that $\|f - \widetilde{f}\| \leq 2^{-\ell_1 + O(D\lg(D)\sigma)}$. To achieve an accuracy of $2^{-\ell}$ in the decomposition, such that $\|f - \widetilde{f}\| \leq 2^{-\ell}$, we should choose $\ell_1 = \ell + O(D\lg(D)\sigma)$. Thus, all the computations should be performed with precision of $\ell + O(D\lg(D)\sigma)$ bits. The bit complexity of computing the decomposition of f up to ℓ bits is dominated by the solving and refining process and it is $\widetilde{O}_B(D\ell + D^2\sigma)$. If we substitute the value for σ, then we arrive at the complexity bound of $\widetilde{O}_B(D\ell + D^4 + D^3\tau)$.

Theorem 45. *Let $f \in \mathbb{Z}[x,y]$ be a homogeneous polynomial of degree D and maximum coefficient bitsize τ. We compute an approximate decomposition of accuracy $2^{-\ell}$ in $\widetilde{O}_B(D\ell + D^4 + D^3\tau)$ bit operations.*

Acknowledgments

The authors thank George Labahn and Vincent Neiger for pointing out important references to derandomize the computation of P_v and P_w. The authors thank the anonymous reviewers for their helpful comments. Matías Bender thanks Joos Heintz for supervising his Master's thesis. The authors are partially supported by ANR JCJC GALOP (ANR-17-CE40-0009), the PGMO grant ALMA and the PHC GRAPE.

References

Ansola, M., Daz-Cano, A., Zurro, M. A., Jun. 2017. Semialgebraic decomposition of real binary forms of a given degree's space. arXiv:1706.04207 [math]ArXiv: 1706.04207.
URL `http://arxiv.org/abs/1706.04207`

Bajaj, C., 1988. The algebraic degree of geometric optimization problems. Discrete & Computational Geometry 3 (1), 177–191.

Bender, M. R., Faugère, J.-C., Perret, L., Tsigaridas, E., 2016. A superfast randomized algorithm to decompose binary forms. In: Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation. ACM, pp. 79–86.

Bernardi, A., Daleo, N. S., Hauenstein, J. D., Mourrain, B., Dec. 2017. Tensor decomposition and homotopy continuation. Differential Geometry and its Applications 55, 78–105.
URL `http://www.sciencedirect.com/science/article/pii/S0926224517301055`

Bernardi, A., Gimigliano, A., Ida, M., 2011. Computing symmetric rank for symmetric tensors. Journal of Symbolic Computation 46 (1), 34–53.

Blekherman, G., 2015. Typical real ranks of binary forms. Foundations of Computational Mathematics 15 (3), 793–798.

Boij, M., Carlini, E., Geramita, A., 2011. Monomials as sums of powers: the real binary case. Proceedings of the American Mathematical Society 139 (9), 3039–3043.

Bostan, A., Chyzak, F., Giusti, M., Lebreton, R., Lecerf, G., Salvy, B., Schost, É., 2017. Algorithmes efficaces en calcul formel. published by the Authors.

Brachat, J., Comon, P., Mourrain, B., Tsigaridas, E., 2010. Symmetric tensor decomposition. Linear Algebra and its Applications 433 (11), 1851–1872.

Cabay, S., Choi, D.-K., 1986. Algebraic computations of scaled Padé fractions. SIAM Journal on Computing 15 (1), 243–270.

Carlini, E., Catalisano, M. V., Chiantini, L., Geramita, A. V., Woo, Y., 2018. Symmetric tensors: rank, Strassen's conjecture and e-computability. Annali della Scuola Normale Superiore di Pisa. Classe di scienze 18 (1), 363–390.
URL `https://dialnet.unirioja.es/servlet/articulo?codigo=6389435`

Carlini, E., Catalisano, M. V., Oneto, A., Jul. 2017. Waring loci and the Strassen conjecture. Advances in Mathematics 314, 630–662.
URL `http://www.sciencedirect.com/science/article/pii/S0001870816305758`

Cassels, J. W. S., 1986. Local fields. Vol. 3. Cambridge University Press Cambridge.

Comas, G., Seiguer, M., 2011. On the rank of a binary form. Foundations of Computational Mathematics 11 (1), 65–78.

Comon, P., 2014. Tensors: a brief introduction. IEEE Signal Processing Magazine 31 (3), 44–53.

Comon, P., Golub, G., Lim, L.-H., Mourrain, B., 2008. Symmetric tensors and symmetric tensor rank. SIAM Journal on Matrix Analysis and Applications 30 (3), 1254–1279.

Comon, P., Mourrain, B., 1996. Decomposition of quantics in sums of powers of linear forms. Signal Processing 53 (2), 93–107.

