On symmetries of and equivalence tests for two polynomial families and a circuit class

A THESIS

SUBMITTED FOR THE DEGREE OF

Doctor of Philosophy

IN THE

Faculty of Engineering

BY

Nikhil Gupta

Computer Science and Automation
Indian Institute of Science
Bangalore – 560 012 (INDIA)

August, 2022

Declaration of Originality

I, **Nikhil Gupta**, with SR No. **04-04-00-14-12-17-1-15373** hereby declare that the material presented in the thesis titled

On symmetries of and equivalence tests for two polynomial families and a circuit class

represents original work carried out by me in the **Department of Computer Science and Automation** at **Indian Institute of Science** during the years **2017-2022**.
With my signature, I certify that:

- I have not manipulated any of the data or results.
- I have not committed any plagiarism of intellectual property. I have clearly indicated and referenced the contributions of others.
- I have explicitly acknowledged all collaborative research and discussions.
- I have understood that any false claim will result in severe disciplinary action.
- I have understood that the work may be screened for any form of academic misconduct.

Date: 29-08-2022 Student Signature

In my capacity as supervisor of the above-mentioned work, I certify that the above statements are true to the best of my knowledge, and I have carried out due diligence to ensure the originality of the report.

Advisor Name: Chandan Saha Advisor Signature

DEDICATED TO

My parents

for their love and support

Acknowledgements

First and foremost, I thank my advisor, Prof. Chandan Saha, for everything he has done for me in the last seven years at IISc. His continuous encouragement to think independently and ask relevant questions while doing the research, and his advice on improving my writing and presentation skills have helped me a lot during my Ph.D. journey. I am thankful for all the lessons I have learned while working with him. I admire his way of doing research, writing and presenting technical content, collaborating with others, curiosity for learning new things, helpful nature, tolerance, and patience. I also thank him for all the help and guidance I have received from him in various non-academic matters. I am indebted for all the valuable advice I got from him, which helped me improve my personal and professional fronts.

I thank Prof. Dilip Patil for teaching me various subjects in mathematics like linear algebra, abstract algebra, commutative algebra, and algebraic geometry. Whatever little I know in these areas of mathematics is mainly because of Prof. Patil. The topics I learned from him in several formal and informal sessions have helped me tremendously in my research. I am grateful for everything I have received from Prof. Patil. I also want to thank all the outstanding teachers of IISc whose courses I have attended in the last seven years.

I thank my collaborators Ankit Garg, Neeraj Kayal, Chandan Saha, and Bhargav Thankey. I have learned a lot from each one of them about how to ask relevant questions when solving a research problem, collaborate, and present the technical content in the simplest possible way. I also thank Anuj Tawari for sitting through my presentations on a special case of equivalence test for the Nisan-Wigderson polynomial and providing valuable feedback and suggestions.

I am thankful to the organizers of the 'workshop in algebraic complexity theory (WACT) 2019', the 'inter-research-institute student seminar (IRISS) 2020', and the 'IISc EECS research students symposium 2021' for inviting me to present our work. I also thank the organizers of the workshop on 'sensitivity, query complexity, communication complexity and Fourier analysis of Boolean function 2020' for inviting me to attend the workshop.

I thank my lab mates - Vineet Nair, Sumant Hedge, Abhijat Sharma, Janaky Murthy, Bhargav Thankey, Anuj Tawari, Arpita Korwar, and Agrim Dewan - for several insightful

discussions on various academic and non-academic matters. My special thanks to Vineet and Bhargav. Vineet has always been there for me in the last seven years and has helped me in various situations. I have always cherished his company and have had several memorable moments with him. I thank him for everything. I also thank Bhargav for being a wonderful friend. I adore his way of solving a research problem, and his ability to come up with simple proofs is worth admiring. I have learned many things while working with him. I also thank Vishakha Patil for being a fantastic friend. It is rare to get friends like her. I would also like to thank my friends Anand Krishna and Shravani Patil. I will always relish the beautiful time spent with Vineet, Vishakha, Anand, and Shravani. I thank Vipul Arora, Protik Paul, Philips George John, Shuprovat Ghoshal, Saravanan Kandasamy and my other friends from the CSA department for their amazing company. I also thank the CSA office for providing all the help during my stay at IISc. I thank the housekeeping staff, mess workers, medical staff, security guards and the administration of IISc for working tirelessly to make our stay at IISc safe during the spread of COVID-19 pandemic.

I thank my music teacher Smt. Geeta Ananth for teaching me Hindustani classical music. I am grateful for all the valuable lessons I learned about music and life from her. I also thank her, Prof. Ananth Ramaswamy, and the coordinators of the SPICMACAY Bangalore chapter for providing an opportunity to be a part of SPICMACAY. Because of them, I had the privilege to meet eminent musicians like Pandit Rajan and Sajan Mishra, Pandit Vikku Vinayakram, Ustad Wasifuddin Dagar, and Pandit Vishwa Mohan Bhatt, to name a few.

I am thankful to all my friends who made my stay at IISc enjoyable. I thank Ishan Rastogi, Parth Verma, Lokesh Mohan, Anubhav Guleria, Ashutosh Mohanty, Jagabandhu Sahoo, Somnath Arjun, Ashish Tolambia, Anupam Bhim, Preetam Kumar, Bhardwaj Pandit, Santosh Wupadrshta, for their company. I have learned a lot from each one of them. I thank Srila Prabhupada for his valuable teachings on spirituality, which were immensely helpful during the tough phases of my Ph.D. journey. I am thankful to Shweta Makhija for trusting and supporting me in every situation. Her constant motivation to always do better has helped me substantially during my Ph.D. Finally, I thank my family for their love and constant support. I thank my parents for always being there for me and for always encouraging me to be a good human being. Whatever I am today is because of them. I dedicate this piece of work to them.

Abstract

Two polynomials $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ over a field $\mathbb{F}$ are said to be *equivalent* if there exists an $n \times n$ invertible matrix A over $\mathbb{F}$ such that $g = f(A\mathbf{x})$, where $\mathbf{x} = (x_1 \cdots x_n)^T$. The *equivalence test* (in short, ET) for a polynomial family $\{f_m\}_{m\in\mathbb{N}}$ (similarly, a circuit class $\mathscr{C}$) is the following algorithmic problem: Given input black-box access to $g \in \mathbb{F}[x_1, \ldots, x_n]$, determine whether there exists an $f \in \{f_m\}_{m\in\mathbb{N}}$ (respectively, a circuit $\mathsf{C} \in \mathscr{C}$) such that $g = f(A\mathbf{x})$ (respectively, $g = \mathsf{C}(A\mathbf{x})$) for some $n \times n$ invertible matrix A over $\mathbb{F}$. If the answer is yes, it also outputs an $f \in \{f_m\}_{m\in\mathbb{N}}$ (respectively, a circuit $\mathsf{C} \in \mathscr{C}$) and an $n \times n$ invertible *certificate* matrix A over $\mathbb{F}$ such that $g = f(A\mathbf{x})$ (respectively, $g = \mathsf{C}(A\mathbf{x})$). In this thesis, we study equivalence tests for two polynomial families, namely the families of Nisan-Wigderson design polynomials (in short, NW) and determinant, and a circuit class, namely the class of *regular* read-once arithmetic formulas. In the process of designing ET for NW, we prove some fundamental structural and algorithmic results related to the *symmetries* of NW, namely *characterization by symmetries*, *characterization by circuit identities*, a *circuit testing* algorithm and a *flip theorem*. An invertible matrix A is called a symmetry of NW if $\mathsf{NW} = \mathsf{NW}(A\mathbf{x})$.

In the first work, we study some useful properties of the symmetries of NW. NW is an important polynomial in algebraic complexity theory (ACT) as it has been used to prove lower bounds for various classes of arithmetic circuits. Similar to NW, other polynomials like the determinant, the permanent, the IMM, etc. have also been used in many lower bound proofs in ACT. Unlike these polynomials, which are well-studied, not much is known about NW. The family of NW is in VNP but it is not known whether it is in VP and or is VNP-complete. In this work, we fill in some gaps in our understanding of NW by answering certain interesting questions related to the symmetries of NW. These questions are quite relevant from the context of geometric complexity theory and have been studied for the permanent. We show that NW is *characterized by its symmetries* over $\mathbb{C}$ but not over $\mathbb{R}$ and $\mathbb{Q}$. Using the symmetries of NW, we show that NW is *characterized by circuit identities* over any field. By exploiting the second property, we give a randomized polynomial time *circuit testing* algorithm and a *flip theorem* for NW. A circuit testing algorithm checks whether a given circuit computes NW and hence

is a natural special case of ET for NW. A circuit testing algorithm is also required for the ET for NW. We give a randomized polynomial time reduction from general ET for NW to the *block-permuted* ET for NW. Further, we also give a randomized polynomial time algorithm for a special case of block-permuted ET for NW, which we call *block-diagonal permutation scaling* ET for NW. These structural and algorithmic results crucially use some special symmetries of NW as well as the structure of the *group of symmetries* of NW, denoted $\mathscr{G}_{\mathsf{NW}}$. The structure of $\mathscr{G}_{\mathsf{NW}}$ was studied in the author's master's thesis [Gup17] and is not included in this thesis.

In the second work, we study ET for the family of determinant (in short, DET) over finite fields and over $\mathbb{Q}$. A randomized polynomial time DET over $\mathbb{C}$ was given in [Kay12]. A randomized polynomial time DET over a finite field $\mathbb{F}_q$ was given in [KNS19], which outputs a certificate matrix over a degree n extension field of $\mathbb{F}_q$, provided the input polynomial is equivalent to the $n \times n$ determinant, denoted Det_n. In this work, we give a randomized polynomial time DET over $\mathbb{F}_q$, which outputs a certificate matrix over the base field. We also give the *first* randomized DET over $\mathbb{Q}$, which takes oracle access to an integer factoring algorithm (IntFact), and outputs a certificate matrix over $\mathbb{Q}$. This DET runs in polynomial time in the Turing machine model if n is bounded. If we remove oracle access to IntFact from DET over $\mathbb{Q}$, then we get a polynomial time randomized DET for every n, but it outputs a certificate matrix over an extension field $\mathbb{L}$ of $\mathbb{Q}$, where $[\mathbb{L} : \mathbb{Q}] \leq n$. The heart of these algorithms is a randomized polynomial time reduction from DET to the *full matrix algebra isomorphism* (FMAI) problem. This reduction exploits the rich structure of the Lie algebra of the determinant and works over almost every field. FMAI is a well-studied problem in computer algebra and FMAI algorithms are known over finite fields and $\mathbb{Q}$. We prove that assuming the Generalized Riemann Hypothesis, there exists a randomized polynomial time reduction from integer factoring to DET for quadratic forms over $\mathbb{Q}$ (i.e., $n = 2$ case). This shows that it is unlikely to get rid of the IntFact oracle from DET over $\mathbb{Q}$. We also give a reduction from FMAI to DET over almost every field, which is efficient if n is bounded. This shows that FMAI and DET are randomized polynomial time reducible to each other whenever n is bounded.

In the third work, we give the *first* randomized polynomial time equivalence test with oracle access to *quadratic form equivalence* (QFE) for the class of *regular* read-once arithmetic formulas (in short, regular ROFs). An arithmetic formula C over a field $\mathbb{F}$ is said to be read-once if every leaf node of C is labelled by either a distinct variable or a constant from $\mathbb{F}$. ROFs are well-studied in the literature. An ROF C is called regular if every variable in C is a child of a $\times$ gate. Thus, the class of regular ROFs is a natural subclass of ROFs. An ET for regular ROFs significantly generalizes QFE over $\mathbb{C}$ and ET algorithms for two previously studied sub-classes of regular ROFs, namely the classes of *sum-product polynomials* and *ROANFs*. Equivalence tests

for these two classes were given recently in [MS21]. Our ET algorithm is based on some useful properties of the Hessian determinant of a regular ROF like its non-zeroness, knowledge of its factors and its essential variables. The arbitrary nature of the underlying tree of a regular ROF makes the analysis of the above mentioned properties of its Hessian determinant technically challenging. We overcome this challenge by studying the structures and coefficients of some *nice monomials* in the Hessian determinant of a regular ROF.

Publications based on this Thesis

1. On the symmetries of and equivalence test for design polynomials,
   Joint work with Chandan Saha,
   Proceedings of 44th International Symposium on Mathematical Foundations of Computer Science (MFCS), 2019.

2. Determinant equivalence test over finite fields and over $\mathbb{Q}$,
   Joint work with Ankit Garg, Neeraj Kayal and Chandan Saha,
   Proceedings of 46th International Colloquium on Automata, Languages and Programming (ICALP), 2019.

3. Equivalence test for read-once arithmetic formulas,
   Joint work with Chandan Saha and Bhargav Thankey,
   Under submission. Available from Electronic Colloquium on Computational Complexity (ECCC), report number TR22-099.

Contents

Chapter 1

Introduction

Isomorphism plays an important role in mathematics. Two mathematical objects A and B of the "same type" are said to be isomorphic if there exists a structure preserving bijective map between A and B. Isomorphisms of various algebraic objects like groups, rings, fields, vector spaces, modules, algebras etc. are well-studied. Let us see an example of an isomorphism between two groups. Let $\mathbb{Z}$ be the set of integers, a be a fixed non-zero integer, and $a\mathbb{Z}$ be the set of integer multiples of a. Then, $(\mathbb{Z}, +)$ and $(a\mathbb{Z}, +)$ are groups (Definition 2.1) under integer addition and hence have the same type. Further, $\varphi : \mathbb{Z} \to a\mathbb{Z}\ ;\ b \mapsto ab$ is a *group isomorphism*, i.e., φ is bijective and for every $b, c \in \mathbb{Z}, \varphi(b + c) = \varphi(b) + \varphi(c)$. Now, we give an example of a vector space isomorphism. Let $\mathbb{F}$ be a field (Definition 2.3). Then, $\mathbb{F}^n$ is an $\mathbb{F}$-vector space (Definition 2.5). Let A be an $n \times n$ invertible matrix over $\mathbb{F}$. Then, $\varphi : \mathbb{F}^n \to \mathbb{F}^n; \mathbf{a} \mapsto A\mathbf{a}$ is a vector space isomorphism from $\mathbb{F}^n$ to $\mathbb{F}^n$, i.e., φ is bijective and for every $\mathbf{a}, \mathbf{b} \in \mathbb{F}^n, \alpha, \beta \in \mathbb{F}, \varphi(\alpha\mathbf{a} + \beta\mathbf{b}) = \alpha\varphi(\mathbf{a}) + \beta\varphi(\mathbf{b})$.

Computational problems related to isomorphism of two similar objects have been studied. For example, the graph isomorphism problem (in short, GI), which determines whether two graphs are same up to permutation of vertices or not, is one of the most important and well-studied algorithmic problem in theoretical computer science. An extensive research spanning several decades on getting an efficient algorithm for GI culminated in a quasi-polynomial time algorithm given by Babai [Bab16], which is one of the breakthroughs of the last decade. Similarly, one can ask if given two objects of the same type, can we determine algorithmically if these are isomorphic? This thesis studies a similar question pertaining to polynomials.

We say that $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ are *isomorphic* (or *equivalent*) if there exist n linearly independent linear forms $\ell_1, \ldots, \ell_n \in \mathbb{F}[x_1, \ldots, x_n]$ such that $f = g(\ell_1, \ldots, \ell_n)$. In other words, there exists an $n \times n$ invertible matrix A over $\mathbb{F}$ such that $f = g(A\mathbf{x})$, where $\mathbf{x} = (x_1 \cdots x_n)^T$ is a column vector. The problem of testing whether two polynomials given as lists of coefficients

are isomorphic is known as the *polynomial equivalence* problem (in short, PE) and is the central theme of this thesis. To set up the context for PE, we briefly talk about polynomials in Section 1.1.1, about *algebraic complexity theory* (ACT) in Section 1.1.2, and about the four main problems in ACT and their relationships with PE in Sections 1.1.3 - 1.1.6.

1.1 Background

1.1.1 Polynomials

Polynomials are extensively used in mathematics. Apart from having numerous applications in various branches of mathematics, polynomials are also widely used in theoretical computer science. For example, the Fourier expansion of a Boolean function f, which is a multilinear polynomial that agrees with f on the Boolean hypercube, plays an important role in the analysis of f (see [O'D14]). Polynomials have been used in the proof of $\mathsf{IP} = \mathsf{PSPACE}$ given in [Sha92]. Polynomials have also been instrumental in answering some of the long standing open questions in combinatorics. For example, a beautiful proof of a near optimal lower bound on the size of *Kakeya sets in finite fields* given by Dvir (see [Dvi08]) is based on the *polynomial method*. We direct interested readers to [Gut16, Dvi12, Tao13] for more applications of the polynomial method in various other problems.

Polynomials also appear in algorithms for algebraic and number theoretic problems. Algorithms for such problems have a rich history. These algorithms can be classified into three categories. The first category contains the algorithms whose outputs are polynomial functions in their inputs, the algorithms in the second category use polynomial functions in the intermediate stages but their outputs are not polynomial functions in their inputs, and the third category consists of algorithms whose inputs and outputs are polynomials.

An interesting example from the first category is a matrix multiplication algorithm. The output of such an algorithm is the product of two input matrices A and B, where every entry of $A \cdot B$ is a quadratic polynomial in the entries of A and B. A long line of research on matrix multiplication algorithms given in [Str69, CW90, LG12, Wil12, DS13, CU13, LG14, AW21] aims to understand the exact complexity of this problem. Another example in this category is an algorithm for computing the determinant of a square matrix A. Using the Leibniz formula, the determinant of A can be expressed as a polynomial function in the entries of A. Efficient parallel algorithms are known for determinant computation [Csa76, Ber84, Pip22].

The second category contains many interesting number theoretic algorithms. For example, [SS71] gave an integer multiplication algorithm, which encodes integers as univariate polynomials and uses an algorithm to multiply two univariate polynomials. This polynomial mul-

tiplication algorithm uses Fast Fourier Transform (FFT), for which an efficient algorithm is known [CT65]. Algorithms for integer multiplication problems have been extensively studied [SS71, Fü09, DKSS08] and a recent result by [HvdH21] led to an $O(n \log n)$ time algorithm to multiply two n bit numbers. Another example in this category is primality testing. A deterministic polynomial time algorithm [AKS02] and several randomized polynomial time algorithms [SS77, Rab80, GK86, AB03] for this problem are known. The algorithms of [AB03, AKS02] are based on testing some polynomial identities. Polynomials also appear in solving special instances of the *sum of square roots problem* (see [KS12]).

An important example in the last category is a polynomial factorization algorithm. A polynomial time randomized algorithm for factoring univariate polynomials over finite fields [CZ81], a polynomial time deterministic univariate polynomial factorization algorithm over rational numbers [LLL82a], and randomized polynomial time reduction from factorization of multivariate polynomials to univariate polynomial factorization [Kal87, Kal89, KT90] are known.

These algorithms have many applications in complexity theory (see [Sha92, GG13, Aar16]), cryptography (see [GG13, Koe21]), coding theory (see [GG13]) etc. We direct the interested reader to [GG13, Coh03, Sho05, Koe21] for an exhaustive exposition to algebraic and number theoretic algorithms. Designing efficient algorithms involving polynomials is one of the main objectives of *computer algebra* and *algebraic complexity theory* (ACT). PE is one such important problems in ACT. Now, we briefly review some important problems in ACT.

1.1.2 Algebraic complexity theory (ACT)

ACT is a branch of computational complexity theory, that deals with understanding the strengths and weaknesses of algebraic computation. A natural model for performing algebraic computation is given by *arithmetic circuits*. An arithmetic circuit takes input a set of variables $\mathbf{x} = \{x_1, \ldots, x_n\}$ and computes a polynomial function in $\mathbf{x}$ (or simply a polynomial in $\mathbf{x}$-variables). An arithmetic circuit is represented as a directed acyclic graph, where the leaf nodes are labelled by $\mathbf{x}$ and elements from $\mathbb{F}$ and other nodes are labelled by basic arithmetic operations such as $+, -, \times$ and $\div$. See Definition 2.35 for a formal description of an arithmetic circuit and Figure 1.1.2 for an example. In an arithmetic circuit, every arithmetic operation on field elements is done in a unit time. If the underlying graph of an arithmetic circuit is a tree then it is called an *arithmetic formula*. The following complexity measures are associated with an arithmetic circuit C: The *size* of C, which is the number of edges in C and the *depth* of C, which is the length of the longest path from an input node to the output node in C. In a certain sense, size and depth capture the serial and the parallel complexities of computing a polynomial by an arithmetic circuit respectively.

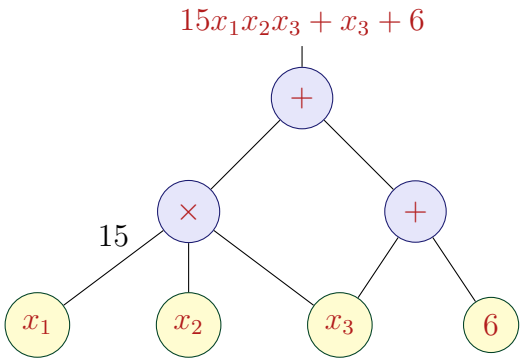


Figure 1.1: An arithmetic circuit computing $15x_1x_2x_3 + x_3 + 6$

1.1.2.1 Valiant's complexity classes

Valiant categorised families of polynomials into two classes, namely *p-computable* and *p-definable* [Val79], which are now popularly known as Valiant's P and Valiant's NP, denoted VP and VNP respectively. A polynomial family $\{f_n\}_{n\in\mathbb{N}} \in \mathsf{VP}$ if and only if for every $n \in \mathbb{N}$, f_n is an n-variate polynomial, the total degree of f_n, denoted $\deg(f_n)$, is $n^{O(1)}$ and f_n is computed by an arithmetic circuit of size $\text{poly}(n)$, where $\text{poly}(n) = n^{O(1)}$. Examples of interesting polynomial families in VP are the families of power symmetric polynomials, elementary symmetric polynomials, determinant polynomials, iterated matrix multiplication (IMM) polynomials.

A polynomial family $\{f_n\}_{n\in\mathbb{N}} \in \mathsf{VNP}$ if and only if there exist $\{h_n\}_{n\in\mathbb{N}} \in \mathsf{VP}$ and a polynomial function $t : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$, f_n is an n-variate polynomial and

$$f_n(x_1, \ldots, x_n) = \sum_{(e_1,\ldots,e_{t(n)})\in\{0,1\}^{t(n)}} h_{n+t(n)}(x_1, \ldots, x_n, e_1, \ldots, e_{t(n)}).$$

Observe that $\mathsf{VP} \subseteq \mathsf{VNP}$. We might sometime abuse the notation and say that a polynomial f is in VP (or VNP), which would mean that there exists a polynomial family $\{f_n\}_{n\in\mathbb{N}}$ in VP (respectively, in VNP) such that $f = f_n$ for some $n \in \mathbb{N}$. It is known that a family $\{f_n\}_{n\in\mathbb{N}}$ is in VNP if there exists an algorithm that takes input $(e_1, \ldots, e_n) \in \mathbb{N}^n$ and outputs the coefficient of $x_1^{e_1} \cdots x_n^{e_n}$ in f_n in $\text{poly}(n)$ time [Val82, Bür00] [1]. This is popularly known as *Valiant's criterion*. It follows immediately from this criterion that the permanent of an $n \times n$ symbolic matrix is in VNP. It was shown in [Str73c, HY11] that over any field $\mathbb{F}$, if a degree d polynomial $f \in \mathbb{F}[\mathbf{x}]$ is computed by an arithmetic circuit C of size s then there exists an arithmetic circuit C' for f such that the size of C' is $\text{poly}(s, d)$ and C' does not contain nodes labelled with $\div$

[1] In fact, if the coefficient of $x_1^{e_1} \cdots x_n^{e_n}$ in f_n can be computed in $\#\mathsf{P}/\mathsf{poly}$ then also $\{f_n\}_{n\in\mathbb{N}}$ is in VNP (see Proposition 2.20 of [Bür00]).

operations[1]. Note that a node labelled with $-$ operation in C' can be replaced by a $+$ gate and the labels of the edges going out of this $+$ gate are multiplied with -1. Henceforth, we will only be interested in the polynomial families in VP and VNP, where the degree of the n-th member in a polynomial family is poly(n). Further, we will assume from now on that non-leaf nodes in an arithmetic circuit are labelled by $+$ and $\times$.

Similar to the concept of reductions in Boolean complexity theory, we have a notion of *projections* in ACT. A polynomial family $\{f_n\}_{n\in\mathbb{N}}$ over a field $\mathbb{F}$ is said to be a *p-projection* of $\{g_n\}_{n\in\mathbb{N}}$ over $\mathbb{F}$ if there exists a polynomial function $t : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$, f_n and g_n are n-variate polynomials and there exists $m \leq t(n)$ such that $f_n(x_1, \ldots, x_n) = g_m(a_1, \ldots, a_m)$, where every $a_i \in \mathbb{F} \cup \{x_1, \ldots, x_n\}$. This notion of p-projection is used to define the concept of completeness in ACT. A family $\{f_n\}_{n\in\mathbb{N}} \in \mathsf{VNP}$ is said to be VNP-complete if every $\{g_n\}_{n\in\mathbb{N}} \in \mathsf{VNP}$ is a p-projection of $\{f_n\}_{n\in\mathbb{N}}$. Valiant showed in [Val79] that the family of permanent is VNP-complete over every field not having characteristic equal to two.[2] Some natural VNP-complete polynomial families corresponding to graphs are given in [Bür00].

Observe that the permanent family is one of the 'hardest' polynomial families in VNP as showing that this family is in VP would immediately imply $\mathsf{VP} = \mathsf{VNP}$. Valiant conjectured that the permanent can not be computed by an arithmetic circuit of polynomial size over any field having characteristic other than two. Proving this would immediately imply that VP is a strict subset of VNP. The VP versus VNP question has been a long standing open problem from more than four decades. It is not only the holy grail of ACT but is also one of the most important open questions in theoretical computer science. Apart from being a natural and important problem, it is also related to the non-uniform version of the P versus NP problem. It was shown in [BÖ0, Bür00] that if $\mathsf{VP} = \mathsf{VNP}$ over finite fields then $\mathsf{P/poly} = \mathsf{NP/poly}$. The same result holds over infinite fields assuming the Generalised Riemann Hypothesis [BÖ0, Bür00]. However, the converse of this is not known. Thus, VP versus VNP can be considered as a stepping stone for the $\mathsf{P/poly}$ versus $\mathsf{NP/poly}$ and it is hoped that the complete understanding of the exact relationship between VP and VNP might shed some light on the $\mathsf{P/poly}$ versus $\mathsf{NP/poly}$ problem. One of the promising approaches to understand VP versus VNP is *geometric complexity theory*.

1.1.2.2 Geometric complexity theory (GCT)

GCT is an approach that aims to resolve the VP versus VNP conjecture with the help of advanced tools and techniques from algebraic geometry and representation theory. It was proposed by Mulmuley and Sohoni [MS01]. Its one of the main objectives is to separate the

[1] Strassen's argument works over fields having sufficiently large size. This constraint was removed in [HY11].

[2] Over the fields of characteristic equal to two, the permanent and the determinant of an $n \times n$ symbolic matrix are the same. Hence, over such fields, the permanent is in VP.

complexities of the determinant and the permanent. GCT aims to show this by proving that *(padded)* permanent of an $n \times n$ matrix, denoted Perm_n^* [1], is not an *affine projection* of the determinant of an $m \times m$ matrix, denoted Det_m, where $m = \text{poly}(n)$ [2]. To understand whether Perm_n^* is an affine projection (Definition 2.33) of a poly(n) size determinant, GCT considers the *orbit closures*[3] of Perm_n^* and Det_m. If one can show that Perm_n^* is not present in the orbit closure of Det_m, for $m = \text{poly}(n)$ then $\mathsf{Perm_n}^*$ is not an affine projection of Det_m. This is because it is a well-known fact that over fields of characteristic zero, affine projections of a polynomial is contained in its orbit closure (see Appendix F of [ST21] for a proof of this fact). Since orbit closures are algebraic varieties, tools from algebraic geometry are potentially useful here. GCT hopes to show that Perm_n^* is not contained in the orbit closure of Det_m for $m = \text{poly}(n)$, by exploiting the *characterisation by symmetries* property (Definition 2.24) possessed by these two polynomials. This property lies at the heart of GCT and it also avoids the *natural proof barriers*. We talk about this property in Section 1.3.1. We direct the interested reader to Chapter 3 of [Gro12] and Section 6.6 of [Aar16] for introductory level exposition to GCT.

GCT suggests to study some algorithmic problems to gain more structural insights on orbit closures of the permanent and the determinant. A natural algorithmic problem in the context of understanding whether the permanent is in the orbit closure of a polynomial size determinant is to test whether a polynomial f is in the orbit (Definition 2.34) of the determinant. Such a question is called *equivalence test* for the determinant. Kayal gave an equivalence test (in short, ET) for the determinant over $\mathbb{C}$ [Kay12]. In this thesis, we give ET for the determinant over $\mathbb{Q}$ and finite fields (see Section 1.3.2 and Chapter 4).

1.1.2.3 Equivalence test

In this section, we give some useful definitions related to the equivalence test for the sake of discussion on connections of ET to other important problems in ACT given in the subsequent sections. A detailed description of equivalence test is given in Section 1.2.2.

Let $\{f_n\}_{n\in\mathbb{N}}$ be a polynomial family and $\mathscr{C}$ be a circuit class. An equivalence test for $\{f_n\}_{n\in\mathbb{N}}$ (similarly, for $\mathscr{C}$) is the following algorithmic task: Given two polynomials $f(\mathbf{x})$ and $g(\mathbf{x})$ as black-boxes [4] where $f \in \{f_n\}_{n\in\mathbb{N}}$ (respectively, f is computed by a circuit in $\mathscr{C}$), determine if

[1] Let Perm_n denote the permanent of an $n \times n$ symbolic matrix X. Then, $\mathsf{Perm}_n^* = z^{m-n}\mathsf{Perm}_n$, where z is a fresh variable not appearing in X.

[2] The non-padded version of the permanent versus the determinant problem is as follows: Is $\mathsf{Perm}_n(\mathbf{x}) = \det(B)$, where B is an $m \times m$ matrix, where $m = \text{poly}(n)$ and the entries of B are affine forms in $\mathbf{x}$-variables? As it is more convenient to deal with homogeneous polynomials, the padded version of the permanent versus determinant problem is studied.

[3] The orbit closure of an n-variate degree d polynomial $f \in \mathbb{C}[\mathbf{x}]$ is the Zariski closure of the orbit of f (Definition 2.34), where polynomials in the orbit of f are identified with their coefficient vectors in $\mathbb{C}^{\binom{n+d}{d}}$.

[4] A black-box of a polynomial $f \in \mathbb{F}[\mathbf{x}]$ takes input an $\mathbf{a} \in \mathbb{F}^{|\mathbf{x}|}$ and outputs $f(\mathbf{a})$.

there exists an invertible matrix A such that $g = f(A\mathbf{x})$. If yes, output an invertible matrix A such that $g = f(A\mathbf{x})$. Hence, ET is a special case of PE, where one of the two input polynomials comes from either a specific polynomial family or a fixed circuit class. In many cases of ET for a polynomial family, for example the families of the determinant and the permanent, there exists a unique $f \in \{f_n\}_{n\in\mathbb{N}}$ such that the number of variables in f is equal to the number of variables in g. Thus, in this case, the polynomial f is implicit and we can only give g as an input to ET for $\{f_n\}_{n\in\mathbb{N}}$. But this is not the case with ET for a circuit class $\mathscr{C}$ because there can be many circuits in $\mathscr{C}$ having the same number of variables as in g. Thus, if we only give g as the input to ET for $\mathscr{C}$, the problem becomes 'harder' than the usual ET for $\mathscr{C}$. This is so because now along with finding an invertible matrix A, the algorithm also has to find a circuit $\mathsf{C} \in \mathscr{C}$ such that $g = \mathsf{C}(A\mathbf{x})$. Thus, this version of ET generalizes the *reconstruction problem* for $\mathscr{C}$ (see Section 1.1.5). Henceforth, we only consider the 'harder version' of ET for $\mathscr{C}$. In this thesis, we study equivalence tests for two polynomial families, namely the families of determinant and *Nisan-Wigderson design polynomial*, and a circuit class, namely the class of *regular read-once arithmetic formulas* (ROFs). The details are given in Section 1.3.

In Sections 1.1.3, 1.1.4 and 1.1.5, we touch upon the three most important problems in ACT, namely *lower bounds*, *PIT* and *arithmetic circuit reconstruction*, and highlight their connections to ET. We give a brief survey of the progress made in these problems in Appendix A. We briefly talk about an important problem in computer algebra called *functional decomposition of polynomials* and its connection to ET in Section 1.1.6.

1.1.3 Lower bounds

Proving a super-polynomial lower bound on the size of arithmetic circuit computing a VNP-complete polynomial is the main objective of ACT. In last four decades, a lot of research has happened on lower bounds for various classes of arithmetic circuits. Although, the best known lower bound on the size of an arithmetic circuit is merely super-linear [Str73a, BS83], several strong lower bounds are known for many sub-classes of arithmetic circuits. We direct the interested reader to Section A.1 for a brief survey of the progress made in lower bounds. In this part, we present some connections of lower bounds to the equivalence testing problem.

1. ET and other algorithmic questions for polynomials used in lower bounds. Many polynomials like the permanent, the determinant, the iterated matrix multiplication polynomial (in short, IMM), the elementary symmetric polynomial, the power symmetric polynomial etc. have been used as hard polynomials in several lower bound results. Apart from these, the

Nisan-Wigderson design polynomial, denoted NW, has also been used extensively in many lower bound results. The definition of NW and a list of lower bounds results which use NW as a hard polynomial is given in Section 1.3.1. It is natural to develop a good understanding of all the polynomial families used in the lower bound proofs by studying various useful properties of these families. In this thesis, we study some interesting properties of NW.

All the families mentioned above except the family of NW are well-studied. The family of NW is in VNP (see Section 1.3.1) but it is neither known to be in VP, nor known to be VNP-complete. The family of permanent is VNP-complete over the fields of characteristic not equal to two. In the absence of a proof that $\mathsf{VP} = \mathsf{VNP}$, we have the following natural and interesting algorithmic problem: Let $\{f_n\}_{n\in\mathbb{N}}$ be a family in VNP, which is not known to be in VP and $f \in \{f_n\}_{n\in\mathbb{N}}$. Given a circuit C determine whether C computes f. Such a problem is called as the *circuit testing* problem for f. Two randomized polynomial time algorithms for circuit testing are known for the permanent [Lip89, Mul10]. In this thesis, we give a randomized polynomial time circuit testing algorithm for NW (see Theorem 1.3).

We can also ask the circuit testing question for the orbit (Definition 2.34) of a polynomial family [1]. Observe that circuit testing for the orbit of $\{f_n\}_{n\in\mathbb{N}}$ is essentially the equivalence testing problem for $\{f_n\}_{n\in\mathbb{N}}$. Randomized polynomial time equivalence testing algorithms are known for the families of permanent, determinant, IMM, power symmetric polynomial and elementary symmetric polynomial (see the subsection on known results on ET in Section 1.2.2). In this thesis, we give an interesting special case of ET for the family of NW (see Theorem 1.5).

Another interesting problem for a family $\{f_n\}_{n\in\mathbb{N}} \in \mathsf{VNP}$ not known to be in VP is a *flip theorem* defined as follows: Suppose $f \in \{f_n\}_{n\in\mathbb{N}}$ is such that it is not computable by an arithmetic circuit of size s. Can we generate a list of certificate points $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ efficiently over the underlying field, where $m = \text{poly}(s)$, such that the for every arithmetic circuit C of size s, there exists an $i \in [m]$ such that $f(\mathbf{a}_i) \neq \mathsf{C}(\mathbf{a}_i)$? Flip theorem is important from the viewpoint of GCT (see [Mul10, Gro12, Aar16]). A flip theorem is known for the permanent [Mul10, Mul11a], and in this thesis we give a flip theorem for NW (see Theorem 1.4).

2. ET for the determinant important from the perspective of GCT. As seen in Section 1.1.2.2, that understanding whether the permanent is in the orbit closure of a polynomial size determinant is enough to separate the complexities of the permanent and the determinant. An ET for the family of determinant is natural first question in this direction as it tests whether a polynomial is in the orbit of the determinant or not. In this thesis, we study ET for the family of determinant (see Section 1.3.2).

[1] The orbit of a polynomial family is the union of the orbits of every member in the family.

3. ET implies lower bounds for orbits of circuit classes. Let $\mathscr{C}$ be a circuit class. It follows from the discussion given in the previous section on the equivalence testing problem that ET for $\mathscr{C}$ is reconstruction of polynomials in the orbit of $\mathscr{C}$. It was shown in [FK09] that an randomized polynomial time reconstruction algorithm for a circuit class implies a lower bound for the same class. This result was derandomized in [Vol16]. Thus, an ET for $\mathscr{C}$ implies lower bounds for the orbit of $\mathscr{C}$. Proving an explicit lower bound for the orbit of *read-once algebraic branching programs* (ROABPs) is mentioned as an open question in [ST21] (see Section 7 of [ST21]). Following [FK09, Vol16] A randomized polynomial time equivalence test for the class of ROABPs would imply a lower bound for the orbit of ROABPs. The class of ROABPs is interesting because the affine projection of ROABPs captures algebraic branching programs.

1.1.4 Polynomial Identity Testing (PIT)

Polynomial identity testing is an algorithmic question that determines whether a given arithmetic circuit computes the identically zero polynomial. If the input is given as a list of coefficients then this is a trivial problem. The input of a PIT algorithm is of two types: either an arithmetic circuit C, in which case the algorithm has access to the whole circuit, or an oracle access (also called black-box access) to C, which outputs evaluations of C at points from the underlying field. The PIT problem in the former and the latter cases are called *white-box PIT* and the *black-box PIT* respectively. A simple polynomial time randomized algorithm is known for black-box PIT due to the Schwartz-Zippel lemma [DL78, Zip79, Sch80] (Fact 2.13) but a sub-exponential time deterministic algorithm for the same has remained elusive. PIT has been used to design many interesting algorithms like algorithms for perfect matchings in graphs [Lá79, KUW85, MVV87, FGT16, ST17], algorithms for primality testing [AB03, AKS02], an algorithm for linear matroid intersection [GT20], etc. PIT has also been used in the proof of IP=PSPACE [Sha92]. See Section A.2 of Appendix A for a brief survey of results in PIT.

Let $\{f_n\}_{n\in\mathbb{N}}$ be a polynomial family. Suppose we have a deterministic ET for $\{f_n\}_{n\in\mathbb{N}}$. In order to determine whether the input polynomial g is equivalent to some f in $\{f_n\}_{n\in\mathbb{N}}$, any *reasonable* ET algorithm would query black-box of g at points from the underlying field, which are not roots of g. Otherwise, it would get no information about g. We can obtain a deterministic PIT algorithm for the orbit of $\{f_n\}_{n\in\mathbb{N}}$ using a deterministic ET for $\{f_n\}_{n\in\mathbb{N}}$ as follows: Suppose g is the input of the PIT algorithm. Simulate ET on g. Suppose at some time, black-box of g returns a non-zero field element. Then, output 'g is not zero'. Otherwise, output 'g is zero'. As a reasonable ET for $\{f_n\}_{n\in\mathbb{N}}$ would query black-box of g at non-roots of g, provided g is non-zero, the output of the PIT algorithm is correct. In this way, a deterministic

ET for $\{f_n\}_{n\in\mathbb{N}}$ yields a deterministic black-box PIT for the orbit of $\{f_n\}_{n\in\mathbb{N}}$.

Now, lets see how an ET of a circuit class $\mathscr{C}$ implies PIT for $\mathscr{C}$. Any *reasonable* ET for $\mathscr{C}$ would output an invertible matrix and the trivial zero circuit [1], i.e., a circuit having only one node labelled with zero, provided the input polynomial g is the identically zero polynomial. Suppose we have a deterministic ET for $\mathscr{C}$, then we get a deterministic PIT for the orbit of $\mathscr{C}$ as follows: Suppose the input of the PIT algorithm is g. We run the ET for $\mathscr{C}$ on g. If the algorithm outputs a trivial zero circuit, we output 'g is zero', otherwise we output 'g is non-zero'. In this way, a deterministic ET for $\mathscr{C}$ implies a deterministic black-box PIT for the orbit of $\mathscr{C}$. However, it is possible that a randomized ET exists for $\{f_n\}_{n\in\mathbb{N}}$ but a deterministic PIT for the orbit of $\{f_n\}_{n\in\mathbb{N}}$ is not known. For instance, we know efficient randomized ET for the family of determinant over different fields but a deterministic PIT for the orbit of the family of determinant is not known (see Section 7 of [ST21]).

Recently, PIT for orbits of various classes of arithmetic circuits have been studied. Quasi-polynomial time PIT algorithms for orbits of sparse polynomials, read-once arithmetic formulas, bounded-width read-once algebraic branching programs (ROABPs) etc. were given in [MS21, ST21, BG21]. After having these black-box PIT algorithms, a natural next question to ask is whether we can also reconstruct polynomials in the orbits of the circuit classes mentioned above, which is basically ET for these circuit classes. In this thesis, we give a randomized polynomial time ET for the class of mildly restricted read-once arithmetic formulas (in short, ROFs), called *regular* ROFs .

1.1.5 Arithmetic circuit reconstruction

Reconstruction (or learning) of arithmetic circuits is the following algorithmic problem: Given black-box access to an arithmetic circuit C of size s computing an n-variate degree d polynomial f, output some arithmetic circuit C', which computes f and has size $\text{poly}(n, d, s)$. If C and C' belong to the same circuit class then the corresponding reconstruction algorithm is said to be *proper*, otherwise it is called an *improper* reconstruction algorithm. Arithmetic circuit reconstruction is an algebraic analog of the exact learning of Boolean functions given in [Ang88] and has been widely studied in the past two decades. We give some connections between the equivalence testing problem and the reconstruction problem below and give a survey of known results on arithmetic circuit reconstruction in Section A.3 of Appendix A.

Recall that an ET for a circuit class $\mathscr{C}$ takes black-box access to a polynomial $g(\mathbf{x})$, determines if g is equivalent to some circuit in $\mathscr{C}$ and if yes, computes an invertible matrix A and constructs a circuit $\mathsf{C} \in \mathscr{C}$ such that $g = \mathsf{C}(A\mathbf{x})$. An ET for $\mathscr{C}$ is more general than the

[1] Usually, the trivial zero circuit is present in every circuit class.

reconstruction problem for $\mathscr{C}$. This is so because given black-box access to a $\mathsf{C} \in \mathscr{C}$, an ET for $\mathscr{C}$ outputs an invertible matrix A and a $\mathsf{C}' \in \mathscr{C}$ such that $\mathsf{C} = \mathsf{C}'(A\mathbf{x})$. Thus, $\mathsf{C}'(A\mathbf{x}) \in \mathscr{C}$. An ET for a polynomial family $\{f_n\}_{n\in\mathbb{N}}$ can also be considered as an algorithm to reconstruct polynomials in the orbits of $\{f_n\}_{n\in\mathbb{N}}$. In the following paragraph, we show how equivalence test algorithms imply *average-case* reconstruction algorithm. An average-case reconstruction algorithm for a circuit class $\mathscr{C}$ reconstructs circuits chosen randomly from $\mathscr{C}$ according to some input distribution. See Section A.3 of Appendix A for more details.

Reconstruction algorithms from equivalence tests. We note here three instances where we obtain *average-case* reconstruction algorithms from equivalence tests. All these algorithms reconstruct random circuits satisfying *high number of variables* property. We will make this notion precise in the three cases discussed below.

1. **Depth 3 powering circuits**. A depth 3 powering circuit computes a polynomial of the type $g = \ell_1^d + \cdots + \ell_s^d$, where $d \in \mathbb{N}$ and every $\ell_i \in \mathbb{F}[x_1, \ldots, x_n]$ is a linear polynomial. If $\ell_1, \ldots, \ell_s$ are $\mathbb{F}$-linearly independent linear forms then observe that g is in the orbit of the power symmetric polynomial $f = x_1^d + \cdots + x_s^d$. Let C be a random depth 3 powering circuit - C is obtained by choosing $s \in \mathbb{N}$ and picking the coefficients of $x_1, \ldots, x_n$ in $\ell_1, \ldots, \ell_s$ independently and uniformly at random from a large enough finite subset of $\mathbb{F}$ - satisfying $n \geq s$. As C is in the orbit of f with high probability, an ET for power symmetric polynomials also serves as a reconstruction algorithm for C.

   A polynomial time ET algorithm for degree three power symmetric polynomials [Har70, LRA93], popularly known as Jennrich's algorithm. Later, [Kay11, GKP18] gave randomized polynomial time equivalence test for power symmetric polynomials. These algorithms reconstruct random depth 3 powering circuits in the high number of variables regime.

2. **Arithmetic formulas**. Recall that an arithmetic formula is an arithmetic circuit, where the underlying graph is a tree. Let C be an arithmetic formula having an arbitrary tree structure, where the layers of C are labelled alternatively with $+$ and $\times$ gates, the leaves of C are labelled with random linear forms, and the number of leaf nodes in C is upper bounded by the number of variables in C. Then, C is in the orbit of an ROF (Definition 2.38) with high probability. Thus, an ET for ROFs would imply an average-case reconstruction algorithm for arithmetic formula in high number of variables setting. In this thesis, we give a randomized polynomial time ET for the class of regular ROF (see Theorem 1.11) and in a follow-up work [GST22], we give a randomized polynomial time ET for the class of general ROFs. Both these ET algorithms work over almost all fields

and take oracle access to PE for quadratic forms over the underlying field.

3. **Algebraic branching programs (ABPs).** An ABP is described in Definition 2.36. Let C be an ABP where the underlying graph is arbitrary, the edges of C are labelled by random linear forms and the number of edges in C is upper bounded by the number of variables in C. Then, C is in the orbit of an *iterated matrix multiplication polynomial* (IMM) [1] with high probability. Thus, an ET for IMM reconstructs random ABPs satisfying the high number of variables condition. A randomized polynomial time ET algorithm was given in [KNST19]. This algorithm works over almost all fields.

1.1.6 Functional decomposition of polynomials

Let $\mathbb{F}$ be a field and $\mathbf{x} = \{x_1, \ldots, x_n\}$ be the set of variables. The *functional decomposition problem* (FDP) is as follows: Given a polynomial $f \in \mathbb{F}[\mathbf{x}]$, determine if there exists a *functional decomposition* of f, i.e., there exists a $g \in \mathbb{F}[y_1, \ldots, y_m]$ and $h_1, \ldots, h_m \in \mathbb{F}[\mathbf{x}]$ such that

$$f = g(h_1(\mathbf{x}), \ldots, h_m(\mathbf{x})).$$

If the answer is yes, output a functional decomposition of f. FDP is a well-studied problem in computer algebra having applications in many interesting problems like root finding (see [BZ85]), the *N-partition problem* (see Page 10 of [Dic89]), the *endomorphism invertibility problem* (see Page 11 of [Dic89]), design of *asymmetric cryptosystem* (see [PG97]) etc. It is easy to see that ET is a special case of FDP: Recall that in case of ET for $\{f_n\}_{n\in\mathbb{N}}$ (similarly, a circuit class $\mathscr{C}$), we are given a polynomial $g \in \mathbb{F}[x_1, \ldots, x_n]$ over a field $\mathbb{F}$ and we want to algorithmically determine whether there exists an $f \in \{f_n\}_{n\in\mathbb{N}}$ (respectively, a circuit $\mathsf{C} \in \mathscr{C}$) such that $g = f(\ell_1, \ldots, \ell_n)$ ($g = \mathsf{C}(\ell_1, \ldots, \ell_n)$), where $\ell_1, \ldots, \ell_n \in \mathbb{F}[x_1, \ldots, x_n]$ are $\mathbb{F}$-linearly independent linear forms. If the answer is yes, we have to output an $f \in \{f_n\}_{n\in\mathbb{N}}$ (respectively, a circuit $\mathsf{C} \in \mathscr{C}$) and $\mathbb{F}$-linearly independent linear forms $\ell_1, \ldots, \ell_n \in \mathbb{F}[x_1, \ldots, x_n]$ such that $g = f(\ell_1, \ldots, \ell_n)$. Thus, $f, \ell_1, \ldots, \ell_n$ (respectively, $\mathsf{C}, \ell_1, \ldots, \ell_n$) is a functional decomposition of g. Hence, ET for a polynomial family or a circuit class is a special case of the FDP problem.

The univariate version of FDP determines whether for an $f \in \mathbb{F}[x]$, there exist $g, h \in \mathbb{F}[x], deg(g) > 1$ such that $f = g(h(x))$ [2]. This version of FDP is well-studied. See Chapter 5 of [Coh03] for a detailed overview of the univariate FDP. The univariate FDP is used to solve univariate polynomial equations in many computer algebra systems (see Section 5.1 of

[1] Let $w, d \in \mathbb{N}$ and for $i \in [d], X_i = (x_{i,j})_{i,j\in[w]}$ be a formal matrix. Then, the iterated matrix multiplication polynomial, denoted $\mathsf{IMM}_{w,d}$, is defined as the $(1,1)$-th entry of $X_1 \cdot X_2 \cdots X_d$.

[2] In this case, it is important that $\deg(g) > 1$. Otherwise, every $f \in \mathbb{F}[x]$ admits a functional decomposition. For example, let $g = ax + b$ and $h = \frac{1}{a}f(x) - \frac{b}{a}$, where $a, b \in \mathbb{F}, a \neq 0$. Then, $f = g(h(x))$.

[Coh03]). Let us understand with an example how univariate FDP can be helpful in solving a polynomial equation. Let $f = x^4 - 3x^2 + 2, h = x^2$ and $g = x^2 - 3x + 2$. Then, $f = g(h(x))$. Thus, solving $f = 0$ is same as solving $g(h(x)) = 0$. As $h(x) = x^2$, from $g(h(x)) = 0$, we get $x^2 = 1$ and $x^2 = 2$. On solving this, we obtain the solutions $x = 1, -1, \sqrt{2}, -\sqrt{2}$, which are also solutions of f. The univariate version of FDP is efficiently solvable and several algorithms are known for the univariate FDP problem [AT85, BZ85, KL89, KLZ96]. In general, FDP is known to be NP-hard [Dic93]. However, efficient algorithms are known for some special cases of multivariate FDP (see [vzG90, von90, FP09b, FP09a, FvzGP10]).

1.2 Polynomial equivalence and equivalence testing

1.2.1 The polynomial equivalence problem

Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}$, $\mathbb{F}$ be a field and $f, g \in \mathbb{F}[\mathbf{x}]$. Recall that the polynomial equivalence problem (PE) is as follows: given f and g as lists of coefficients, determine if f is equivalent to g or not. Further, if the answer is yes then output an $n \times n$ invertible matrix A over $\mathbb{F}$ such that $g = f(A\mathbf{x})$. Polynomial equivalence is a natural and an important problem in ACT. It is a special case of testing whether out of the two given polynomials, one is an affine projection (Definition 2.33) of the other. Many important problems in ACT like the permanent versus determinant problem, the matrix multiplication etc. are instances of the affine projection problem (see [Kay12]). It was shown by Kayal in [Kay12] that the task of determining whether one polynomial is an affine projection of the other is NP-hard. Since PE is a special case of testing whether one polynomial is an affine projection of the other, it is natural to ask if PE is efficiently solvable over the underlying field $\mathbb{F}$.

An immediate solution for PE over $\mathbb{F}$ is obtained from an algorithm for polynomial solvability over $\mathbb{F}$ - treat the entries of A as formal variables and then solve the system of polynomial equations in the entries of A originating from $f = g(A\mathbf{x})$. The polynomial solvability has time complexity exponential in the input parameters over finite fields [HW99], over $\mathbb{R}$ [GV88] and over $\mathbb{C}$ [Ier89], and it is not even known to be decidable over $\mathbb{Q}$. However, PE could be an easier problem than polynomial solvability. It was shown in [Thi98, Sax06] that over finite fields, PE is in NP $\cap$ coAM and is unlikely to be NP-complete unless the polynomial hierarchy collapses. But over $\mathbb{C}$ and $\mathbb{R}$, the best known time complexity of PE is same as that of polynomial solvability over these fields, and PE is not even known to be decidable over $\mathbb{Q}$. PE is also related to the graph isomorphism problem. It was shown in [AS05] that graph isomorphism reduces in polynomial time to PE for cubic forms (i.e., homogeneous degree 3 polynomials) over any field.

PE for quadratic and cubic forms. Efficient PE algorithms for quadratic forms (i.e., homogeneous degree 2 polynomials) over $\mathbb{C}, \mathbb{R}$, finite fields having characteristic other than 2 and over $\mathbb{Q}$ with oracle access to integer factoring are known (see Section 2.2.3). These algorithms are based on the well-known classification results of quadratic forms over these fields (see [Ser73, Ara11]). On the contrary, it was shown in [AS05] that over any field, graph isomorphism reduces in polynomial time to PE for cubic forms. Thereafter, [AS06] showed that over any $\mathbb{F}$, *commutative* $\mathbb{F}$*-algebra isomorphism* reduces in polynomial time to PE for cubic forms over $\mathbb{F}$. In [AS05, Sax06], the converse of this was shown over the fields containing third roots of every element of $\mathbb{F}$ [1]. Recently, [GQ21] improved this result by showing that over $\mathbb{F}$ having $char(\mathbb{F}) = 0$ or ≥ 3, PE for cubic forms reduces to the $\mathbb{F}$-algebra isomorphism problem. They proved this by showing that many isomorphism problems like *group isomorphisms for p-groups, matrix space isometry, matrix space conjugacy, algebra isomorphism, trilinear form equivalence*, and PE for cubic forms (over fields having characteristic other than 2 or 3) are equivalent under polynomial time reduction. [GQT21] gave an *average-case algorithm* having running time $q^{O(n)}$ for deciding if two n-variate cubic forms $f(\mathbf{x}), g(\mathbf{x})$ are equivalent over the finite field $\mathbb{F}_q$. If f and g are equivalent, they also output an $A \in \mathrm{GL}(n, \mathbb{F}_q)$ such that $f = g(A\mathbf{x})$. Since their algorithm is average-case, it works for a large fraction of cubic forms in $\mathbb{F}_q[\mathbf{x}]$. Over $\mathbb{Q}$, it is not even known if PE for cubic forms is decidable.

PE in cryptography. Consider the following problem, known as *isomorphism of polynomials with 1 secret* (IP1S): Given tuples of polynomials $\mathbf{f} = (f_1, \ldots, f_m), \mathbf{g} = (g_1, \ldots, g_m)$, where every $f_i, g_j \in \mathbb{F}[\mathbf{x}]$, determine if there exists an invertible matrix A over $\mathbb{F}$ such that $f_i = g_i(A\mathbf{x})$ for every $i \in [m]$. Note that when $m = 1$, IP1S is same as PE. IP1S was first introduced and used in an authentication scheme by Patrin [Pat96]. This authentication scheme relies on the hardness of PE for cubic polynomials. After that, IP1S has been extensively studied in cryptography (see [BFP15] and the references therein). Recently, efficient algorithms were given in [BFP15, IQ19] for the variant of IP1S over finite fields where $m > 1$ and every polynomial in $\mathbf{f}, \mathbf{g}$ is a quadratic form[2]. For $m = 1$, this variant of IP1S is same as the PE for quadratic forms, for which efficient algorithms are known over different fields.

[1] A more general result was shown in [Sax06], which is as follows: For $d \in \mathbb{N}$, if $\mathbb{F}$ contains d-th roots of every element in $\mathbb{F}$ then PE for homogeneous degree d polynomials over $\mathbb{F}$ reduces to $\mathbb{F}$-algebra isomorphism problem.

[2] The algorithm in [BFP15] works when the quadratic forms satisfy some "regularity conditions" and the characteristic of the underlying field is not equal to 2. [IQ19] improved this result and gave an efficient algorithm over finite fields of odd size for any tuples of quadratic forms $\mathbf{f}$ and $\mathbf{g}$.

1.2.2 Equivalence testing

As PE is hard even for cubic forms, one can ask if there are interesting instances of this problem other than PE for quadratic forms, which can be solved efficiently. In this direction, Kayal initiated a new line of work in [Kay11], where one of the two input polynomials given to a PE algorithm comes from an important polynomial family $\{f_n\}_{n\in\mathbb{N}}$ and other polynomial is given as black-box. This problem is known as *equivalence test* (or ET) for $\{f_n\}_{n\in\mathbb{N}}$. We first recall the formal definition of ET from Section 1.1.2.3, then recall some of the motivations for ET discussed before, then compare PE and ET, and finally give a brief overview of the progress made in the equivalence testing problems.

ET comes in two flavours - ET for a polynomial family $\{f_n\}_{n\in\mathbb{N}}$ and ET for a circuit class $\mathscr{C}$. An ET for $\{f_n\}_{n\in\mathbb{N}}$ (similarly, $\mathscr{C}$) takes inputs as black-box access to $g(\mathbf{x})$ and $f(\mathbf{x})$, where $f \in \{f_n\}_{n\in\mathbb{N}}$ (respectively, f is computed by a circuit in $\mathscr{C}$) and determines if g is equivalent to f. If the answer is yes, then it outputs an invertible matrix A such that $g = f(A\mathbf{x})$. Thus, ET is a special case of PE. We consider the version of this problem where only black-box access to g is given as input to the ET algorithm and it has to decide if there exists an $f \in \{f_n\}_{n\in\mathbb{N}}$ (respectively, an f computed by a circuit $\mathsf{C} \in \mathscr{C}$) such that f is equivalent to g. In the case of ET for many important polynomial families, like the families of the permanent and the determinant, this version of ET is same as the original one because there is a unique f in the family which has the same number of variables as in g. So, even if f is not given as an input, the algorithm implicitly knows f. But this is not the case with ET for $\mathscr{C}$ as there can be many circuits in $\mathscr{C}$ having the same number of variables as in g. So, in this sense this version of ET for $\mathscr{C}$ is harder than the original ET for $\mathscr{C}$ and here the difficulty is twofold as the algorithm has to output an invertible matrix A and a circuit $\mathsf{C} \in \mathscr{C}$ satisfying $g = \mathsf{C}(A\mathbf{x})$. In this thesis, we consider the harder version of ET for a circuit class.

Motivation. An ET for $\{f_n\}_{n\in\mathbb{N}}$ (similarly, $\mathscr{C}$) reconstructs orbits (Definition 2.34) of polynomials in $\{f_n\}_{n\in\mathbb{N}}$ (respectively, circuits in $\mathscr{C}$). The orbit of a polynomial f, denoted $\mathrm{orb}(f)$, is contained in the set of affine projections of f. As mentioned above, affine projections of polynomials play an important role in ACT. Affine projections of some polynomial families or apparently *weak* circuit classes contain some powerful classes of circuits. For example, the class of arithmetic formulas is contained in the set of affine projections of ROFs (Definition 2.38). The set of affine projection of the iterated matrix multiplication polynomial captures the class of ABPs. One of the reasons for studying the orbit of a polynomial f is that the affine

projections of f are contained in the *orbit closure*[1] of f. Thus, the study of various properties of orbits of polynomials can possibly give us crucial insights about their orbit closures.

In this thesis, we study equivalence testing problem for a *hard* polynomial family, i.e., a polynomial family not known to be in VP; an *easy* polynomial family, i.e., the family in VP; and a circuit class, namely the class of regular ROFs. Here we recall the motivations discussed at multiple places before to study ET for such polynomial families and circuit classes. As discussed in Section 1.1.3, ET for a *hard* polynomial family generalizes the circuit testing algorithm for this family. In the absence of a proof that $\mathsf{VP} = \mathsf{VNP}$, a circuit testing algorithm for a hard polynomial family is an interesting question. In this thesis, we study ET for the family of Nisan-Wigderson polynomial. We saw in Section 1.1.2.2 that an ET for easy polynomial families like the determinant are important from the standpoint of GCT. Also, we saw in Section 1.1.5 that ET algorithms for the families of power symmetric polynomial and IMM imply average-case reconstruction algorithm in the *high number of variables* setting for the classes of depth 3 powering circuits and ABPs. In this thesis, we study ET for the family of determinant. We saw in Section 1.1.5 that an ET for ROFs gives an average-case reconstruction algorithm for arithmetic formulas in the high number of variables setting. Sub-exponential time black-box PIT algorithms for orbits of ROFs were given in simultaneous works of [MS21] and [ST21]. Apart from this, sub-exponential time black-box PIT algorithms for orbits of sparse polynomials and bounded-width ROABPs have also been studied recently [MS21, ST21, BG21]. Since the equivalence testing problem is about reconstruction of orbits, it becomes natural to ask if there exist efficient equivalence tests for the circuit classes considered in [MS21, ST21, BG21]. In this thesis, we give an equivalence test for the class of *regular* ROFs.

PE and ET. As noted above that the original versions of ET for a polynomial family or a circuit class are special instances of PE. Also, the other version of ET for a polynomial family, where only one polynomial is given as input is also a special case of PE. But the other version of ET for a circuit class is not exactly similar to PE as now the ET has to also reconstruct a circuit along with finding a certificate matrix. However, ET for $\mathscr{C}$ can be helpful in designing *PE for orbits of circuits in* $\mathscr{C}$. Let us first try to understand this with an example of PE for quadratic forms over $\mathbb{C}$ and then we formally define this problem.

Suppose f is a quadratic form over $\mathbb{C}$ such that f has no *redundant variables* (Definition 2.32) and the number of variables in f is even. Then, f is equivalent to $x_1x_2 + \cdots + x_{n-1}x_n$ over $\mathbb{C}$, where n is an even number. If f has odd number of variables, then it is equivalent to $x_1x_2 + \cdots + x_{n-2}x_{n-1} + x_n^2$ over $\mathbb{C}$. With the help of Witt's cancellation theorem [Wit37], this

[1]Recall from Section 1.1.2.2 that the orbit closure of f is the closure of $\mathrm{orb}(f)$ in the Zariski topology.

case can be reduced to the case when n is even [1]. A PE for quadratic forms over $\mathbb{C}$ takes black-box access to two quadratic forms f, g and check one by one whether f and g are equivalent to $h := x_1x_2 + \cdots + x_{n-1}x_n$ and if these are equivalent to h then computes invertible matrices A_1, A_2 such that $f(A_1\mathbf{x}) = h$ and $g(A_2\mathbf{x}) = h$. This implies $f = g(A_2A_1^{-1}\mathbf{x})$. Thus, PE for quadratic forms is based on ET for quadratic forms. Now, we formally define the polynomial equivalence problem for orbits of circuits in $\mathscr{C}$.

Let $\mathscr{C}$ be a circuit class. Given black-box access to two polynomials $f(\mathbf{x}), g(\mathbf{x})$, which are in the orbits of two unknown circuits, say $\mathsf{C}_1, \mathsf{C}_2 \in \mathscr{C}$ respectively, decide if f is equivalent to g. One of the promising approaches to solve this problem is to first solve ET for $\mathscr{C}$ on inputs f, g, which return invertible matrices A_1, A_2 and $\mathsf{C}_1', \mathsf{C}_2' \in \mathscr{C}$ such that $f = \mathsf{C}_1'(A_1\mathbf{x})$ and $g = \mathsf{C}_2'(A_2\mathbf{x})$. Now the problem reduces to determining whether there exists an invertible matrix A such that $\mathsf{C}_1' = \mathsf{C}_2'(A\mathbf{x})$ and the structures of C_1' and C_2' can be helpful in finding such an A. In [GST22], we give PE for the orbits of *additive-constant-free ROFs*, which is based on the equivalence test for general ROFs given in the same work. This ET for general ROFs is the generalization of ET for regular ROFs (Theorem 1.11) studied in this thesis. The ET for general ROFs given in [GST22] is not a part of this thesis.

Known results on ET

ET for important polynomial families. In [Kay12], a randomized polynomial time equivalence test for the family of permanent was given. In this work, we give a special case of ET for the family of *Nisan-Wigderson design polynomial* (see Section 1.3.1 and Chapter 3).

In [Kay12], Kayal gave a randomized polynomial time equivalence test for the family of determinant over $\mathbb{C}$. [KNST19] gave a randomized polynomial time equivalence test for the family of IMM, which holds over $\mathbb{C}, \mathbb{Q}$ and finite fields. A polynomial time randomized equivalence test for the family of determinant over finite fields was given in [KNS19], where if the given n^2-variate polynomial f is equivalent to the $n \times n$ determinant over a finite field $\mathbb{F}_q$ then the algorithm outputs a certificate matrix A over a degree n *extension field* of $\mathbb{F}_q$. In this work, we give equivalence tests for the determinant over $\mathbb{Q}$ and finite fields (see Section 1.3.2 and Chapter 4). Our algorithm over $\mathbb{F}_q$ outputs a certificate matrix over the base field and *not* over an extension field. [MNS20] studied the equivalence test for the family of *trace-IMM*

[1] Witt's cancellation theorem says that over any $\mathbb{F}$ with $char(\mathbb{F}) \neq 2$, if $\alpha_1x_1^2 + \alpha_2x_2^2 + \cdots + \alpha_nx_n^2$ and $\alpha_1x_1^2 + \beta_2x_2^2 + \cdots + \beta_nx_n^2$ are equivalent over $\mathbb{F}$, where $n > 1$ and $\alpha_1, \alpha_i, \beta_i \in \mathbb{F}^\times, i \in \{2, \cdots, n\}$ then $\alpha_2x_2^2 + \cdots + \alpha_nx_n^2$ and $\beta_2x_2^2 + \cdots + \beta_nx_n^2$ are also equivalent over $\mathbb{F}$ (a proof is given in Section 2 of [CMM17]). Thus, over $\mathbb{C}$, it is enough to test whether $f + x_{n+1}^2$ and $g + x_{n+1}^2$ are equivalent to $q := x_1x_2 + \cdots + x_{n-2}x_{n-1} + x_n^2 + x_{n+1}^2$. It is easy to see that q is equivalent to $x_1x_2 + \cdots + x_nx_{n+1}$ over $\mathbb{C}$.

polynomial and building on our work, they showed that the equivalence tests for the families of determinant and trace-IMM are reducible to each other in randomized polynomial time over $\mathbb{C}, \mathbb{Q}$ and finite fields.

In [Kay11], Kayal gave polynomial time randomized equivalence tests for the families of power symmetric polynomial and elementary symmetric polynomials. In [GKP18], a randomized polynomial time equivalence test was given for the *sum of univariates polynomials*. This problem generalizes the equivalence test for the power symmetric polynomial. Recently, [KS21a, KS21b] gave a polynomial time randomized equivalence test for the power symmetric polynomial, which only performs basic arithmetic operations and equality tests. On the other hand, the ET for the same polynomial given in [Kay11] requires that roots of univariate polynomials can be extracted in a unit time over the underlying field.

ET for circuit classes. In [MS21], randomized polynomial time equivalence tests were given for the class of *sum-product polynomials*, i.e., polynomials of the type $\sum_{i\in[s]} x_{i,1}\cdots x_{i,d}$, which is contained in the class of depth 2 read-once arithmetic formulas (Definition 2.38) and the class of *ROANFs* (Definition 2.41). The sum-product polynomials and ROANFS are examples of regular read-once arithmetic formulas (Definition 2.40). In this thesis, we give a randomized polynomial time ET for the class of regular ROFs. This result was generalized in a follow-up work [GST22], where we give a randomized polynomial time ET for the class of general ROFs.

1.3 Motivation and our results

Now, we describe our main contributions. We will follow the notations given in the beginning of Chapter 2. The content of this section is divided into three parts - the first one contains some structural and algorithmic results for the Nisan-Wigderson design polynomial, the second one is about ET for the determinant, and the last part contains an ET for the class of *regular* read-once arithmetic formulas. In each part, we first motivate the problems, then state the main theorems and then give a summary of our main technical contributions. We elaborate these summaries by giving detailed proof ideas in Section 1.4. Based on these proof ideas, the complete proofs are given in Chapters 3, 4 and 5.

