Universidad de Buenos Aires

Facultad de Ciencias Exactas y Naturales

Departamento de Computación

A Superfast Algorithm for the Decomposition of Binary Forms

Tesis presentada para optar al título de
Licenciado en Ciencias de la Computación

Matías Rafael Bender

Director: Joos Heintz (UBA-CONICET)

Codirector: Jean-Charles Faugère (INRIA)

Otros: Elias Tsigaridas (INRIA)

Ludovic Perret (UPMC)

Buenos Aires, 2015

UN ALGORITMO SUPERFAST PARA DESCOMPONER FORMAS BINARIAS

Descomponer una *Forma Binaria* consiste en reescribir un polinomio homogéneo en dos variables de grado D como una combinación lineal de D-esimas potencias de factores lineales. En este trabajo nos concentraremos en las combinaciones lineales con la mínima cantidad posible de sumandos, valor conocido como el *Rango* de la forma binaria. Nuestro problema es equivalente al de la *Descomposición de Tensores Simétricos* cuando el tensor simétrico tiene dimensión 2.

En esta tesis proponemos un algoritmo para la descomposición de formas binarias, el cual se basa en el trabajo de Sylvester del siglo XIX. Retomamos su aporte utilizando técnicas del *Álgebra Lineal* y resultados sobre *Secuencias Linealmente Recurrentes*. De esta manera ofrecemos un nuevo enfoque para la descomposición de formas binarias con una complejidad aritmética *cuasi-lineal* en el grado de la forma dada, óptima si no consideramos los factores poli-logarítmicos. La descomposición involucra números algebraicos sobre el cuerpo original, por lo que demostramos una cota superior para el grado de la extensión algebraica necesaria, la cual es $\texttt{Min}(rango; D - rango + 1)$.

Palabras claves: Formas Binarias, Descomposición de Tensores, Rango Tensorial, Algoritmos Superfast, Matrices de Hankel.

A SUPERFAST ALGORITHM FOR THE DECOMPOSITION OF BINARY FORMS

To decompose a *Binary Form* we write an homogeneous polynomial on two variables and degree D as a linear combination of D-powers of linear forms. In this work we focus on the smallest possible number of summands in the linear combination, a quantity known as *Rank*. Our problem is equivalent to the *Symmetric Tensor Decomposition* problem when the symmetric tensor has dimension 2.

In this thesis we focus on an algorithm for the decomposition of binary forms, which relies on the work from Sylvester in the 19th century. We revisit this work using *linear algebra* techniques and results from *linear recurrent sequences*. We propose a new approach for the decomposition of binary forms with *soft linear* arithmetic complexity in the degree of the given form, and hence optimal, up to poly-logarithmic factors. The solution of the decomposition problem requires to deal with algebraic numbers over the ground field whose degree we surprisingly succeed to bound by $\texttt{Min}(rank; D - rank + 1)$.

Keywords: Binary Form, Tensor Decomposition, Tensor Rank, Superfast Algorithm, Hankel Matrix.

DISCLAIMER

Due to the thesis regulation, we can not include in the front page all the people involved in this work. Still we want to make clear that, besides Jean-Charles, Joos and Matías, Elias Tsigaridas (INRIA) and Ludovic Perret (UPMC) were involved in this thesis, which was partly made during Matías' internship at LIP6, funded by INRIA.

AGRADECIMIENTOS

Escribo mis agradecimientos con felicidad y un poco de miedo. Feliz de poder finalizar mi licenciatura que luego de tanto esfuerzo, sin duda, valió la pena. A lo largo de los años fueron muchísimas las personas que estuvieron a mi lado, me ayudaron, me enseñaron, me guiaron, me contuvieron y por eso temo no poder hacerles justicia en este corto texto, asi como también temo por los errores de cohesión en él.

Los primeros agradecimientos quiero dedicárselos a mi familia, con cuyas enseñanzas y apoyo llegué hasta acá y espero llegar más lejos. Mis padres, Javier y Mónica, son los responsables de todo esto. Fue su trabajo y esfuerzo constante lo que inspiró el mío. Son las personas a las que admiro y a las que intento parecerme. Junto a ellos, mis abuelos Clara, Israel y María, y mis tíos Ale (x2), Gardo, Gisi y Myriam, me criaron. Si alguien merece sentirse orgulloso de este logro, son ellos. Sin embargo, fue la compañía de mis hermanos Ariel y Tami la que determinó quién soy. Ellos transitaron un sendero semejante al mío y juntos llegamos a este momento, con peleas y encontronazos, pero con un amor incondicional e invaluable. De ambos estoy muy orgulloso.

Luego quiero agradecerle a Joos. Al volver a la Argentina me enfrentaba con el desafío de graduarme rápidamente y para eso debía presentar mi tesis. Joos, de manera totalmente desinteresada, aceptó ser mi director y me ayudó a lograr mi objetivo. Si bien sus enseñanzas y ayuda académica fueron fundamentales, lo que más valoro fueron las charlas que hemos tenido sobre los asuntos más diversos (política, historia, filosofía y religión).

Ici, je parlerai de mon expérience à Paris mais comme mon français est mauvais, j'écrirai en anglais. Désolé. I arrived at the PolSys team mostly by chance. When I signed up for my internship there I had no idea what to expect. A year later, and about to start my PhD as part of the team, I am grateful and glad to say that I had an amazing working experience and a spectacular guidance throughout the whole stay. I want to specially thank Elias Tsigaridas, Jean-Charles Faugère and Ludovic Perret for their hospitality, patience, "les bieres", and their invaluable teachings and help, not only in what we've done but also in the things we'll surely accomplish over the coming years.

La educación pública de Argentina me dio mis mejores herramientas gracias a la labor y al esfuerzo de los docentes y co-docentes (mal llamados no-docentes). Estos son los que con el viento en contra, condiciones precarias y trabajos mal pagos, se ponen al hombro esta tarea tan fundamental que es la de defender y hacer realidad la educación de calidad pública, laica y gratuita. Son muchas las personas que me formaron en estos años, pero quiero destacar particularmente a Esteban Feuerstein y Fernando Schapachnik, de quienes he aprendido infinidad de cosas y a quienes tomo como ejemplo del profesional en el que espero convertirme.

Al referirme a la defensa de la educación pública, no puedo dejar de nombrar a mis

compañeros de [En Acción] y de El Transformador. Junto a ellos asumimos la difícil tarea de luchar por una educación y un sistema científico público de verdad. No quiero explayarme más sobre lo que pienso de nuestro objetivo para no sonar bobamente romántico, sin embargo, quiero reconocer a mis compañeros que supieron transformarse en amigos. Por si algún servicio lee esto, ellos son: Dami, Gre, Jesi, Jota, Juancito, Julito, Lean, Lu, Manu (Chaky), Marian, Nico, Pablito, Pato, Pedrito, Pedro (Chiva), Rodri, Tommy y Zippo. También una mención especial merecen el resto de mis compañeros del Partido Obrero. Parafraseando a Marx, los trabajadores no tenemos nada más que perder que nuestras cadenas, en cambio, ¡Tenemos un mundo que ganar!

El año 2014 fue el más peculiar de mi vida, sin dudas. Mis andanzas lejos de casa me llevaron a conocer nuevas personas, lugares y experiencias. Con orgullo puedo decir que son muchísimas las personas a las que quiero agradecer, pero me conformaré con algunos pocos. Primero que nada debo agradecerle a mis coinquilinos madrilenses. Fue mi primera experiencia viviendo "con extraños" y, más allá de algunos desencuentros, fue excelente. Realmente fue un placer vivir con ellos y de cada uno puedo decir que aprendí un buen par de cosas. Luego quiero agradecerle a mis compañeros de AEC Malasaña, con quienes realmente entendí el sentido y el valor del internacionalismo. Con ellos tuve discusiones, aunque acaloradas, muy fructíferas que me cambiaron la forma de ver algunas cosas. Siguiendo, debo agradecerle a Yaiza, quien también supo ocupar un lugar muy importante durante mi estadía en Madrid. En tierras galas conocí a mi querida banda del Cono Sur. Nuestras aventuras y la pasión exacerbada por la comida y el alcohol me hicieron sentir en casa del otro lado del charco. De todas las personas que conocí, hay particularmente dos en quienes quiero hacer hincapié, Marce y Meli. Ambos fueron mi sustento emocional y mis confidentes durante esta odisea, sin ellos hubiera sido imposible sobrevivirla. Hoy en día son dos personas indispensables en mi vida que, aunque no veo tanto como me gustaría, sin duda voy a echar mucho en falta.

Indudablemente mi paso por Exactas fue un punto de inflexión y la etapa más linda que transité. Encontré mi lugar en el mundo, conocí muchísimas personas, aprendí en demasía y me hice de amistades invaluables. Es mucha la gente, lamento si olvido a alguien. Creo que mi primer grupo de amigos en la facu fue el que integramos con Javo, JP, Lucho, Manu, Mirko, Nacho y Anita. Con poca gente me reí tanto en la vida. Nos hemos matado haciendo TPs o jugando juegos de mesa, hemos inventado los "Casos de Abuso", hemos comido cantidades obscenas de pizza y hemos pasado horas tomando mate y filosofando sobre la vida. Me han salvado las papas del fuego muchísimas veces y sé que lo seguirán haciendo. Quiero mencionar también a Euge y a Fabri, posiblemente las únicas personas que conozco que son tan humildes como inteligentes. Junto a ellos tuve el honor de integrar el equipo BBB. No se si tuvimos los mejores resultados, pero esta experiencia fue sumamente interesante y entretenida.

Durante mi cursada de Algo 2 tuve la suerte de conocer a un grupo de excelentes personas: Mauro, Nico, Pablo, Santi, Tincher y Juan Ma. No estoy seguro de haber cursado más materias con ellos, pero forjamos una amistad muy especial. Aunque esporádicas, nuestras reuniones son algo que aprecio muchísimo. Son gente de fierro que siempre está cuando la necesito. Si hay caras que vi durante muchas cursadas fueron las de Soifer, Ale, Uri, Leo, Vale, Manolo, Kuja, Iván, Pablo y Gabi. Con el tiempo aprendimos a querernos y puedo decir que termino la carrera con un gran grupo de amigos. Falta mencionar a los muchachos

del Óvalo, Juli, Gasti y Marian, con los que, birra y pizza de por medio, nos dimos a la tarea de pensar un DC mejor. Quedó mucho por hacer, pero creo que dejamos nuestra marca.

No quería dejar de nombrar a todos esos amigos de viejas épocas que sobreviven al día de hoy. Son de grupos muy diversos como las historias que nos unen. A algunos no los veo tanto como quisiera, pero "we'll always have Paris". Estos son: Mori, Pani, Petro, Singer, Vain, Zela, Diuk, Huevo, Leo, Menta, Sebi, Ari, Marcos, Marto y Santi.

Por último quiero agradecerle al lector que, si se tomó el trabajo de llegar hasta acá, tuvo la oportunidad de conocer un poco a las personas espectaculares que me hicieron quien soy.

A mis abuelos, Pa, Ma y Ma Clara

A mis viejos, Mónica y Javier

A mis hermanos, Tami y Ari

CONTENTS

1. INTRODUCTION

"El futuro es nuestro, por prepotencia de trabajo."
– Roberto Arlt

In this work we introduce a new algorithm for the decomposition of binary forms (homogeneous polynomials with two variables). Given a binary form $f(x,y) = \sum_{i=0}^{D} a_i x^i y^{D-i}$, with $a_i \in \mathbb{F}$ and $\mathbb{F}$ some field, finding a decomposition means get
$\lambda_1, \ldots, \lambda_r, \alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_r \in \overline{\mathbb{F}}$, with $\overline{\mathbb{F}}$ the algebraic closure of $\mathbb{F}$, such that

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

We are interested in getting the minimal r such that a decomposition with r summands exists. We call this value the **rank** of the binary form and we say that a decomposition is minimal if it has as many summands as its rank.

The problem we are considering is a special case of "Symmetric Decomposition Problem". A symmetric tensor of dimension n and order D, whose coefficients belong to a field, can always be decomposed as a sum of rank-1 symmetric tensors. As in our problem, the minimal quantity such that a decomposition exists is known as **rank** of the tensor. [1]

Finding a minimal decomposition is one of the fundamental problems in the theory of the symmetric tensors. It is a very important issue and particular cases had been intensively studied. For example, for symmetric matrices, that is for tensors of order 2, the decomposition problem is equivalent to the Singular Values Decomposition. Thus tensor decomposition could be seen as an extension of SVD to higher order tensors. Under different formulations, this problem can be found in many different areas. For example, in Statistics it appears with the use of cumulants. In the Blind Source Separation (BSS) problem appears when we assume that the source mixture is linear, [9]. In Data Analysis it can be found in Independent Component Analysis, [16]. It also appears in Electrical Engineering, for example in problems in the Antenna Array Processing [21]. Much more applications appear in the survey of Comon [8].

There is a isomorphism between the symmetric tensors of dimension n and order D and the homogeneous polynomials in $\mathbb{F}[x_1, \ldots, x_n]$ of degree D, which allow us to work directly with homogeneous polynomials. For further details about this relationship we refer the reader

[1] Some authors (e.g. Comon et al. [11]) make a distinction between the **rank** and the **symmetric rank**. In the bibliography, when no distinction is made, the rank is understood as what those authors call the symmetric rank. This work is not the exception as we just refer to the rank.

to Comon et al. [11]. It is interesting to note that the formulation that involves polynomials can be thought as a kind of Waring's problem. Because of its multiples formulations, different authors worked in this problem at the same time, without being aware of the previous results, so many of them were rediscovered through the years.

Coming back to the decomposition of binary forms, if we $\mathbb{F} = \mathbb{C}$, this problem was mathematically solved by Sylvester in 1851 [23] when he proved the necessary and sufficient conditions for a decomposition to exist. This idea leaded straightforward to the algorithm of Comon and Mourrain [10], which, as far as we know, is the first algorithm for getting a minimal decomposition.

This last algorithm can be improved if we observe that the rank has just two possible values. This last thing was rediscovered through the years. The first proof, as far as we know, comes from the Control Theory field and is thanks to Helmke [14]. After that it was proved using analysis of secant varieties and it appears in the works of Comas and Seiguer [7], Comon et al. [11], Bernardi et al. [3].

What all those approaches have in common is the use of Hankel matrices (see Section 2.1). Over the years, many authors, as Iohvidov [15] and Heinig and Rost [13], worked with this special kind of matrices, and nowadays they are deeply understood. Also in the algorithmic world those matrices have been deeply studied, and there are many superfast algorithms (whose arithmetic complexity is almost linear in the size of the generator vector) induced by the analysis of their displacement rank, [4].