Draisma, J., Horobeţ, E., Ottaviani, G., Sturmfels, B., Thomas, R. R., 2016. The euclidean distance degree of an algebraic variety. Foundations of computational mathematics 16 (1), 99–149.

Dür, A., Oct 1989. On computing the canonical form for a binary form of odd degree. Journal of Symbolic Computation 8 (4), 327–333.

García-Marco, I., Koiran, P., Pecatte, T., 2017. Reconstruction algorithms for sums of affine powers. In: Proceedings of the 2017 ACM on International Symposium on Symbolic and Algebraic Computation. ISSAC '17. ACM, New York, NY, USA, pp. 317–324.

Gathen, J. v. z., Gerhard, J., 2013. Modern computer algebra. Cambridge University Press, Cambridge.

Giesbrecht, M., Kaltofen, E., Lee, W.-s., 2003. Algorithms for computing sparsest shifts of polynomials in power, chebyshev, and pochhammer bases. Journal of Symbolic Computation 36 (3-4), 401–424.

Giesbrecht, M., Roche, D. S., 2010. Interpolation of shifted-lacunary polynomials. Computational Complexity 19 (3), 333–354.

Gundelfinger, S., 1887. Zur theorie der binären formen. Journal für die reine und angewandte Mathematik 100, 413–424.

Hauenstein, J. D., Oeding, L., Ottaviani, G., Sommese, A. J., 2016. Homotopy techniques for tensor decomposition and perfect identifiability. Journal fr die reine und angewandte Mathematik (Crelles Journal) 0 (0).
URL `https://www.degruyter.com/view/j/crll.ahead-of-print/crelle-2016-0067/crelle-2016-0067.xml`

Heinig, G., Rost, K., 1984. Algebraic methods for Toeplitz-like matrices and operators. Springer.

Helmke, U., 1992. Waring's problem for binary forms. Journal of pure and applied algebra 80 (1), 29–45.

Iarrobino, A., Kanev, V., 1999. Power sums, Gorenstein algebras, and determinantal loci. Springer.

Kaltofen, E., Yagati, L., 1989. Improved sparse multivariate polynomial interpolation algorithms. In: Symbolic and Algebraic Computation. Springer, pp. 467–474.

Kung, J. P., 1990. Canonical forms of binary forms: variations on a theme of Sylvester. Institute for Mathematics and Its Applications 19, 46.

Kung, J. P., Rota, G.-C., 1984. The invariant theory of binary forms. Bulletin of the American Mathematical Society 10 (1), 27–85.

Landsberg, J. M., 2012. Tensors: geometry and applications. American Mathematical Society.

McNamee, J. M., Pan, V. Y., 2013. Numerical methods for roots of polynomials (II). Elsevier.

Nie, J., Ranestad, K., Sturmfels, B., 2010. The algebraic degree of semidefinite programming. Mathematical Programming 122 (2), 379–405.

Oeding, L., Ottaviani, G., 2013. Eigenvectors of tensors and algorithms for waring decomposition. Journal of Symbolic Computation 54, 9–35.

Pan, V., 2001. Structured matrices and polynomials: unified superfast algorithms. Springer.

Pan, V. Y., 2002. Univariate polynomials: Nearly optimal algorithms for numerical factorization and root-finding. Journal of Symbolic Computation 33 (5), 701 – 733.

Pan, V. Y., Tsigaridas, E., 2017a. Accelerated approximation of the complex roots and factors of a univariate polynomial. Theoretical Computer Science 681, 138–145.

Pan, V. Y., Tsigaridas, E. P., 2017b. Nearly optimal computations with structured matrices. Theoretical Computer Science 681, 117–137.

Reznick, B., 1996. Homogeneous polynomial solutions to constant coefficient pde's. Advances in Mathematics 117 (2), 179–192.

Reznick, B., 2013a. On the length of binary forms. In: Quadratic and Higher Degree Forms. Springer, pp. 207–232.

Reznick, B., 2013b. Some new canonical forms for polynomials. Pacific Journal of Mathematics 266 (1), 185–220.

Reznick, B., Tokcan, N., 2017. Binary forms with three different relative ranks. Proceedings of the American Mathematical Society 145 (12), 5169–5177.

Reznick, B. A., 1992. Sums of even powers of real linear forms. Vol. 463. American Mathematical Society.

Sylvester, J. J., 1851. On a remarkable discovery in the theory of canonical forms and of hyperdeterminants. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2 (12), 391–410.

Sylvester, J. J., 1904. An essay on canonical forms, supplement to a sketch of a memoir on elimination, transformation and canonical forms. In: The collected papers of James Joseph Sylvester. Vol. 1. Cambridge University Press, pp. 203–216, the original paper dates back to 1851.