1.3.1 Some structural and algorithmic results for NW

In this section, we consider the family of the Nisan-Wigderson design polynomial (in short, $\mathscr{NW}$), which was introduced in [KSS14]. The results present here about the Nisan-Wigderson polynomial are from [GS19], which is a joint work with Chandan Saha. We first define the Nisan-Wigderson design polynomial. Let d be a prime number, $k \in \mathbb{N}, k << d$, $\mathbb{F}_d$ be the finite

field of size d, $\mathbf{x} = \{x_{i,j} : i, j \in \mathbb{F}_d\}$, and $\mathbb{F}_d[z]_k = \{h \in \mathbb{F}_d[z] \ : \ \deg(h) \leq k\}$. Then,

$$\mathsf{NW}_{d,k}(\mathbf{x}) = \sum_{h \in \mathbb{F}_d[z]_k} \prod_{i \in \mathbb{F}_d} x_{i,h(i)}. \tag{1.1}$$

The Nisan-Wigderson design polynomial was inspired by the set-system given by Nisan and Wigderson (see Section 2.3 of [NW94]), where every pair of sets in the set-system has low intersection. A variant of the Nisan-Wigderson polynomial was also used in [Raz10]. $\mathsf{NW}_{d,k}$ has been used as a hard polynomial in many lower bound proofs (see below). In these works, the value of k is taken as d^ϵ for some $\epsilon \in (0,1)$. Although our results hold for $k \in \{1, \ldots, \frac{d}{4} - 5\}$, it is best to think of $k = d^\epsilon$, where $\epsilon \in (0,1)$ is an arbitrary constant. Then, $\mathscr{NW} = \{\mathsf{NW}_{d,k} : d \text{ is a prime}\}$. When the value of d is clear from the context, we drop the subscripts from $\mathsf{NW}_{d,k}$ for notational simplicity. We first record some important properties of this polynomial.

1. **Set-multilinearity and homogeneity**: Consider the partition $\mathbf{x} = \biguplus_{i \in \mathbb{F}_d} \mathbf{x}_i$, where for every $i \in \mathbb{F}_d, \mathbf{x}_i = \{x_{i,j} : j \in \mathbb{F}_d\}$. Then, $\mathsf{NW}_{d,k}(\mathbf{x})$ is *set-multilinear* with respect to the partition $\mathbf{x} = \biguplus_{i \in \mathbb{F}_d} \mathbf{x}_i$, i.e., every monomial of $\mathsf{NW}_{d,k}$ contains exactly one variable from every $\mathbf{x}_i$ [1]. Thus, $\mathsf{NW}_{d,k}$ is a homogeneous degree d polynomial in d^2 variables.

2. $\mathscr{NW}$ **is in** VNP: Suppose we are given a monomial m in $\mathbf{x}$ variables, where $\mathbf{x} = \uplus_{i \in \mathbb{F}_d} \mathbf{x}_i$ and d is a prime number. If m is not set-multilinear then the coefficient of m in $\mathsf{NW}_{d,k}$ is zero. Oherwise $m = x_{0,l_0} \cdots x_{d-1,l_{d-1}}$. Let $h \in \mathbb{F}[z]$ be obtained by interpolating $\{(i, l_i) : i \in \mathbb{F}_d\}$. If $\deg(h) \leq k$ then the coefficient of m in $\mathsf{NW}_{d,k}(\mathbf{x})$ is 1, otherwise it is zero. Since these checks can be done in $\text{poly}(d)$ time and number of variables in $\mathsf{NW}_{d,k}$ is d^2, it follows from the Valiant's criterion given in Section 1.1.2.1 that $\mathscr{NW} \in$ VNP.

3. **Low-intersection property**: Let m_1 and m_2 be two arbitrary monomials of $\mathsf{NW}_{d,k}$. Then, the number of common variables between m_1 and m_2 is at most k. This follows from the fact that distinct $h_1, h_2 \in \mathbb{F}_d[z]_k$ agree on at most k elements of $\mathbb{F}_d$.

NW as a hard polynomial in lower bounds. The low-intersection property of NW has been exploited many times to give strong lower bounds for various classes of arithmetic circuits. Different variants of the Nisan-Wigderson polynomials have been used as hard polynomials in several lower bound proofs given in [Raz10, KSS14, CM14, KS14a, KLSS17, KS16a, KS16b, KST16, KS17a, GST20, KS22]. Thus, from the perspective of proving lower bounds in ACT, $\mathscr{NW}$ is an important family. Similar to the family of permanent, we know that $\mathscr{NW} \in$ VNP.

[1] Whenever we talk about set-multilinearity in this section, Section 1.4.1 or Chapter 3, it is always with respect to the partition $\mathbf{x} = \biguplus_{i \in \mathbb{F}_d} \mathbf{x}_i$.

But apart from this, we do not know much about other interesting structural and algorithmic properties of NW, which have been studied for the permanent and other polynomials like the determinant, IMM, the power symmetric polynomial, the elementary symmetric polynomial, etc. These polynomials also have been crucially used in many lower bound results.

In this thesis, we study some structural properties related to the *symmetries* (Definition 2.23) of NW like *characterization by symmetries* (Definition 2.24). We also give some useful algorithmic results for NW like a *circuit testing* algorithm, a *flip theorem* and a special case of ET for the family of NW. These results crucially use the symmetries of NW. Such results have been studied for the permanent. We first motivate the problems we study about NW and then state them formally. The study of most of these problems originate from GCT (Section 1.1.2.2).

Characterization by symmetries. Our first result is related to the *characterisation by symmetries* property (Definition 2.24) of NW over different fields. We mentioned in Section 1.1.2.2 that this property plays an important role in GCT. Let us first see why it is so. There are some *barrier type results* in complexity theory, which say that certain approaches can never yield strong enough lower bounds for a specific model of computation, provided some widely believed assumptions hold. An example of such a result is the concept of *natural proofs* introduced by Razborov and Rudich in [RR97]. At a high level, this result says that assuming the existence of pseudo-random functions, if any lower bound technique for general Boolean circuits is based on a property that satisfies the *constructivity* and *largeness* criteria (see Section 2.1 of [RR97] for their definitions) then we can not get super-polynomial lower bounds for general Boolean circuits using this technique. Algebraic variants of natural proofs have also been studied (see [FSV17, GKSS17, CKSV20]). Thus, if a lower bound technique is based on a property that does not satisfy either of these two criteria and if there exists an *explicit function* that satisfies this property then it may be possible to obtain super-polynomial lower bounds using this technique. Characterization by symmetries is one such property for polynomials, which violates the largeness criterion, i.e., this property is not satisfied by a large fraction of polynomials. This is so because it is known that a *random polynomial* is not characterised by its symmetries (see Proposition 3.4.9 in [Gro12]). GCT aims to show a super-polynomial lower bound on the size of arithmetic circuits computing the permanent by exploiting the characterization by symmetries properties of the permanent and the determinant.

Characterization of the permanent and the determinant by their symmetries are classical results in mathematics shown in [MM62] and [Fro97] respectively. Simpler proofs of these facts are given in Chapter 3 of [Gro12]. It is also known that IMM is characterized by its symmetries [Ges16, KNST19]. This property also holds for the power symmetric polynomial (see Section

2 in [CKW11]) but is not possessed by the elementary symmetric polynomial [Hüt16]. As the family of NW has also been used in lower bound proofs, it is natural to ask whether NW is characterized by its symmetries over the underlying field. The following two theorems are devoted to the characterization by symmetries property for NW. This property says that if a polynomial f is characterized by its symmetries then f is uniquely identified (up to a constant multiple) by its symmetries. For these theorems, we need the notion of the *group of symmetries* of a polynomial f, denoted as $\mathscr{G}_f$ (see Definition 2.23).

Theorem 1.1 (NW characterized by its symmetries over $\mathbb{C}$) *Let d be a prime number, $\mathbb{F}$ be a field containing a d-th primitive root of unity*[1]*, and f be a homogeneous degree d polynomial in d^2 variables over $\mathbb{F}$. If $\mathscr{G}_{\mathsf{NW}_{d,k}} \subseteq \mathscr{G}_f$ then $f = \alpha \cdot \mathsf{NW}_{d,k}$ for some $\alpha \in \mathbb{F}$.*

Theorem 1.1 holds over $\mathbb{C}$ and a finite field $\mathbb{F}$ having a d-th root $\zeta \neq 1$ and $|\mathbb{F}| \neq d+1$. One can ask if we can get the same result over $\mathbb{R}$ and $\mathbb{Q}$. We answer it in the following theorem.

Theorem 1.2 (NW not characterized by its symmetries over $\mathbb{R}$) *Let d be a prime number and $\mathbb{F}$ be either $\mathbb{R}, \mathbb{Q}$ or a finite field satisfying $d \nmid |\mathbb{F}| - 1$. Then, $\mathsf{NW}_{d,k}$ is not characterized by its symmetries over $\mathbb{F}$.*

Thus, NW is characterized by its symmetries over $\mathbb{C}$ but not over $\mathbb{R}$ and $\mathbb{Q}$. On the other hand, the permanent of an $n \times n$ symbolic matrix is characterized by its symmetries over fields having more than n elements (see Proposition 3.4.5 in [Gro12]). In the proof of Theorem 1.2, the structures of symmetries of NW play a very important role. We showed in the author's master's thesis [Gup17] that every element of $\mathscr{G}_{\mathsf{NW}}$ is a product of a permutation matrix and a diagonal matrix (see Theorem 3.3). A similar result also holds for the group of symmetries of the permanent. Certain symmetries of NW used in the proofs of Theorems 1.1 and 1.2 (see Section 1.4.1) also come in handy to show some important algorithmic results for NW mentioned below.

Algorithmic results for NW. As seen before, we do not know whether $\mathscr{NW} \in \mathsf{VP}$, i.e., $\mathsf{NW}_{d,k}$ is computable by an arithmetic circuit of size $\mathrm{poly}(d)$. In the absence of a proof that $\mathsf{VP} = \mathsf{VNP}$, the algorithmic problem of testing whether a given circuit computes NW is an interesting and non-trivial problem. This problem is known as the *circuit testing* problem and makes sense for every polynomial family not known to be in VP. Two efficient algorithms for circuit testing are known for the permanent - one based on the *self-reducibility* of the permanent [Lip89] and other based on its symmetries [Mul10]. In this thesis, we give a circuit testing algorithm for NW

[1] An $\alpha \in \mathbb{F}$ is called a d-th primitive root of unity if $\alpha^d = 1$ and for every $1 \leq r < d, \alpha^r \neq 1$.

using its symmetries. In particular, this algorithm uses the property that NW is *characterized by circuit identities* (see Definition 2.25 and Lemma 3.2). Before stating the result, we give a useful notation. In the following two theorems, whenever we say a size-s arithmetic circuit, we would mean an arithmetic circuit C of size s, where the degree of the polynomial computed by C (also called the degree of C) is bounded by $\delta(s)$, where $\delta : \mathbb{N} \to \mathbb{N}$ is a polynomial function. We state these two theorems over a finite field $\mathbb{F}$ satisfying a mild condition on its size. Suitable versions of these theorems also hold over $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$. In these theorems, d is a prime number and we assume without loss of generality that $\delta(s) \geq d$.

Theorem 1.3 (Circuit testability) *There is a randomized algorithm that takes input black-box access to a size-s arithmetic circuit C over a finite field $\mathbb{F}$, where $|\mathbb{F}| \geq 4 \cdot \delta(s)$, and determines whether or not $\mathsf{C} = \mathsf{NW}_{d,k}$ with probability $1 - \exp(-s)$, using $\mathrm{poly}(s)$ field operations.*

The next theorem addresses the question of the following type: Assuming a polynomial (or a Boolean function) f is not computable by a size-s arithmetic circuit (respectively, a size-s Boolean circuit), is it easy to "certify" this hardness algorithmically? A result of this kind is called a *flip theorem* for f. Flip theorems are known for 3SAT [FPS08, Ats06] and the permanent [Mul10, Mul11b]. Flip theorem is important from the perspective of GCT (see [Mul07, Mul10]). We give a flip theorem for NW below, which is also based on the property that NW is characterized by circuit identities (Lemma 3.2).

Theorem 1.4 (Flip theorem) *Suppose $\mathsf{NW}_{d,k}$ is not computable by size-s arithmetic circuits over $\mathbb{F}$, where $|\mathbb{F}| \geq 4 \cdot \delta(s)$. Then, there exist $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{F}^n$, where $m = \mathrm{poly}(s)$ such that for every arithmetic circuit C of size at most s, there is an $\ell \in [m]$ satisfying $\mathsf{C}(\mathbf{a}_\ell) \neq \mathsf{NW}_{d,k}(\mathbf{a}_\ell)$. A set of randomly generated points $\mathbf{a}_1, \ldots, \mathbf{a}_m \in_r \mathbb{F}^n$ has this property with probability $1 - \exp(-s)$. Moreover, black-box derandomization of PIT for size-$(10s)$ circuits over $\mathbb{F}$ using $\mathrm{poly}(s)$ field operations implies $\mathbf{a}_1, \ldots, \mathbf{a}_m$ can be computed deterministically using $\mathrm{poly}(s)$ field operations.*

In the above theorem, $\mathbf{a}_1, \ldots, \mathbf{a}_m$ is a short list of certificate points against all arithmetic circuits of size at most s. At the end, we give a special case of ET for NW, called the *BD-PS equivalence test*. This checks whether the given polynomial f is *BD-PS equivalent* to NW, i.e., whether there exist a *block-diagonal* permutation matrix[1] $A \in \mathrm{GL}(d^2, \mathbb{F})$ and a diagonal (also called scaling) matrix $B \in \mathrm{GL}(d^2, \mathbb{F})$ such that $f = \mathsf{NW}(AB\mathbf{x})$. We call such a matrix AB a block-diagonal permutation scaling (BD-PS) matrix.

[1] An $A \in M_{d^2}(\mathbb{F})$ is called a block-diagonal permutation matrix if it looks like $A = \mathrm{diag}(A_{\sigma_0}, \ldots, A_{\sigma_{d-1}})$, where every σ_i is a permutation on $\mathbb{F}_d$ and A_{σ_i} is the $d \times d$ matrix, where for every $l \in \mathbb{F}_d$, the $(l, \sigma_i(l))$-th entry of A_{σ_i} is 1 and every other entry is 0. We say A_{σ_i} is the permutation matrix corresponding to σ_i.

The motivation for looking into the BD-PS equivalence test for is that it is a special case of the *block-permuted* equivalence test (in short, BP ET) for NW. A BP ET for NW does the following: Given black-box access to an $f \in \mathbb{F}[\mathbf{x}]$, it determines whether there exists an invertible *block-permuted matrix*[1] A over $\mathbb{F}$, such that $f = \mathsf{NW}(A\mathbf{x})$. The reason BP ET for NW is interesting is that we show that the general ET for NW reduces to its BP ET in randomized polynomial time over almost any field (see Lemma 3.3 of Chapter 3). Thus, a randomized polynomial time BP ET for NW would immediately imply a randomized polynomial time equivalence test for NW. We hope that a BD-PS equivalence test for NW can give us crucial insights on the BP ET for NW. Another motivation for looking into the BD-PS ET for NW is that it is a special case of the *permutation scaling* (PS) equivalence test for NW, i.e., the underlying matrix is a product of a permutation matrix and an invertible scaling matrix. PS equivalence test played a very crucial role for the permanent over any field and the determinant over $\mathbb{C}$. In particular, Kayal in [Kay12] gave randomized polynomial time reduction from general ET for permanent over any field (similarly, determinant over $\mathbb{C}$) to PS equivalence test for the permanent (respectively, PS equivalence test for the determinant over $\mathbb{C}$).

Now, we state the theorem on BD-PS equivalence test for NW.

Theorem 1.5 (Block-diagonal permutation scaling ET for NW) *Let d be a prime number, $\mathbb{F}$ be a finite field such that $d \nmid (|\mathbb{F}|-1)$ and $|\mathbb{F}| \geq 4d$. There is a randomized $\mathrm{poly}(d, \log|\mathbb{F}|)$ time algorithm that takes input black-box access to a degree d polynomial $f \in \mathbb{F}[\mathbf{x}]$ and correctly decides if f is BD-PS equivalent to NW with high probability. If the answer is yes then it outputs a BD-PS matrix $C \in \mathrm{GL}(d^2, \mathbb{F})$ such that $f = \mathsf{NW}(C\mathbf{x})$.*

An appropriate version of the above theorem holds over $\mathbb{R}$. The proof ideas of Theorems 1.1-1.5 are given in Section 1.4.1 and their proofs are given in Chapter 3.

Summary of our main technical contributions. We give a high level overview of the key technical results here. A detailed overview of the proofs of the above mentioned theorems are given in Section 1.4.1. In this work, we build on the structural results related to the symmetries and the Lie algebra of NW given in the author's master's thesis [Gup17]. There we obtained a complete understanding of the Lie algebra of NW. Using this, the *Hessian* of NW, and the *evaluation dimension* measure, we showed that every symmetry of NW is a product of a permutation and an invertible scaling matrix. In this thesis, we gave some explicit symmetries of NW over $\mathbb{C}$ (see Claim 3.1.1), which imply the characterization by symmetries result for NW

[1] A $d^2 \times d^2$ matrix A is said to be block-permuted with block size d if there exists a $d^2 \times d^2$ block-diagonal matrix B with block size d and a $d \times d$ permutation matrix P such that $A = B \cdot (P \otimes I_d)$.

over $\mathbb{C}$. We also show that some specific symmetries of NW present over $\mathbb{C}$ are not present over $\mathbb{R}$ and $\mathbb{Q}$, and using this along with the structure of symmetries of NW mentioned above, we show that NW is not characterized over $\mathbb{R}$ and $\mathbb{Q}$. We also prove that NW is characterized by circuit identities over any field (see Lemma 3.2). This result uses some symmetries of NW. By exploiting the fact that NW is characterized by circuit identities, we obtain two algorithmic results for NW: A randomized polynomial time circuit testing algorithm and a flip theorem. We also give the block-diagonal permutation scaling ET for NW. This ET uses some symmetries of NW along with a structural insight obtained from the analysis of the Lie algebra of NW given Chapter 3 of [Gup17] (see Claim 3.1.2 in this regard).

1.3.2 ET for the determinant over finite fields and over $\mathbb{Q}$

In this section, we present our results on the determinant equivalence test (in short, DET). These results are taken from [GGKS19], which is a joint work with Ankit Garg, Neeraj Kayal and Chandan Saha.

Let $n \in \mathbb{N}$, $X = (x_{i,j})_{i,j\in[n]}$, where $x_{i,j}$ is a variable for every $i, j \in [n]$ and $\mathbf{x} = \{x_{1,1}, \ldots, x_{n,n}\}$. Let $\mathsf{Det}_n(\mathbf{x}) := \det(X)$. Then, Det_n is a degree n polynomial. Apart from being a well-studied object in linear algebra, the determinant also plays an important role in ACT. Det_n has been used as a hard polynomial in many lower bound proofs [Kal85, SW01, Raz09, RY09, GKKS14a] and it is also crucial from the perspective of GCT. The family of determinant is complete for the class of ABPs (Definition 2.36) under p-projections (see [MV97] for a proof). Another polynomial family, which is also complete for the class of ABPs is the family of IMM. A polynomial time randomized equivalence test for IMM over almost any field was given in [KNST19]. One might ask whether the family of determinant also possesses an efficient equivalence test over most fields. As mentioned before in Section 1.1.2, DET is also interesting from the viewpoint of GCT as it allows us to test whether a given polynomial is in the orbit of the determinant.

We noted in Section 1.2.2 that a randomized polynomial time DET over $\mathbb{C}$ was given in [Kay12]. We also saw that [KNS19] gave a randomized polynomial time DET over a finite field $\mathbb{F}_q$. But one shortcoming of this result is that if the polynomial f given as input to DET is equivalent to Det_n over $\mathbb{F}_q$ then the algorithm outputs a "certificate matrix" A over a degree n extension field of $\mathbb{F}_q$ such that $f = \mathsf{Det}_n(A\mathbf{x})$. In this work, we give a randomized polynomial time DET algorithm over finite fields, where the output certificate matrix is also over the base field $\mathbb{F}_q$ and not over an extension of $\mathbb{F}_q$. It was not known before this work if DET is decidable over $\mathbb{Q}$. We give the *first* DET algorithm over $\mathbb{Q}$, which takes oracle access to an integer factoring algorithm IntFact and outputs a certificate matrix over $\mathbb{Q}$. This algorithm is randomized and runs in polynomial time if n is bounded. If we remove oracle access to IntFact

from the DET algorithm then it outputs a certificate matrix over an extension field $\mathbb{L}$ of $\mathbb{Q}$ satisfying $[\mathbb{L} : \mathbb{Q}] \leq n$, where $[\mathbb{L} : \mathbb{Q}]$ is the degree of field extension (see Definition 2.12). This variant of the DET is also randomized but runs in polynomial time for every value of n. Now, we state the main theorems related to DET over finite fields and $\mathbb{Q}$.

Theorem 1.6 (DET over finite fields) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_{1,1}, \ldots, x_{n,n}\}$, $\mathbb{F}$ be a finite field such that $|\mathbb{F}| \geq 10n^4$ and $char(\mathbb{F}) \nmid n(n-1)$, and $f \in \mathbb{F}[\mathbf{x}]$ be a degree n polynomial. Then, there exists a randomized algorithm that takes black-box access to f and decides if f is equivalent to Det_n or not with high probability. If yes, it outputs an $A \in \mathrm{GL}(n^2, \mathbb{F})$ such that $f = \mathsf{Det}_n(A\mathbf{x})$, otherwise it outputs 'Fail'. The running time of this algorithm is $\mathrm{poly}(n, \log |\mathbb{F}|)$.*

Theorem 1.7 (DET over $\mathbb{Q}$) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_{1,1}, \ldots, x_{n,n}\}$, $f \in \mathbb{Q}[\mathbf{x}]$ be a degree n polynomial, and β be the bit length of coefficients of f. Suppose we have black-box access to f.*

1. *There exists a randomized algorithm, which takes oracle access to an integer factoring algorithm* IntFact *and decides if f is equivalent to Det_n over $\mathbb{Q}$ with high probability. If yes, it outputs an $A \in \mathrm{GL}(n^2, \mathbb{Q})$ such that $f = \mathsf{Det}_n(A\mathbf{x})$, otherwise outputs 'Fail'. If n is bounded, the algorithm runs in $\mathrm{poly}(n, \beta)$ time.*

2. *There exists a randomized $\mathrm{poly}(n, \beta)$ time algorithm, which decides if f is equivalent to Det_n over $\mathbb{Q}$ with high probability. If yes, it outputs an $A \in \mathrm{GL}(n^2, \mathbb{L})$ such that $f = \mathsf{Det}_n(A\mathbf{x})$, where $\mathbb{L}$ is an extension field of $\mathbb{Q}$ satisfying $[\mathbb{L} : \mathbb{Q}] \leq n$.*

The DET algorithms in Theorems 1.6 and 1.7 have two main steps - The first step is to reduce DET to the *full matrix algebra isomorphism* (FMAI) problem in randomized polynomial time. FMAI is an algorithmic problem that determines whether an $\mathbb{F}$-algebra (Definition 2.18) $\mathscr{A} \subseteq M_{n^2}(\mathbb{F})$ is isomorphic to $M_n(\mathbb{F})$ and if yes, it returns an $\mathbb{F}$-algebra isomorphism (Definition 2.19) $\varphi : \mathscr{A} \to M_n(\mathbb{F})$. In the second step, we invoke the FMAI algorithm, which are known over finite fields and $\mathbb{Q}$ (see Section 2.2.4). Reduction of DET to FMAI is the main technical contribution of our work and this reduction works over almost every field. We formally state this result as the following theorem.

Theorem 1.8 (Reduction of DET to FMAI) *Let $n \geq 2$, $|\mathbb{F}| > 10n^4$ and $char(\mathbb{F}) \nmid n(n-1)$. Then, there exists a polynomial time randomized algorithm, with oracle access to FMAI, that takes input black-box access to an $f \in \mathbb{F}[\mathbf{x}]$ of degree n and solves DET for f over $\mathbb{F}$ with high probability.*

One might wonder if we can get rid of IntFact oracle from the DET algorithm over $\mathbb{Q}$ given in the first part of Theorem 1.7. In the following theorem, we show that it is unlikely.

Theorem 1.9 (IntFact reduces to DET for quadratic forms over $\mathbb{Q}$) *Assuming the Generalized Riemann Hypothesis (GRH), there exists a randomized polynomial time reduction from the problem of factoring square-free integers to computing an $A \in \mathrm{GL}(4,\mathbb{Q})$ such that $f = \mathsf{Det}_2(A\mathbf{x})$, provided f is equivalent to Det_2.*

In the following theorem, we give a reduction from FMAI to DET, which is polynomial time if n is bounded.

Theorem 1.10 (FMAI reduces to DET) *Let $n \in \mathbb{N}$ and $\mathbb{F}$ be a field that satisfies $\mathrm{char}(\mathbb{F}) \nmid n$. There exists an algorithm, which takes input a basis of an $\mathbb{F}$-algebra $\mathscr{A}$, has oracle access to DET over $\mathbb{F}$ and decides if $\mathscr{A}$ is isomorphic as an $\mathbb{F}$-algebra to $M_n(\mathbb{F})$ or not using $n^{O(n)}$ many field operations. If the answer is yes, it outputs an $\mathbb{F}$-algebra isomorphism from $\mathscr{A}$ to $M_n(\mathbb{F})$.*

Remark 1.1 *In a follow up work, Theorem 1.10 was improved significantly in* [MNS20]. *The authors showed that FMAI reduces in randomized polynomial time to DET.*

The proof ideas of these theorems are given in Section 1.4.2 and their proofs are given in Chapter 4. Now, we give a brief overview of the main technical contribution of this work.

Summary of our main technical contributions. The main result of this work is Theorem 1.8, which gives a randomized polynomial time reduction from DET to FMAI over almost any field. We give a high level idea of this reduction here and a detailed overview in Section 1.4.2. This reduction is based on the rich structure of the Lie algebra of the determinant, denoted $\mathfrak{g}_{\mathsf{Det}}$ (see Section 1.4.2 for the description of $\mathfrak{g}_{\mathsf{Det}}$). If the input polynomial f to the DET algorithm satisfies $f = \mathsf{Det}(A\mathbf{x})$ for some invertible matrix A then $\mathfrak{g}_f = A^{-1} \cdot \mathfrak{g}_{\mathsf{Det}} \cdot A$ (Fact 2.10). The algorithm computes a basis of $\mathfrak{g}_f$ using Fact 2.16, then decomposes $\mathfrak{g}_f$ into two sub-spaces and picks one of these subspace, which we call $\mathscr{F}_{\mathrm{col}}$. This decomposition of $\mathfrak{g}_f$ is at the core of the reduction from DET to FMAI and works over almost any field. To obtain this decomposition, we analyse *irreducible invariant spaces* (Definition 2.16) of a set of carefully chosen linear operators on $\mathfrak{g}_f$. After obtaining $\mathscr{F}_{\mathrm{col}}$, the algorithm computes the algebra $\mathscr{A}$ generated by $\mathscr{F}_{\mathrm{col}}$ and invokes FMAI algorithm to obtain an $\mathbb{F}$-algebra isomorphism $\varphi : \mathscr{A} \to M_n(\mathbb{F})$. Using φ and the *Skolem-Noether theorem* (see Theorem 2.1), the algorithm computes an invertible matrix B such that $f = \mathsf{Det}(B\mathbf{x})$. We also give a reduction from FMAI to DET, which is efficient if n is bounded. This reduction exploits the fact that the determinant is *characterized by its Lie algebra* (see Lemma 4.5).

1.3.3 Equivalence test for regular ROFs

An arithmetic formula C over a field $\mathbb{F}$ is said to be a read-once arithmetic formula (in short, an ROF) if every leaf node of C is labelled by either a distinct variable or a constant from $\mathbb{F}$ (Definition 2.38). The class of ROFs is well-studied in the literature. Efficient black-box PIT algorithms are known for ROFs [SV15, SV14, AFS+18, MV18] and this class also has efficient randomized and deterministic reconstruction algorithms [HH91, BHH95, SV14, MV18]. It is one of the few classes of arithmetic circuits for which we have a deterministic polynomial time black-box PIT algorithm and a deterministic polynomial time reconstruction algorithm. In this section, we talk about ET for the class of ROFs. Recall that an ET for ROFs reconstructs polynomials in the orbits of ROFs. In this thesis, we give a randomized polynomial time ET for the class of restricted (yet quite expressive) ROFs, which we call *regular ROFs* (Definition 2.40). The class of regular ROFs contains some interesting circuit classes like the classes of *sum-product polynomials* and *ROANFs* considered in [MS21]. The ET for regular ROFs takes oracle access to PE for quadratic forms (in short, QFE) over the underlying field $\mathbb{F}$. This is a joint work with Chandan Saha and Bhargav Thankey. Before presenting the main theorem, we motivate why ET for the class of ROFs is an interesting problem in ACT.

1. **Relationship with PE for quadratic forms**. Efficient QFE algorithms over $\mathbb{C}, \mathbb{R}$, and finite fields having characteristic not equal to 2 and over $\mathbb{Q}$ with oracle access to an integer factoring algorithm are known (see Fact 2.19). These algorithms are based on well-known results on the classification of quadratic forms (see [Ser73, Ara11]). Suppose f is a quadratic form over $\mathbb{C}$ such that f has no redundant variables (Definition 2.32) and the number of variables in f is even. Then, f is equivalent to $x_1x_2 + \cdots + x_{n-1}x_n$ over $\mathbb{C}$, where n is an even number. If f has odd number of variables, then it is equivalent to $x_1x_2 + \cdots + x_{n-2}x_{n-1} + x_n^2$ over $\mathbb{C}$. Recall from Section 1.2.2 that with the help of Witt's cancellation theorem [Wit37], this case can be reduced to the case when n is even. A PE for quadratic forms over $\mathbb{C}$ takes black-box access to two quadratic forms f, g and check one by one whether f and g are equivalent to $h := x_1x_2 + \cdots + x_{n-1}x_n$ and if these are equivalent to h then computes invertible matrices A_1, A_2 such that $f(A_1\mathbf{x}) = h$ and $g(A_2\mathbf{x}) = h$. This implies $f = g(A_2A_1^{-1}\mathbf{x})$. Thus, internally a QFE algorithm is solving ET for quadratic forms, which is a special case of ET for ROFs as h is a depth-2 ROF. Thus, an ET for the class of ROFs would clearly generalize QFE over $\mathbb{C}$.

2. **Reconstructing orbits of well-studied circuit classes**. As mentioned in Section 1.2.2, recently [MS21], [ST21] and [BG21] gave deterministic quasi-polynomial time PIT

algorithms for the orbits of many interesting circuit classes including sparse polynomials, ROFs and bounded-width ROABPs. So, it is natural to ask if we can also reconstruct/learn orbits of these circuit classes efficiently or in other words, design efficient ET algorithms for these circuit classes.

3. **Relationship with non-degenerate formula reconstruction**. Let C be an n-variate *random arithmetic formula*, i.e., the C has an arbitrary tree structure, alternate layers of $+$ and $\times$ gates and the leaves of C are labelled by s many random linear forms in n variables over $\mathbb{F}$. If $n \geq s$ then with high probability, C is in the orbit of an ROF over a sufficiently large field $\mathbb{F}$. Thus, an efficient ET for ROFs leads to an efficient algorithm to reconstruct an n-variate random arithmetic formula in the *high number of variables* regime. A randomized polynomial time algorithm for reconstruction of random arithmetic formulas in the alternating normal form (ANF) was given in [GKQ13]. This algorithm reconstructs random arithmetic formulas whose underlying tree is a complete binary tree. Such formulas need not satisfy $n \geq s$ whereas an ET for ROFs gives an algorithm to reconstruct random formulas satisfying $n \geq s$ and having an arbitrary tree structure.

4. **Generalizes reconstruction algorithm for ROFs**. As mentioned above, a polynomial time deterministic reconstruction algorithm is known for ROFs. ET for ROFs is clearly a strict generalization of the ROF reconstruction problem as here we aim to find an invertible matrix A and an ROF C such that the given polynomial f satisfies $f = \mathsf{C}(A\mathbf{x})$.

5. **Generalization of ET for two sub-classes of ROFs**. As mentioned in Section 1.2.2, [MS21] gave efficient ETs for the classes of sum-product polynomials and ROANFs. It is easy to see that these are sub-classes of the class of regular ROFs. Thus, an efficient ET for regular ROFs would generalize these two results.

We now state our main result. In the following theorem, we consider a slightly general definition of the orbit of a polynomial g, which is $\text{orb}(g) = \{g(B\mathbf{x}+\mathbf{d}) : B \in \text{GL}(n,\mathbb{F}), \mathbf{d} \in \mathbb{F}^n\}$.

Theorem 1.11 (ET for regular ROFs) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}$, $\mathbb{F}$ be a field satisfying $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) \geq n^2$ and $|\mathbb{F}| \geq n^{13}$, and $f \in \mathbb{F}[\mathbf{x}]$ be in the orbit of an unknown regular ROF C. Then, there exists a randomized $\text{poly}(n)$ time algorithm that takes input black-box access to f, has oracle access to QFE over $\mathbb{F}$ and does the following with high probability: It outputs an $A \in \text{GL}(n,\mathbb{F})$ such that $f(A\mathbf{x}) = \mathsf{C}(PS\mathbf{x} + \mathbf{b})$, where $P, S \in \text{GL}(n,\mathbb{F})$ are permutation and scaling matrices respectively and $\mathbf{b} \in \mathbb{F}^n$.*

Remark 1.2 *1. Note that $f(A\mathbf{x})$ is an ROF. So, we can find an ROF C' using any of the polynomial time ROF reconstruction algorithm* [HH91, BHH95, MV18] *such that C' computes $f(A\mathbf{x})$.*

2. There are efficient algorithms for QFE over $\mathbb{C}, \mathbb{R}$, finite fields not having characteristic two and over $\mathbb{Q}$ with oracle access to an integer factoring algorithm. As ET for regular ROFs takes oracle access to QFE, our ET algorithm is efficient over these fields.

3. Our ET algorithm extensively uses the black-box polynomial factorization by [KT90] *(see Fact 2.17). This is one of the main reasons for constraints on the size and characteristic of the underlying field in the above theorem.*

4. In a follow-up work [GST22]*, we have removed the regularity condition from Theorem 1.11 and have given a randomized polynomial time ET with oracle access to QFE for the class of general ROFs. This algorithm uses several new ideas along with the ones used in the ET for the class regular ROFs. This improvement is mainly due to my other two co-authors of this work and hence is not a part of my thesis.*

The proof idea and a proof of Theorem 1.11 are given in Section 1.4.3 and Chapter 5 respectively.

Summary of the main technical contributions. We give a high-level overview of the main technical parts of our work. The equivalence test for the class of regular ROFs extensively uses some important properties related to the Hessian determinant of a regular ROF - non-zeroness, knowledge of the essential variables and knowlegde of the factors of the Hessian determinant of a regular ROF. We study these properties in this thesis. Let C be a regular ROF. We show that $\det(H_{\mathsf{C}})$ is non-zero over a field $\mathbb{F}$ satisfying either $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) \geq |\text{var}(\mathsf{C})|$. This is not true over a field $\mathbb{F}$ having $char(\mathbb{F}) < |\text{var}(\mathsf{C})|$. For example, if $\mathsf{C} = x_1x_2x_3$ then $\det(H_{\mathsf{C}}) = 0$ over fields having characteristic equal to two. We also study the factors of $\det(H_{\mathsf{C}})$ and this allows us to show that all the variables present in C except the variables present in the 'top-quadratic term' of C are essential for $\det(H_{\mathsf{C}})$. Using this knowledge along with QFE over the underlying field $\mathbb{F}$, we compute an invertible matrix A such that $f(A\mathbf{x})$ is the sum of variable disjoint polynomials, where f is the input of the ET algorithm. After this, we obtain black-box access to the sub-formulas of f[1] from black-box of f and solve the problem recursively. It is important that at every step of recursion, black-box access to a sub-formula of f, say f_1, is computed using just one black-box query to the parent of f_1, otherwise the running time of our

[1]By this, we mean sub-formulas of the formula of f obtained from C by replacing leaves labelled with variables by linearly independent affine forms.

ET algorithm may be exponential. This is so because the height of the formula of f can be as large as n. We show how to compute black-box access *efficiently* at every step of recursion by using some factors of $\det(H_f)$, which are also $+$-rooted sub-formulas of f. Elaborating on these key ideas, we give a detailed overview of the proof of Theorem 1.11 in Section 1.4.3.

1.4 Proof ideas

In this section, we give high level proof overviews of the theorems stated in Section 1.3. The detailed proofs of these theorems are given in subsequent chapters (Chapters 3 - 5). Like Section 1.3, this section also has three main subsections - the first one devoted to the results on NW, the second one is on ET for the determinant and the last one is on ET for regular ROFs.

1.4.1 Some structural and algorithmic results for NW

We saw five theorems in Section 1.3.1 - the first two were about the characterization by symmetries property of NW over different fields, the next two were about the circuit testability and a flip theorem for NW and the last one was about a special case of ET for NW. We present the proof overviews of these theorems below. The detailed proofs based on these overviews are given in Chapter 3. Recall that d is a prime number, $k = d^\epsilon$ for some $\epsilon \in (0,1)$. For simplicity of notations, we drop the subscripts from $\mathsf{NW}_{d,k}$ whenever the value of d is clear. Also, recall that whenever we mention a set-multilinear polynomial in $\mathbb{F}[\mathbf{x}]$, it is always with respect to the partition $\mathbf{x} = \biguplus_{i\in\mathbb{F}_d} \mathbf{x}_i$, where for every $i \in \mathbb{F}_d, \mathbf{x}_i = \{x_{i,j} : j \in \mathbb{F}_d\}$.

Characterization by symmetries: Proof ideas of Theorems 1.1 and 1.2

We first show how NW is characterized by its symmetries over fields containing a d-th primitive root of unity. Recall that $\mathscr{G}_{\mathsf{NW}}$ denotes the group of symmetries (Definition 2.23) of NW. The rows and columns of matrices in $\mathscr{G}_{\mathsf{NW}}$ are labelled by the ordered set $((0,0),\ldots,(d-1,d-1))$. Suppose $\mathbb{F}$ contains a d-th primitive root of unity ζ. We show that the following symmetries are present in $\mathscr{G}_{\mathsf{NW}}$. This helps in proving that NW is characterized by its symmetries over $\mathbb{F}$.

1. A set of diagonal matrices of the type $A = \mathrm{diag}(\beta_0,\ldots,\beta_0,\ldots,\beta_{d-1},\ldots,\beta_{d-1})$, where $\beta_i \in \mathbb{F}^\times$, for every $i \in \mathbb{F}_d$, every β_i appears exactly d times on the diagonal and $\prod_{i\in\mathbb{F}_d} \beta_i = 1$.
2. A set of *special* permutation matrices *corresponding to* the polynomials in $\mathbb{F}_d[z]_k$.
3. A set of *special* invertible diagonal matrices *corresponding to* ζ.

The exact descriptions of these symmetries are given in Claim 3.1.1. Suppose $f \in \mathbb{F}[\mathbf{x}]$ is a degree d homogeneous polynomial such that $\mathscr{G}_{\mathsf{NW}} \subseteq \mathscr{G}_f$. Then, the three types of matrices

mentioned above are also symmetries of f. The first type ensures that f is a set-multilinear polynomial. Thus, naturally every monomial of f looks like $x_{0,l_0} \cdots x_{d-1,l_{d-1}}$ and if h is obtained by interpolating $\{(i, l_i) : i \in \mathbb{F}_d\}$ then $\deg(h) \leq d-1$. The symmetries of the second type ensure that if a monomial $x_{0,l_0} \cdots x_{d-1,l_{d-1}}$ is present in f then f contains $x_{0,l_0+h(0)} \cdots x_{d-1,l_{d-1}+h(d-1)}$ for every $h \in \mathbb{F}_d[z]_k$ and the coefficients of all these monomials in f are same. The symmetries of the last type show that f does not contain a monomial $x_{0,h(0)} \cdots x_{d-1,h(d-1)}$, where $h \in \mathbb{F}_d[z]$ and $\deg(h) > k$. Hence, $f = \alpha \cdot \mathsf{NW}(\mathbf{x})$ for some $\alpha \in \mathbb{F}$.

Now suppose $\mathbb{F}$ is either $\mathbb{R}, \mathbb{Q}$ or a finite field satisfying $d \nmid |\mathbb{F}| - 1$. In this case, the matrices of first two types are still symmetries of NW but the matrices of the last type are not. We prove that in the absence of matrices of the third type from $\mathscr{G}_{\mathsf{NW}}$, NW is not characterized by its symmetries over $\mathbb{F}$. This proof is based on two results: One is the structure of $\mathscr{G}_{\mathsf{NW}}$ given in Chapter 4 of [Gup17], where we showed that $\mathscr{G}_{\mathsf{NW}}$ is generated by certain permutation and diagonal matrices (see Theorem 3.3), and the other is a structural insight about NW (see Claim 3.1.2) obtained from the analysis of the Lie algebra of NW given in Chapter 3 of [Gup17]. We exploit these properties to show that there exists a set-multilinear polynomial $f \in \mathbb{F}[\mathbf{x}]$ (the definition of f is given in Equation (3.2)), which is not a scalar multiple of NW but $\mathscr{G}_{\mathsf{NW}} \subseteq \mathscr{G}_f$.

The proofs of Theorems 1.1 and 1.2 are given in Section 3.1.1 of Chapter 3.

Circuit testing and flip theorem for NW: Proof ideas of Theorems 1.3 and 1.4

Theorem 1.3 gives a randomized polynomial time circuit testing algorithm for NW and Theorem 1.4 proves a flip theorem for NW. These two algorithmic results hinge on a neat structural property possessed by NW, known as the *characterization by circuit identities* (see Definition 2.25). This property says that there exists $\mathrm{poly}(d)$ many polynomial identities satisfied by NW, where every polynomial involved in these identities can be computed by a $\mathrm{poly}(d)$ size arithmetic circuit, and an $f \in \mathbb{F}[\mathbf{x}]$ satisfies these identities if and only if f is a scalar multiple of NW. We show in Lemma 3.2 of Chapter 3 that NW is characterized by circuit identities over any field. Some symmetries of NW play a crucial role in showing this characterization result.

The input of a circuit testing algorithm is black-box access to an arithmetic circuit C and the task is to determine whether $\mathsf{C} = \mathsf{NW}$. Suppose NW is computable by C. Since NW is characterized by circuit identities, there exist $\mathrm{poly}(d)$ many polynomial identities, which are also satisfied by C. In the algorithm, we check if C satisfies these identities by evaluating them on random points from $\mathbb{F}^{d^2}$. If it fails in any of these checks, we output 'Fail', and if all of these tests go through, $f = \alpha \cdot \mathsf{NW}$ for some $\alpha \in \mathbb{F}$. If it does not fail, we check the coefficient of an arbitrary monomial in f and if it is not 1, we again output 'Fail', else we output that C computes NW. The Schwartz-Zippel lemma (Fact 2.13) ensures that the algorithm works correctly with

high probability. The source of randomness in this algorithm is the Schwartz-Zippel lemma.

In case of a flip theorem, we are given that NW is not computed by any arithmetic circuit of size at most s. We pick $\text{poly}(d)$ many points, called *witness points*, uniformly at random from $\mathbb{F}^{d^2}$ and using the Schwartz-Zippel lemma we show that if C is an arithmetic circuit of size s then there exists a point $\mathbf{a}$ in the above collection, such that $\mathsf{C}(\mathbf{a}) \neq \mathsf{NW}(\mathbf{a})$ with high probability. Further, to show that a polynomial time black-box PIT for size-$10s$ arithmetic circuits implies that such witness points can be computed in deterministic polynomial time, we again use the fact that NW is characterized by circuit identities. As no size-s arithmetic circuit C computes NW, clearly $\alpha \cdot \mathsf{NW}$ is also not computed by a size-s arithmetic circuit for any $\alpha \in \mathbb{F}^{\times}$. Thus there exists a circuit identity, which is satisfied by NW, but not by C. On exploiting this and using a black-box PIT algorithm for size-$10s$ arithmetic circuits, we get a deterministic algorithm that computes a set of $\text{poly}(d)$ many witness points in polynomial time.

The proofs of Theorems 1.3 and 1.4 are given in Section 3.2 of Chapter 3.

Equivalence test for NW: Proof idea of Theorem 1.5

In this theorem, we give a randomized polynomial time block-diagonal permutation scaling equivalence test (in short, BD-PS equivalence test) for NW over finite fields satisfying $d \nmid (|\mathbb{F}|-1)$ and over $\mathbb{R}$. This test determines if there exists a BD-PS matrix (recall the definition of a BD-PS matrix from Section 1.3.1) C such that the input polynomial f satisfies $f = \mathsf{NW}(C\mathbf{x})$.

Now, we give a high level overview of the BD-PS equivalence test for NW given in Section 3.2 of Chapter 3. We assume that the input polynomial f is BD-PS equivalent to NW. Otherwise, we can detect with high probability that f is not equivalent to NW using the circuit testing algorithm for NW given in Theorem 1.3. The BD-PS ET for NW has two main steps: In the first step, it recovers a block-diagonal permutation matrix A (recall the definition of a block-diagonal permutation matrix from Section 1.3.1) such that $f(A^{-1}\mathbf{x})$ is *scaling equivalent* to NW, i.e., there exists an invertible scaling (or diagonal) matrix B such that $f(A^{-1}B^{-1}\mathbf{x}) = \mathsf{NW}$. This step *works over any field*. In the next step, we recover an invertible scaling matrix B such that $f(A^{-1}B^{-1}\mathbf{x}) = \mathsf{NW}$. This step works either over finite fields satisfying $d \nmid (|\mathbb{F}| - 1)$ or over $\mathbb{R}$.

The objective of Step 1 of the BD-PS ET is to construct d permutations $\sigma_0, \ldots, \sigma_{d-1}$ on $\mathbb{F}_d$ such that if A_{σ_i} is the $d \times d$ permutation matrix corresponding to σ_i for every $i \in \mathbb{F}_d$ and $A = \text{diag}(A_{\sigma_0}, \ldots, A_{\sigma_{d-1}})$, then $f(A^{-1}\mathbf{x})$ is scaling equivalent to NW. The algorithm starts by considering a canonical form of $\sigma_0, \ldots, \sigma_{d-1}$ (see Claim 3.2.3) and efficiently constructs a set of "nice polynomials" in $\mathbb{F}_d[z]_k$ (see Definition 3.1 and Claim 3.2.4). Using these, we first compute $d-k$ entries of every σ_i [1] efficiently (see Claim 3.2.5). Then, we compute the remaining entries

[1] We can treat a permutation on $\mathbb{F}_d$ by a size-d tuple.

of these permutations (see Claims 3.2.6 and 3.2.7). This is how we compute $\sigma_0, \ldots, \sigma_{d-1}$. This step of the BD-PS ET crucially uses symmetries and the low-intersection property of NW.

In the second step of the BD-PS ET, we again pick a useful subset of $\mathbb{F}_d[z]_k$ (see Step 2 of Algorithm 8), using which we compute a scaling matrix B such that $f(B^{-1}A^{-1}\mathbf{x}) = \mathsf{NW}$. Using this set, we compute a 0-1 matrix, which helps us obtaining an important system of linear equations corresponding to NW. We get this system either over finite fields satisfying $d \nmid (|\mathbb{F}| - 1)$ (see Equation (3.10)) or $\mathbb{F} = \mathbb{R}$ (see Equation (3.11)) and solve it to compute the required invertible scaling matrix B. It is due to this reason that the BD-PS ET holds over these fields. A similar system of linear equation was immensely helpful in obtaining the complete understanding of the Lie algebra of NW given in [Gup17] (see Claim 3.1.2).

1.4.2 ET for the determinant over finite fields and over $\mathbb{Q}$

We give proof overviews of the theorems stated in Section 1.3.2. The proofs of these theorems are given in Chapter 4. Recall that $\mathsf{Det}_n(\mathbf{x}) = \det(X)$, where $X = (x_{i,j})_{i,j\in[n]}, x_{i,j}$ is a variable for every $i, j \in [n]$, $\mathbf{x} = \{x_{1,1}, \ldots, x_{n,n}\}$, and $M_n(\mathbb{F})$ is the set of $n \times n$ matrices over $\mathbb{F}$.

Determinant equivalence test: Proof ideas of Theorems 1.6, 1.7 and 1.8

The input of DET is black-box access to $f \in \mathbb{F}[\mathbf{x}], \deg(f) = n$, where $\mathbb{F}$ is either a finite field satisfying the conditions given in Theorem 1.6 or $\mathbb{F} = \mathbb{Q}$. We can assume without loss of generality that f is equivalent to Det_n. Otherwise, it would be detected with high probability that f is not equivalent to Det_n using the Schwartz-Zippel lemma. Let $A \in \mathrm{GL}(n^2, \mathbb{F})$ such that $f = \mathsf{Det}_n(A\mathbf{x})$. The main component of our DET algorithms is Theorem 1.8, which gives a randomized polynomial time reduction from DET to the FMAI problem (recall FMAI from Section 1.3.2). This reduction works over *almost any field*. We first give a proof idea of this reduction and then complete the DET algorithm by invoking FMAI algorithm. FMAI algorithms are known over finite fields and $\mathbb{Q}$ (see Section 2.2.4). For discussing the reduction of DET to FMAI, let $\mathbb{F}$ be a field satisfying mild conditions given in Theorem 1.8.

Reducing DET to FMAI. The reduction of DET to FMAI is obtained by exploiting the Lie algebra (Definition 2.30) of the input polynomial f, denoted $\mathfrak{g}_f$. The Lie algebra of a polynomial is a useful tool for ET as the Lie algebras of equivalent polynomials are conjugates of each other, i.e., $\mathfrak{g}_f = A^{-1} \cdot \mathfrak{g}_{\mathsf{Det}_n} \cdot A$ (see Fact 2.10). $\mathfrak{g}_{\mathsf{Det}_n}$ is well-studied and has a nice structure (see the following paragraph). The same structure gets induced to $\mathfrak{g}_f$ via the above mentioned conjugacy relation. We start with talking about the structure of $\mathfrak{g}_{\mathsf{Det}_n}$.

It is known that over a field $\mathbb{F}$ satisfying $char(\mathbb{F}) \nmid n$, $\mathfrak{g}_{\mathsf{Det}_n} = \mathscr{L}_{\text{row}} \oplus \mathscr{L}_{\text{col}}$, where $\mathscr{L}_{\text{row}} :=$

$\{A \otimes I_n : A \in Z_n\}$ and $\mathscr{L}_{\text{col}} := \{I_n \otimes A : A \in Z_n\}$ (see Definition 2.20), where $Z_n = \{A \in M_n(\mathbb{F}) : \text{trace}(A) = 0\}$. A proof of this fact can be found in Section 3.2 of Chapter 3 in [Nai19]. Then, $\mathfrak{g}_f = \mathscr{F}_{\text{row}} \oplus \mathscr{F}_{\text{col}}$, where $\mathscr{F}_{\text{row}} = A^{-1} \cdot \mathscr{L}_{\text{row}} \cdot A$ and $\mathscr{F}_{\text{col}} = A^{-1} \cdot \mathscr{L}_{\text{col}} \cdot A$. Clearly, $\dim \mathscr{F}_{\text{row}} = \dim \mathscr{F}_{\text{col}} = n^2 - 1$. Let $r := n^2 - 1$. In the algorithm, we exploit this structural richness of $\mathfrak{g}_f$ by decomposing it and getting hold of $\mathscr{F}_{\text{row}}$ and $\mathscr{F}_{\text{col}}$. An important property of the Lie algebra of a polynomial that is useful here is that given black-box access to a polynomial f, we can compute a basis of $\mathfrak{g}_f$ in randomized polynomial time (see Fact 2.16). Now, we give an overview of the reduction algorithm, which works over almost every field.

We first compute a basis $\{B_1, \ldots, B_{2r}\}$ of $\mathfrak{g}_f$ using Fact 2.16 from black-box of f. Now, using $\{B_1, \ldots, B_{2r}\}$, we compute a "special set" of matrices $\mathscr{P} = \{P_1, \ldots, P_{2r}\} \subseteq M_{2r}(\mathbb{F})$, which correspond to some specific linear operators on $\mathfrak{g}_f$ (the description of these linear operators are given in Section 4.3.1 of Chapter 4). Such an operator looks as follows: For $E \in \mathfrak{g}_f, \rho_E : \mathfrak{g}_f \to \mathfrak{g}_f$ maps an $F \in \mathfrak{g}_f$ to $EF - FE$. Then, $\mathscr{P}$ contains the matrices of $\rho_E, E \in \mathfrak{g}_f$, where the rows and columns of these matrices are labelled by the basis $(B_1, \ldots, B_{2r})$. The set $\mathscr{P}$ is intimately related to $\mathfrak{g}_f$ as follows - we prove that $\mathscr{F}_{\text{col}}$ and $\mathscr{F}_{\text{row}}$ are the only *irreducible invariant subspaces* (Definition 2.16) of $\mathscr{P}$ (see Lemma 4.1 in this regard). Thus, the objective now is to get hold of the irreducible invariant subspaces of $\mathscr{P}$. This is done as follows: We pick a random matrix Q from $\mathscr{P}$ and compute its characteristic polynomial, which is denoted $h(z)$. After that, we factorize $h(z)$ and let $h_1, \ldots, h_k$ be irreducible factors of $h(z)$ such that none of the h_i is a variable. For every $i \in [k]$, we compute a basis of the null space of $h_i(Q)$, denoted $\mathscr{N}_i$. Then, we compute the $\mathscr{P}$-closure (Definition 2.17) of bases vectors of $\mathscr{N}_i, i \in [k]$. The set of these closures only contains the spaces $\mathscr{F}_{\text{row}}$ and $\mathscr{F}_{\text{col}}$ (see Lemma 4.1).

The reason $\mathscr{F}_{\text{col}}$ is interesting at this point because the $\mathbb{F}$-algebra $\mathscr{A}$ generated by an $\mathbb{F}$-basis of $\mathscr{F}_{\text{col}}$ is isomorphic to $M_n(\mathbb{F})$. Let $L_1, \ldots, L_{n^2}$ be an $\mathbb{F}$-basis of $\mathscr{A}$ computed from a basis of $\mathscr{F}_{\text{col}}$. Hence we invoke FMAI algorithm on $L_1, \ldots, L_{n^2}$. As $\mathscr{A}$ is isomorphic as an $\mathbb{F}$-algebra to $M_n(\mathbb{F})$, the FMAI algorithm returns an $\mathbb{F}$-algebra isomorphism $\varphi : \mathscr{A} \to M_n(\mathbb{F})$. The Skolem-Noether theorem (Theorem 2.1) gives us useful information about the structure of φ (see Claim 4.3.8), using which we compute an $A \in \text{GL}(n^2, \mathbb{F})$ such that $f = \mathsf{Det}_n(A\mathbf{x})$. This is how the reduction from FMAI to DET works and this completes the proof overview of Theorem 1.8. After reducing DET to FMAI, we invoke known FMAI algorithms over finite fields and $\mathbb{Q}$ and this is how we get DET over finite fields and $\mathbb{Q}$.

The detailed proof of Theorems 1.6, 1.7 and 1.8 are given in Section 4.2 of Chapter 4.

Reduction of Intfact to DET over $\mathbb{Q}$: Proof idea of Theorem 1.9

In this theorem, assuming GRH, we give a randomized polynomial time reduction from the problem of factoring square-free integers to DET for quadratic forms over $\mathbb{Q}$. The proof of this theorem is based on a result from [Ron87], which assuming GRH gives a randomized polynomial time reduction from the problem of factoring square-free integers to the following problem: Given $a, b \in \mathbb{Q}^\times$, find $x, y, z \in \mathbb{Q}$ (not all zero) such that $x^2 - ay^2 - bz^2 = 0$, if such a solution exists. Let $a, b \in \mathbb{Q}^\times$ and $f(\mathbf{x}) := x_{1,1}^2 - ax_{1,2}^2 - bx_{2,1}^2 + abx_{2,2}^2$. We show that f satisfies $f = \mathsf{Det}_2(A\mathbf{x})$ for some $A \in \mathrm{GL}(4, \mathbb{Q})$ if and only if $x^2 - ay^2 - bz^2 = 0$ has a non-zero solution over $\mathbb{Q}$. This gives a randomized polynomial time reduction from factoring square-free integers to ET for Det_2 over $\mathbb{Q}$. The proof of Theorem 1.9 is given in Section 4.4 of Chapter 4.

Reduction of FMAI to DET: Proof idea of Theorem 1.10

Let $n \in \mathbb{N}$ and $\mathbb{F}$ be a field satisfying $char(\mathbb{F}) \nmid n$. We are given a basis of an $\mathbb{F}$-algebra $\mathscr{A} \subseteq M_{n^2}(\mathbb{F})$, oracle access to DET over $\mathbb{F}$ and we want to decide whether $\mathscr{A}$ is isomorphic as an $\mathbb{F}$-algebra to $M_n(\mathbb{F})$. If yes, we also want to output an $\mathbb{F}$-algebra isomorphism $\varphi : \mathscr{A} \to M_n(\mathbb{F})$.

The algorithm first compute a "special set" $\{L_{i,j} \in M_{n^2}(\mathbb{F}) : i, j \in [n]\}$, where the entries of every $L_{i,j}$ correspond to the *structural constants* of $\mathscr{A}$ [1]. The details of matrices in this set are given in Step 2 of Algorithm 11. Using the Skolem-Noether theorem, we show in Claim 4.5.1 that if $\mathscr{A}$ is isomorphic to $M_n(\mathbb{F})$ then there exists a $K \in \mathrm{GL}(n^2, \mathbb{F})$ such that for every $i, j \in [n], L_{i,j} = K^{-1} \cdot (I_n \otimes C_{i,j}) \cdot K$, where $C_{1,1}, \ldots, C_{n,n}$ is an $\mathbb{F}$-basis of $M_n(\mathbb{F})$. Let $\tilde{L} = \{\tilde{L}_{i,j} : i, j \in [n]\}$, where for every $i, j \in [n], \tilde{L}_{i,j}$ is the traceless part of $L_{i,j}$, i.e. $\tilde{L}_{i,j} := L_{i,j} - \frac{\mathrm{trace}(L_{i,j})}{n^2} I_{n^2}$. Then, it follows from above that if $\mathscr{A}$ is isomorphic to $M_n(\mathbb{F})$ then $\langle \tilde{L} \rangle = K^{-1} \cdot \mathscr{L}_{\mathrm{col}} \cdot K$. Another important result we show is that if $f \in \mathbb{F}[\mathbf{x}]$ is such that $\mathscr{L}_{\mathrm{col}} \subseteq \mathfrak{g}_f$ then $f = \alpha \cdot \mathsf{Det}_n$, where $\alpha \in \mathbb{F}^\times$ (see Lemma 4.5), i.e., the determinant is *characterized by its Lie algebra*. We now show how these two results imply a reduction from FMAI to DET.

Compute $L_{i,j}, i, j \in [n]$ using a basis of $\mathscr{A}$ and then using these, compute a basis of $\langle \tilde{L} \rangle$. Then, we construct a polynomial $f \in \mathbb{F}[\mathbf{x}]$ such that $\langle \tilde{L} \rangle$ is an $\mathbb{F}$-subspace of $\mathfrak{g}_f$. This step takes $n^{O(n)}$ many field operations as we compute f in a brute force manner. Using the facts that determinant is characterized by its Lie algebra and if $f' = \mathsf{Det}_n(B\mathbf{x})$ for some $B \in \mathrm{GL}(n^2, \mathbb{F})$ then $\mathfrak{g}_{f'} = (B^{-1} \cdot \mathscr{L}_{\mathrm{row}} \cdot B) \oplus (B^{-1} \cdot \mathscr{L}_{\mathrm{col}} \cdot B)$, we get that $\mathscr{A}$ is isomorphic as an $\mathbb{F}$-algebra to $M_n(\mathbb{F})$ if and only if $f = \alpha \cdot \mathsf{Det}_n$. Thus, by running DET on f, we get either an $A \in \mathrm{GL}(n^2, \mathbb{F})$ such that $f = \mathsf{Det}_n(A\mathbf{x})$ or 'Fail' depending on whether f is isomorphic to Det_n or not. In the former case, we use A to compute an $\mathbb{F}$-algebra isomorphism from $\mathscr{A}$ to $M_n(\mathbb{F})$, and in the

[1] Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be an $\mathbb{F}$-basis of an $\mathbb{F}$-algebra $\mathscr{A}$. Then, for every $i, j \in [m]$ there exist $\alpha_{i,j,k}, k \in [m]$ such that $\mathbf{u}_i \cdot \mathbf{u}_j = \sum_{k \in [m]} \alpha_{i,j,k} \mathbf{u}_k$. Then, $\alpha_{i,j,k}, i, j, k \in [m]$ are called as the structure constants of $\mathscr{A}$.

latter case, we output '$\mathscr{A}$ is not isomorphic to $M_n(\mathbb{F})$'.

A complete proof of Theorem 1.10 is given in Section 4.5 of Chapter 4.

1.4.3 ET for regular ROFs: Proof idea of Theorem 1.11

We are given black-box access to an n-variate polynomial $f(\mathbf{x})$, which is in the orbit of a regular ROF C, i.e., $f = \mathsf{C}(B\mathbf{x} + \mathbf{d})$, where $B \in \mathrm{GL}(n, \mathbb{F})$ and $\mathbf{d} \in \mathbb{F}^n$. The objective is to compute an $A \in \mathrm{GL}(n, \mathbb{F})$ such that $f(A\mathbf{x}) = \mathsf{C}(PS\mathbf{x} + \mathbf{b})$, where $P, S \in \mathrm{GL}(n, \mathbb{F})$ are permutation and scaling matrices respectively and $\mathbf{b} \in \mathbb{F}^n$.

The most important ingredient of the ET for regular ROFs is the Hessian determinant (Definition 2.27) of a regular ROF. Let us first try to understand with an example how the Hessian determinant can be used in an ET for regular ROFs. Suppose $\mathsf{C} = x_1x_2x_3 + x_4x_5x_6$. Then, it is easy to show that over fields not having characteristic equal to 2, $\det(H_{\mathsf{C}}) \neq 0$ and every x_i is a factor of $\det(H_{\mathsf{C}})$. Let $f = \mathsf{C}(B\mathbf{x} + \mathbf{d})$, where $B \in \mathrm{GL}(n, \mathbb{F})$ and $\mathbf{d} \in \mathbb{F}^n$. Suppose $B\mathbf{x}+\mathbf{d} = (\ell_1 \cdots \ell_6)^T$. Then, $f = \ell_1\ell_2\ell_3+\ell_4\ell_5\ell_6$. An important property of the Hessian determinants of equivalent polynomials given in Corollary 2.1 implies that

$$\det(H_f) = (\det(B))^2 \det(H_{\mathsf{C}})(B\mathbf{x} + \mathbf{d}).$$

Then it follows that over fields not having characteristic equal to 2, $\det(H_f) \neq 0$ and for every $i \in [6]$, ℓ_i is a factor of $\det(H_f)$. This information is good enough for designing an ET for C. From black-box of f, we compute black-box access to $\det(H_f)$, which can be done efficiently due to Fact 2.14 and then factorize $\det(H_f)$ using the algorithm in [KT90] (see Fact 2.17). This algorithm gives us black-box access to $\alpha_i\ell_i, i \in [6]$, where $\alpha_i \in \mathbb{F}^\times$ and $\alpha_1 \cdots \alpha_6 = 1$. We reconstruct $\alpha_i\ell_i$ for every $i \in [6]$ and then compute an $A \in \mathrm{GL}(6, \mathbb{F})$, which maps every $\alpha_i\ell_i$ to a distinct variable. Then, $f(A\mathbf{x}) = \mathsf{C}(PS\mathbf{x} + \mathbf{d})$, where $P, S \in \mathrm{GL}(6, \mathbb{F})$ are permutation and scaling matrices and $\mathbf{d} \in \mathbb{F}^6$.

One can ask how to use this idea to design an equivalence test for arbitrary regular ROFs, where it may no longer be true that every variable appearing in a regular ROF C is a factor of $\det(H_{\mathsf{C}})$. In [Kay11], Kayal gave a promising approach in this direction. We first mention this approach, and then show how to adapt it in our setting.

A basic approach: Let $h = h_1(\mathbf{x}_1) + h_2(\mathbf{x}_2)$, where $\mathbf{x}_1 \cap \mathbf{x}_2 = \emptyset$. Kayal gave a randomized polynomial time algorithm (see Theorem 7.2 in [Kay11]) that takes black-box access to $g := h(B\mathbf{x})$, where $B \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ and computes an $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ such that $g(A\mathbf{x})$ is the sum of two variable disjoint polynomials, provided the *number of essential variables* (Definition 2.31)

in $\det(H_h)$ is equal to $|\mathbf{x}|$. We denote the number of essential variables of a $p \in \mathbb{F}[\mathbf{x}]$ by $N_{ess}(p)$.

Let $\mathsf{C} = T_1 + \cdots + T_s + \gamma$ [1] where for every $k \in [s]$, T_k is a $\times$-rooted regular ROF, i.e., the root of T_k is a $\times$ gate, and $\gamma \in \mathbb{F}$. Then, $f = \widehat{T}_1 + \cdots + \widehat{T}_s + \gamma$, where for every $k \in [s]$, $\widehat{T}_k = T_k(B\mathbf{x}+\mathbf{d})$. We call $T_1, \ldots, T_s$ and $\widehat{T}_1, \ldots, \widehat{T}_s$ as the terms of C and f respectively. Without loss of generality, assume that there exists an $s_1 \in [s]$ such that for every $k \in [s_1], \deg(\widehat{T}_k) \geq 3$ and for every $l \in \{s_1+1, \ldots, s\}, \deg(\widehat{T}_l) = 2$. Let $q = T_{s_1+1} + \cdots + T_s$ and $\widehat{q} = \widehat{T}_{s_1+1} + \cdots + \widehat{T}_s$. We call q and $\widehat{q}$ as the quadratic terms of C and f respectively, and $T_1, \ldots, T_{s_1}$ and $\widehat{T}_1, \ldots, \widehat{T}_{s_1}$ as non-quadratic terms of C and f respectively. As C is the sum of variable disjoint polynomials and we are given black-box access to f in the orbit of C, the basic approach mentioned above seems useful. This is so because if we are able to make the terms of f variable disjoint, then after that we can get hold of its terms one by one, factorize the terms and solve the problem for each of these factors recursively, as each of these factors is an input instance having product-depth[2] less than the product-depth of f. However, there is a challenge in obtaining *efficient* black-box access to these factors even after making the terms of f variable disjoint. We talk about this challenge later and show how we handle it. Let us first see how to make the terms of f variable disjoint.

Making terms variable disjoint. The first phase of our ET algorithm computes an invertible matrix A such that the terms of $f(A\mathbf{x})$ are variable disjoint. However, there are some technical challenges in implementing the basic approach, which we list below.

1. $\det(H_{\mathsf{C}})$ *can be equal to 0*: It might happen that the Hessian determinant of an arbitrary regular ROF is zero over the underlying field. For example, if $\mathsf{C} = x_1x_2x_3$ then $\det(H_{\mathsf{C}}) = 0$ over the fields having characteristic equal to two.

2. $N_{ess}(\det(H_{\mathsf{C}})) < |\mathrm{var}(\mathsf{C})|$: Let $\mathsf{C} = x_1x_2x_3 + x_4x_5$. Then, $N_{ess}(\det(H_{\mathsf{C}})) = 3$.

In the presence of these two hurdles, we can not implement the basic approach directly. We now show how we handle these challenges. We overcome the first challenge by proving in Lemma 5.1 of Chapter 5 that if C is a regular ROF and $\mathbb{F}$ satisfies either $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) \geq |\mathrm{var}(\mathsf{C})|$ then $\det(H_{\mathsf{C}}) \neq 0$. We argue this by studying the structures and coefficients of some "nice monomials" in $\det(H_{\mathsf{C}})$. We address the second challenge as follows: We prove in Claim 5.1.3 of Chapter 5 that if T is a $\times$-rooted regular ROF having at least three variables

[1] For the equivalence testing problem, we assume without loss of generality that C is $+$-rooted. Otherwise, using the polynomial factorization algorithm in [KT90], we can reduce an ET for a $\times$-rooted regular ROF to an ET for a $+$-rooted regular ROF.