In this work we rediscover important properties about the rank by using a new approach related with linear algebra. Based on the properties of the kernel of the Hankel matrices, we deduce a new superfast algorithm to get the rank and we extend it to get a minimal decomposition efficiently. Unlike the previous works, we prove the arithmetic complexity of our algorithm which is almost linear on the degree of the binary form. As far as we know, this is the first superfast algorithm known for this problem. We give an algorithm which does not compute numerically the solution, it just compute an efficient expression of it.

Is important to note that when $\mathbb{F}$ is not algebraically closed, we have a very important difference with the classical formulation of the problem. This difference comes from that we allow the decomposition to have elements in the closure of the original field, and the classical definition asks all the coefficients in the decomposition to be of the same field. An important work about decompositions over the same field is the one by Reznick [20]. In the particular case of $\mathbb{F} = \mathbb{R}$, Helmke [14] characterized all the possible decompositions and the necessary conditions for them to exist. About this distinction we have to remark two things. When the field is algebraically closed (e.g. $\mathbb{F} = \mathbb{C}$), our results are valid for the classical definition of the problem. When it is not, our rank is a lower bound for the classical definition of rank.

For the cases when the $\mathbb{F}$ is not algebraically closed, we show that we can give a minimal decomposition over some extension field, whose algebraic degree we bound. Moreover, we express the solution as the addition of a rational function evaluated over all the roots of a polynomial, where both functions have all their coefficients in the original field. We do not assume that the characteristic is zero, but still we need it to be "big enough".

The paper is organized as follows. In Chapter 2 we introduce the notation of the paper and some results that we use. In Chapter 3 we present the main algorithm and we prove its correctness. In the following sections we explain the details of the main algorithm and prove its complexity. In Chapter 4, we show how to compute efficiently the kernels of the matrices of Equation (2.3). After, in Chapter 5 we bound the algebraic degree of the problem. Following, in Chapter 6, we talk about how to solve linear systems related with transpose of Vandermonde matrices. In Chapter 7, we sum up the results and analyze the arithmetic complexity of the main algorithm. In Chapter 8 we discuss briefly the relation between our results and the ones from Helmke [14] and Comas and Seiguer [7], giving some new proofs. Finally, in Chapter 9 we have a little discussion about the decomposition of general symmetric tensors.

2. PRELIMINARIES

In this section we introduce the notation of the paper and some results that we use. In Section 2.1 we introduce our notation for the binary forms. Following, in Section 2.2, we define what we understand by a decomposition of a binary form. In Section 2.3, we introduce Sylvester's Theorem, which is the basis of our analysis. Latter, in Section 2.4 we introduce some notation for the Hankel matrices and the theorems that we will use. In Section 2.5 we present our notation for the Linear Recurrence Sequence and we recall the arithmetic complexities of the associated problems. Finally, in Section 2.6, we come back to the binary forms to introduce the changes of coordinates.

In the following we refer to $\mathbb{F}$ as an arbitrary field and to $\overline{\mathbb{F}}$ as the algebraic closure of $\mathbb{F}$.

2.1 Binary Forms

In this section we recall some definitions related to the binary forms. Particularly we mention their relationship with the univariate polynomials. Our aim is to extended the definition of being square-free to the binary forms.

Definition 2.1.1. A binary form f of degree D is an homogeneous polynomial in $\mathbb{F}[x, y]$ that can be written as

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

Notation 2.1.2. We call $\mathbb{F}[x, y]_D$ to the set of all the binary forms in $\mathbb{F}[x, y]$ of degree D.

Always it is possible to write a binary form as a product of linear forms.

Proposition 2.1.3. *Given a binary form $f \in \mathbb{F}[x, y]$ of degree D, it can be expressed as*

$$f(x, y) = \prod_{j=1}^{D} (\beta_j x - \alpha_j y)$$

*Where $(\beta_j x - \alpha_j y) \in \overline{\mathbb{F}}[x, y]$. We say that this expression is a **Factorization** of f.*

As we claim, these polynomials are deeply related to the univariate polynomials. In fact, we can rewrite them as a product between the univariate polynomial $f(x, 1)$, composed with $\frac{x}{y}$, and y^D. The actual relation between the binary forms and the univariate polynomials is

that first ones are the homogeneous projection of the second ones. In a few words, the points were the evaluation of the binary form f is zero, belongs to a finite set of lines described by each factor in the factorization of f. If we take the direction of those lines, we observe that they are the homogeneous coordinates of the roots of the associated polynomial $f(x,1)$, and its value at infinite. This allow us to talk about roots of a univariate polynomial and to think the binary forms as their projection. This way, each time that we refer to a **root**, we are talking about the direction of a line where f is zero. With that on mind, we extend the definition of **square-free polynomial**.

Notation 2.1.4. A binary form f is said to be **square-free** when all the linear factors of the factorization of f, taken pairwise, are not multiples.

Using the above observations, it is possible to check if a binary form f has square-roots using the Euclidean Algorithm. The superfast implementation of that algorithm takes $O(\mathtt{M}(D)\cdot\log(D))$ ops, [12, Section 11.1].

2.2 Decomposition of a binary form

As we explained in the introduction, the main objective for this work is to find a decomposition for any binary form. In this section we introduce what we understand by a decomposition for a binary form, and the difference between our definition and the classical one.

First, let us begin with a fundamental theorem which proves that a decomposition always exists.

Theorem 2.2.1 ([20, Theorem 4.2])**.** *Any set $\{(\alpha_j x+\beta_j y)^D : 0\le j\le D\}$, with $\alpha_j,\beta_j\in\mathbb{F}$ of pairwise distinct D-th powers is linearly independent and spans the binary forms of degree D with coefficients in $\mathbb{F}$.*

Proof. The matrix of this set with respect to the basis $\binom{D}{i}x^iy^{D-i}$ is $\left[\alpha_j^i\beta_j^{D-i}\right]_{i,j}$, whose determinant is Vandermonde:

$$\prod_{0\le j<k\le D}\begin{vmatrix}\alpha_k & \beta_k\\ \alpha_j & \beta_j\end{vmatrix}$$

This determinant is a product of non-zero terms by hypothesis.

□

The Theorem 2.2.1 proves that for any binary form f of degree D we can find a finite set of binary forms $\{(\alpha_j x+\beta_j y)^D : 1\le j\le r\}$ and constants $\lambda_1,\dots,\lambda_r$ such that,

$$f(x,y) = \sum_{i=1}^{r} \lambda_j (\alpha_j x + \beta_j y)^D \tag{2.1}$$

Definition 2.2.2 (Decomposition for a Binary Form)**.** A **decomposition for a binary form** $f \in \mathbb{F}[x,y]_D$ is a set $\{(\alpha_j x + \beta_j y)^D : 1 \leq j \leq r\} \subset \overline{\mathbb{F}}[x,y]_D$ and constants $\lambda_1, \ldots, \lambda_r \in \overline{\mathbb{F}}$ such that Equation (2.1) holds.

Observation 2.2.3. In most of the texts the definition of decomposition is different. In all of them, for a decomposition is necessary a set $\{(\alpha_j x + \beta_j y)^D : 1 \leq j \leq r\} \subset \mathbb{F}[x,y]_D$ and constants $\lambda_1, \ldots, \lambda_r \in \mathbb{F}$. Note that the difference is that in our definition we have a "relaxed condition", we work with decompositions over the algebraic closure and not over the original field. As many authors work over $\mathbb{C}$, this distinction is not necessary, and all the results of this work apply. However, when the field is not algebraically closed, it is mandatory to make this distinction. In Chapter 6 we show that the terms involved in a decomposition belong to extension of the field, a subfield of the closure of the field, and we prove a bound for the degree of the field extension needed.

It is important to note that, given a binary form of degree D, there is a decomposition such that the amount of summands is minimal with respect to all other possible decompositions. For any $f \in \mathbb{F}[x,y]_D$, this minimal amount of summands is upper bounded by $D+1$, by Theorem 2.2.1. However, for each f the minimal amount can be different. Consider, for example, x^D and $x^D + y^D$ in $\mathbb{C}[x,y]_D$. It is clear that, for the first polynomial, the minimal amount is 1 and for the second one, it is 2.

Definition 2.2.4. Given $f \in \mathbb{F}[x,y]_D$, the **rank** of f is the minimal r such that there is a decomposition for f that involves r summands.

Observation 2.2.5. Again, here is necessary to make a statement. Our definition of the rank differs with the classical one because the terms in the decomposition are not the same. In particular, our rank is a lower bound for what is usually called the rank of a binary form. Once more, if we work over the complex field, this distinction is not necessary.

This way, when we refer to a **minimal decomposition** of a binary form, we talk about a decomposition of such polynomial where the number of summands is its rank.

2.3 Sylvester's Theorem

Our algorithm can be considered a corollary of the 1851 Sylvester's Theorem [23]. This theorem gives the necessary and sufficient conditions for a binary form to have a decomposition over its algebraic closure.

Theorem 2.3.1 (Sylvester, 1851)**.** *Let*

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

with $a_i \in \mathbb{F} \subseteq \mathbb{C}$. *Also, let*

$$Q(x,y) = \sum_{i=0}^{r} c_i x^i y^{r-i} = \prod_{j=1}^{r} (\beta_j x - \alpha_j y) \tag{2.2}$$

be a square-free polynomial. There are $\lambda_j \in \overline{\mathbb{F}}$ *such that*

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

if and only if,

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_r \\ a_1 & a_2 & \cdots & a_{r+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{D-r} & a_{D-r+1} & \cdots & a_D \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_r \end{pmatrix} = 0 \tag{2.3}$$

Where $\overline{\mathbb{F}}$ *is the algebraic closure of* $\mathbb{F}$.

For a proof of the theorem when $\mathbb{F} = \mathbb{C}$ we refer to Reznick [20, Theorem 2.1]. For an arbitrary $\mathbb{F}$, we consider the same proof, knowing that the ring $\mathbb{F}[X]$ is an Euclidean Domain. As a unique partial fraction decomposition always exist for the quotient ring $\mathbb{F}(X)$, [6, Section 3], the proof over $\mathbb{C}$ can be easily adapted.

We introduce the notation we use through the text to manipulate the matrices of Equation (2.3). These matrices are known as Hankel matrices.

Definition 2.3.2. Given a vector $a = (a_0, \ldots, a_D)$, let $\{H_a^k\}_{1 \le k \le D}$ be the family of Hankel matrices indexed by k, such that $H_a^k \in \mathbb{F}^{(D-k+1)\times(k+1)}$ and

$$H_a^k = \begin{pmatrix} a_0 & a_1 & \cdots & a_{k-1} & a_k \\ a_1 & a_2 & \cdots & a_k & a_{k+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{D-k-1} & a_{D-k} & \cdots & a_{D-2} & a_{D-1} \\ a_{D-k} & a_{D-k+1} & \cdots & a_{D-1} & a_D \end{pmatrix} \tag{2.4}$$

We refer indifferently to the family of Hankel matrices of a vector $(a_0, \ldots, a_D)$ and to the family of Hankel matrices of a binary form $\sum_{i=0}^{D} \binom{D}{i} a_j x^i y^{D-i}$. When it is clear from the context, we skip the subindex.

The binary forms whose coefficients belong to the kernel of the matrices H^i play an important roll in Sylvester's Theorem. For that reason, we call such polynomials **kernel polynomials**. To relate the vector $u = (u_0, \ldots, u_k)$ in the kernel of H^k, with the polynomial with those coefficients, we define the following,

Definition 2.3.3. Given a vector $u = (u_0, \ldots, u_k)$, we define P_u as

$$P_u := \sum_{i=0}^{k} u_i x^i y^{k-i}$$

Notation 2.3.4. A binary form $G(x, y)$ of degree k is said to be a **kernel polynomial** of a family $\{H_G^i\}_{i \leq D}$ if there is a vector $g \in \texttt{Ker}(H^k)$ such that $P_g = G$.

The next corollary summarizes the relationship between a minimal decomposition and Sylvester's Theorem.

Corollary 2.3.5. *Given an binary form $f = \sum_{j=1}^{r} \lambda_j (\alpha_j x + \beta_j y)^D$, its rank is r if and only if there is a non-zero vector u in the kernel of H_f^r such that,*

- $P_u = \prod_{j=1}^{r} (\beta_j x - \alpha_j y)$
- *P_u is a square-free kernel polynomial.*
- *For $1 \leq k < r$, for all non-zero $\hat{u} \in ker(H_f^k)$, the polynomial $P_{\hat{u}}$ is not square-free.*

2.4 Kernel of a Hankel matrix

To compute the minimal decomposition, we find the minimum r such that the Equation (2.3) holds. This approach demands a better inside into the family of matrices of such equation. In this section we characterize the kernels of the Hankel matrices. All the following results can be found in [13, Section 5].

Definition 2.4.1. A **Hankel matrix** is a matrix $\{\{a_{i,j}\}\}$ with constant skew-diagonals (positive sloping diagonals). That means, $(\forall i, j)\ a_{i,j} = a_{(i-1),(j+1)}$.

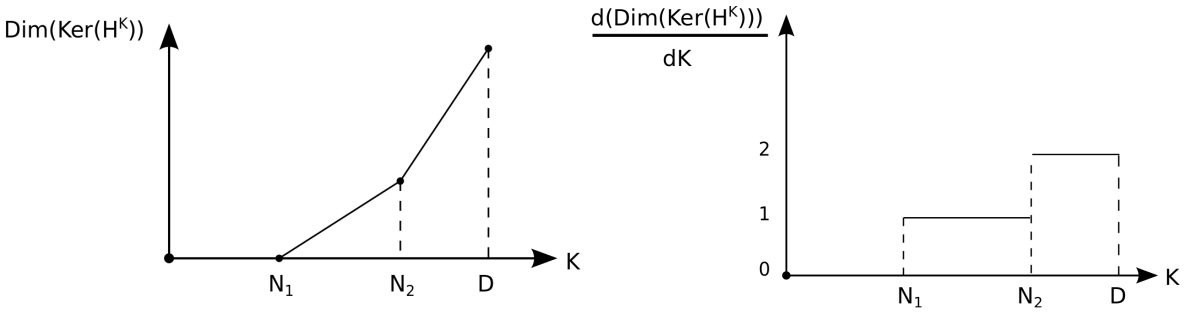
For each family of Hankel matrices defined by Definition 2.3.2 there are two constants that describe the dimension of all the kernels of those matrices.