[2] The product-depth of an ROF C having alternate layers of $+$ and $\times$ gates is the maximum number of $\times$ gates in any root to leaf path in C. The product-depth of a polynomial in the orbit of C is same as the product-depth of C.

then every variable appearing in T is essential for $\det(H_T)$. This result also follows from the structure of nice monomials and uses the regularity of T crucially. Our analysis of $\det(H_{\mathsf{C}})$ is quite long due to the study of the structures and the coefficients of some explicit monomials in it. Soon after this analysis, Bhargav Thankey independently came up with a shorter analysis of the above mentioned properties of $\det(H_{\mathsf{C}})$. This analysis is given in [GST22]. The proof approach of Thankey is totally different from ours. The main difference between our proof and Thankey's proof is that we study the structures and coefficients of some *explicit monomials* of $\det(H_{\mathsf{C}})$ whereas Thankey showed that $\det(H_{\mathsf{C}})$ contains non-zero monomials without giving the details of the structures and coefficients of such monomials. However, this information is good enough for obtaining all the properties of the Hessian determinant of a regular ROF required for designing an ET for regular ROFs. See Remark 5.1 in this regard.

Having the above two results in hand, and using the basic approach mentioned before, we first compute an $A_0 \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ such that the non-quadratic terms of $f(A_0\mathbf{x})$ are variable disjoint. Let the variables in non-quadratic terms of $f(A_0\mathbf{x})$ be $\mathbf{z}$ and $\mathbf{y} := \mathbf{x} \setminus \mathbf{z}$. Let $|\mathbf{y}| = 2m$. We compute black-box access to the homogeneous degree two part q' of $f(A_0\mathbf{x})$ in $\mathbf{y}$-variables. Then, we invoke a QFE algorithm over the underlying field on q' and $q'' := y_{1,1}y_{1,2}+\cdots+y_{m,1}y_{m,2}$. It returns an $A_1' \in \mathrm{GL}(|\mathbf{y}|, \mathbb{F})$ such that $q'(A_1\mathbf{y}) = q''$. We extend A_1' to $A_1 \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ such that A_1 and A_1' maps every $\mathbf{y}$-variable to the same linear form in $\mathbf{y}$-variables and A_1 maps every $\mathbf{z}$-variable to itself. It turns out that $\widehat{q}(A_1A_0\mathbf{x}) = (y_{1,1} + h_{1,1})(y_{1,2} + h_{1,2}) + \cdots + (y_{m,1} + h_{m,1})(y_{m,2} + h_{m,2})$, where for every $i \in [m], j \in [2], h_{i,j} \in \mathbb{F}[\mathbf{z}]$ is a linear polynomial. Thereafter, for every $i \in [m], j \in [2]$, we compute first order partial derivative of $f(A_1A_0\mathbf{x})$ with respect to $y_{i,j}$ and get access to $y_{i,j'} + h_{i,j'}$, where $j' \in [2] \setminus \{j\}$. For $i \in [m], j \in [2]$, let $h_{i,j} = h'_{i,j} + \beta_{i,j}$, where $h'_{i,j} \in \mathbb{F}[\mathbf{z}]$ is a linear form and $\beta_{i,j} \in \mathbb{F}$. We compute an $A_2 \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$, which maps every $y_{i,j} + h'_{i,j}$ to $y_{i,j}$ and every $\mathbf{z}$-variable to itself. Let $A = A_2A_1A_0$. Thus, at the end of this phase, $f(A\mathbf{x})$ is the sum of variable disjoint terms.

Now, we give an overview of the other steps of our ET algorithm.

Computing black-box access to the terms of $f(A\mathbf{x})$. The matrix A computed in the previous step ensures that for every $l \neq k \in [s], \widehat{T}_l(A\mathbf{x})$ and $\widehat{T}_k(A\mathbf{x})$ are variable disjoint and $\widehat{q}(A\mathbf{x}) = (y_{1,1} + \beta_{1,1})(y_{1,2} + \beta_{1,2}) + \cdots + (y_{m,1} + \beta_{m,1})(y_{m,2} + \beta_{m,2})$, where $\beta_{i,j} \in \mathbb{F}$ for every $i \in [m], j \in [2]$. At this point, we want to get black-box access to $\widehat{T}_1(A\mathbf{x}), \ldots, \widehat{T}_{s_1}(A\mathbf{x})$ from black-box of f. We only need black-box access to the non-quadratic terms of f as the quadratic term of f has already been handled. We have the following third technical challenge in this context (the other two challenges have been mentioned above).

3. How to compute black-box access to $\widehat{T}_k(A\mathbf{x})$ *efficiently* for every $k \in [s_1]$?

Let us first explain the meaning of efficient computation of black-box access to $\widehat{T}_k(A\mathbf{x})$. We want to compute black-box access to $\widehat{T}_k(A\mathbf{x})$ using *only one* black-box query to f. Our algorithm recurses on the product-depth of C. Due to this, if we make more that one black-box query to f, the algorithm might have exponential running time in n as the product-depth of C can be as large as n. We now show how to obtain black-box access to $\widehat{T}_k(A\mathbf{x})$ using only one query to f.

First, the algorithm learns the variable sets $\mathbf{z}_1, \ldots, \mathbf{z}_{s_1}$ of $\widehat{T}_1(A\mathbf{x}), \ldots, \widehat{T}_{s_1}(A\mathbf{x})$ using double derivatives (see Observation 5.4). Then, it picks a $k \in [s_1]$ arbitrarily and substitutes every variable present in $\mathbf{x} \setminus \mathbf{z}_k$ with 0. Let $T'_k = \widehat{T}_k(A\mathbf{x})$. Thus, we get black-box access to

$$g := T'_k + \gamma',$$

where $\gamma' \in \mathbb{F}$ is unknown. Now, the goal is to learn γ' because after learning γ' we can subtract it from black-box of g and get black-box access to T'_k. The following useful observation comes in handy here: Since g is a $+$-rooted ROF, Observation 2.7 implies that g is irreducible. Hence, $g - \beta'$ is reducible for some $\beta' \in \mathbb{F}$ if and only if $\beta' = \gamma'$. This uniqueness of γ' will play a key role in the discovery of γ'. Now, we show how to compute γ' efficiently. At this point, the Hessian determinant becomes useful again.

Now, lets see how the algorithm gets hold of γ'. First it computes $\det(H_g)$ using Fact 2.14, then factorizes $\det(H_g)$ using the algorithm in [KT90]. We classify the factors of $\det(H_g)$ into two categories, which we call "good factors" and "bad factors". We will describe these two one by one. First we talk about good factors. Note that $\det(H_g) = \det(H_{T'_k})$. As $\deg(T'_k) \geq 3$, Corollary 5.1 of Chapter 5 implies that there exists a child $Q'_{k,j}$ of the topmost $\times$ gate of T'_k [1] such that $\alpha_{k,j}Q'_{k,j}$ is an irreducible factor of $\det(H_g)$ for some $\alpha_{k,j} \in \mathbb{F}^\times$. Such factors of $\det(H_g)$ as good factors. All the remaining non-constant factors of $\det(H_g)$ are called bad factors.

After factorizing $\det(H_f)$, the algorithm picks a non-constant factor q of $\det(H_g)$. We first set up a useful notation. Let t be a fresh variable. For every $i \in [|\mathbf{z}_k|]$, let c_i be chosen randomly from a large enough finite set of $\mathbb{F}$. Let $\mathbf{z}_k = \{z_{k,i} : i \in [|\mathbf{z}_k|]\}$. Define $\pi : \mathbb{F}[\mathbf{z}_k] \to \mathbb{F}[t]$ as follows: For every $p(z_{k,1}, \ldots, z_{k,|\mathbf{z}_k|}) \in \mathbb{F}[\mathbf{z}_k]$, $\pi(p) := p(c_1 t, \ldots, c_{|\mathbf{z}_k|}t)$. It follows from the Schwartz-Zippel lemma that $\pi(g)$ and $\pi(q)$ are non-constant polynomials with high probability, provided $\mathbb{F}$ is large enough. Now, the algorithm applies π to black-boxes of g and q and interpolates $\pi(g)$ and $\pi(q)$. This can be done efficiently as $\pi(g)$ and $\pi(q)$ are univariate polynomials. Now,

[1] As T'_k is in the orbit of the $\times$-rooted regular ROF T_k, we can view T'_k as a $\times$-rooted formula, which is obtained from T_k be replacing the leaf nodes labelled with variables with the nodes labelled with corresponding affine forms obtained from $BA\mathbf{x} + \mathbf{d}$.

consider the following equation, where the coefficients of $p$and a_0 are formal variables.

$$\pi(g) = p \cdot \pi(q) + a_0. \tag{1.2}$$

Suppose $\mathbf{a}$ is the set containing the coefficients of p and a_0. The algorithm solves the system of linear equations in $\mathbf{a}$-variables formed by comparing the coefficients of monomials in t in the L.H.S. and the R.H.S. of Equation (1.2). Now, let us see how the algorithm gets access to γ'. We analyse the behaviour of the algorithm in the following two steps.

Case 1: q is a good factor. Suppose $q = \alpha_{k,j}Q'_{k,j}$. In this case, the system of linear systems mentioned above has a solution: One of the solutions is obtained by setting $p = \pi\left(\alpha_{k,j}^{-1}\frac{T'_k}{Q'_{k,j}}\right)$ and $a_0 = \gamma'$. Using the fact that $\deg(p) > 1$ with high probability, it is not a difficult task to show that this system of linear equations in $\mathbf{a}$-variables has a unique solution. Hence, when the algorithm picks a good factor of $\det(H_g)$, it always gets hold of γ'.

Case 2: q is a bad factor. If the algorithm picks a bad factor then it solves the above system of linear equations in $\mathbf{a}$-variables, computes some p and a_0, and checks whether $g - a_0$ is reducible or not using the algorithm in [KT90]. If yes, it returns black-box access to $g - a_0$. The fact that $g - a_0$ is reducible if and only if $a_0 = \gamma'$ ensures that the algorithm works correctly.

The presence of a good factor in the list of factors of $\det(H_g)$ ensures that the algorithm computes γ' at some point of time. Once the algorithm has γ', it subtracts γ' from black-box of g to get black-box-access to T'_k.

Recursively solving factors of T'_k: Suppose $T'_k = Q'_{k,1}\cdots Q'_{k,m_k}$. After getting black-box access to T'_k, the algorithm factorizes it using the algorithm in [KT90] and obtains black-box access to $\alpha'_{k,1}Q'_{k,1}, \ldots, \alpha'_{k,m_k}Q'_{k,m_k}$, where $\alpha'_{k,1}, \ldots, \alpha'_{k,m_k} \in \mathbb{F}^\times$ and $\alpha'_{k,1}\cdots\alpha'_{k,m_k} = 1$. Using Claim 2.2.2, the algorithm computes an $A' \in \mathrm{GL}(|\mathbf{z}_k|, \mathbb{F})$ such that $\alpha'_{k,1}Q'_{k,1}(A'\mathbf{z}_k), \ldots, \alpha'_{k,m_k}Q'_{k,m_k}(A'\mathbf{z}_k)$ are variable disjoint. Let $j \in [m_k]$ be picked arbitrarily and φ be the map on $\mathbf{z}_k$, which substitutes variables of $\alpha'_{k,i}Q'_{k,i}(A'\mathbf{z}_k)$ for every $i \in [m_k]\setminus\{j\}$ uniformly at random from a large enough finite subset of $\mathbb{F}$. For $i \in [m_k] \setminus \{j\}$, let $\varphi(\alpha'_{k,i}Q'_{k,i}(A'\mathbf{z}_k)) = \beta_{k,i}$. The Schwartz-Zippel lemma ensures that with high probability, $\beta_{k,i} \neq 0$ for every $i \in [m_k] \setminus \{j\}$. Note that we know $\beta_{k,i}$ for every $i \in [m_k] \setminus \{j\}$. The algorithm first computes black-box access to $T'_k(A'\mathbf{z}_k)$ from black-box of T'_k. Since $\alpha'_{k,j}Q'_{k,j}(A'\mathbf{z}_k) = \varphi(T'_k(A'\mathbf{z}_k))\prod_{i\in[m_k]\setminus\{j\}}\beta_{k,i}^{-1}$, by applying φ to $T'_k(A'\mathbf{z}_k)$ and then multiplying it with $\prod_{i\in[m_k]\setminus\{j\}}\beta_{k,i}^{-1}$, the algorithm gets black-box access to $\alpha'_{k,j}Q'_{k,j}(A'\mathbf{z}_k)$. After this, the algorithm recurses on $\alpha'_{k,j}Q'_{k,j}(A'\mathbf{z}_k)$.

The detailed proof of Theorem 1.11 is given in Chapter 5.

1.5 Organization

The remaining part of this thesis is organized as follows: In Chapter 2, we state the notations and give some preliminary results. These include some useful definitions, results from linear algebra and algebra, results related to partial derivatives, Hessian, symmetries, Lie algebra, essential and redundant variables of a polynomial and some important algorithmic facts. In Chapter 3, we prove the theorems about the Nisan-Wigderson polynomial stated in Section 1.3.1. Chapter 4 contains the proofs of the theorems on equivalence test for the family of determinant stated in Section 1.3.2. In Chapter 5, we give a randomized polynomial time ET for the class of regular ROFs. This ET uses some important properties of the Hessian determinant of a regular ROF like its non-zeroness, knowledge of its essential variables and factors. Chapter 6 is devoted to the study of these properties of the Hessian determinant of a canonical ROF (Definition 2.39) and as a regular ROF is canonical, all these properties hold for the Hessian determinant of a regular ROF. We give a conclusion of the main contributions of this thesis and mention some open questions in Chapter 7.

Chapter 2

Preliminaries

In this chapter, we present some notations and recall some basic mathematical and algorithmic results, which would be required for the later chapters. This chapter has two sections - the first one contains useful structural results and the next one is devoted to some important algorithmic results.

Notations. The sets of natural numbers, integers, rational numbers, real numbers and complex numbers are denoted by $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ respectively. Let $\mathbb{N}^\times = \mathbb{N} \setminus \{0\}$. Unless otherwise specified, for $n \in \mathbb{N}^\times, [n] := \{1, \ldots, n\}$ and for $m, n \in \mathbb{N}, m < n, [m, n] = \{m, \ldots, n\}$. For a field $\mathbb{F}$, $\mathbb{F}^\times = \mathbb{F} \setminus \{0\}$, elements of $\mathbb{F}$ are represented by lower case Greek letters like α, β, γ, and vectors over $\mathbb{F}$ are denoted by boldface letters like $\mathbf{a}, \mathbf{b}, \mathbf{d}$. For $n \in \mathbb{N}, M_n(\mathbb{F})$ and $\mathrm{GL}(n, \mathbb{F})$ denote the set of $n \times n$ matrices and the set of $n \times n$ invertible matrices over $\mathbb{F}$ respectively, and I_n denotes the $n \times n$ identity matrix. The elements of $M_n(\mathbb{F})$ are denoted by upper case Roman alphabets like A, B, C. Vector spaces over $\mathbb{F}$ are represented by U, W and $\mathrm{End}_{\mathbb{F}}(U) := \{\varphi : U \to U\}$ is the set of $\mathbb{F}$-linear maps (Definition 2.10). Given a set S, $\langle S \rangle$ denotes the vector space generated by S over the underlying field. The sets of variables are represented by $\mathbf{x}, \mathbf{y}, \mathbf{z}$, $\mathbb{F}[\mathbf{x}]$ is the set of polynomials in $\mathbf{x}$ variables over $\mathbb{F}$, and polynomials in $\mathbb{F}[\mathbf{x}]$ are denoted by f, g, h. For $f \in \mathbb{F}[\mathbf{x}], \mathrm{var}(f)$ denotes the set of variables appearing in f, $\deg(f)$ denotes the total degree of f and for $x \in \mathbf{x}, \deg_x(f)$ denotes the highest degree of x in f. A set of variables $\mathbf{x}$ would often be treated as a column vector and for $A \in \mathrm{GL}(n, \mathbb{F})$, $A\mathbf{x}$ means that A is left multiplied to $\mathbf{x}$. For $n \in \mathbb{N}, \mathrm{poly}(n)$ means $n^{O(1)}$.

2.1 Structural preliminaries

This section is devoted to some useful definition and structural results. We have classified these into eight parts.

2.1.1 Algebraic and linear algebraic preliminaries

Definition 2.1 (Group) *Let G be a set and $\circ : G \times G \to G$ be a binary operation on G. Then, $(G, \circ)$ is called a group if it satisfies the following properties.*

1. *(Associativity): For every $g_1, g_2, g_3 \in G, (g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)$.*
2. *(Identity): There exists $e \in G$, such that for every $g \in G, e \circ g = g \circ e = g$.*
3. *(Inverse): For every $g \in G$, there exists a $g' \in G$ such that $g \circ g' = g' \circ g = e$.*

Further, for every $g_1, g_2 \in G$, if $g_1 \circ g_2 = g_2 \circ g_1$ then $(G, \circ)$ is called a commutative group.

For example, $(\mathbb{Z}, +)$ and $(\mathbb{R}, \times)$ are commutative groups but $(\mathrm{GL}(n, \mathbb{R}), *)$ is a non-commutative group, where $*$ denotes the matrix multiplication operation.

Definition 2.2 (Ring) *Let R be a set and $+, \cdot$ be binary operations on R. Then, $(R, +, \cdot)$ is called a ring if $(R, +)$ is a commutative group and the following properties are satisfied.*

1. *(Associativity of $\cdot$): For every $r_1, r_2, r_3 \in R, (r_1 \cdot r_2) \cdot r_3 = r_1 \cdot (r_2 \cdot r_3)$.*
2. *(Multiplicative identity): There exists an $e' \in R$ such that for every $r \in R, r \cdot e' = e' \cdot r = r$.*
3. *(Distributive property): For every $r_1, r_2, r_3 \in R, (r_1 + r_2) \cdot r_3 = (r_1 \cdot r_3) + (r_2 \cdot r_3)$ and $r_1 \cdot (r_2 + r_3) = (r_1 \cdot r_2) + (r_1 \cdot r_3)$.*

Further, $(R, +, \cdot)$ is called a commutative ring if for every $r_1, r_2 \in R, r_1 \cdot r_2 = r_2 \cdot r_1$.

For example, $(\mathbb{Z}, +, \times)$ is a commutative ring, where as $(M_n(\mathbb{R}), +, *)$ is a non-commutative ring, where $*$ the denotes matrix multiplication operation.

Definition 2.3 (Field) *A ring $(\mathbb{F}, +, \cdot)$ is called a field if $(\mathbb{F}, \cdot)$ is a commutative group. The identities of $(\mathbb{F}, +)$ and $(\mathbb{F}, \cdot)$ are called additive and multiplicative identities of $\mathbb{F}$ respectively.*

For examples, $(\mathbb{Q}, +, \times), (\mathbb{R}, +, \times), (\mathbb{C}, +, \times)$ and $(\mathbb{Z}_n, +_n, \times_n)$ are fields, where n is a prime number, $\mathbb{Z}_n = \{0, \ldots, n-1\}$ and $+_n, \times_n$ are addition and multiplication modulo n respectively.

Remark 2.1 *For simplicity of notations, we would often denote an other algebraic object like a group or a ring with the underlying set without mentioning the associated operations.*

Definition 2.4 (Characteristic of a field) *Let $\mathbb{F}$ be a field having 0 and 1 as the additive and multiplicative identities respectively. We say that $\mathbb{F}$ has characteristic n for some $n \in \mathbb{N}^{\times}$ if it is the smallest number such that $n \cdot 1 = 0$. If such an n does not exist, we say $\mathbb{F}$ has characteristic zero. The characteristic of $\mathbb{F}$ is denoted as $char(\mathbb{F})$.*

It is an easy exercise to prove the following fact.

Fact 2.1 (Value of the characteristic of a field) *Let $\mathbb{F}$ be a field. Then, $char(\mathbb{F})$ is either 0 or a prime number.*

Definition 2.5 (Vector space and subspace) *Let $(\mathbb{F}, +, \cdot)$ be a field and $(U, +)$ be a commutative group. Then, U is said to be a vector space over $\mathbb{F}$ (or an $\mathbb{F}$-vector space), if there exists $\circ : \mathbb{F} \times U \to U$ such that the following properties are satisfied.*

1. *For every $\alpha \in \mathbb{F}, u, v \in U, \alpha \circ (u + v) = (\alpha \circ u) + (\alpha \circ v)$.*
2. *For every $\alpha, \beta \in \mathbb{F}, u \in U, (\alpha \cdot \beta) \circ u = \alpha \circ (\beta \circ u)$.*
3. *For every $\alpha, \beta \in \mathbb{F}, u \in U, (\alpha + \beta) \circ u = (\alpha \circ u) + (\beta \circ u)$, where $+$ on the L.H.S. and on the R.H.S. are operations of $\mathbb{F}$ and U respectively.*
4. *For every $u \in U, 1 \circ u = u$, where 1 is the multiplicative identity of $\mathbb{F}$.*

Elements of a vector space are called vectors. A subset $W \subseteq U$ is said to be an $\mathbb{F}$-subspace of U if $(W, +)$ is an $\mathbb{F}$-vector space, where $+$ is the binary operation of U.

For example, $(M_n(\mathbb{F}), +)$ is a $\mathbb{F}$-vector space. Let $U = \{f \in \mathbb{R}[x_1, \ldots, x_n] : \deg(f) \leq d\}$, where $\deg(f)$ denotes the total degree of f. Then, $(U, +)$ is an $\mathbb{R}$-vector space.

In Definitions 2.6 - 2.8, $\mathbb{F}$ is a field, U is an $\mathbb{F}$-vector space, and $W \subseteq U$.

Definition 2.6 (Linearly dependent and independent sets) *W is said to be $\mathbb{F}$-linearly dependent if there exist $w_1, \ldots, w_n \in W$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that for some $i \in [n]$, $\alpha_i \neq 0$ and $\alpha_1 w_1 + \cdots + \alpha_n w_n = 0$. Otherwise, W is said to be $\mathbb{F}$-linearly independent.*

Definition 2.7 (Generating system of a vector space) *W is said to be a generating system (or a spanning set) of U if for every $u \in U$, there exist $w_1, \ldots, w_n \in W$, and $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $\alpha_1 w_1 + \cdots + \alpha_n w_n = u$.*

Definition 2.8 (Basis and the dimension of a vector space) *W is said to be an $\mathbb{F}$-basis of U (or simply a basis of U) if it is $\mathbb{F}$-linearly independent and also a generating system of U. The dimension of U, denoted $\dim U$, is defined as the cardinality of a basis of U*[1].

For example, for $i, j \in [n]$, let $E_{i,j} \in M_n(\mathbb{F})$ such that the (i,j)-th entry of $E_{i,j}$ is 1 and every other entry is 0. Then, $\{E_{i,j} : i, j \in [n]\}$ is an $\mathbb{F}$-basis of $(M_n(\mathbb{F}), +)$ and thus $\dim(M_n(\mathbb{F})) = n^2$. Let $U = \{f \in \mathbb{F}[x] : \deg(f) \leq d\}$. Then, $\{1, x, \ldots, x^d\}$ is an $\mathbb{F}$-basis of U and hence $\dim(U) = d + 1$.

It is easy to prove the following.

Observation 2.1 (Linearly independent set extends to a basis) *Let $\mathbb{F}$ be a field and U be an $\mathbb{F}$-vector space such that $\dim U = n$. Let $W \subseteq U$ be an $\mathbb{F}$-linearly independent set. Then, there exists a basis of U, which contains W. Thus, if $|W| = n$ then W is a basis of U.*

Definition 2.9 (Sum and direct sum of subspaces) *Let $\mathbb{F}$ be a field, U be an $\mathbb{F}$-vector space, and W_1, W_2 be subspaces of U. Then, the sum of W_1 and W_2 is defined as $W_1 + W_2 := \{w_1 + w_2 : w_1 \in W_1, w_2 \in W_2\}$. Further, $W_1 + W_2$ is said to be a direct sum, denoted $W_1 \oplus W_2$, if $W_1 \cap W_2 = \{0\}$, where $0 \in U$ is the zero vector.*

It is a simple exercise to prove the following.

Observation 2.2 (Dimension of the direct sum of subspaces) *Let $\mathbb{F}$ be a field, U be an $\mathbb{F}$-vector space, and W_1, W_2 be subspaces of U such that $W_1 + W_2$ is a direct sum. Then, $W_1 \oplus W_2$ is a subspace of U and $\dim W_1 \oplus W_2 = \dim W_1 + \dim W_2$.*

Definition 2.10 (Linear map) *Let $\mathbb{F}$ be a field and U, W be $\mathbb{F}$-vector spaces. A map $\varphi : U \to W$ is said to be $\mathbb{F}$-linear if for every $u_1, u_2 \in U, \alpha, \beta \in \mathbb{F}, \varphi(\alpha u_1 + \beta u_2) = \alpha\varphi(u_1) + \beta\varphi(u_2)$. If φ is bijective then it is called an isomorphism of $\mathbb{F}$-vector spaces.*

For example, let $U = \{f \in \mathbb{F}[x] : \deg(f) \leq d\}$. Then, $\varphi : U \to \mathbb{F}^{d+1}, \alpha_0 + \alpha_1 x + \cdots + \alpha_d x^d \mapsto (\alpha_0, \ldots, \alpha_d)$ is an isomorphism of $\mathbb{F}$-vector spaces.

Definition 2.11 (Matrix associated with a linear map) *Let $n, m, r, s \in \mathbb{N}, \mathbb{F}$ be a field and U_1 and U_2 be $\mathbb{F}$-subspaces of $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively such that $\dim U_1 = r$ and $\dim U_2 = s$.*

[1] It is a well-known fact in linear algebra that every vector space over any field $\mathbb{F}$ has a basis and any two bases of an $\mathbb{F}$-vector space have same cardinalities.

Let $\varphi : U_1 \to U_2$ be an $\mathbb{F}$-linear map and $\mathbf{u}_1 := (\mathbf{u}_{1,1}, \ldots, \mathbf{u}_{1,r})$ and $\mathbf{u}_2 := (\mathbf{u}_{2,1}, \ldots, \mathbf{u}_{2,s})$ be $\mathbb{F}$-bases of U_1 and U_2 respectively. For every $i \in [r]$, let $a_{i,1}, \ldots, a_{i,s} \in \mathbb{F}$ such that

$$\varphi(\mathbf{u}_{1,i}) = \sum_{j=1}^{s} a_{i,j}\mathbf{u}_{2,j}.$$

Then, $A = (a_{i,j})_{i\in[r],j\in[s]}$ is said to be the matrix associated with φ with respect to the ordered bases $\mathbf{u}_1$ and $\mathbf{u}_2$ of U_1 and U_2 respectively.

Definition 2.12 (Extension field and its degree) *Let $\mathbb{F}, \mathbb{L}$ be fields such that $\mathbb{F} \subseteq \mathbb{L}$ and the operations of $\mathbb{F}$ and same as operation of $\mathbb{L}$ restricted to $\mathbb{F}$. Then, $\mathbb{L}$ is called an extension field of $\mathbb{F}$. Note that $\mathbb{L}$ is also an $\mathbb{F}$-vector space and the degree of the field extension, denoted $[\mathbb{L} : \mathbb{F}]$, is the dimension of $\mathbb{L}$ over $\mathbb{F}$.*

For example, $\mathbb{C}$ is a field extension of $\mathbb{R}$ of degree 2 as $\{1, i\}$ is an $\mathbb{R}$-basis of $\mathbb{C}$.

Definition 2.13 (Algebraically closed field) *A field $\mathbb{F}$ is said to be algebraically closed if for every non-constant polynomial $f \in \mathbb{F}[x]$, there exists an $\alpha \in \mathbb{F}$ such that $f(\alpha) = 0$.*

It is a well-known fact in mathematics that $\mathbb{C}$ is an algebraically closed field.

Definition 2.14 (Algebraic closure of a field) *Let $\mathbb{F}$ be a field. Then, the algebraic closure of $\mathbb{F}$, denoted $\overline{\mathbb{F}}$, is an extension field of $\mathbb{F}$ such that for every non-constant $f \in \mathbb{F}[x]$, there exists an $\alpha \in \overline{\mathbb{F}}$ such that $f(\alpha) = 0$.*

For example, $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$. The following is a well-known fact in algebra, whose proof can be found in any standard textbook on field theory or abstract algebra.

Fact 2.2 *Every field $\mathbb{F}$ has an algebraic closure $\overline{\mathbb{F}}$.*

Definition 2.15 (Invariant space) *Let U be an $\mathbb{F}$-vector space, $\mathscr{T} \subseteq \mathrm{End}_{\mathbb{F}}(U)$, and $W \subseteq U$ be an $\mathbb{F}$-subspace. Then, U is said to be $\mathscr{T}$-invariant if for every $\varphi \in \mathscr{T}, w \in W, \varphi(w) \in W$. If $\mathscr{T} \subseteq M_n(\mathbb{F})$ then the 'invariant space of $\mathscr{T}$' means a $\mathscr{T}$-invariant subspace of $\mathbb{F}^n$.*

Definition 2.16 (Irreducible invariant space) *Let U be an $\mathbb{F}$-vector space, $\mathscr{T} \subseteq \mathrm{End}_{\mathbb{F}}(U)$, and $W \subseteq U$ be a $\mathscr{T}$-invariant space. Then, W is called an irreducible $\mathscr{T}$-invariant space if there do not exist $\mathscr{T}$-invariant subspaces $W_1, W_2 \subseteq W$ such that $W = W_1 \oplus W_2$.*

Definition 2.17 (Closure of a vector) *Let U be an $\mathbb{F}$-vector space, $\mathscr{T} \subseteq \mathrm{End}_{\mathbb{F}}(U)$, and $\mathbf{u} \in U$. Then, the closure of $\mathbf{u}$ with respect to $\mathscr{T}$, denoted $\mathrm{closure}_{\mathscr{T}}(\mathbf{u})$, is the smallest $\mathscr{T}$-invariant subspace of U containing $\mathbf{u}$.*

Definition 2.18 (Algebra over a field) *Let $\mathbb{F}$ be a field and $(\mathscr{A}, +)$ be an $\mathbb{F}$-vector space. Then, $(\mathscr{A}, +)$ is called an algebra over $\mathbb{F}$ (or an $\mathbb{F}$-algebra) if there exists $\circ : \mathscr{A} \times \mathscr{A} \to \mathscr{A}$ such that the following properties are satisfied:*

1. *For every $u_1, u_2, u_3 \in \mathscr{A}$, $(u_1+u_2)\circ u_3 = u_1 \circ u_3 + u_2 \circ u_3$ and $u_1 \circ (u_2+u_3) = u_1 \circ u_2 + u_1 \circ u_3$.*
2. *For every $\alpha, \beta \in \mathbb{F}, u_1, u_2 \in \mathscr{A}$, $(\alpha u_1) \circ (\beta u_2) = (\alpha\beta)(u_1 \circ u_2)$.*

Further, if $\mathscr{A}$ also contains the identity with respect to $\circ$ (called the multiplicative identity) then $\mathscr{A}$ is called as a unital algebra.

For example, $(M_n(\mathbb{F}), +, *)$ is an $\mathbb{F}$-algebra and popularly known as *matrix algebra*. Another example of an $\mathbb{F}$-algebra is $(\mathbb{F}[\mathbf{x}], +, *)$.

Remark 2.2 *Let $\mathbb{F}$ be a field, U be an $\mathbb{F}$-vector space, and $\mathbf{u}_1, \ldots, \mathbf{u}_r \in U$. Then, the $\mathbb{F}$-algebra generated by $\mathbf{u}_1, \ldots, \mathbf{u}_r$, is the set containing finite $\mathbb{F}$-linear combinations of powers of $\mathbf{u}_1, \ldots, \mathbf{u}_r$.*

Remark 2.3 *Whenever we say a basis of an $\mathbb{F}$-algebra $\mathscr{A}$, we mean an $\mathbb{F}$-basis of the $\mathbb{F}$-vector space $\mathscr{A}$.*

Definition 2.19 (Algebra homomorphism and isomorphism) *Let $\mathscr{A}, \mathscr{B}$ be $\mathbb{F}$-algebras and $\varphi : \mathscr{A} \to \mathscr{B}$. Then, φ is said to be an $\mathbb{F}$-algebra homomorphism if for every $u_1, u_2 \in \mathscr{A}, \alpha, \beta \in \mathbb{F}, \varphi(\alpha u_1 + \beta u_2) = \alpha\varphi(u_1) + \beta\varphi(u_2)$ and $\varphi(u_1 \circ u_2) = \varphi(u_1) \circ \varphi(u_2)$. If $\mathscr{A}, \mathscr{B}$ are unital with multiplicative identities $1_{\mathscr{A}}$ and $1_{\mathscr{B}}$ then $\varphi(1_{\mathscr{A}}) = 1_{\mathscr{B}}$. Further, φ is said to be an $\mathbb{F}$-algebra isomorphism if it is an $\mathbb{F}$-algebra homomorphism and bijective.*

For example, $(M_n(\mathbb{F}), +, *)$ and $(I_n \otimes M_n(\mathbb{F}), +, *)$ are $\mathbb{F}$-algebras, where $I_n \otimes M_n(\mathbb{F}) = \{I_n \otimes A : A \in M_n(\mathbb{F})\}$ (see Definition 2.20 below). Then, $\varphi : M_n \to I_n \otimes M_n, A \mapsto I_n \otimes A$ is an $\mathbb{F}$-algebra isomorphism.

Definition 2.20 (Tensor product of matrices) *Let $n \in \mathbb{N}^{\times}$, $\mathbb{F}$ be a field, $A = (a_{i,j})_{i,j\in[n]}, B = (b_{i,j})_{i,j\in[n]} \in M_n(\mathbb{F})$. Then, the tensor product of A and B, denoted $A \otimes B$, is the following $n^2 \times n^2$ matrix over $\mathbb{F}$*

$$A \otimes B = \begin{pmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{n,1}B & \cdots & a_{n,n}B \end{pmatrix},$$

where for $i, j \in [n], a_{i,j}B = (a_{i,j} \cdot b_{k,l})_{k,l \in [n]}$.

Now, we present an important result, which will be used in the equivalence test for the determinant given in Chapter 4. This result is a special case of the *Skolem-Noether theorem* (see page 173 of [Lor08] for the general statement of the Skolem-Noether theorem).

Theorem 2.1 (Skolem-Noether) *Let* $n, s \in \mathbb{N}^{\times}$ *such that* $n \mid s$, *and* $\mathscr{A} \subseteq M_s$ *be an* $\mathbb{F}$*-algebra (containing* I_s*) that is isomorphic to* M_n *via a map* $\phi : M_n \to \mathscr{A}$. *Then, there exists a* $K \in \mathrm{GL}(s, \mathbb{F})$ *such that for every* $C \in M_n$,

$$\phi(C) = K^{-1} \cdot (I_{s/n} \otimes C) \cdot K.$$

Definition 2.21 (Characteristic polynomial and eigenvalues) *Let* $n \in \mathbb{N}, \mathbb{F}$ *be a field and* $A \in M_n(\mathbb{F})$. *Then, the characteristic polynomial of* A *is defined as the determinant of* $(xI_n - A)$, *where* x *is a formal variable. The roots of the characteristic polynomial of* A *are called as the eigenvalues of* A.

It is easy to prove the following.

Fact 2.3 (Similar matrices have same characteristic polynomials) *Let* $n \in \mathbb{N}, \mathbb{F}$ *be a field,* $A \in M_n(\mathbb{F})$ *and* $B \in \mathrm{GL}(n, \mathbb{F})$. *Then, the characteristic polynomials of* A *and* $B \cdot A \cdot B^{-1}$ *are the same.*

The proofs of the following fact can be found in any standard textbook of linear algebra.

Fact 2.4 (Cayley-Hamilton theorem) *Let* $n \in \mathbb{N}, \mathbb{F}$ *be a field,* $A \in M_n(\mathbb{F})$ *and* $p(x)$ *be the characteristic polynomial of* A. *Then,* $p(A) = 0$.

Fact 2.5 (Jordan normal form) *Let* $n \in \mathbb{N}$, $\mathbb{F}$ *be a field and* $\overline{\mathbb{F}}$ *be the algebraic closure of* $\mathbb{F}$. *Let* $A \in M_n(\mathbb{F})$, $\alpha_1, \ldots, \alpha_r \in \overline{\mathbb{F}}$ *be distinct eigenvalues of* A *and* α_i *appears* n_i *times for every* $i \in [r]$. *Then,* $n_1 + \cdots + n_r = n$. *For* $i \in [r]$, *let* $J_i \in M_{n_i}(\overline{\mathbb{F}})$ *be an upper triangular matrix*[1], *where every diagonal entry is* α_i, *for* $l \in [n_i - 1]$, *the* $(l, l+1)$*-th entry is 1 and every other entry is 0. Then, there exists a* $B \in \mathrm{GL}(n, \overline{\mathbb{F}})$ *such that*

$$A = B^{-1} \cdot J \cdot B,$$

where $J = \mathrm{diag}(J_1, \ldots, J_r)$ *is a block-diagonal matrix having* $J_1, \ldots, J_r$ *as the diagonal block.*

[1] The matrix J_i is called as the Jordan block corresponding to α_i.

Definition 2.22 (Sylvester matrix) *Let $\mathbb{F}$ be a field, $f = \alpha_m x^m + \cdots + \alpha_1 x + \alpha_0$ and $g = \beta_d x^d + \cdots + \beta_1 x + \beta_0$, where for every $i \in [m], j \in [d], \alpha_i, \beta_j \in \mathbb{F}, \alpha_m \neq 0, \beta_d \neq 0$. Then, the Sylvester matrix of f and g, denoted $S_{f,g}$, is a $(m+d) \times (m+d)$ matrix, that looks as follows*

$$S_{f,g} = \begin{pmatrix} \alpha_m & \alpha_{m-1} \cdots & \alpha_1 & \alpha_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \alpha_m \cdots & \alpha_2 & \alpha_1 & \alpha_0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 \cdots & 0 & 0 & \alpha_m & \alpha_{m-1} & \cdots & \alpha_1 & \alpha_0 \\ \beta_d & \beta_{d-1} \cdots & \beta_0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 \cdots & 0 & 0 & 0 & \beta_d & \cdots & \beta_1 & \beta_0 \end{pmatrix},$$

The following important well-known property associated with $S_{f,g}$ gives us a way to test if the $\gcd(f, d)$ is non-constant or not.

Fact 2.6 (Determinant of the Sylvester matrix) *Let $\mathbb{F}$ be a field and $f, g \in \mathbb{F}[x]$. Then, $\gcd(f, g)$ is non-constant if and only if $\det(S_{f,g}) = 0$.*

2.1.2 Symmetries of a polynomial

Definition 2.23 (Symmetries of a polynomial) *Let $n \in \mathbb{N}, \mathbb{F}$ be a field, $\mathbf{x} = \{x_1, \ldots, x_n\}$ and $f \in \mathbb{F}[\mathbf{x}]$. Then, $A \in \mathrm{GL}(n, \mathbb{F})$ is called a symmetry of f if $f = f(A\mathbf{x})$. The set of symmetries of f is called as the group of symmetries of f, denoted $\mathscr{G}_f$, as it forms a group with respect to matrix multiplication.*

The following two definitions would be used in Chapter 3.

Definition 2.24 (Characterisation by symmetries) *Let $\mathbb{F}$ be a field, $\mathbf{x}$ be a set of variables and $g \in \mathbb{F}[\mathbf{x}]$ be a homogeneous polynomial of degree d, i.e., every monomial of g has degree d. Then, g is said to be characterised by its symmetries if for every degree d homogeneous polynomial $f \in \mathbb{F}[\mathbf{x}]$, $\mathscr{G}_g \subseteq \mathscr{G}_f$ implies that $f(\mathbf{x}) = \alpha \cdot g(\mathbf{x})$ for some $\alpha \in \mathbb{F}$.*

Definition 2.25 (Characterization by circuit identities) *Let $g \in \mathbb{F}[\mathbf{x}]$ be an n-variate polynomial, and $\mathbf{z}, \mathbf{u}$ be two sets of constantly many variables and $|\mathbf{z}| = c$. Suppose that there exist $m = \mathrm{poly}(n)$ polynomials $q_1(\mathbf{z}, \mathbf{u}), \ldots, q_m(\mathbf{z}, \mathbf{u})$ over $\mathbb{F}$ such that for every $i \in [m]$, q_i is computable by a constant size arithmetic circuit and there are matrices $A_{i1}, \ldots, A_{ic} \in \mathbb{F}[\mathbf{u}]^{n \times n}$ computable by $\mathrm{poly}(n)$ size circuits, and the following condition is satisfied: For $f \in \mathbb{F}[\mathbf{x}]$, $q_i(f(A_{i1}\mathbf{x}), \ldots, f(A_{ic}\mathbf{x}), \mathbf{u}) = 0$ for every $i \in [m]$ if and only if $f = \alpha \cdot g$ for some $\alpha \in \mathbb{F}$. Then, g is characterized by circuit identities over $\mathbb{F}$.*

Definition 2.25 has been taken after slightly modifying Definition 3.4.7 in [Gro12], to suit our purpose.

2.1.3 Partial derivatives of a polynomial

Definition 2.26 (Partial derivative) *Let $n, d \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ be a field and $f \in \mathbb{F}[\mathbf{x}]$ be a degree d polynomial, where $f = \sum_{\mathbf{e} \in \mathbb{N}^n} \alpha_{\mathbf{e}} x_1^{e_1} \cdots x_n^{e_n}$ and $\mathbf{e} = (e_1, \ldots, e_n)$. Then, the first-order partial derivative of f with respect to x_i for some $i \in [n]$ is defined as*

$$\frac{\partial f}{\partial x_i} = \sum_{\mathbf{e} \in \mathbb{N}^n} \alpha_{\mathbf{e}} e_i \prod_{j \in [n] \setminus \{i\}} x_j^{e_j} x_i^{e_i - 1}.$$

Note that if f is not a constant and $\mathrm{char}(\mathbb{F}) > d$ then $\frac{\partial f}{\partial x_i} \neq 0$ for every $i \in [n]$. Now we record some useful properties of partial derivatives. We direct interested reader to [CKW11] for many interesting applications of partial derivatives in ACT.

Important properties of partial derivatives. Let $f, g \in \mathbb{F}[\mathbf{x}]$, $\alpha, \beta \in \mathbb{F}$, and $x \in \mathbf{x}$ be an arbitrary variable.

1. (Linearity). $\frac{\partial}{\partial x}(\alpha f + \beta g) = \alpha \frac{\partial f}{\partial x} + \beta \frac{\partial g}{\partial x}$.
2. (Derivative of the product). $\frac{\partial}{\partial x} fg = f \frac{\partial g}{\partial x} + g \frac{\partial f}{\partial x}$.
3. (Chain rule). Let $h \in \mathbb{F}[y_1, \ldots, y_r]$ and $\mathbf{g} = (g_1, \ldots, g_r)$, where $g_i \in \mathbb{F}[\mathbf{x}]$ for every $i \in [r]$. Let $h \circ \mathbf{g} := h(g_1(\mathbf{x}), \ldots, g_r(\mathbf{x}))$. Then,

$$\frac{\partial}{\partial x}(h \circ \mathbf{g}) = \sum_{i \in [r]} \frac{\partial h}{\partial g_i} \frac{\partial g_i}{\partial x},$$

   where $\frac{\partial h}{\partial g_i}$ means $\frac{\partial h}{\partial y_i}(\mathbf{g})$. Clearly, $\frac{\partial h}{\partial g_i} \in \mathbb{F}[\mathbf{x}]$ for every $i \in [r]$.

The chain rule of partial derivatives implies the following observation.

Observation 2.3 *Let $n \in \mathbb{N}, \mathbb{F}$ be a field, $\mathbf{x} = \{x_1, \ldots, x_n\}$, $f \in \mathbb{F}[\mathbf{x}], A \in \mathrm{GL}(n, \mathbb{F})$ and $\nabla f = \left(\frac{\partial f}{\partial x}\right)^T_{x \in \mathbf{x}}$. Then,*

$$\nabla f(A\mathbf{x}) = A^T [\nabla f](A\mathbf{x}).$$

2.1.4 Hessian of a polynomial

Definition 2.27 (Hessian and the Hessian determinant) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ be a field and $f \in \mathbb{F}[\mathbf{x}]$. Then, the Hessian of f, denoted H_f, is the following matrix*

$$H_f := \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j \in [n]},$$

where $\frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)$ for every $i, j \in [n]$. The determinant of H_f is called as the Hessian determinant of f. Clearly, $\det(H_f) \in \mathbb{F}[\mathbf{x}]$.

The following properties are very useful in the equivalence test for regular ROFs given in Chapter 5.

Fact 2.7 (Lemma 2.6 of [CKW11]) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ be a field, $f \in \mathbb{F}[\mathbf{x}]$, $A \in M_n(\mathbb{F})$, and $g = f(A\mathbf{x})$. Then,*

$$H_g = A^T \cdot H_f(A\mathbf{x}) \cdot A.$$

It is easy to prove the following.

Fact 2.8 *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ be a field, $\mathbf{b} \in \mathbb{F}^n$, and $f \in \mathbb{F}[\mathbf{x}]$. Then,*

$$H_{f(\mathbf{x}+\mathbf{b})} = H_f.$$

These two facts immediately imply the following.

Corollary 2.1 *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ be a field, $g \in \mathbb{F}[\mathbf{x}], A \in \mathrm{GL}(n, \mathbb{F}), \mathbf{b} \in \mathbb{F}^n$ and $f = g(A\mathbf{x} + \mathbf{b})$. Then,*

$$\det(H_f) = (\det(A))^2 \det(H_g)(A\mathbf{x} + \mathbf{b}).$$

2.1.5 Lie algebra of a polynomial

We start this section with the following definition, which would be crucially used in Chapter 4.

Definition 2.28 (Lie bracket) *Let $n \in \mathbb{N}, \mathbb{F}$ be a field and $A, B \in M_n(\mathbb{F})$. Then, the Lie bracket of A, B, denoted $[A, B]$, is defined as $[A, B] = AB - BA$.*

Consider the following definition of the Lie algebra of a polynomial. For a more abstract definition of the Lie algebra, that is based on the Lie brackets, we direct the reader to Chapter 2 of [Hal03]. The equivalence of this following definition of the Lie algebra of a polynomial and the definition given in [Hal03] follows from Theorem 2.27 of [Hal03].

Definition 2.29 (Lie algebra) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ be a field, $f \in \mathbb{F}[\mathbf{x}]$ and ϵ be a formal variable satisfying $\epsilon^2 = 0$. Then, the Lie algebra of f, denoted $\mathfrak{g}_f$, is a subset of $M_n(\mathbb{F})$ such that every $A \in \mathfrak{g}_f$ satisfies the following*

$$f((I_n + \epsilon A)\mathbf{x}) - f(\mathbf{x}) = 0.$$

It follows from the definition that $\mathfrak{g}_f$ is an $\mathbb{F}$-vector space. The following fact is important and is easy to prove.

Fact 2.9 (Lie algebra closed under Lie bracket) *Let $f \in \mathbb{F}[\mathbf{x}]$ and $A, B \in \mathfrak{g}_f$. Then, $[A, B] \in \mathfrak{g}_f$.*

The following definition, which we call the working definition of the Lie algebra of a polynomial, uses first order partial derivatives of a polynomial. This definition is taken from [Kay12]. Claim 58 in [Kay12] shown the equivalence of Definitions 2.29 and 2.30.

Definition 2.30 (The working definition of $\mathfrak{g}_f$) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ be a field, and $f \in \mathbb{F}[\mathbf{x}]$. Then, $\mathfrak{g}_f$ is the set of matrices in $M_n(\mathbb{F})$, such that every $A = (a_{i,j})_{i,j\in[n]} \in \mathfrak{g}_f$ satisfies the following equation.*

$$\sum_{i,j\in[n]} a_{i,j} x_j \cdot \frac{\partial f}{\partial x_i} = 0. \tag{2.1}$$

The following fact relates Lie algebras of two equivalent polynomials. This fact is Proposition 58 in [Kay12]. This fact will be extensively useful for a special case of the equivalence test of NW given in Chapter 3 and the equivalence test for the determinant given in Chapter 4.

Fact 2.10 (Conjugacy relation of Lie algebras) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ be a field, $f \in \mathbb{F}[\mathbf{x}], A \in \mathrm{GL}(n, \mathbb{F})$ and $h = f(A\mathbf{x})$. Then,*

$$\mathfrak{g}_h = A^{-1} \cdot \mathfrak{g}_f \cdot A.$$

2.1.6 Essential and redundant variables of a polynomial

Definition 2.31 (Number of essential variables) *Let $n, s \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ be a field and $f \in \mathbb{F}[\mathbf{x}]$. We say that f has s essential variables if there exists an $A \in \mathrm{GL}(n, \mathbb{F})$ such that $|\mathrm{var}(f(A\mathbf{x}))| = s$ and there does not exist an $A' \in \mathrm{GL}(n, \mathbb{F})$ such that $|\mathrm{var}(f(A'\mathbf{x}))| < s$. In this case, the number of redundant variables in f is $(n - s)$.*

We borrow the notation $N_{ess}(f)$ from [Car06] to denote the number of essential variables in f. The following fact relates $N_{ess}(f)$ with the dimension of $\left\langle \frac{\partial f}{\partial x} : x \in \mathbf{x} \right\rangle$. The proof of this fact follows from the proof of Claim 2.3 in [KNST19].

Fact 2.11 (Relation between essential variables and partials) *Let $n, d \in \mathbb{N}, \mathbb{F}$ be a field such that char$(\mathbb{F}) = 0$ or $> d$, $\mathbf{x} = \{x_1, \ldots, x_n\}$, and $f \in \mathbb{F}[\mathbf{x}]$ be a polynomial having individual degree at most d. Let $U := \left\langle \frac{\partial f}{\partial x} : x \in \mathbf{x} \right\rangle$. Then, $N_{ess}(f) = \dim U$. Further, there exists $\mathbf{z} \subseteq \mathbf{x}$ such that $\left\{ \frac{\partial f}{\partial z} : z \in \mathbf{z} \right\}$ is a basis of U if and only if there exists an $A \in \mathrm{GL}(n, \mathbb{F})$ which maps every variable in $\mathbf{x} \setminus \mathbf{z}$ to itself, $\mathrm{var}(f(A\mathbf{x})) = \mathbf{z}$, and $N_{ess}(f) = |\mathbf{z}|$.*

By exploiting the relationship given in the above fact, [Car06] gave a polynomial time algorithm to remove redundant variables from f, when the input is the coefficient vector of f. [Kay11] gave a randomized algorithm for the same when f is given as a black-box. We talk more about algorithms to remove redundant variables from a polynomial in Section 2.2.2.

Definition 2.32 (Essential and redundant variables) *Let $n, d \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}$, and $\mathbb{F}$ be a field. Then, $\mathbf{z} \subseteq \mathbf{x}$ is called a set of essential variables of f if $N_{ess}(f) = |\mathbf{z}|$ and there exists an $A \in \mathrm{GL}(n, \mathbb{F})$ such that $f(A\mathbf{x}) \in \mathbb{F}[\mathbf{z}]$. Once we fix such a $\mathbf{z}$, every other variable in $\mathbf{x}$ is said to be redundant for f.*

Observe that a set of essential variables of a polynomial need not be unique. The following three observations would be required for the ET for regular ROFs given in Chapter 5.

Observation 2.4 *Let $n, d \in \mathbb{N}$, $\mathbf{x}$ and $\mathbf{y}$ be disjoint sets of variables such that $|\mathbf{x}|+|\mathbf{y}| = n$, and $\mathbb{F}$ be a field satisfying either char$(\mathbb{F}) = 0$ or char$(\mathbb{F}) > d$. Let $h \in \mathbb{F}[\mathbf{x}]$ be such that $\deg(h) \leq d$ and $N_{ess}(h) = |\mathbf{x}|$. Let $\mathbf{z} \subseteq \mathbf{x} \uplus \mathbf{y}$ and $A \in \mathrm{GL}(n, \mathbb{F})$ such that $|\mathbf{z}| = |\mathbf{x}|$ and $h(A(\mathbf{x}, \mathbf{y})) \in \mathbb{F}[\mathbf{z}]$. Then, A maps every $\mathbf{x}$-variable to a linear form in $\mathbf{z}$-variables.*

Proof: First, we assume that $\mathbf{x} = \mathbf{z}$. Suppose the rows and columns of A are labelled by the ordered tuple $(\mathbf{x}, \mathbf{y})$ and A looks like

$$A = \begin{bmatrix} A_{\mathbf{x}} & A_1 \\ A_2 & A_{\mathbf{y}} \end{bmatrix},$$

where the rows and columns of $A_{\mathbf{x}}, A_{\mathbf{y}}$ are labelled by $\mathbf{x}$ and $\mathbf{y}$ respectively. Note that it is sufficient to show that A_1 is the zero matrix. Let $g(\mathbf{x}, \mathbf{y}) = h(A(\mathbf{x}, \mathbf{y}))$. Then, it follows from Observation 2.3 that

$$\nabla g(\mathbf{x}, \mathbf{y}) = A^T [\nabla h](A(\mathbf{x}, \mathbf{y})), \tag{2.2}$$

Let $\nabla g = ([\nabla g]_{\mathbf{x}}, [\nabla g]_{\mathbf{y}})^T$, where $[\nabla g]_{\mathbf{x}} = \left(\frac{\partial g}{\partial x}\right)_{x\in\mathbf{x}}$ and $[\nabla g]_{\mathbf{y}} = \left(\frac{\partial g}{\partial y}\right)^T_{y\in\mathbf{y}}$. Similarly, let $\nabla h = ([\nabla h]_{\mathbf{x}}, [\nabla h]_{\mathbf{y}})^T$. As $g, h \in \mathbb{F}[\mathbf{x}]$, for every $y \in \mathbf{y}$, $\frac{\partial g}{\partial y} = \frac{\partial h}{\partial y} = 0$. This implies $[\nabla g]_{\mathbf{y}} = [\nabla h]_{\mathbf{y}} = \mathbf{0}$. Since $\mathbf{x}$ is the set of essential variables of h, all the entries in $[\nabla h]_{\mathbf{x}}$ are $\mathbb{F}$-linearly independent and thus all the entries in $[\nabla h]_{\mathbf{x}}(A(\mathbf{x}, \mathbf{y}))$ are also $\mathbb{F}$-linearly independent. Now, it is easy to see from Equation (2.2) and the structure of A that every entry of A_1^T should be equal to 0. Otherwise, we get a non-zero linear combination of $\frac{\partial h}{\partial x}(A(\mathbf{x}, \mathbf{y})), x \in \mathbf{x}$, which is equal to 0 and this is a contradiction.

Now, suppose $\mathbf{x} \neq \mathbf{z}$. Let $P \in \mathrm{GL}(n, \mathbb{F})$ be a permutation matrix, that maps $\mathbf{x}$ to $\mathbf{z}$, $\mathbf{z}$ to $\mathbf{x}$ and every other variable to itself. We know $h(A(\mathbf{x}, \mathbf{y})) \in \mathbb{F}[\mathbf{z}]$. Then, note that $h(AP(\mathbf{x}, \mathbf{y})) \in \mathbb{F}[\mathbf{x}]$. Now, it follows from the above argument that AP maps every variable in $\mathbf{x}$ to a linear form in $\mathbf{x}$-variables. It is not difficult to see that this implies that A maps every variable in $\mathbf{x}$ to a linear form in $\mathbf{z}$-variables. □

Observation 2.5 *Let $d \in \mathbb{N}$, $\mathbf{x}$ and $\mathbf{y}$ be disjoint sets of variables and $\mathbb{F}$ be a field satisfying $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$. Let $h(\mathbf{x}, \mathbf{y}) = g(\mathbf{x})^e \cdot p(\mathbf{x}, \mathbf{y})$, where $\deg(h) \leq d$, $g(\mathbf{x}), p(\mathbf{x}, \mathbf{y})$ are co-prime and $e \geq 1$. Suppose every variable in $\mathbf{x}$ is essential for g and*

$$\sum_{x\in\mathbf{x}} \alpha_x \frac{\partial h}{\partial x} + \sum_{y\in\mathbf{y}} \beta_y \frac{\partial h}{\partial y} = 0,$$

where $\alpha_x, \beta_y \in \mathbb{F}$ for every $x \in \mathbf{x}, y \in \mathbf{y}$. Then, $\alpha_x = 0$ for every $x \in \mathbf{x}$.

Proof: Since $h = g(\mathbf{x})^e \cdot p(\mathbf{x}, \mathbf{y})$ and $e \geq 1$, we get

$$\sum_{x\in\mathbf{x}} \alpha_x \left(g^e \frac{\partial p}{\partial x} + e \cdot g^{e-1} \cdot p \frac{\partial g}{\partial x} \right) + \sum_{y\in\mathbf{y}} \beta_y g^e \frac{\partial p}{\partial y} = 0.$$

On dividing the above equation by g^{e-1} and rearranging the terms we get

$$g \left(\sum_{x\in\mathbf{x}} \alpha_x \frac{\partial p}{\partial x} + \sum_{y\in\mathbf{y}} \beta_y \frac{\partial p}{\partial y} \right) + e \cdot p \left(\sum_{x\in\mathbf{x}} \alpha_x \frac{\partial g}{\partial x} \right) = 0.$$

As $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d, \frac{\partial g}{\partial x} \neq 0, \frac{\partial p}{\partial x} \neq 0, \frac{\partial p}{\partial y} \neq 0$ for every $x \in \mathbf{x}, y \in \mathbf{y}$. Let $g' = \left(\sum_{x\in\mathbf{x}} \alpha_x \frac{\partial p}{\partial x} + \sum_{y\in\mathbf{y}} \beta_y \frac{\partial p}{\partial y} \right)$ and $p' = \left(\sum_{x\in\mathbf{x}} \alpha_x \frac{\partial g}{\partial x} \right)$. If $g' = 0$, then we get $p' = 0$. Otherwise, g divides $e \cdot p \cdot p'$. Since g and p are co-prime polynomials, g divides p', which is not possible

unless $p' = 0$ as $\deg(p') < \deg(g)$. Thus, in both the cases we get

$$\sum_{x \in \mathbf{x}} \alpha_x \frac{\partial g}{\partial x} = 0.$$

As every variable in $\mathbf{x}$ is essential for g, Fact 2.11 we get $\alpha_x = 0$ for every $x \in \mathbf{x}$. □

Observation 2.6 *Let $d \in \mathbb{N}$, $\{x_1, x_2\}$ and $\mathbf{y}$ be disjoint sets of variables and $\mathbb{F}$ be a field satisfying $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$. Let $h(x_1, x_2, \mathbf{y}) = \sum_{i \geq 0} p_i(\mathbf{y}) \cdot (x_1 x_2)^i$ be a polynomial of degree at most d such that $p_i(\mathbf{y}) \neq 0$ for some $i \geq 1$. Suppose*

$$\alpha_1 \frac{\partial h}{\partial x_1} + \alpha_2 \frac{\partial h}{\partial x_2} + \sum_{y \in \mathbf{y}} \beta_y \frac{\partial h}{\partial y} = 0,$$

where $\alpha_1, \alpha_2, \beta_y \in \mathbb{F}$ for every $y \in \mathbf{y}$. Then, $\alpha_1 = \alpha_2 = 0$.

Proof: As $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$, $\frac{\partial h}{\partial x_1} \neq 0$, $\frac{\partial h}{\partial x_2} \neq 0$, and $\frac{\partial h}{\partial y} \neq 0$ for every $y \in \mathbf{y}$. Since $h(\mathbf{x}, \mathbf{y}) = \sum_{i \geq 0} p_i(\mathbf{y})(x_1 x_2)^i$, the above equation can be written as

$$\alpha_1 \left(\sum_{i \geq 1} i \cdot p_i \cdot x_1^{i-1} x_2^i \right) + \alpha_2 \left(\sum_{i \geq 1} i \cdot p_i \cdot x_1^i x_2^{i-1} \right) + \sum_{y \in \mathbf{y}} \beta_y \left(\sum_{i \geq 0} (x_1 x_2)^i \frac{\partial p_i}{\partial y} \right) = 0.$$

Note that in this equation, the polynomials $\alpha_1 \left(\sum_{i \geq 1} i \cdot p_i \cdot x_1^{i-1} x_2^i \right)$, $\alpha_2 \left(\sum_{i \geq 1} i \cdot p_i \cdot x_1^i x_2^{i-1} \right)$, and $\sum_{y \in \mathbf{y}} \beta_y \left(\sum_{i \geq 0} (x_1 x_2)^i \frac{\partial p_i}{\partial y} \right)$ do not have common monomials in x_1, x_2 variables. Thus, each of these three polynomials should be equal to zero. Suppose $\alpha_1 \neq 0$. Then,

$$\sum_{i \geq 1} i \cdot p_i \cdot x_1^{i-1} x_2^i = 0.$$

As $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$ and for every $i \geq 1, p_i \in \mathbb{F}[\mathbf{y}]$, we get from the above equation that $p_i = 0$ for every $i \geq 1$, which is a contradiction. Thus, $\alpha_1 = 0$. Similarly, $\alpha_2 = 0$. □

2.1.7 Orbit of a polynomial

First consider the following definition.

Definition 2.33 (Affine projection) *An n-variate polynomial $f(\mathbf{x})$ is said to be an* affine projection *of an m-variate polynomial g if there exist an $A \in \mathbb{F}^{m \times n}$ and a $\mathbf{b} \in \mathbb{F}^m$ such that $f = g(A\mathbf{x} + \mathbf{b})$.*

Definition 2.34 (Orbit) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n, \}, \mathbb{F}$ be a field and $f \in \mathbb{F}[\mathbf{x}]$. Then, the orbit of f, denoted* $\mathrm{orb}(f)$*, is the set $\{f(A\mathbf{x}) : A \in \mathrm{GL}(n, \mathbb{F})\}$.*

It is easy to prove the following fact.

Fact 2.12 *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}$, $f \in \mathbb{F}[\mathbf{x}]$ and $g = f(A\mathbf{x} + \mathbf{b})$, where $A \in \mathrm{GL}(n, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^n$. Then, $N_{ess}(f) = N_{ess}(g)$.*

2.1.8 Algebraic models of computation

Definition 2.35 (Arithmetic circuit) *Let $n \in \mathbb{N}, \mathbb{F}$ be a field and $\mathbf{x} = \{x_1, \ldots, x_n\}$. An arithmetic circuit C over $\mathbb{F}$ is a directed acyclic graph, where leaf nodes are labelled by $\mathbf{x}$ and constants from $\mathbb{F}$, other nodes are labelled by $+, \times, -, \div$ and edges are labelled by constants from $\mathbb{F}$. A node labelled by $+$ or $\times$ has arbitrary fan-in and labelled by $-$ or $\div$ has fan-in two. Computation in C happens as follows: Nodes having in-degree zero compute their labels; if v is a $+$ node (respectively, a $\times$ node) having children $v_1, \ldots, v_r$ such that for every $i \in [r]$, the edge connecting v_i to v is labelled by α_i then v computes $\sum_{i\in[r]} \alpha_i v_i$ (respectively, $\prod_{i\in[r]} \alpha_i v_i$); if v is a $-$ node (respectively, a $\div$ node) having children v_1, v_2 such that edge connecting v_i to v has label α_i then v computes $\alpha_1 v_1 - \alpha_2 v_2$ (respectively, $\frac{\alpha_1 v_1}{\alpha_2 v_2}$, provided $\alpha_2 v_2 \neq 0$). The rational function computed by the node having out-degree 0 is said to be the output of C.*

An example of an arithmetic circuit is given in Figure 1.1.2. We restrict our attention to arithmetic circuits computing polynomials. As argued in Section 1.1.2, we will assume that the non-leaf nodes in an arithmetic circuit are labelled by $+$ and $\times$ operations.

Definition 2.36 (ABP) *Let $d \in \mathbb{N}$, $\mathbb{F}$ be a field and $\mathbf{x}$ be a set of variables. An algebraic branching program (ABP for short) is a directed acyclic graph having $d+1$ layers labelled by $0, \ldots, d$, where the first and the last layer have exactly one node each, called the source (denoted s), and the sink (denoted t) respectively and for $i \in [d-1]$, the i-th layer has w_i many vertices. The edges are present only between vertices of adjacent layers and every edge is labelled by a linear form in $\mathbb{F}[\mathbf{x}]$. Suppose $p = (e_1, \ldots, e_d)$ is an s-t path, where e_i is an edge from layer $(i-1)$ to layer i and the label of e_i is the linear form ℓ_i. Let $g_p := \ell_1 \cdots \ell_d$. Then $\sum_{p:s\text{-}t\ path} g_p$ is the polynomial computed by this ABP.*

Definition 2.37 (Arithmetic formula) *An arithmetic circuit C is called an arithmetic formula if the underlying graph of C is a tree.*

Remark 2.4 *Let* C *be an arithmetic formula. Without loss of generality, we assume from now on that* C *has alternate layers of* $+$ *and* $\times$ *nodes, every non-leaf node in* C *has fan-in at least two and every child of a* $\times$ *gate in* C *computes a non-constant polynomial.*

The *product-depth* of C is the number of $\times$ gate in a longest path in C from a leaf node to the root node of C.

Remark 2.5 *Let* C *be an arithmetic circuit. For convenience, we would also use* C *to denote the polynomial computed by* C*. Similarly, if* v *is a node of* C *then* v *would also denote the polynomial it computes.*

Definition 2.38 (ROF) *An arithmetic formula* C *over a field* $\mathbb{F}$ *is said to be a read-once arithmetic formula (in short, an ROF) if every leaf of* C *is labelled by either a distinct variable or a constant from* $\mathbb{F}$*. If the root node of* C *is a* $+$ *node then we call it a* $+$*-rooted ROF, otherwise it is called a* $\times$*-rooted ROF.*

It follows immediately from the above definition that every ROF computes a mutlilinear polynomial. The following useful property of ROF would be frequently used in Chapter 5.

Observation 2.7 (Irreducibility of a +-rooted ROF) *Let* $\mathbb{F}$ *be a field and* C *be a* $+$*-rooted ROF over* $\mathbb{F}$*. Then,* C *is irreducible over* $\mathbb{F}$*.*

Proof: Let $\mathsf{C} = T_1 + \cdots + T_s + \gamma$, where for every $l \in [s], T_l$ is either a variable or a $\times$-rooted sub-ROF of C and $\gamma \in \mathbb{F}$. Suppose $\mathrm{var}(\mathsf{C}) = \mathbf{x}$. We know that C is a multilinear polynomial in $\mathbb{F}[\mathbf{x}]$. We prove the result in two cases.

Case 1: $s \geq 2$. Suppose C is reducible. Then, there exist non-constant polynomials $g_1, g_2 \in \mathbb{F}[\mathbf{x}]$ such that $\mathsf{C} = g_1 g_2$. Since C is multilinear, $\mathrm{var}(g_1) \cap \mathrm{var}(g_2) = \emptyset$. As g_1 is non-constant, $\mathrm{var}(g_1)$ contains a variable from $\mathrm{var}(T_l)$ for some $l \in [s]$. Then, for every $k \in [s] \setminus \{l\}$, $\mathrm{var}(T_k) \subseteq \mathrm{var}(g_1)$, otherwise we get a monomial in C containing variables from $\mathrm{var}(T_k)$ and $\mathrm{var}(T_l)$, which is not possible. As g_2 is non-constant and $\mathrm{var}(g_1) \cap \mathrm{var}(g_2) = \emptyset$, $\mathrm{var}(g_2)$ must contain some variable in $\mathrm{var}(T_l) \setminus \mathrm{var}(g_1)$. Thus, we get a monomial in C, which contains variables from $\mathrm{var}(T_l)$ and $\mathrm{var}(T_k)$ for some $k \in [s] \setminus \{l\}$. This is a contradiction.

Case 2: $s = 1$. If $\gamma = 0$ then C is not a $+$-rooted ROF, so $\gamma \neq 0$. Then, T_1 is either a variable or $T_1 = Q_1 \cdots Q_m$, where $m \geq 2$ and for every $i \in [m]$, Q_i is either a $+$-rooted sub-ROF of C or a variable. If T_1 is a variable then C is clearly irreducible. Otherwise, let $p_1 = Q_1$ and

$p_2 = Q_2 \cdots Q_m$. Then, $\mathsf{C} = p_1 p_2 + \gamma$ and $\text{var}(p_1) \cap \text{var}(p_2) = \emptyset$. Suppose there exist non-constant polynomials $g_1, g_2 \in \mathbb{F}[\mathbf{x}]$ such that $\mathsf{C} = g_1 g_2$. As C is multilinear, g_1, g_2 are also multilinear and $\text{var}(g_1) \cap \text{var}(g_2) = \emptyset$. Then,

$$p_1 p_2 + \gamma = g_1 g_2. \tag{2.3}$$

Assume without loss of generality that there exists $x_1 \in \text{var}(p_1) \cap \text{var}(g_1)$. Suppose there exists $y_1 \in \text{var}(p_2) \cap \text{var}(g_2)$. In this case, we substitute every variable in $\text{var}(p_1) \setminus \{x_1\}$ and $\text{var}(p_2) \setminus \{y_1\}$ with random $\mathbb{F}$-constants.[1] After this, Equation (2.3) looks like

$$(\alpha_1 x_1 + \alpha_0)(\alpha_3 y_1 + \alpha_2) + \gamma = (\beta_1 x_1 + \beta_0)(\beta_3 y_1 + \beta_2),$$

where $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_0, \beta_1, \beta_2, \beta_3 \in \mathbb{F}$ and it follows from the Schwartz-Zippel lemma (Fact 2.13) that $\alpha_1, \beta_3 \neq 0$ with high probability. Now, on substituting $x_1 = \frac{-\alpha_0}{\alpha_1}$ in the above equation, we get $\gamma = \left(\frac{-\beta_1 \cdot \alpha_0}{\alpha_1} + \beta_0\right)(\beta_3 y_1 + \beta_2)$. This can never happen as $\left(\frac{-\beta_1 \cdot \alpha_0}{\alpha_1} + \beta_0\right)(\beta_3 y_1 + \beta_2)$ is either zero or it contains y_1, whereas $\gamma \in \mathbb{F} \setminus \{0\}$. Thus we get a contradiction. Now, suppose $\text{var}(g_2) \cap \text{var}(p_2) = \emptyset$. Thus, $\text{var}(p_2) \subseteq \text{var}(g_1)$. As g_2 is non-constant, there exists $x_1' \in \text{var}(g_2) \cap \text{var}(p_1)$. Pick a $y_1' \in \text{var}(p_2)$ arbitrarily. We substitute all variables of $\text{var}(p_1) \setminus \{x_1'\}$ and $\text{var}(p_2) \setminus \{y_1'\}$ with random $\mathbb{F}$-constants. Now, using the same argument as before, we get a contradiction. Hence, C is irreducible over $\mathbb{F}$. □

Definition 2.39 (Canonical ROF) *Let $\mathbb{F}$ be a field and C be an ROF over $\mathbb{F}$. Then, C is said to be a canonical ROF if the label associated with every edge in C is one and every $+$ gate C has at most 1 variable child or at most one constant child but not both.*

It is not difficult to show that every ROF is in the orbit of a canonical ROF. Thus, from the viewpoint of equivalence testing, canonical ROFs capture general ROFs.

Definition 2.40 (Regular ROF) *A canonical ROF C over a field is said to be regular if the parent of every variable in C is a $\times$ gate.*

As mentioned in Sections 1.2 and 1.3.3, a useful example of a regular ROF is an arithmetic circuit in a *read-once alternating normal form* (ROANF), which is defined below.

Definition 2.41 (ROANF) *A canonical ROF C over a field $\mathbb{F}$ is said to be in the read-once alternating normal form (in short, ROANF) if the underlying tree of C is a complete binary tree, the bottom-most layer contains $\times$ nodes, and every leaf node is labelled by a distinct variable. Thus, if C is an ROANF and has product-depth Δ then $|\text{var}(\mathsf{C})| = 4^{\Delta}$.*

[1] For this, we need $|\mathbb{F}| > n$. If it is not the case, we work with a large enough extension field $\mathbb{L}$ of $\mathbb{F}$ and give the argument over $\mathbb{L}$. Clearly, if C is irreducible over $\mathbb{L}$, it has to be irreducible over $\mathbb{F}$.

The following property of canonical ROFs is useful for Chapters 5 and 6.

Observation 2.8 (Canonical ROFs are free from redundant variables) *Let* C *be a canonical ROF over a field* $\mathbb{F}$ *and* $\mathbf{x} = \mathrm{var}(\mathsf{C})$*. Then, every variable in* $\mathbf{x}$ *is essential for* C*.*

Proof: Fact 2.11 implies that it is sufficient to show that $\left\{\frac{\partial \mathsf{C}}{\partial x} : x \in \mathbf{x}\right\}$ is $\mathbb{F}$-linearly independent. As C is multilinear, for every $x \in \mathbf{x}$, $\frac{\partial \mathsf{C}}{\partial x} \neq 0$. Consider the following equation

$$\sum_{x \in \mathbf{x}} \alpha_x \frac{\partial \mathsf{C}}{\partial x} = 0, \tag{2.4}$$

where for every $x \in \mathbf{x}, \alpha_x \in \mathbb{F}$. Let $x \in \mathbf{x}$ be an arbitrary variable. Then, path(x) denotes the path starting from the root of C to x and the product-depth of x, denoted Δ_x, is defined as the number of product gates appearing on path(x). For $i \in [\Delta_x]$, let $I_{x,i}$ denote the set of indices of those children of the i-th product gate on path(x), which do not lie on this path. Let $N_x = \{Q_{x,i,j} : i \in [\Delta_x], j \in I_{x,i}\}$. Then, for every $i \in [\Delta_x], j \in I_{x,i}, Q_{x,i,j}$ is a sibling of a $+$-rooted ROF, which lies on path(x) and is a child of the i-th product gate on path(x). Thus, $Q_{x,i,j}$ is either a variable or a $+$-rooted sub-ROF of C. Observe that

$$\frac{\partial \mathsf{C}}{\partial x} = \prod_{i \in [\Delta_x]} \prod_{j \in I_{x,i}} Q_{x,i,j}.$$

Let $x \in \mathbf{x}$ be arbitrary. Then, we claim that there does not exist $x' \in \mathbf{x} \setminus \{x\}$ satisfying $\Delta_{x'} \leq \Delta_x$ such that $N_x \subseteq N_{x'}$. Suppose this is not true. Then, note that $\Delta_x = \Delta_{x'}$. As $N_x \subseteq N_{x'}$ and $\Delta_x = \Delta_{x'}$, observe that this means that x and x' are children of the same $+$ gate in C. But this contradicts the fact that C is a canonical ROF. Hence, such an x' does not exist.

Now, we show that there exists a monomial p_x in $\frac{\partial \mathsf{C}}{\partial x}$ such that for every $x' \in \mathbf{x}$, where $\Delta_{x'} \leq \Delta_x$, p_x is not a monomial of $\frac{\partial \mathsf{C}}{\partial x'}$. Let $i \in [\Delta_x], j \in I_{x,i}$ be arbitrary, $p_{x,i,j}$ be the largest monomial of $Q_{x,i,j}$ under a degree lexicographic order on $\mathbb{F}[\mathbf{x}]$ and $p_x = \prod_{i \in [\Delta_x], j \in I_{x,i}} p_{x,i,j}$. Let $x' \in \mathbf{x} \setminus \{x\}$ be an arbitrary variable such that $\Delta_{x'} \leq \Delta_x$. Suppose the first common ancestor gate of x and x' in C is a $+$ gate. As C is canonical, it is not difficult to see that there exists a $Q_{x,i,j} \in N_x \setminus N_{x'}$. Otherwise, there exists a $Q_{x,i,j} \in N_x \setminus N_{x'}$ such that $N_{x'}$ contains the factors of $\frac{\partial Q_{x,i,j}}{\partial x'}$ but not $Q_{x,i,j}$. As $p_{x,i,j}$ is the largest monomial of $Q_{x,i,j}$ according to the degree lexicographic order, it is not difficult to see that in both the cases, we can not get a monomial in $\frac{\partial \mathsf{C}}{\partial x'}$, which is divisible by $p_{x,i,j}$ and hence p_x can not be a monomial of $\frac{\partial \mathsf{C}}{\partial x'}$. We repeat this for every variable in $\mathbf{x}$ in the non-increasing order of their product-depths. Then, it is not difficult to see that since C is canonical, for every $x \in \mathbf{x}, \alpha_x = 0$ in Equation (2.4). □

Corollary 2.2 (Regular ROFs do not have redundant variables) *Let* C *be a regular ROF over a field* $\mathbb{F}$ *and* $\mathbf{x} = \text{var}(\mathsf{C})$*. Then, every variable in* $\mathbf{x}$ *is essential for* C*.*

2.2 Algorithmic preliminaries

We first give some basic algorithmic results, then give some algorithms related to removal of redundant variables from a polynomial, then give known PE algorithms for quadratic forms over different fields, and finally state known algorithmic results on *full matrix isomorphism problem* over finite fields and $\mathbb{Q}$.

2.2.1 Basic algorithmic facts

Let $\mathbb{F}$ be a field, $\mathbf{x}$ be a set of n variables and $f \in \mathbb{F}[\mathbf{x}]$. Then, black-box of f takes $\mathbf{a} \in \mathbb{F}^n$ and outputs $f(\mathbf{a})$. We first record an important result, which implies a polynomial time randomized black-box PIT algorithm.

Fact 2.13 (Schwartz-Zippel lemma) [DL78, Zip79, Sch80] *Let* $n, d \in \mathbb{N}, \mathbb{F}$ *be a field such that* $|\mathbb{F}| > d$*,* $S \subseteq \mathbb{F}$ *be a finite subset satisfying* $|S| > d$*,* $\mathbf{x} = \{x_1, \ldots, x_n\}$ *and* $f \in \mathbb{F}[\mathbf{x}]$ *be a non-zero polynomial. Then,*

$$\Pr_{a_1,\ldots,a_n \in_r S}(f(a_1, \ldots, a_n) = 0) \leq \frac{d}{|S|},$$

where $a_1, \ldots, a_n \in_r S$ *means that* $a_1, \ldots, a_n$ *are chosen independently and uniformly at random from* S*.*

Now, suppose we are given input n, d and a black-box of f and we want to decide whether f is identically zero or not. We pick a finite set $S \subseteq \mathbb{F}$ satisfying $|S| > d$ (provided $|\mathbb{F}| > d$) and pick $a_1, \ldots, a_n \in_r S$. If $f(a_1, \ldots, a_n) = 0$ we output $f = 0$ otherwise output $\mathsf{C} \neq 0$. The Schwartz-Zippel lemma implies that our result is correct with high probability. Now, we state other useful results about a polynomial given as black-box.

Fact 2.14 (Computing black-box access to derivatives) *Let* $n, d \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ *be a field satisfying* $char(\mathbb{F}) = 0$ *or greater than* d *and* $f \in \mathbb{F}[\mathbf{x}]$ *be a degree* d *polynomial. Given an* $x \in \mathbf{x}$ *and black-box access to* f*, we can compute black-box access to* $\frac{\partial f}{\partial x}$ *in* $\text{poly}(n, d)$ *time.*

A proof of the above well-known fact is given in Section 2.2 of [KNST19]. The following fact follows from Clam 2.2 of [KNST19], whose proof is based on the Schwartz-Zippel lemma.

Fact 2.15 (Computing black-box access to a basis) *Let $n, d, m \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}$, and $\mathbb{F}$ be a field satisfying $|\mathbb{F}| > (dm)^2$. Suppose we are given black-box access to $f_1, \ldots, f_m \in \mathbb{F}[\mathbf{x}]$. Then, there exists a randomized algorithm which computes black-box access to an $\mathbb{F}$-basis of $\langle f_1, \ldots, f_m \rangle$ in time $\mathrm{poly}(n, m, d)$.*

The following fact allows us to compute black-box access to an $\mathbb{F}$-basis of the Lie algebra of a polynomial given as black-box. It was proven in [Kay12] and the proof is based on Equation (2.1) and Facts 2.14 and 2.15.

Fact 2.16 (Computing a basis of $\mathfrak{g}_f$) *Let $n, d \in \mathbb{N}, \mathbb{F}$ be a field satisfying $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$ and $f \in \mathbb{F}[\mathbf{x}]$ be a degree d polynomial. Then, there exists a randomized algorithm that takes black-box access to f and outputs a basis of $\mathfrak{g}_f$ with high probability in $\mathrm{poly}(n, d)$ time.*

The following fact is extensively used in the equivalence test given in Chapter 5.

Fact 2.17 (Black-box polynomial factorization algorithm [KT90]) *Let $n, d \in \mathbb{N}, \mathbb{F}$ be a field satisfying $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$ and $|\mathbb{F}| \geq d^6$, $\mathbf{x} = \{x_1, \ldots, x_n\}$, and $f \in \mathbb{F}[\mathbf{x}]$ be a degree d polynomial. Then, there exists a randomized algorithm that takes black-box access to f, oracle access to a univariate polynomial factorization algorithm over $\mathbb{F}$ and outputs black-boxes of non-zero scalar multiples of irreducible factors of f with high probability in $\mathrm{poly}(n, d)$ time.*

Remark 2.6 *In Chapter 5, we assume that univariate polynomials can be factorized efficiently over the underlying field $\mathbb{F}$. This implies that the factorization algorithm given in Fact 2.17 is also efficient over $\mathbb{F}$. This assumption holds over finite fields and $\mathbb{Q}$ as efficient univariate polynomial factorization algorithms are known over these fields (see [Ber70, LLL82b]).*

Fact 2.18 (Computing closure of a vector) *Let $m, n \in \mathbb{N}, \mathbb{F}$ be a field and $\mathscr{T} \subseteq M_n$ be an $\mathbb{F}$-vector space of dimension m. Then, there exists a polynomial time algorithm, that takes a $\mathbf{v} \in \mathbb{F}^n$ and a basis $\{M_1, \ldots, M_m\}$ of $\mathscr{T}$ and outputs a basis of $\mathrm{closure}_{\mathscr{T}}(\mathbf{v})$.*

See Section 4.2 of [KNST19] for a proof of the above fact.

2.2.2 Removal of redundant variables

In this section, we talk about algorithms to get rid of redundant variables from a polynomial. These algorithms would be used in the equivalence test given in Chapter 5. As mentioned in Section 2.1.6, algorithms to remove redundant variables are known in two settings - when input

is a list of coefficients [Car06] and when the input is a black-box [Kay11] (see also Claim 2.3 of [KNST19]). In this section, whenever we say that a set of variables $\mathbf{z}$ is redundant for an n-variate polynomial $f \in \mathbb{F}[\mathbf{x}]$, we would mean that there exists $A \in \mathrm{GL}(n, \mathbb{F})$ such that $f(A\mathbf{x})$ does not contain any $\mathbf{z}$-variable. In Claim 2.2.1, we present a slightly general version of the algorithm [Kay11].

Observation 2.9 *Let $n, d \in \mathbb{N}, \mathbf{x}$ be a set of n variables and $\mathbb{F}$ be a field satisfying either $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$ and $|\mathbb{F}| \geq 2n^2d$. Suppose we are given $\mathbf{y}' \subseteq \mathbf{x}, \mathbf{x}' \subseteq \mathbf{x}$ and black-box access to $g \in \mathbb{F}[\mathbf{x}]$ such that $\deg(g) \leq d$, $\left\{\frac{\partial g}{\partial y} : y \in \mathbf{y}'\right\}$ is $\mathbb{F}$-linearly independent, there exists $\mathbf{y} \subseteq \mathbf{x}'$ such that $\mathbf{y}' \subseteq \mathbf{y}$ and $\mathbf{y}$ is a set of essential variables of g. Then, such a $\mathbf{y}$ can be computed in randomized $\mathrm{poly}(n, d)$ time.*

Proof:

Procedure 1 Compute-Essential-Vars$(g, \mathbf{y}', \mathbf{x}')$

Input: Black-box access to $g \in \mathbb{F}[\mathbf{x}], \mathbf{y}' \subseteq \mathbf{x}, \mathbf{x}' \subseteq \mathbf{x}$, s.t. $\left\{\frac{\partial g}{\partial y} : y \in \mathbf{y}'\right\}$ is $\mathbb{F}$-linearly independent and $\exists\, \mathbf{y} \subseteq \mathbf{x}'$, s.t. $\mathbf{y}' \subseteq \mathbf{y}$ and $\mathbf{y}$ is a set of essential variables of g.

Output: $\mathbf{y} \subseteq \mathbf{x}'$, s.t. $\mathbf{y}' \subseteq \mathbf{y}$ and $\mathbf{y}$ is a set of essential variables of g.

1. $F \leftarrow$ a subset of $\mathbb{F}$ of size at least $2n^2d$. $\mathbf{y} \leftarrow \mathbf{y}'$ and $\mathbf{z} \leftarrow \mathbf{x}' \setminus \mathbf{y}$.
2. **for** $z \in \mathbf{z}$ **do**
3. Compute black-box access to $\frac{\partial g}{\partial z}$ and $\frac{\partial g}{\partial y}$ for every $y \in \mathbf{y}$.
4. For $x \in \mathbf{y} \cup \{z\}$, let $\mathbf{a}_x$ be chosen uniformly at random from F^n. $C \leftarrow \left(\frac{\partial g}{\partial x_2}(\mathbf{a}_{x_1})\right)_{x_1, x_2 \in \mathbf{y} \cup \{z\}}$. If $\det(C) \neq 0$, then $\mathbf{y} \leftarrow \mathbf{y} \cup \{z\}$ and $\mathbf{z} \leftarrow \mathbf{z} \setminus \{z\}$.
5. **end for**
6. Return $\mathbf{y}$.

It follows from Fact 2.14 that this algorithm runs in $\mathrm{poly}(n, d)$ time. Now, we argue its correctness. Consider a specific iteration of the loop of lines 2 - 5. For $x \in \mathbf{y} \cup \{z\}$, let $\mathbf{a}_x = (a_{x,i})_{i \in [n]}$. For every $i \in [n]$, treat $a_{x,i}$ as a variable. Note that $\det(C) \neq 0$ if and only if $N := \left\{\frac{\partial g}{\partial x} : x \in \mathbf{y} \cup \{z\}\right\}$ is $\mathbb{F}$-linearly independent. Since $\deg(\det(C)) < nd$, the Schwartz-Zippel implies that after substituting $a_{x,i}$s with random values, the probability that N is linearly independent but $\det(C) = 0$ is at most $\frac{nd}{2n^2d}$. Now, by using the union bound on the error probability, we get that $\mathbf{y}$ computed at the end of the algorithm satisfies the desired property with probability at least $\frac{1}{2}$. □

Claim 2.2.1 (Eliminating redundant variables) *Let $n, d \in \mathbb{N}, \mathbb{F}$ be a field satisfying $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$ and $|\mathbb{F}| \geq 2nd^2$, $\mathbf{x} = \{x_1, \ldots, x_n\}$, $g \in \mathbb{F}[\mathbf{x}]$ a degree d polynomial and $\mathbf{z} \subseteq \mathbf{x}$ be a set of redundant variables of g. Then there exists a randomized $\mathrm{poly}(n, d)$ time algorithm that takes input $\mathbf{z}$ and black-box access to g and outputs an $A \in \mathrm{GL}(n, \mathbb{F})$ that maps every $\mathbf{z}$-variable to itself such that $g(A\mathbf{x}) \cap \mathbf{z} = \emptyset$ and every variable in $\mathrm{var}(g(A\mathbf{x}))$ is essential.*

Proof: The following algorithm is obtained from algorithms given in [Kay11, KNST19] for removing redundant variables from a polynomial given as black-box.

Algorithm 2 Remove-Redundant-Vars$(g, \mathbf{z})$

Input: Black-box access to $g \in \mathbb{F}[\mathbf{x}]$ and $\mathbf{z} \subseteq \mathbf{x}$ such that all variables in $\mathbf{z}$ are redundant for g.
Output: $A \in \mathrm{GL}(n, \mathbb{F})$ such that A maps every variable in $\mathbf{z}$ to itself and $g(A\mathbf{x})$ does not contain a redundant variable, including the $\mathbf{z}$-variables.

1. $\mathbf{y} \leftarrow$ Compute-Essential-Vars$(g, \emptyset, \mathbf{x} \setminus \mathbf{z})$ (Procedure 1), $\mathbf{z}' \leftarrow \mathbf{x} \setminus \mathbf{y}$.
2. $F \leftarrow$ a subset of $\mathbb{F}$ of size at least $2n^2d$.
3. **for** $z \in \mathbf{z}'$ **do**
4. Compute the values of $\alpha_{y,z}, y \in \mathbf{y}$ by solving a system of linear equations obtained by evaluating Equation (2.5) at $\{\mathbf{a}_y \in F^n : y \in \mathbf{y}\}$, where $\mathbf{a}_y$ is chosen independently and uniformly at random from F^n for every $y \in \mathbf{y}$.

$$\sum_{y \in \mathbf{y}} \alpha_{y,z} \frac{\partial g}{\partial y} + \frac{\partial g}{\partial z} = 0. \tag{2.5}$$

5. **end for**
6. Let $A \in \mathbb{F}^{n \times n}$ such that $\forall y \in \mathbf{y}, z \in \mathbf{z}'$, the (y, z)-th entry of A, is $\alpha_{y,z}$, the (y, y)-th and (z, z)-th entry of A is 1 and every other entry is 0. Return A.

It follows from Observation 2.9 that its running time is $\mathrm{poly}(n, d)$. Now, we argue its correctness. Observation 2.9 implies that $\mathbf{y}$ is a set of essential variables of f and $\mathbf{z} \cap \mathbf{y} = \emptyset$. Hence, $\mathbf{z}'$ is a set of redundant variables of g.

As $\left\{\frac{\partial g}{\partial y} : y \in \mathbf{y}\right\}$ forms a basis of $\left\{\frac{\partial g}{\partial x} : x \in \mathbf{x}\right\}$, clearly for every $z \in \mathbf{z}'$, there exist $\alpha_{y,z} \in \mathbb{F}, y \in \mathbf{y}$ such that Equation (2.5) is satisfied. Fix $z \in \mathbf{z}'$ arbitrarily. Let C be a matrix, whose rows and columns are labelled by $\mathbf{y}$ and for $y_1, y_2 \in \mathbf{y}$, the (y_1, y_2)-th entry of C is $\frac{\partial g}{\partial y_1}(\mathbf{a}_{y_2})$. Let $\boldsymbol{\beta}_z := \left(-\frac{\partial g}{\partial z}(\mathbf{a}_y)\right)_{y \in \mathbf{y}}$ and $\boldsymbol{\alpha}_z \in \mathbb{F}^{|\mathbf{y}|}$ be such that for $y \in \mathbf{y}$, the y-th entry is $\alpha_{y,z}$. Then, Equation (2.5) implies $C \cdot \boldsymbol{\alpha}_z = \boldsymbol{\beta}_z$. Now, the Schwartz-Zippel lemma implies that C is invertible with high probability and thus we get correct values of $\alpha_{y,z}, y \in \mathbf{y}$ with high probability.

Consider the matrix A computed in Step 6. Note that is invertible and has the following structure: A maps every variable in $\mathbf{z}'$ to itself and every $y \in \mathbf{y}$ to a linear form in y and $\mathbf{z}'$. It follows from Fact 2.11 that $g(A\mathbf{x})$ does not have redundant variables including the $\mathbf{z}$-variables. □

Suppose $g_1, \ldots, g_m \in \mathbb{F}[\mathbf{x}]$ are pairwise variable disjoint polynomials. Then, it is easy to see that $N_{ess}(g_1 \cdots g_m) = N_{ess}(g_1) + \cdots + N_{ess}(g_m)$. In Claim 2.2.2, we prove the converse of this and give an algorithm to compute an $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ such that $f_1(A\mathbf{x}), \ldots, f_m(A\mathbf{x})$ are pairwise variable disjoint. Before that, we note some useful observations needed for this claim.

Observation 2.10 *Let $d, m, n \in \mathbb{N}$, $\mathbb{F}$ be a field satisfying either $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$, $\mathbf{x} = \{x_1, \ldots, x_n\}$ and $g_1, \ldots, g_m \in \mathbb{F}[\mathbf{x}]$ such that $\deg(g_1 \cdots g_m) \leq d$.*

1. *$N_{ess}(g_1 \cdots g_m) \leq N_{ess}(g_1) + \cdots + N_{ess}(g_m)$.*
2. *Let $N_{ess}(g_1 \cdots g_m) = N_{ess}(g_1) + \cdots + N_{ess}(g_m)$ and $I \subseteq [m]$ be a non-empty set. Then,*

$$N_{ess}\left(\prod_{i \in I} g_i\right) = \sum_{i \in I} N_{ess}(g_i).$$

Proof:

1. We know from Fact 2.11 that for every $g \in \mathbb{F}[\mathbf{x}]$, $N_{ess}(g) = \dim \left\langle \frac{\partial g}{\partial x} \,:\, x \in \mathbf{x} \right\rangle$. For $i \in [m]$, let $U_i = \left\langle \frac{\partial g_i}{\partial x} \,:\, x \in \mathbf{x} \right\rangle$ and $W_i = \left\langle \prod_{j \in [m] \setminus \{i\}} g_j \cdot \frac{\partial g_i}{\partial x} \,:\, x \in \mathbf{x} \right\rangle$. As $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$, $\frac{\partial g_i}{\partial x} \neq 0$ for every $i \in [m], x \in \mathbf{x}$. Note that for every $i \in [m]$, $\dim U_i = \dim W_i$. Let $g = g_1 \cdots g_m$ and $W = \left\langle \frac{\partial g}{\partial x} \,:\, x \in \mathbf{x} \right\rangle$. Then, it is easy to see that W is an $\mathbb{F}$-subspace of $W_1 + \cdots + W_m$. Hence, $\dim W \leq \dim W_1 + \cdots + \dim W_m$, which implies $N_{ess}(g_1 \cdots g_m) \leq N_{ess}(g_1) + \cdots + N_{ess}(g_m)$.
2. Let $I \subseteq [m], h_1 := \prod_{i \in I} g_i$ and $h_2 := \prod_{i \in [m] \setminus I} g_i$. We know from Part 1 that $N_{ess}(h_1 h_2) \leq N_{ess}(h_1) + N_{ess}(h_2)$. Suppose $N_{ess}(h_1) < \sum_{i \in I} N_{ess}(g_i)$. Then, using Part 1 we get $N_{ess}(h_2) \leq \sum_{i \in [m] \setminus I} N_{ess}(g_i)$, which implies $N_{ess}(h_1 h_2) < N_{ess}(g_1) + \cdots + N_{ess}(g_m)$. This is a contradiction as $h_1 h_2 = g_1 \cdots g_m$. Hence, $N_{ess}(\prod_{i \in I} g_i) = \sum_{i \in I} N_{ess}(g_i)$.

□

Observation 2.11 *Let $d, n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}$ and $\mathbb{F}$ be a field satisfying $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$. Let $\mathbf{y} \subseteq \mathbf{x}, p \in \mathbb{F}[\mathbf{y}]$, and $q \in \mathbb{F}[\mathbf{x}]$ be such that $\mathbf{y}$ is a set of essential variables of p, $\deg(pq) \leq d$ and $N_{ess}(pq) = N_{ess}(p) + N_{ess}(q)$. Then, $\mathbf{y}$ is redundant for q.*

Proof: Let $\mathbf{x}' \subseteq \mathbf{x}$ be a set of essential variables of q and $\mathbf{z} = \mathbf{x} \setminus \mathbf{y}$. If $\mathbf{x}' \cap \mathbf{y} = \emptyset$ then $\mathbf{y}$ is redundant for q. Suppose $\mathbf{x}' \cap \mathbf{y} \neq \emptyset$. In this case, we show that there exists $C \in \mathrm{GL}(n, \mathbb{F})$, such that C maps every $\mathbf{y}$-variable to a linear form in $\mathbf{y}$-variables, $q(C\mathbf{x})$ does not have redundant variables and $\mathrm{var}(q(C\mathbf{x})) \cap \mathbf{y} = \emptyset$. Let $\mathbf{y}' \subseteq \mathbf{y}$ and $\mathbf{z}' \subseteq \mathbf{z}$ such that $\mathbf{x}' = \mathbf{y}' \uplus \mathbf{z}'$. Then, for every $x \in \mathbf{x} \setminus \mathbf{x}'$, there exist $\alpha_{x,y'}, \beta_{x,z'} \in \mathbb{F}$ for every $y' \in \mathbf{y}', z' \in \mathbf{z}'$ such that

$$\sum_{y' \in \mathbf{y}'} \alpha_{x,y'} \frac{\partial q}{\partial y'} + \sum_{z' \in \mathbf{z}'} \beta_{x,z'} \frac{\partial q}{\partial z'} = \frac{\partial q}{\partial x}.$$

We call this as *the equation corresponding to* x. As $\mathrm{char}(\mathbb{F}) = 0$ or $> d$, $\frac{\partial q}{\partial x} \neq 0$ for every $x \in \mathbf{x}$. Suppose $z \in \mathbf{z} \setminus \mathbf{z}'$. Consider the equation corresponding to z. If there exists $y' \in \mathbf{y}'$, such that $\alpha_{z,y'} \neq 0$ then we swap $\alpha_{z,y'} \frac{\partial q}{\partial y'}$ and $\frac{\partial q}{\partial z}$ in this equation and divide the equation by $\alpha_{z,y'}$. Now, we substitute $\frac{\partial q}{\partial y'}$ with $-\left(\sum_{y'' \in \mathbf{y}' \setminus \{y'\}} \frac{\alpha_{z,y''}}{\alpha_{z,y'}} \frac{\partial q}{\partial y''} + \sum_{z' \in \mathbf{z}'} \frac{\beta_{z,z'}}{\alpha_{z,y'}} \frac{\partial q}{\partial z'} - \frac{1}{\alpha_{z,y'}} \frac{\partial q}{\partial z}\right)$ in every equation other that the equation corresponding to z. We iteratively do this for every $\mathbf{z} \setminus \mathbf{z}'$ satisfying the property that after the substitutions in the previous iterations, for some $y' \in \mathbf{y}'$ the coefficient of $\frac{\partial q}{\partial y'}$ in the equation corresponding to z is still non-zero.

After doing this, we have two sets of equations: in the first set, the R.H.S. of every equation is $\frac{\partial q}{\partial y}$ for some $y \in \mathbf{y}$ and in the second set, the R.H.S. of every equation is $\frac{\partial q}{\partial z}$ for some $z \in \mathbf{z}$. It is easy to note that in every equation in the second set, the coefficient of $\frac{\partial q}{\partial y}$ is equal to 0 for every $y \in \mathbf{y}$. Let $\mathbf{y}_1 \subseteq \mathbf{y}$ be such that all the equations in the first set look like

$$\sum_{x \in \mathbf{x} \setminus \{y_1\}} \alpha'_{x,y_1} \frac{\partial q}{\partial x} = \frac{\partial q}{\partial y_1}, \quad y_1 \in \mathbf{y}_1, \tag{2.6}$$

where $\alpha'_{x,y_1} \in \mathbb{F}$ for every $y_1 \in \mathbf{y}_1, x \in \mathbf{x} \setminus \{y_1\}$. Similarly, let $\mathbf{z}_1 \subseteq \mathbf{z}$ be such that all the equations in the second set look like

$$\sum_{x \in \mathbf{x} \setminus \{z_1\}} \beta'_{x,z_1} \frac{\partial q}{\partial x} = \frac{\partial q}{\partial z_1}, \quad z_1 \in \mathbf{z}_1, \tag{2.7}$$

where for every $z_1 \in \mathbf{z}_1, x \in \mathbf{x} \setminus \{z_1\}, \beta'_{x,z_1} \in \mathbb{F}$. Further, we know that for every $z_1 \in \mathbf{z}_1, x \in \mathbf{y}$, $\beta'_{x,z_1} = 0$. Let C be an $n \times n$ matrix defined as follows: for every $y_1 \in \mathbf{y}_1, z_1 \in \mathbf{z}_1, x \in \mathbf{x} \setminus \{y_1\}, x' \in \mathbf{x} \setminus \{z_1\}$, the $(y_1, y_1), (z_1, z_1)$-th entries of C are -1, the (x, y_1)-th and (x', z_1)-th entries of C are α'_{x,y_1} and β'_{x',z_1} given in Equations (2.6) and (2.7) respectively; for every $x \in \mathbf{x} \setminus (\mathbf{y}_1 \cup \mathbf{z}_1)$, the (x, x)-th entry of A is 1 and other entries are zero.

Note that for $y_1 \in \mathbf{y}_1, z_1 \in \mathbf{z}_1$ the columns of C labelled by y_1 and z_1 are associated with the coefficients involved in the Equations (2.6) and (2.7) corresponding to y_1 and z_1 respectively.

Also, notice that $\mathbf{z}_1 \subseteq \mathbf{z} \setminus \mathbf{z}'$. It is not difficult to see that for every $x \in \mathbf{y}_1 \cup \mathbf{z}_1$, there exists exactly one equation in the group of equations given in (2.6) and (2.7), where the coefficient of $\frac{\partial q}{\partial x}$ is non-zero. Using this we can show that $C \in \mathrm{GL}(n, \mathbb{F})$. Observe that C maps a $\mathbf{y}$-variable to a linear form in $\mathbf{y}$-variables. Thus, $p(C\mathbf{x}) \in \mathbb{F}[\mathbf{y}]$ and every variable in $\mathbf{y}$ is essential for $p(C\mathbf{x})$. Further, it follows from the proof of Fact 2.11 given in [KNST19] (see Claim 2.3 in [KNST19]) that $q(C\mathbf{x})$ is free from redundant variables. It is also easy to see that $\mathrm{var}(q(C\mathbf{x})) \cap \mathbf{y} = \emptyset$, otherwise $N_{ess}(pq) < N_{ess}(p) + N_{ess}(q)$. Thus, $\mathbf{y}$ is redundant for q. □

Claim 2.2.2 (Making polynomials variable disjoint) *Let $n, d, m \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ be a field satisfying $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) > d$ and $|\mathbb{F}| \geq 2 \cdot nd$, and $g_1, \ldots, g_m \in \mathbb{F}[\mathbf{x}]$ such that $N_{ess}(g_1 \cdots g_m) = N_{ess}(g_1) + \cdots N_{ess}(g_m)$. Then there exists a randomized $\mathrm{poly}(n, d)$ time algorithm that takes black-box access to $g_1, \ldots, g_m$ and outputs an $A \in \mathrm{GL}(n, \mathbb{F})$ such that for every $i \in [m]$, g_i does not contain redundant variables and for $i \neq j \in [m]$, $\mathrm{var}(g_i(A\mathbf{x})) = \mathrm{var}(g_j(A\mathbf{x}))$.*

Proof: We first give the algorithm and then argue its correctness.

Algorithm 3 Make-Polys-Var-Disjoint$(g_1, \ldots, g_m)$

Input: Black-box access to $g_1, \ldots g_m \in \mathbb{F}[\mathbf{x}]$ such that $N_{ess}(g_1 \cdots g_m) = N_{ess}(g_1) + \cdots + N_{ess}(g_m)$.

Output: $A \in \mathrm{GL}(n, \mathbb{F})$ such that $\forall i \in [m], g_i(A\mathbf{x})$ does not have redundant variables and for every $i, j \in [m], i \neq j, \mathrm{var}(g_i(A\mathbf{x})) \cap \mathrm{var}(g_j(A\mathbf{x})) = \emptyset$.

1. $A \leftarrow I_{n \times n}$, $\mathbf{y} \leftarrow \emptyset$.
2. **for** $i = 1, \ldots, m$ **do**
3. $\quad A_i \leftarrow$ Remove-Redundant-Vars$(g_i(A\mathbf{x}), \mathbf{y}), \mathbf{y}_i \leftarrow \mathrm{var}(g_i(AA_i\mathbf{x}))$.
4. $\quad A \leftarrow AA_i, \mathbf{y} \leftarrow \mathbf{y} \cup \mathbf{y}_i$.
5. **end for**
6. Return A.

It follows from Claim 2.2.1 that the above algorithm runs in time $\mathrm{poly}(n, d)$. The correctness of the algorithm follows from Subclaim 2.2.1. □

Subclaim 2.2.1 *Let $i \in [m]$ be arbitrarily chosen and g, A, and $\mathbf{y}$ be as after the i-th iteration of Algorithm 3. Then, for every $j \leq i, g_j(A\mathbf{x})$ does not have redundant variables and for every $j, k \in [i], j \neq k, g_j(A\mathbf{x})$ and $g_k(A\mathbf{x})$ are variable disjoint. Further, $g(A\mathbf{x}) \in \mathbb{F}[\mathbf{y}]$ and every variable in $\mathbf{y}$ is essential for $g(A\mathbf{x})$.*

Proof: We prove the subclaim by induction on i. Suppose $i = 1$. Then, it follows from the proof of correctness of Algorithm 2 that $g_1(A\mathbf{x})$ does not have redundant variables and $\mathbf{y}$

is a set of essential variables of $g(A\mathbf{x})$. Now, suppose $i \geq 2$ and the claim holds for $i-1$. Let $g' = g_1 \cdots g_{i-1}$, $A' = A_1 \cdots A_{i-1}$ and $\mathbf{y}' = \mathbf{y}_1 \cup \cdots \cup \mathbf{y}_{i-1}$. Then, we know from the induction hypothesis that for every $j \in [i-1], g_j(A'\mathbf{x})$ does not have redundant variables, for every $j, k \in [i-1], j \neq k, \text{var}(g_j(A'\mathbf{x})) \cap \text{var}(g_k(A'\mathbf{x})) = \emptyset$ and $g'(A'\mathbf{x}) \in \mathbb{F}[\mathbf{y}']$, such that every variable in $\mathbf{y}'$ is essential for $g'(A'\mathbf{x})$.

Recall $g = g'g_i$. Since $N_{ess}(g_1 \cdots g_m) = N_{ess}(g_1) + \cdots + N_{ess}(g_m)$, Observation 2.10 implies that $N_{ess}(g') = N_{ess}(g_1) + \cdots + N_{ess}(g_{i-1})$ and $N_{ess}(g) = N_{ess}(g') + N_{ess}(g_i)$. As A' is an invertible matrix, Fact 2.12 implies that $N_{ess}(g(A'\mathbf{x})) = N_{ess}(g'(A'\mathbf{x})) + N_{ess}(g_i(A'\mathbf{x}))$. Since from the induction hypothesis, we have that $g'(A'\mathbf{x}) \in \mathbb{F}[\mathbf{y}']$ and it has no redundant variables, it follows from the proof of Observation 2.11 that $\mathbf{y}'$ is a set of redundant variables for $g_i(A'\mathbf{x})$. Then from Claim 2.2.1, $A_i \in \text{GL}(n, \mathbb{F})$ maps every variable in $\mathbf{y}'$ to itself and $g_i(A'A_i\mathbf{x}) \in \mathbb{F}[\mathbf{y}_i]$ is free from redundant variables. Let $A = A'A_i$ and $\mathbf{y} = \mathbf{y}' \cup \mathbf{y}_i$. Since A_i maps every variable in $\mathbf{y}'$ to itself, $g'(A\mathbf{x}) = g'(A'\mathbf{x})$. This immediately implies that for every $j \in [i], g_j(A\mathbf{x})$ does not have redundant variables, for every $j, k \in [i], j \neq k, g_j(A\mathbf{x})$ and $g_k(A\mathbf{x})$ are variable disjoint and $g(A\mathbf{x}) \in \mathbb{F}[\mathbf{y}]$ does not contain redundant variables. □

The next claim is used in the ET for regular ROFs and it depends on Claim 2.2.2.

Claim 2.2.3 (Making factors variable disjoint) *Let $d, m, n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}$ and $\mathbb{F}$ be a field satisfying either $\text{char}(\mathbb{F}) = 0$ or $\text{char}(\mathbb{F}) > d$ and $|\mathbb{F}| \geq \max\{2n^2d, d^6\}$. Let $g, g_1, \ldots, g_m \in \mathbb{F}[\mathbf{x}]$ be such that $g = g_1 \cdots g_m$, where $g_1, \ldots, g_m$ need not be irreducible, $\deg(g) \leq d$ and for every $i, j \in [m], i \neq j, \text{var}(g_i) \cap \text{var}(g_j) = \emptyset$. There is a randomized $\text{poly}(n, d)$ time algorithm that takes black-box access to $g(B\mathbf{x} + \mathbf{d})$, where $B \in \text{GL}(n, \mathbb{F}), \mathbf{d} \in \mathbb{F}^n$ and does the following:*

1. *It computes $A \in \text{GL}(n, \mathbb{F})$ such that for every $i \in [m], g_i(BA\mathbf{x}+\mathbf{d})$ does not have redundant variables and for every $i, j \in [m], i \neq j, \text{var}(g_i(BA\mathbf{x} + \mathbf{d})) \cap \text{var}(g_j(BA\mathbf{x} + \mathbf{d})) = \emptyset$.*

2. *It also computes a set $V = \{\text{var}(h_{i,l}(A\mathbf{x})) : i \in [m], l \in [m_i]\}$, where for every $i \in [m]$, there exist m_i polynomials $h_{i,1}, \ldots, h_{i,m_i}$, such that $\prod_{l \in [m_i]} h_{i,l} = g_i(B\mathbf{x} + \mathbf{d})$ and for distinct $l, l' \in [m_i], \text{var}(h_{i,l}(A\mathbf{x})) \cap \text{var}(h_{i,l'}(A\mathbf{x})) = \emptyset$.*

Proof: We give the algorithm and then argue its correctness.

Algorithm 4 Make-Factors-Var-Disjoint($g(B\mathbf{x} + \mathbf{d})$)

Input: Black-box access to $g(B\mathbf{x} + \mathbf{d})$, where $g = g_1 \cdots g_m$ and $\forall i, j \in [m], i \neq j, \text{var}(g_i) \cap \text{var}(g_j) = \emptyset$.

Output: $A \in \text{GL}(n, \mathbb{F})$ and a set V as given in Claim 2.2.3

1. Factorize $g(B\mathbf{x}+\mathbf{d})$ using Fact 2.17 and $N \leftarrow \{h_1, \ldots, h_{s'}\}$ is the set of black-boxes of the irreducible factors of $g(B\mathbf{x}+\mathbf{d})$.
2. **while** $N_{ess}(\prod_{h\in N} h) \neq \sum_{h\in N} N_{ess}(h)$ **do** ,
3. For the first $i \in [|N|]$, such that $N_{ess}(h_1 \cdots h_i) \neq N_{ess}(h_1)+\cdots+N_{ess}(h_i)$, find a $k \in [i-1]$, such that $N_{ess}(h_1 \cdots h_{k-1} \cdot h_i) = N_{ess}(h_1)+\cdots+N_{ess}(h_{k-1})+N_{ess}(h_i)$ but $N_{ess}(h_1 \cdots h_k \cdot h_i) < N_{ess}(h_1)+\cdots+N_{ess}(h_k)+N_{ess}(h_i)$.
4. $N \leftarrow N \cup \{h_k \cdot h_i\}, N \leftarrow N \setminus \{h_k, h_i\}$.
5. **end while**
6. Let $N = \{h_1, \cdots, h_s\}$. $A \leftarrow$ Make-Polys-Var-Disjoint$(h_1, \ldots, h_s)$ (Algorithm 3).
7. $V \leftarrow \{\text{var}(h_1(A\mathbf{x})), \ldots, \text{var}(h_s(A\mathbf{x}))\}$.
8. Return A, V.

We first figure out the running time of Algorithm 4. Fact 2.17 ensures that Step 1 runs in randomized polynomial time. Note that there can be at most $s' \leq d$ iterations of the the loop of lines 2-5 and that each iteration executes in poly(n, d) time. Claim 2.2.2 implies that Step 6 also gets executed in randomized poly(n, d) time. In the next step, we can compute V in randomized polynomial time by using the first-order partial derivatives and a randomized PIT algorithm. Thus, the above algorithm has running time poly(n, d).

Now, we argue the correctness of this algorithm. We know that $g = g_1 \cdots g_m$ and for every $i, j \in [m], i \neq j, \text{var}(g_i) \cap \text{var}(g_j) = \emptyset$. Then, $N_{ess}(g) = N_{ess}(g_1) + \cdots + N_{ess}(g_m)$. Fact 2.11 implies that $N_{ess}(g(B\mathbf{x}+\mathbf{d})) = N_{ess}(g_1(B\mathbf{x}+\mathbf{d}))+\cdots+N_{ess}(g_m(B\mathbf{x}+\mathbf{d}))$. We first collect the black-boxes of the irreducible factors of $g(B\mathbf{x}+\mathbf{d})$ in N and then keep on updating N in Step 2 until $N_{ess}(\prod_{h\in N} h) = \sum_{h\in N} N_{ess}(h)$. The following subclaim ensures that Step 2 is correct.

Subclaim 2.2.2 *Suppose there exists an $i \in [|N|]$, such that $N_{ess}(h_1 \cdots h_{i-1}) = N_{ess}(h_1) + \cdots + N_{ess}(h_{i-1})$ but $N_{ess}(h_1 \cdots h_i) < N_{ess}(h_1)+\cdots+N_{ess}(h_i)$. Then, there exists a $k \in [i-1]$, such that $N_{ess}(h_1 \cdots h_{k-1} \cdot h_i) = N_{ess}(h_1)+\cdots+N_{ess}(h_{k-1})+N_{ess}(h_i)$ but $N_{ess}(h_1 \cdots h_k \cdot h_i) < N_{ess}(h_1)+\cdots+N_{ess}(h_k)+N_{ess}(h_i)$. Further, there exists a $j \in [m]$, such that h_i and h_k are the factors of $g_j(B\mathbf{x}+\mathbf{d})$.*

A proof of Subclaim 2.2.2 is given after this proof. This subclaim ensures that after the execution of the loop of lines 2-5, $N = \{h_1, \ldots, h_s\}$ satisfies $N_{ess}(h_1 \cdots h_s) = N_{ess}(h_1)+\cdots+N_{ess}(h_s)$. So, we can invoke Algorithm 3 on N, which returns $A \in \text{GL}(n, \mathbb{F})$, such that $h_1(A\mathbf{x}), \ldots, h_s(A\mathbf{x})$ are pairwise variable disjoint.

Now, we have to show that $g_1(BA\mathbf{x}+\mathbf{d}),\ldots,g_m(BA\mathbf{x}+\mathbf{d})$ are pairwise variable disjoint polynomials. We claim that for every $h \in N$, there exists $j \in [m]$, such that for every $j' \in [m]\setminus\{j\}, \gcd(h, g_{j'}(B\mathbf{x}+\mathbf{d})) = 1$. Observe that this is true before the first iteration of the loop of lines 2-5. Subclaim 2.2.2 implies that if it is ture before an iteration of this loop, then it is also true after that iteration. For every $j \in [m]$, let $I_j \subseteq [s]$, such that for every $l \in I_j$, h_l is a factor of $g_j(B\mathbf{x}+\mathbf{d})$, which implies $h_l(A\mathbf{x})$ is a factor of $g_j(BA\mathbf{x}+\mathbf{d})$. Let $j_1, j_2 \in [m]$ be distinct and arbitrary. As $\text{var}(g_{j_1}) \cap \text{var}(g_{j_2}) = \emptyset$, we get that g_{j_1} and g_{j_2} are co-prime, which implies $I_{j_1} \cap I_{j_2} = \emptyset$. Since $h_1(A\mathbf{x}),\ldots,h_s(A\mathbf{x})$ are pairwise variable disjoint, we get that $g_1(BA\mathbf{x}+\mathbf{d}),\ldots,g_m(BA\mathbf{x}+\mathbf{d})$ are also pairwise variable disjoint polynomials. Note that $V = \{\text{var}(h_1(A\mathbf{x})),\ldots,\text{var}(h_s(A\mathbf{x}))\}$ is a required set. □

Proof: [Proof of Subclaim 2.2.2] Suppose for every $k \in [i-1]$, $N_{ess}(h_1\cdots h_{k-1}\cdot h_i) = N_{ess}(h_1)+\cdots+N_{ess}(h_{k-1})+N_{ess}(h_i)$ and $N_{ess}(h_1\cdots h_k\cdot h_i) = N_{ess}(h_1)+\cdots+N_{ess}(h_k)+N_{ess}(h_i)$. Then, by setting $k = i-1$, we get $N_{ess}(h_1\cdots h_{i-1}\cdot h_i) = N_{ess}(h_1)+\cdots+N_{ess}(h_{i-1})+N_{ess}(h_i)$, which is a contradiction. Thus, there exists $k \in [i-1]$, such that $N_{ess}(h_1\cdots h_{k-1}\cdot h_i) = N_{ess}(h_1)+\cdots+N_{ess}(h_{k-1})+N_{ess}(h_i)$ and $N_{ess}(h_1\cdots h_k\cdot h_i) < N_{ess}(h_1)+\cdots+N_{ess}(h_k)+N_{ess}(h_i)$.

Now, we want to argue that there exists $j \in [m]$, such that h_k and h_i are factors of $g_j(B\mathbf{x}+\mathbf{d})$. Suppose for every $l \in [|N|], h'_l \in \mathbb{F}[\mathbf{x}]$ is such that $h_l = h'_l(B\mathbf{x}+\mathbf{d})$. Then, it is sufficient to show that there exists $j \in [m]$, such that h'_k and h'_i are factors of g_j. Suppose this is not true and there exist $j_1, j_2 \in [m], j_1 \neq j_2$, such that h'_k and h'_i are factors of g_{j_1} and g_{j_2} respectively. For $j \in \{j_1, j_2\}$, let $I_j \subseteq [k-1]$, such that for every $l \in I_j, h'_l$ is a factor of g_j and $p_j := \prod_{l\in I_j} h'_l$. Further, let $q := \prod_{l\in[k-1]\setminus(I_{j_1}\cup I_{j_2})} h'_l$. Then, note that

$$h'_1\cdots h'_{k-1}\cdot h'_i = q\cdot p_{j_1}\cdot p_{j_2}\cdot h'_i \quad \text{and} \quad h'_1\cdots h'_{k-1}\cdot h'_k = q\cdot p_{j_1}\cdot p_{j_2}\cdot h'_k.$$

Since q, p_{j_1} and p_{j_2} are factors of $\left(\prod_{j\in[m]\setminus\{j_1,j_2\}} g_j\right)$, g_{j_1} and g_{j_2} respectively, q, p_{j_1} and p_{j_2} are pairwise variable disjoint polynomials. Also, as $N_{ess}(h_1\cdots h_{k-1}\cdot h_i) = N_{ess}(h_1)+\cdots+N_{ess}(h_{k-1})+N_{ess}(h_i)$, Fact 2.12 implies that

$$N_{ess}(h'_1\cdots h'_{k-1}\cdot h'_i) = N_{ess}(h'_1)+\cdots+N_{ess}(h'_{k-1})+N_{ess}(h'_i).$$

On using Observation 2.10, the above equation can be written as

$$N_{ess}(q\cdot p_{j_1}\cdot p_{j_2}\cdot h'_i) = N_{ess}(q)+N_{ess}(p_{j_1})+N_{ess}(p_{j_2})+N_{ess}(h'_i). \tag{2.8}$$

Similarly, we get $N_{ess}(h'_1 \cdots h'_{k-1} \cdot h'_k) = N_{ess}(h'_1) + \cdots + N_{ess}(h'_{k-1}) + N_{ess}(h'_k)$, which implies

$$N_{ess}(q \cdot p_{j_1} \cdot p_{j_2} \cdot h'_k) = N_{ess}(q) + N_{ess}(p_{j_1}) + N_{ess}(p_{j_2}) + N_{ess}(h'_k). \tag{2.9}$$

Recall that h'_i and p_{j_2} are the factors of g_{j_2}. Then, it follows from Equation (2.8) and Observation 2.10 that $N_{ess}(h'_i p_{j_2}) = N_{ess}(h'_i) + N_{ess}(p_{j_2})$. Then, Claim 2.2.2 implies that there exists $A_{j_2} \in \mathrm{GL}(n, \mathbb{F})$, which maps every variable in $\mathrm{var}(g_{j_2})$ to a linear form in $\mathrm{var}(g_{j_2})$ and maps every variable in $\mathbf{x} \setminus \mathrm{var}(g_{j_2})$ to itself, such that $h'_i(A_{j_2}\mathbf{x})$ and $p_{j_2}(A_{j_2}\mathbf{x})$ are variable disjoint. Similarly, as h'_k, p_{j_1} are factors of g_{j_1}, there exists $A_{j_1} \in \mathrm{GL}(n, \mathbb{F})$, which maps every variable in $\mathrm{var}(g_{j_1})$ to a linear form in $\mathrm{var}(g_{j_1})$ and maps every variable in $\mathbf{x} \setminus \mathrm{var}(g_{j_1})$ to itself, such that $h'_k(A_{j_1}\mathbf{x})$ and $p_{j_1}(A_{j_1}\mathbf{x})$ are variable disjoint. Hence,

$$q \cdot p_{j_1} \cdot p_{j_2} \cdot h'_k \cdot h'_i(A_{j_1}A_{j_2}\mathbf{x}) = q \cdot p_{j_1}(A_{j_1}\mathbf{x}) p_{j_2}(A_{j_2}\mathbf{x}) h'_k(A_{j_1}\mathbf{x}) h'_i(A_{j_2}\mathbf{x}).$$

The polynomials h'_k and h'_i are factors of g_{j_1} and g_{j_2} respectively, which are variable disjoint as $j_1 \neq j_2$. Thus, $q, p_{j_1}(A_{j_1}\mathbf{x}), p_{j_2}(A_{j_2}\mathbf{x})$, $h'_k(A_{j_1}\mathbf{x})$ and $h'_i(A_{j_2}\mathbf{x})$ are variable disjoint and we get

$$N_{ess}(q \cdot p_{j_1} \cdot p_{j_2} \cdot h'_k \cdot h'_i) = N_{ess}(h'_1 \cdots h'_k h'_i) = \sum_{l \in [k]} N_{ess}(h'_l) + N_{ess}(h'_i),$$

which is a contradiction. Thus, $j_1 = j_2$. □

2.2.3 PE for quadratic forms

A polynomial f is a quadratic form if it is homogeneous and $\deg(f) = 2$. Recall that PE for quadratic form (also called quadratic form equivalence and denoted QFE) is the following problem: Given two n-variate quadratic forms $f, g \in \mathbb{F}[\mathbf{x}]$, check if there exists an $A \in \mathrm{GL}(n, \mathbb{F})$ such that $f = g(A \cdot \mathbf{x})$. If the answer is yes, output $A \in \mathrm{GL}(n, \mathbb{F})$ such that $f = g(A \cdot \mathbf{x})$. QFE over $\mathbb{C}, \mathbb{R}, \mathbb{Q}$ and finite fields are well-studied and these algorithms are based on well-known results on classification of quadratic forms. We direct the interested reader to [Ser73, Lam04, Ara11] for a comprehensive discussion on these classification results. In the following fact, we state the complexity of QFE over different fields. Recall from Theorem 1.11 that the equivalence test for regular ROFs require oracle access to QFE. Then, the following fact implies that the equivalence test for regular ROF given in Chapter 5 is efficient.

Over $\mathbb{R}$ and $\mathbb{C}$, the model of computation is an arithmetic circuit with oracle access to a root finding algorithm; every arithmetic operation in this circuit takes a unit time. Whereas, over $\mathbb{Q}$ and finite fields, the model of computation is a Turing machine, where the running time is measured in terms of bit operations.

Fact 2.19 (QFE over standard fields) *Let $n \in \mathbb{N}$ be the number of variables present in each of the two quadratic forms given as input to a QFE algorithm.*

1. (Over $\mathbb{R}$ and $\mathbb{C}$). *There exists a* $\mathrm{poly}(n, d)$ *time algorithm over* $\mathbb{R}$ *and* $\mathbb{C}$.
2. (Over a finite field $\mathbb{F}_q$). *Let* $\mathbb{F}_q$ *be such that* $char(\mathbb{F}_q) \neq 2$. *There is a randomized* $\mathrm{poly}(n, \log q)$ *time QFE algorithm over* $\mathbb{F}_q$.
3. (Over $\mathbb{Q}$) [Wal13]. *There is a deterministic* $\mathrm{poly}(n, \beta)$ *time QFE algorithm over* $\mathbb{Q}$ *with oracle access to integer factoring, where* β *is the bit length of the coefficients of the input quadratic forms.*

2.2.4 Full matrix algebra isomorphism

The full matrix algebra isomorphism (FMAI for short) over a field $\mathbb{F}$ is the following algorithmic problem: given a basis B of an $\mathbb{F}$-algebra $\mathscr{A}$, where $\dim \mathscr{A} = n^2$, determine whether $\mathscr{A}$ is isomorphic as an $\mathbb{F}$-algebra to $M_n(\mathbb{F})$. If yes, output an $\mathbb{F}$-algebra from $\mathscr{A}$ to $M_n(\mathbb{F})$. In case of a yes instance, the output of an FMAI algorithm is a basis $A_1, \ldots, A_{n^2}$ of $M_n(\mathbb{F})$ such that if $B = \{B_1, \ldots, B_{n^2}\}$ then the required $\mathbb{F}$-algebra isomorphism $\varphi : \mathscr{A} \to M_n(\mathbb{F})$ is obtained as follows: for every $i \in [n^2], \varphi(B_i) = A_i$. Recall that the equivalence test for the determinant given in Chapter 4 takes oracle access to FMAI over finite fields and over $\mathbb{Q}$. In this section, we record the results on FMAI over finite fields and over $\mathbb{Q}$.

Theorem 2.2 [Theorem 5.1 of [Rón90]] *Let $\mathbb{F}$ be a finite field. Given a basis of a $\mathbb{F}$-algebra $\mathscr{A} \subseteq M_m$ such that $\mathscr{A}$ and M_n are isomorphic $\mathbb{F}$-algebras, an isomorphism $\varphi : \mathscr{A} \to M_n$ can be constructed in randomized* $\mathrm{poly}(m, \log |\mathbb{F}|)$ *time.*

Theorem 2.3 [Theorem 1 of [IRS12]] *There is a randomized algorithm with oracle access to IntFact that takes input a basis of a $\mathbb{Q}$-algebra $\mathscr{A} \subseteq M_m$ such that $\mathscr{A}$ and M_n are isomorphic $\mathbb{F}$-algebras, and outputs an isomorphism $\varphi : \mathscr{A} \to M_n$ with high probability. The algorithm runs in time polynomial in the bit length of the input, if n is bounded.*

Theorem 2.4 [Lemma 2.5 of [BR90]] *There is a randomized algorithm that takes input a basis of a $\mathbb{Q}$-algebra $\mathscr{A} \subseteq M_m$ such that $\mathscr{A}$ and M_n are isomorphic $\mathbb{F}$-algebras, and outputs an isomorphism $\varphi : \mathscr{A} \otimes_{\mathbb{Q}} L \to M_n(L)$ with high probability, where L is an extension field of $\mathbb{Q}$ satisfying $[L : \mathbb{Q}] \leq n$. The algorithm runs in time polynomial in the bit length of the input.*

Chapter 3

Structural and algorithmic results on the NW polynomial

In this chapter, we prove the theorems given in Section 1.3.1. The content of this chapter are present in [GS19], which is a joint work with Chandan Saha. There are two main sections here: In Section 3.1, we present the characterization by symmetries and characterization by circuit identities properties of NW and Section 3.2 contains three algorithmic results for NW, namely a circuit testing algorithm, a flip theorem and a BD-PS equivalence test for NW. These results are based on the symmetries of NW.

We recall the definition of the Nisan-Wigderson polynomial from Section 1.3.1. Let d be a prime, $k \in \mathbb{N}, k << d$, $\mathbf{x} = \{x_{i,j} : i, j \in \mathbb{F}_d\}$ and $\mathbb{F}_d[z]_k = \{h \in \mathbb{F}_d[z] : \deg(h) \leq k\}$. Then,

$$\mathsf{NW}_{d,k}(\mathbf{x}) = \sum_{h \in \mathbb{F}_d[z]_k} \prod_{i=0}^{d-1} x_{i,h(i)},$$

The number of variables in $\mathsf{NW}_{d,k}$ is d^2. Although, our results hold for any $k \in [1, \frac{d}{4} - 5]$, we fix $k = d^\epsilon$ for some arbitrarily chosen $\epsilon \in (0, 1)$ as in most of the lower bound proofs using $\mathsf{NW}_{d,k}$ as a hard polynomial, k is chosen to be d^ϵ. Henceforth, we would drop the subscripts from $\mathsf{NW}_{d,k}$ whenever the value of d is clear from the context. We would denote the elements of $\mathbb{F}$ and $\mathbb{F}_d$ by α, β, γ and a, b, c respectively, and polynomial in $\mathbb{F}[\mathbf{x}]$ and $\mathbb{F}_d[z]$ by f, g, q and h, p respectively. We fix $n = d^2$. For $m \in \mathbb{N}^\times, [m] = \{0, \ldots, m-1\}$. In this chapter, whenever we mention a set-multilinear polynomial in $\mathbb{F}[\mathbf{x}]$, it is always with respect to the partition $\mathbf{x} = \uplus_{i \in \mathbb{F}_d} \mathbf{x}_i$, i.e., every monomial of a set-mulitlinear polynomial has one variable from each $\mathbf{x}_i$.

In Section 3.1, we give the proofs of Theorems 1.1 and 1.2 along with a lemma that shows that NW is characterized by circuit identities (Definition 2.25) over all fields. This lemma is

immensely helpful for the circuit testing algorithm for NW (Theorem 1.3) and a flip theorem for NW (Theorem 1.4), which are given in Section 3.2. The circuit testing algorithm also has an application in the BD-PS equivalence test for NW (Theorem 1.5) given in Section 3.2. This equivalence test uses symmetries and some structural insights from the Lie algebra of NW. The structure of the Lie algebra of NW, the structure of the symmetries of NW and some *continuous* and *discrete symmetries* of NW were studied in the author's master's thesis [Gup17].

3.1 Structural results

This section is further divided into two subsections: The first one is on the *characterization by symmetries* (Definition 2.24) property of NW over different fields and the next one is about the *characterization by circuit identities* (Definition 2.25) property of NW.

3.1.1 Characterization by symmetries

We first recall Theorem 1.1 and then give its proof. Also, recall that $\mathscr{G}_f$ denotes the group of symmetries of f . The rows and columns of matrices in $\mathscr{G}_{\mathsf{NW}}$ are labelled by the ordered set $((0,0),(0,1),\ldots,(d-1,d-1))$. Also, recall that an $\alpha \in \mathbb{F}$ is called a d-th primitive root of unity if $\alpha^d = 1$ and for every $1 \leq r < d, \alpha^r \neq 1$.

Theorem 3.1 (NW characterized by its symmetries over $\mathbb{C}$) *Let d be a prime number, $\mathbb{F}$ be a field containing a d-th primitive root of unity, and f be a homogeneous degree d polynomial in d^2 variables over $\mathbb{F}$. If $\mathscr{G}_{\mathsf{NW}_{d,k}} \subseteq \mathscr{G}_f$ then $f = \alpha \cdot \mathsf{NW}_{d,k}$ for some $\alpha \in \mathbb{F}$.*

Before coming to the proof of the above theorem, we give the following useful claim.

Claim 3.1.1 (Useful symmetries of NW) *Let d be a prime number and $\mathbb{F}$ be a field mentioned in Theorem 3.1. Then, the following matrices in $M_n(\mathbb{F})$ are in $\mathscr{G}_{\mathsf{NW}}$:*

1. *A diagonal matrix $A_{\boldsymbol{\beta}}$ with $A_{\boldsymbol{\beta}}((i,j),(i,j)) = \beta_i \in \mathbb{F}^\times$ for $i,j \in [d]$ such that $\prod_{i\in\mathbb{F}_d} \beta_i = 1$.*
2. *For $h \in \mathbb{F}_d[z]_k$, A_h satisfying $A_h((i,j),(i,j+h(i))) = 1$ for $i,j \in [d]$ and other entries 0.*
3. *A diagonal matrix A_ℓ with $((i,j),(i,j))$-th entry as $\zeta^{i^\ell \cdot j}$ for $i,j \in [d]$, where $\ell \in [d-k-1]$.*

Proof: By definition, $A_{\boldsymbol{\beta}}, A_\ell \in \mathrm{GL}(n,\mathbb{F})$. Note that for every $h \in \mathbb{F}_d[z]_k$, A_h is a permutation matrix, which implies $A_h \in \mathrm{GL}(n,\mathbb{F})$. Observe that the polynomials $\mathsf{NW}(A_{\boldsymbol{\beta}}\mathbf{x}), \mathsf{NW}(A_\ell\mathbf{x})$ and $\mathsf{NW}(A_h\mathbf{x})$ are obtained from $\mathsf{NW}(\mathbf{x})$ by replacing the variable $x_{i,j}$ with $\beta_i \cdot x_{i,j}$, $\zeta^{i^\ell \cdot j} \cdot x_{i,j}$ and $x_{i,j+h(i)}$ respectively, for $i,j \in [d]$. Let $p \in \mathbb{F}_d[z]_k$ and $m_p = \prod_{i\in\mathbb{F}_d} x_{i,p(i)}$. When $A_{\boldsymbol{\beta}}$ is applied on $\mathbf{x}$, m_p gets mapped to $\prod_{i\in[d]} \beta_i \cdot m_p = m_p$ as $\prod_{i\in[d]} \beta_i = 1$, implying $A_{\boldsymbol{\beta}} \in \mathscr{G}_{\mathsf{NW}}$. When A_h

is applied on $\mathbf{x}$, m_p gets mapped to m_{p+h}; in other words, the monomials of NW are 'shifted around' and so $A_h \in \mathscr{G}_{\mathsf{NW}}$. When A_ℓ is applied on $\mathbf{x}$, m_p is mapped to $\prod_{i\in[d]} \zeta^{i^\ell \cdot p(i)} \cdot m_p$. We show below that $\prod_{i\in[d]} \zeta^{i^\ell \cdot p(i)} = 1$ for every $\ell \in [d-k-1]$, thereby implying $A_\ell \in \mathscr{G}_{\mathsf{NW}}$.

Observation 3.1 *For every $p \in \mathbb{F}_d[\mathbf{x}]_k$ and $\ell \in [d-k-1]$, $\prod_{i\in[d]} \zeta^{i^\ell \cdot p(i)} = 1$.*

Proof: As $\zeta \neq 1$ is a d-th root of unity, $\prod_{i\in[d]} \zeta^{i^\ell \cdot p(i)} = \zeta^{\sum_{i\in\mathbb{F}_d} i^\ell \cdot p(i)}$, it is sufficient to show that $\sum_{i\in\mathbb{F}_d} i^\ell \cdot p(i) = 0$. Suppose $p(z) = a_r z^r + \cdots + a_0$, where $r \leq k$ and $a_r, \ldots, a_0 \in \mathbb{F}_d$. Then

$$\sum_{i\in\mathbb{F}_d} i^\ell \cdot p(i) = a_r \left(\sum_{i\in\mathbb{F}_d} i^{r+\ell}\right) + \cdots + a_0 \left(\sum_{i\in\mathbb{F}_d} i^\ell\right).$$

Each summand in the R.H.S. of the above equation is of the form $a \cdot \left(\sum_{i\in\mathbb{F}_d} i^s\right)$, where $0 \leq s \leq d-2$. As $\sum_{i\in\mathbb{F}_d} i^0 = 0$, assume that $1 \leq s \leq d-2$. Let b be a generator of $\mathbb{F}_d^\times$. Then

$$\sum_{i\in\mathbb{F}_d} i^s = \sum_{i\in\mathbb{F}_d^\times} i^s = \sum_{t\in[d-1]} b^{t\cdot s} = \frac{1-b^{(d-1)\cdot s}}{1-b^s} = 0, \quad \text{as } b^{d-1} = 1 \text{ in } \mathbb{F}_d. \tag{3.1}$$

Hence, $\sum_{i\in\mathbb{F}_d} i^\ell \cdot p(i) = 0$ implying $\prod_{i\in[d]} \zeta^{i^\ell \cdot p(i)} = 1$. □

□

Now we are ready to prove Theorem 3.1.