Proposition 2.4.2. *Given the family of Hankel matrices $\{H_a^k\}_{1\leq k\leq D}$, defined by Definition 2.3.2, there are two constants, N_1^a, N_2^a, such that,*

1. $0 \leq N_1^a \leq N_2^a \leq D$
2. $(\forall k : 1 \leq k \leq D)\ \mathtt{dim}(\mathtt{Ker}(H_a^k)) = \max(0; k - N_1^a) + \max(0; k - N_2^a)$
3. $N_1^a + N_2^a = D$

Notation 2.4.3. Through the text, every time we refer to a family of Hankel matrices, we are talking about the family defined by Definition 2.3.2. For the constants, N_1 and N_2, when it is clear from the context, we skip the superindexes.

Figure 2.1 illustrates the relation between the kernels of the Hankel matrices and those constants. There we can observe how the dimension of the kernel variates when the index increases.



(a) Size of the Kernel

(b) Variation of the size of the Kernel

Figure 2.1: Relationship between H^k and N_1 and N_2

Also it is worth to consider the variation of the rank of those matrices. In the Figure 2.2, it is possible to see a "plateau". That means that, from N_1 up to N_2, the rank stays invariant. Note that if $N_1 = N_2$, this "plateau" fails to exist.

Remark 2.4.4. The maximum rank of the matrices $\{H^i\}_{0\leq i\leq D}$ is $N_1 + 1$.

To understand which vectors characterize the kernels of a family of Hankel matrices, we define the **U-chains**.

Definition 2.4.5. An **U-chain** of a vector $v = (v_0, \ldots, v_n) \in \mathbb{F}^{n+1}$ of length k is a family of vectors $U_k^0 v, U_k^1 v, \ldots, U_k^{k-1} v \in \mathbb{F}^{n+k}$, where the i-th element ($i \in [0; k-1]$) is

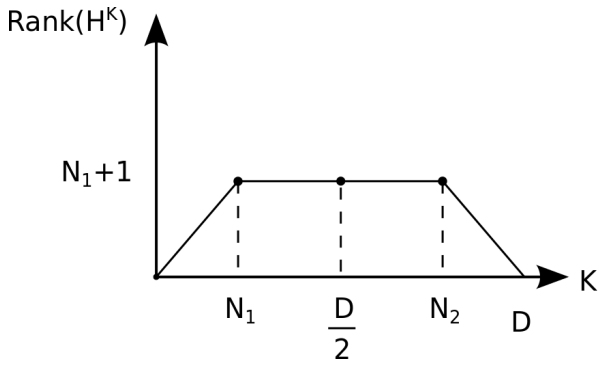


Figure 2.2: Rank of H^k

$$U_k^i v = (\underbrace{0 \dots 0}_{i}, \overbrace{v_0 \dots v_n}^{n+1}, \underbrace{0 \dots 0}_{k-1-i}) \tag{2.5}$$

Note that if v is not zero, then all the elements in an **U-chain** of v are linearly independent. The following theorem explains the relationship between the families of Hankel matrices and the U-chains. It gives an easy way to manipulate the kernels of those matrices.

Proposition 2.4.6 (Definition of **v** and **w**)**.** *Given the family of Hankel matrices $\{H^k\}_{1 \leq k \leq D}$, let N_1 and N_2 be the constants defined by Proposition 2.4.2. There are two vectors, $v \in \mathbb{F}^{N_1+1}$ and $w \in \mathbb{F}^{N_2+1}$, such that,*

For $N_1 < k \leq N_2$, the U-chain of v of length $(k - N_1)$ form a basis for $\texttt{Ker}(H^k)$.

$$\langle U_{k-N_1}^0 v, \dots, U_{k-N_1}^{k-N_1-1} v \rangle = \texttt{Ker}(H^k)$$

For $N_2 < k \leq D$, the U-chain of v of length $(k - N_1)$ together with the U-chain of w of length $(k - N_2)$ form a basis for $\texttt{Ker}(H^k)$.

$$\langle U_{k-N_1}^0 v, \dots, U_{k-N_1}^{k-N_1-1} v, U_{k-N_2}^0 w, \dots, U_{k-N_2}^{k-N_2-1} w \rangle = \texttt{Ker}(H^k)$$

Moreover, v and w are not unique. The vector v could be any vector in $\texttt{Ker}(H^{N_1+1})$, *and w could be any vector in* $\texttt{Ker}(H^{N_2+1})$ *linearly independent to 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 from Proposition 2.4.6. To relate the previous theorem with the values N_1 and N_2 we note the following.

Remark 2.4.7.

- 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 $(k - N_2)$ form a linearly independent set.
- If $i \leq N_1$, then $\texttt{Ker}(H^i) = \{0\}$
- If $N_1 < N_2$, then $\texttt{Ker}(H^{N_1+1}) = \langle v \rangle$.
- If $N_1 = N_2$, then $\texttt{Ker}(H^{N_1+1}) = \langle v, w \rangle$.
- In general, $\texttt{Ker}(H^{N_2+1}) = \langle U^0_{N_2-N_1+1}v, \ldots, U^{N_2-N_1}_{N_2-N_1+1}v, w \rangle$

Now we have an powerful way to manipulate the kernels of the family $\{H^k\}_k$ using v and w. If we consider the kernel polynomials, see Notation 2.3.4, then they can be expressed as "polynomial combinations" of P_v and P_w of degree k. The following proposition is a corollary of Heinig and Rost [13, Proposition 5.1].

Proposition 2.4.8. *The kernel polynomials of H^k are*

- *If $0 < k \leq N_1$, $\{0\}$*
- *If $N_1 < k \leq N_2$, $\{P_\mu \cdot P_v : \mu \in \mathbb{F}^{k-N_1}\}$*
- *If $N_2 < k \leq D$, $\{P_\mu \cdot P_v + P_\rho \cdot P_w : \mu \in \mathbb{F}^{k-N_1}, \rho \in \mathbb{F}^{k-N_2}\}$*

Proof. The first case is trivial. The second and the third are a consequences of the [13, Proposition 5.1]. For the sake of completeness, we sketch the proof. Given a vector v, let $U^j_k v$ be the j-th element of a U-chain of v of length k. Hence,

$$P_{U^j_k v} := x^j y^{k-1-j} P_v$$

Note, also, that $P_{\alpha \cdot u + \beta \cdot w} = \alpha P_u + \beta P_w$. Using this two facts, the proof is straightforward.

□

To conclude, the polynomials P_v and P_w do not share any root over $\overline{\mathbb{F}}$.

Proposition 2.4.9 ([13, Proposition 5.5])**.** *P_v and P_w don't share any root.*

2.5 Linear Recurrence Sequences

As we show in Chapter 4, the kernel of the Hankel matrices and the **Linear Recurrence Sequences** are deeply related. In this section we define the linear recurrence sequences, we talk about the minimal generating sequences and the arithmetic complexity of compute them.

Definition 2.5.1. A sequence S (finite or not) is said to be **linearly recurrent** when there is a finite sequence $(v_0, \ldots, v_n)$, also known as the **generating sequence**, such that:

$$S_{n+1+i} = \sum_{k=0}^{n} v_k \cdot S_{i+k} \qquad (0 \leq i)$$

Definition 2.5.2. A **minimal generating sequence** is a generating sequence of S whose length is the shortest with respect to the length of all the generating sequences of S.

The length, the uniqueness and the existence of those generating sequences change depending on the generated sequence. When the generated sequence is not finite, there is at most one minimal generating sequence. When it is finite, the minimal generating sequence may not be unique and its length could be as long as the original sequence, but it always exists.

Remark 2.5.3. Another way of defining the generating sequences is using matrices. A vector $(v_0, \ldots, v_n)$ is a generating sequence of a sequence S if and only if it is a solution to the system Equation (2.6).

$$\begin{pmatrix} S_0 & S_1 & \cdots & S_n \\ S_1 & S_2 & \cdots & S_{n+1} \\ S_2 & S_3 & \cdots & S_{n+2} \\ \vdots & \vdots & & \vdots \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} S_{n+1} \\ S_{n+2} \\ S_{n+3} \\ \vdots \end{pmatrix} \tag{2.6}$$

Note that the matrix of this system is finite if and only if S is finite.

When there is a bound for the length of the minimal generating sequences, it is possible to compute it efficiently by means of the **Berlekamp-Massey algorithm**, [2, 19]. In [12, Section 12.3] it is possible to find a proof for the Proposition 2.5.4.

Proposition 2.5.4. *Let S be a linear recurrence sequence of length n. If the length of its minimal generating sequence is at most $\lfloor \frac{n}{2} \rfloor$, then it can be computed in $O(\mathtt{M}(n) \cdot \log(n))$ ops. Where $\mathtt{M}(n)$ is the arithmetic cost of multiply two polynomial of degree n.*

2.6 Linear Change of Coordinates

A **linear change of coordinates**, in the particular case of the binary forms, can be thought as an automorphism $L_T : F[x, y] \to F[x, y]$ associated to an invertible matrix $T = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, where

$$L_T(F) = F\left(T \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right) = F\big((ax + by), (cx + dy)\big)$$

The linear change of coordinates preserves the degree of the original function. Furthermore, it preserves its rank. The relation between decompositions of different linear changes of coordinates allows us to work with a change of coordinates, instead of using the original polynomial.

Lemma 2.6.1. *Let T be a non-singular 2×2 matrix. Given a minimal decomposition for the binary form $F(x,y) = \sum_{j=1}^{r} \lambda_j(\alpha_j x + \beta_j y)^D$, a minimal decomposition for $L_T(F)$ is*

$$L_T(F) = \sum_{j=1}^{r} \lambda_j \left((\alpha_j, \beta_j) \cdot T \cdot \begin{pmatrix} x \\ y \end{pmatrix} \right)^D \tag{2.7}$$

Proof. By definition, Equation (2.7) brings a decomposition. As T is invertible, $L_{T^{-1}}(L_T(F)) = F$. Hence the rank of a binary form is preserved after a linear changes of coordinates, so the decomposition is minimal.

□

By Lemma 2.6.1, if we have a decomposition for F, then we can compute easily another decomposition for $L_T(F)$, for any invertible matrix T. This way, for getting the minimal decomposition of F we compute the minimal decomposition for $L_T(F)$, and after we recover the minimal decomposition of F.

It is possible to compute such linear change of coordinates in $O(\mathtt{M}(D) \cdot \log(D))$ ops using multi-point evaluation and interpolation algorithms.

Proposition 2.6.2. *Given a binary form F of degree D and an invertible 2×2 matrix T, we can compute $L_T(F)$ in $O(\mathtt{M}(D) \cdot \log(D))$ ops.*

Proof. To perform a linear change of coordinates we can evaluate $F(x,y)$ and then interpolate it. For univariate polynomials the multi- point evaluation and the interpolation can be achieved in $O(\mathtt{M}(D) \cdot \log(D))$ ops, [12, Chapter 10]. We show how we use the univariate algorithms to solve this problem.

If $F(x,y)$ is a binary form, then $F(x,1)$ is an univariate polynomial. We have to evaluate our target function, $L_T(F)$, in $(x_i, 1)$ for D different points x_i. Suppose that $T \cdot \binom{x_i}{1} = \binom{x_i'}{y_i'}$

- If $y_i' = 0$, then $F(x_i', 0) = a_D \cdot (x_i')^D$
- If $y_i' \neq 0$, then $F(x_i', y_i') = {y_i'}^D \cdot F(\frac{x_i'}{y_i'}, 1)$

So, for the evaluation of $L_T(F)(x,1)$, first we can use the fast multipoint evaluation algorithm for $\{\frac{x_i'}{y_i'} : y_i' \neq 0\}$ and after compute the remaining values.

To interpolate $L_T(F)$, we can interpolate the univariate polynomial $L_T(F)(x, 1)$ knowing that $L_T(F)(x_i, 1) = F(x'_i, y'_i)$, and after homogenize the result. If the degree of the interpolated polynomial is not the degree of the original polynomial, we should multiply by necessaries y to obtain the original degree.

□

3. ALGORITHM

Considering the previous chapter, the following Algorithm 1 computes a minimal decomposition. In the subsequent section we prove its correctness and termination. To achieve a better complexity bound, in the following chapters we explain in detail how to perform some of these steps. Still, this description is made to facilitate the complete overview of the algorithm.

3.1 Correctness

For the proof of correctness, consider the following lemma. By Sylvester's Theorem, we need a square-free polynomial. If P_v is not square-free, then no kernel polynomial of degree less than $N_2 + 1$ is square-free. Hence, we do not need to check this property over those polynomials.

Lemma 3.1.1. *If P_v is not square-free, then all the kernel polynomials of $\{H^k\}_{N_1<k\leq N_2}$, are not square-free.*

Proof. By Proposition 2.4.8, all the kernel polynomials of H^k can be express as $P_v \cdot P_\mu$ where $\mu \in \mathbb{F}^{k-N_1}$. Therefore, if P_v is not square-free, then $P_v \cdot P_\mu$ is not square-free either.

□

As P_v and P_w do not share any roots, it is always possible to take a square-free kernel polynomial of degree $(N_2 + 1)$. Hence, the rank is either $(N_1 + 1)$ or $(N_2 + 1)$. Moreover, P_v is square-free if and only if the rank is $(N_1 + 1)$.

Theorem 3.1.2. *If P_v is square-free, then the rank of the binary form is (N_1+1). Otherwise, the rank is $(N_2 + 1)$.*

Proof. Following the Remark 2.4.7, if $i \leq N_1$ then the kernel of H^i is trivial. Hence, by Theorem 2.3.1, there is not a decomposition. If P_v is square-free, then we take P_v and, by Corollary 2.3.5, the rank is $N_1 + 1$.

If P_v is not square-free, then by Lemma 3.1.1, all the kernel polynomials of degree less than $(N_2 + 1)$ have square roots. By Heinig and Rost [13, Proposition 5.5], P_v and P_w don't share any root, so there is a polynomial P_μ of degree $(N_2 - N_1)$ such that $Q := P_v \cdot P_\mu + P_w$ is square-free. Hence, by Proposition 2.4.8, the polynomial Q is a kernel polynomial of H^{N_2+1}. Therefore, by Corollary 2.3.5, the rank of the binary form is $(N_2 + 1)$.

□

Algorithm 1 Scheme to get a minimal decomposition for a binary form

Input: A binary form $f(x,y) = \sum_{i=0}^{D} \binom{D}{i} a_i x^i y^{D-i}$ of degree D
Output: A decomposition of f as $f(x,y) = \sum_{j=1}^{r} \lambda_j (\alpha_j x + \beta_j y)^D$

1. **Get v and w**

   - Taking the vector $a = (a_0, \ldots, a_D)$, we consider the vector v as the first vector in the kernel of the family $\{H_a^k\}_{0 \le k \le D}$. This means that $v \in \texttt{Ker}(H_a^{N_1+1})$, and $(\forall i < |v| - 1)\, \texttt{Ker}(H_a^i) = \{0\}$.
   - The vector w is a vector in the kernel of $H_a^{N_2+1}$ which is "linearly independent" to v.