Proof: Let $f \in \mathbb{F}[\mathbf{x}]$ be a non-zero homogeneous degree d polynomial such that $\mathscr{G}_{\mathsf{NW}} \subseteq \mathscr{G}_f$. If $f = 0$, there is nothing to prove. Claim 3.1.1 implies that $A_{\boldsymbol{\beta}}, A_\ell, A_h \in \mathscr{G}_f$ for all choices of $\boldsymbol{\beta}, \ell, h$ mentioned in Claim 3.1.1. The presence of $A_{\boldsymbol{\beta}}$ in $\mathscr{G}_f$ implies that f is a set-multilinear polynomial. If not then there is a term $\alpha \cdot m$ in f, where $\alpha \in \mathbb{F}^\times$ and m is a degree-d monomial with no $\mathbf{x}_t$-variables for some $t \in [d]$. As $\mathbb{F}$ contains a d-th primitive root of unity, it is easy to verify that $|\mathbb{F}| \neq d+1$. Thus, there exists a $\gamma \in \mathbb{F}^\times$ such that $\gamma^d \neq 1$. Pick such a γ. Now, set $\beta_i = \gamma$ for $i \in [d]\backslash\{t\}$ and $\beta_t = \gamma^{-(d-1)}$ so that $\prod_{i\in[d]} \beta_i = 1$ is satisfied. When $A_{\boldsymbol{\beta}}$ is applied on $\mathbf{x}$, the term $\alpha \cdot m$ maps to $\alpha\gamma^d \cdot m \neq \alpha \cdot m$, implying that $f(A_{\boldsymbol{\beta}}\mathbf{x}) \neq f(\mathbf{x})$.

As f is set-multilinear, every term of f is of the kind $\alpha_p \cdot m_p$, where $\alpha_p \in \mathbb{F}^\times$, $m_p = \prod_{i\in\mathbb{F}_d} x_{i,p(i)}$ and $p \in \mathbb{F}_d[z]$ with $\deg(p) \leq d-1$. This is because any function from $\mathbb{F}_d$ to $\mathbb{F}_d$ can be represented by a univariate polynomial of degree at most $d-1$. We now show that $\deg(p) \leq k$ for every term $\alpha_p \cdot m_p$ in f. Suppose not. Then, there is a term $\alpha_p \cdot m_p$ such that $p = a_r z^r + \cdots + a_0$, $r > k$ and $a_r \neq 0$. When A_ℓ is applied on $\mathbf{x}$, the term $\alpha_p \cdot m_p$ gets mapped to $\prod_{i\in[d]} \zeta^{i^\ell \cdot p(i)} \cdot \alpha_p \cdot m_p$. Now choose $\ell = d-r-1 \leq d-k-2$. That $\prod_{i\in[d]} \zeta^{i^\ell \cdot p(i)} \neq 1$ for this choice of ℓ can be argued

as follows: Since $\prod_{i\in[d]} \zeta^{i^\ell \cdot p(i)} = \zeta^{\sum_{i\in\mathbb{F}_d} i^\ell \cdot p(i)}$, it is sufficient to show that $\sum_{i\in\mathbb{F}_d} i^\ell \cdot p(i) \neq 0$. Expanding the sum,

$$\sum_{i\in\mathbb{F}_d} i^\ell \cdot p(i) = a_r \left(\sum_{i\in\mathbb{F}_d} i^{d-1}\right) + a_{r-1}\left(\sum_{i\in\mathbb{F}_d} i^{d-2}\right) + \cdots + a_0 \left(\sum_{i\in\mathbb{F}_d} i^{d-r-1}\right).$$

As argued in Equation (3.1), the above sum is $a_r \cdot (d-1)$. As $char(\mathbb{F}_d) = d$, $a_r \cdot (d-1) \neq 0$, which implies $f(A_\ell \mathbf{x}) \neq f(\mathbf{x})$. Hence, every term $\alpha_p \cdot m_p$ of f must have $\deg(p) \leq k$. When A_h is applied on $\mathbf{x}$, a term $\alpha_p \cdot m_p$ maps to $\alpha_p \cdot m_{p+h}$ which implies $\alpha_p = \alpha_{p+h}$. Running over all $h \in \mathbb{F}_d[z]_k$, we get $\alpha_p = \alpha$ for every $p \in \mathbb{F}_d[z]_k$, for some $\alpha \in \mathbb{F}^\times$. Hence, $f = \alpha \cdot \mathsf{NW}$. □

Now, we show that if $\mathbb{F}$ does not have a d-th primitive root of unity, the above result does not hold. In particular, we prove Theorem 1.2, which we recall below.

Theorem 3.2 (NW not characterized by its symmetries over $\mathbb{R}$) *Let d be a prime number and $\mathbb{F}$ be either $\mathbb{R}, \mathbb{Q}$ or a finite field satisfying $d \nmid |\mathbb{F}| - 1$. Then, $\mathsf{NW}_{d,k}$ is not characterized by its symmetries over $\mathbb{F}$.*

Proof: As $\mathbb{F}$ does not contain a d-th primitive root of unity, it is easy to see that the matrices A_ℓ for $\ell \in [d-k-1]$ mentioned in Claim 3.1.1 are not contained in $\mathscr{G}_{\mathsf{NW}}$ over $\mathbb{F}$. To prove this theorem, we need the following lemma, which says that every diagonal symmetry of NW over $\mathbb{F}$ is of the first kind mentioned in Claim 3.1.1. We exploit this property and show that no matter which other symmetry is present in $\mathscr{G}_{\mathsf{NW}}$ over $\mathbb{F}$, NW is not characterized by its symmetries over $\mathbb{F}$. We first complete this proof assuming Lemma 3.1, which is proved later.

Lemma 3.1 (Diagonal symmetries over $\mathbb{F}$) *Let $\mathbb{F}$ be a field mentioned in Theorem 3.2. If $D \in \mathscr{G}_{\mathsf{NW}}$ is a diagonal matrix over $\mathbb{F}$ then $D = \mathrm{diag}(\beta_0, \ldots, \beta_{d-1}) \otimes I_d$, where $\beta_i \in \mathbb{F}^\times$ for every $i \in [d]$ and $\prod_{i\in[d]} \beta_i = 1$.*

Along with the above lemma, we also need the following result about the structure of $\mathscr{G}_{\mathsf{NW}}$.

Theorem 3.3 ([Gup17]) *Let d be a prime number and $\mathbb{F}$ be a field satisfying $|\mathbb{F}| > \binom{d}{2}$ and $char(\mathbb{F}) \neq d$. If $A \in \mathscr{G}_{\mathsf{NW}}$ then $A = DP$, where $D, P \in \mathscr{G}_{\mathsf{NW}}$ are diagonal and permutation matrices respectively.*

A proof of the above theorem is given in Chapter 4 of [Gup17]. Let $P_1, \ldots, P_r$ be all the permutation matrices in $\mathscr{G}_{\mathsf{NW}}$. We now show that there exists a set-multilinear polynomial $f \in \mathbb{F}[\mathbf{x}]$ such that f is not a non-zero scalar multiple of NW but $\mathscr{G}_{\mathsf{NW}} \subseteq \mathscr{G}_f$. Let $h \in \mathbb{F}_d[z]$ has degree $k+1$ and $m_h := \prod_{i\in[d]} x_{i,h(i)}$. Let S be the smallest set of monomials containing m_h

such that for every monomial $m \in S$, $m(P_i\mathbf{x}) \in S$ for every $i \in \{1, \ldots, r\}$. Clearly, S is a set of set-multilinear monomials. Suppose $f \in \mathbb{F}[\mathbf{x}]$ is defined as follows

$$f = \sum_{m \in S} m. \tag{3.2}$$

It follows from Lemma 3.1 that if $D \in \mathscr{G}_{\mathsf{NW}}$ is a diagonal matrix and $g \in \mathbb{F}[\mathbf{x}]$ is a set-multilinear polynomial then $D \in \mathscr{G}_g$. As f is set-multilinear, it is easy to see that all the diagonal symmetries of NW are contained in $\mathscr{G}_f$. By definition, all the permutation symmetries of NW are also contained in $\mathscr{G}_f$. Thus, $\mathscr{G}_{\mathsf{NW}} \subseteq \mathscr{G}_f$ but f is not a scalar multiple of NW. □

Proof of Lemma 3.1

Before proving Lemma 3.1, we give some useful results, one of which is the following claim. This result forms the core of the analysis of the structure of the Lie algebra of NW and has been proven in Chapter 3 of [Gup17].

Claim 3.1.2 ([Gup17]) *Let $\mathbb{F}$ be a field such that $char(\mathbb{F}) \neq d$. Consider the linear system over $\mathbb{F}$ obtained from the equations $\sum_{i \in [d]} y_{i,h(i)} = 0$ for all $h \in \mathbb{F}_d[z]_k$, where $\{y_{i,j} : i, j \in \mathbb{F}_d\}$ are the variables. The solution space of the system consists of the solutions $y_{i,0} = y_{i,1} = \ldots = y_{i,d-1} = \alpha_i$ for every $i \in \mathbb{F}_d$, where $\alpha_0, \ldots, \alpha_{d-1} \in \mathbb{F}$ satisfy $\sum_{i \in [d]} \alpha_i = 0$, and these are the only solutions.*

Some useful matrices. Let $D \in M_{d^2}(\mathbb{F})$ such that the rows of D are labelled with $\{(a, b) : a, b \in \mathbb{F}_d\}$, where (a, b) is interpreted as $bz + a \in \mathbb{F}_d[z]$ and the columns of D are labelled by $\{(l, r) : l, r \in \mathbb{F}_d\}$, where (l, r) corresponds to $x_{l,r}$. Then, the $((a, b), (l, r))$-th entry of D is 1 if $x_{l,r}$ is present in $\prod_{i \in \mathbb{F}_d} x_{i,bi+a}$, else it is 0. The dimension of the Lie algebra of NW was studied in [Gup17] by analysing the rank of D. In particular, we showed that over a field $\mathbb{F}$ satisfying $char(\mathbb{F}) \neq d$, the $d^2 - d + 1$ rows of D labelled by $\{(a, b) : a \in [d-1], b \in [d]\}$ are $\mathbb{F}$-linearly independent. Now, we define two useful matrices using D, which would be needed for the proof of Lemma 3.1. Let $u := d^2 - d + 1$ and $\mathbf{x}' = \mathbf{x} \setminus \{x_{1,0}, \ldots, x_{d-1,0}\}$. Let B be the $u \times d^2$ size matrix obtained by restricting D to the rows indexed by $\{bz + a : b \in [d], a \in [d-1]\} \cup \{d-1\}$ and C be the $u \times u$ matrix obtained by restricting B to the columns labelled by $\mathbf{x}'$.

Claim 3.1.3 *The absolute value of the determinant of C over $\mathbb{Z}$ is d^r, where $r = O(d^2)$.*

Proof: We know that

$$B\mathbf{x} = \left(\sum_{i \in \mathbb{F}_d} x_{i,0} \cdots \sum_{i \in \mathbb{F}_d} x_{i,(d-1)i} \cdots \sum_{i \in \mathbb{F}_d} x_{i,d-2} \cdots \sum_{i \in \mathbb{F}_d} x_{i,(d-1)i+(d-2)} \sum_{i \in \mathbb{F}_d} x_{i,(d-1)} \right)^T,$$

which gives the following set of linear polynomials in $\mathbf{x}$.

$$S_1 = \left\{ \sum_{i \in \mathbb{F}_d} x_{i,h(i)} \ : \ h \in \{bz + a \ : \ b \in [d], a \in [d-1]\} \cup \{d-1\} \right\}.$$

Let S_2 be the set of $d^2 - d + 1$ distinct linear polynomials in $\mathbf{x}$ defined as

$$S_2 = \{x_{i,j} - x_{i,0} \ : \ i \in [d], j \in [d] \setminus \{0\}\} \cup \left\{ \sum_{i \in \mathbb{F}_d} x_{i,0} \right\}.$$

Consider the following easy to prove fact.

Fact 3.1 *Suppose S_1, S_2 are two sets of linear polynomials in n variables over $\mathbb{F}$ having same solution spaces. Then, $\langle S_1 \rangle = \langle S_2 \rangle$ over $\mathbb{F}$.*

Then, from Claim 3.1.2 and Fact 3.1, we get

$$\langle S_1 \rangle = \langle S_2 \rangle. \tag{3.3}$$

Let A be the $u \times d^2$ coefficient matrix of the polynomials in S_2. Then, Equation (3.3) implies that there is an $M \in \mathrm{GL}(u, \mathbb{F})$ such that

$$M \cdot B = A. \tag{3.4}$$

Let A_1 be the $u \times u$ matrix obtained by restricting A to the columns indexed by $\mathbf{x}'$. It is not difficult to see from the structures of polynomials in S_2 that A_1 is invertible. Hence, Equation (3.4) implies that $M \cdot C = A_1$ and so C is also invertible.

We claim that $\det_{\mathbb{Z}}(C) = d^r$ for some $r \in \mathbb{N}$. Suppose not. Then there exists a prime number $d' \neq d$ such that d' divides $\det_{\mathbb{Z}}(C)$. Then, the determinant of C is 0 over the finite field $\mathbb{F}_{d'}$, which is a contradiction as $char(\mathbb{F}_{d'}) \neq d$. Since C is a 0/1 matrix, $|\det_{\mathbb{Z}}(C)| \leq (d^2 - d + 1)!$, which implies $r = O(d^2)$. □

Now we generalize Claim 3.1.2 over the rings containing the inverse of d.

Claim 3.1.4 *Let d be a prime number, $\mathscr{R}$ be a ring with multiplicative identity such that d is invertible in $\mathscr{R}$. Consider the linear system over $\mathscr{R}$ obtained from the equations $\sum_{i \in \mathbb{F}_d} y_{i,h(i)} = 0$ for all $h \in \mathbb{F}_d[z]_k$, where $y_{i,j} : i, j \in [d]$ are variables. The solution space of the system consists of the solutions $y_{i,0} = y_{i,1} = \ldots = y_{i,d-1} = \alpha_i$ for every $i \in [d]$, where $\alpha_0, \ldots, \alpha_{d-1} \in \mathscr{R}$ satisfy $\sum_{i \in [d]} \alpha_i = 0$, and these are the only solutions.*

Proof: Recall matrices B and C. Observe that $B \cdot \mathbf{y} = 0$ implies the following

$$C \cdot \mathbf{y}' = \mathbf{v},$$

where $\mathbf{y}' = \mathbf{y} \setminus \{y_{1,0}, \ldots, y_{d-1,0}\}$ and the entries of $\mathbf{v}$ are linear forms in $y_{1,0}, \ldots, y_{d-1,0}$. Let $\mathrm{Adj}(C)$ be the adjoint of C. (Observe that entries of $\mathrm{Adj}(C)$ are integers and are well-defined in $\mathscr{R}$.) On multiplying the above equation with $\mathrm{Adj}(C)$, we get

$$\mathrm{Adj}(C) \cdot C \cdot \mathbf{y}' = \mathrm{Adj}(C) \cdot \mathbf{v},$$

which implies

$$\det(C) \cdot \mathbf{y}' = \mathbf{v}', \tag{3.5}$$

where $\mathbf{v}' = \mathrm{Adj}(C) \cdot \mathbf{v}$. Clearly, every entry of $\mathbf{v}'$ is a linear form in $y_{1,0}, \ldots, y_{d-1,0}$. This equation holds over any commutative ring $\mathscr{R}$ with multiplicative identity. In particular, it also holds over a field $\mathbb{F}$ such that $char(\mathbb{F}) \neq d$. From Claim 3.1.2 we know $y_{i,0} = y_{i,1} = \ldots = y_{i,d-1}$ for every $i \in [d]$ and $\sum_{i\in[d]} y_{i,0} = 0$. Thus, for $i \in \{1, \ldots, d-1\}, j \in [d] \setminus \{0\}$ the entry of $\mathbf{v}'$ indexed by $y_{i,j}$ must be $\det(C) \cdot y_{i,0}$, and for $j \in [d]$ the entry indexed by $y_{0,j}$ must be $\det(C) \cdot (-(\sum_{i=1}^{d-1} y_{i,0}))$. From Claim 3.1.3, we know that $\det(C) = d^r$. As d is invertible in $\mathscr{R}$, on multiplying Equation (3.5) with $(\det(C))^{-1}$, we get the result. □

Now, we are ready to prove Lemma 3.1.

Proof: Let $\mathbb{F} = \mathbb{R}$. Let $D \in \mathscr{G}_{\mathsf{NW}}$ be a diagonal matrix, and the $((i,j),(i,j))$-th entry of D be $\beta_{i,j} \in \mathbb{R}$ for $i, j \in [d]$. We can express $\beta_{i,j}$ as $\beta_{i,j} = (-1)^{\lambda_{i,j}} \cdot 2^{\gamma_{i,j}}$, where $\lambda_{i,j} \in \{0,1\}$ and $\gamma_{i,j} \in \mathbb{R}$. On applying D to $\mathbf{x}$, a monomial $m_h = \prod_{i\in\mathbb{F}_d} x_{i,h(i)}$ of NW maps to $\left(\prod_{i\in\mathbb{F}_d} (-1)^{\lambda_{i,h(i)}} \cdot 2^{\gamma_{i,h(i)}}\right) \cdot m_h$, implying $\prod_{i\in\mathbb{F}_d} (-1)^{\lambda_{i,h(i)}} = \prod_{i\in\mathbb{F}_d} 2^{\gamma_{i,h(i)}} = 1$. In other words,

$$\begin{aligned}
\sum_{i\in[d]} \lambda_{i,h(i)} &= 0 \quad \text{over } \mathbb{F}_2, \text{ for all } h \in \mathbb{F}_d[z]_k, \text{ and} \\
\sum_{i\in[d]} \gamma_{i,h(i)} &= 0 \quad \text{over } \mathbb{R}, \text{ for all } h \in \mathbb{F}_d[z]_k.
\end{aligned}$$

By invoking Claim 3.1.2 (over $\mathbb{F} = \mathbb{F}_2$ and over $\mathbb{F} = \mathbb{R}$) for the above two linear systems, we get $\lambda_{i,0} = \cdots = \lambda_{i,d-1} = \lambda_i$ and $\gamma_{i,0} = \cdots = \gamma_{i,d-1} = \gamma_i$ for every $i \in [d]$, where $\lambda_i \in \mathbb{F}_2, i \in [d]$ (similarly, $\gamma_i \in \mathbb{R}, i \in [d]$) satisfy $\sum_{i\in[d]} \lambda_i = 0$ in $\mathbb{F}_2$ (similarly, $\sum_{i\in[d]} \gamma_i = 0$ in $\mathbb{R}$). This implies $\beta_{i,0} = \cdots = \beta_{i,d-1} = \beta_i$ for every $i \in [d]$, where $\beta_0, \ldots, \beta_{d-1} \in \mathbb{R}$ satisfy $\prod_{i\in[d]} \beta_i = 1$. As $\mathbb{Q}$ is a sub-field of $\mathbb{R}$, NW can not have a diagonal symmetry other that $\mathrm{diag}(\beta_0, \ldots, \beta_{d-1}) \otimes I_d$ over $\mathbb{Q}$.

Let $\mathbb{F}$ be a finite field not containing a d-th primitive root of unity. Then, $d \nmid |\mathbb{F}| - 1$. Let $\mathscr{R} = \mathbb{Z}_{|\mathbb{F}|-1}$. We first argue that d is invertible over $\mathscr{R}$. It follows from the pigeonhole principle that to show d is invertible in $\mathscr{R}$, it is sufficient to show that the map $\varphi_d : \mathscr{R} \to \mathscr{R}; a \mapsto da$ is injective. Let $a_1, a_2 \in \mathscr{R}$ such that $da_1 = da_2$, which implies $d(a_1 - a_2) = 0$ in $\mathscr{R}$. We also have $(|\mathbb{F}| - 1) \cdot 1 = 0$ in $\mathscr{R}$. If $a_1 \neq a_2$ then we get a contradiction as $d \nmid |\mathbb{F}| - 1$.

Let $D = \mathrm{diag}(\beta_{0,0}, \ldots, \beta_{d-1,d-1})$, where $\beta_{i,j} \in \mathbb{F}$ for every $i, j \in [d]$. Then for every $i, j \in [d]$, $\beta_{i,j}$ can be written as $\beta_{i,j} = \tau^{\delta_{i,j}}$, where τ is a generator of $\mathbb{F}^\times$. When D is applied to $\mathbf{x}$, a monomial $m_h = \prod_{i \in \mathbb{F}_d} x_{i,h(i)}$ of NW gets mapped to $(\prod_{i \in \mathbb{F}_d} \tau^{\delta_{i,j}}) \cdot m_h$. As $D \in \mathscr{G}_{\mathsf{NW}}$, $\prod_{i \in \mathbb{F}_d} \tau^{\delta_{i,j}} = 1$, which implies

$$\sum_{i \in [d]} \delta_{i,h(i)} = 0 \quad \text{over } \mathscr{R}, \quad \text{for all } h \in \mathbb{F}_d[z]_k.$$

By invoking Claim 3.1.4 over $\mathscr{R}$ for the above system, we get the desired result. □

3.1.2 Characterization by circuit identities

In this section, we prove the following lemma. Recall Definition 2.25 from Chapter 2.

Lemma 3.2 *The polynomial* NW *is characterized by circuit identities over any field* $\mathbb{F}$.

Proof: Recall, $n = d^2$. We show that if an n-variate polynomial $f \in \mathbb{F}[\mathbf{x}]$ satisfies the following polynomial identities then $f = \alpha \cdot \mathsf{NW}$ for some $\alpha \in \mathbb{F}$. The rows and columns of the $n \times n$ matrices in the identities below are indexed by the ordered set $((0,0), (0,1), \ldots, (d-1, d-1))$.

1. $q_1(f(A_i(u)\mathbf{x}), f(\mathbf{x}), u) = 0$, for $i \in [d]$, where $q_1(z_1, z_2, u) := z_1 - u \cdot z_2$. Here, $A_i(u) \in \mathbb{F}[u]^{n \times n}$ is a diagonal matrix with the $((i,j),(i,j))$-th entry as u, for every $j \in [d]$, and the other diagonal entries as 1.

2. $q_2(f(A_{a,r}\mathbf{x}), f(\mathbf{x})) = 0$, for $a \in \mathbb{F}_d^\times$ and $r \in [k+1]$, where $q_2(z_1, z_2) := z_1 - z_2$. Here, $A_{a,r} \in \mathbb{F}^{n \times n}$ with the $((i,j),(i, j + a \cdot i^r))$-th entry as 1, for every $i, j \in \mathbb{F}_d$, and the other entries as 0.

3. $q_3(f(A_t\mathbf{x})) = 0$, for $t \in [d] \backslash [k+1]$, where $q_3(z) := z$. Here, $A_t \in \mathbb{F}^{n \times n}$ is a diagonal matrix with the $((t,0),(t,0))$-th and the $((i,j),(i,j))$-th entries as 0, for every $i \in [k+1], j \in [d] \backslash \{0\}$, and the remaining diagonal entries as 1.

Observe that there are poly(n) many identities above: d many under item 1, $(d-1)(k+1)$ many under item 2, and $(d-k-1)$ many under item 3. Also, it is clear that every q_i is computable by a constant size circuit, and the matrices $A_i(u), A_{a,r}$ and A_t are computable by poly(n) size

circuits. The identities under item 1 imply that f is a set-multilinear, homogeneous, degree-d polynomial. If not then f contains a term $\beta \cdot m$, where the degree of the $\mathbf{x}_i$-variables in m is $e \neq 1$ for some $i \in [d]$. On applying $A_i(u)$ to $\mathbf{x}$, the term $\beta \cdot m$ gets mapped to $u^e \beta \cdot m \neq u\beta \cdot m$, implying $f(A_i(u) \cdot \mathbf{x}) \neq u \cdot f(\mathbf{x})$, i.e., $q_1(f(A_i(u) \cdot \mathbf{x}), f(\mathbf{x}), u) \neq 0$.

As f is set-multilinear and homogeneous, every term of f looks like $\alpha_p \cdot m_p$, where $\alpha_p \in \mathbb{F}^\times$ and $m_p = \prod_{i \in \mathbb{F}_d} x_{i,p(i)}$ for some $p \in \mathbb{F}_d[z]$ with $\deg(p) \leq d-1$. When $A_{a,r}$ is applied on $\mathbf{x}$, for some $a \in \mathbb{F}_d^\times$ and $r \in [k+1]$, a term $\alpha_p \cdot m_p$ maps to $\alpha_p \cdot m_{p+h}$, where $h = az^r \in \mathbb{F}_d[z]_k$. Since, f satisfies the identities in item 2, $f(A_{a,r} \cdot \mathbf{x}) = f(\mathbf{x})$ and so $\alpha_p \cdot m_{p+h}$ is also a term in f. By varying $a \in \mathbb{F}_d^\times$ and $r \in [k+1]$, we see that f contains the term $\alpha_p \cdot m_{p+h}$ for every $h \in \mathbb{F}_d[z]_k$. Thus, there is a set $S_1 \subseteq \mathbb{F}_d[z]_{d-1}$ such that f is of the form,

$$f = \sum_{p \in S_1} \alpha_p \cdot \sum_{h \in \mathbb{F}_d[z]_k} m_{p+h}. \tag{3.6}$$

If $f \neq \alpha \cdot \mathsf{NW}$ for all $\alpha \in \mathbb{F}$, then there is a $p \in \mathbb{F}_d[z]$ with $\deg(p) > k$ such that f contains a term $\alpha_p \cdot m_p$ for some $\alpha_p \in \mathbb{F}^\times$. Let $h \in \mathbb{F}_d[z]_k$ such that $h(i) = -p(i)$ for all $i \in [k+1]$. From Equation (3.6), f contains the term $\alpha_p \cdot m_{p+h}$. As $\deg(p) > k$, $h(z) \neq -p(z)$. So, there is a $t \in [d] \backslash [k+1]$ such that $p(t) + h(t) \neq 0$. On applying A_t to $\mathbf{x}$, only those terms of f survive that contain the variables $x_{0,0}, \ldots, x_{k,0}$ but do not contain $x_{t,0}$, and $\alpha_p \cdot m_{p+h}$ is such a term. Hence, $q_3(f(A_t \cdot \mathbf{x})) = f(A_t \cdot \mathbf{x}) \neq 0$. This contradicts f satisfying the identities in item 3. Therefore, $f = \alpha \cdot \mathsf{NW}$, for some $\alpha \in \mathbb{F}$. On the other hand, any $f = \alpha \cdot \mathsf{NW}$ satisfies all the identities.

□

3.2 Algorithmic results

Now, we present three algorithmic results for NW, namely a circuit testing algorithm, a flip theorem and a BD-PS equivalence test in three subsections. In the following two sections, whenever we say a size-s arithmetic circuit, we would mean an arithmetic circuit C of size s, where the degree of the polynomial computed by C (also called the degree of C) is bounded by $\delta(s)$, where $\delta : \mathbb{N} \to \mathbb{N}$ is a polynomial function.

3.2.1 Circuit testability

We recall Theorem 1.3 from Section 1.3.1.

Theorem 3.4 (Circuit testing) *There is a randomized algorithm that takes input black-box access to a size-s arithmetic circuit C over a finite field $\mathbb{F}$, where $|\mathbb{F}| \geq 4 \cdot \delta(s)$, and determines whether or not $\mathsf{C} = \mathsf{NW}$ with probability $1 - \exp(-s)$, using $\mathrm{poly}(s)$ field operations.*

Proof: Consider the following algorithm.

Algorithm 5 Circuit-testing(C)

Input: Black-box access to a circuit C of size s over $\mathbb{F}$.

Output: 'True' if $\mathsf{C}(\mathbf{x}) = \mathsf{NW}$, else 'False'.

1. Pick $\mathbf{a} \in_r \mathbb{F}^n$ and $\mu \in_r \mathbb{F}$.
2. **for** $i \in [d], a \in \mathbb{F}_d^{\times}, r \in [k+1], t \in [d] \backslash [k+1]$ **do**
3. **if** $(\mathsf{C}(A_i(\mu) \cdot \mathbf{a}) - \mu \cdot \mathsf{C}(\mathbf{a}) \neq 0)$ or $(\mathsf{C}(A_{a,r} \cdot \mathbf{a}) - \mathsf{C}(\mathbf{a}) \neq 0)$ or $(\mathsf{C}(A_t \cdot \mathbf{a}) \neq 0)$ **then**
4. Return 'False'.
5. **end if**
6. **end for**
7. Let $\mathbf{b} \in \mathbb{F}^n$ be an assignment obtained by setting $x_{i0} = 1$, for $i \in [d]$, and all other variables to zero. If $f(\mathbf{b}) \neq 1$, return 'False'. Else, return 'True'.

Let C be a given circuit of size s over $\mathbb{F}$ that computes an n-variate polynomial $f = \mathsf{C}(\mathbf{x})$. Naturally, $\deg(f) \leq \delta(s)$. Algorithm 5 intends to check, in the for loop 2-6, if f satisfies the identities given in the proof of Lemma 3.2. If $f \neq \alpha \cdot \mathsf{NW}$ for all $\alpha \in \mathbb{F}$, then at least one of the identities is not satisfied. For the polynomials q_1, q_2 and q_3 defined in the proof of Lemma 3.2, observe that the degree of $q_1(f(A_i(u) \cdot \mathbf{x}), f(\mathbf{x}), u)$ is bounded by $2 \cdot \delta(s)$, whereas the degrees of $q_2(f(A_{a,r}\mathbf{x}), f(\mathbf{x}))$ and $q_3(f(A_t\mathbf{x}))$ are at most $\delta(s)$. As $|\mathbb{F}| \geq 4 \cdot \delta(s)$, by Schwartz-Zippel lemma [Zip79, Sch80], step 4 returns 'False' with probability at least $\frac{1}{2}$. If $f = \alpha \cdot \mathsf{NW}$ for some $\alpha \in \mathbb{F}$ then all the identities are satisfied, and step 7 ensures that $\alpha = 1$. Clearly, the algorithm uses $\text{poly}(s)$ field operations. The success probability is boosted from $\frac{1}{2}$ to $1 - \exp(-s)$ by repeating the algorithm $\text{poly}(s)$ times. □

3.2.2 A flip theorem

Theorem 3.5 (Flip theorem) *Suppose* NW *is not computable by size-*s *arithmetic circuits over* $\mathbb{F}$*, where* $|\mathbb{F}| \geq 4 \cdot \delta(s)$*. Then, there exist* $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{F}^n$*, where* $m = \text{poly}(s)$ *such that for every size-*s *arithmetic circuit* C*, there is an* $\ell \in [m]$ *satisfying* $\mathsf{C}(\mathbf{a}_\ell) \neq \mathsf{NW}(\mathbf{a}_\ell)$*. A set of randomly generated points* $\mathbf{a}_1, \ldots, \mathbf{a}_m \in_r \mathbb{F}^n$ *has this property with probability* $1 - \exp(-s)$*. Moreover, black-box derandomization of PIT for size-*$(10s)$ *circuits over* $\mathbb{F}$ *using* $\text{poly}(s)$ *field operations implies* $\mathbf{a}_1, \ldots, \mathbf{a}_m$ *can be computed deterministically using* $\text{poly}(s)$ *field operations.*

Proof: Let C be a circuit of size s over a finite field $\mathbb{F}$. As NW is not computable by size-s circuits over $\mathbb{F}$ (by assumption), $\mathsf{C}(\mathbf{x}) - \mathsf{NW} \neq 0$. The polynomial $\mathsf{C}(\mathbf{x}) - \mathsf{NW}$ has degree

bounded by $\delta(s)$, as $\delta(s) \geq d$. By Schwartz-Zippel lemma, for any $m \in \mathbb{N}$,

$$\Pr_{\mathbf{a}_1,\ldots,\mathbf{a}_m \in_r \mathbb{F}^n} [\mathsf{C}(\mathbf{a}_\ell) = \mathsf{NW}(\mathbf{a}_\ell), \text{ for all } \ell \in [m]] \leq \left(\frac{\delta(s)}{|\mathbb{F}|}\right)^m.$$

The number of size-s circuits over $\mathbb{F}$ is at most $2^{s^2+s} \cdot |\mathbb{F}|^s$ (as there are 2^s ways to label the nodes as $+$ and $\times$ gates, at most 2^{s^2} ways to choose the adjacency matrix of the underlying directed graph, and $|\mathbb{F}|^s$ ways to label the edges of a given graph). Therefore,

$$\Pr_{\mathbf{a}_1,\ldots,\mathbf{a}_m \in_r \mathbb{F}^n} [\exists \text{ a size-}s \text{ circuit } \mathsf{C} \text{ such that } \mathsf{C}(\mathbf{a}_\ell) = \mathsf{NW}(\mathbf{a}_\ell), \text{ for all } \ell \in [m]] \leq |\mathbb{F}|^s \cdot 2^{s^2+s} \cdot \left(\frac{\delta(s)}{|\mathbb{F}|}\right)^m.$$

By fixing $m = s^2 + 2s$, the above probability can be upper bounded by $\exp(-s)$ as $|\mathbb{F}| \geq 4 \cdot \delta(s)$.

Now, let us show that derandomization of black-box PIT implies $\mathbf{a}_1, \ldots, \mathbf{a}_m$ can be computed deterministically. Consider the class $\mathscr{C}$ of size-$(10s)$ circuits over $\mathbb{F}$ on $n+1$ variables $\mathbf{x} \uplus u$. Assume that $\mathscr{H} = \{(\mathbf{b}_0, \mu_0), \ldots, (\mathbf{b}_{w-1}, \mu_{w-1})\} \subseteq \mathbb{F}^{n+1}$ is a *hitting set*[1] for $\mathscr{C}$, and $\mathscr{H}$ is computable using $\text{poly}(s)$ field operations. Let $\mathscr{P} \subseteq \mathbb{F}^n$ be the set of points that includes $\mathbf{b}_0, \ldots, \mathbf{b}_{w-1}$ along with $A_i(\mu_\ell) \cdot \mathbf{b}_\ell$, $A_{a,r} \cdot \mathbf{b}_\ell$ and $A_t \cdot \mathbf{b}_\ell$, where $A_i, A_{a,r}, A_t$ are the matrices considered in Lemma 3.2 for every $\ell \in [w], i \in [d], a \in \mathbb{F}_d^\times, r \in [k+1]$ and $t \in [d] \backslash [k+1]$. Finally, $\mathscr{P}$ also contains the point $\mathbf{b} \in \mathbb{F}^n$ obtained by setting $x_{i,0} = 1$, for $i \in [d]$, and all other variables to zero. Observe that $|\mathscr{P}| = \text{poly}(s)$ as $|\mathscr{H}| = \text{poly}(s)$.

Claim 3.2.1 *Let C be a size-s arithmetic circuit over $\mathbb{F}$. Then, there exists an $\mathbf{a}$ in $\mathscr{P}$ such that $\mathsf{C}(\mathbf{a}) \neq \mathsf{NW}(\mathbf{a})$.*

Proof: As NW is not computable by size-s circuits, $f = \mathsf{C}(\mathbf{x}) \neq \alpha \cdot \mathsf{NW}$ for all $\alpha \in \mathbb{F}^\times$ [2]. Hence, at least one of the identities, in the proof of Lemma 3.2, is not satisfied by f unless $f = 0$. If $f = 0$ then $f(\mathbf{b}) \neq \mathsf{NW}(\mathbf{b}) = 1$, and so let $f \neq 0$. The degrees of the polynomials $q_1(f(A_i(u)\mathbf{x}), f(\mathbf{x}), u)$, $q_2(f(A_{a,r}\mathbf{x}), f(\mathbf{x}))$ and $q_3(f(A_t\mathbf{x}))$ mentioned in Lemma 3.2 are upper bounded by $2 \cdot \delta(s)$. Also, it can be verified that the polynomials $q_1(f(A_i(u)\mathbf{x}), f(\mathbf{x}), u)$, $q_2(f(A_{a,r}\mathbf{x}), f(\mathbf{x}))$ and $q_3(f(A_t\mathbf{x}))$ are computable by size-$(10s)$ circuits on $n+1$ variables $\mathbf{x} \uplus u$. Hence, $\mathscr{H}$ is a hitting-set for these polynomials. Without loss of generality, let $q_1(f(A_i(u)\mathbf{x}), f(\mathbf{x}), u) = 0$ be an identity that is not satisfied by f. Then, there is a $(\mathbf{b}_\ell, \mu_\ell) \in \mathscr{H}$ such that $q_1(f(A_i(\mu_\ell)\mathbf{b}_\ell), f(\mathbf{b}_\ell), \mu_\ell) \neq 0$ implying $f(A_i(\mu_\ell)\mathbf{b}_\ell) \neq \mu_\ell \cdot f(\mathbf{b}_\ell)$. On the other hand,

[1] A set of points $\mathscr{H}$ is a hitting-set for a circuit class $\mathscr{C}$ if for every circuit $\mathsf{C} \in \mathscr{C}$ computing a non-zero polynomial, there exists a point $\mathbf{b} \in \mathscr{H}$ such that $\mathsf{C}(\mathbf{b}) \neq 0$. Black-box derandomization of identity testing for a circuit class amounts to constructing a hitting-set for the class.

[2] If $\alpha \cdot \mathsf{NW}$ is computable by a size-s circuit C, for some $\alpha \in \mathbb{F}^\times$, then NW is also computable by a size-s circuit by appropriately scaling some of the edges feeding into the output gate of C by α^{-1}.

$\mathsf{NW}(A_i(\mu_\ell)\mathbf{b}_\ell) = \mu_\ell \cdot \mathsf{NW}(\mathbf{b}_\ell)$ as NW satisfies all the identities. Therefore, either $f(A_i(\mu_\ell)\mathbf{b}_\ell) \neq \mathsf{NW}(A_i(\mu_\ell)\mathbf{b}_\ell)$ or $f(\mathbf{b}_\ell) \neq \mathsf{NW}(\mathbf{b}_\ell)$. This implies the claim as $A_i(\mu_\ell)\cdot\mathbf{b}_\ell$ and $\mathbf{b}_\ell$ belong to $\mathscr{P}$. □

The proof of the theorem follows from the above claim and by observing that $\mathscr{P}$ can be constructed from $\mathscr{H}$ using poly(s) field operations. □

3.2.3 Equivalence test for NW

First, we show a randomized polynomial time reduction of equivalence test for NW to block-permuted ET (in short, BP ET) in Lemma 3.3. Recall from Section 1.4.1 that a BP-equivalence test for NW checks if there exists an invertible block-permuted matrix A such that the given polynomial f satisfies $f = \mathsf{NW}(A\mathbf{x})$. A $d^2 \times d^2$ matrix A is said to be block-permuted with block size d if there exists a $d^2 \times d^2$ block-diagonal matrix B with block size d and a $d \times d$ permutation matrix P such that $A = B \cdot (P \otimes I_d)$. For the following lemma, we assume that univariate polynomial factorization over $\mathbb{F}$ can be done in polynomial time.

Lemma 3.3 (Reduction to BP ET) *Let $\mathbb{F}$ be such that char$(\mathbb{F}) \neq d$ and $|\mathbb{F}| \geq 2d^2$. There is a randomized algorithm that takes input black-box access to a degree d polynomial $f \in \mathbb{F}[\mathbf{x}]$ and does the following with high probability: It outputs black-box access to a degree d polynomial $g \in \mathbb{F}[\mathbf{x}]$ such that f is equivalent to NW if and only if g is BP equivalent to NW. Moreover, the transformation for f can be recovered efficiently from the transformation for g. This algorithm uses* poly(d) *many field operations.*

Proof:

Algorithm 6 Reduce-ET-to-BP-ET(f)

Input: Black-box access to $f \in \mathbb{F}[\mathbf{x}]$.

Output: If f is equivalent to NW then black-box access to a polynomial which is BP equivalent to NW, otherwise 'Fail'.

1. Compute a basis $L_1, \ldots, L_r$ of $\mathfrak{g}_f$ using Fact 2.16. If $r \neq d-1$, output 'Fail'.
2. Let $S \subseteq \mathbb{F}, |\mathbb{F}| = d^2$. Let $L = a_1L_1 + \cdots + a_rL_r$, where $a_i \in_r S$. Compute $D \in \mathrm{GL}(d^2, \mathbb{F})$ s.t. $D^{-1} \cdot L \cdot D = \mathrm{diag}(\beta_1, \ldots, \beta_d) \otimes I_d$, where $\beta_j \in \mathbb{F}$. If no such D exists, output 'Fail.'
3. Output black-box access to $f(D\mathbf{x})$.

For arguing the correctness of the above algorithm, we need the following lemma from [Gup17]. See Chapter 3 of [Gup17] for a proof of this lemma.

Lemma 3.4 [Gup17] *Let d be a prime number, $k \geq 1$ and $\mathbb{F}$ be a field having characteristic not equal to d. Then,* $\dim \mathfrak{g}_{\mathsf{NW}} = d-1$ *over $\mathbb{F}$ and the diagonal matrices $B_1, \ldots, B_{d-1}$ (defined below) form an $\mathbb{F}$-basis of $\mathfrak{g}_{\mathsf{NW}}$. For $\ell \in \{1, \ldots, d-1\}$,*

$$(B_\ell)_{(i,j),(i,j)} = \begin{cases} 1, & \text{if } i = 0, j \in [d] \\ -1, & \text{if } i = \ell, j \in [d] \\ 0, & \text{otherwise.} \end{cases}$$

Proof of correctness. It follows from Lemma 3.4 and Fact 2.10 that if f is equivalent to NW then $\dim \mathfrak{g}_f = d-1$. The correctness of Steps 2 and 3 follows from the next claim.

Claim 3.2.2 *With high probability, matrix D can be computed using* $\mathrm{poly}(d)$ *field operations. Moreover, f is equivalent to NW if and only if $f(D\mathbf{x})$ is BP equivalent to NW.*

□

Analysis of the running time. Fact 2.16 implies that a basis of $\mathfrak{g}_f$ can be computed in randomized $\mathrm{poly}(d, \rho)$ time. We can compute a D mentioned in the algorithm by solving a system of linear equations in the entries of D originating from $D^{-1} \cdot L \cdot D = \mathrm{diag}(\beta_1, \ldots, \beta_d) \otimes I_d$.

Proof of Claim 3.2.2. Suppose $f = \mathsf{NW}(A\mathbf{x})$ for $A \in \mathrm{GL}(d^2, \mathbb{F})$. Then, $R_1, \ldots, R_{d-1}$ is a basis of $\mathfrak{g}_{\mathsf{NW}}$, where $L_i = A^{-1} \cdot R_i \cdot A$ (Fact 2.10). We know that $L = a_1 L_1 + \cdots + a_{d-1} L_{d-1}$, where $a_1, \ldots, a_{d-1}$ are chosen uniformly at random from S. Pretend that $\mathbf{a} = \{a_1, \ldots, a_{d-1}\}$ are formal variables. Then, $L = A^{-1} \cdot R \cdot A$, where $R = a_1 R_1 + \cdots + a_{d-1} R_{d-1}$. Lemma 3.4 implies that $R = \mathrm{diag}(\alpha_1, \ldots, \alpha_d) \otimes I_d$, where $\alpha_1, \ldots, \alpha_d$ are linear forms in $\mathbf{a}$-variables, and $\alpha_d = -(\sum_{i=1}^{d-1} \alpha_i)$. As $|\mathbb{F}| \geq d^2$, Lemma 3.4 implies that there is a setting of the $\mathbf{a}$-variables that makes $\alpha_1, \ldots, \alpha_d$ distinct field elements. In other words, $\alpha_1, \ldots, \alpha_d$ are pairwise distinct linear forms in $\mathbf{a}$-variables. Hence, from the Schwartz-Zippel lemma, on setting $a_1, \ldots, a_d$ uniformly at random from S, $\alpha_1, \ldots, \alpha_d$ become distinct elements of $\mathbb{F}$ with high probability.

Compute the characteristic polynomial (Definition 2.21) of L, denoted $h_L(z)$ and factorize it. As f is equivalent to NW, L and R are similar matrices and their characteristic polynomials are the same. Then $h_L(z)$ factorizes as $h_L(z) = (z-\beta_1)^d \cdots (z-\beta_d)^d$, for distinct $\beta_1, \ldots, \beta_d \in \mathbb{F}$ such that there is an (unknown) permutation σ on $[d]$ such that $\beta_i = \alpha_{\sigma(i)}$ for $i \in [d]$. Suppose $B = \mathrm{diag}(\beta_1, \ldots, \beta_d) \otimes I_d$. Let D be a $d^2 \times d^2$ size formal matrix such that

$$L \cdot D = D \cdot B. \tag{3.7}$$

Solve the system of linear equations obtained from Equation (3.7) (by treating the entries of

D as variables) and pick a random matrix from the solution space; call this solution matrix D. With high probability D is invertible (as $D = A^{-1}P$ is also in the solution space for a suitable permutation matrix P). Equation (3.7) implies that

$$R \cdot A \cdot D = A \cdot D \cdot B.$$

Recall that $R = \mathrm{diag}(\alpha_1, \ldots, \alpha_d) \otimes I_d$ and $B = \mathrm{diag}(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(d)}) \otimes I_d$. As $\alpha_1, \ldots, \alpha_d$ are distinct, it is an easy exercise to show that AD is a block permuted matrix. Hence $f(D\mathbf{x})$ is BP equivalent to NW. ■

A BD-PS equivalence test for NW

We saw in Lemma 3.3 that an efficient BP equivalence test for NW immediately implies an efficient equivalence test for it. In this section, we give a special case of the BP equivalence test, called BD-PS equivalence test (or block-diagonal permutation scaling equivalence test). Recall that in the BD-PS equivalence test, the underlying matrix is a product of a block-diagonal permutation matrix (recall the definition of a block-diagonal permutation matrix from Section 1.3.1) and an invertible scaling (or diagonal) matrix. We hope that this variant of ET would give us some useful insights on the BP ET for NW. We first recall Theorem 1.5.

Theorem 3.6 (BD-PS ET for NW) *Let d be a prime number, $\mathbb{F}$ be a finite field such that $d \nmid (|\mathbb{F}| - 1)$ and $|\mathbb{F}| \geq 4d$. There is a randomized $\mathrm{poly}(d, \log|\mathbb{F}|)$ time algorithm that takes input black-box access to a degree d polynomial $f \in \mathbb{F}[\mathbf{x}]$ and correctly decides if f is BD-PS equivalent to NW with high probability. If the answer is yes then it outputs a BD-PS matrix $C \in \mathrm{GL}(d^2, \mathbb{F})$ such that $f = \mathsf{NW}(C\mathbf{x})$.*

The BD-PS ET has two steps: First we reduce the BD-PS equivalence test to scaling equivalence test and then solves the scaling equivalence test. In the scaling equivalence test, it is determined whether the given polynomial is equivalent to NW via an invertible scaling matrix. We assume that the given polynomial f is BD-PS equivalent to NW and the equivalence test computes a block-diagonal permutation matrix A and an invertible scaling matrix B. Thereafter, it uses a circuit testing algorithm (Theorem 3.4) to determine whether $f(A^{-1}B^{-1}\mathbf{x}) = \mathsf{NW}$.

1. Reduction of BD-PS equivalence test to scaling equivalence test.
Assume $f = \mathsf{NW}(BA\mathbf{x})$, where A is a block-diagonal permutation matrix and B is an invertible scaling matrix. Algorithm 7 does not explicitly use the knowledge of the entries of B. Thus, we may assume without loss of generality that $B = I_{d^2}$. Then, the task reduces to solving the BD permutation equivalence test for NW. We identify matrix A with d permutations $\sigma_0, \ldots, \sigma_{d-1}$

on $[d]$ as $A = \text{diag}(A_{\sigma_0}, \ldots, A_{\sigma_{d-1}})$, where A_{σ_i} is the $d \times d$ permutation matrix corresponding to σ_i ,i.e., for $i, r, s \in [d]$, $A_{\sigma_i}(r, s) = 1$ if and only if $\sigma_i(r) = s$.

Observation 3.2 *Suppose f is BD permutation equivalent to* NW*, i.e., $f = \mathsf{NW}(A\mathbf{x})$. Then, a monomial $\prod_{i \in \mathbb{F}_d} x_{i,h(i)}$ of* NW *gets mapped to a unique monomial $\prod_{i \in \mathbb{F}_d} x_{i,\sigma_i(h(i))}$ of f.*

Algorithm 7 starts by assuming that $\sigma_0(0) = \cdots = \sigma_k(0) = 0$ and $\sigma_0(1) = 1$. The symmetries of NW allow us to make this assumption without loss of generality (see Claim 3.2.3). The aim is to figure out all the entries of σ_i [1]. This is done by carefully picking a bunch of polynomials from $\mathbb{F}_d[z]_k$ (which we call *nice* polynomials) and then exploiting the association between f and NW mentioned in Observation 3.2 using these polynomials. The algorithm works over every field. The following algorithm gradually discovers the entries of $\sigma_0, \ldots, \sigma_{d-1}$.

Algorithm 7 BD-Permutation-Equivalence(f)

Input: Black-box access to $f \in \mathbb{F}[\mathbf{x}]$.

Output: Black-box access to $g \in \mathbb{F}[\mathbf{x}]$ s.t. if f is BD-PS equivalent to NW then g is scaling equivalent to NW.

1. Assume that $\sigma_0(0) = \cdots = \sigma_k(0) = 0$ and $\sigma_0(1) = 1$ (Claim 3.2.3).
2. Construct a list of nice polynomials in $\mathbb{F}_d[z]_k$ (Definition 3.1) as mentioned in Claim 3.2.4.
3. Recover $(d - k)$ distinct entries of each permutation $\sigma_0, \ldots, \sigma_{d-1}$ as mentioned in Claim 3.2.5.
4. Let N be a $d \times d$ matrix, where the columns and rows are indexed by $(\sigma_0, \ldots, \sigma_{d-1})$ and $(0, \ldots, d-1)$ respectively and for $l, i \in [d]$, $N(l, i) := \sigma_i(l)$. Pick $l_0, \ldots, l_k \in [d]$ such that in each of the rows indexed by $l_0, \ldots, l_k$ at least $k + 1$ entries are known (Claim 3.2.6).
5. Use $l_0, \ldots, l_k \in [d]$ to recover all the entries of the rows of N as mentioned in Claim 3.2.7. Compute $A = \text{diag}(A_{\sigma_0}, \ldots, A_{\sigma_{d-1}})$ and return black box access to $f(A^{-1}\mathbf{x})$

Proof of correctness. The correctness of Algorithm 7 follows from the following chain of claims, which are proved immediately after this proof. In these claims, ρ is the bit complexity of the coefficients of f.

Claim 3.2.3 (Canonical form of $\sigma_0, \ldots, \sigma_{d-1}$) *Suppose $f \in \mathbb{F}[\mathbf{x}]$ is BD permutation equivalent to* NW*. Then, there exist permutations $\sigma_0, \ldots, \sigma_{d-1}$ on $[d]$ such that $\sigma_0(0) = \cdots = \sigma_k(0) = 0, \sigma_0(1) = 1$ and $A = \text{diag}(A_{\sigma_0}, \ldots, A_{\sigma_{d-1}})$ satisfies $f = \mathsf{NW}(A\mathbf{x})$.*

[1] σ_i is treated as an ordered tuple $(\sigma_i(0), \ldots, \sigma_i(d-1))$

The above claim helps to kick-start Algorithm 7. After that, we compute certain nice polynomials, defined below.

Definition 3.1 (Nice polynomials) *A set $\{h_0, \ldots, h_{d-k-1}\} \subseteq \mathbb{F}_d[z]_k$ is called a list of nice polynomials if the following properties are satisfied:*

1. *For distinct $r_1, r_2 \in [d-k]$, $h_{r_1}(\ell) = h_{r_2}(\ell)$ for every $\ell \in [k]$ and $h_{r_1}(\ell) \neq h_{r_2}(\ell)$ for every $\ell \in \{k, \ldots, d-1\}$.*

2. *For every $r \in [d-k], \sigma_0(h_r(0)), \ldots, \sigma_k(h_r(k))$ can be computed in poly(d, ρ) time.*

Claim 3.2.4 *A list of nice polynomials $\{h_0, \ldots, h_{d-k-1}\}$ can be computed in* poly(d, ρ) *time.*

Using the list of nice polynomials, we recover $d-k$ distinct entries of $\sigma_0, \ldots, \sigma_{d-1}$.

Claim 3.2.5 *Given a list of nice polynomials $\{h_0, \ldots, h_{d-k-1}\}$, we can recover $d-k$ distinct entries in each of $\sigma_0, \ldots, \sigma_{d-1}$ in* poly(d, ρ) *time.*

The matrix N defined in the algorithm is filled with some known entries and some unknowns. The goal is to recover all the entries of N which is accomplished by the following claims.

Claim 3.2.6 *Suppose $k \in [1, \frac{d}{3}]$. Then, there exist $k+1$ rows in N such that in each of these rows at least $k+1$ entries are known.*

Claim 3.2.7 *Using $k+1$ rows of N indexed by $l_0, \ldots, l_k$ (as mentioned in Step 4), we can recover all the entries of N in* poly(d, ρ) *time.*

This completes the proof of correctness of Algorithm 7. Now, we give the proofs of these claims one by one.

Proof of Claim 3.2.3. Since $f \in \mathbb{F}[\mathbf{x}]$ is BD permutation equivalent to NW, there exist permutations $\pi_0, \ldots, \pi_{d-1}$ on $[d]$, such that $A' = \mathrm{diag}(A_{\pi_0}, \ldots, A_{\pi_{d-1}})$ satisfies $f = \mathsf{NW}(A'\mathbf{x})$. Let $h \in \mathbb{F}_d[z]_k$ such that $\pi_0(0) = h(0), \ldots, \pi_k(0) = h(k)$. For $i \in [d]$, define $\sigma_i : \mathbb{F}_d \to \mathbb{F}_d$ as

$$\sigma_i(l) := \alpha \cdot (\pi_i(l) - h(i)) \text{ for all } l \in [d],$$

where $\alpha := \frac{1}{\pi_0(1) - h(0)}$. Note that for every $i \in [d]$, σ_i is well defined as $\pi_0(1) \neq h(0)$. The following observation can be verified easily.

Observation 3.3 *$\sigma_0, \ldots, \sigma_{d-1}$ are permutations on $\mathbb{F}_d$. Also, $\sigma_0(0) = \cdots = \sigma_k(0) = 0$ and $\sigma_0(1) = 1$.*

For $i \in [d]$, let $\tau_i : \mathbb{F}_d \to \mathbb{F}_d$ be defined as $\tau_i(l) := \alpha \cdot (l - h(i))$ for every $l \in \mathbb{F}_d$. Observe that $\tau_0, \ldots, \tau_{d-1}$ are permutations on $\mathbb{F}_d$ and for every $i \in [d]$

$$\sigma_i = \tau_i \circ \pi_i. \tag{3.8}$$

Let $A = \mathrm{diag}(A_{\sigma_0}, \ldots, A_{\sigma_{d-1}}), B' = \mathrm{diag}(B'_{\tau_0}, \ldots, B'_{\tau_{d-1}})$. As A, A', C are block diagonal matrices, the above equation implies

$$A = B' \cdot A'.$$

Observation 3.4 *Let B' be the matrix defined above. Then, $B' \in \mathscr{G}_{\mathsf{NW}}$.*

Proof: On applying B' on $\mathbf{x}$, $x_{i,j}$ gets mapped to $x_{i,\alpha\cdot(j-h(i))}$ for every $i, j \in [d]$. This shows $B' \in \mathscr{G}_{\mathsf{NW}}$ (similar to item 2 of Claim 3.1.1). □

Since $\mathsf{NW}(\mathbf{x}) = \mathsf{NW}(B'\mathbf{x})$, we get $f = \mathsf{NW}(B'A'\mathbf{x}) = \mathsf{NW}(A\mathbf{x})$. This completes the proof.

Proof of Claim 3.2.4. We create two lists of $d-k$ distinct polynomials in $\mathbb{F}_d[z]_k$, namely the p-list and the h-list as described below. Then we show that the h-list is a list of nice polynomials.

A procedure to create h-list and p-list:

1. Interpolate $(0,0), \ldots, (k,0)$ to get $p_0 \in \mathbb{F}_d[z]_k$ and then interpolate $(0,1), (1,0), \ldots, (k-1,0), (k,0)$ to get $h_0 \in \mathbb{F}_d[z]_k$. (In this case, $p_0 = 0$ and $h_0 \neq 0$.)
2. Interpolate $(0,0), \ldots, (k-1,0), (k+1, h_0(k+1))$ to get $p_1 \in \mathbb{F}_d[z]_k$ and then interpolate $(0,1), (1,0) \ldots, (k-1,0), (k, p_1(k))$ to get $h_1 \in \mathbb{F}_d[z]_k$.
3. For $r \in \{2, \ldots, d-k-1\}$ do the following.
   (a) For $r_1 = 1$ to r, interpolate $(0,0), \ldots, (k-1,0), (k+r_1, h_{r-1}(k+r_1))$ to get $\tilde{p}_{r_1} \in \mathbb{F}_d[z]_k$. (It is argued in Observation 3.6 that $\tilde{p}_1, \ldots, \tilde{p}_r$ are distinct polynomials.) Pick a polynomial from $\tilde{p}_1, \ldots, \tilde{p}_r$ that is different from each of $p_0, \ldots, p_{r-2}$. Set that polynomial to be p_r. (It is argued in Observation 3.7 that no polynomial amongst $\tilde{p}_1, \ldots, \tilde{p}_r$ is equal to p_{r-1}, and so $p_r \neq p_i$ for all $i \in [r]$.)
   (b) Interpolate $(0,1), (1,0), \ldots, (k-1,0), (k, p_r(k))$ to get $h_r \in \mathbb{F}_d[z]_k$.

We note some easy-to-verify observations about these lists.

Observation 3.5 *1. The p-list and h-list can be computed in* $\text{poly}(d)$ *time and they do not have a polynomial in common.*

2. All polynomials in the p-list (similarly in the h-list) agree on k points, namely $0, \ldots, k-1$.

3. For distinct $r, r' \in [d-k]$, p_r *and* h_r *agree on k points* $1, \ldots, k$, *and* p_r *and* $h_{r'}$ *agree on* $k-1$ *points* $1, \ldots, k-1$.

The following two sub claims imply that $\{h_0, \ldots, h_{d-k-1}\}$ is a list of distinct nice polynomials.

Subclaim 3.2.1 *Each of the p-list and h-list contains* $d-k$ *distinct polynomials.*

Proof: For some $r \in [d-k]$, item 2 of Observation 3.5 implies that if $p_0, \ldots, p_r$ are pairwise distinct then $h_0, \ldots, h_r$ are also pairwise distinct. We show that $p_0, \ldots, p_r$ are pairwise distinct polynomials by induction on r. The base case, i.e. $r = 0$ is trivially satisfied. Suppose the hypothesis holds for $r-1$, i.e. $p_0, \ldots, p_{r-1}$ are pairwise distinct. This implies $h_0, \ldots, h_{r-1}$ are also pairwise distinct. We construct r polynomials $\tilde{p}_1, \ldots \tilde{p}_r$ in $\mathbb{F}_d[z]_k$ by interpolating $(0,0), \ldots, (k-1,0), (k+1, h_{r-1}(k+1)); \ldots; (0,0), \ldots, (k-1,0), (k+r, h_{r-1}(k+r))$ respectively. Consider the following observations.

Observation 3.6 $\tilde{p}_1, \ldots, \tilde{p}_r$ *are distinct polynomials in* $\mathbb{F}_d[z]_k$.

Proof: Suppose not. Then, there exist distinct $r_1, r_2 \in \{1, \ldots, r\}$, such that $\tilde{p}_{r_1} = \tilde{p}_{r_2}$. This implies that the polynomials $\tilde{p}_{r_1}$ and h_{r-1} agree on $k+1$ points $1, \ldots, k-1, k+r_1$ and $k+r_2$, which is a contradiction because $\tilde{p}_{r_1}$ and h_{r-1} are distinct polynomials (recall that $\tilde{p}_{r_1}(0) = 0$ whereas $h_{r-1}(0) = 1$). □

Observation 3.7 *For every* $r_1 \in \{1, \ldots, r\}$, $\tilde{p}_{r_1} \neq p_{r-1}$.

Proof: Suppose not. Then, there exists $r_1 \in \{1, \ldots, r\}$ such that, $\tilde{p}_{r_1} = p_{r-1}$. Then, $p_{r-1}(k+r_1) = \tilde{p}_{r_1}(k+r_1) = h_{r-1}(k+r_1)$, which along with item 3 of Observation 3.5 implies that h_{r-1} and p_{r-1} agree on $k+1$ points $1, \ldots, k, k+r_1$, which can not happen as p_{r-1} and h_{r-1} are distinct polynomials. □

Hence, $p_0, \ldots, p_r$ are distinct polynomials. This completes the proof of Subclaim 3.2.1 □

The following fact would be required to prove that $\{h_0, \ldots, h_{d-k-1}\}$ is a list of nice polynomials.

Fact 3.2 *Suppose* $h \in \mathbb{F}_d[z]_k$ *and* $i_0, \ldots, i_k \in [d]$ *be distinct elements. Given* $\sigma_{i_0}(h(i_0)), \ldots, \sigma_{i_k}(h(i_k))$, *we can compute* $\sigma_i(h(i))$ *for every* $i \in [d] \setminus \{i_0, \ldots, i_k\}$ *in* $\text{poly}(d, \rho)$ *time.*

Proof: Since f is BD permutation equivalent to NW, Observation 3.2 implies that on setting $x_{i_0,\sigma_{i_0}(h(i_0))} = \cdots = x_{i_k,\sigma_{i_k}(h(i_k))} = 1$ and other variables $\mathbf{x}$ equal to zero, f reduces to

$$c \cdot \prod_{i \in [d] \setminus \{i_0,\ldots,i_k\}} x_{i,\sigma_i(h(i))}, \qquad \text{where } c \in \mathbb{F}.$$

It is easy to show that in this case $\sigma_i(h(i))$ for $i \in [d] \setminus \{i_0, \ldots, i_k\}$ can be recovered in $\mathrm{poly}(d, \rho)$ time from black-box access to f. □

Subclaim 3.2.2 *For every $r \in [d-k], \sigma_0(h_r(i)), i \in [k]$ can be computed in $\mathrm{poly}(d, \rho)$ time.*

Proof: For every $r \in [d-k], \sigma_0(h_r(0)) = 1, \sigma_1(h_r(1)) = \cdots = \sigma_{k-1}(h_r(k-1)) = 0$ from Step 1 of Algorithm 7. We show that $\sigma_k(h_r(k))$ can be computed efficiently by induction on r. When $r = 0$, we know that $\sigma_k(h_0(k)) = \sigma_k(0) = 0$. Thus, the base case holds. Suppose that the hypothesis holds for $r-1$, i.e., we can efficiently compute $\sigma_k(h_{r-1}(k))$. Recall that p_r is computed by interpolating $(0,0), \ldots, (k-1,0), (k+r_1, h_{r-1}(k+r_1))$ for some $r_1 \in \{1, \ldots, r\}$. Using Fact 3.2 on $\sigma_0(h_{r-1}(0)), \ldots, \sigma_{k-1}(h_{r-1}(k-1)), \sigma_k(h_{r-1}(k))$ we compute $\sigma_{k+r_1}(h_{r-1}(k+r_1)) = \sigma_{k+r_1}(p_r(k+r_1))$ and then using Fact 3.2 again on $\sigma_0(0), \ldots, \sigma_{k-1}(0), \sigma_{k+r_1}(p_r(k+r_1))$, we compute $\sigma_k(p_r(k))$, which is equal to $\sigma_k(h_r(k))$. □

This completes the proof of Claim 3.2.4.

Proof of Claim 3.2.5. We first show that using $S := \{h_0, \ldots, h_{d-k-1}\}$, we can recover $(d-k)$ distinct entries of each of the permutations $\sigma_{k+1}, \ldots, \sigma_{d-1}$. Fix an $i \in \{k+1, \ldots, d-1\}$. As $h_0, \ldots, h_{d-k-1}$ are nice polynomials, for every $h \in \{h_0, \ldots, h_{d-k-1}\}$, $\sigma_0(h(0)), \ldots, \sigma_k(h(k))$ can be computed efficiently. By invoking Fact 3.2 on $\sigma_0(h(0)), \ldots, \sigma_k(h(k))$ for every such h, we get $\sigma_i(h_0(i)), \ldots, \sigma_i(h_{d-k-1}(i))$. Since $h_1(i) = h_2(i)$ for every $h_1 \neq h_2 \in S$, and for every $i \in [k]$, we get that for every $\ell \in \{k, \ldots, d-1\}, h_1(\ell) \neq h_2(\ell)$. From item 2 of Observation 3.5 and Subclaim 3.2.1 and the fact that σ_i is a permutation, $\sigma_i(h_0(i)), \ldots, \sigma_i(h_{d-k-1}(i))$ are $d-k$ distinct entries of σ_i.

Now using the $d-k$ known entries of σ_{k+1}, we recover $d-k$ distinct entries of each of $\sigma_0, \ldots, \sigma_k$. Suppose there exist distinct $l_0, \ldots, l_{d-k-1} \in [d]$, such that $\sigma_{k+1}(l_0), \ldots, \sigma_{k+1}(l_{d-k-1})$ are known. Fix an $i \in [k+1]$. For $s \in [d-k]$, let p_s be a polynomial in $\mathbb{F}_d[z]_k$ obtained by interpolating $(i', 0), (k+1, l_s)$ for $i' \in [k+1] \setminus \{i\}$. Observe that these are $d-k$ distinct polynomials. Further, for $s_1 \neq s_2, p_{s_1}$ and p_{s_2} agree on k points $i' \in [k+1] \setminus \{i\}$ and $p_{s_1}(i) \neq p_{s_2}(i)$, which implies that $(\sigma_i(p_0(i)), \ldots, \sigma_i(p_{d-k-1}(i)))$ is a tuple of distinct entries. Using Fact 3.2 on $\sigma_{i'}(p_s(i')), \sigma_{k+1}(p_s(k+1))$ for $i' \in [k+1] \setminus \{i\}$, we obtain $d-k$ distinct values $\sigma_i(p_s(i))$ for every $s \in [d-k]$. This shows that for every $i \in [k+1]$, we can compute $d-k$ distinct entries of σ_i

efficiently. This completes the proof. ■

Proof of Claim 3.2.6. Suppose this is not true. Then, N has at most k rows such that in each row at least $k+1$ entries are known, and in the remaining at least $d-k$ rows at most k entries are known. This implies that at most $d \cdot k + (d-k)k$ entries are known in N. We know exactly $d(d-k)$ entries in N due to Claim 3.2.5. Thus, $d(d-k) \leq 2dk - k^2$, which implies $k > \frac{d}{3}$. This is a contradiction. ■

Proof of Claim 3.2.7. First we show how to recover all the entries of the rows of N indexed by $l_0, \ldots, l_k$. Given that in the rows of N indexed by $l_0, \ldots, l_k$, at least $k+1$ entries are known. For $l \in \{l_0, \ldots, l_k\}$, there exist distinct $i_0, \ldots, i_k \in [d]$, such that $\sigma_{i_0}(l), \ldots, \sigma_{i_k}(l)$ are known. Using Fact 3.2 on $\sigma_{i_0}(l), \ldots, \sigma_{i_k}(l)$, we recover $\sigma_i(l)$ for every $i \in [d] \setminus \{i_0, \ldots, i_k\}$.

Now we show how to recover $\sigma_i(l)$ for every $l \in [d] \setminus \{l_0, \ldots, l_k\}$ and $i \in [d]$. Let $h = z + (l-i)$. Clearly, $h(i) = l$. Let $i_0, \ldots, i_k \in [d]$ be such that $l_0 = i_0 + l - i, \ldots, l_k = i_k + l - i$. Then, $h(i_0) = l_0, \ldots, h(i_k) = l_k$. Use Fact 3.2 on the points $\sigma_{i_0}(h(i_0)), \ldots, \sigma_{i_k}(h(i_k))$ to recover $\sigma_i(h(i))$, which is $\sigma_i(l)$. Thus, we recover all the entries of N. ■

This completes the proofs of Claims 3.2.3-3.2.7.

2. Scaling equivalence test for NW. We present the scaling ET for NW over a finite field $\mathbb{F}$, where $d \nmid |\mathbb{F}| - 1$. We also give a scaling equivalence test for NW over $\mathbb{R}$ by appropriately modifying this algorithm. Assume that f is scaling equivalent to NW.

Algorithm 8 Scaling-ET(f)

Input: Black box access to $f \in \mathbb{F}[\mathbf{x}]$.

Output: An invertible diagonal matrix B such that $f = \mathsf{NW}(B\mathbf{x})$.