   In Section 2.4 we define v and w. The Algorithm 2 compute these values.

2. **IF** $P_v(x,y)$ **is square-free**

   $Q \longleftarrow P_v$

   ELSE

   Get a square-free binary form Q

   We look for a vector μ of length $(N_2 - N_1 + 1)$, such that $(P_\mu \cdot P_v + P_w)$ is square-free.

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

   See Chapter 5 for more details.

3. **Factorize Q**

   Write Q as a product $\prod_{j=1}^{r} (\beta_j x - \alpha_j y)$

4. **Get the Lambdas**

   Get the coefficients $\lambda_i, 1 \le i \le r$, by solving the linear system from Equation (3.1)

$$\begin{pmatrix} \beta_1^D & \beta_2^D & \cdots & \beta_r^D \\ \beta_1^{D-1}\alpha_1 & \beta_2^{D-1}\alpha_2 & \cdots & \beta_r^{D-1}\alpha_r \\ \beta_1^{D-2}\alpha_1^2 & \beta_2^{D-2}\alpha_2^2 & \cdots & \beta_r^{D-2}\alpha_r^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^D & \alpha_2^D & \cdots & \alpha_r^D \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_r \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_D \end{pmatrix} \tag{3.1}$$

   In Chapter 6 we explain how to do this.

5. **Return**

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

Theorem 3.1.3 (Correctness of the Algorithm)**.** *The Algorithm 1 computes the minimal decomposition.*

Proof. Straightforward from Sylvester's Theorem and Theorem 3.1.2. The correctness of the Step 4 is proved in Chapter 6.

□

4. GETTING v AND w VIA LINEAR RECURRENCE SEQUENCES

In this section we prove that, generically, after performing a random linear change of coordinates, v is related to a minimal generating sequence of the linear recurrence sequence $(a_0, \ldots, a_D)$. Moreover, w is related to the minimal generating sequence of $(a_0, \ldots, a_{2N_1-1})$. By Proposition 2.5.4, we can find those generating sequences and compute v and w in $O(\mathtt{M}(n) \cdot \log(n))$ ops.

4.1 Algorithm

Algorithm 2 Getting v and w via Linear Recurrence Sequences

Input: A family of Hankel matrices $\{H_a^k\}_{0 \leq k \leq D}$ with **Generic Rank Profile**.
Output: Vectors v and w as Theorem 3.1.2

- $p \leftarrow 2 * \lceil \frac{D-1}{2} \rceil$
- $(u_0, \ldots, u_{N_1}) \leftarrow$ Minimal generating sequence of $(a_0, \ldots, a_p)$
- **IF** $(u_0, \ldots, u_{N_1})$ is the generating sequence of $(a_0, \ldots, a_D)$
  - $v \leftarrow (u_0, \ldots, u_{N_1}, -1)$
  - $(w_0, \ldots, w_{N_1-1}) \leftarrow$ Minimal generating sequence of $(a_0, \ldots, a_{2N_1-1})$
  - $w \leftarrow (w_0 \ldots w_{N_1-1}, -1, \underbrace{0, \ldots, 0}_{N_2-N_1+1})$

  ELSE
  - $w \leftarrow (u_0, \ldots, u_{N_1}, -1, 0)$
  - $(v_0, \ldots, v_{N_1}) \leftarrow$ Minimal generating sequence of $(a_0, \ldots, a_D, c)$, for some c.
  - $v \leftarrow (v_0, \ldots, v_{N_1}, -1)$
- **Return** v and w

In the sequel we prove the correctness and the complexity of the Algorithm 2.

4.2 Computing v as a minimal generating sequence

When the last position of v is (-1), this vector can be computed as the minimal generating sequence of $(a_0, \ldots, a_D)$. In this section we prove that statement.

Lemma 4.2.1. *The vector $(u_0, \ldots, u_r)$ is a generating sequence of $(a_0, \ldots, a_D)$ if and only if $(v_0, \ldots, v_r, -1) \in \mathtt{Ker}(H^{r+1})$*

Proof. Observe the following implications,

$$\begin{pmatrix} a_0 & \cdots & a_{r-1} & a_r \\ a_1 & \cdots & a_r & a_{r+1} \\ \vdots & \ddots & \vdots & \vdots \\ a_{D-r-1} & \cdots & a_{D-2} & a_{D-1} \\ a_{D-r} & \cdots & a_{D-1} & a_D \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_{r-1} \\ -1 \end{pmatrix} = 0 \iff \begin{pmatrix} a_0 & \cdots & a_{r-1} \\ a_1 & \cdots & a_r \\ \vdots & \ddots & \vdots \\ a_{D-r} & \cdots & a_{D-1} \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_{r-1} \end{pmatrix} = \begin{pmatrix} a_r \\ a_{r+1} \\ \vdots \\ a_D \end{pmatrix} \tag{4.1}$$

By Remark 2.5.3, the left-side of the implication indicates that for each generating sequence $(u_0, \ldots, u_r)$, the vector $(u_0, \ldots, u_r, -1)$ belongs to kernel of H^r. From the right-side of the implication follows that, if $u \in \texttt{Ker}(H^{r+1})$ and $(u_{r+1} = -1)$, the vector $(u_0, \ldots, u_r)$ is a generating sequence of $(a_0, \ldots, a_D)$.

□

Corollary 4.2.2. *The vector $(v_0, \ldots, v_{N_1})$ is a minimal generating sequence of $(a_0, \ldots, a_D)$ if and only if $(v_0, \ldots, v_{N_1}, -1) \in \texttt{Ker}(H^{N_1+1})$. If $N_1 \neq N_2$, then it is the unique minimal generating sequence.*

Proof. If $N_1 \neq N_2$, then the dimension of the kernel of $\texttt{Ker}(H^{N_1+1})$ is one. So, all the elements in $\texttt{Ker}(H^{N_1+1})$ are multiples and just one has (-1) at its last position.

□

Remark 4.2.3. Let $v \in \texttt{Ker}(H^{N_1+1})$, if the element at the last position of v different from zero, it is always possible to get a $\hat{v} \in \texttt{Ker}(H^{N_1+1})$ such that its last position is (-1).

We use the Proposition 2.5.4 to prove that, if $v_{N_1+1} = -1$, we can compute v in $O(\texttt{M}(D) \cdot \log(D))$ ops. For doing that, we need to consider the special case when $N_1 = N_2$. When $N_1 < N_2$, the length of the minimal generating sequence of $(a_0, \ldots, a_D)$ is bounded by $\lfloor \frac{D+1}{2} \rfloor$, so the hypothesis of that proposition holds. When $N_1 = N_2$, that is not true.

Lemma 4.2.4. *If $N_1 = N_2$, then for any $c \in \mathbb{F}$ the sequence $(a_0, \ldots, a_D, c)$ has a unique minimal generating sequence of length $N_1 + 1$.*

Proof. First note that H^{N_1} is invertible. As $(D = N_1 + N_2)$ and $(N_1 = N_2)$, $H^{N_1} \in \mathbb{F}^{(D-N_1+1)\times(N_1+1)}$ is a square matrix. By Proposition 2.4.2, the matrix H^{N_1} has trivial kernel. Hence, it is invertible. This implies that the following system has a unique solution,

$$H^{N_1} \cdot \begin{pmatrix} x_0 \\ \vdots \\ x_{N_1} \end{pmatrix} = \begin{pmatrix} a_{N_1+1} \\ \vdots \\ a_D \\ c \end{pmatrix} \iff \begin{pmatrix} a_0 & \cdots & a_{N_1} & a_{N_1+1} \\ a_1 & \cdots & a_{N_1+1} & a_{N_1+2} \\ \vdots & \ddots & \vdots & \vdots \\ a_{D-N_1-1} & \cdots & a_{D-1} & a_D \\ a_{D-N_1} & \cdots & a_D & c \end{pmatrix} \cdot \begin{pmatrix} x_0 \\ \vdots \\ x_{N_1} \\ -1 \end{pmatrix} = 0 \qquad (4.2)$$

By Lemma 4.2.1, for every c the solution of the system in Equation (4.2) is the unique generating sequence of $(a_0, \ldots, a_D, c)$ with length $(N_1 + 1)$. It is the minimal generating sequence because if there is another generating sequence $(u_0, \ldots, u_k)$, with $k < N_1$, then the sequences $(0, \ldots, 0, u_0, \ldots, u_k), (0, \ldots, 0, -u_0, u_0 - u_1, \ldots, u_k - 1) \in \mathbb{F}^{N_1+1}$ are generating sequences too, and hence they are different solutions for the system in Equation (4.2), which it is not possible because the solution is unique.

□

Theorem 4.2.5 (Complexity of getting v)**.** *If $v \in \texttt{Ker}(H^{N_1+1})$ and $v_{N_1+1} = -1$, then v can be computed in $O(\texttt{M}(D) \cdot \log(D))$ ops.*

Proof. The Proposition 2.5.4 proves that we can compute the minimal generating sequence of a sequence in $O(\texttt{M}(D) \cdot \log(D))$ ops when the length of the minimal generating sequence is less or equal than half of the length of the original sequence. We show how this hypothesis holds in these cases.

Consider the case when $N_1 \neq N_2$. Recalling Proposition 2.4.2, $D = N_1 + N_2$. As we assumed that $N_1 < N_2$, then $2 \cdot N_1 < D$, or equivalently, $2 \cdot N_1 + 1 \leq D$. Hence, $2 \cdot (N_1 + 1) \leq D + 1$, where, by Corollary 4.2.2, $(2 \cdot (N_1 + 1))$ is two times the length of the minimal recurrence sequence of $(a_0, \ldots, a_D)$. Therefore, $(v_0, \ldots, v_{N_1})$ can be computed in $O(\texttt{M}(D) \cdot \log(D))$ ops.

When $N_1 = N_2$, we can not deduce that $2(N_1 + 1) \leq D + 1$. In this case we consider some c and compute the minimal generating sequence of $(a_0, \ldots, a_D, c)$. By Lemma 4.2.4, its length is $(N_1 + 1)$. Hence, we have a sequence of length $(D + 2)$ and its minimal generating sequence of length $N_1 + 1 = \frac{D}{2} + 1$. So, $2(N_1 + 1) \leq D + 2$, and by Proposition 2.5.4 the sequence $(v_0, \ldots, v_{N_1})$ can be computed in $O(\texttt{M}(D) \cdot \log(D))$ ops.

Given a minimal generating sequence $(v_0, \ldots, v_{N_1})$ of $(a_0, \ldots, a_D)$, it is associated to the vector $(v_0, \ldots, v_{N_1}, -1) \in \texttt{Ker}(H^{N_1+1})$. From this follows that, when $v_{N_1+1} = -1$, v can be computed in $O(\texttt{M}(D) \cdot \log(D))$ ops.

□

4.3 Generic Rank Profile on Binary Forms

As we prove in the Section 2.6, given a decomposition for a binary form, it is easy to compute the decomposition of any other binary form obtained by applying a linear change of coordinates to the original polynomial. Up to now we have an effective method to compute v when $v_{N_1+1} \neq 0$. But this is not always true. In this section we show that after performing a **random linear change of coordinates** this holds generically. We refer to the work by Manthey et al. [18, Section 2] where it is proved that all the square principal submatrices of each matrix in the family $\{H^k\}_k$, coming from the random linear change of coordinates, are invertible if their dimensions are lower or equal to its rank. This property is known as **Generic Rank Profile**. Using this, in Section 4.4, we relate w to another linear recurrence sequence.

Definition 4.3.1. A matrix is said to have **Generic Rank Profile** if all its square principal submatrices, of dimensions lower or equal to its rank, are non-singular.

In the following we show that the Hankel matrices ${H^k}_k$ coming from a random linear change of coordinates have generic rank profile. For doing that, first we prove that for each of those matrices of rank i, its i-th principal minor is not zero. Second, we note that all of them share the same principal submatrices. Finally, we prove that if i is the rank of some of those matrices, then for all $j < i$ there is another of those matrices of rank j. Hence, as they all share the same square principal submatrices, they have Generic Rank Profile.

Proposition 4.3.2 ([18, Theorem 2.8])**.** *Let F be a binary form of degree D and $\{H_F^k\}_{1\leq k\leq D}$ its family of Hankel matrices, as in Definition 2.3.2. Let T_t be $\left(\begin{smallmatrix}1 & t\\ 0 & 1\end{smallmatrix}\right)$. Similarly, let $\{H_{L(F)}^k\}_{1\leq k\leq D}$ be the family of Hankel matrices associated to the change of coordinates $L_{T_t}(F)$. So, for each k, the determinant of the $\left(\mathtt{rk}\left(H_{L(F)}^k\right) \times \mathtt{rk}\left(H_{L(F)}^k\right)\right)$ principal minor of $H_{L(F)}^k$ is a non-zero univariate polynomial (with t the variable) of degree at most $\mathtt{rk}\left(H_{L(F)}^k\right)\cdot\left(D - 2\cdot\mathtt{rk}\left(H_{L(F)}^k\right)\right)$.*[1]

The proof of this theorem can be found in Manthey et al. [18, Theorem 2.8].[2]

Remark 4.3.3. All the Hankel matrices H^i which contain a submatrix of dimension $(j \times j)$, share the same $(j \times j)$ principal submatrix.

Lemma 4.3.4. *Given a family of Hankel matrices $\{H^i\}_i$, for all $0 < i \leq N_1 + 1$, the rank of the matrix H^{i-1} is i.*

Proof. By definition of N_1, for all $1 \leq i \leq (N_1 + 1)$, the $dim(\mathtt{Ker}(H^{i-1})) = 0$. This implies that all the matrices $H^{i-1} \in \mathbb{F}^{(D-i)\times i}$ have full rank. In those cases, $i \leq D - i$ and the rank of H^{i-1} is i.

□

[1] Notation: The matrices $H_{L(F)}^k$ are $H_{L_{T_t}(F)}^k$

[2] There is a typo in the paper. The binary forms should be $\sum_{j=0}^{d}\binom{d}{j}\gamma_{j+1}X^{d-j}Y^j$, following the notation of the paper. In the paper of the same authors, [14], they correct it.

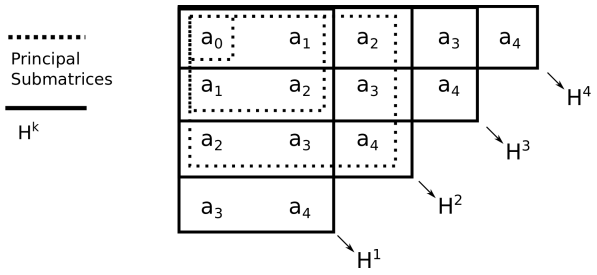


Figure 4.1: Principal submatrices of the Hankel family induced by $(a_0, \ldots, a_4)$

The previous implies that, generically, after performing a random change of coordinates, all the new matrices $\{H^i\}_i$ have generic rank profile.

Proposition 4.3.5. *Let F be a binary form of degree D and $t \in \Lambda$ where $\Lambda \subseteq \mathbb{F}$ and Λ is a finite set. Let $\{H^k_{L_{T_t}(F)}\}_{1 \leq k \leq D}$ be the family of Hankel matrices associated to $L_{T_t}(F)$. Hence, the probability of taking a t such that for all $0 < i \leq N_1 + 1$ the $(i \times i)$ principal submatrices of $H^k_{L_{T_t}(F)}$ are invertible is bounded by*

$$Prob\left((\forall k \leq D)\ H^k_{L_{T_t}(F)} \text{ has Generic Rank Profile} \mid t \in \Lambda\right) \geq 1 - \sum_{i=1}^{N_1+1} \frac{i \cdot (D - 2 \cdot i)}{\#\Lambda}$$

Proof. Recalling Remark 4.3.3, all the matrices in a family share the same principal submatrices. We identify each principal submatrix with a matrix in the family. Recalling Remark 2.4.4, the maximum possible rank in the family is $N_1 + 1$, therefore we take the matrices $\{H^{i-1}_{L_{T_t}(F)}\}_{1 \leq i \leq (N_1+1)}$. By Lemma 4.3.4, in those family, the i-th Hankel matrix have rank i. By Proposition 4.3.2, for each $1 \leq i \leq (N_1 + 1)$, the determinant of the $(i \times i)$ principal submatrix of $H^{i-1}_{L_{T_t}(F)}$ is a non-zero univariate polynomial of degree at most $i \cdot (D - 2i)$, because the rank of that matrix is i.

For each $1 \leq i \leq (N_1 + 1)$, by Schwartz–Zippel lemma [22, 24], the probability that, taking randomly and uniformly a $t \in \Lambda$, the determinant of the $(i \times i)$ principal submatrix of i-th matrix vanishes is bounded by,

$$Prob\left((i \times i) \text{ principal submatrix of } H^{i-1}_{L_{T_t}(F)} \text{ is singular} \mid t \in \Lambda\right) \leq \frac{i \cdot (D - 2i)}{\#\Lambda}$$

As the probability of the union is smaller, or equal, than the sum of the probabilities, we have

$$Prob\left((\exists i \leq (N_1+1))\ (i \times i) \text{ principal submatrix is singular} \mid t \in \Lambda\right) \leq \sum_{i=1}^{N_1+1} \frac{i \cdot (D-2\cdot i)}{\#\Lambda}$$

Observe that the last equation is equivalent to our bound.

□

This proposition allows us to perform a random linear change of coordinates and bound the probability of having a family of Hankel matrices with generic rank profile. If after performing the change that property holds, then there is a v whose last position is (-1).

Theorem 4.3.6. *Given a family of Hankel matrices with generic rank profile, if $N_1 \neq N_2$, then there is a vector $v \in \mathtt{Ker}(H^{N_1+1})$ such that $v_{N_1+1} = -1$. Such vector is unique.*

Proof. By Lemma 4.3.4, the matrix H^{N_1} has rank N_1+1. As we assume that it has generic rank profile, its $((N_1+1)\times(N_1+1))$ principal submatrix, from now on called J, is invertible. Recalling Remark 4.3.3, $H^{N_1+1} \in \mathbb{F}^{(D-N_1)\times(N_1+2)}$ and J is the principal submatrix of H^{N_1+1}, because $N_1 < N_2$ and $(D-N_1) \geq (N_1+1)$.

The kernel of the matrix H^{N_1+1}, by definition, is not trivial. As the first N_1+1 columns of H^{N_1+1} are linearly independent, because the columns of J are subvectors of those columns, the last position of any non-trivial vector in the kernel of H^{N_1+1} is not zero. Therefore, there is a vector $v \in \mathtt{Ker}(H^{N_1+1})$ such that $v_{N_1+1} = -1$.

As $N_1 \neq N_2$, the dimension of $\mathtt{Ker}(H^{N_1+1})$ is one, so there is just one v like that.

□

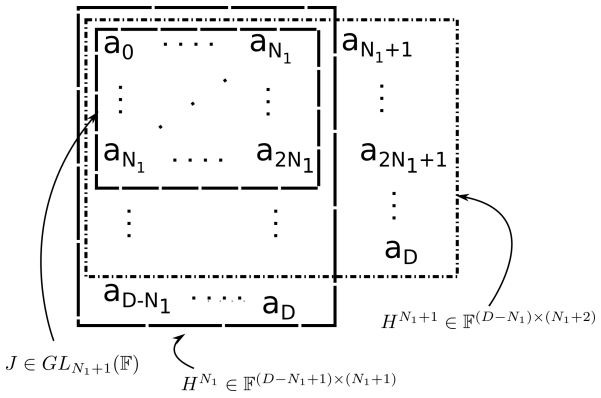


Figure 4.2: Relation between H^{N_1} and H^{N_1+1}

4.4 Computing w assuming Generic Rank Profile

To compute w we can assume that the family of Hankel matrices have generic rank profile. We show that w can be obtained computing the minimal generating sequence of $(a_0, \ldots, a_{2N_1-1})$.

Lemma 4.4.1. *Given a family of Hankel matrices with generic rank profile, there is a vector $w \in \texttt{Ker}(H^{N_2+1})$, linearly independent to the U-chain of v of length $N_2 - N_1 + 1$, such that,*

$$w = (w_0 \ldots w_{N_1-1}, -1, \overbrace{0 \ldots 0}^{N_2-N_1+1}) \tag{4.3}$$

Where $(w_0, \ldots, w_{N_1-1})$ is the minimal generating sequence of $(a_0, \ldots, a_{2N_1-1})$.

Proof. First of all, note that if such vector exists, then it has to be linearly independent to the U-chain of v of length $N_2 - N_1 + 1$ because $w_{N_1+1} = 0$ and, as the family has generic rank profile, by Theorem 4.3.6, $v_{N_1+1} \neq 0$.

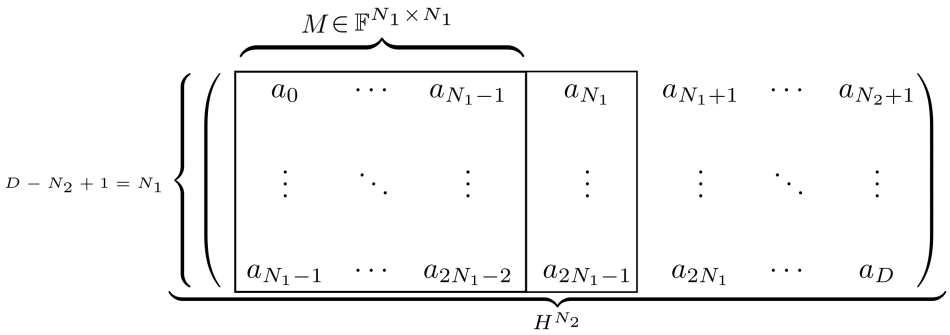
The matrix H^{N_2+1} has dimensions $(D - N_2) \times (N_2 + 2)$. As $N_1 = D - N_2 < N_2 + 2$ and we assumed having generic rank profile, the $N_1 \times N_1$ principal submatrix of H^{N_2+1}, from now on called M, is invertible. So, the following system has a unique solution,

$$\underbrace{\begin{pmatrix} a_0 & \cdots & a_{N_1-1} \\ \vdots & \ddots & \vdots \\ a_{N_1-1} & \cdots & a_{2N_1-2} \end{pmatrix}}_{M} \cdot \begin{pmatrix} w_0 \\ \vdots \\ w_{N_1-1} \end{pmatrix} = \begin{pmatrix} a_{N_1} \\ \vdots \\ a_{2N_1-1} \end{pmatrix} \tag{4.4}$$

Reasoning as in the proof of Corollary 4.2.2, the vector w belongs to $\texttt{Ker}(H^{N_2+1})$, where w is defined as,

$$w = (w_0 \ldots w_{N_1-1}, -1, \overbrace{0 \ldots 0}^{N_2-N_1+1}) \tag{4.5}$$

The sequence $(w_0 \ldots w_{N_1-1})$ is the minimal generating sequence of $(a_0, \ldots, a_{2N_1-1})$. Suppose that there is another generating sequence, $(u_0, \ldots, u_n)$, with $n < N_1$. Hence, as in Corollary 4.2.2, the vector $(u_0, \ldots, u_n, -1, 0, \ldots, 0) \in \mathbb{F}^{N_2+2}$ belongs to $\texttt{Ker}(H^{N_2+1})$. Therefore, the first $(n+1)$ columns of H^{N_2+1} are not linearly independent. The first N_1 columns of H^{N_2+1} are linearly independent, because they are the columns of the invertible matrix M. Therefore, $n = (N_1 - 1)$ and $u_i = w_i$ because the solution of the Equation (4.4) is unique.



$$v = \begin{pmatrix} v_0 & \dots & v_{N_1-1} & v_{N_1} & -1 \end{pmatrix}$$
$$w = \begin{pmatrix} w_0 & \dots & w_{N_1-1} & -1 & 0 & \dots & 0 \end{pmatrix}$$

Figure 4.3: Relation between w, M and H^{N_2+1}

Figure 4.3 illustrates the proof.

□

Corollary 4.4.2. *Given a family of Hankel matrices with generic rank profile where N_1 is known, the vector w from Lemma 4.4.1 can be computed in $O(\mathtt{M}(D) \cdot \log(D))$ ops.*

Proof. The arithmetic complexity of computing such w comes from getting the minimal generating sequence of $(a_0, \dots, a_{2N_1-1})$. As is proved in the previous lemma, the length of such generating sequence is N_1, which is equal to half of the length of the original sequence. By Proposition 2.5.4, such minimal generating sequence can be computed in $O(\mathtt{M}(D) \cdot \log(D))$ ops.

□

4.5 Complexity of computing v and w

In this section we prove the complexity and the correctness of the Algorithm 2. Depending if $N_1 = N_2$ holds or not, we have different approaches to compute v and w but, up to now, we can not decide in which case we are. The following lemma gives a solution for this issue. Note that as $(N_1 + N_2) = D$, an odd D implies $N_1 \neq N_2$.

Lemma 4.5.1. *Let $f = \sum_i \binom{D}{i} a_i x^i y^{D-i}$ be a binary form of even degree D such that the family $\{H_a^k\}_k$, defined by the sequence $a = (a_0, \dots, a_D)$, have generic rank profile. The minimal generating sequence of $b = (a_0, \dots, a_{D-1})$ is a generating sequence of a, if and only if, $N_1^a < N_2^a$. Where N_1^a and N_2^a are defined by Proposition 2.4.2.*

Proof. Let $(v_0, \dots, v_n)$ a minimal generating sequence of b.

First note that if $\{H_a^k\}_k$ has generic rank profile, then $\{H_b^k\}_k$ has generic rank profile too because they share the same submatrices. As D is even, by Proposition 2.4.2, $D-1 = N_1^b + N_2^b$, which implies that $N_1^b < N_2^b$.

By Theorem 4.3.6, there is a unique vector $v = (v_0, \ldots, v_{N_1^b}, -1)$ in the kernel of $H_b^{N_1^b+1}$, so by the Corollary 4.2.2, $(v_0, \ldots, v_{N_1^b})$ is the unique minimal generating sequence of b. If $(v_0, \ldots, v_{N_1^b})$ is generating sequence of a, by Lemma 4.2.1, the vector $(v_0, \ldots, v_{N_1^b}, -1)$ belongs to the kernel of $H_a^{N_1^b+1}$ and hence, by definition of N_1^a, $N_1^a \leq N_1^b$. If $N_1^a = N_2^a$, then

$$D = N_1^a + N_2^a = 2 \cdot N_1^a \leq 2 \cdot N_1^b < N_1^b + N_2^b = D - 1$$

Therefore, if the minimal generating sequence of b is a generating sequence of a, $N_1^a < N_2^a$.

As H_b^i is a submatrix of H_a^i, if $u \in \mathtt{Ker}(H_a^i)$, then $u \in \mathtt{Ker}(H_b^i)$. So, $N_1^a \geq N_1^b$. Note that if $u \in \mathtt{Ker}(H_b^i)$, then $(u, 0) \in \mathtt{Ker}(H_a^{i+1})$. So, $(N_1^b + 1) \geq N_1^a$. Hence, $(N_1^b + 1) \geq N_1^a \geq N_1^b$.

If $N_1^a < N_2^a$, by Proposition 2.4.2, the dimension of the kernel of $H_a^{N_1^a+1}$ and $H_b^{N_1^b+1}$ is one.

If $(N_1^b + 1) = N_1^a$, and $u \in \mathtt{Ker}(H_b^{N_1^b+1})$, then $(u, 0) \in \mathtt{Ker}(H_a^{N_1^a+1})$. As the dimension of $\mathtt{Ker}(H_a^{N_1^a+1})$ is one and $N_1^a < N_2^a$, all the vectors in $\mathtt{Ker}(H_a^{N_1^a+1})$ have a zero last position. But this is a contradiction to Theorem 4.3.6, as we assumed generic rank profile.

Hence, if $N_1^a < N_2^a$, then $N_1^a = N_1^b$. As one is a submatrix of the other, $\mathtt{Ker}(H_b^{N_1^b}) = \mathtt{Ker}(H_a^{N_1^a})$. By Theorem 4.3.6, as we assumed generic rank profile, there is a vector $(v_0, \ldots, v_{N_1^b}, -1) \in \mathtt{Ker}(H_a^{N_1^a})$. Therefore, if $N_1^a < N_2^a$, then the minimal generating sequence of a and b is $(v_0, \ldots, v_{N_1^b})$.

□

Lemma 4.5.2. *Let $f = \sum_i \binom{D}{i} a_i x^i y^{D-i}$ be a binary form of degree D defined by the sequence $a = (a_0, \ldots, a_D)$ such that $\{H_a^k\}_k$ has generic rank profile. It is possible to decide if $N_1^a = N_2^a$ in $O(\mathtt{M}(D) \cdot \log(D))$ ops.*

Proof. By Proposition 2.4.2, $D = N_1^a + N_2^a$. If the D is odd, then $N_1^a < N_2^a$, so deciding takes $O(1)$ ops.

If D is even, by Lemma 4.5.1, the minimal generating sequence of $b = (a_0, \ldots, a_{D-1})$ is a generating sequence of a, if and only if, $N_1^a < N_2^a$.