1. Let $B = \mathrm{diag}(\alpha_{0,0}, \ldots, \alpha_{d-1,d-1})$, where $\{\alpha_{i,j} \ : \ i,j \in [d]\}$ are unknown. Set $\alpha_{1,0} = \cdots = \alpha_{d-1,0} = 1$ (Claim 3.2.8).
2. Let $S = (0, z, \ldots, (d-1)z,\ 1, z+1 \ldots, (d-1)z+1,\ \ldots\ , d-2, z+d-2 \ldots, (d-1)z+d-2,\ d-1)$ be the ordered set of $d^2 - d + 1$ polynomials in $\mathbb{F}[z]$. For every $h \in S$, query the coefficient c_h of the monomial $\prod_{i \in \mathbb{F}_d} x_{i,h(i)}$ from the black-box of f (Observation 3.8).
3. Let C be a 0/1 matrix of size $(d^2 - d + 1) \times (d^2 - d + 1)$ whose rows and columns are indexed by S and $\mathbf{y} = (y_{0,0}, \ldots, y_{0,d-1}, y_{1,1}, \ldots, y_{1,d-1}, \ldots, y_{d-1,1}, \ldots, y_{d-1,d-1})$, respectively, such that for $h \in S$ and $y_{i,j} \in \mathbf{y}$, the $(h, y_{i,j})$-th entry of C is 1 if $h(i) = j$. (It is argued

in Claim 3.1.3 that $|\det(C)|$ is a power of d). Compute the inverse of $\det(C)$ in $\mathbb{Z}_{|\mathbb{F}|-1}$ and denote it by γ. (Note that $\mathbf{y}$ does not contain the variables $\{y_{1,0}, \ldots, y_{d-1,0}\}$.)

4. Fix $\alpha_{i,j} \in \{\alpha_{0,0}, \ldots, \alpha_{d-1,d-1}\} \setminus \{\alpha_{1,0}, \ldots, \alpha_{d-1,0}\}$ arbitrarily. For every $h \in S$, compute the minor of C with respect to the row and column indexed by h and $y_{i,j}$ respectively and call it δ_h. Set $\alpha_{i,j} = \prod_{h \in S} c_h^{(\delta_h \cdot \gamma) \mod (|\mathbb{F}|-1)}$.
5. Return $B = \mathrm{diag}(\alpha_{0,0}, \ldots, \alpha_{d-1,d-1})$. (see Claim 3.2.9)

Proof of correctness: The following claims and observations argue the correctness of the algorithm. We first complete this proof and then prove Claims 3.2.8 and 3.2.9.

Claim 3.2.8 *We can assume that* $\alpha_{1,0} = \cdots = \alpha_{d-1,0} = 1$ *without loss of generality.*

The following observation can be proved easily.

Observation 3.8 *Given a monomial* m *in* $\mathbf{x}$ *variables, we can recover the coefficient of* m *in* f *in* $\mathrm{poly}(d, \rho)$ *time.*

Claim 3.2.9 *In Step 4,* $\alpha_{i,j}$ *can be computed in* $\mathrm{poly}(d, \rho)$ *time. Further,* $f = \mathsf{NW}(B\mathbf{x})$.

This completes the proof of correctness of Algorithm 8.

Proof of Claim 3.2.8. As f is scaling equivalent to NW, there exists a $C = \mathrm{diag}(\beta_{0,0}, \ldots, \beta_{d-1,d-1})$ such that $f = \mathsf{NW}(C\mathbf{x})$. Suppose $D = \mathrm{diag}(a, \beta_{1,0}^{-1}, \ldots \beta_{d-1,0}^{-1}) \otimes I_d$, where $a = \prod_{i=1}^{d-1} \beta_{i,0}$. Then, from Claim 3.1.1, $D \in \mathscr{G}_{\mathsf{NW}}$, which implies $f = \mathsf{NW}(DC\mathbf{x})$. Set $B = DC$. Hence $\alpha_{1,0} = \ldots = \alpha_{d-1,0} = 1$.

Proof of Claim 3.2.9. For $i, j \in [d]$, suppose $\alpha_{i,j} = \tau^{y_{i,j}}$, where τ is a generator of $\mathbb{F}^\times$. Claim 3.2.8 implies that $y_{1,0} = \ldots = y_{d-1,0} = 0$. If $f = \mathsf{NW}(B\mathbf{x})$, then a monomial $m_h = \prod_{i \in \mathbb{F}_d} x_{i,h(i)}$ of NW gets mapped to $c_h \cdot m_h$, where $c_h = \prod_{i \in \mathbb{F}_d} \alpha_{i,h(i)}$. Let $c_h = \tau^{e_h}$. Then, we get the following system of linear equations for every $h \in \mathbb{F}_d[z]_k$ over the ring $\mathbb{Z}_{|\mathbb{F}|-1}$.

$$\sum_{i \in \mathbb{F}_d} y_{i,h(i)} = e_h. \tag{3.9}$$

Recall C, S and $\mathbf{y}$ from Step 3 of the algorithm. On restricting to the polynomials in S, we get

$$C \cdot \mathbf{y}^T = \boldsymbol{e},$$

where $\boldsymbol{e} = (e_0\ e_z \ldots e_{(d-1)z}\ e_1\ e_{z+1} \ldots e_{(d-1)z+1} \ldots e_{d-2}\ e_{z+d-2} \ldots\ e_{(d-1)z+d-2}\ e_{d-1})^T$. Recall γ and δ_h from Step 3 and 4. From Cramer's rule, we get

$$y_{i,j} = \gamma \cdot \left(\sum_{h \in S} e_h \cdot \delta_h \right) \mod (|\mathbb{F}| - 1), \tag{3.10}$$

This immediately implies,

$$\alpha_{i,j} = \tau^{y_{i,j}} = \tau^{\gamma \cdot \left(\sum_{h \in S} e_h \cdot \delta_h\right) \mod (|\mathbb{F}|-1)},$$

As $c_h = \tau^{e_h}$, $\alpha_{i,j} = \prod_{h \in S} c_h^{(\delta_h \cdot \gamma) \mod (|\mathbb{F}|-1)}$. As C is a 0/1 matrix, $|\det(C)|$ is bounded by $(d^2 - d + 1)!$, which implies the bit complexity of $\det(C)$ is $\mathrm{poly}(d)$. This implies that the above calculations can be done in $\mathrm{poly}(d, \rho)$ time using repeated squaring. ■

Scaling equivalence test for NW over $\mathbb{R}$. We first state the model of computation over $\mathbb{R}$. We assume that addition, subtraction, multiplication and division of two real numbers can be done in unit time. In addition, we also assume that the positive real root of a univariate real polynomial $y^r - \delta$ can be computed in $\mathrm{poly}(\log r)$ time (see [Bre76, Ye94]).

Suppose a degree d polynomial $f \in \mathbb{F}[\mathbf{x}]$ is scaling equivalent to NW. We wish to find a $B = \mathrm{diag}(\alpha_{0,0}, \ldots, \alpha_{d-1,d-1}) \in \mathrm{GL}(d^2, \mathbb{R})$, such that $f = \mathsf{NW}(B\mathbf{x})$. Note that every $\alpha_{i,j}$ can be written as $\alpha_{i,j} = (-1)^{s_{i,j}} \cdot 2^{\beta_{i,j}}$, where $s_{i,j} \in \mathbb{F}_2$ and $\beta_{i,j} \in \mathbb{R}$. Assume $s_{i,j}, \beta_{i,j}, i, j \in [d]$ are formal variables. Here also, we can assume without loss of generality that $\alpha_{1,0} = \ldots = \alpha_{d-1,0} = 1$, which sets $s_{i,0} = \beta_{i,0} = 0$ for $i \in \{1, \ldots, d-1\}$. For $h \in \{az + b : a \in [d], b \in [d-1]\} \cup \{d-1\}$, let $c_h = (-1)^{\delta_h} \cdot 2^{\gamma_h}$ be the coefficient of $\prod_{i \in \mathbb{F}_d} x_{i,h(i)}$ in f. This gives us the following system of linear equations in $\boldsymbol{\beta}$ and $\boldsymbol{s}$ variables over $\mathbb{R}$ and $\mathbb{F}_2$ respectively.

$$\sum_{i \in \mathbb{F}_d} \beta_{i,h(i)} = \gamma_h \qquad \text{and} \qquad \sum_{i \in \mathbb{F}_d} s_{i,h(i)} = \delta_h. \tag{3.11}$$

Hereon, the scaling equivalence test for NW over $\mathbb{R}$ can be obtained by easily adapting Algorithm 8 to solve the system of linear equations mentioned in Equation (3.11) and compute B.

Chapter 4

Determinant equivalence test over finite fields and over $\mathbb{Q}$

In this chapter, we prove the theorems stated in Section 1.3.2. This chapter is based on [GGKS19], which is a joint work with Ankit Garg, Neeraj Kayal and Chandan Saha. The content of the chapter is divided into four sections. In the first section, we state some useful properties related to the Lie algebra of the determinant. Using these properties along with other tools, we prove prove Theorems 1.6, 1.7 and 1.8 in the second section. Our main technical contribution is Theorem 1.8, which reduces DET to FMAI over almost any field. In the third section, we show that assuming GRH, the integer factoring problem reduces in randomized polynomial time to DET for quadratic forms over $\mathbb{Q}$. In the last section, we give a reduction from FMAI over $\mathbb{F}$ to DET over $\mathbb{F}$.

We start this chapter by introducing some notations. Let $n \in \mathbb{N}$, $\mathbb{F}$ be a field, $X = (x_{i,j})_{i,j\in[n]}$ be such that for every $i, j \in [n], x_{i,j}$ is a variable and $\mathbf{x} = \{x_{i,j} : i, j \in [n]\}$. Then, $\mathsf{Det}_n(\mathbf{x}) := \det(X)$. Clearly, $\mathsf{Det}_n(\mathbf{x})$ is a homogeneous polynomial of degree n. Whenever the value of n is clear from the context, we would drop the subscript of Det_n. Recall that $M_n(\mathbb{F})$ is the set of $n \times n$ matrices over $\mathbb{F}$. Then, $M_n(\mathbb{F})$ is an $\mathbb{F}$-algebra (Definition 2.18) with respect to matrix addition and matrix multiplication, and this is known as the *full matrix algebra*. Let $Z_n(\mathbb{F})$ be the set of traceless matrices in $M_n(\mathbb{F})$. Then, $Z_n(\mathbb{F})$ is clearly an $\mathbb{F}$-vector space. Whenever $\mathbb{F}$ is clear from the context, we will drop $\mathbb{F}$ from $M_n(\mathbb{F})$ and $Z_n(\mathbb{F})$. For $i, j \in [n], i \neq j$, let $E_{i,j} \in M_n$ be such that the (i, j)-th entry of $E_{i,j}$ is 1 and other entries are 0. For $\ell \in \{2, \ldots, n\}$, let $E_\ell \in M_n$ be the diagonal matrix, where the $(1, 1)$-th and (ℓ, ℓ)-th entries are 1 and -1 respectively and other entries are 0. Then, it is easy to see that $\{E_{i,j}, E_\ell : i, j \in [n], i \neq j, \ell \in [2, n]\}$ is an $\mathbb{F}$-basis of Z_n and hence $\dim Z_n = n^2 - 1$. We fix $r = n^2 - 1$ and $m = n^2$ for the rest of this chapter.

4.1 The Lie algebra of the determinant

In this section, we highlight the structure of the Lie algebra of Det, denoted $\mathfrak{g}_{\mathsf{Det}}$, which is well-studied. This structure is very crucially used in designing our DET algorithms. We first setup the notations required for describing the structure of $\mathfrak{g}_{\mathsf{Det}}$.

Let $I_n \in M_n$ be the identity matrix. Let $\mathscr{M}_{\text{col}} := I_n \otimes M_n$, i.e., $\mathscr{M}_{\text{col}} = \{I_n \otimes A : A \in M_n\}$, where $\otimes$ denote the tensor product of matrices (Definition 2.20). Similarly, we define $\mathscr{M}_{\text{row}} = M_n \otimes I_n, \mathscr{L}_{\text{col}} = I_n \otimes Z_n$ and $\mathscr{L}_{\text{row}} = Z_n \otimes I_n$. It is easy to see that $\mathscr{M}_{\text{col}}$ and $\mathscr{M}_{\text{row}}$ are $\mathbb{F}$-algebras (Definition 2.18) with respect to matrix addition and matrix multiplication and these are isomorphic as $\mathbb{F}$-algebras (see Definition 2.19) to M_n via the maps $\varphi : \mathscr{M}_{\text{col}} \to M_n; I_n \otimes A \mapsto A$ and $\phi : \mathscr{M}_{\text{row}} \to M_n; A \otimes I_n \mapsto A$ respectively. Further, $\mathscr{L}_{\text{col}}$ (similarly, $\mathscr{L}_{\text{row}}$) is an $\mathbb{F}$-subspace of $\mathscr{M}_{\text{col}}$ (respectively, $\mathscr{M}_{\text{row}}$). The fact that $\dim Z_n = r$ implies $\dim \mathscr{L}_{\text{col}} = \dim \mathscr{L}_{\text{row}} = r$. In fact, we also readily get bases of $\mathscr{L}_{\text{col}}$ and $\mathscr{L}_{\text{row}}$ from the basis of Z_n given before. We record this property as the following observation.

Observation 4.1 (Standard bases of $\mathscr{L}_{\text{col}}$ and $\mathscr{L}_{\text{row}}$) *For $i, j \in [n], i \neq j$, let $E_{ij} \in M_n$ be such that the (i,j)-th entry is 1 and other entries are 0, and for $\ell \in [2,n]$, let $E_\ell \in M_n$ be a diagonal matrix with the $(1,1)$-th and (ℓ,ℓ)-th entries as 1 and -1 respectively and other entries as 0. Then,*

1. *$\{I_n \otimes E_{ij},\ I_n \otimes E_\ell\ :\ i,j \in [n], i \neq j,\ \text{and}\ \ell \in [2,n]\}$ is a basis of $\mathscr{L}_{col}$. We call this as the standard basis of $\mathscr{L}_{col}$ and denote it as $\{S_1, \ldots, S_r\}$.*
2. *$\{E_{ij} \otimes I_n,\ E_\ell \otimes I_n\ :\ i,j \in [n], i \neq j,\ \text{and}\ \ell \in [2,n]\}$ is a basis of $\mathscr{L}_{row}$. We call this as the standard basis of $\mathscr{L}_{row}$ and denote it as $\{S_{r+1}, \ldots, S_{2r}\}$.*

The following well-known fact shows that $\mathscr{L}_{\text{col}}$ and $\mathscr{L}_{\text{row}}$ are the only subspaces of $\mathfrak{g}_{\mathsf{Det}}$. See Section 3.2 of [Nai19] for a proof of this fact. This fact will be crucially used later.

Fact 4.1 (Structure of $\mathfrak{g}_{\mathsf{Det}}$) *Let $n \in \mathbb{N}^\times$ and $\mathbb{F}$ be a field satisfying $char(\mathbb{F}) \nmid n$. Then, $\mathfrak{g}_{Det_n} = \mathscr{L}_{row} \oplus \mathscr{L}_{col}$.*

Let $f = \mathsf{Det}_n(A\mathbf{x})$ for some $A \in \mathrm{GL}(m, \mathbb{F})$. Claim 2.10 and Fact 4.1 imply the following.

Corollary 4.1 *Let $A \in \mathrm{GL}(m, \mathbb{F})$ and $f = \mathsf{Det}_n(A\mathbf{x})$. Let $\mathscr{F}_{col} := A^{-1} \cdot \mathscr{L}_{col} \cdot A$ and $\mathscr{F}_{row} := A^{-1} \cdot \mathscr{L}_{row} \cdot A$. Then, $\mathfrak{g}_f = \mathscr{F}_{row} \oplus \mathscr{F}_{col}$.*

Henceforth, we would refer to $\mathscr{L}_{\text{col}}$ and $\mathscr{L}_{\text{row}}$ (similarly, $\mathscr{F}_{\text{col}}$ and $\mathscr{F}_{\text{row}}$) as the Lie sub-algebras of $\mathfrak{g}_{\mathsf{Det}_n}$ (respectively, $\mathfrak{g}_f$). Now, we note some useful properties of $\mathfrak{g}_f$.

Properties of $\mathfrak{g}_f$

Recall the definition of the Lie bracket of matrices (Definition 2.28) from Chapter 2. In this section, we first show that $\mathfrak{g}_f$ is closed under the Lie bracket operation. For this, we require Observations 4.2 and 4.3, which are stated and proved below.

Observation 4.2 (Lie bracket of matrices in $\mathscr{M}_{\mathbf{col}}$ and $\mathscr{M}_{\mathbf{row}}$) *Let $F \in \mathscr{M}_{row}$ and $L \in \mathscr{M}_{row}$. Then, $[F, L] = 0$.*

Proof: Let $A = (a_{i,j})_{i,j\in[n]}, B = (b_{l,r})_{l,r\in[n]} \in M_n$ such that $F = A \otimes I_n$ and $L = I_n \otimes B$. Note that for $i, j, l, r \in [n]$, the $((i, j), (l, r))$-th entries of FL and LF are $a_{j,l}b_{i,r}$. Hence, $[F, L] = FL - LF = 0$. □

In other words, Observation 4.2 says that matrices in $\mathscr{M}_{\text{row}}$ commute with matrices in $\mathscr{M}_{\text{col}}$. It is easy to prove the following.

Observation 4.3 ($\mathscr{L}_{\mathbf{row}}$ and $\mathscr{L}_{\mathbf{col}}$ closed under Lie bracket) *For every $L_1, L_2 \in \mathscr{L}_{col}$ (similarly, $\mathscr{L}_{row}$), $[L_1, L_2] \in \mathscr{L}_{col}$ (respectively, $\mathscr{L}_{row}$).*

Note that Corollary 4.1, Observations 4.2 and 4.3 imply the following.

Observation 4.4 ($\mathfrak{g}_f$ closed under Lie bracket) *Let $n \in \mathbb{N}$, $A \in \mathrm{GL}(n, \mathbb{F})$ and $f = \mathsf{Det}_n(A\mathbf{x})$. Then, for every $E, F \in \mathfrak{g}_f$, $[E, F] \in \mathfrak{g}_f$.*

The second important property associated with $\mathfrak{g}_f$ is the structure of the $\mathbb{F}$-algebra generated by an $\mathbb{F}$-basis of $\mathscr{F}_{\text{col}}$ (see Remark 2.2 in this context). We note it in the following observation and this property will be very helpful for the DET algorithm.

Observation 4.5 (Algebra generated by $\mathscr{F}_{\mathbf{col}}$) *Let $n \in \mathbb{N}$, $A \in \mathrm{GL}(n, \mathbb{F})$ and $f = \mathsf{Det}_n(A\mathbf{x})$. Let $\mathscr{A} \subseteq M_m$ be the $\mathbb{F}$-algebra generated by a basis of $\mathscr{F}_{col}$. Then, $\mathscr{A} = A^{-1} \cdot (I_n \otimes M_n) \cdot A$.*

Proof: Consider the standard basis given in Observation 4.1. It is not difficult to see that the $\mathbb{F}$-algebra generated by this basis of $\mathscr{L}_{\text{col}}$ is equal to $I_n \otimes M_n$. This along with the fact that $\mathscr{F}_{\text{col}} = A^{-1} \cdot \mathscr{L}_{\text{col}} \cdot A$ immediately implies that $\mathscr{A} = A^{-1} \cdot (I_n \otimes M_n) \cdot A$. □

4.2 Reduction from DET to FMAI: The algorithm

In this section, we prove Theorem 1.8. This theorem gives a randomized polynomial time reduction from DET to FMAI (recall FMAI from Section 2.2.4), which works over any field satisfying mild conditions on the size and the characteristic. As seen in Section 2.2.4, FMAI algorithms over finite fields and $\mathbb{Q}$ are known. By invoking these algorithms, we get Theorems 1.6 and 1.7 readily from Theorem 1.8. We recall Theorems 1.6 and 1.7 below.

Theorem 4.1 (DET over finite fields) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_{1,1}, \ldots, x_{n,n}\}$, $\mathbb{F}$ be a finite field such that $|\mathbb{F}| \geq 10n^4$ and $char(\mathbb{F}) \nmid n(n-1)$, and $f \in \mathbb{F}[\mathbf{x}]$ be a degree n polynomial. Then, there exists a randomized algorithm that takes black-box access to f and decides if f is equivalent to Det_n or not with high probability. If yes, it outputs an $A \in \mathrm{GL}(n^2, \mathbb{F})$ such that $f = \mathsf{Det}_n(A\mathbf{x})$, otherwise it outputs 'Fail'. The running time of this algorithm is $\mathrm{poly}(n, \log |\mathbb{F}|)$.*

Theorem 4.2 (DET over $\mathbb{Q}$) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_{1,1}, \ldots, x_{n,n}\}$ and $f \in \mathbb{Q}[\mathbf{x}]$ be a degree n polynomial. Suppose we have black-box access to f. Let β be the bit length of coefficients of f.*

1. *There exists a randomized algorithm, which takes oracle access to an integer factoring algorithm* IntFact *and decides if f is equivalent to Det_n over $\mathbb{Q}$ with high probability. If yes, it outputs an $A \in \mathrm{GL}(n^2, \mathbb{Q})$ such that $f = \mathsf{Det}_n(A\mathbf{x})$, otherwise outputs 'Fail'. If n is bounded, the algorithm runs in $\mathrm{poly}(n, \beta)$ time.*

2. *There exists a randomized $\mathrm{poly}(n, \beta)$ time algorithm, which decides if f is equivalent to Det_n over $\mathbb{Q}$ with high probability. If yes, it outputs an $A \in \mathrm{GL}(n^2, \mathbb{L})$ such that $f = \mathsf{Det}_n(A\mathbf{x})$, where $\mathbb{L}$ is an extension field of $\mathbb{Q}$ satisfying $[\mathbb{L} : \mathbb{Q}] \leq n$* [1].

As mentioned above, the heart of the algorithms in the above two theorems is Theorem 1.8, which we recall below.

Theorem 4.3 (Reduction of DET to FMAI) *Let $n \geq 2$, $|\mathbb{F}| > 10n^4$ and $char(\mathbb{F}) \nmid n(n-1)$. Then, there exists a randomized polynomial time algorithm, with oracle access to FMAI, that takes input black-box access to an $f \in \mathbb{F}[\mathbf{x}]$ of degree n and solves DET for f over $\mathbb{F}$ with high probability.*

So, now the task is to prove Theorem 4.3. We first give an overview of the reduction from DET to FMAI, then present the algorithm and then argue its correctness.

An overview of the reduction from DET to FMAI. The algorithm takes black-box access to a polynomial f of degree n. We assume that $f = \mathsf{Det}_n(A\mathbf{x})$ for some $A \in \mathrm{GL}(n, \mathbb{F})$, otherwise the algorithm will detect with high probability that f is not equivalent to Det_n. The algorithm has two phases: In the first phase, it computes an $\mathbb{F}$-basis of the Lie algebra (Definition 2.30) of f, denoted $\mathfrak{g}_f$, using Fact 2.16. We know from Corollary 4.1 that $\mathfrak{g}_f = \mathscr{F}_{\text{row}} \oplus \mathscr{F}_{\text{col}}$. In this phase, the algorithm decomposes $\mathfrak{g}_f$ into $\mathscr{F}_{\text{col}}$ and $\mathscr{F}_{\text{row}}$. The details of this decomposition is given in Section 4.3.1. After decomposing $\mathfrak{g}_f$ and obtaining $\mathscr{F}_{\text{row}}$ and $\mathscr{F}_{\text{col}}$, the algorithm

[1] See Definition 2.12 for the meaning of $[\mathbb{L} : \mathbb{Q}]$.

computes the $\mathbb{F}$-algebra $\mathscr{A}$ generated by an $\mathbb{F}$-basis $\mathscr{F}_{\text{col}}$. Then, it follows from Observation 4.5 that $\mathscr{A}, M_n$ are isomorphic $\mathbb{F}$-algebras. In the second phase, we invoke FMAI on an $\mathbb{F}$-basis $\{L_1, \ldots, L_m\}$ of $\mathscr{A}$ and it returns an $\mathbb{F}$-algebra isomorphism $\varphi : \mathscr{A} \to M_n$ in the form of an $\mathbb{F}$-basis $\{C_1, \ldots, C_m\}$ of M_n, where for every $i \in [m], \varphi(L_i) = C_i$. Then, using the Skolem-Noether theorem (Theorem 2.1) we compute a $B \in \mathrm{GL}(n, \mathbb{F})$ from φ such that $f = \mathsf{Det}_n(B\mathbf{x})$. Now, we give the formal description of the algorithm. It uses Decompose-Lie-Algebra() (Procedure 10) as a sub-routine. This procedure takes input black-box access to f and returns a set of r many $\mathbb{F}$-linearly independent matrices. If f is equivalent to Det then this set is an $\mathbb{F}$-basis of $\mathscr{F}_{\text{col}}$. We give an overview, formal description and the analysis of Procedure 10 in Section 4.3.1.

Algorithm 9 Reduce-DET-to-FMAI(f)

Input: Black-box access to an $f \in \mathbb{F}[\mathbf{x}]$ of degree n, and oracle access to FMAI.

Output: A $B \in \mathrm{GL}(m, \mathbb{F})$ s.t. $f = \mathsf{Det}(B\mathbf{x})$, if such a B exists. Else, output 'Fail'.

/* Decomposing the Lie algebra of f */

1. Suppose $\{U_1, \ldots, U_r\}$ be the output of Decompose-Lie-Algebra(f) (Procedure 10).

/* Invoke FMAI algorithm */

2. Compute an $\mathbb{F}$-basis $\{L_1, \ldots, L_k\}$ of the $\mathbb{F}$-algebra $\mathscr{A}$ generated by $U_1, \ldots, U_r$. If $k \neq m$, output 'Fail'.
3. Invoke the FMAI oracle on $(L_1, \ldots, L_m)$ which returns a basis $(C_1, \ldots, C_m)$ of M_n.

/* Computing the matrix */

4. Pick a *random* $M \in M_m$ satisfying $L_i \cdot M = M \cdot (I_n \otimes C_i)$ for every $i \in [m]$.
5. Let b be the evaluation of $f(M\mathbf{x})$ obtained by setting $x_{1,1} = \cdots = x_{n,n} = 1$ and remaining $x_{i,j}$'s equal to 0.
6. If $M \notin \mathrm{GL}(m, \mathbb{F})$ or $b = 0$, output 'Fail'. Else, set $D = \mathrm{diag}(b, 1, \ldots, 1) \in M_n$.
7. Return $(I_n \otimes D) \cdot M^{-1}$.

4.3 Analysis of the algorithm

In this section, we argue the correctness of Algorithm 9. This will be done in two stages: In Section 4.3.1 we describe Procedure 10 and argue its correctness. Then, in Section 4.3.2, we argue the correctness of Steps 2 - 6 of Algorithm 9.

4.3.1 Decomposition of the Lie algebra of f in the orbit of Det_n

In this subsection, we prove the following theorem.

Theorem 4.4 (Decomposition of $\mathfrak{g}_f$) *Let $n \geq 2, \mathbf{x} = \{x_{1,1}, \ldots, x_{n,n}\}$, $\mathbb{F}$ be a field satisfying $|\mathbb{F}| \geq 10n^4$, and $char(\mathbb{F}) \nmid n(n-1)$ and $f \in \mathbb{F}[\mathbf{x}]$ be a degree n polynomial. There is a randomized algorithm, which takes input black-box access to f and if f is equivalent to Det, it outputs bases of $\mathscr{F}_{row}$ and $\mathscr{F}_{col}$ with high probability. The running time of this algorithm is $\mathrm{poly}(n, \gamma)$, where γ is the bit length of the coefficients of f.*

We first give a high level overview of the decomposition algorithm and then state it formally.

An overview of the decomposition algorithm. We assume $f = \mathsf{Det}(A\mathbf{x})$ for some $A \in \mathrm{GL}(n, \mathbb{F})$, otherwise it will be detected with high probability that f is not equivalent to Det because of the rich structure of $\mathfrak{g}_{\mathsf{Det}}$. This algorithm first computes black-box access to a basis $\{B_1, \ldots, B_{2r}\}$ of $\mathfrak{g}_f$ using Fact 2.15. Then, using this basis, we compute a basis $\{P_{B_1}, \ldots, P_{B_{2r}}\}$ of a set $\mathscr{P} \subseteq M_{2r}(\mathbb{F})$, which correspond to a 'special set' of $\mathbb{F}$-linear operators on $\mathfrak{g}_f$. $\mathscr{P}$ and the corresponding set of $\mathbb{F}$-linear operators on $\mathfrak{g}_f$ are described below. Then, we pick a random matrix Q in $\mathscr{P}$, compute its characteristic polynomial, denoted $h(z)$. Then, we factorize $h(z)$ using the algorithms in [Ber70] or [LLL82b] depending on whether $\mathbb{F}$ is a finite field or $\mathbb{Q}$. As f is equivalent to Det_n, the irreducible factors of $h(z)$ guide us to bases of $\mathscr{F}_{\mathrm{col}}$ and $\mathscr{F}_{\mathrm{row}}$ as follows: We compute an $\mathbb{F}$-basis of the null space of $h'(Q)$ for every factor h' of $h(z)$ such that h' is not a variable and then compute $\mathscr{P}$-closure (Definition 2.17) of every vector in bases of each of these null spaces. Because of the richness of $\mathfrak{g}_f$ induced by $\mathfrak{g}_{\mathsf{Det}_n}$, it turns out that this set of $\mathscr{P}$-closures of vectors only contains $\mathscr{F}_{\mathrm{row}}$ and $\mathscr{F}_{\mathrm{col}}$ as these are the only irreducible invariant spaces of $\mathscr{P}$. This is how we get access to bases of $\mathscr{F}_{\mathrm{row}}$ and $\mathscr{F}_{\mathrm{col}}$. Now, we first describe the set $\mathscr{P}$ and then present the algorithm.

The description of $\mathscr{P}$ and its properties. Suppose $f = \mathsf{Det}(A\mathbf{x})$ for some $A \in \mathrm{GL}(n, \mathbb{F})$.

Consider the following $\mathbb{F}$-linear operators on $\mathfrak{g}_f$: For every $F \in \mathfrak{g}_f$,

$$\begin{aligned} \rho_F \ : \ & \mathfrak{g}_f \to \mathfrak{g}_f \\ & E \mapsto [E, F]. \end{aligned}$$

Observation 4.4 implies that for every $E \in \mathfrak{g}_f, [E, F] \in \mathfrak{g}_f$. It is easy to see that ρ_F is an $\mathbb{F}$-linear map. As ρ_F is $\mathbb{F}$-linear, we can associate a matrix $P_F \in M_{2r}$ with ρ_F (see Definition 2.11), after fixing an ordering of the basis $(B_1, \ldots, B_{2r})$ of $\mathfrak{g}_f$ computed by the algorithm. Let $\mathscr{P} := \{P_F \ : \ F \in \mathfrak{g}_f\}$. Then, $\mathscr{P}$ is an $\mathbb{F}$-vector space.

Claim 4.3.1 ($\mathfrak{g}_f$ and $\mathscr{P}$ are isomorphic vector spaces) *Let $\mathbb{F}$ be a field such that $char(\mathbb{F}) \nmid n$. Then, $\mathfrak{g}_f$ and $\mathscr{P}$ are isomorphic as $\mathbb{F}$-vector spaces via the map $F \mapsto P_F$ for every $F \in \mathfrak{g}_f$.*

Proof: It is easy to see that $\mathscr{P}$ is an $\mathbb{F}$-vector space. Consider the following map

$$\begin{aligned} \tau : & \mathfrak{g}_f \to \mathscr{P} \\ & F \mapsto P_F \end{aligned}$$

Observe that τ is $\mathbb{F}$-linear and onto. Let $F \in \mathrm{Ker}(\tau)$[1]. Then $P_F = 0$, i.e., $[E, F] = 0$ for every $E \in \mathfrak{g}_f$, and hence $L := A \cdot F \cdot A^{-1} \in \mathfrak{g}_{\mathsf{Det}}$ commutes with every element of $\mathfrak{g}_{\mathsf{Det}}$. Recall the basis $\{S_1, \ldots, S_r\}$ and $\{S_{r+1}, \ldots, S_{2r}\}$ of $\mathscr{L}_{\mathrm{col}}$ and $\mathscr{L}_{\mathrm{row}}$ given in Observation 4.1. It is not difficult to show that as if L commutes with $\{S_1, \ldots, S_{2r}\}$ then $L = \alpha \cdot I_{n^2}$ for some $\alpha \in \mathbb{F}$. As $\mathrm{trace}(L) = 0$ and $char(\mathbb{F}) \nmid n$, we have $L = 0$. Hence, τ is injective. □

The above claim implies the following.

Observation 4.6 *The matrices $\{P_{B_1}, \ldots, P_{B_{2r}}\}$ is a basis of $\mathscr{P}$, which can be efficiently computed from $\{B_1, \ldots, B_{2r}\}$ (by considering the elements $[B_i, B_j]$, for $i, j \in [2r]$).*

We intend to study the irreducible invariant subspaces of $\mathscr{P}$ in order to compute bases of $\mathscr{F}_{\mathrm{row}}$ and $\mathscr{F}_{\mathrm{col}}$. Claim 4.3.2 would be useful in this regard as it relates $\mathscr{P}$ and $\mathfrak{g}_{\mathsf{Det}}$. For that, we need the following: For $i \in [2r]$, let $J_i := A \cdot B_i \cdot A^{-1}$. Then, it follows from Fact 2.10 that $J_i, i \in [2r]$, is an $\mathbb{F}$-basis of $\mathfrak{g}_{\mathsf{Det}}$. Like ρ_F, we can associate a $\mathbb{F}$-linear map χ_L with every $L \in \mathfrak{g}_{\mathsf{Det}}$ as follows:

$$\begin{aligned} \chi_L \ : \ & \mathfrak{g}_{\mathsf{Det}} \to \mathfrak{g}_{\mathsf{Det}} \\ & K \mapsto [K, L]. \end{aligned}$$

[1] $\mathrm{Ker}(\tau)$ is called the kernel of τ and is defined as $\mathrm{Ker}(\tau) := \{F \in \mathfrak{g}_f : \tau F = 0\}$.

Let $Q_L \in M_{2r}$ be the matrix corresponding to the linear map χ_L, with respect to the (ordered) basis $(J_1, \ldots, J_{2r})$. The following claim implies that it is sufficient to focus on $\mathfrak{g}_{\mathsf{Det}}$ while analysing the invariant subspaces of $\mathscr{P}$.

Claim 4.3.2 ($\mathscr{P}$ and $\mathfrak{g}_{\mathsf{Det}}$) *For every $i \in [2r]$, $Q_{J_i} = P_{B_i}$ and so $\mathscr{P} = \{Q_L : L \in \mathfrak{g}_{\mathsf{Det}}\}$.*

Proof: Let $E \in \mathfrak{g}_f, K \in \mathfrak{g}_{\mathsf{Det}}$ and $E = AKA^{-1}$. Observe that $\mathbf{u}_E = \mathbf{v}_K$, where $\mathbf{u}_E, \mathbf{v}_K$ are the coordinate vectors of E, K with respect to the bases $(B_1, \ldots, B_{2r})$ and $(J_1, \ldots, J_{2r})$ respectively[1]. Hence, $Q_{J_i}\mathbf{v}_K = \mathbf{v}_{[K,J_i]} = \mathbf{u}_{[E,B_i]} = P_{B_i}\mathbf{u}_E = P_{B_i}\mathbf{v}_K$, implying $Q_{J_i} = P_{B_i}$. □

Like Claim 4.3.1, $\mathfrak{g}_{\mathsf{Det}}$ and $\mathscr{P}$ are isomorphic as $\mathbb{F}$-vector spaces via the map $L \mapsto Q_L$, for $L \in \mathfrak{g}_{\mathsf{Det}}$. These connections of $\mathscr{P}$ with $\mathfrak{g}_f$ and $\mathfrak{g}_{\mathsf{Det}}$ will be very helpful. We will compute a basis of $\mathscr{P}$ using a basis of $\mathfrak{g}_f$ in the algorithm and will analyse some important properties of $\mathscr{P}$ using $\mathfrak{g}_{\mathsf{Det}}$. The algorithm computes two invariant subspaces $\mathscr{V}_1$ and $\mathscr{V}_2$ of $\mathscr{P}$ defined below.

$$\begin{aligned} \mathscr{V}_1 &= \left\{ \mathbf{v} = (a_1, \ldots, a_{2r})^T \in \mathbb{F}^{2r} \ : \ \sum_{i\in[2r]} a_i \cdot J_i \in \mathscr{L}_{\text{col}} \right\}, \\ \mathscr{V}_2 &= \left\{ \mathbf{v} = (b_1, \ldots, b_{2r})^T \in \mathbb{F}^{2r} \ : \ \sum_{i\in[2r]} b_i \cdot J_i \in \mathscr{L}_{\text{row}} \right\}. \end{aligned} \tag{4.1}$$

Clearly, $\dim(\mathscr{V}_1) = \dim(\mathscr{V}_2) = r$. As $B_i = A^{-1} \cdot J_i \cdot A$, for $i \in [2r]$, we get

$$\begin{aligned} \mathscr{V}_1 &= \left\{ \mathbf{v} = (a_1, \ldots, a_{2r})^T \in \mathbb{F}^{2r} \ : \ \sum_{i\in[2r]} a_i \cdot B_i \in \mathscr{F}_{\text{col}} \right\}, \\ \mathscr{V}_2 &= \left\{ \mathbf{v} = (b, \ldots, b_{2r})^T \in \mathbb{F}^{2r} \ : \ \sum_{i\in[2r]} b_i \cdot B_i \in \mathscr{F}_{\text{row}} \right\}. \end{aligned} \tag{4.2}$$

From bases of $\mathscr{V}_1$ and $\mathscr{V}_2$, and $(B_1, \ldots, B_{2r})$, we get bases of $\mathscr{F}_{\text{col}}$ and $\mathscr{F}_{\text{row}}$ readily. The aspects of the space $\mathscr{P}$ that help in computing $\mathscr{V}_1$ and $\mathscr{V}_2$ are the facts that these are the only two irreducible invariant subspaces of $\mathscr{P}$ and bases of these can be computed from a random element of $\mathscr{P}$. These facts are proved and elaborated upon in Section 4.3.1.2.

[1] Let $E = \alpha_1 B_1 + \cdots + \alpha_{2r} B_{2r}$, where $\alpha_i \in \mathbb{F}$ for every $i \in [2r]$. Then, the coordinate vector of W with respect to $(B_1, \ldots, B_{2r})$ is $(\alpha_1, \ldots, \alpha_{2r})$. Similarly, we define the coordinate vector of K with respect to $(J_1, \ldots, J_r)$.

4.3.1.1 The decomposition algorithm

Now, we present the algorithm and argue its correctness in the next section.

Procedure 10 Decompose-Lie-Algebra(f)

Input: Black box access to f.

Output: Either bases of spaces $\mathscr{V}_1$ and $\mathscr{V}_2$ (as in Equation (4.2)) or 'Fail'.

1. Compute a basis $\{B_1, \ldots, B_{2r}\}$ of $\mathfrak{g}_f$ (Fact 2.16), and form the basis $\{P_{B_1}, \ldots, P_{B_{2r}}\}$ of $\mathscr{P}$.
2. Pick a random element $Q = a_1 P_{B_1} + \cdots + a_{2r} P_{B_{2r}}$ from $\mathscr{P}$, where every a_i is chosen uniformly and independently at random from a fixed subset of $\mathbb{F}$ of size $10n^4$.
3. Compute the characteristic polynomial $h(z)$ of Q.
4. Factor $h(z)$ into irreducible factors over $\mathbb{F}$. Let $h(z) = z^{2(n-1)} \cdot h_1(z) \cdots h_k(z)$, where $z, h_1, \ldots, h_k$ are mutually coprime and irreducible. If h is not as above, output 'Fail'.
5. For every $i \in [k]$, compute a basis of the null space $\mathscr{N}_i$ of $h_i(Q)$, pick a vector $\mathbf{v}$ from the basis of $\mathscr{N}_i$ and compute a basis of $\mathscr{C}_i := \text{closure}_{\mathscr{P}}(\mathbf{v})$ (using Fact 2.18).
6. Remove repetitive spaces from the set $\{\mathscr{C}_1, \ldots, \mathscr{C}_k\}$. After this, if we are *not* left with exactly two spaces $\mathscr{U}_1$ and $\mathscr{U}_2$ then output 'Fail'. Else, output the basis of $\mathscr{U}_1$.

4.3.1.2 Analysis of the procedure

We first analyse $\mathscr{P}$ through the lens of a convenient basis of $\mathfrak{g}_{\mathsf{Det}}$, namely the basis $\{S_1, \ldots, S_{2r}\}$ given in Observation 4.1. After that, we argue the correctness of Algorithm 10 in Lemma 4.3.

For $K \in \mathfrak{g}_{\mathsf{Det}}$, let $\mathbf{w}_K, \mathbf{v}_K \in \mathbb{F}^{2r}$ be the coordinate vectors of K with respect to the ordered bases $(S_1, \ldots, S_{2r})$ and $(J_1, \ldots, J_{2r})$ respectively, where $J_i = A \cdot B_i \cdot A^{-1}$ for every $i \in [2r]$. There is a basis change matrix $H \in \mathrm{GL}(2r, \mathbb{F})$ such that for every $K \in \mathfrak{g}_{\mathsf{Det}}$,

$$\mathbf{v}_K = H \cdot \mathbf{w}_K. \tag{4.3}$$

Recall Q_L from Claim 4.3.2. Let $R_L := H^{-1} \cdot Q_L \cdot H$, for every $L \in \mathfrak{g}_{\mathsf{Det}}$, and

$$\mathscr{R} := \{R_L \ : \ L \in \mathfrak{g}_{\mathsf{Det}}\} = H^{-1} \cdot \mathscr{P} \cdot H. \tag{4.4}$$

Observe that $\{R_{S_1}, \ldots, R_{S_{2r}}\}$ is a basis of $\mathscr{R}$. Also,

$$R_L \cdot \mathbf{w}_K = \mathbf{w}_{[K,L]}, \tag{4.5}$$

for every $L, K \in \mathfrak{g}_{\mathsf{Det}}$. Let us note a few important properties of $\mathscr{R}$.

Observation 4.7 (Structure of matrices in $\mathscr{R}$) *Every $R \in \mathscr{R} \subseteq M_{2r}$ is a block diagonal matrix having two blocks of size $r \times r$ each, i.e., the non-zero entries of R are confined to the entries $\{(S_i, S_j) \ : \ i, j \in [r]\}$ and $\{(S_i, S_j) \ : \ i, j \in [r+1, 2r]\}$.*

Proof: Let $L = L_1 + L_2 \in \mathfrak{g}_{\mathsf{Det}}$, where $L_1 \in \mathscr{L}_{\text{col}}, L_2 \in \mathscr{L}_{\text{row}}$. From Equation (4.5), $R_L \cdot \mathbf{w}_{S_i} = \mathbf{w}_{[S_i,L]} = \mathbf{w}_{[Si,L_1]+[S_i,L_2]}$. Thus, $R_L \cdot \mathbf{w}_{S_i}$ is either $\mathbf{w}_{[S_i,L_1]}$ if $i \in [r]$, or $\mathbf{w}_{[S_i,L_2]}$ if $i \in [r+1, 2r]$. By Observation 4.3, $[S_i, L_1] \in \mathscr{L}_{\text{col}}$ for $i \in [r]$ and $[S_i, L_2] \in \mathscr{L}_{\text{row}}$ for $i \in [r+1, 2r]$. Hence R_L is block diagonal. □

We refer to the two blocks of R as $R^{(1)}$ and $R^{(2)}$, corresponding to $\{S_1, \ldots, S_r\}$ and $\{S_{r+1}, \ldots, S_{2r}\}$, respectively. A snapshot of R is given in Figure 4.1 below. The next observation follows directly from the definition of $\mathscr{R}$.

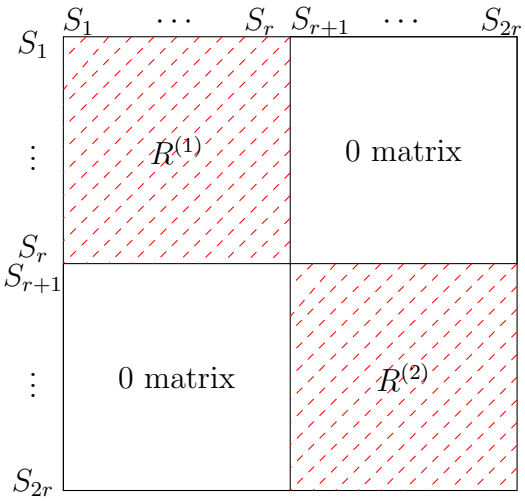


Figure 4.1: Structure of a matrix $R \in \mathscr{R}$

Observation 4.8 (Invariant subspaces of $\mathscr{P}$ and $\mathscr{R}$) *$\mathscr{W}$ is an invariant subspace of $\mathscr{R}$ if and only if $H \cdot \mathscr{W}$ is an invariant subspace of $\mathscr{P}$, where $H \cdot \mathscr{W} = \{H \cdot \mathbf{w} : \mathbf{w} \in \mathscr{W}\}$.*

Observation 4.8 allows us to switch from $\mathscr{P}$ to $\mathscr{R}$ while studying the invariant subspaces of $\mathscr{P}$. The following lemmas on the invariant subspaces of $\mathscr{R}$ are crucial in arguing the correctness of Algorithm 10. Their proofs are given in Sections 4.3.1.3 and 4.3.1.4 respectively. Lemma 4.1 implies that the decomposition of $\mathbb{F}^{2r}$ into irreducible invariant subspaces of $\mathscr{R}$ is unique.

Lemma 4.1 (Irreducible invariant subspaces) *Let $\mathbf{w}_K \in \mathbb{F}^{2r}$ for a nonzero K in $\mathscr{L}_{col}$ or*

in $\mathscr{L}_{row}$. *Then,*

$$\begin{aligned}\text{closure}_{\mathscr{R}}(\mathbf{w}_K) &= \{\mathbf{w}_L \ : \ L \in \mathscr{L}_{col}\} =: \mathscr{W}_1, \quad \text{if } K \in \mathscr{L}_{\text{col}},\\ \text{closure}_{\mathscr{R}}(\mathbf{w}_K) &= \{\mathbf{w}_L \ : \ L \in \mathscr{L}_{row}\} =: \mathscr{W}_2, \quad \text{if } K \in \mathscr{L}_{\text{row}}.\end{aligned}$$

Moreover, $\mathscr{W}_1$ *and* $\mathscr{W}_2$ *are the only two irreducible invariant subspaces of* $\mathscr{R}$, *and* $\mathbb{F}^{2r} = \mathscr{W}_1 \oplus \mathscr{W}_2$.

Lemma 4.2 (Characteristic polynomial) *Let* $R = \sum_{i\in[2r]} \ell_i(a_1, \ldots, a_{2r}) \cdot R_{S_i}$, *where* $\ell_1, \ldots, \ell_{2r}$ *are* $\mathbb{F}$*-linearly independent linear forms and* $a_1, \ldots, a_{2r}$ *are picked uniformly and independently at random from a fixed subset of* $\mathbb{F}$ *of size* $10n^4$. *Then, with high probability, the characteristic polynomial* $h_R(z)$ *of* R *factors as* $z^{2(n-1)} \cdot h_1(z) \cdots h_k(z)$, *where* $z, h_1(z), \ldots, h_k(z)$ *are mutually coprime irreducible polynomials over* $\mathbb{F}$.

The following lemma argues the correctness of Procedure 10.

Lemma 4.3 (Correctness of Procedure 10) *Suppose* f *is equivalent to* Det. *Then, with high probability,* $\{\mathscr{U}_1, \mathscr{U}_2\} = \{\mathscr{F}_{col}, \mathscr{F}_{row}\}$. *In fact, we can assume without loss of generality that* $\mathscr{U}_1 = \mathscr{F}_{col}$ *and* $\mathscr{U}_2 = \mathscr{F}_{row}$.

Proof: In Step 2, we choose a random $Q \in \mathscr{P}$. By Equation (4.4), there is an $R \in \mathscr{R}$ satisfying the following equation

$$R = H^{-1} \cdot Q \cdot H \ = \ a_1 R_{J_1} + \cdots + a_{2r} R_{J_{2r}} \ = \ \ell_1(a_1, \ldots, a_{2r}) \cdot R_{S_1} + \cdots + \ell_{2r}(a_1, \ldots, a_{2r}) \cdot R_{S_{2r}},$$

where $\ell_1, \ldots, \ell_{2r}$ are $\mathbb{F}$ linear forms in $a_1, \ldots, a_{2r}$. As $(R_{J_1}, \ldots, R_{J_{2r}})$ and $(R_{S_1}, \ldots, R_{S_{2r}})$ are $\mathbb{F}$-bases of $\mathscr{R}$, it is easy to verify that $\ell_1(\mathbf{a}), \ldots, \ell_{2r}(\mathbf{a})$ are $\mathbb{F}$-linearly independent. By Lemma 4.2, Step 4 of Procedure 10 holds with high probability. We know from Fact 2.3 that the characteristic polynomial of R is same as the characteristic polynomial of Q. From Observation 4.7, R is a block diagonal matrix with blocks $R^{(1)}$ and $R^{(2)}$. Let $h(z) = g_1(z) \cdot g_2(z)$, where $g_1(z)$ and $g_2(z)$ are the characteristic polynomials of $R^{(1)}$ and $R^{(2)}$, respectively. There are a couple of factors of h, say h_1 and h_2, that divide g_1 and g_2, respectively. In Step 5, we compute the null spaces $\mathscr{N}_1$ and $\mathscr{N}_2$ of $h_1(Q)$ and $h_2(Q)$ respectively. As $h_1(R) = H^{-1} \cdot h_1(Q) \cdot H$ and $h_2(R) = H^{-1} \cdot h_2(Q) \cdot H$, the null spaces of $h_1(R)$ and $h_2(R)$, denoted by $\mathscr{O}_1$ and $\mathscr{O}_2$ respectively, satisfy the following (due to Equation (4.3)): $\mathscr{O}_1 = H^{-1} \cdot \mathscr{N}_1$ and $\mathscr{O}_2 = H^{-1} \cdot \mathscr{N}_2$.

Claim 4.3.3 *If* $\mathbf{w}_K \in \mathscr{O}_1$ *(similarly,* $\mathbf{w}_K \in \mathscr{O}_2$*) then* $K \in \mathscr{L}_{col}$ *(respectively,* $K \in \mathscr{L}_{row}$*).*

Proof: We give the proof for $\mathscr{O}_1$, a similar proof holds for $\mathscr{O}_2$. Recall $\mathbf{w}_K$ is the coordinate vector of K with respect to the ordered basis $(S_1, \ldots, S_{2r})$ of $\mathfrak{g}_{\mathsf{Det}}$. Let $\mathbf{w}_K^{(1)}, \mathbf{w}_K^{(2)} \in \mathbb{F}^r$ be the sub-vectors obtained from $\mathbf{w}_K$ by restricting it to the indices $S_1, \ldots, S_r$ and $S_{r+1}, \ldots, S_{2r}$ respectively. It is sufficient to show $\mathbf{w}_K^{(2)} = 0$ to prove $K \in \mathscr{L}_{\text{col}}$. Let $R \in \mathscr{R}$. Then, R is a block diagonal matrix with $R^{(1)}, R^{(2)}$ as the blocks. By definition, $h_1(R) \cdot \mathbf{w}_K = 0$, which implies

$$h_1(R^{(1)}) \cdot \mathbf{w}_K^{(1)} = h_1(R^{(2)}) \cdot \mathbf{w}_K^{(2)} = 0.$$

As $g_2(z)$ is the characteristic polynomial of $R^{(2)}$, from the Cayley Hamilton theorem (Fact 2.4), $g_2(R^{(2)}) = 0$, which implies

$$g_2(R^{(2)}) \cdot \mathbf{w}_K^{(2)} = 0.$$

Since $h_1(z)$ and $g_2(z)$ are coprime polynomials, there exist $p_1, p_2 \in \mathbb{F}[z]$ such that

$$h_1(z) \cdot p_1(z) + g_2(z) \cdot p_2(z) = 1.$$

This implies

$$h_1(R^{(2)}) \cdot p_1(R^{(2)}) + g_2(R^{(2)}) \cdot p_2(R^{(2)}) = I_r.$$

On multiplying the above equation with $\mathbf{w}_K^{(2)}$, we get $\mathbf{w}_K^{(2)} = 0$ showing $K \in \mathscr{L}_{\text{col}}$. □

In Step 5, we pick a vector $\mathbf{v}$ from a null space, say $\mathscr{N}_1$, and compute $\text{closure}_{\mathscr{P}}(\mathbf{v})$. Clearly, $\mathbf{v} = \mathbf{v}_K$ for some $K \in \mathfrak{g}_{\mathsf{Det}}$. So, $\mathbf{v}_K \in \mathscr{N}_1$ if and only if $\mathbf{w}_K = H^{-1} \cdot \mathbf{v}_K \in \mathscr{O}_1$. As $\mathscr{R} = H^{-1} \cdot \mathscr{P} \cdot H$, Observation 4.8 implies that

$$\begin{aligned}
\text{closure}_{\mathscr{P}}(\mathbf{v}_K) &= H \cdot \text{closure}_{\mathscr{R}}(\mathbf{w}_K) \\
&= H \cdot \mathscr{W}_1 \quad \text{(by Claim 4.3.3 and Lemma 4.1)} \\
&= \mathscr{V}_1 \qquad \text{(by Equations (4.1) and (4.3), as } \mathscr{V}_1 = \{\mathbf{v}_L \ : \ L \in \mathscr{L}_{\text{col}}\}).
\end{aligned}$$

Similarly, if we pick a $\mathbf{v} \in \mathscr{N}_2$ then $\text{closure}_{\mathscr{P}}(\mathbf{v}) = \mathscr{V}_2$. Thus, in Step 6, one of $\mathscr{U}_1$ and $\mathscr{U}_2$ is $\mathscr{V}_1$ and the other is $\mathscr{V}_2$. Finally, we can take $\mathscr{U}_1 = \mathscr{V}_1$ and $\mathscr{U}_2 = \mathscr{V}_2$ without loss of generality: Let $P \in M_m$ be the permutation matrix such that P maps $x_{i,j}$ to $x_{j,i}$ when multiplied to $\mathbf{x}$. Clearly, $P^{-1} = P$. Note that P is a symmetry of Det, i.e.,

$$\mathsf{Det}(\mathbf{x}) = \mathsf{Det}(P\mathbf{x}) \quad \text{and hence} \quad f(\mathbf{x}) = \mathsf{Det}(A\mathbf{x}) = \mathsf{Det}(PA\mathbf{x}).$$

Observe that $\mathscr{L}_{\text{col}} = P^{-1} \cdot \mathscr{L}_{\text{row}} \cdot P$. Hence,

$$\mathscr{F}_{\text{col}} = A^{-1}P^{-1} \cdot \mathscr{L}_{\text{row}} \cdot PA \quad \text{and} \quad \mathscr{F}_{\text{row}} = A^{-1}P^{-1} \cdot \mathscr{L}_{\text{col}} \cdot PA.$$

As the underlying matrix is unknown to Algorithm 10, we can take it to be either A or PA. □

Before proving Lemmas 4.1 and 4.2, we give a comparison of our decomposition algorithm with the algorithm for decomposing *semisimple* Lie algebras given in [de 97] and with the algorithm for decomposition of modules over finite algebras given in [CIK97].

Comparison with decomposition algorithms in [de 97] and [CIK97]. A polynomial time algorithm for decomposition of *semisimple* Lie algebra into direct sum of *simple* Lie subalgebras was given in [de 97]. This algorithm works over fields having characteristic zero. Our decomposition algorithm is a special case of the decomposition algorithm given in [de 97] as $\mathfrak{g}_f$ is a direct sum of $\mathscr{F}_{\text{col}}$ and $\mathscr{F}_{\text{row}}$. However, our algorithm works over any field satisfying mild conditions on its size and characteristic. It is not clear to us how to adapt the algorithm in [de 97] in our case. In [CIK97], a randomized polynomial time algorithm was given to decompose modules over a finite dimensional algebra into indecomposable submodules. This algorithm works over finite fields. As the decomposition of $\mathbb{F}^{2r}$ into irreducible invariant subspaces of $\mathscr{R}$ is unique (follows from Lemma 4.1), the algorithm in [CIK97] can be used to compute $\mathscr{F}_{\text{col}}$ and $\mathscr{F}_{\text{row}}$ over finite fields in randomized polynomial time. However, in case of $\mathbb{Q}$, the algorithm in [CIK97] decomposes modules into submodules in polynomial time, where each submodule is over an extension field of $\mathbb{Q}$.

4.3.1.3 Proof of Lemma 4.1

We first complete the proof of the lemma assuming the following three claims and then prove these. The proof of these claims are given for $\mathscr{L}_{\text{col}}$, and similar proofs hold for $\mathscr{L}_{\text{row}}$. Recall $\mathbf{w}_K$ is the coordinate vector of $K \in \mathscr{L}_{\text{col}}$ with respect to the ordered basis $(S_1, \ldots, S_{2r})$ of $\mathscr{L}_{\text{col}}$.

Claim 4.3.4 *Let the entry indexed by $I_n \otimes E_{ij}$ (similarly, $E_{ij} \otimes I_n$) in $\mathbf{w}_K$ is nonzero for some $i, j \in [n], i \neq j$. Then* $\text{closure}_{\mathscr{R}}(\mathbf{w}_K)$ *contains the unit vector $\mathbf{w}_{I_n \otimes E_{ij}}$ (respectively, $\mathbf{w}_{E_{ij} \otimes I_n}$).*

Claim 4.3.5 *Let $p, q \in [n]$ and $p \neq q$. Then*

$$\text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}}) = \{\mathbf{w}_L \ : \ L \in \mathscr{L}_{col}\} = \mathscr{W}_1.$$

Similarly, $\text{closure}_{\mathscr{R}}(\mathbf{w}_{E_{pq} \otimes I_n}) = \{\mathbf{w}_L \ : \ L \in \mathscr{L}_{row}\} = \mathscr{W}_2$.

Claim 4.3.6 *Suppose $\mathbf{w}_K \in \mathbb{F}^{2r}$ is such that the entry indexed by $I_n \otimes E_\ell$ (similarly, $E_\ell \otimes I_n$) for $\ell \in [2,n]$ is nonzero, and the entries indexed by $I_n \otimes E_{ij}$ (similarly, $E_{ij} \otimes I_n$) are zero for every $i,j \in [n], i \neq j$. Then, for some $i \neq \ell$,*

$$\mathbf{w}_{I_n \otimes E_{i\ell}} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_K) \quad (\text{respectively, } \mathbf{w}_{E_{i\ell} \otimes I_n} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_K)).$$

Claims 4.3.4, 4.3.5 and 4.3.6 imply that for a nonzero $K \in \mathscr{L}_{\text{col}}$, $\text{closure}_{\mathscr{R}}(\mathbf{w}_K) = \mathscr{W}_1$ (similarly, for a nonzero $K \in \mathscr{L}_{\text{row}}$, $\text{closure}_{\mathscr{R}}(\mathbf{w}_K) = \mathscr{W}_2$). This completes the proof of the lemma. Now, we prove these claims one by one.

Proof of Claim 4.3.4. First, consider the following subclaim.

Subclaim 4.3.1 *There is a diagonal matrix $R \in \mathscr{R}$ such that $R(I_n \otimes E_\ell, I_n \otimes E_\ell) = R(E_\ell \otimes I_n, E_\ell \otimes I_n) = 0$ for every $\ell \in [2,n]$, and the remaining $2n^2 - 2n$ diagonal entries are distinct nonzero field elements.*

Let $R \in \mathscr{R}$ be the diagonal matrix given in Subclaim 4.3.1. Consider the following equation in the variables $a_1, \ldots, a_{2n^2-2n}$,

$$\mathbf{w}_{I_n \otimes E_{ij}} = \sum_{i=1}^{2n^2-2n} a_i \cdot R^i \cdot \mathbf{w}_K.$$

As the resulting system is a Vandermonde system, there is a solution over $\mathbb{F}$. Before coming to the proof of Subclaim 4.3.1, we prove some important facts. We state these facts for $\mathscr{L}_{\text{col}}$, similar statements hold for $\mathscr{L}_{\text{row}}$.

Fact 4.2 *Let $S = I_n \otimes E_\ell$ for $\ell \in [2,n]$. Then $R_S \in \mathscr{R}$ is a diagonal matrix that satisfies:*

1. *$R_S^{(2)}$ is an all zero matrix.*
2. *If $S_t = I_n \otimes E_{\ell'}, \ell' \in [2,n]$, then the (S_t, S_t)-th entry of R_S is 0.*
3. *If $S_t = I_n \otimes E_{ij}, i,j \in [n]$ and $i \neq j$, then the (S_t, S_t)-th entry of R_S is*
   - *(a) -1 if $i = 1$ and $j \notin \{1, \ell\}$, or $j = \ell$ and $i \notin \{1, \ell\}$,*
   - *(b) 1 if $i = \ell$ and $j \notin \{1, \ell\}$, or $j = 1$ and $i \notin \{1, \ell\}$,*
   - *(c) -2 if $(i,j) = (1, \ell)$,*
   - *(d) 2 if $(i,j) = (\ell, 1)$,*

(e) 0 *otherwise.*

Proof: Recall that $S = I_n \otimes E_\ell$ for $\ell \in [2,n]$. It follows from Observation 4.9 that $R_S^{(2)} = 0$. To prove other parts of the fact, let us consider a generic element $T = I_n \otimes Z$ in $\mathscr{L}_{\text{col}}$ such that $Z = (a_{ij})_{i,j\in[n]}$. Clearly, $[T,S] = I_n \otimes [Z, E_\ell]$.

$$[Z,E_\ell] = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{\ell 1} & \dots & a_{\ell n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} 1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & -1 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix} - \begin{bmatrix} 1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & -1 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix} \cdot \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{\ell 1} & \dots & a_{\ell n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

From this, we get

$$[Z,E_\ell] = \begin{bmatrix} a_{11} & 0 & \dots & -a_{i\ell} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{\ell 1} & 0 & \dots & -a_{\ell\ell} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & 0 & \dots & -a_{n\ell} & \vdots & 0 \end{bmatrix} - \begin{bmatrix} a_{11} & a_{12} & \dots & a_{i\ell} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -a_{\ell 1} & -a_{\ell 2} & \dots & -a_{\ell\ell} & \dots & -a_{\ell n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix}$$

This implies

$$[Z,E_\ell] = \begin{bmatrix} 0 & -a_{12} & \dots & -2a_{i\ell} & \dots & -a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 2a_{\ell 1} & a_{\ell 2} & \dots & 0 & \dots & a_{\ell n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & 0 & \dots & -a_{n\ell} & \dots & 0 \end{bmatrix}$$

Restricting Z to $E_{\ell'}$ and E_{ij} for different settings of i, j, ℓ' imply the result. □

The following fact immediately follows from Fact 4.2.

Fact 4.3 *Let* $R_1 = \sum_{\ell\in[2,n]} a_\ell \cdot R_{I_n\otimes E_\ell}$ *, where* $a_2, \dots, a_n \in \mathbb{F}$*. Then* R_1 *is a diagonal matrix satisfying the following properties:*

1. $R_1^{(2)}$ *is a zero block.*
2. *If* $S_t = I_n \otimes E_{\ell'}, \ell' \in [2,n]$*, then the* (S_t, S_t)*-th entry of* R_1 *is 0.*
3. *If* $S_t = I_n \otimes E_{ij}, i, j \in [n], i \neq j$*, then the* (S_t, S_t)*-th entry of* R_1 *is*

(a) $a_i - a_j$, *if* $i, j \in [2, n]$,

(b) $-(\sum_{k=2}^{n} a_k + a_j)$ *if* $i = 1$,

(c) $(\sum_{k=2}^{n} a_k + a_i)$ *if* $j = 1$.

In the next fact, we argue the structures of matrices $R_{I_n \otimes E_{ij}}$ for $i, j \in [n], i \neq j$.

Fact 4.4 *Let* $S = I_n \otimes E_{ij}$ *for* $i, j \in [n], i \neq j$. *Then,* R_S *satisfies the following properties:*

1. $R_S^{(2)}$ *is an all zero matrix.*

2. *A column indexed by* $I_n \otimes E_{pq}, p, q \in [n], p \neq q$ *has the following structure:*

   (a) *If* $p \neq j$ *and* $q = i$ *then the column contains exactly one nonzero entry, namely a* 1 *at the row indexed by* $I_n \otimes E_{pj}$.

   (b) *If* $q \neq i$ *and* $p = j$ *then the column contains exactly one nonzero entry, namely a* -1 *at the row indexed by* $I_n \otimes E_{iq}$.

   (c) *If* $(p, q) = (j, i)$ *and* $i, j \neq 1$ *then the column has exactly two nonzero entries, namely a* 1 *and a* -1 *at the rows indexed by* $I_n \otimes E_i$ *and* $I_n \otimes E_j$ *respectively.*

   (d) *If* $(p, q) = (j, i)$ *and* $j = 1$ *(similarly,* $(p, q) = (j, i)$ *and* $i = 1$*) then the column has exactly one nonzero entry, a* 1 *(respectively, a* -1*) at the row indexed by* $I_n \otimes E_i$ *(respectively,* $I_n \otimes E_j$*).*

   (e) *Otherwise the entire column is zero.*

3. *A column indexed by* $I_n \otimes E_\ell, \ell \in [2, n]$ *has the following structure:*

   (a) *If* $i, j \neq 1$, *and* $\ell = i$ *then the column has exactly one nonzero entry, namely a* -1 *at the row indexed by* $I_n \otimes E_{ij}$.

   (b) *If* $i, j \neq 1$, *and* $\ell = j$ *then the column has exactly one nonzero entry, namely a* 1 *at the row indexed by* $I_n \otimes E_{ij}$.

   (c) *If* $i = 1$ *and* $\ell = j$ *then the column has exactly one nonzero entry, namely a* 2 *at the row indexed by* $I_n \otimes E_{ij}$. *If* $i = 1$ *and* $\ell \neq j$, *then the column exactly one nonzero entry, a* 1 *at the row indexed by* $I_n \otimes E_{ij}$.

   (d) *If* $j = 1$ *and* $\ell = i$, *then it has exactly one nonzero entry, a* -2 *at the row indexed by* $I_n \otimes E_{ij}$. *If* $j = 1$ *and* $\ell \neq i$, *then the column contains exactly one nonzero entry, a* -1 *at the row indexed by* $I_n \otimes E_{ij}$.

   (e) *Otherwise the column has all zero entries.*

Proof: First we note a useful observation, which follows from the proof of Observation 4.7.

Observation 4.9 *For all* $i \in [r], R^{(2)}_{S_i} = 0$. *Similarly, for all* $i \in [r+1, 2r], R^{(1)}_{S_i} = 0$.

Part 1 follows from Observation 4.9. Let us consider a generic element $T = I_n \otimes Z$ in $\mathscr{L}_{\text{col}}$ such that $Z = (a_{ij})_{i,j\in[n]}$. Clearly, $[T, S] = I_n \otimes [Z, E_{ij}]$. A derivation similar to that in the proof of Fact 4.2, implies the following.

$$[Z, E_{ij}] = \begin{bmatrix} 0 & 0 & \dots & a_{1i} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -a_{ji} & -a_{j2} & \dots & a_{ii} - a_{jj} & \dots & -a_{jn} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{ni} & \dots & 0 \end{bmatrix},$$

where the rows and columns other than the i-th row and the j-th column are 0. Restricting Z to E_{pq} and E_ℓ for various settings of p, q, ℓ imply the result. □

Now we are ready to prove Subclaim 4.3.1.

Proof of Subclaim 4.3.1. We wish to show that $\mathscr{R}$ contains a diagonal matrix R such that $R(I_n \otimes E_\ell, I_n \otimes E_\ell) = R(E_\ell \otimes I_n, E_\ell \otimes I_n) = 0$ for every $\ell \in [2, n]$, and the remaining $2n^2 - 2n$ entries of R are distinct nonzero field elements. Let

$$R = \sum_{\ell\in[2,n]} (a_\ell \cdot R_{I_n\otimes E_\ell} + b_\ell \cdot R_{E_\ell\otimes I_n}),$$

where $a_\ell, b_\ell \in \mathbb{F}$. From Fact 4.3 (for both $\mathscr{L}_{\text{col}}$ and $\mathscr{L}_{\text{row}}$), R is a diagonal matrix with exactly $2(n-1)$ zero diagonal entries and the remaining diagonal entries are distinct nonzero linear forms in $a_2, \dots, a_n$ and $b_2, \dots, b_n$ (as $char(\mathbb{F}) \neq 2$). As $|\mathbb{F}| > \binom{2n^2-2n}{2}$, the Schwartz-Zippel lemma implies that if we substitute $a_2, \dots, a_n$ and $b_2, \dots, b_n$ randomly from a fixed subset of $\mathbb{F}$ of size $10n^4$, then R has the desired property. ■

This also completes the proof of Claim 4.3.4.

Proof of Claim 4.3.5. We would show that the vectors $\mathbf{w}_{S_1}, \dots, \mathbf{w}_{S_r}$ are in closure$_{\mathscr{R}}(\mathbf{w}_{I_n\otimes E_{pq}})$. The three observations below follow from the structure of matrices in $\mathscr{R}$ mentioned in Fact 4.4.

1. If $S = I_n \otimes E_{qj}$, where $j \neq p$ then $R_S \cdot \mathbf{w}_{I_n\otimes E_{pq}} = \mathbf{w}_{I_n\otimes E_{pj}}$. (from Fact 4.4 item 2(a))
2. If $S = I_n \otimes E_{ip}$, where $i \neq q$ then $R_S \cdot \mathbf{w}_{I_n\otimes E_{pq}} = -\mathbf{w}_{I_n\otimes E_{iq}}$. (from Fact 4.4 item 2(b))

3. If $q \neq 1, p = 1$ then for $S = I_n \otimes E_{q1}$, $R_S \cdot \mathbf{w}_{I_n \otimes E_{pq}} = \mathbf{w}_{I_n \otimes E_q}$. Similarly, if $p \neq 1, q = 1$ then for $S = I_n \otimes E_{1p}$, $R_S \cdot \mathbf{w}_{I_n \otimes E_{pq}} = -\mathbf{w}_{I_n \otimes E_p}$. (From Fact 4.4 item 2(d))

These properties immediately imply that

$$\begin{aligned}
\mathbf{w}_{I_n \otimes E_{pj}} &\in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}}) \quad \text{for } j \in [n], j \neq p, \\
\mathbf{w}_{I_n \otimes E_{iq}} &\in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}}) \quad \text{for } i \in [n], i \neq q, \\
\mathbf{w}_{I_n \otimes E_q} &\in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}}) \quad \text{for } q \neq 1, p = 1, \\
\mathbf{w}_{I_n \otimes E_p} &\in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}}) \quad \text{for } p \neq 1, q = 1.
\end{aligned} \tag{4.6}$$

Now we show that for $S = I_n \otimes E_{st}, \mathbf{w}_S \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}})$ for any $s, t \in [n]$, $s \neq t$. If $(s,t) = (p,q)$, there is nothing to prove. Suppose $(s,t) \neq (p,q)$.

Case 1: Suppose $t \neq p$, then from Equation (4.6), $\mathbf{w}_{I_n \otimes E_{pt}} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}})$. Further, applying Equation (4.6) on $\mathbf{w}_{I_n \otimes E_{pt}}$, we get $\mathbf{w}_{I_n \otimes E_{st}} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pt}})$ as $s \neq t$.

Case 2: Suppose $s \neq q$ then from Equation (4.6), $\mathbf{w}_{I_n \otimes E_{sq}} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}})$. Further, applying Equation (4.6) on $\mathbf{w}_{I_n \otimes E_{sq}}$, we get $\mathbf{w}_{I_n \otimes E_{st}} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{sq}})$ as $s \neq t$.

Case 3: Let $(s,t) = (q,p)$. If $n \geq 3$ then pick a $j \in [n] \backslash \{p,q\}$. By applying Equation (4.6) repeatedly, we have $\mathbf{w}_{I_n \otimes E_{pj}} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}})$, $\mathbf{w}_{I_n \otimes E_{qj}} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pj}})$ and $\mathbf{w}_{I_n \otimes E_{qp}} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{qj}})$. If $n = 2$ then either p or q is 1. Suppose $p = 1$ and $s = q \neq 1$, then $\mathbf{w}_{I_n \otimes E_q} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}})$ (from Equation (4.6)). On applying Fact 4.4 item 3(d), $\mathbf{w}_{I_n \otimes E_{qp}} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_q})$ (note that $char(\mathbb{F}) \neq 2$ as $char(\mathbb{F}) \nmid n(n-1)$).

To complete the proof of the claim, we would like to show that $\mathbf{w}_{I_n \otimes E_\ell} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}})$ for every $\ell \in [2, n]$. It follows from what we have shown so far that $\mathbf{w}_{I_n \otimes E_{1\ell}} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{pq}})$. We conclude from Equation (4.6) that $\mathbf{w}_{I_n \otimes E_\ell} \in \text{closure}_{\mathscr{R}}(\mathbf{w}_{I_n \otimes E_{1\ell}})$. ■

Proof of Claim 4.3.6. Let $K \in \mathscr{L}_{\text{col}}$ and $\mathbf{w}_K = \sum_{p \in [2,n]} a_p \cdot \mathbf{w}_{I_n \otimes E_p}$, where $a_p \in \mathbb{F}$ and $a_\ell \neq 0$. Then, for $i \notin \{1, \ell\}$,

$$R_{I_n \otimes E_{i\ell}} \mathbf{w}_K = \sum_{p \in [2,n]} a_p R_{I_n \otimes E_{i\ell}} \mathbf{w}_{I_n \otimes E_p} = (a_\ell - a_i) \mathbf{w}_{I_n \otimes E_{i\ell}}, \text{ from Fact 4.4 items 3(a) and 3(b), and}$$

$$R_{I_n \otimes E_{1\ell}} \mathbf{w}_K = \sum_{p \in [2,n]} a_p R_{I_n \otimes E_{1\ell}} \mathbf{w}_{I_n \otimes E_p} = (a_2 + \cdots + 2a_\ell + \cdots + a_n) \mathbf{w}_{I_n \otimes E_{1\ell}}, \text{ from Fact 4.4 item 3(c).}$$

If $R_{I_n \otimes E_{i\ell}} \cdot \mathbf{w}_K = 0$ for all $i \in [n] \setminus \{1, \ell\}$ then $a_i = a_\ell$ for all $i \in [n] \setminus \{1, \ell\}$, implying $R_{I_n \otimes E_{1\ell}} \cdot \mathbf{w}_K = n \cdot a_\ell \cdot \mathbf{w}_{I_n \otimes E_{1\ell}}$, which is non-zero as $char(\mathbb{F}) \nmid n$. ■

4.3.1.4 Proof of Lemma 4.2

Let $R = R_L$ for some $L \in \mathfrak{g}_{\mathsf{Det}}$ and e be the maximum power of z dividing the characteristic polynomial $h_R(z)$ of R. Clearly, e is greater than equal to the dimension of the null space of R_L. Let us now lower bound the dimension of this null space. Suppose $\mathbf{w}_K$ is in the null space of R_L, where $K \in \mathfrak{g}_{\mathsf{Det}}$. Then,

$$R_L \cdot \mathbf{w}_K = 0,$$

which along with Equation (4.5) implies $\mathbf{w}_{[K,L]} = 0$. This means $[K, L] = 0$, i.e., K commutes with L. Thus, the dimension of the null space of R_L is exactly equal to the dimension of the subspace of $\mathfrak{g}_{\mathsf{Det}}$, that commute with L. We know that $L = L_1 + L_2$ and $K = K_1 + K_2$, where $L_1, K_1 \in \mathscr{L}_{\text{col}}$ and $L_2, K_2 \in \mathscr{L}_{\text{row}}$. Observation 4.2 implies that $[K, L] = 0$ if and only if $[K_1, L_1] = [K_2, L_2] = 0$. The following claim is helpful in this regard. We first complete the proof of this lemma assuming the claim below and then prove the claim.

Claim 4.3.7 (Dimension of the subspaces of M_n and Z_n commuting with B) *Let $n \in \mathbb{N}$ and $B \in M_n$. Then, the dimension of the space of matrices in M_n (similarly, in Z_n) that commute with B is at least n (respectively, at least $n-1$).*

It follows from Claim 4.3.7 that $e \geq 2(n-1)$. We know

$$R = \sum_{i \in [2r]} \ell_i(a_1, \ldots, a_{2r}) \cdot R_{S_i}.$$

Treat $a_1, \ldots, a_{2r}$ as formal variables. Then, from the above discussion, we get

$$h_R(z) = z^{2(n-1)} \cdot g(z),$$

where the coefficients of $g(z)$, which is a monic polynomial of degree $2n(n-1)$, are polynomials in $a_1, \ldots, a_{2r}$ of degree at most $2r$. As the linear forms $\ell_i(a_1, \ldots, a_{2r}), i \in [2r]$, are $\mathbb{F}$-linearly independent, Subclaim 4.3.1 implies that there is a way to set the $\mathbf{a}$-variables to field constants such that $g(z)$ is square-free, has only linear factors and is not divisible by z. This means that the determinant of the Sylvester matrix (Definition 2.22) of $g(z)$ and $\frac{\partial g(z)}{\partial z}$ is a nonzero polynomial in $\mathbf{a}$-variables of degree at most $8n^4$ (see Fact 2.6 in this context). As g is monic, i.e., the leading constant of g with respect to the underlying variable ordering is 1, and $char(\mathbb{F}) \nmid n(n-1)$, the

dimension of the Sylvester matrix does not change with various settings of the **a**-variables to field constants. Hence, from the Schwartz-Zippel lemma, if we plug $a_1, \ldots, a_{2r}$ with random values from a subset of $\mathbb{F}$ of size $10n^4$, then with high probability the characteristic polynomial $h_R(z)$ factors as

$$h_R(z) = z^{2(n-1)} \cdot h_1(z) \cdots h_k(z),$$

where $z, h_1, \ldots, h_k$ are mutually coprime irreducible polynomials over $\mathbb{F}$. ■

Proof of Claim 4.3.7. Let $\overline{\mathbb{F}}$ be the algebraic closure (Definition 2.14) of $\mathbb{F}$ and $\lambda_1, \ldots, \lambda_t$ be distinct eigenvalues of B, where for $i \in [t]$, λ_i appears n_i times. Then, it follows from Fact 2.5 that there exists a $G \in \mathrm{GL}(n, \overline{\mathbb{F}})$ such that $B = G \cdot J \cdot G^{-1}$, where $J = \mathrm{diag}(J_1, \ldots, J_t)$, J_i is the $n_i \times n_i$ Jordan block corresponding to λ_i for every $i \in [t]$, and $n_1 + \cdots + n_t = n$. For a fixed $i \in [t]$, the Jordan block $J_i \in M_{n_i}(\overline{\mathbb{F}})$ looks like

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_i & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \ldots & \lambda_i & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda_i \end{bmatrix}. \tag{4.7}$$

Let $\mathscr{S}, \tilde{\mathscr{S}}$ be the spaces of $n \times n$ matrices that commute with B, J over $\mathbb{F}$ and $\overline{\mathbb{F}}$ respectively. We claim that $\tilde{\mathscr{S}} = G^{-1} \cdot \mathscr{S} \cdot G$. This is so because if $S \in \mathscr{S}$ then $SB = BS$. Using $B = G \cdot J \cdot G^{-1}$, we get $G^{-1}SG \cdot J = J \cdot G^{-1}SG$. Thus, $G^{-1}SG \in \tilde{\mathscr{S}}$. Thus, $G^{-1} \cdot \mathscr{S} \cdot G \subseteq \tilde{\mathscr{S}}$. Similarly, it is easy to show that $\tilde{\mathscr{S}} \subseteq G^{-1} \cdot \mathscr{S} \cdot G$. Hence, $\tilde{\mathscr{S}} = G^{-1} \cdot \mathscr{S} \cdot G$. Thus, to prove the claim, it is sufficient to show that the dimension $\tilde{S}$ is at least n. The structure of J_i given above implies

$$J_i = \lambda_i \cdot I_{n_i} + N_i,$$

where N_i is a nilpotent matrix[1] and looks like

$$N_i = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{bmatrix}.$$

[1] $A \in M_n$ is said to be nilpotent if there exists an $r \in \mathbb{N}^{\times}$ such that $A^r = \mathbf{0}$, where $\mathbf{0}$ is the all zero matrix.

It is easy to see that $I_{n_i}, N_i, \ldots, N_i^{n_i-1}$ are $\overline{\mathbb{F}}$-linearly independent and they commute with J_i. Since J is a block diagonal matrix, the dimension of the space of matrices commuting with J over $\overline{\mathbb{F}}$ is at least $\sum_{i\in[t]} n_i = n$. This proves that the dimension of the space of matrices in M_n that commutes with B is at least n.

Let $B_1, \ldots, B_s$ be a basis of the space of matrices in $M_n(\mathbb{F})$ commuting with B. We are interested in the space $\mathscr{C}$ of traceless matrices that commute with B. Then,

$$\mathscr{C} := \left\{ a_1B_1 + \cdots + a_sB_s \ : \ a_1, \ldots, a_s \in \mathbb{F} \quad \text{and} \quad \text{trace}\left(\sum_{i\in[s]} a_iB_i\right) = 0 \right\}.$$

Observe that the dimension of $\mathscr{C}$ is $s-1$, which is at least $n-1$ as $s \geq n$. ■

4.3.2 Invoking FMAI

In this section, we argue the correctness of Steps 2 - 6 of Algorithm 9. Recall the FMAI problem from Section 2.2.4. An algorithm for FMAI takes input an ordered basis $(L_1, \ldots, L_m)$ of an $\mathbb{F}$-algebra $\mathscr{A} \subseteq M_{n^2}$ such that $\mathscr{A}$ is isomorphic as an $\mathbb{F}$-algebra to M_n, and outputs a $\mathbb{F}$-algebra isomorphism $\phi : \mathscr{A} \to M_n$ in the form of an ordered basis $(C_1, \ldots, C_m)$ of M_n, where $C_i = \phi(L_i)$ for $i \in [m]$. Recall $m = n^2$.

If f is not equivalent to Det then it can be detected with high probability by checking if $f(\mathbf{a}) = b \cdot \mathsf{Det}(M^{-1}\mathbf{a})$ at a random point $\mathbf{a} \in_r S^m$, where $S \subseteq \mathbb{F}$ is sufficiently large. So, assume that $f = \mathsf{Det}(A\mathbf{x})$ for some $A \in \mathrm{GL}(m, \mathbb{F})$. The correctness of Algorithm 10 ensure that $\mathscr{U}_1 = \mathscr{F}_{\text{col}}$ without loss of generality. Step 2 can be executed efficiently by checking if $U_iU_j \in \langle U_1, \ldots, U_r\rangle$ for $i, j \in [r]$. Observation 4.5 implies that $\mathscr{A}, M_n$ are isomorphic $\mathbb{F}$-algebras, i.e., $L_i = A^{-1} \cdot (I_n \otimes B_i') \cdot A$ for every $i \in [m]$, where $\{B_1', \ldots, B_m'\}$ is a basis of M_n. In Step 3, the FMAI oracle returns an $\mathbb{F}$-algebra isomorphism $\phi : \mathscr{A} \to M_n$ such that $\{C_i := \phi(L_i) : i \in [m]\}$ is an $\mathbb{F}$-basis of M_n. Consider the following claim.

Claim 4.3.8 *There exists an $S \in \mathrm{GL}(n, \mathbb{F})$ such that $B_i' = S^{-1} \cdot C_i \cdot S$ for every $i \in [m]$.*

Proof: Recall $L_i = A^{-1} \cdot (I_n \otimes B_i') \cdot A$, for $i \in [m]$, where $\{L_1, \ldots, L_m\}$ and $\{B_1', \ldots, B_m'\}$ are bases of $\mathscr{A}$ and M_n respectively. Consider the following $\mathbb{F}$-algebra isomorphism from M_n to $\mathscr{A}$

$$\begin{aligned} \tau : M_n \to \quad & \mathscr{A} \\ B' \mapsto \quad & A^{-1} \cdot (I_n \otimes B') \cdot A. \end{aligned}$$

Let $\Gamma = \phi \circ \tau$, where $\phi : \mathscr{A} \to M_n$ is the $\mathbb{F}$-algebra isomorphism constructed in Step 3 of

Algorithm 9. Clearly, $\Gamma : M_n \to M_n$ is an $\mathbb{F}$-algebra isomorphism. On applying the Skolem-Noether theorem (Theorem 2.1) on Γ, we get an $S \in \mathrm{GL}(n, \mathbb{F})$ such that for every $i \in [m]$,

$$B_i = S^{-1} \cdot C_i \cdot S, \tag{4.8}$$

where $\Gamma(B_i) = \phi(L_i) = C_i$. □

The above claim implies that $L_i = A^{-1}(I_n \otimes S^{-1})(I_n \otimes C_i)(I_n \otimes S)A$. Then, it is easy to verify that $(I_n \otimes S)$ and $(I_n \otimes S^{-1})$ are inverses of each others, which implies $L_i = A^{-1}(I_n \otimes C_i)A$. Hence, the matrix M mentioned in Step 4 of the algorithm exists. Consider the linear system defined by the equation $L_i \cdot M = M \cdot (I_n \otimes C_i)$, where the entries of M are taken as variables. Step 4 is executed by picking the free variables of the solution space of the system from a sufficiently large subset of $\mathbb{F}$. Finally, the correctness of Step 6 is argued in the following claim.

Claim 4.3.9 *Suppose $f = \mathsf{Det}(A\mathbf{x})$, where $A \in \mathrm{GL}(m, \mathbb{F})$. Then, $f = \mathsf{Det}((I_n \otimes D)M^{-1}\mathbf{x})$ with high probability.*

Proof: Recall that $L_i = A^{-1} \cdot (I_n \otimes B_i) \cdot A$, where $L_1, \ldots, L_m$ and $B'_1, \ldots, B'_m$ are bases of the $\mathbb{F}$-algebras $\mathscr{A}$ and M_n respectively, and M satisfies the following equation for every $i \in [m]$,

$$L_i \cdot M = M \cdot (I_n \otimes C_i).$$

This implies, for all $i \in [m]$,

$$(I_n \otimes B'_i) \cdot AM = AM \cdot (I_n \otimes C_i). \tag{4.9}$$

We view AM as a block matrix of block size $n \times n$. Let $M_{\ell k} \in M_n$ be the (ℓ, k)-th block of AM. Then, from Equation (4.9), we get the Equation 4.10 for every $\ell, k \in [n]$ and $i \in [m]$:

$$B'_i \cdot M_{\ell k} = M_{\ell k} \cdot C_i \tag{4.10}$$

Observation 4.10 *The block $M_{11} \in M_n$ is an invertible matrix with high probability.*

Claim 4.3.8 implies that $A^{-1} \cdot (I_n \otimes S^{-1})$ is a candidate for M, and for this choice of M, $M_{11} = S^{-1}$. The Schwartz-Zippel lemma then implies the above observation. From Observation 4.10 and Equation (4.10), we get the next equation for every $\ell, k \in [n]$ and $i \in [m]$,

$$B'_i \cdot M_{\ell k} \cdot M_{11}^{-1} = M_{\ell k} \cdot M_{11}^{-1} \cdot B'_i.$$

As $B'_1, \ldots, B'_m$ is a basis of M_n, the above equation implies that $M_{\ell k} \cdot M_{11}^{-1}$ commutes with every matrix in M_n. Thus, according to Observation 4.11, $M_{\ell k} \cdot M_{11}^{-1} = b_{\ell k} \cdot I_n$, for some $b_{\ell k} \in \mathbb{F}$.

Observation 4.11 *If $C \in M_n$ commutes with every $B \in M_n$ then $C = c \cdot I_n$ for some $c \in \mathbb{F}$.*

Observation 4.11 can be easily proved by considering the basis $\{E_{ij} : i, j \in [n]\}$ of M_n, where E_{ij} is the matrix having (i, j)-th entry 1 and other entries 0. Thus, we get the following

$$AM = G \otimes M_{11} = (G \otimes I_n) \cdot (I_n \otimes M_{11}),$$

where $G = (b_{\ell k})_{\ell, k \in [n]}$. As $f = \mathsf{Det}(A \cdot \mathbf{x})$, we get

$$\begin{aligned} f(M\mathbf{x}) &= \mathsf{Det}(AM\mathbf{x}) \\ &= \mathsf{Det}((G \otimes I_n) \cdot (I_n \otimes M_{11}) \cdot \mathbf{x}) \\ &= \det(G \cdot X \cdot M_{11}^T) \\ &= b \cdot \det(X) \\ &= b \cdot \mathsf{Det}(\mathbf{x}) \\ &= \mathsf{Det}((I_n \otimes D) \cdot \mathbf{x}), \end{aligned}$$

where $D = \mathrm{diag}(b, 1, \ldots, 1) \in M_n$. This implies

$$f(\mathbf{x}) = \mathsf{Det}((I_n \otimes D)M^{-1}\mathbf{x}).$$

□

4.4 Reduction from integer factoring to DET over $\mathbb{Q}$

We first recall Theorem 1.9 and then give a proof. In this section, GRH means the Generalized Riemann Hypothesis.

Theorem 4.5 (IntFact reduces to DET for quadratic forms) *Assuming GRH, there exists a randomized polynomial time reduction from the problem of factoring square-free integers to computing an $A \in \mathrm{GL}(4, \mathbb{Q})$ such that $f = \mathsf{Det}_2(A\mathbf{x})$, provided f is equivalent to Det_2.*

Consider the following result from [Ron87], which is required for the proof of Theorem 4.5.

Fact 4.5 ([Ron87]) *Assuming GRH, there is a randomized polynomial time reduction from the problem of factoring square-free integers to the following problem: Given non-zero $a, b \in \mathbb{Q}$, find rational numbers x, y, z (not all zero) such that $x^2 - ay^2 - bz^2 = 0$, if such a solution exists.*

The following fact cited in [Ron87] is also required for the proof of Theorem 4.5. We give a proof for the sake of completeness.

Fact 4.6 *Let $a, b \in \mathbb{Q}$ be non-zero. Then the equation $x^2 - ay^2 - bz^2 = 0$ has a non-zero rational solution if and only if the equation $x^2 - ay^2 - bz^2 + abw^2 = 0$ has a non-zero rational solution.*

Proof: If $x^2 - ay^2 - bz^2 = 0$ has a non-zero solution then the same solution gets extended to $x^2 - ay^2 - bz^2 + abw^2 = 0$. Now, suppose $x^2 - ay^2 - bz^2 + abw^2 = 0$ has a non-zero solution $(x, y, z, w) \in \mathbb{Q}^4$. Using this, we will construct a non-zero solution of $x^2 - ay^2 - bz^2 = 0$.

Suppose a is a perfect square. Then, we immediately get the following solution for $x^2 - ay^2 - bz^2 + abw^2 = 0$: $z = 0, w = 0, y = 1, x = \sqrt{a}$. Thus, we can assume that none of a and b is a perfect square. We have $x^2 - ay^2 = b(z^2 - aw^2)$. Suppose $z^2 - aw^2 = 0$. As b is not a perfect square, we get $x = y = z = w = 0$. This contradicts the assumption that $x^2 - ay^2 - bz^2 + abw^2 = 0$ has a non-zero solution. Hence, $z^2 - aw^2 \neq 0$. Then, we get

$$\begin{aligned} b &= \frac{(x^2 - ay^2)}{(z^2 - aw^2)} \\ &= \frac{(x^2 - ay^2)(z^2 - aw^2)}{(z^2 - aw^2)^2} \\ &= \frac{(xz + ayw)^2}{(z^2 - aw^2)^2} - \frac{a(wx + yz)^2}{(z^2 - aw^2)^2}. \end{aligned}$$

The above equation can be rewritten as $x_1^2 - ay_1^2 - bz_1^2 = 0$, where $x_1 = \frac{(xz+ayw)}{(z^2-aw^2)}, y_1 = \frac{a(wx+yz)}{(z^2-aw^2)}$ and $z_1 = 1$. As $b \neq 0$, we get that $x^2 - ay^2 - bz^2 = 0$ has a non-zero rational solution. □

The proof of Theorem 4.5 follows immediately from the Lemma 4.4 and Fact 4.5.

Lemma 4.4 *Let $a, b \in \mathbb{Q}^\times, \mathbf{x} = \{x_{1,1}, \ldots, x_{2,2}\}$ and $f_{a,b} = x_{1,1}^2 - ax_{1,2}^2 - bx_{2,1}^2 + abx_{2,2}^2$. Then, $f_{a,b}(\mathbf{x}) = \mathsf{Det}_2(A\mathbf{x})$ for some $A \in \mathrm{GL}(4, \mathbb{Q})$ if and only if the equation $x^2 - ay^2 - bz^2 = 0$ has a non-zero solution over $\mathbb{Q}$. Moreover, such a solution can be efficiently computed using A.*

Proof: Suppose, $f_{a,b}(\mathbf{x}) = \mathsf{Det}_2(A\mathbf{x})$, where $A \in \mathrm{GL}(4, \mathbb{Q})$. This means

$$f_{a,b}(A^{-1}\mathbf{x}) = x_{1,1}x_{2,2} - x_{1,2}x_{2,1}.$$

Let $\mathbf{a} = (1, 0, 0, 0)$ and $A^{-1} = (\alpha_{i,j})_{i,j\in[4]}$. Then,

$$f_{a,b}(A^{-1}\mathbf{a}) = \alpha_{1,1}^2 - a\alpha_{2,1}^2 - b\alpha_{3,1}^2 + ab\alpha_{4,1}^2 = 0.$$

As $A \in \mathrm{GL}(4,\mathbb{Q})$, there exists an $i \in [4]$ such that $\alpha_{i,1} \neq 0$. This gives a non-zero rational solution for the equation $x^2 - ay^2 - bz^2 + abw^2 = 0$. Then, Fact 4.6 implies that $x^2 - ay^2 - bz^2 = 0$ also has a non-zero solution over $\mathbb{Q}$. Now, suppose $x, y, z \in \mathbb{Q}$, not all zero, such that $x^2 - ay^2 - bz^2 = 0$. We want to construct an $A \in \mathrm{GL}(4,\mathbb{Q})$ such that $f_{a,b}(\mathbf{x}) = \mathsf{Det}_2(A\mathbf{x})$. Note that both y and z can not be simultaneously zero. Otherwise, we get $x = y = z = 0$, which is a contradiction. This implies that either $u^2 - av^2 = b$ or $u^2 - bv^2 = 0$ depending on whether $y \neq 0$ or $z \neq 0$. Assume without loss of generality that $u^2 - av^2 = b$. Let

$$A = \begin{bmatrix} 1 & 0 & u & -av \\ 0 & 1 & v & -u \\ 0 & 1 & -av & au \\ 1 & 0 & -u & av \end{bmatrix}$$

It is easy to verify that $f_{a,b} = \mathsf{Det}_2(A\mathbf{x})$. Now, we show that $A \in \mathrm{GL}(4,\mathbb{Q})$ by arguing that the columns $C_1, \ldots, C_4$ of A are $\mathbb{Q}$-linearly independent. Let $\beta_1, \ldots, \beta_4 \in \mathbb{Q}$ such that

$$\sum_{i \in [4]} \beta_i C_i = 0.$$

Observe that the above equation gives us the following equations:

$$\beta_1 = \beta_2 = 0, u\beta_3 = av\beta_4 \text{ and } v\beta_3 = u\beta_4.$$

From the last two equation, we get either get $\beta_4 = 0$ or $u^2 - av^2 = 0$. Recall that $u^2 - av^2 = b \neq 0$. Hence, $\beta_4 = 0$, which implies $\beta_3 = 0$. Thus, $A \in \mathrm{GL}(4,\mathbb{Q})$. This completes the proof. □

4.5 Reduction from FMAI to DET

This section is devoted to the proof of Theorem 1.10, which we recall below.

Theorem 4.6 (FMAI reduces to DET) *Let $n \in \mathbb{N}$ and $\mathbb{F}$ be a field that satisfies $\mathrm{char}(\mathbb{F}) \nmid n$. There exists an algorithm, which takes input a basis of an $\mathbb{F}$-algebra $\mathscr{A}$, has oracle access to DET over $\mathbb{F}$ and decides if A is isomorphic as an $\mathbb{F}$-algebra to $M_n(\mathbb{F})$ or not using $n^{O(n)}$ many field operations. If the answer is yes, it outputs an $\mathbb{F}$-algebra isomorphism from $\mathscr{A}$ to $M_n(\mathbb{F})$.*

For this proof, we need a structural result about $\mathfrak{g}_{\mathsf{Det}}$ given in the following subsection.

4.5.1 Deteminant characterized by its Lie algebra

The following lemma is a well-known fact over $\mathbb{C}$. Its proof given below holds over any field satisfying $char(\mathbb{F}) \nmid n$. Recall from Fact 4.1 that $\mathfrak{g}_{\mathsf{Det}} = \mathscr{L}_{\text{row}} \oplus \mathscr{L}_{\text{col}}$.

Lemma 4.5 (Characterization of Det by its Lie algebra) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_{1,1}, \ldots, x_{n,n}\}$, $\mathbb{F}$ be a field satisfying $char(\mathbb{F}) \nmid n$ and $f \in \mathbb{F}[\mathbf{x}]$ be a degree n homogeneous polynomial. If $\mathscr{L}_{col} \subseteq \mathfrak{g}_f$ then $f = \alpha \mathsf{Det}_n(\mathbf{x})$ for some $\alpha \in \mathbb{F}$.*

Proof: Recall the definition of the lie algebra of a polynomial (Definition 2.30). Let $B = (b_{(i,j),(k,l)})_{i,j,k,l,\in[n]} \in \mathfrak{g}_f$. Then,

$$\sum_{i,j,k,l\in[n]} b_{((i,j),(k,l))} \cdot x_{k,l} \cdot \frac{\partial f}{\partial x_{i,j}} = 0. \tag{4.11}$$

Since $\mathscr{L}_{\text{col}} \subseteq \mathfrak{g}_f$, we will pick some specific matrices from $\mathscr{L}_{\text{col}}$ to argue that f is a scalar multiple of Det_n. The basis given in Observation 4.1 becomes helpful here. Let $j, l \in [n]$ be distinct and $B = I_n \otimes E_{j,l}$. Then, it follows from Equation (4.11) that for every $j, l \in [n], j \neq l$,

$$\sum_{i\in[n]} x_{i,l} \cdot \frac{\partial f}{\partial x_{i,j}} = 0. \tag{4.12}$$

For $j \in [n]$, let B_j be the matrix where the (j,j)-th entry is 1 and every other entry is 0 and $B = I_n \otimes (B_j - n^{-1} I_n)$. As $char(\mathbb{F}) \nmid n, n^{-1}$ exists in $\mathbb{F}$. Then, it is easy to verify that $B \in \mathscr{L}_{\text{col}}$. Then, Equation (4.11) implies that for every $j \in [n]$,

$$\sum_{i\in n} x_{i,j} \cdot \frac{\partial f}{\partial x_{i,j}} = n^{-1} \left(\sum_{i',j'\in[n]} x_{i',j'} \cdot \frac{\partial f}{\partial x_{i',j'}} \right) = f(\mathbf{x}), \tag{4.13}$$

where the second equation follows from Euler's identity and the fact that n^{-1} is present in $\mathbb{F}$. Let L be the $n \times n$ matrix, where for every $i, j \in [n]$, the (j,i)-th entry is $\frac{\partial f}{x_{i,j}}$. Let $X = (x_{i,j})_{i,j\in[n]}$. Then, it follows from Equations (4.12) and (4.13) that

$$LX = f(\mathbf{x}) \cdot I_n.$$

Hence

$$L = \frac{f(\mathbf{x})}{\mathsf{Det}_n} \cdot (\text{Adj}(X)),$$

where $\mathrm{Adj}(X)$ is the adjoint of X. Then, every entry in $\mathrm{Adj}(X)$ is a degree $n-1$ polynomial. As every entry in L is a homogeneous degree $n-1$ polynomial, Det_n is irreducible, and $\deg(f) = \deg(\mathsf{Det}_n) = n$, we get that $f = \alpha\mathsf{Det}_n(\mathbf{x})$ for some $\alpha \in \mathbb{F}$. □

If we remove the condition $char(\mathbb{F}) \nmid n$ from he above theorem then Det_n not characterized by its Lie algebra. A counter example is $f = x_{1,1}^n + \mathsf{Det}_n(\mathbf{x})$. Observe that over $\mathbb{F}$ satisfying $char(\mathbb{F}) \nmid n$, $\mathfrak{g}_f = \mathfrak{g}_{\mathsf{Det}_n}$. The above lemma immediately implies the following.

Corollary 4.2 *Let $n \in \mathbb{N}, \mathbf{x} = \{x_{1,1}, \ldots, x_{n,n}\}, \mathbb{F}$ be a field satisfying $char(\mathbb{F}) \nmid n$, and $f \in \mathbb{F}[\mathbf{x}]$ be a homogeneous polynomial of degree n. If there exists $A \in \mathrm{GL}(n^2, \mathbb{F})$ such that $A^{-1} \cdot \mathscr{L}_{col} \cdot A \subseteq \mathfrak{g}_f$ then $f = \alpha\mathsf{Det}_n(A\mathbf{x})$ for some $\alpha \in \mathbb{F}$.*

Proof: It is given that $A^{-1} \cdot \mathscr{L}_{\mathrm{col}} \cdot A \subseteq \mathfrak{g}_f$. Let $h = f(A^{-1}\mathbf{x})$. We know from Fact 2.10 that $\mathfrak{g}_h = A^{-1} \cdot \mathfrak{g}_f \cdot A$. Then, $\mathscr{L}_{\mathrm{col}} \subseteq \mathfrak{g}_h$. As $char(\mathbb{F}) \nmid n$, Lemma 4.5 implies that $g = \alpha\mathsf{Det}_n(\mathbf{x})$ for some $\alpha \in \mathbb{F}$, which implies $f = \alpha\mathsf{Det}_n(A\mathbf{x})$. □

We first give the algorithm for Theorem 1.10 and then argue its correctness. Throughout the following discussion, $n \in \mathbb{N}$ is fixed and $char(\mathbb{F}) \nmid n$.

The algorithm

Algorithm 11 Reduce-FMAI-to-DET($\mathscr{A}$)

Input: A basis $\{B_1, \ldots, B_r\}$ of an $\mathbb{F}$-algebra $\mathscr{A} \subseteq M_m$, and access to an algorithm for DET.
Output: If $\mathscr{A}$ is isomorphic to M_n as an $\mathbb{F}$-algebra for some $n \in \mathbb{N}$ then 1 and an $\mathbb{F}$-algebra isomorphism from $\mathscr{A}$ to M_n, otherwise 0.

1. If $r \neq n^2$ for any $n \in \mathbb{N}$, output 0 and halt. Else, rename the basis elements as $B_{1,1}, \ldots, B_{n,n}$.
2. For $i, j \in [n]$, let $L_{i,j} \in M_{n^2}$ be the matrix corresponding to the left-multiplication action of $B_{i,j}$ on $B_{1,1}, \ldots, B_{n,n}$. That is $B_{i,j} \cdot B_{i_2,j_2} = \sum_{i_1,j_1} L_{i,j}((i_1, j_1), (i_2, j_2)) \cdot B_{i_1,j_1}$.
3. Compute a basis for the traceless parts of $L_{i,j}$'s, i.e., compute a basis $\tilde{L}_1, \ldots, \tilde{L}_s$ of the space spanned by $L_{1,1} - \frac{\mathrm{tr}(L_{1,1})}{n^2} I_{n^2}, \ldots, L_{n,n} - \frac{\mathrm{tr}(L_{n,n})}{n^2} I_{n^2}$. If $s \neq n^2 - 1$, output 0 and halt.
4. Find a non-zero homogeneous polynomial $f(\mathbf{x})$ of degree n, satisfying the following equations for every $M \in \{\tilde{L}_1, \ldots, \tilde{L}_{n^2-1}\}$ (these give linear equations in the coefficients of f). If no such non-zero polynomial exists then output 0 and halt.

$$\sum_{i_1,j_1,i_2,j_2 \in [n]} M((i_1, j_1), (i_2, j_2)) \cdot x_{i_2,j_2} \cdot \frac{\partial f}{\partial x_{i_1,j_1}} = 0 \tag{4.14}$$

5. Run DET on f. If it outputs 'Fail' then output 0 and halt. Else, it outputs an $A \in \mathrm{GL}(n, \mathbb{F})$ such that $f(\mathbf{x}) = \mathsf{Det}_n(A\mathbf{x})$.
6. Check if there exist $F_{1,1}, \ldots, F_{n,n} \in M_n$ such that $A \cdot L_{i,j} \cdot A^{-1} = I_n \otimes F_{i,j}$ for all i, j. If yes, output 1 and the isomorphism $\phi(B_{i,j}) = F_{i,j}$ (extended linearly to whole of $\mathscr{A}$). If no, check if there exist $F_{1,1}, \ldots, F_{n,n} \in M_n$ such that $A \cdot L_{i,j} \cdot A^{-1} = F_{i,j} \otimes I_n$ for all i, j. If yes, output 1 and the isomorphism $\phi(B_{i,j}) = F_{i,j}$ (extended linearly to whole of $\mathscr{A}$). If no, output 0.

Analysis of the algorithm

Claim 4.5.1 *Suppose the algebra $\mathscr{A}$ spanned by $B_{1,1}, \ldots, B_{n,n}$ is isomorphic as an $\mathbb{F}$-algebra to M_n for some $n \in \mathbb{N}$. Let $L_{i,j}, i, j \in [n]$ be the matrices computed in Step 2 of the above algorithm. Then, there exist a $K \in \mathrm{GL}(n^2, \mathbb{F})$ and $C_{1,1}, \ldots, C_{n,n} \in M_n$ such that for every $i, j \in [n], L_{i,j} = K^{-1} \cdot (I_n \otimes C_{i,j}) \cdot K$.*

Proof: Recall the matrices $L_{1,1}, \ldots, L_{n,n}$ generated in the algorithm and let $\mathscr{L}$ be the $\mathbb{F}$-algebra generated by these n^2 matrices. It is not difficult to show that $\mathscr{L}$ is isomorphic as an $\mathbb{F}$-algebra to $\mathscr{A}$. As $\mathscr{A}$ is isomorphic to M_n, we get that $\mathscr{L}$ and M_n are isomorphic as $\mathbb{F}$-algebras. $\mathscr{L}$ also contains the identity matrix I_{n^2}. Then, the Skolem-Noether theorem (Theorem 2.1) implies that there exist a $K \in \mathrm{GL}(n^2, \mathbb{F})$ and $C_{1,1}, \ldots, C_{n,n} \in M_n$ such that for every $i, j \in [n], L_{i,j} = K^{-1} \cdot (I_n \otimes C_{i,j}) \cdot K$. □

Correctness of Algorithm 11. Since for every $n \in \mathbb{N}, \dim M_n = n^2$, first step of the algorithm is correct. As $\{B_{1,1}, \ldots, B_{n,n}\}$ is an $\mathbb{F}$-basis of $\mathscr{A}$, the matrices $L_{1,1}, \ldots, L_{n,n}$ mentioned in Step 2 exist. Suppose $\mathscr{A}$ is isomorphic as an $\mathbb{F}$-algebra to M_n for some $n \in \mathbb{N}$. Then, Claim 4.5.1 implies that there exists a $K \in \mathrm{GL}(n^2, \mathbb{F})$ and matrices $C_{1,1}, \ldots, C_{n,n} \in M_n$ such that for every $i, j \in [n], L_{1,j} = K^{-1} \cdot (I_n \otimes C_{i,j}) \cdot K$. As $L_{1,1}, \ldots, L_{n,n}$ is an $\mathbb{F}$-basis of $\mathscr{A}$, it is easy to verify from above that $C_{1,1}, \ldots, C_{n,n}$ is an $\mathbb{F}$-basis of M_n. In the next step, we compute the $\mathbb{F}$-vector space spanned by the traceless parts $\tilde{L}_{i,j}, i, j \in [n]$ of $L_{i,j}, i, j \in [n]$. Then, it follows from the above discussion that $\langle \tilde{L}_1, \ldots, \tilde{L}_{n^2-1} \rangle = K^{-1} \cdot \mathscr{L}_{\mathrm{col}} \cdot K$.

In Step 4, we compute a polynomial f by solving the set of linear equations in the coefficients of f obtained from Equation (4.14). As $K^{-1} \cdot \mathscr{L}_{\mathrm{col}} \cdot K \subseteq g_f$, it follows from Corollary 4.2 that $f(\mathbf{x}) = \alpha \mathsf{Det}_n(K\mathbf{x})$. It is this step which takes $n^{O(n)}$ field operations and dominates the overall time complexity of Algorithm 11. As f is equivalent to Det_n, on invoking Step 5, we get an $A \in \mathrm{GL}(n, \mathbb{F})$ such that $f = \mathsf{Det}_n(A\mathbf{x})$. As $\tilde{L}_1, \ldots, \tilde{L}_{n^2-1}$ spans a Lie algebra of dimension

n^2-1, it follows from the decomposition of $\mathfrak{g}_{\mathsf{Det}_n}$ given in Fact 4.1 that these n^2-1 either span $A^{-1}\cdot\mathscr{L}_{\mathrm{col}}\cdot A$ or $A^{-1}\cdot\mathscr{L}_{\mathrm{row}}\cdot A$. This immediately implies that either of the following conditions should be true:

1. There exist $F_{1,1},\ldots,F_{n,n}\in M_n$ such that $A\cdot L_{i,j}\cdot A^{-1}=I_n\otimes F_{i,j}$ for every $i,j\in[n]$.
2. There exist $F_{1,1},\ldots,F_{n,n}\in M_n$ such that $A\cdot L_{i,j}\cdot A^{-1}=F_{i,j}\otimes I_n$ for every $i,j\in[n]$.

This immediately argues the correctness of Step 6. It is easy to show that whatever the algorithm outputs is actually an $\mathbb{F}$-algebra isomorphism from $\mathscr{A}$ to M_n.

Chapter 5

Equivalence test for regular ROFs

This chapter is devoted to the proof of Theorem 1.11. This is a joint work with Chandan Saha and Bhargav Thankey. There are three sections in this chapter - in the first one, we list some important properties of the Hessian determinant of an ROF, which are used in the equivalence test for regular ROFs, the second section contains the ET algorithm and the last section contains the analysis of the ET algorithm. The properties of the Hessian determinant of an ROF stated in the first section are proved in Chapter 6.

Notations. Let C be a regular ROF (Definition 2.40) over a field $\mathbb{F}$. Let $\mathbf{x} = \text{var}(\mathsf{C})$, where $\text{var}(\mathsf{C})$ refers to the set of variables appearing in C. Then, every variable in $\mathbf{x}$ is attached to a product gate in C. Corollary 2.2 implies that every variable in $\mathbf{x}$ is essential (Definition 2.32) for C. Throughout this section, we will identify a node of C with the polynomial it computes. It follows from Remark 2.4 that we can assume without loss of generality that C has alternate layers of $+$ gates and $\times$ gates, every non-leaf node in C has at least two children, and none of the children of a $\times$ gate is a constant. Further, as C is regular, there are no edge labels in C. We will denote $+$ gates of C with $Q, Q_i, Q_{i,j}$ etc. and $\times$ gates of C with $T, T_i, T_{i,j}$ etc. We say that C is $+$-rooted (similarly, $\times$-rooted) if the topmost gate of C (also called the root of C) is a $+$ gate (respectively, a $\times$ gate). In this chapter, we consider a slightly more general definition of the orbit of a polynomial. We say that an n-variate polynomial f is in the orbit of C, denoted $\text{orb}(\mathsf{C})$, if there exist an $A \in \text{GL}(n, \mathbb{F})$ and a $\mathbf{b} \in \mathbb{F}^n$ such that $f = \mathsf{C}(A\mathbf{x} + \mathbf{b})$.

We recall Theorem 1.11 below. In this theorem, QFE means PE for quadratic forms.

Theorem 5.1 (ET for regular ROFs) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}$, $\mathbb{F}$ be a field satisfying $\text{char}(\mathbb{F}) = 0$ or $\text{char}(\mathbb{F}) \geq n^2$ and $|\mathbb{F}| \geq n^{13}$, and $f \in \mathbb{F}[\mathbf{x}]$ be in the orbit of an unknown regular ROF C. Then, there exists a randomized $\text{poly}(n)$ time algorithm that takes input black-box access to f, has oracle access to QFE over $\mathbb{F}$ and does the following with high probability: it outputs*

an $A \in \mathrm{GL}(n, \mathbb{F})$ *such that* $f(A\mathbf{x}) = \mathtt{C}(PS\mathbf{x} + \mathbf{b})$*, where* $P, S \in \mathrm{GL}(n, \mathbb{F})$ *are permutation and scaling matrices respectively and* $\mathbf{b} \in \mathbb{F}^n$*.*

As f is in the orbit of a regular ROF $\mathtt{C}$ and as $\mathtt{C}$ does not have redundant variables, it follows from Fact 2.12 that f also does not contain any redundant variable. We introduce a useful definition here. We say that an n-variate polynomial g is in the PS-orbit of $\mathtt{C}$, denoted PS-orb($\mathtt{C}$), if there exist a permutation matrix $P \in \mathrm{GL}(n, \mathbb{F})$, a scaling matrix $S \in \mathrm{GL}(n, \mathbb{F})$, and a vector $\mathbf{d} \in \mathbb{F}^n$, such that $g = \mathtt{C}(PS\mathbf{x} + \mathbf{d})$. Using this terminology, we say that the algorithm in Theorem 5.1 outputs an $A \in \mathrm{GL}(n, \mathbb{F})$ such that $f(A\mathbf{x}) \in$ PS-orb($\mathtt{C}$).

In the equivalence test for ROFs (Algorithm 12), we will extensively use some important properties of the Hessian determinant (Definition 2.27) of $\mathtt{C}$, denoted $\det H_{\mathtt{C}}$. In Section 5.1, we state these properties for a canonical ROF (Definition 2.39) and give their proofs in Chapter 6. Since every regular ROF $\mathtt{C}$ by definition is canonical, these properties hold for $\det H_{\mathtt{C}}$.

5.1 The Hessian determinant of an ROF

This section is devoted to some important properties of the Hessian determinant of a canonical ROF, which are needed for Algorithm 12. We only present the statements of these properties here for the sake of completeness and give their proofs in Chapter 6. The reason for pushing these proofs to a separate chapter is their long lengths due to a detailed analysis of the Hessian determinant of a canonical ROF.

Lemma 5.1 (Non-zeroness of $\det(H_{\mathtt{C}})$**)** *Let* $n \in \mathbb{N}$ *and* $\mathbb{F}$ *be a field such that either* $char(\mathbb{F}) = 0$ *or* $\geq n$*. Let* $\mathtt{C} = T_1 + \cdots + T_s + \gamma$ *be a canonical ROF over* $\mathbb{F}$*, where for every* $l \in [s], T_l$ *is a* $\times$*-rooted child of the root node in* $\mathtt{C}$*,* $|\mathrm{var}(T_l)| \leq n, \deg(T_l) \geq 2$*, and* $\gamma \in \mathbb{F}$*. Then, the Hessian determinant of* $\mathtt{C}$ *is non-zero over* $\mathbb{F}$*.*

As $\mathtt{C}$ is an ROF, all T_l's are pairwise variable disjoint, which implies that the Hessian of $\mathtt{C}$ is a block-diagonal matrix, where the blocks on the diagonal are Hessians of $T_1, \ldots, T_s$. Thus

$$\det(H_{\mathtt{C}}) = \prod_{l \in [s]} \det(H_{T_l}). \tag{5.1}$$

Hence, it is sufficient to argue that every $\det(H_{T_l})$ is non-zero over fields of characteristic either zero or greater than equal to n. So, we focus on an arbitrary term $T \in \{T_1, \ldots, T_s\}$. If T is a product of $+$-rooted ROFs such that each of these $+$-rooted ROFs has product-depth at most 1, then we give the complete description of $\det(H_T)$ in Claim 6.4.1 of Chapter 6. Otherwise, we show in Lemma 6.1 of Chapter 6 that $\det(H_T)$ is non-zero whenever $char(\mathbb{F}) = 0$

or $char(\mathbb{F}) \geq n$. This statement may not hold over $\mathbb{F}$, where $|\mathbb{F}| < n$. For example, let $\mathtt{C} = x_1x_2x_3$. Then, $\det(H_{\mathtt{C}}) = 0$ over fields having characteristic two.

We prove the non-zeroness of $\det(H_T)$ by understanding the structure of some *nice* monomials in $\det(H_T)$. We show that these nice monomials have non-zero coefficients when $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) \geq n$. The analysis of the coefficients of these nice monomials is long and involved because of the two reasons: First, the *exact* values of these coefficients are helpful in showing that $\det(H_T)$ is non-zero even over fields of relatively small characteristic. Second, the precise structures of these nice monomials are useful in getting the complete understanding of the essential variables of $\det(H_T)$ in Claim 5.1.2 and Claim 5.1.3, when T is a regular ROF.

In the following claim, we study the factors of the Hessian determinant of $\mathtt{C}$, which are obtained by carefully analysing the structure of the Hessian of $\mathtt{C}$. The knowledge of these factors is very crucially used in Algorithm 12.

Claim 5.1.1 (Factors of $\det(H_{\mathtt{C}})$) *Let $\mathtt{C}$ a canonical ROF over an arbitrary field $\mathbb{F}$.*

1. *Let $x \in \mathbf{x}$ be such that x is a child of a $\times$ gate in $\mathtt{C}$ that computes a polynomial of degree at least 3. Then, x is a factor of $\det(H_{\mathtt{C}})$.*

2. *Let Q be a $+$-rooted sub-ROF of $\mathtt{C}$ and $Q_1, \cdots, Q_m$ the siblings of Q in $\mathtt{C}$, i.e., for every $i \in [m]$, Q_i is either a variable or a $+$-rooted sub-ROF of $\mathtt{C}$, and Q and Q_i have the same parent in $\mathtt{C}$. Let $|\mathrm{var}(Q_1)| + \cdots + |\mathrm{var}(Q_m)| = k$. Then, the multiplicity of Q in $\det(H_{\mathtt{C}})$ is at least $(k-1)$.*

The proof of Claim 5.1.1 follows from Claim 6.2.3 of Section 6.2 in Chapter 6, where the factorization of the Hessian determinant of an arbitrary term T of $\mathtt{C}$ is studied. Corollary 5.1 follows from Corollary 6.1 of Section 6.2 in Chapter 6.

Corollary 5.1 *Let $\mathbb{F}$ be an arbitrary field and $\mathtt{C} = T_1 + \cdots + T_s + \gamma$ a canonical ROF over $\mathbb{F}$, where for every $l \in [s], T_l$ is a $\times$-rooted canonical ROF and $\gamma \in \mathbb{F}$. If $l \in [s]$ is such that T_l computes a polynomial of degree at least 3, then there is a $+$-rooted child Q of T_l such that Q is a factor of $\det(H_{\mathtt{C}})$.*

Now, we discuss the essential variables in the Hessian determinant of a $\times$-rooted regular ROF T. We have a complete understanding of the essential variables of $\det(H_T)$.

Claim 5.1.2 (Variables of $\det(H_T)$ for a regular T) *Let $n \in \mathbb{N}$ and $\mathbb{F}$ be a field satisfying either $char(\mathbb{F}) = 0$ or $\geq n$. Let T be a $\times$-rooted regular ROF over $\mathbb{F}$ such that $|\mathrm{var}(T)| = n$. If $n \geq 3$ then all the variables present in T appear in the Hessian determinant of T.*

Claim 5.1.3 (Essential variables of $\det(H_T)$ for a regular T) *Let $n \in \mathbb{N}$ and $\mathbb{F}$ be a field satisfying either $char(\mathbb{F}) = 0$ or $\geq n$. Let T be a $\times$-rooted regular ROF over $\mathbb{F}$ such that $|\text{var}(T)| = n$. If $n \geq 3$ then all the variables present in T are essential for the Hessian determinant of T.*

Claims 5.1.2 and 5.1.3 are Claims 6.5.3 and 6.5.4 of Section 6.5.7 in Chapter 6 respectively. Claim 5.1.3 is used in Algorithm 12 to make the terms of the input polynomial f variable disjoint (see Section 5.3.1). Its proof uses Claim 5.1.2, Observation 2.5 and Observation 2.6.

Remark 5.1 *After we obtained this involved analysis of the Hessian determinant of $\mathtt{C}$, Bhargav Thankey independently came up with a different analysis, which is much shorter and proves all the results stated here. His analysis is given in Section 3 and its associated appendix of* [GST22]. *The main difference between our and Thankey's proofs is as follows: We analyse coefficients and structures of some explicit monomials in $\det(H_{\mathtt{C}})$, which we call* nice monomials. *The construction of such monomials and analysis of their coefficients make our proof longer. On the other hand, Thankey's proof shows that there exists some high degree monomials (without giving the explicit description of such monomials or their coefficients in $\det(H_{\mathtt{C}})$) having non-zero coefficients in $\det(H_{\mathtt{C}})$.*

However, we feel that the knowledge of the structure and coefficients of nice monomials can be helpful in understanding the Hessian determinant of univariate-substituted ROFs (see Point 1). As the class of univariate-substituted ROFs generalise ROFs, it is natural to ask if we can use the ideas given here to design an efficient ET for univariate-substituted ROFs. As in the ET for regular ROFs given in this chapter, the Hessian determinant of a univariate-substituted ROF can also turn out to be instrumental in designing an ET for this model. See Section 7.3 of Chapter 6 for motivations to study ET for univariate-substituted ROFs.

5.2 Equivalence test

We start by giving an overview of the equivalence test before stating it formally. Recall Fact 2.17, Claims 2.2.2 and 2.2.3 from Chapter 2. These would be used in the algorithm. We are given that f is in the orbit of an unknown regular ROF $\mathtt{C}$. We can assume without loss of generality that $\mathtt{C}$ is a $+$-rooted regular ROF. This is so because we can reduce an ET for a $\times$-rooted ROF to an ET for a $+$-rooted ROF as follows: Suppose $\mathtt{C} = Q_1 \cdots Q_s$, where for every $k \in [s], Q_k$ is either a variable or a $+$-rooted regular ROF. Let $f \in \text{orb}(\mathtt{C})$, i.e., there exist a $B \in \text{GL}(n, \mathbb{F})$ and a $\mathbf{d} \in \mathbb{F}^n$ such that $f = \mathtt{C}(B\mathbf{x} + \mathbf{d})$. Let $\widehat{Q}_k = Q_k(B\mathbf{x} + \mathbf{d})$ for every $k \in [s]$. Using the algorithm in Claim 2.2.2, we compute a matrix $A_0 \in \text{GL}(|\mathbf{x}|, \mathbb{F})$ such

that $\widehat{Q}_1(A_0\mathbf{x}), \ldots, \widehat{Q}_s(A_0\mathbf{x})$ are variable disjoint. For $k \in [s]$, let $\mathbf{x}_k = \mathrm{var}(\widehat{Q}_k(A_0\mathbf{x}))$. Now, we compute black-box access to irreducible factors of $f(A_0\mathbf{x})$ using Fact 2.17. It follows from Observation 2.7 that after factorization, we get black-box access to $\alpha_k\widehat{Q}_k(A_0\mathbf{x}), k \in [s]$, where every $\alpha_k \in \mathbb{F}^\times$ and $\alpha_1 \cdots \alpha_s = 1$. Now, suppose we have an ET for $+$-rooted regular ROFs, which when invoked on $\alpha_k\widehat{Q}_k(A_0\mathbf{x})$, returns an $A_k \in \mathrm{GL}(|\mathbf{x}_k|, \mathbb{F})$ for every $k \in [s]$, such that $\alpha_k\widehat{Q}_k(A_0(A_k\mathbf{x}_k, \mathbf{x} \setminus \mathbf{x}_k)) \in \text{PS-orb}(Q_k)$. Let $A = \mathrm{diag}(A_1, \ldots, A_s)$. Then, $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ and it is easy to that $f(A_0A\mathbf{x}) \in \text{PS-orb}(\mathtt{C})$.

5.2.1 An overview of the algorithm

Let $\mathtt{C} = T_1 + \cdots + T_s + \gamma$, where for every $k \in [s], T_k$ is a $\times$-rooted regular ROF and $\gamma \in \mathbb{F}$. Let $f = \mathtt{C}(B\mathbf{x} + \mathbf{d})$, where $B \in \mathrm{GL}(n, \mathbb{F})$ and $\mathbf{d} \in \mathbb{F}^n$. Then, $f = \widehat{T}_1 + \cdots + \widehat{T}_s + \gamma$, where for every $k \in [s], \widehat{T}_k = T_k(B\mathbf{x} + \mathbf{d})$. We refer to $T_1, \ldots, T_s$ and $\widehat{T}_1, \ldots, \widehat{T}_s$ as the terms of $\mathtt{C}$ and f respectively. Without loss of generality, assume that there exists an $s_1 \in [s]$ such that for every $k \in [s_1], \deg(T_k) = \deg(\widehat{T}_k) \geq 3$ and for every $k \in \{s_1 + 1, \ldots, s\}, \deg(T_k) = \deg(\widehat{T}_k) = 2$. Let $q := T_{s_1+1} + \cdots + T_s$ and $\widehat{q} = \widehat{T}_{s_1+1} + \cdots + \widehat{T}_s$. Then, we call q and $\widehat{q}$ as the quadratic terms of $\mathtt{C}$ and f respectively. The algorithm has two phases, which are described below:

Phase 1: Making terms of f variable disjoint. The objective of this phase is to compute an $A_0 \in \mathrm{GL}(n, \mathbb{F})$ such that $\widehat{T}_1(A_0\mathbf{x}), \ldots, \widehat{T}_s(A_0\mathbf{x})$ are variable disjoint (see Procedure 13). This phase further has the following two steps.

1. **Handling non-quadratic terms of f**: We first compute the Hessian determinant of f using Fact 2.14. Lemma 5.1 and Corollary 2.1 ensure that $\det(H_f) \neq 0$ over $\mathbb{F}$. Let $k \in \{s_1 + 1, \ldots, s\}$. Then, it is easy to see that $\det(H_{T_k}) \in \mathbb{F}^\times$. This along with Equation (5.1) and Corollary 2.1 implies that

$$\det(H_f) = \alpha \cdot \prod_{k \in [s_1]} \det(H_{T_k})(B\mathbf{x} + \mathbf{d}),$$

   where $\alpha \in \mathbb{F}^\times$. Let $\mathtt{C}_1 = T_1 + \cdots + T_{s_1}$ and $f_1 = \widehat{T}_1 + \cdots + \widehat{T}_{s_1}$. Since $\deg(T_k) \geq 3$ for every $k \in [s_1]$, it follows from Claim 5.1.3 that the number of essential variables in $\det(H_{\mathtt{C}_1})$ is equal to the number of variables appearing in $\mathtt{C}_1$. Then, using the basic approach mentioned in Section 1.4.3 (see Claim 2.2.3), we compute an $A_0 \in \mathrm{GL}(n, \mathbb{F})$ such that $\widehat{T}_1(A_0\mathbf{x}), \ldots, \widehat{T}_{s_1}(A_0\mathbf{x})$ are variable disjoint.

2. **Handling the quadratic term of f**: In the previous step, we computed an $A_0 \in$

$\mathrm{GL}(n,\mathbb{F})$ such that $f_1(A_0\mathbf{x})$ is a sum of variable disjoint polynomials. However, the terms of $\widehat{q}(A_0\mathbf{x})$ need not be variable disjoint. At this point, we invoke QFE over $\mathbb{F}$ and compute an $A_0' \in \mathrm{GL}(n,\mathbb{F})$ such that A_0' maps every variable in $f_1(A_0\mathbf{x})$ to itself and $\widehat{q}(A_0'\mathbf{x}) = (y_1+\beta_1)(y_2+\beta_2)+\cdots+(y_{m-1}+\beta_{m-1})(y_m+\beta_m)$, where for every $i \in [m]$, $y_i \notin \mathrm{var}(f_1(A_0\mathbf{x}))$ and $\beta_i \in \mathbb{F}$. We update $A_0 = A_0 \cdot A_0'$. Then,

$$f(A_0\mathbf{x}) = \sum_{k\in[s_1]} \widehat{T}_k(A_0\mathbf{x}) + (y_1+\beta_1)(y_2+\beta_2)+\cdots+(y_{m-1}+\beta_{m-1})(y_m+\beta_m),$$

where for every $k \neq k' \in [s], \mathrm{var}(\widehat{T}_k(A_0\mathbf{x})) \cap \mathrm{var}(\widehat{T}_{k'}(A_0\mathbf{x})) = \emptyset$. Let $f' := f(A_0\mathbf{x})$ and for $k \in [s], T_k' := \widehat{T}_k(A_0\mathbf{x})$.

Phase 2: Recursively performing ET on the factors of the terms of f'. The objective of this phase is to first get black-box access to a term T_k' of f' using *only one black-box query* to f and if $T_k' = Q_{k,1}' \cdots Q_{k,m_k}'$ then obtain black-box access to a factor $Q_{k,j}'$ using *only one black-box query* to T_k'. The algorithm first learns the variable sets of $T_1', \ldots, T_{s_1}'$, say $\mathbf{z}_1, \ldots, \mathbf{z}_{s_1}$. Then, it picks a $k \in [s_1]$ and for every $k' \in [s_1] \setminus \{k\}$, it substitutes every variable in $\mathbf{z}_{k'}$ equal to 0. Thus, we get black-box access to

$$g := T_k' + \gamma',$$

where $\gamma' \in \mathbb{F}$. As shown in the proof overview of Theorem 1.11 given in Section 1.4.3, we compute γ' using a "good factor" of $\det(H_g)$ and after that we subtract γ' from the black-box of g, which allows us to compute black-box of T_k' using only one query to black-box of f. Once we have access to T_k', we factorize it using Fact 2.17, and then recursively solve the problem for factors of T_k' one by one as each of these factors is an instance of f and has product-depth less than the product-depth of f.

5.2.2 The algorithm

We now give a formal description of the algorithm.

Algorithm 12 Find-Equivalence($f(\mathbf{x})$)

Input: Black-box access to an n-variate polynomial f in the orbit of an unknown $+$-rooted regular ROF C such that every $x \in \mathbf{x}$ is essential for f.

Output: An $A \in \mathrm{GL}(n,\mathbb{F})$ such that $f(A\mathbf{x}) \in \text{PS-orb}(\mathsf{C})$.

/* The base case. */

1. If $\deg(f) = 1$, return $I_{n\times n}$.

/* Making the terms of f variable disjoint */

2. Let $(A_0, \mathbf{y}, \mathbf{z}_1, \ldots, \mathbf{z}_{s'})$ be the output of Make-Terms-Var-Disjoint(f). Let $\mathbf{z} = \uplus_{k\in[s']}\mathbf{z}_k$.

/* Learning $\mathrm{var}(\widehat{T}_1(A_0\mathbf{x})), \ldots, \mathrm{var}(\widehat{T}_{s_1}(A_0\mathbf{x}))$ */

3. Let $E = \emptyset, V = \{\mathbf{z}_1, \ldots, \mathbf{z}_{s'}\}$ and $G = (V, E)$ be a graph.
4. **for** $i, j \in [s']$ **do**
5. If there exist $z_i \in \mathbf{z}_i$ and $z_j \in \mathbf{z}_j$ such that $\frac{\partial^2 f(A_0\mathbf{x})}{\partial z_i \partial z_j} \neq 0$ then add the edge $(\mathbf{z}_i, \mathbf{z}_j)$ to E.
6. **end for**
7. Let $\mathbf{z}_1, \ldots, \mathbf{z}_{s_1}$ be the variable sets corresponding to the different connected component of G. Let $\mathscr{C} = \{\mathbf{z}_1, \ldots, \mathbf{z}_{s_1}\}$.
8. **for** $k \in [s_1]$ **do**
9. Let $\widehat{T}$ be the output of Compute-Term-Black-Box($f(A_0(\mathbf{z}_k, \mathbf{x} \setminus \mathbf{z}_k = 0))$).

/* Making factors of $\widehat{T}(A_0\mathbf{x})$ variable disjoint */

10. Use algorithm in Fact 2.17 to obtain black-box access to factors $\widehat{Q}_1, \ldots, \widehat{Q}_m$ of $\widehat{T}(A_0\mathbf{x})$.
11. Let $A_{k,0} \in \mathrm{GL}(|\mathbf{z}_k|, \mathbb{F})$ be the output of the Make-Polys-Var-Disjoint($\widehat{Q}_1, \ldots, \widehat{Q}_m$) (Algorithm 3 in Claim 2.2.2). For $l \in [m], \mathbf{z}_{k,l} := \mathrm{var}(\widehat{Q}_l(A_{k,0}\mathbf{z}_k))$ and $\overline{\mathbf{z}}_{k,l} := \mathbf{z}_k \setminus \mathbf{z}_{k,l}$.

/* Performing ET on $\widehat{Q}_1(A_{k,0}\mathbf{z}_k), \ldots, \widehat{Q}_m(A_{k,0}\mathbf{z}_k)$ */

12. **for** $l \in [m]$ **do**
13. Let $\mathbf{a}$ be a size-$|\overline{\mathbf{z}}_{k,l}|$ vector of random field elements and $\widehat{Q}_l = \widehat{T}_k(A_{k,0}(\mathbf{z}_{k,l}, \overline{\mathbf{z}}_{k,l} = \mathbf{a}))$.
14. Let $A_{k,l} \in \mathrm{GL}(|\mathbf{z}_{k,l}|, \mathbb{F})$ be the output of Find-Equivalence($\widehat{Q}_l$).
15. **end for**
16. Construct an $A'_{k,0} \in M_{|\mathbf{z}_k|}(\mathbb{F})$, which maps every $z \in \mathbf{z}_{k,l}$ to $A_{k,l} \circ z$ for every $l \in [m]$.
17. Let $A_k = A_{k,0}A'_{k,0}$.
18. **end for**
19. Construct an $A'_0 \in M_n(\mathbb{F})$ such that $A'_0 \circ z = A_k \circ z, \forall z \in \mathbf{z}_k, k \in [s_1]$ and $A'_0 \circ y = y, \forall y \in \mathbf{y}$.
20. Return $A_0A'_0$.

Here, we give the input-output behaviours of the procedures Make-Terms-Var-Disjoint() (Procedure 13) and Compute-Term-Black-Box() (Procedure 14), which are used as subroutines in the above algorithm. These procedures are formally described in Section 5.3. The procedure

Make-Terms-Var-Disjoint() takes input black-box access to f in the orbit of an unknown $+$-rooted regular ROF such that every variable in $\mathbf{x}$ is essential for f and computes a matrix $A \in \mathrm{GL}(n, \mathbb{F})$ such that $f(A\mathbf{x})$ is a sum of variable disjoint terms. The procedure Compute-Term-Black-Box() takes black-box access to $\widehat{T}_k(A_0\mathbf{x}) + \gamma'$, where $\widehat{T}_k$ is a term of $f', \gamma' \in \mathbb{F}$, and A_0 is the matrix mentioned in Step 2 and outputs black-box access to $\widehat{T}_k(A_0\mathbf{x})$ using only one query to the black-box of f.

5.3 Analysis of the algorithm

The following lemma will establish the correctness of Algorithm 12.

Lemma 5.2 (Correctness of Algorithm 12) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}, \mathbb{F}$ be a field such that $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) \geq n^2$ and $|\mathbb{F}| \geq n^{13}$, and $f(\mathbf{x})$ be in the orbit of a $+$-rooted regular ROF $\mathtt{C}$ such that every $x \in \mathbf{x}$ is essential for f. Let A be the matrix returned by Algorithm 12. Then, there exist a permutation matrix $P \in M_n(\mathbb{F})$, an invertible scaling matrix $S \in M_n(\mathbb{F})$, and a $\mathbf{b} \in \mathbb{F}^n$ such that $f(A\mathbf{x}) = \mathtt{C}(PS\mathbf{x} + \mathbf{b})$.*

We first give the formal description of Procedure Make-Terms-Var-Disjoint() below and argue its correctness in Sections 5.3.1 and 5.3.2. Then, we give Procedure Compute-Term-Black-Box() in Section 5.3.3. After that, we prove Lemma 5.2.