In that case, the length of b is even, so $N_1^b < N_2^b$. By Theorem 4.3.6, there exist a vector $(v_0, \ldots, v_{N_1^b}, -1) \in \mathtt{Ker}(H^{N_1^b+1})$, and by Corollary 4.2.2, $(v_0, \ldots, v_{N_1^b})$ is a minimal generating sequence of b. By Proposition 2.5.4, as the length of that minimal generating sequence is $(N_1^b + 1) \leq \lfloor \frac{D}{2} \rfloor$, it can be computed in $O(\mathtt{M}(D) \cdot \log(D))$ ops.

□

Theorem 4.5.3 (Correctness and Complexity)**.** *Given a binary form f of degree D, if its family of Hankel matrices has generic rank profile, the Algorithm 2 computes v and w in $O(\mathtt{M}(D) \cdot \log(D))$ ops.*

Proof. The correctness and the complexity follows from Lemma 4.5.2, Theorem 4.2.5 and from Corollary 4.4.2.

□

5. ALGEBRAIC DEGREE OF THE PROBLEM

In this section we show that, when the rank of the binary form is $N_2 + 1$, we can take a square-free kernel polynomial Q of degree $N_2 + 1$ whose bigger irreducible divisor over $\mathbb{F}[x]$ has degree at most N_1. Moreover, we prove that for almost all the choices of $(N_2 - N_1 + 1)$ elements in $\mathbb{F}$, we can take a square-free kernel polynomial whose roots include those elements.

Lemma 5.1. *Let f we a binary form whose rank is $N_2 + 1$. Given a set $\Lambda \subset \mathbb{F} \setminus \{\rho : P_v(\rho, 1) = 0\}$ of size $(N_2 - N_1 + 1)$, there is a unique polynomial Q in the kernel of $H_f^{N_2+1}$ such that for all the $\alpha_j \in \Lambda$, $Q(\alpha_j, 1) = 0$.*

Proof. By Proposition 2.4.8, all the polynomials in the kernel of $H_f^{N_2+1}$ are written as $Q_\mu := P_\mu \cdot P_v + P_w$, with P_μ of degree $(N_2 - N_1)$. As the elements of the set are not roots of $P_v(x, 1)$ and we want $Q(x, 1)$ to be zero over those point, we can interpolate P_μ knowing that

$$P_\mu(\alpha_j, 1) = -\frac{P_w(\alpha_j, 1)}{P_v(\alpha_j, 1)} \tag{5.1}$$

As the degree of P_μ is $(N_2 - N_1)$, and the set Λ has $(N_2 - N_1 + 1)$ elements, there is just one polynomial that interpolate the points of the equation (5.1). Therefore, we know that the interpolated polynomial is the unique polynomial of degree equal or less to $(N_2 - N_1)$ such that the vector of its coefficient belongs to the kernel of $H_f^{N_2+1}$. Given that polynomial, we homogenize it to get a binary form of degree $(N_2 + 1)$.

□

If we choose randomly and uniformly $(N_2 - N_1 + 1)$ roots for Q, that polynomial, generically, is square-free.

Theorem 5.2. *Let f be a binary form whose rank is $N_2 + 1$ and let $\Gamma \subset \mathbb{F} \setminus \{\rho : P_v(\rho, 1) = 0\}$ be a set of cardinal G. Taking randomly and uniformly $(N_2 - N_1 + 1)$ elements from Γ, the probability that the unique polynomial Q in the kernel of $H_f^{N_2+1}$, as in Lemma 5.1, has not **square-roots** is bounded by,*

$$Prob\,(Q \text{ is a square-free polynomial }) \geq 1 - \frac{(N_1 + 1)(3N_2 - N_1 + 1)}{G - N_2 + N_1}$$

For the proof we refer to Appendix A.

This means that, if the rank is $(N_2 + 1)$, $(N_2 - N_1 + 1)$ of those roots can be chosen. This implies that the biggest irreducible factor of Q has, at most, degree N_1.

Theorem 5.3. *The degree of the biggest irreducible factor of Q has, at most, degree $N_1 + 1$ when the rank is $N_1 + 1$, or degree N_1 when the rank is $N_2 + 1$.*

Proof. When the rank is $N_1 + 1$, the degree of any kernel polynomial is $N_1 + 1$, so the degree of the biggest irreducible factor could be as big as the rank. When the rank is $N_2 + 1$, as Theorem 5.2 assures, there are kernel polynomials Q where $(N_2 - N_1 + 1)$ of the roots of $Q(x, 1)$ belongs to certain subset of the $\mathbb{F}$. Hence, $(N_2 - N_1 + 1)$ of the irreducible factors of those kernel polynomials are lineal factors, and therefore the biggest irreducible factor has, at most, degree N_1.

□

Remark 5.4. When the rank is $N_2 + 1$, if we choose randomly and uniformly the roots of the kernel polynomial, generically, the binary form $g(x, y) = y$ does not divide that kernel polynomial. The proof is similar to the one from Theorem 5.2.

6. COMPUTING THE λS VIA POLYNOMIAL DIVISION IN $\mathbb{F}[X]$

In this section we show how to compute the lambdas from Step 4 of Algorithm 1. We prove that we can express them as a rational function evaluated over the roots of the polynomial Q. Also, we show that the arithmetic complexity of computing the denominator and numerator of such rational function is $O(\mathtt{M}(D))$ ops.

Lemma 6.1. *If $\sum_{j=1}^{r} \lambda_j \left(\alpha_j x + \beta_j y\right)^D$ is a minimal decomposition of $f = \sum_{i=0}^{D} \binom{D}{i} a_j x^i y^{D-i}$,then the vector $(\lambda_1, \ldots, \lambda_r)$ is the unique solution of the system*

$$\begin{pmatrix} \beta_1^D & \beta_2^D & \cdots & \beta_r^D \\ \beta_1^{D-1}\alpha_1 & \beta_2^{D-1}\alpha_2 & \cdots & \beta_r^{D-1}\alpha_r \\ \beta_1^{D-2}\alpha_1^2 & \beta_2^{D-2}\alpha_2^2 & \cdots & \beta_r^{D-2}\alpha_r^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^D & \alpha_2^D & \cdots & \alpha_r^D \end{pmatrix} \begin{pmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_D \end{pmatrix} \tag{6.1}$$

Proof. By Sylvester's theorem, we know that the (α_j, β_j) are pairwise linearly independent. If we expand $\sum_{j=1}^{r} \lambda_j \left(\alpha_j x + \beta_j y\right)^D$, then we obtain,

$$\sum_{j=1}^{r} \lambda_j \left(\alpha_j x + \beta_j y\right)^D = \sum_{j=1}^{r} \lambda_j \left(\sum_{i=0}^{D} \binom{D}{i} \alpha_j^i \beta_j^{D-i} x^i y^{D-i} \right) = \sum_{i=0}^{D} \binom{D}{i} \left(\sum_{j=1}^{r} \lambda_j \alpha_j^i \beta_j^{D-i} \right) x^i y^{D-i}$$

Hence, if f is equal to that polynomial

$$f(x,y) = \sum_{i=0}^{D} \binom{D}{i} a_j x^i y^{D-i} = \sum_{i=0}^{D} \binom{D}{i} \left(\sum_{j=1}^{r} \lambda_j \alpha_j^i \beta_j^{D-i} \right) x^i y^{D-i}$$

Therefore, $a_i = \sum_{j=1}^{r} \lambda_j \alpha_j^i \beta_j^{D-i}$, which is equivalent to Equation (6.1).

For the uniqueness of the lambdas, let us assume that β_j is different to zero.

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \left(\frac{\alpha_1}{\beta_1}\right)^1 & \left(\frac{\alpha_2}{\beta_2}\right)^1 & \cdots & \left(\frac{\alpha_r}{\beta_r}\right)^1 \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{\alpha_1}{\beta_1}\right)^D & \left(\frac{\alpha_2}{\beta_2}\right)^D & \cdots & \left(\frac{\alpha_r}{\beta_r}\right)^D \end{pmatrix} \cdot \begin{pmatrix} \beta_1^D & 0 & \cdots & 0 \\ 0 & \beta_2^D & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta_r^D \end{pmatrix} = \begin{pmatrix} \beta_1^D & \beta_2^D & \cdots & \beta_r^D \\ \beta_1^{D-1}\alpha_1 & \beta_2^{D-1}\alpha_2 & \cdots & \beta_r^{D-1}\alpha_r \\ \beta_1^{D-2}\alpha_1^2 & \beta_2^{D-2}\alpha_2^2 & \cdots & \beta_r^{D-2}\alpha_r^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^D & \alpha_2^D & \cdots & \alpha_r^D \end{pmatrix}$$

As the (α_j, β_j) are pairwise linearly independent, the coefficients $\left(\frac{\alpha_j}{\beta_j}\right)$ are all different. Note that the first matrix is a Vandermonde matrix whose coefficients are all different, so it is full rank. The diagonal matrix is invertible because in its diagonal there are not zeros. Hence, the matrix from Equation (6.1) has full rank. Sylvester's Theorem assures that there is a solution to that system and therefore, the solution is unique.

If β_i is zero, then for $j \neq i$, $\beta_j \neq 0$ because the (α_j, β_j) are pairwise linearly independent. In that case, adapting the above argument is straightforward.

□

Remark 6.2. As a corollary from the Theorem 4.3.6, after a random linear change of coordinates, generically the last position of the vector v is not zero, so the polynomial P_v is not divisible by y. In Remark 5.4 we observed that when the rank is N_2+1, generically the chosen square-free kernel polynomial Q is not divisible by y neither. This means that we expect all the β_i to be different from zero. In the following, we assume that. Anyway, our approach is easily extensible to the case when some β_i is zero.

For this reason, **all the following propositions assume that all the β_i are one**.

Lemma 6.3. *If all the β_j are not zero, then they can be taken as* 1.

Proof. As all the β_j are not zero, $Q(x, 0) \neq 0$. In that case, Q can be rewritten as,

$$Q(x, y) := \prod_{j=1}^{r} (\beta_j x - \alpha_j y) = c \cdot \prod_{j=1}^{r} (x - \frac{\alpha_j}{\beta_j} y)$$

If we take we take $\frac{Q}{c}$, then we can just consider the coefficients $\hat{\beta}_i = 1$ and $\hat{\alpha}_i = \frac{\alpha_j}{\beta_j}$.

□

Corollary 6.4. *The lambdas can be taken as the unique solution of*

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_r \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^r & \alpha_2^r & \cdots & \alpha_r^r \end{pmatrix} X = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_r \end{pmatrix} \tag{6.2}$$

Proof. Note that the matrix of Equation (6.2) is the principal $((r+1)\times(r+1))$ of the matrix of Lemma 6.1, which is invertible.

□

To be able to write clearly the inversion formula for the transpose of the Vandermonde matrix we must introduce some notation.

Definition 6.5. Given a polynomial $P(x) := \sum_{i=0}^{n} a_i x^i$, the reverse polynomial of P is,

$$rev(P)(x) := \sum_{i=0}^{n} a_{r-i} x^i$$

Definition 6.6. Given a polynomial $P(x) := \sum_{i=0}^{n} a_i x^i$ and $0 < k \leq n$, let Quo and Rem be,

$$Quo(P,k)(x) := \sum_{i=k}^{n} a_i x^{i-k} \qquad Rem(P,k)(x) := \sum_{i=0}^{k-1} a_i x^i$$

Proposition 6.7. *Let Q be a square-free binary form of degree r, obtained after the Step 4 of Algorithm 1 for a given form f. Let the Q' be the derivative of $Q(x,1)$ and the polynomial $T(x)$,*

$$T(x) := Quo\Big(\Big(Q(x,1) \ \cdot \ rev\Big(Rem\big(f(x,1),r\big)\Big)\Big), r\Big) \tag{6.3}$$

Hence, each λ_j from Equation (6.1) can be written as

$$\lambda_j = \frac{T}{Q'}(\alpha_j)$$

For a proof of Proposition 6.7 we refer to Kaltofen and Yagati [17, Section 5]. Consider that the previous proposition solves the linear system of the Equation (6.2), which involves a transpose of a Vandermonde matrix.

Corollary 6.8. *Given a Q related with the kernel polynomial of a minimal decomposition of binary form f of degree D, f can be written as*

$$f(x,y) = \sum_{\{\alpha \in \overline{\mathbb{F}} | Q(\alpha,1)=0\}} \frac{T}{Q'}(\alpha) \cdot (\alpha x + y)^D$$

Lemma 6.9. *Given a kernel polynomial Q of degree r, obtained after the Step 4 for a binary form f of degree D, the polynomials T and Q' from Proposition 6.7 can be computed in $O(\mathtt{M}(D))$ ops.*

Proof. The functions rev, Quo, Rem and the derivative have a linear arithmetic complexity with respect to the degree of the polynomial. In this case, such degree is bounded by $2D$, because the degree of Q is at most D. The only operation involved there whose complexity is not linear, is in the multiplication $\Big(Q(x,1) \cdot rev\Big(Rem\big(f(x,1),r\big)\Big)\Big)$. As the degree of both polynomials is bounded by D, the multiplication can be computed in $O(\mathtt{M}(D))$ ops.

□

7. ARITHMETIC COMPLEXITY AND FORM OF THE SOLUTIONS

In the previous sections we prove that the Algorithm 1 is correct and we analyze the arithmetic complexity of each step. In this section we summarize all the assumptions that we make above to conclude that the arithmetic complexity of getting an algebraic solution is bounded by $O(\mathtt{M}(D) \cdot log(D))$ ops, where D is the degree of the original polynomial. Moreover, we show the special form that has the minimal decomposition obtained. It can be expressed as an addition of a rational polynomial $\mathbb{F}[x, y](z)$ evaluated over all the roots of a univariate polynomial $Q \in \mathbb{F}[x]$ with a bounded algebraic degree.

Algorithm 3 Computing the algebraic formulation of the minimal decomposition

Input: A binary form $f \in \mathbb{F}[x, y]$ of degree D.
Output: A minimal decomposition for $f(x, y)$

1. **Apply a random linear change of coordinates to** f

   $G \longleftarrow L_C(f)$

   Where C is a nonsingular random matrix in $\mathbb{F}^{2\times 2}$

   And G the binary form obtained after the change of coordinates of f with C

2. **Apply Algorithm 1 to** G

   Where the output from the Algorithm 1 is,

$$\sum_{\{\alpha\in\overline{\mathbb{F}}|Q(\alpha,1)=0\}} \frac{T}{Q'}(\alpha) \cdot \left((\alpha, 1) \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right)^D$$

   With $T, Q', Q(x, 1) \in \mathbb{F}[x]$.

3. **Return the decomposition for** f

$$\sum_{\{\alpha\in\overline{\mathbb{F}}|Q(\alpha,1)=0\}} \frac{T}{Q'}(\alpha) \cdot \left((\alpha, 1) \cdot C^{-1} \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right)^D$$

Theorem 7.1. *The algorithm 3 computes an algebraic formulation of a minimal decomposition for a binary form f of degree D in $O(\mathtt{M}(D) \cdot log(D))$ ops.*

Proof. For the arithmetic complexity, it is important to emphasize the application of a random linear change of coordinates to the original binary form. The complexity of computing such change of coordinates (Step 1) is $O(\mathtt{M}(D) \cdot log(D))$ ops, by Proposition 2.6.2.

In Step 2, we should analyze each step of Algorithm 1. First of all, as we explain in Chapter 4, after performing a random linear change of coordinates, we can compute, generically, the vectors v and w (Proposition 2.4.6) in $O(\mathtt{M}(D) \cdot log(D))$ ops.

Second, as we explain in Chapter 5, we can obtain a square-free kernel polynomial in $O(\mathtt{M}(D) \cdot log(D))$ ops. Moreover, that kernel polynomial has the algebraic degree bounded by $\min\Big(rank(f)\ ,\ \big(D - rank(f) + 1\big)\Big)$ as is explained in Theorem 5.3.

Third, we have to solve the system of the equation (3.1), which we can do in $O(\mathtt{M}(D))$ ops, by Lemma 6.9. So Step 2 takes $O(\mathtt{M}(D) \cdot log(D))$ ops.

The step 3 has a constant complexity. Therefore, we conclude that Algorithm 3 computes an algebraic formulation for a minimal decomposition of a binary form of degree D in $O(\mathtt{M}(D) \cdot log(D))$ ops.

□

Finally, note the form of the output of the Algorithm 3.

Corollary 7.2. *Given a binary form $f \in \mathbb{F}[x, y]$ of degree D, the Algorithm 3 decomposes that binary form as*

$$f(x, y) = \sum_{\{\alpha \in \overline{\mathbb{F}} | Q(\alpha,1)=0\}} \frac{T}{Q'}(\alpha) \cdot \left((\alpha, 1) \cdot C^{-1} \cdot \binom{x}{y}\right)^D$$

Where C is a 2×2 invertible matrix and $Q'(x), Q(x, 1), T(x) \in \mathbb{F}[x]$ have a degree of at most D. The degree of the minimal algebraic extension of $\mathbb{F}$ that contains the set $\{\alpha \in \overline{\mathbb{F}} \mid Q(\alpha, 1) = 0\}$ is upper bounded by $\mathtt{Min}\Big(rank(f)\ ,\ \big(D - rank(f) + 1\big)\Big)$.

8. NEW PROOFS FOR CLASSIC RESULTS

In this chapter we prove some results by Helmke [14, Theorem B] and Comas and Seiguer [7, Theorem 2] using our approach. Moreover, those papers just work over the complex numbers. Under our formulation of the problem, we extend those results for any field (we consider the decompositions where the coefficients belong to the algebraic closure of $\mathbb{F}$).

Sylvester's Theorem proves that every possible decomposition is associated to a square-free polynomial Q, and moreover, to its roots. Hence, any multiple of Q has the same decomposition associated. Therefore, we say that we have an "unique" minimal decomposition when all the polynomials associated to all the minimal decompositions are multiples.

Corollary 8.1. *If $N_1 \neq N_2$ and P_v is square-free, then the minimal decomposition is "unique".*[1]

Proof. This follows from the Remark 2.4.7. If $N_1 \neq N_2$, then the dimension of the kernel of H^{N_1+1} is one. Let v be any vector in H^{N_1+1}. All the polynomials in the kernel of H^{N_1+1} are multiples of P_v. Hence, by Theorem 3.1.2, as P_v is square-free, the rank of the binary form is $N_1 + 1$. So all the candidates polynomials for Sylvester's Theorem are multiples. Therefore, given two minimal decompositions, for each term in the first decomposition, there is a multiple term in second one, and vice versa.

□

As a corollary we can prove [14, Theorem B] and the [7, Theorem 2], which relates the rank of a binary form with the rank of a Hankel matrix.

Consider the binary forms $f_1 := \sum_{i=0}^{2n} \binom{2n}{i} a_i x^i y^{2n-i}$ and $f_2 := \sum_{i=0}^{2n+1} \binom{2n}{i} a_i x^i y^{2n+1-i}$. Regard the Hankel matrices $H^n_{f_1}$ and $H^n_{f_2}$,

$$H^n_{f_1} := \begin{pmatrix} a_0 & a_1 & \cdots & a_n \\ a_1 & a_2 & \cdots & a_{n+1} \\ \vdots & \ddots & \vdots & \vdots \\ a_n & a_{n+1} & \cdots & a_{2n} \end{pmatrix} \quad \textit{and} \quad H^n_{f_2} := \begin{pmatrix} a_0 & a_1 & \cdots & a_n \\ a_1 & a_2 & \cdots & a_{n+1} \\ \vdots & \ddots & \vdots & \vdots \\ a_{n+1} & a_{n+2} & \cdots & a_{2n+1} \end{pmatrix} \tag{8.1}$$

Note that the rank of the matrix $H^n_{f_i}$ is $(N_1^{f_i} + 1)$. Therefore, the rank of the binary form f_i of degree D is either $(N_1^{f_i} + 1) = \mathtt{rk}(H^n_{f_i})$ or $(N_2^{f_i} + 1) = D - \mathtt{rk}(H^n_{f_i}) + 2$.

[1] This means that for any minimal decomposition each term is a multiple of another term in any other decomposition

Lemma 8.2. *Let f be a binary form of degree D. Hence,* $\mathtt{rk}\left(H_f^{\lfloor \frac{D}{2} \rfloor}\right) = N_1 + 1$.

Proof. By Proposition 2.4.2, as $D = N_1 + N_2$ and $N_1 \leq N_2$,

$$\mathtt{dim}(H_f^{\lfloor \frac{D}{2} \rfloor}) = \underbrace{\min\left(\left(\left\lfloor \frac{D}{2} \right\rfloor - N_1\right); 0\right)}_{\lfloor \frac{D}{2} \rfloor - N_1} + \underbrace{\min\left(\left(\left\lfloor \frac{D}{2} \right\rfloor - N_2\right); 0\right)}_{0}$$

By Rank–Nullity theorem, as $H_f^{\lfloor \frac{D}{2} \rfloor} \in \mathbb{F}^{(D - \lfloor \frac{D}{2} \rfloor + 1) \times (\lfloor \frac{D}{2} \rfloor + 1)}$

$$\left\lfloor \frac{D}{2} \right\rfloor + 1 = \mathtt{rk}\left(H_f^{\lfloor \frac{D}{2} \rfloor}\right) + \left\lfloor \frac{D}{2} \right\rfloor - N_1$$

□

Proposition 8.3 ([14, Theorem B] and [7, Theorem 2])**.** *The rank of a binary form f of degree D is either* $\mathtt{rk}\left(H_f^{\lfloor \frac{D}{2} \rfloor}\right)$ *or* $\left(D - \mathtt{rk}\left(H_f^{\lfloor \frac{D}{2} \rfloor}\right) + 2\right)$

9. THE GENERAL CASE

As we mentioned in the introduction, the problem that we solved in this thesis is a particular case of a bigger problem called "Symmetric Tensor Decomposition". Now we are going to talk a little about this general case, and our formulation will be just in terms of homogeneous polynomials. To get more details between these two formulation we recommend the paper from Comon et al. [11]. In the following we will discuss some known results, working over the complex numbers.

Given a homogeneous form $g(x_1, \ldots, x_n)$ of degree D, $g \in \mathbb{C}[x_1, \ldots, x_n]_D$, we say that we have a decomposition of it, if we have $u_1, \ldots, u_r \in \mathbb{C}^n$ such that Equation (9.1) holds.

$$g(x_1, \ldots, x_n) = \sum_{i=1}^{r} (u_{i,1}x_1 + \cdots + u_{i,n}x_n)^D \tag{9.1}$$

There is always a decomposition for each homogeneous polynomial, [11, Lemma 4.2]. As in the binary form case, the **rank** of an homogeneous polynomial is the minimal r such that there is a decomposition with just r summands.

A hard and interesting question that arises in this context is the determination of the generic rank. Instead of considering particular polynomials, we will analyze the expected rank for "almost all" the homogeneous polynomials of given degree. For example, if we just consider the forms in $\mathbb{C}[x_1, \ldots, x_n]_D$, there is only one expected rank for "almost all" of them.

Formally, we split $\mathbb{C}[x_1, \ldots, x_n]_D$ in subsets of polynomials where all of them have the same rank, $Z_r = \{f \in \mathbb{C}[x_1, \ldots, x_n]_D : rank(f) = r\}$. There is just one r such that Z_r is dense with the Zariski topology over $\mathbb{C}[x_1, \ldots, x_n]_D$. We say that the rank of those polynomials, r, is the **generic rank**. The determination of the generic rank was one of the most important open questions in this area up to the work of Alexander and Hirschowitz [1] in 1995. There they proved the following theorem,

Theorem 9.1. *The generic rank of a symmetric tensor of order $D > 2$ and dimension n is equal to*

$$\left\lceil \frac{1}{n} \binom{n + D - 1}{D} \right\rceil$$

Except for the following cases: $(D, n) \in \{(3, 5), (4, 3), (4, 4), (4, 5)\}$, where generic rank should be the increased by 1.

It is good to say that the generic rank of the binary forms had been solved by Sylvester. Using our approach for Hankel matrices it is easy to prove Theorem 9.1 in case of $n = 2$.

To conclude, let us talk about a potential general algorithm to decompose any symmetric tensor. The most important point to remark about this issue is that the complexity is unknown. This issue is important because, as the rank of a symmetric tensor is always bounded, it is always possible to get a minimal decomposition by solving a polynomial equation system. We can perform a binary search over the rank r and get a polynomial system from coefficients of each monomial in Equation (9.1) taking the unknowns as $u_1, \ldots, u_n$. Using Gröebner basis it is possible to solve those systems, but the complexity is too big to be affordable (just consider that every permutation of the basis leads to a different solution to that system).

Iterative algorithms as Alternate Least Squares or gradient descents have been used to solve this rank problem, but they lack of a proof for their global convergence. Extending the work of Sylvester, Brachat et al. [5] introduced a better algorithm which is efficient when the tensor to computation has a sub-generic rank which always converges. The main idea was to analyze the dual problem and to use Hankel operators. This algorithm is practically more efficient than the one proposed before, but still its complexity is unknown.

APPENDIX

A. PROOF OF THEOREM 5.2

In this appendix we prove that given a binary form f with rank (N_2+1), there is a square-free kernel polynomial such that $(N_2 - N_1 + 1)$ of its roots belong to a chosen set. For this proof we use Lagrange polynomials for interpolating univariate polynomials and the Pigeonhole principle. In this appendix, for simplicity, we consider all the binary forms as univariate polynomials.

First we prove that if we fix $(N_2 - N_1)$ of the roots, we can always get a square-free kernel polynomial whose $(N_2 - N_1 + 1)$-th root belongs to a chosen subset of $\mathbb{F}$. We find the minimal cardinal that such subset should have. Using those facts, we show what happens when $(N_2 - N_1 + 1)$ roots are chosen randomly.

Reminding the Proposition 2.4.8, the polynomials in the kernel of H^{N_2+1} can be written as $P_v \cdot P_\mu + P_w$, where P_μ is a binary form of degree $(N_2 - N_1)$. As we prove in Lemma 5.1, given $(N_2 - N_1 + 1)$ values which are not roots of P_v, there is an unique polynomial P_μ such that those values belongs to the roots of $P_v \cdot P_\mu + P_w$. Let $\beta_1, \ldots, \beta_{N_2-N_1} \in \mathbb{F} \setminus \mathit{RootsOf}(P_v)$ be $(N_2 - N_1)$ different values. Given $\alpha \in \mathbb{F} \setminus \big(\mathit{RootsOf}(P_v) \cup \{\beta_1, \ldots, \beta_{N_2-N_1}\}\big)$, we define $P_{(\alpha)}$ as the unique binary form of degree $(N_2 - N_1)$ such that $\alpha, \beta_1, \ldots, \beta_{N_2-N_1}$ are roots of the polynomial $Q_{(\alpha)}$, where

$$Q_{(\alpha)} := P_v \cdot P_{(\alpha)} + P_w$$

.

Using Lagrange polynomials we can write $P_{(\alpha)}$ as,

$$P_{(\alpha)}(x) = -\sum_{i=1}^{N_2-N_1} \frac{P_w(\beta_i)}{P_v(\beta_i)} \frac{(x-\alpha)}{\beta_i - \alpha} \prod_{j \neq i} \frac{(x-\beta_j)}{\beta_i - \beta_j} - \frac{P_w(\alpha)}{P_v(\alpha)} \prod_{i=1}^{N_2-N_1} \frac{(x-\beta_i)}{\alpha - \beta_i}$$

Lemma A.1. *Let $\alpha, \rho \in \mathbb{F} \setminus \big(\mathit{RootsOf}(P_v) \cup \{\beta_1, \ldots, \beta_{N_2-N_1}\}\big)$. $P_{(\alpha)} = P_{(\rho)}$, if and only if, $Q_{(\alpha)}(\rho) = 0$.*

Proof. If we consider the polynomial $P_{(\alpha)} - P_{(\rho)}$, its degree is at most $(N_2 - N_1)$. Note that $\beta_1, \ldots, \beta_{N_2-N_1}, \rho$ are $(N_2 - N_1 + 1)$ different roots. Hence, that polynomial is identically zero.

□

We show that there is a bound for the possibles λs such that $Q_{(\alpha)}$ is not square-free. We split the proof in two parts. Without loss of generality, in the following lemma we bound the

possibles αs such that β_1 is a square-root of $Q_{(\alpha)}$,

Lemma A.2. *There are at most* (N_1+1) *values for* $\alpha \in \mathbb{F}\setminus\big(RootsOf(P_v)\cup\{\beta_1,\ldots,\beta_{N_2-N_1}\}\big)$ *such that* β_1 *is a square-root of* $Q_{(\alpha)}$.

Proof. If β_1 is a square-root of $Q_{(\alpha)}$, then

$$\begin{cases} Q_{(\alpha)}(\beta_1) = & P_v(\beta_1)\cdot P_{(\alpha)}(\beta_1) + P_w(\beta_1) & = 0 \\ Q'_{(\alpha)}(\beta_1) = & P'_v(\beta_1)\cdot P_{(\alpha)}(\beta_1) + P_v(\beta_1)\cdot P'_{(\alpha)}(\beta_1) + P'_w(\beta_1) & = 0 \end{cases}$$

So,

$$P'_{(\alpha)}(\beta_1) = P'_v(\beta_1)\cdot\frac{P_w}{P_v}(\beta_1) - P'_w(\beta_1) \tag{A.1}$$

At the same time, we have that,

$$\begin{aligned} P'_{(\alpha)}(\beta_1) = & -\frac{P_w}{P_v}(\beta_1)\frac{1}{\beta_1-\alpha} - \frac{P_w}{P_v}(\beta_1)\sum_{j=2}^{N_2-N_1}\frac{1}{\beta_1-\beta_j} \\ & -\sum_{i=2}^{N_2-N_1}\frac{P_w}{P_v}(\beta_i)\frac{\beta_1-\alpha}{(\beta_i-\beta_1)(\beta_i-\alpha)}\prod_{j\notin\{1,i\}}\frac{\beta_1-\beta_j}{\beta_i-\beta_j} \\ & -\frac{P_w}{P_v}(\alpha)\frac{1}{\alpha-\beta_1}\prod_{j=2}^{N_2-N_1}\frac{\beta_0-\beta_j}{\alpha-\beta_j} \end{aligned}$$

We can rewrite the previous equations as

$$\begin{aligned} P'_{(\alpha)}(\beta_1) = & -A(\beta_1)\frac{1}{\beta_1-\alpha} - B(\beta_1) \\ & -\sum_{i=2}^{N_2-N_1}C_i(\beta_1)\frac{\beta_1-\alpha}{(\beta_i-\alpha)} \\ & -\frac{P_w}{P_v}(\alpha)E(\beta_1)\prod_{j=1}^{N_2-N_1}\frac{1}{\alpha-\beta_j} \end{aligned} \tag{A.2}$$

We rewrite Equation (A.1) as $P'_{(\alpha)}(\beta_1) = F(\beta_1)$ where,

$$F(\beta_1) := P_v'(\beta_1) \cdot \frac{P_w}{P_v}(\beta_1) - P_w'(\beta_1)$$

Therefore, joining Equation (A.1) and Equation (A.2),

$$F(\beta_1) = A(\beta_1)\frac{1}{\beta_1 - \alpha} - B(\beta_1) - \sum_{i=2}^{N_2-N_1} C_i(\beta_1)\frac{\beta_1 - \alpha}{(\beta_i - \alpha)} - \frac{P_w}{P_v}(\alpha)E(\beta_1)\prod_{j=1}^{N_2-N_1}\frac{1}{\alpha - \beta_j}$$

$$\begin{aligned} P_v(\alpha)(F+B)(\beta_1)\prod_{j=1}^{N_2-N_1}(\alpha-\beta_j) = {} & P_v(\alpha)A(\beta_1)\prod_{j\neq 1}(\alpha-\beta_j) \\ & + P_v(\alpha)\sum_{i=2}^{N_2-N_1} C_i(\beta_1)(\beta_1-\alpha)\prod_{j\neq i}(\alpha-\beta_j) \\ & + P_w(\alpha)E(\beta_1) \end{aligned} \tag{A.3}$$

Each sides of the last equation can be consider as univariate polynomials, where α is the variable. As the degree of both of sides of Equation (A.3) is $N_2 + 1$, if there were more than $N_2 + 1$ values for α such that β_1 is a square-root of $Q_{(\alpha)}$, both polynomials would be the same. That would mean that P_v divides P_w. By Proposition 2.4.9, we know that this is not true. Therefore, there are at most $N_2 + 1$ values for α such that β_1 is a square-root of $Q_{(\alpha)}$.

□

For each α, the square-roots of $Q_{(\alpha)}$, if any, could be a β_i or not. By Lemma A.2 we proved that just a bounded amount values of α makes β_i a square-root of $Q_{(\alpha)}$. In Lemma A.3 we show that just for a few values, α is a square-root of $Q_{(\alpha)}$.

Lemma A.3. *There are at most* $(2N_1+1)$ *values for* $\alpha \in \mathbb{F}\backslash\big(RootsOf(P_v)\cup\{\beta_1,\ldots,\beta_{N_2-N_1}\}\big)$ *such that* α *is a square-root of* $Q_{(\alpha)}$.

Proof. The proof is similar to the Lemma A.2. If α is a square-root of $Q_{(\alpha)}$, then

$$\left(P_v^2 P_{(\alpha)}'\right)(\alpha) = \big(P_v' P_w - P_w' P_v\big)(\alpha)$$

At the same time,

$$P'_{(\alpha)}(\alpha) = -\sum_{i=1}^{N_2-N_1} \frac{P_w}{P_v}(\beta_i)\frac{1}{\beta_i-\alpha}\prod_{j\neq i}\frac{\alpha-\beta_j}{\beta_i-\beta_j} - \frac{P_w}{P_v}(\alpha)\sum_i \frac{1}{\alpha-\beta_i}$$

Therefore,

$$P_v(\alpha)\left(-P_v(\alpha)\sum_{i=1}^{N_2-N_1}\frac{P_w}{P_v}(\beta_i)\prod_{j\neq i}\frac{(\alpha-\beta_j)^2}{\beta_j-\beta_i} - P_w(\alpha)\right) = \big(P'_vP_w - P'_wP_v\big)(\alpha)\prod_i(\alpha-\beta_i)$$

Once again, we can consider the equations of both sides as polynomials in α of degree $(2N_2+1)$. If there were more than $(2N_2+1)$ values for α such that this equality holds, then the polynomials would be equal. By definition of β_i, $P_v(\beta_i) \neq 0$, so P_v must divide P'_vP_w, which is not true because, by Proposition 2.4.9, P_v and P_w do not share any root.

□

Theorem A.4. *There are at most* $(N_1+1)(3N_2-N_1+1)$ *values for* $\alpha \in \mathbb{F}\setminus\big(RootsOf(P_v)\cup\{\beta_1,\ldots,\beta_{N_2-N_1}\}\big)$ *such that* $Q_{(\alpha)}$ *has square-roots.*

Proof. By the Lemma A.2, for each $1 \leq i \leq (N_2-N_1)$, there are at most N_1+1 values for α such that β_i is a square-root of $Q_{(\alpha)}$. Therefore, there are at most $(N_2-N_1)(N_1+1)$ values for α such that any β_i is a square-root of $Q_{(\alpha)}$.

Suppose that ρ is a square-root of $Q_{(\alpha)}$, and $\rho \neq \beta_i$. By Lemma A.1, $Q_{(\alpha)} = Q_{(\rho)}$. Hence, by Lemma A.3, there are at most $(2N_2+1)$ different possible values for ρ. As the polynomial $Q_{(\alpha)}$ has degree N_2+1, there are at most (N_1+1) roots of $Q_{(\alpha)}$ which are not a β_i. Therefore, there are at most $(N_1+1)(2N_2+1)$ values for $\alpha \neq \beta_i$ such that $Q_{(\alpha)}$ has square-roots different from a β_i.

Therefore, there are at most $(N_1+1)(2N_2+1)+(N_2-N_1)(N_1+1)$ values for $\alpha \in \mathbb{F}\setminus\big(RootsOf(P_v)\cup\{\beta_1,\ldots,\beta_{N_2-N_1}\}\big)$ such that $Q_{(\alpha)}$ has square-roots.

□

The Theorem A.4 gives a bound for the quantity of αs that makes $Q_{(\alpha)}$ a polynomial with square-roots. Hence, using the pigeonhole principle, if we choose the α randomly and uniformly from a set, then we can bound the probability of having a square-free polynomial $Q_{(\alpha)}$.

Corollary A.5. *Let* $\Gamma \subset \mathbb{F}\setminus\big(RootsOf(P_v)\cup\{\beta_1,\ldots,\beta_{N_2-N_1}\}\big)$ *be a finite set.*
If we choose randomly and uniformly an element $\alpha \in \Gamma$*, we can bound the probability of getting a square-free kernel polynomial by*

$$Prob\left(Q_{(\alpha)} \text{ is a polynomial square-free} \mid \alpha \in \Gamma\right) \geq \frac{\#\Gamma - (N_1+1)(3N_2 - N_1 + 1)}{\#\Gamma}$$

Up to now, we assumed that the $\beta_1, \ldots, \beta_{N_2-N_1}$ are fixed. Appendix A bounds the probability of getting a square-free polynomial where all the $\alpha, \beta_1, \ldots, \beta_{N_2-N_1}$ are chosen randomly and uniformly. Let $\Lambda \subseteq \mathbb{F}$ be a finite set whose cardinal is $(N_2 - N_1 + 1)$. We define $P_{(\Lambda)}$ as the unique polynomial such that $\Lambda \subseteq RootsOf(Q_{(\Lambda)})$, where $Q_{(\Lambda)} := P_v \cdot P_{(\Lambda)} + P_w$.

Theorem (5.2)**.** *Let $\Gamma \subset \mathbb{F} \setminus \left(RootsOf(P_v) \cup \{\beta_1, \ldots, \beta_{N_2-N_1}\}\right)$ be a finite set.*
If we choose randomly and uniformly a set $\Lambda \subseteq \Gamma$ whose cardinal is $(N_2 - N_1 + 1)$, then the probability that $Q_{(\Lambda)}$ is square-free is bounded by,

$$Prob\left(Q_{(\Lambda)} \text{ is a square-free polynomial} \mid \Lambda \subseteq \Gamma\right) \geq 1 - \frac{(N_1+1)(3N_2 - N_1 + 1)}{\#\Gamma - N_2 + N_1}$$

Proof. By Theorem A.4, for each set $\overline{\Lambda} \subseteq \Gamma$ with cardinal $N_2 - N_1$, there are at most $(N_1 + 1)(3N_2 - N_1 + 1)$ different values for $\lambda \in \Gamma$, such that $Q_{\left(\overline{\Lambda} \cup \{\alpha\}\right)}$ has square-roots. Hence, there are at most the possibles Λ such that $Q_{(\Lambda)}$ has square-roots is bounded by,

$$\#\{\Lambda \mid Q_{(\Lambda)} \text{ has square-roots } \} \leq \binom{\#\Gamma}{N_2 - N_1}(N_1+1)(3N_2 - N_1 + 1)$$

Note that this bound is not tight because we are considering the same sets $(N_2 - N_1 + 1)$ times. If $Q_{(\Lambda)}$ has square-roots, with $\Lambda = \{\gamma_0, \ldots, \gamma_{N_2-N_1}\}$, then we are counting this set for each subset $\Lambda_i = \{\gamma_0, \ldots, \gamma_{i-1}, \gamma_{i+1}, \ldots, \gamma_{N_2-N_1}\}$ because $Q_{(\Lambda_i \cup \{\gamma_i\})}$ has always square-roots. This way, a tighter bound is,

$$\#\{\Lambda \mid Q_{(\Lambda)} \text{ has square-roots } \} \leq \frac{\binom{\#\Gamma}{N_2-N_1}(N_1+1)(3N_2 - N_1 + 1)}{N_2 - N_1 + 1}$$

There are $\binom{\#\Gamma}{N_2-N_1+1}$ different possible sets. If we take each one with the same probability,

$$Prob\big(Q_{(\Lambda)} \text{ has square-roots} \mid \Lambda \subseteq \Gamma\big) \leq \frac{\binom{\#\Gamma}{N_2-N_1}(N_1+1)(3N_2 - N_1 + 1)}{\binom{\#\Gamma}{N_2-N_1+1}(N_2 - N_1 + 1)}$$

$$= \frac{(N_1+1)(3N_2 - N_1 + 1)}{\#\Gamma - N_2 + N_1}$$

□

BIBLIOGRAPHY

[1] J. Alexander and A. Hirschowitz. Polynomial interpolation in several variables. *Journal of Algebraic Geometry*, 4(2):201–222, 1995.

[2] E. R. Berlekamp. *Nonbinary BCH decoding*. University of North Carolina. Department of Statistics, 1966.

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

[4] D. Bini and V. Y. Pan. *Polynomial and matrix computations (vol. 1): fundamental algorithms*. Birkhauser Verlag, 1994.

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

[6] W. T. Bradley and W. J. Cook. Two proofs of the existence and uniqueness of the partial fraction decomposition. In *International Mathematical Forum*, volume 7, pages 1517–1535, 2012.

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

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

[9] P. Comon and C. Jutten. *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.

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

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

[12] J. v. z. Gathen. *Modern computer algebra*. Cambridge University Press, Cambridge, 2013. ISBN 9781139856065 1139856065 9781299772717 1299772714 9781107248052 1107248051. URL `http://dx.doi.org/10.1017/CBO9781139856065`.

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

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

[15] I. S. Iohvidov. *Hankel and Toeplitz Matrices and Forms*. Birkhäuser Boston, 1 edition edition, Jan 1982. ISBN 9780817630904.

[16] T. Jiang and N. D. Sidiropoulos. Kruskal's permutation lemma and the identification of candecomp/parafac and bilinear models with constant modulus constraints. *Signal Processing, IEEE Transactions on*, 52(9):2625–2636, 2004.

[17] E. Kaltofen and L. Yagati. Improved sparse multivariate polynomial interpolation algorithms. In *Symbolic and Algebraic Computation*, pages 467–474. Springer, 1989.

[18] W. Manthey, U. Helmke, and D. Hinrichsen. Topological aspects of the partial realization problem. *Mathematics of Control, Signals and Systems*, 5(2):117–149, 1992.

[19] J. L. Massey. Shift-register synthesis and bch decoding. *Information Theory, IEEE Transactions on*, 15(1):122–127, 1969.

[20] B. Reznick. On the length of binary forms. *arXiv:1007.5485 [math]*, Jul 2010. URL `http://arxiv.org/abs/1007.5485`. arXiv: 1007.5485.

[21] S. Sahnoun and P. Comon. Tensor polyadic decomposition for antenna array processing. In *21st International Conference on Computational Statistics (CompStat'2014)*, 2014.

[22] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, Oct. 1980. ISSN 0004-5411. doi: 10.1145/322217.322225. URL `http://doi.acm.org/10.1145/322217.322225`.

[23] Sylvester. An essay on canonical forms. In *The collected mathematical papers of James Joseph Sylvester*, volume 1, pages 203–216. Cambridge University Press, 1904.

[24] R. Zippel. Probabilistic algorithms for sparse polynomials. In *Proceedings of the International Symposiumon on Symbolic and Algebraic Computation*, EUROSAM '79, pages 216–226, London, UK, UK, 1979. Springer-Verlag. ISBN 3-540-09519-5. URL `http://dl.acm.org/citation.cfm?id=646670.698972`.