Procedure 13 Make-Terms-Var-Disjoint($f(\mathbf{x})$)

Input. Black-box access to an n-variate polynomial f in the orbit of an unknown $+$-rooted regular ROF such that every $x \in \mathbf{x}$ is essential for f.

Output. $(A, \mathbf{y}, \mathbf{z}_1, \ldots, \mathbf{z}_{s'})$, where $A \in \mathrm{GL}(n, \mathbb{F})$, s.t. the terms of $f(A\mathbf{x})$ are variable disjoint, $\mathbf{y} = \mathrm{var}(\widehat{q}(A\mathbf{x}))$, for $i \neq j \in [s'], \mathbf{z}_i \cap \mathbf{z}_j = \emptyset$ and $\forall \widehat{T}_k$ satisfying $\deg(\widehat{T}_k) \geq 3, \exists J_k \subseteq [s']$, s.t. $\mathrm{var}(\widehat{T}_k(A\mathbf{x})) = \uplus_{i \in J_k} \mathbf{z}_i$.

/* Handling the terms of f computing polynomials of degree at least 3 */

1. Let $h = \det(H_f)$ and $(A_1, \{\mathbf{z}_1, \ldots, \mathbf{z}_{s'}\})$ be the output of Make-Factors-Var-Disjoint(h) (Algorithm 4 in Claim 2.2.3). Let $\mathbf{z} = \mathrm{var}(h(A_1\mathbf{x}))$ and $\mathbf{y} = \mathbf{x} \setminus \mathbf{z}$.
2. If $\mathbf{z} = \mathrm{var}(f(A_1\mathbf{x}))$, return $(A_1, \emptyset, \mathbf{z}_1, \ldots, \mathbf{z}_{s'})$.

/* Handling the terms of f_1 computing degree 2 polynomials */

3. Let $f_1 = f(A_1\mathbf{x})$, $\mathbf{y} = \{y_1, \ldots, y_{2m}\}$, and $\widetilde{q} = y_1y_2 + \cdots + y_{m-1}y_m$.
4. Compute black-box access to the degree-2 homogeneous part q' of f_1 in $\mathbf{y}$-variables.

5. Let $\widehat{A}_2$ be the output of QFE $(q', \widetilde{q})$. Extend $\widehat{A}_2 \in \mathrm{GL}(|\mathbf{y}|, \mathbb{F})$ to $A_2 \in \mathrm{GL}(n, \mathbb{F})$, such that $A_2 \circ \mathbf{z} = \mathbf{z}$ and for every $y \in \mathbf{y}$, A_2 and $\widehat{A}_2$ map y to the same linear form in $\mathbf{y}$-variables.
6. Let $f_2 = f_1(A_2\mathbf{x})$.
7. **for** $y \in \mathbf{y}$ **do**
8. Compute the black-box access to $\frac{\partial f_2}{\partial y}$ and interpolate it. Let $\frac{\partial f_2}{\partial y} = y' + \ell_{\mathbf{z},y'} + \alpha_{y'}$, where $y' \in \mathbf{y} \setminus \{y\}, \ell_{\mathbf{z},y'} \in \mathbb{F}[\mathbf{z}]$ is a linear form and $\alpha_{y'} \in \mathbb{F}$.
9. **end for**
10. Compute $A_3 \in \mathrm{GL}(n, \mathbb{F})$, that maps every $y' \in \mathbf{y}$ to $y' - \ell_{\mathbf{z},y'}$ and every $z \in \mathbf{z}$ to itself.
11. Return $(A_1A_2A_3, \mathbf{y}, \mathbf{z}_1, \ldots, \mathbf{z}_{s'})$.

Recall $f = \widehat{T}_1 + \cdots + \widehat{T}_s + \gamma$ and $\widehat{q} = \widehat{T}_{s_1+1} + \cdots + \widehat{T}_s$, where $\widehat{q}$ is the quadratic term of f. The objective of the above procedure is to compute an $A_0 \in \mathrm{GL}(n, \mathbb{F})$ such that $\widehat{T}_1(A_0\mathbf{x}), \ldots, \widehat{T}_s(A_0\mathbf{x})$ become variable disjoint. Let $I_1 = [s_1]$ and $I_2 = \{s_1 + 1, \ldots, s\}$. Then, we know that for every $k \in I_1$, $\widehat{T}_k$ computes a polynomial of degree at least 3 and for every $k \in I_2, \widehat{T}_k$ computes a degree 2 polynomial. In Section 5.3.1 we first show how to make the terms of f corresponding to I_1 variable disjoint and then show in Section 5.3.2 how to handle the terms of f corresponding to I_2 by using oracle access to QFE over $\mathbb{F}$.

5.3.1 Making terms variable disjoint

In the procedure, we first compute $h = \det(H_f)$. Lemma 5.1 and Corollary 2.1 imply that $h \neq 0$. Then, it follows from Claim 5.1.2 and Fact 2.7 that for every $k \in I_1, \det\left(H_{\widehat{T}_k}\right)$ is a non-constant polynomial and it is easy to see that for every $k \in I_2, \det\left(H_{\widehat{T}_k}\right) \in \mathbb{F}^\times$. Thus, when we invoke Algorithm 4 in Step 1 and factorize h inside this algorithm, the non-constant irreducible factors of h are only contributed by $\det\left(H_{\widehat{T}_k}\right)$, $k \in I_1$. It follows from Claim 5.1.3 that for all $k \in I_1$, all variables appearing in $\det\left(H_{\widehat{T}_k}\right)$ are essential. So, Claim 2.2.3 implies that for every $k, k' \in I_1, k \neq k', \det\left(H_{\widehat{T}_k}\right)(A_1\mathbf{x})$ and $\det\left(H_{\widehat{T}_{k'}}\right)(A_1\mathbf{x})$ are variable disjoint, where $A_1 \in \mathrm{GL}(n, \mathbb{F})$ is the matrix obtained in Step 1. The following observation ensures that A_1 makes $\widehat{T}_k, k \in I_1$ variable disjoint.

Observation 5.1 *For every $k, k' \in I_1$, $k \neq k'$, $\widehat{T}_k(A_1\mathbf{x})$ and $\widehat{T}_{k'}(A_1\mathbf{x})$ are variable disjoint. Further for every $k \in I_1$, $\widehat{T}_k(A_1\mathbf{x})$ has no redundant variables.*

Proof: Fix $k \in I_1$ arbitrarily. Let $\mathbf{x}_k := \mathrm{var}(\det(H_{T_k}))$ and $\mathbf{z}_k := \mathrm{var}\left(\det\left(H_{\widehat{T}_k}\right)(A_1\mathbf{x})\right)$. We know from Claim 5.1.2 that $\mathbf{x}_k$ is the set of variables appearing in T_k and as T_k is a regular ROF, Observation 2.8 and Claim 5.1.3 imply that every variable in $\mathbf{x}_k$ is essential for $\det(H_{T_k})$.

Claim 2.2.3 implies that $\mathbf{z}_k$ is a set of essential variables of $\det(H_{\widehat{T}_k})(A_1\mathbf{x})$. Then, it follows from Fact 2.12 that $|\mathbf{x}_k| = |\mathbf{z}_k|$ for every $k \in I_1$.

Let $k \in I_1$. We know $\widehat{T}_k = T_k(B\mathbf{x} + \mathbf{d})$, which implies $\widehat{T}_k(A_1\mathbf{x}) = T_k(BA_1\mathbf{x} + \mathbf{d})$. We know $\det(H_{T_k}) \in \mathbb{F}[\mathbf{x}_k]$ and $\det\left(H_{\widehat{T}_k}\right)(A_1\mathbf{x}) \in \mathbb{F}[\mathbf{z}_k]$, which implies $\det(H_{T_k})(BA_1\mathbf{x} + \mathbf{d}) \in \mathbb{F}[\mathbf{z}_k]$. As $|\mathbf{z}_k| = |\mathbf{x}_k|$ and $\mathbf{x}_k$ is the set of essential variables of T_k, it follows from Observation 2.4 that BA_1 maps every variable in $\mathbf{x}_k$ to a linear form in $\mathbf{z}_k$. As $\mathbf{x}_k = \mathrm{var}(T_k)$ and BA_1 maps every variable in $\mathbf{x}_k$ to $\mathbf{z}_k$, clearly $\mathrm{var}\left(\widehat{T}_k(A_1\mathbf{x})\right) = \mathbf{z}_k$. Because every variable in $\mathbf{x}_k$ is essential for T_k, we get that $\mathbf{z}_k$ is the set of essential variables of $\widehat{T}_k(A_1\mathbf{x})$. Since for every $k, k' \in I_1, k \neq k', \mathbf{z}_k \cap \mathbf{z}_{k'} = \emptyset, \widehat{T}_k(A_1\mathbf{x})$ and $\widehat{T}_{k'}(A_1\mathbf{x})$ are variable disjoint. □

For $k \in I_1$, let $\mathbf{z}_k = \mathrm{var}\left(\det\left(H_{\widehat{T}_k}\right)(A_1\mathbf{x})\right)$. As noted in the proof of Observation 5.1, $\widehat{T}_k(A_1\mathbf{x})$ computes a polynomial in $\mathbb{F}[\mathbf{z}_k]$. Let $\mathbf{z} = \uplus_{k\in I_1}\mathbf{z}_k$ and $\mathbf{y} = \mathbf{x} \setminus \mathbf{z}$. Then, $\mathbf{z}$ is the set of variables appearing in $\sum_{k\in I_1}\widehat{T}_k(A_1\mathbf{x})$. If $I_1 = [s]$ then $\mathbf{z} = \mathbf{x}$ and we are done. Otherwise, it might happen that there exist a $k' \in I_2$ and $z \in \mathbf{z}$, such that $z \in \mathrm{var}\left(\widehat{T}_{k'}(A_1\mathbf{x})\right)$. We show how to handle the terms corresponding to I_2 in the following section.

5.3.2 Handling the top quadratic term

Now, we handle the terms corresponding to I_2. Recall $f_1 = f(A_1\mathbf{x})$. Then,

$$f_1 = \sum_{k\in I_1}\widehat{T}_k(A_1\mathbf{x}) + \sum_{k\in I_2}(\ell_{k,1,\mathbf{y}} + \ell_{k,1,\mathbf{z}} + \alpha_{k,1})(\ell_{k,2,\mathbf{y}} + \ell_{k,2,\mathbf{z}} + \alpha_{k,2}) + \gamma,$$

where for every $k \in I_2, j \in [2]$, $\ell_{k,j,\mathbf{y}} \in \mathbb{F}[\mathbf{y}]$, $\ell_{k,j,\mathbf{z}} \in \mathbb{F}[\mathbf{z}]$ are linear forms and $\alpha_{k,j} \in \mathbb{F}$.

It is easy to verify that the following observation is true. If not, it can be shown the number of essential variables of f_1 (which is equal to the number of essential variables of f) is strictly less than n, which gives a contradiction as f_1 is in the orbit of an n-variate regular ROF.

Observation 5.2 *The set* $\{\ell_{k,j,\mathbf{y}} : k \in I_2, j \in [2]\}$ *is* $\mathbb{F}$*-linearly independent.*

In Step 4, we compute black-box access to q', which is the homogeneous degree 2 part of f_1 in $\mathbf{y}$-variables. This is done by multiplying every variable in $\mathbf{y}$ with a fresh variable t and then interpolating the coefficient of t^2 from the black-box of f_1. Note that this coefficient is the black-box of q', which is equal to $\sum_{k\in I_2}\ell_{k,1,\mathbf{y}} \cdot \ell_{k,2,\mathbf{y}}$. We rename the set $\mathbf{y} = \{y_1, \ldots, y_m\}$ to $\mathbf{y} = \{y_{k,1}, y_{k,2} : k \in I_2\}$. Let $\widetilde{q} = \sum_{k\in I_2} y_{k,1} \cdot y_{k,2}$. Observation 5.2 implies that we can invoke QFE on $(q', \widetilde{q})$, which returns $\widehat{A}_2 \in \mathrm{GL}(|\mathbf{y}|, \mathbb{F})$, such that

$$q'\left(\widehat{A}_2 \cdot \mathbf{y}\right) = \sum_{k\in I_2}\ell_{k,1,\mathbf{y}} \cdot \ell_{k,2,\mathbf{y}}\left(\widehat{A}_2\mathbf{y}\right) = \sum_{k\in I_2} y_{k,1} \cdot y_{k,2}. \tag{5.2}$$

Let A_2 be the extension of $\widehat{A}_2$ in the following manner: For every $z \in \mathbf{z}$, A_2 maps z to itself and for every $y \in \mathbf{y}$, A_2 and $\widehat{A}_2$ map y to the same linear form in $\mathbf{y}$-variables. Clearly, $A_2 \in \mathrm{GL}(n, \mathbb{F})$. Let $f_2 = f_1(A_2\mathbf{x})$. The following observation is helpful in understanding f_2.

Observation 5.3 *Let $A_2 \in \mathrm{GL}(n, \mathbb{F})$ be the matrix computed in Step 5. Then, for every $k \in I_2$, there exist linear polynomials $h_{k,1}, h_{k,2} \in \mathbb{F}[\mathbf{z}]$, such that*

$$\sum_{k \in I_2} (\ell_{k,1,\mathbf{y}} + \ell_{k,1,\mathbf{z}} + \alpha_{k,1})(\ell_{k,2,\mathbf{y}} + \ell_{k,2,\mathbf{z}} + \alpha_{k,2})(A_2\mathbf{x}) = \sum_{k \in I_2} (y_{k,1} + h_{k,1})(y_{k,2} + h_{k,2}).$$

Proof: For every $k \in I_2, j \in [2]$, let $p_{k,j} = \ell_{k,j,\mathbf{z}} + \alpha_{k,j}$. Then,

$$\sum_{k \in I_2} (\ell_{k,1,\mathbf{y}} + p_{k,1})(\ell_{k,2,\mathbf{y}} + p_{k,2})(A_2\mathbf{x}) = \sum_{k \in I_2} (\ell'_{k,1,\mathbf{y}} + p_{k,1})(\ell'_{k,2,\mathbf{y}} + p_{k,2}),$$

where for $k \in I_2, j \in [2], \ell'_{k,j,\mathbf{y}} = \ell_{k,j,\mathbf{y}}(A_2\mathbf{x})$ and as A_2 maps every variable in $\mathbf{z}$ to itself, for $k \in I_2, j \in [2], p_{k,j}(A_2\mathbf{x}) = p_{k,j}$. Since $A_2 \in \mathrm{GL}(n, \mathbb{F})$, Observation 5.2 implies that $\{\ell'_{k,j,\mathbf{y}} : k \in I_2, j \in [2]\}$ is $\mathbb{F}$-linearly independent. On expanding the R.H.S. of the above equation, we get

$$\sum_{k \in I_2} (\ell'_{k,1,\mathbf{y}} + p_{k,1})(\ell'_{k,2,\mathbf{y}} + p_{k,2}) = \sum_{k \in I_2} \ell'_{k,1,\mathbf{y}}\ell'_{k,2,\mathbf{y}} + \sum_{k \in I_2} (\ell'_{k,1,\mathbf{y}}p_{k,2} + \ell'_{k,2,\mathbf{y}}p_{k,1}) + \sum_{k \in I_2} p_{k,1}p_{k,2}. \quad (5.3)$$

For $k \in I_2$, let $h_{k,1}, h_{k,2}$ be the coefficients of $y_{k,2}, y_{k,1}$ in $\sum_{k \in I_2} (\ell'_{k,1,\mathbf{y}}p_{k,2} + \ell'_{k,2,\mathbf{y}}p_{k,1})$ respectively. Then, for every $k \in I_2, h_{k,1}, h_{k,2} \in \mathbb{F}[\mathbf{z}]$ are linear polynomials and

$$\sum_{k \in I_2} (\ell'_{k,1,\mathbf{y}}p_{k,2} + \ell'_{k,2,\mathbf{y}}p_{k,1}) = \sum_{k \in I_2} (y_{k,1}h_{k,2} + y_{k,2}h_{k,1}).$$

Equation (5.2) implies that $\sum_{k \in I_2} \ell'_{k,1,\mathbf{y}}\ell'_{k,2,\mathbf{y}} = \sum_{k \in I_2} y_{k,1} \cdot y_{k,2}$. Using this and on adding and subtracting $\sum_{k \in I_2} h_{k,1}h_{k,2}$ from Equation (5.3), we get

$$\sum_{k \in I_2} (\ell'_{k,1,\mathbf{y}} + p_{k,1})(\ell'_{k,2,\mathbf{y}} + p_{k,2}) = \sum_{k \in I_2} (y_{k,1} + h_{k,1})(y_{k,2} + h_{k,2}) + \sum_{k \in I_2} (p_{k,1}p_{k,2} - h_{k,1}h_{k,2}).$$

Now, we show that $\sum_{k \in I_2} (p_{k,1}p_{k,2} - h_{k,1}h_{k,2}) = 0$. Substitute $y_{k,j} = y_{k,j} - h_{k,j}$ for every $k \in I_2, j \in [2]$ in the above equation. Then, we get

$$\sum_{k \in I_2} (\ell'_{k,1,\mathbf{y}} + p'_{k,1})(\ell'_{k,2,\mathbf{y}} + p'_{k,2}) = \sum_{k \in I_2} y_{k,1}y_{k,2} + \sum_{k \in I_2} (p_{k,1}p_{k,2} - h_{k,1}h_{k,2}),$$

where $p'_{k,j} \in \mathbb{F}[\mathbf{z}]$ is a linear polynomial for every $k \in I_2, j \in [2]$. Note that as the R.H.S. of the above equation does not have a monomial containing variables from both $\mathbf{y}$ and $\mathbf{z}$, we get

$$\sum_{k \in I_2} (\ell'_{k,1,\mathbf{y}} p'_{k,2} + \ell'_{k,2,\mathbf{y}} p'_{k,1}) = 0.$$

Since $\{\ell'_{k,j,\mathbf{y}} : k \in I_2, j \in [2]\}$ is $\mathbb{F}$-linearly independent, it is easy to see that for every $k \in I_2, p'_{k,1} = p'_{k,2} = 0$, which implies $\sum_{k \in I_2} (p_{k,1} p_{k,2} - h_{k,1} h_{k,2}) = 0$. Hence

$$\sum_{k \in I_2} (\ell_{k,1,\mathbf{y}} + \ell_{k,1,\mathbf{z}} + \alpha_{k,1})(\ell_{k,2,\mathbf{y}} + \ell_{k,2,\mathbf{z}} + \alpha_{k,2})(A_2 \cdot \mathbf{x}) = \sum_{k \in I_2} (y_{k,1} + h_{k,1})(y_{k,2} + h_{k,2}).$$

□

This observation implies that f_2 looks as follows.

$$f_2 = \sum_{k \in I_1} \widehat{T}_k(A_1 A_2 \mathbf{x}) + \sum_{k \in I_2} (y_{k,1} + h_{k,1})(y_{k,2} + h_{k,2}) + \gamma.$$

When we take partial derivatives of f_2 with respect to $y_{k,1}, y_{k,2}$ for $k \in I_2$, we get black-box access to $y_{k,2} + h_{k,2}$ and $y_{k,1} + h_{k,1}$ respectively. For $k \in I_2, j \in [2]$ let $h_{k,j} = \widetilde{\ell}_{k,j,\mathbf{z}} + \beta_{k,j}$, where $\widetilde{\ell}_{k,j,\mathbf{z}} \in \mathbb{F}[\mathbf{z}]$ is a linear form and $\beta_{k,j} \in \mathbb{F}$. Now, we compute $A_3 \in \mathrm{GL}(n, \mathbb{F})$, which maps $y_{k,j}$ to $y_{k,j} - \widetilde{\ell}_{k,j,\mathbf{z}}$ for every $k \in I_2, j \in [2]$ and every other variable to itself. Let $f_3 = f_2(A_3 \mathbf{x})$. Let $A = A_1 A_2 A_3$. As noted before, $A_2 A_3$ maps every variable in $\mathbf{z}$ to itself. Thus, for every $k \in I_1, \widehat{T}_k(A\mathbf{x}) = \widehat{T}_k(A_1 \mathbf{x})$ and hence $\widehat{T}_k(A\mathbf{x}) \in \mathbb{F}[\mathbf{z}_k]$. This implies

$$f_3 = f(A\mathbf{x}) = \sum_{k \in I_1} \widehat{T}_k(A\mathbf{x}) + \sum_{k \in I_2} (y_{k,1} + \beta_{k,1})(y_{k,2} + \beta_{k,2}) + \gamma.$$

Since $\widehat{T}_k(A\mathbf{x}) \in \mathbb{F}[\mathbf{z}_k]$ for every $k \in I_1$, the above equation immediately implies that for every $k, k' \in [s]$, $k \neq k', \widehat{T}_k(A\mathbf{x})$ and $\widehat{T}_{k'}(A\mathbf{x})$ are variable disjoint. Further, it follows from Claim 2.2.3 and the proof of Observation 5.1 that for every $k \in [s_1]$, there exists $J_k \subseteq [s']$, such that $\mathrm{var}(\widehat{T}_k(A\mathbf{x})) = \uplus_{i \in J_k} \mathbf{z}_i$.

5.3.3 Computing efficient black-box access to a term

Before describing Procedure 14, we show in the following observation that the for loop in Steps 4 - 6 of Algorithm 12 is correct.

Observation 5.4 (Learning variable sets) *After execution of the for loop in Steps 4 - 6, for every $k \in [s_1]$, $\mathbf{z}_k = \mathrm{var}\left(\widehat{T}_k(A_0 \mathbf{x})\right)$.*

Proof: Recall $\mathscr{C} = \{\mathbf{z}_1, \ldots, \mathbf{z}_{s_1}\}$ is the set of variable sets corresponding to the connected components of the graph $G = (V, E)$, where $V = \{\mathbf{z}_1, \ldots, \mathbf{z}_{s'}\}$. Pick a $k \in [s_1]$ arbitrarily and let $\widehat{T}_k = \widehat{Q}_{k,1} \cdots \widehat{Q}_{k,m_k}$. If there exists an $i \in [s']$ such that for every $l \in [m_k]$, $\mathrm{var}(\widehat{Q}_{k,l}(A_0\mathbf{x})) \subseteq \mathbf{z}_i$ then $\mathrm{var}(\widehat{T}_k(A_0\mathbf{x})) \subseteq \mathbf{z}_i$. Suppose this is not the case and there exist $i_1 \neq i_2 \in [s'], I_1, I_2 \subseteq [m_k]$ such that $I_1 \cap I_2 = \emptyset$, $\mathrm{var}(\prod_{l \in I_1} \widehat{Q}_{k,l}) \subseteq \mathbf{z}_{i_1}$ and $\mathrm{var}(\prod_{l \in I_2} \widehat{Q}_{k,l}) \subseteq \mathbf{z}_{i_2}$. Then, we show in the following two cases that we add an edge between $\mathbf{z}_{i_1}$ and $\mathbf{z}_{i_2}$.

Case 1. Either $m_k \geq 3$ or $\deg(Q_{k,1}) \geq 2$ and $\deg(Q_{k,2}) \geq 2$. In this case, we know from Corollary 5.1 and Observation 2.7 that for every $l \in [m_k], \widehat{Q}_{k,l}$ is an irreducible factor of $\det(H_{\widehat{T}_k})$. It follows from Claim 2.2.3 that there exist $i_1, i_2 \in [s'], i_1 \neq i_2$ and $I_1, I_2 \subseteq [m_k]$ such that $\cup_{l \in I_1} \mathrm{var}(\widehat{Q}_{k,l}) \subseteq \mathbf{z}_{i_1}$ and $\cup_{l \in I_2} \mathrm{var}(\widehat{Q}_{k,l}) \subseteq \mathbf{z}_{i_2}$. Let $z_1 \in \mathbf{z}_{i_1}, z_2 \in \mathbf{z}_{i_2}$ be arbitrary. As $\mathbf{z}_{i_1} \cap \mathbf{z}_{i_2} = \emptyset$,

$$\frac{\partial^2 f(A_0\mathbf{x})}{\partial z_1 \partial z_2} = \frac{\partial^2 \widehat{T}_k(A_0\mathbf{x})}{\partial z_1 \partial z_2} = \prod_{l \in [m_k] \setminus (I_1 \uplus I_2)} \widehat{Q}_{k,l} \left(\frac{\partial \prod_{l_1 \in I_1} \widehat{Q}_{k,l_1}(A_0\mathbf{x})}{\partial z_1} \right) \left(\frac{\partial \prod_{l_2 \in I_2} \widehat{Q}_{k,l_2}(A_0\mathbf{x})}{\partial z_2} \right).$$

Clearly, $\frac{\partial^2 f(A_0\mathbf{x})}{\partial z_1 \partial z_2} \neq 0$ and we add an edge between $\mathbf{z}_{i_1}$ and $\mathbf{z}_{i_2}$.

Case 2. $m_k = 2$ and exactly one of $Q_{k,1}$ and $Q_{k,2}$ is a variable. Without loss of generality, let $Q_{k,1}$ is a variable. It follows from Corollary 5.1 that $\widehat{Q}_{k,1}$ is a factor of $\det(H_{\widehat{T}_k})$. Then, Claim 2.2.3 implies that there exist distinct $i_1, i_2 \in [s']$ such that $\mathrm{var}(\widehat{Q}_{k,1}(A_0\mathbf{x})) \subseteq \mathbf{z}_{i_1}$ and $\mathrm{var}(\widehat{Q}_{k,2}(A_0\mathbf{x})) \cap \mathbf{z}_{i_2} \neq \emptyset$. Pick $z_1 \in \mathbf{z}_{i_1}$ and $z_2 \in \mathbf{z}_{i_2} \cap \mathrm{var}(\widehat{Q}_{k,2}(A_0\mathbf{x}))$ arbitrarily.

$$\frac{\partial^2 f(A_0\mathbf{x})}{\partial z_1 \partial z_2} = \frac{\partial^2 \widehat{T}_k(A_0\mathbf{x})}{\partial z_1 \partial z_2} = \frac{\partial}{\partial z_1} \left(\widehat{Q}_{k,1}(A_0\mathbf{x}) \frac{\partial \widehat{Q}_{k,2}(A_0\mathbf{x})}{\partial z_2} \right).$$

As $z_1 \in \mathrm{var}(\widehat{Q}_{k,1}(A_0\mathbf{x}))$, clearly $\frac{\partial^2 f(A_0\mathbf{x})}{\partial z_1 \partial z_2} \neq 0$ and we add an edge between $\mathbf{z}_{i_1}$ and $\mathbf{z}_{i_2}$.

It follows from these two cases that for every $k \in [s_1], \mathrm{var}(\widehat{T}_k(A_0\mathbf{x}))$ is contained in one connected component of G. Now, suppose $z_1 \in \mathrm{var}(\widehat{T}_{k_1}(A_0\mathbf{x}))$ and $z_2 \in \mathrm{var}(\widehat{T}_{k_2}(A_0\mathbf{x}))$. Then, clearly $\frac{\partial^2 f(A_0\mathbf{x})}{\partial z_1 \partial z_2} = 0$. This implies that we add an edge between $\mathbf{z}_{i_1}$ and $\mathbf{z}_{i_2}$ if and only if $\mathbf{z}_{i_1} \uplus \mathbf{z}_{i_2} \subseteq \mathrm{var}(\widehat{T}_k(A_0\mathbf{x}))$ for some $k \in [s_1]$. Thus, $\mathscr{C} = \{\mathrm{var}(\widehat{T}_1(A_0\mathbf{x})), \ldots, \mathrm{var}(\widehat{T}_{s_1}(A_0\mathbf{x}))\}$. □

Now, we formally describe Procedure 14.

Procedure 14 Compute-Term-Black-Box(g)

Input: Black-box access to $\widehat{T}(A_0\mathbf{x}) + \gamma'$, where $\widehat{T}$ is a term of f and $\gamma' \in \mathbb{F}$.

Output: Black-box access to $\widehat{T}(A_0\mathbf{x})$ using just one black-box query to f.

1. Compute black-box access to $\det(H_g)$ with respect to $\text{var}(g)$ and factorize it using Fact 2.17. Let N be the set of black-boxes of the irreducible factors of $\det(H_g)$.
2. **for** $p \in N$ **do**
3. Let $\mathbf{a}$ be a size-$|\text{var}(g)|$ vector containing random field elements. For a fresh variable t, interpolate $p(\mathbf{a} \cdot t)$ and $g(\mathbf{a} \cdot t)$.
4. Compute $p'(t)$ and $\beta \in \mathbb{F}$ such that $p(\mathbf{a} \cdot t)p'(t) + \beta = g(\mathbf{a} \cdot t)$ by solving a system of linear equations in the coefficients of p' and β.
5. If $g - \beta$ is reducible, then return black-box access to $g - \beta$.
6. **end for**

The correctness of the above procedure is argued in the following claim.

Claim 5.3.1 *Let $k \in [s_1], \mathbf{z}_k = \text{var}(\widehat{T}_k(A_0\mathbf{x}))$. Then, Compute-Term-Black-Box($A_0(\mathbf{z}_k, \mathbf{z} \setminus \mathbf{z}_k = 0)$) returns black-box access to $\widehat{T}_k(A_0\mathbf{x})$ with high probability. Moreover, one query to $\widehat{T}_k(A_0\mathbf{x})$ requires just one black-box query to f.*

Proof: We know that

$$f(A_0\mathbf{x}) = \widehat{T}_1(A_0\mathbf{x}) + \cdots + \widehat{T}_{s_1}(A_0\mathbf{x}) + \widehat{q}(A_0\mathbf{x}) + \gamma.$$

As $\widehat{T}_1(A_0\mathbf{x}), \ldots, \widehat{T}_{s_1}(A_0\mathbf{x}), \widehat{q}(A_0\mathbf{x})$ are pairwise variable disjoint,

$$f(A_0(\mathbf{z}_k, \mathbf{z} \setminus \mathbf{z}_k = 0)) = \widehat{T}_k(A_0\mathbf{x}) + \gamma',$$

where $\gamma' \in \mathbb{F}$. Let $g = f(A_0(\mathbf{z}_k, \mathbf{z} \setminus \mathbf{z}_k = 0))$. Note that black-box access to g can be directly computed from the black-box of f. The objective is to learn γ' and then subtract it from the black-box of g to get black-box access to $\widehat{T}_k(A_0\mathbf{x})$.

The procedure computes N, which is the set of irreducible factors of $\det(H_g) = \det(H_{\widehat{T}_k(A_0\mathbf{x})})$. As $\deg(\widehat{T}_k(A_0\mathbf{x})) \geq 3$, it follows from Corollary 5.1 and Corollary 2.1 that N contains non-zero constant multiples of at least one child of the root of $f(A_0\mathbf{x})$. We call such factors as "good factors" and other factors are called "bad factors". Suppose p is a factor of $\det(H_g)$ picked by the algorithm. Then, p can either be good or bad. We analyse these two cases separately.

Case 1: p is a good factor. In this case, p is also a factor of $\widehat{T}_k(A_0\mathbf{x})$. The procedure computes p' and β, which satisfy

$$p(\mathbf{a} \cdot t)p'(t) + \beta = g(\mathbf{a} \cdot t), \tag{5.4}$$

where $\mathbf{a}$ is a vector of $|\mathbf{z}_k|$ many random field elements and t is a fresh variable. One choice for $p'(t)$ and β are $\frac{\widehat{T}_k(A_0\mathbf{x})}{p}(\mathbf{a}\cdot t)$ and γ' respectively. We can compute $p'(t)$ and β as follows: Interpolate $p(\mathbf{a}\cdot t)$ and $g(\mathbf{a}\cdot t)$ as univariate polynomials in $\mathbb{F}[t]$ and treat the coefficients of $p'(t)$ and β as formal variables. Now, solve the system of linear equations in the coefficients of $p'(t)$ and β by comparing the coefficients of different monomials in t variables in the L.H.S. and the R.H.S. of Equation (5.4). We now argue that we get a unique solution.

Suppose $(p'_1(t), \beta_1)$ and $(p'_2(t), \beta_2)$ satisfy Equation (5.4). Then, we get

$$p(\mathbf{a}\cdot t)(p'_1(t) - p'_2(t)) = \beta_1 - \beta_2.$$

As $\mathbf{a}$ contains random field elements, the Schwartz-Zippel lemma implies that with high probability, $\deg(p(\mathbf{a}\cdot t)) \geq 1$. Thus, $\beta_1 = \beta_2$, which further implies $p'_1(t) = p'_2(t)$. Because of this, $\beta = \gamma'$. Then, $g - \beta$ is reducible and the procedure halts by returning black-box access to $g - \beta$.

Case 2: p is a bad factor. In this case, if $\beta = \gamma'$ then we are done. Let $\beta \neq \gamma'$. Then, note that $g - \beta$ is in the orbit of $T_k + \gamma' - \beta$. It follows from Observation 2.7 that $g - \beta$ is irreducible and the procedure picks another factor from N.

For the 'moreover' part, observe that after learning β, black-box access to $\widehat{T}_k(A_0(\mathbf{z}_k = \mathbf{a}_k, \mathbf{x} \setminus \mathbf{z}_k = 0))$ can be computed for any $\mathbf{a}_k \in \mathbb{F}^{|\mathbf{z}_k|}$ using just one black-box query to f. □

Now, we are ready to prove Lemma 5.2.

Proof of Lemma 5.2

We prove the lemma by induction on the product-depth Δ of $\mathtt{C}$. Recall that we want to show that if A is the matrix computed by Find-Equivalence(f) then $f(A\mathbf{x}) \in$ PS-orb($\mathtt{C}$), where $\mathtt{C}$ is an n-variate regular ROF and f does not have redundant variables. Let $\Delta = 0$. Then, $\mathtt{C} = x_1$ and f is an affine form. As all the variables in f are essential, $f = \alpha x_1 + \beta$ for some $\alpha, \beta \in \mathbb{F}, \alpha \neq 0$. As $n = 1, f(I_{n\times n}\mathbf{x}) \in$ PS-orb($\mathtt{C}$). This proves the base case.

Now, suppose the lemma holds for $\Delta > 0$ and the product-depth of $\mathtt{C}$ is $\Delta + 1$. Recall that $\mathtt{C} = T_1 + \cdots + T_s + \gamma, f = \mathtt{C}(B\mathbf{x} + \mathbf{d})$ for some $B \in \mathrm{GL}(n, \mathbb{F})$ and $\mathbf{d} \in \mathbb{F}^n$. Then, $f = \widehat{T}_1 + \cdots + \widehat{T}_s + \gamma$, where for every $k \in [s], \widehat{T}_k = T_k(B\mathbf{x}+\mathbf{d})$. Also recall that $q = T_{s_1+1} + \cdots + T_s$ is a quadratic form and $\widehat{q} = q(B\mathbf{x} + \mathbf{d})$. It follows from the analysis of Procedure 13 given in Section 5.3.1 that after Step 2 is completed, $\widehat{T}_1(A_0\mathbf{x}), \ldots, \widehat{T}_{s_1}(A_0\mathbf{x}), \widehat{q}(A_0\mathbf{x})$ are variable disjoint polynomials. Further, $\widehat{q}(A_0\mathbf{x}) = (y_1 + \beta_1)(y_2 + \beta_2) + \cdots + (y_{m-1} + \beta_{m-1})(y_m + \beta_m)$, where for every $i \in [m], \beta_i \in \mathbb{F}$. Observation 5.4 implies that after execution of the for loop in Steps 2-6,

for every $k \in [s_1], \widehat{T}_k(A_0\mathbf{x}) = \mathbf{z}_k$.

Let $P_0 \in M_n(\mathbb{F})$ be a permutation matrix such that for every $k \in [s_1], \text{var}(T_k(P_0\mathbf{x})) = \mathbf{z}_k$ and $\text{var}(q(P_0\mathbf{x})) = \mathbf{y}$, where $\mathbf{y} = \{y_1, \ldots, y_m\}$. Then, there exist a $B' \in \text{GL}(n, \mathbb{F})$ and a $\mathbf{d}' \in \mathbb{F}^n$ such that $f(A_0(B'\mathbf{x} + \mathbf{d}')) = \mathtt{C}(P_0\mathbf{x})$. Thus, it suffices to prove that $f(A_0\mathbf{x}) \in \text{PS-orb}(\mathtt{C}(P_0\mathbf{x}))$. The following claim argues the correctness of the for loop in lines 8-18 for some $k \in [s_1]$. We first complete the proof of this lemma assuming the claim below and then prove Claim 5.3.2.

Claim 5.3.2 *Let $k \in [s_1]$. After the execution of the k-th iteration of the for loop in lines 8-18, there exist a permutation matrix $P_k \in M_{|\mathbf{z}_k|}(\mathbb{F})$, an invertible diagonal matrix $S_k \in M_{|\mathbf{z}_k|}(\mathbb{F})$ and a $\mathbf{b}_k \in \mathbb{F}^{|\mathbf{z}_k|}$ such that $\widehat{T}_k(A_0(A_k\mathbf{z}_k, \mathbf{x} \setminus \mathbf{z}_k)) = T_k(P_0(P_kS_k\mathbf{z}_k + \mathbf{b}_k, \mathbf{x} \setminus \mathbf{z}_k))$.*

As $\widehat{q}(A_0\mathbf{x}) = (y_1+\beta_1)(y_2+\beta_2)+\cdots+(y_{m-1}+\beta_{m-1})(y_m+\beta_m)$, clearly there exist a permutation matrix $P_{s_1+1} \in M_{|\mathbf{y}|}(\mathbb{F})$, an invertible diagonal matrix $S_{s_1+1} \in M_{|\mathbf{y}|}(\mathbb{F})$ and a $\mathbf{b}_{s_1+1} \in \mathbb{F}^{|\mathbf{y}|}$ such that $\widehat{q}(A_0\mathbf{x}) = q(P_0(P_{s_1+1}S_{s_1+1}\mathbf{y} + \mathbf{b}_{s_1+1}, \mathbf{x} \setminus \mathbf{y}))$. Let $P \in M_n(\mathbb{F})$ be a permutation matrix such that for every $k \in [s_1], z \in \mathbf{z}_k, P \circ z = P_k \circ z$ and for every $y \in \mathbf{y}, P \circ y = P_{s_1+1} \circ y$. Let $S \in M_n(\mathbb{F})$ be an invertible diagonal matrix such that for every $k \in [s_1], z \in \mathbf{z}_k, S \circ z = S_k \circ z$ and for every $y \in \mathbf{y}, S \circ y = S_{s_1+1} \circ y$. Let $\mathbf{b} \in \mathbb{F}^n$ such that for $k \in [s_1]$, the coordinates of $\mathbf{b}$ labelled by $\mathbf{z}_k$ are $\mathbf{b}_k$ and the coordinates labelled by $\mathbf{y}$ are $\mathbf{b}_{s_1+1}$. Since A'_0 maps every $z \in \mathbf{z}_k$ to $A_k \circ z$ for every $k \in [s_1]$ and maps every $y \in \mathbf{y}$ to itself, we have $f(A_0A'_0\mathbf{x}) = \mathtt{C}(P_0(PS\mathbf{x} + \mathbf{b})) \in \text{PS-orb}(\mathtt{C})$. ■

Proof of Claim 5.3.2

Fix a $k \in [s']$. The correctness of Procedure 14 given in Claim 5.3.1 ensures that after Step 9 is executed, $\widehat{T}$ is the black-box of $\widehat{T}_k(A_0\mathbf{x})$. Suppose that $\widehat{T}_k(A_0\mathbf{x}) = \prod_{l\in[m_k]} \widehat{Q}_{k,l}(A_0\mathbf{x})$, the corresponding term $T_k(P_0\mathbf{x})$ of $\mathtt{C}(P_0\mathbf{x})$ is $T_k(P_0\mathbf{x}) = \prod_{l\in[m_k} Q_{k,l}$, for every $l \in [m_k]$ $Q_{k,l}$ is either a variable or a $+$-rooted sub-ROF of $\mathtt{C}(P_0\mathbf{x})$ and $\widehat{Q}_{k,l}(A_0(B'\mathbf{x} + \mathbf{d}')) = Q_{k,l}$. Then, the factors $\widehat{Q}_1, \ldots, \widehat{Q}_{m_k}$ of $\widehat{T}$ computed in Step 10 are non-zero constant multiples of $\widehat{Q}_{k,l}(A_0\mathbf{x}), l \in [m_k]$, respectively. Since $Q_{k,1}, \ldots, Q_{k,m_k}$ are variable disjoint ROFs,

$$N_{ess}\left(Q_{k,1} \cdots Q_{k,m_k}\right) = N_{ess}\left(Q_{k,1}\right) + \cdots + N_{ess}\left(Q_{m_k}\right).$$

Also, for all $l \in [m_k]$, $N_{ess}\left(\widehat{Q}_l\right) = N_{ess}\left(\widehat{Q}_{k,l}(A_0\mathbf{x})\right) = N_{ess}\left(Q_{k,l}\right)$. Similarly, $N_{ess}\left(\widehat{Q}_1 \cdots \widehat{Q}_{m_k}\right) = N_{ess}\left(Q_{k,1} \cdots Q_{k,m_k}\right)$. Thus,

$$N_{ess}\left(\widehat{Q}_1 \cdots \widehat{Q}_{m_k}\right) = N_{ess}\left(\widehat{Q}_1\right) + \cdots + N_{ess}\left(\widehat{Q}_{m_k}\right).$$

So, from Claim 2.2.2, there exists an $A_{k,0} \in \mathrm{GL}(|\mathbf{z}_k|, \mathbb{F})$ such that $\widehat{Q}_1(A_{k,0}\mathbf{z}_k), \ldots, \widehat{Q}_{m_k}(A_{k,0}\mathbf{z}_k)$ are variable disjoint. Claim 2.2.2 also implies that $\widehat{Q}_1(A_{k,0}\mathbf{z}_k), \ldots, \widehat{Q}_{m_k}(A_{k,0}\mathbf{z}_k)$ do not contain any redundant variable. This means that for all $l \in [m_k]$, $|\mathrm{var}\,(Q_{k,l})\,| = |\mathbf{z}_{k,l}|$, where $\mathbf{z}_{k,l} = \mathrm{var}\left(\widehat{Q}_l(A_{k,0}\mathbf{z}_k)\right)$. So, there exists a permutation matrix $P_{k,0} \in M_{|\mathbf{z}_k|}(\mathbb{F})$ such that for all $l \in [m_k]$, $\mathrm{var}\,(Q_{k,l}(P_{k,0}\mathbf{z}_k)) = \mathbf{z}_{k,l}$.

We now analyse the l-th iteration of the inner loop of lines 12-15 for some $l \in [m_k]$. As $\mathbf{a}$ is chosen randomly, with high probability $\widehat{Q}_l = \widehat{T}\,(A_{k,0}\,(\mathbf{z}_{k,l}, \mathbf{z}_k \setminus \mathbf{z}_{k,l} = \mathbf{a}))$ is $c_l \cdot \widehat{Q}_{k,l}(A_0(A_{k,0}\mathbf{z}_k, \mathbf{x} \setminus \mathbf{z}_k))$ for some $c_l \in \mathbb{F}\setminus\{0\}$. Let Q_l be a regular ROF obtained by multiplying c_l with $Q_{k,l}(P_{k,0}\mathbf{z}_k)$, pushing it down to the leaves, and removing it from any non-constant leaf.

Let $B'' = A'^{-1}_{k,0} B' P'_{k,0}$, where $A'_{k,0} \in \mathrm{GL}(n, \mathbb{F})$ maps every $z \in \mathbf{z}_k$ to $A_{k,0} \circ z$ and every other variable to itself, while $P'_{k,0} \in M_n(\mathbb{F})$ maps every $z \in \mathbf{z}_k$ to $P_{k,0} \circ z$ and every other variable to itself. Also, let $\mathbf{d}'' = A'^{-1}_{k,0}\mathbf{d}'$. It is not difficult to see that $\widehat{Q}_l\,(B''\mathbf{x} + \mathbf{d}'') = Q_l$. Note that the product-depth of Q_l is at most Δ. To recursively perform equivalence test on $\widehat{Q}_l$ we shall show that there exist a $B_l \in \mathrm{GL}(|\mathbf{z}_{k,l}|, \mathbb{F})$ and a $\mathbf{d}_l \in \mathbb{F}^{|\mathbf{z}_{k,l}|}$ such that $\widehat{Q}_l\,(B_l\mathbf{z}_{k,l} + \mathbf{d}_l) = Q_l(\mathbf{z}_{k,l})$.

As $\widehat{Q}_l \in \mathbb{F}[\mathbf{z}_{k,l}]$, every variable in $\mathbf{z}_{k,l}$ is essential for $\widehat{Q}_{k,l}$ and $\widehat{Q}_l(B''\mathbf{x} + \mathbf{d}'') \in \mathbb{F}[\mathbf{z}_{k,l}]$, it follows from Observation 2.4 that B'' maps every $\mathbf{z}_{k,l}$-variable to a linear form in $\mathbf{z}_{k,l}$. Let $[B'']_{\mathbf{z}_{k,l}\times\mathbf{z}_{k,l}}$ and $[B'']_{\mathbf{z}_{k,l}\times\mathbf{x}\setminus\mathbf{z}_{k,l}}$ be obtained by restricting the rows and columns of B'' to $\mathbf{z}_{k,l}, \mathbf{z}_{k,l}$ and $\mathbf{z}_{k,l}, \mathbf{x} \setminus \mathbf{z}_{k,l}$ respectively. Then, $[B'']_{\mathbf{z}_{k,l}\times\mathbf{x}\setminus\mathbf{z}_{k,l}} = 0$ and as $B'' \in \mathrm{GL}(n, \mathbb{F})$, we get that $[B'']_{\mathbf{z}_{k,l}\times\mathbf{z}_{k,l}} \in \mathrm{GL}(|\mathbf{z}_{k,l}|, \mathbb{F})$. Using this and the fact that $\mathrm{var}\left(\widehat{Q}_l\right) = \mathbf{z}_{k,l}$, we get

$$Q_l(\mathbf{z}_{k,l}) = \widehat{Q}_l\left([B'']_{\mathbf{z}_{k,l}\times\mathbf{z}_{k,l}}\,\mathbf{z}_{k,l} + [\mathbf{d}'']_{\mathbf{z}_{k,l}}\right),$$

where $[\mathbf{d}'']_{\mathbf{z}_{k,l}}$ is obtained by restricting $\mathbf{d}''$ to $\mathbf{z}_{k,l}$. As $[B'']_{\mathbf{z}_{k,l}\times\mathbf{z}_{k,l}}$ is invertible, we can set $B_l = [B'']_{\mathbf{z}_{k,l}\times\mathbf{z}_{k,l}}$ and $\mathbf{d}_l = [\mathbf{d}'']_{\mathbf{z}_{k,l}}$.

Thus, by the induction hypothesis, $A_{k,l}$ computed in Step 14, is such that there exist a permutation matrix $P_{k,l} \in M(|\mathbf{z}_{k,l}|, \mathbb{F})$, an invertible scaling matrix $S_{k,l} \in M(|\mathbf{z}_{k,l}|, \mathbb{F})$ and a $\mathbf{b}_{k,l} \in \mathbb{F}^{|\mathbf{z}_{k,l}|}$ satisfying $\widehat{Q}_l(A_{k,l}\mathbf{z}_{k,l}) = Q_l(P_{k,l}S_{k,l}\mathbf{z}_{k,l} + \mathbf{b}_{k,l})$. As $\widehat{Q}_{k,l}(A_0(A_{k,0}\mathbf{z}_k, \mathbf{x} \setminus \mathbf{z}_k)) = c_l^{-1} \cdot \widehat{Q}_l(\mathbf{z}_k)$ and $Q_l(\mathbf{z}_k) = c_l^{-1} \cdot Q_{k,l}(P_{k,0}\mathbf{z}_k)$ we get,

$$\widehat{Q}_{k,l}(A_0(A_{k,0}(A_{k,l}\mathbf{z}_{k,l}, \mathbf{z}_k \setminus \mathbf{z}_{k,l}), \mathbf{x} \setminus \mathbf{z}_k)) = Q_{k,l}(P_{k,0}(P_{k,l}S_{k,l}\mathbf{z}_{k,l} + \mathbf{b}_{k,l}, \mathbf{z}_k \setminus \mathbf{z}_{k,l})).$$

Since this is true for all $l \in [m_k]$, after the execution of the for loop of lines 12-15 and Step 16, for all $l \in [m_k]$,

$$\widehat{Q}_{k,l}(A_0(A_k\mathbf{z}_k, \mathbf{x} \setminus \mathbf{z}_k)) = Q_{k,l}(P_{k,0}(P_kS_k\mathbf{z}_k + \mathbf{b}_k)),$$

where for all $l \in [m_k]$ and $z \in \mathbf{z}_{k,l}$, P_k maps z to $P_{k,l} \circ z$, S_k maps z to $S_{k,l} \circ z$ and the z-th

coordinate of $\mathbf{b}_k$ is the same as that of the z-th coordinate of $\mathbf{b}_{k,l}$. Hence,

$$\widehat{T}_k(A_0(A_k\mathbf{z}_k, \mathbf{x} \setminus \mathbf{z}_k)) = T_k(P_0(P_kS_k\mathbf{z}_k + \mathbf{b}_k, \mathbf{x} \setminus \mathbf{z}_k)).$$

This completes the proof. ■

Running time analysis

We first show that Procedures 13 and 14 run in polynomial time and then analyse Algorithm 12.

Procedure 13. Using Fact 2.14, we can compute $\det(H_f)$ efficiently. Claim 2.2.3 ensures that Make-Factors-Var-Disjoint() also runs in polynomial time. It is easy to see that all the other steps of this procedure also run in polynomial time.

Procedure 14. Here also $\det(H_g)$ can be computed efficiently using Fact 2.14 and then we can factorize $\det(H_g)$ in polynomial time using Fact 2.17. As the number of irreducible factors of $\det(H_g)$ is at most n, univariate polynomials can be interpolated in polynomial time and a system of linear equations can also be solved in polynomial time, we get that this procedure also runs in polynomial time.

Algorithm 12. It follows from the running time analysis of Procedures 13 and 14 that the time spent by Find-Equivalence() outside the recursive calls is $\text{poly}(n)$. First, observe that there are at most n recursive calls made to Find-Equivalence(). This is so because every recursive call is made to a polynomial in the orbit of a distinct $+$-rooted sub-ROF of C. As there are at most n such sub-ROFs, the number of recursive calls made are also at most n. Now, we show that each recursive call also takes polynomial time.

Suppose that during the execution of Algorithm 12, at some recursion depth, recursive call is made to $f_1(\mathbf{x}_1)$, where f_1 is in the orbit of a $+$-rooted sub-ROF C_1 of C. Then, we obtain black-box of f_1 by evaluating black-box of f at some known points from $\mathbb{F}^{|\mathbf{x}_1|}$ and then subtracting a constant from the black-box of f. Hence, one query to f_1 can be computed from one black-box-query to f in $\text{poly}(n)$ time and not in $\text{poly}(|\mathbf{x}_1|)$ time. Now suppose a call to Find-Equivalence($f_2(\mathbf{x}_2)$) is made from Find-Equivalence(f_1). In this case, the time required to prepare a black-box for f_2 is also $\text{poly}(n)$ because here we evaluate f on some known points from $\mathbb{F}^{|\mathbf{x}_2|}$ and then subtract an appropriate constant from black-box of f to obtain black-box of f_2. Thus, the running time to prepare a black-box of f_2 is independent of the recursion depth of the call for f_1. Hence, Algorithm 12 runs in $\text{poly}(n)$ time.

Chapter 6

Hessian determinant of an ROF

In this chapter, we analyse some important properties of the Hessian determinant of a canonical ROF (Definition 2.39) mentioned in Section 5.1 of Chapter 5. This is a joint work with Chandan Saha and Bhargav Thankey. These properties are crucially used in designing an efficient equivalence test for the class of regular ROF (Algorithm 12) given in Chapter 5. The content of this chapter is divided into five sections. In the first section, we give a set of useful notations exclusively for this chapter. The second section is devoted to understanding the structure of the Hessian determinant of a canonical ROF C and the third section contains the *Laplace's expansion* of the Hessian determinant of C. In the fourth section, we give a complete description of the Hessian determinant of a product-depth 2 ROF. The last section is devoted to understanding some important properties of the Hessian determinant of a canonical ROF of arbitrary product-depth.

Let C be a canonical ROF (Definition 2.39) and $\mathrm{var}(\mathsf{C})$ denote the set of variables appearing in C. As C is canonical, it has alternate layers of $+$ gates and $\times$ gates, every gate in C has at least two children, every child of a $\times$ gate computes a non-constant polynomial, and every $+$ gate has at most one variable child. Recall that the product-depth of C, denoted Δ, is equal to the number of $\times$ gates in a longest path from a leaf to the root of C. We identify the polynomial computed by any node v of C with v. In this chapter, we analyse the Hessian determinant (Definition 2.27) of a $\times$-rooted canonical ROF. We first show that it is sufficient to understand the Hessian determinant of a $\times$-rooted canonical ROF. Let

$$\mathsf{C} = T_1 + \cdots + T_s + \gamma,$$

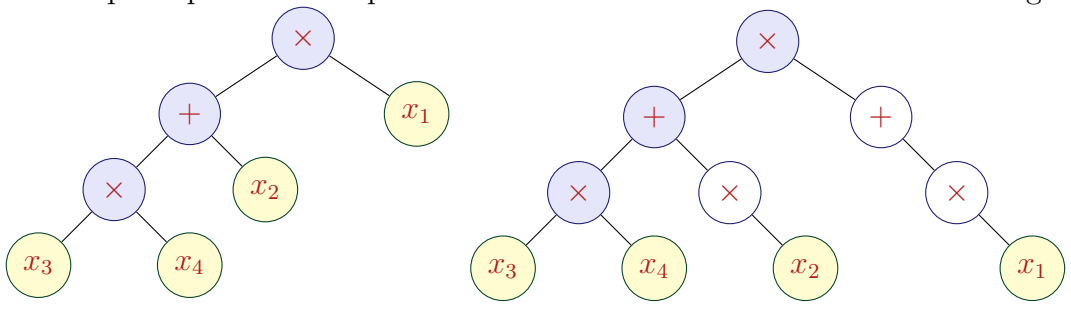
where there exists at most one $l \in [s]$ such that T_l is a variable, for every $k \in [s] \setminus \{l\}, T_k$ is a $\times$-

rooted canonical ROF, and $\gamma \in \mathbb{F}$. Then, the Hessian of C, denoted $H(\mathsf{C})$, [1] is a block-diagonal matrix, with the diagonal blocks being $H_{T_1}, \ldots, H_{T_s}$, where the rows and columns of every H_{T_k} are labelled by $\mathrm{var}(T_k)$. Then, the Hessian determinant of C, denoted $\det(H(\mathsf{C}))$, is as follows:

$$\det(H(\mathsf{C})) = \prod_{l=1}^{s} \det(H(T_l)).$$

Thus, to study $\det(H(\mathsf{C}))$, it is sufficient to focus on $\det(H(T_l))$ for every $l \in [s]$. If for any $l \in [s], |\mathrm{var}(T_l)| = 1$ then $\det(H(\mathsf{C})) = 0$. Thus, we assume that for every $l \in [s], \deg(T_l) \geq 2$. Suppose the product-depth of C is $\Delta + 1$. Henceforth, we focus on an arbitrary term $T \in \{T_1, \ldots, T_s\}$. The following view of T would be helpful in understanding $\det(H(T))$.

The 'extended' version of T. For the sake of analysis of $H(T)$, we transform T as follows: Let p be an arbitrary path in T starting from the topmost $\times$ gate and ending at a non-leaf node w, which is connected to at least one variable. Suppose the length of p, i.e., the number of nodes in p, is ℓ. If $\ell < (2\Delta+1)$ then we disconnect all the variable children from w, add a path of alternate 'dummy' $+$ and $\times$ gates, such that the length of this path is $(2\Delta+1)-\ell$. Further, if w is a $+$ gate then the starting node of this path is a $\times$ gate and vice-versa. Thereafter, we connect all the variable children of w to the bottom-most gate of this path, which is a $\times$ gate. Now, the length of p from root to the last dummy gate is $2\Delta + 1$. Since C is canonical, if w is a $+$ gate then it has at most one variable child. Because of this, it is easy to see that the node w in the original T and the node w in the transformed T compute the same polynomial. We would work with this variant of T in this section and we call it as the *extended canonical form of* T. So, if $T = Q_1 \cdots Q_m$ then for every $u \in [m], Q_u$ is a $+$-rooted sub-ROF of T having product-depth equal to Δ. A pictorial view of the extended canonical form is given below.



[1] In this chapter, we have used $H(\mathsf{C})$ instead of the standard notation H_{C} for denoting the Hessian of C. This is so because we would be using H along with subscripts to denote some 'special' submatrices of the Hessian of C in the later part of this chapter.

One of the most important properties of $\det(H(T))$ needed for the equivalence test given in Chapter 5 is that $\det(H(T)) \neq 0$. We prove that if $\mathbb{F}$ satisfies $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) \geq |\text{var}(T)|$ then $\det(H(T)) \neq 0$ over $\mathbb{F}$. However, it can happen that if $\mathbb{F}$ is a finite field and $char(\mathbb{F}) < |\text{var}(T)|$ then $\det(H(T)) = 0$. For example, suppose $m \in \mathbb{N}$ is such that $m-1$ is a prime number and $T = x_1 \cdots x_m$ then we show in Observation 6.3 that $\det(H(T))$ is divisible by $m-1$, which implies that $\det(H(T)) = 0$ over the fields of characteristic $m-1$.

Now, we briefly talk about how to show that $\det(H(T))$ is non-zero. We first give a set of useful notations in Section 6.1. Then, we analyse the Hessian of T in Section 6.2 and observe some important properties about the structure of $H(T)$. Now, we preprocess $H(T)$ as follows: We first take out all the variables and $+$-rooted sub-ROFs of T common from the rows and columns of $H(T)$. Let the residual matrix be $H'(T)$. We call $\det(H'(T))$ as the 'spurious term' of $\det(H(T))$. $\det(H'(T))$ is a rational function. We study the denominator of $\det(H'(T))$ in Observation 6.1 and denote the numerator of $\det(H'(T))$ as g_T.

Recall that $T = Q_1 \cdots Q_m$, where for every $u \in [m]$, Q_u is a $+$-rooted ROF of product-depth Δ. If each Q_u is a variable (i.e., $\Delta = 0$) then it follows from Claim 6.2.1 and Observation 6.3 that the spurious term of $\det(H(T))$ is equal to $m-1$. But, with an increase in the value of Δ, the spurious term starts becoming more complex. Thus, to understand this, we use the Laplace's expansion of the determinant (Theorem 6.1). In Section 6.3, we study the *Laplace's expansion* of the spurious term of $\det(H(T))$ for a general $\times$-rooted canonical ROF T. After that, we give the complete description of the spurious term of $\det(H(T))$ when $\Delta = 1$ in Section 6.4. In case of $\Delta = 0$, the spurious term of $\det(H(T))$ is an integer whereas for $\Delta = 1$, g_T is a multilinear polynomial. So, it appears to us that giving the complete description of the spurious term of $\det(H(T))$ for higher value of Δ can be quite challenging because of the complex combinatorial structure of g_T. Thus, for $\Delta \geq 2$, we focus on some special monomials of g_T, which we call as the *nice monomials* of g_T.

We show that the coefficients of these nice monomials in g_T are non-zero over the fields of characteristic either 0 or greater than equal to $|\text{var}(T)|$. We are able to do this because the coefficients of these nice monomials are integers and we get neat factorizations of these coefficients, where each factor of the coefficient of any nice monomial is at most $|\text{var}(T)| - 1$. We want to mention here that although there can be $\mathbb{F}$-constants attached to the $+$ gates in T but the coefficients of the nice monomials are independent of these $\mathbb{F}$-constants. A large part of this chapter comprises of Section 6.5, where we understand the structures and coefficients of nice monomials in g_T, where T is a $\times$-rooted canonical ROF having arbitrary product-depth. As mentioned before, this detailed analysis helps us in showing that $\det(H(T))$ is non-zero over fields of characteristic zero and finite fields satisfying $char(\mathbb{F}) \geq |\text{var}(T)|$.

6.1 Notations

Due to the detailed analysis of the Hessian determinant of a $\times$-rooted canonical ROF, we need a set of useful notations. The notations given in this chapter is divided into two parts: We present the first part here and give the second part in Section 6.5.

1. $\mathbf{x}$ denotes the set of variables appearing in T and $|\mathbf{x}| = n$.

2. The top-most $\times$ gate in T is denoted by v_0 and this layer is called as the 0-th layer of T. The fan-in of v_0 is denoted by s_{v_0}, which is equal to m.

3. For $\ell \in [\Delta]$,

   (a) Let Σ_ℓ and $\prod_\ell$ represent the sets of $+$ gates and $\times$ gates in the ℓ-th layer of sum gates and the ℓ-th layer of product gates respectively, starting from the top.[1] The gates in Σ_ℓ and $\prod_\ell$ are denoted by $u_\ell, u'_\ell, \hat{u}_\ell, \ldots$ and $v_\ell, v'_\ell, \hat{v}_\ell, \ldots$ respectively.

   (b) For a fixed $u_\ell \in \Sigma_\ell$, r_{u_ℓ} denotes *the number of non-constant children of* u_ℓ in T and Q_{u_ℓ} represents the sub-ROF of T rooted at the $+$ gate u_ℓ. Similarly, for a fixed $v_\ell \in \prod_\ell$, s_{v_ℓ} and T_{v_ℓ} denote the number of children of v_ℓ[2] and the sub-ROF of T rooted at the $\times$ gate v_ℓ respectively. Further, for $u_\ell \in \Sigma_\ell, v_\ell \in \prod_\ell, n_{u_\ell}$ and n_{v_ℓ} represent the number of variables in Q_{u_ℓ} and T_{v_ℓ} respectively.

   (c) $r(\Sigma_\ell) := \sum_{u_\ell \in \Sigma_\ell} r_{u_\ell}$ and $s(\prod_\ell) := \sum_{v_\ell \in \prod_\ell} s_{v_\ell}$.

   (d) Let $u_\ell \in \Sigma_\ell$. Then, $v \in [r_{u_\ell}]$ means that v is a non-constant child of u_ℓ and $v' \in [r_{u_\ell}] \setminus \{v\}$ means that v' is a non-constant child of u_ℓ other that v.

   (e) Let $v_\ell \in \prod_\ell$. Then, $u \in [s_{v_\ell}]$ means that u is a child of v_ℓ and $u' \in [s_{v_\ell}] \setminus \{u\}$ means that u' is a child of v_ℓ other than u.

4. Let $u \in [m]$ be chosen arbitrarily. Then, $A_u := [m] \setminus \{u\}$.

5. Let $S = \big\{(\mathbf{u}, \mathbf{v}) := (u_\ell, v_\ell)_{\ell \in [\Delta]} : \text{ for every } \ell \in [\Delta], u_\ell \in [s_{v_{\ell-1}}], v_\ell \in [r_{u_\ell}]\big\}$. Notice that S is the set of all paths in T starting from the first layer of sum gates to the last layer of product gates in T. For $u \in [m]$, let $S_u := \{(\mathbf{u}, \mathbf{v}) \in S : u_1 = u\}$.[3]

[1] The layers of gates in T are always labelled from the top-most gate, which is in the 0-th layer.

[2] Recall that in an extended canonical ROF, a multiplication gate does not have a constant child.

[3] One of the reasons to consider the extended canonical form of T is that it gives a uniform description to every path in S. This is helpful in defining the nice monomials.

6. For $(\mathbf{u},\mathbf{v}) \in S$, let $\mathbf{x}_{(\mathbf{u},\mathbf{v})} := \{x_{(\mathbf{u},\mathbf{v},k)} : k \in [n_{(\mathbf{u},\mathbf{v})}]\}$, where $n_{(\mathbf{u},\mathbf{v})} = |\mathbf{x}_{(\mathbf{u},\mathbf{v})}|$. Let $R := \big\{(\mathbf{u},\mathbf{v},k) : (\mathbf{u},\mathbf{v}) \in S, k \in [n_{(\mathbf{u},\mathbf{v})}]\big\}$[1] and for $u \in [m]$, $R_u := \{(\mathbf{u},\mathbf{v},k) \in R : (\mathbf{u},\mathbf{v}) \in S_u\}$.

6.2 The structure of the Hessian of an ROF

This is our first step towards understanding $\det(H(T))$, which is a polynomial in $\mathbb{F}[\mathbf{x}]$. In this section, we first investigate the structure of $H(T)$ and then apply some elementary row and column operations on $H(T)$ to obtain certain factors of $\det(H(T))$. Let $(\mathbf{u},\mathbf{v},k),(\mathbf{u}',\mathbf{v}',k') \in R$. Then, it is easy to observe that the $((\mathbf{u},\mathbf{v},k),(\mathbf{u}',\mathbf{v}',k'))$-th entry of $H(T)$ is one of the following:

1. If $(\mathbf{u},\mathbf{v},k) = (\mathbf{u}',\mathbf{v}',k')$ then the entry is 0.
2. If $(\mathbf{u},\mathbf{v}) = (\mathbf{u}',\mathbf{v}') = (u_\ell,v_\ell)_{\ell\in[\Delta]}$ and $k \neq k'$ then the entry is
$$\frac{\mathbf{x}_{(\mathbf{u},\mathbf{v})}}{x_{(\mathbf{u},\mathbf{v},k)} \cdot x_{(\mathbf{u},\mathbf{v},k')}} \prod_{\ell\in[\Delta]} \left(\prod_{\widehat{u}_\ell \in [s_{v_{\ell-1}}]\setminus\{u_\ell\}} Q_{\widehat{u}_\ell} \right).$$
3. For every $i = 1,\ldots,\Delta$, let $(\mathbf{u},\mathbf{v})_{i-1} = (\mathbf{u}',\mathbf{v}')_{i-1} = (u_\ell,v_\ell)_{\ell\in[i-1]}$.
   (a) If $u_i = u_i'$ and $v_i \neq v_i'$ then the entry is 0.
   (b) If $u_i \neq u_i'$ then it is equal to
$$\frac{\mathbf{x}_{(\mathbf{u},\mathbf{v})} \cdot \mathbf{x}_{(\mathbf{u}',\mathbf{v}')}}{x_{(\mathbf{u},\mathbf{v},k)} \cdot x_{(\mathbf{u}',\mathbf{v}',k')}} \prod_{\substack{\ell\in[i-1],\\ \widehat{u}_\ell\in[s_{v_{\ell-1}}]\setminus\{u_\ell\}}} Q_{\widehat{u}_\ell} \prod_{\hat{u}_i\in[s_{v_{i-1}}]\setminus\{u_i,u_i'\}} Q_{\hat{u}_i} \prod_{\substack{t\in\{i+1,\Delta\},\\ \hat{u}_t\in[s_{v_{t-1}}]\setminus\{u_t\},\\ \hat{u}_t'\in[s_{v'_{t-1}}]\setminus\{u_t'\}}} Q_{\hat{u}_t} Q_{\hat{u}_t'}.$$

Observe that the condition $\widehat{u}_\ell \in [s_{v_{\ell-1}}]\setminus\{u_\ell\}$ implies that $s_{v_{\ell-1}} \geq 2$. Since the original T and its extended canonical form compute the same polynomial, it is sufficient to analyse the Hessian determinant of the extended canonical form. It follows from the structure of $H(T)$ that we can take out the following things common from the rows and the columns of $H(T)$: For every $(\mathbf{u},\mathbf{v},k) \in R$, take out $x_{(\mathbf{u},\mathbf{v},k)}$ from the denominators of each entry of the row and columns indexed by $(\mathbf{u},\mathbf{v},k)$ and $\mathbf{x}_{(\mathbf{u},\mathbf{v})} \cdot \prod_{\ell\in[\Delta]} \left(\prod_{\widehat{u}_\ell \in [s_{v_{\ell-1}}]\setminus\{u_\ell\}} Q_{\widehat{u}_\ell} \right)$ from the numerator of each entry of the $(\mathbf{u},\mathbf{v},k)$-th row of $H(T)$. Note that the polynomials we have taken out common from the rows and the columns of $\det(H(T))$ become factors of $\det(H(T))$. Let $H'(T)$ be the residual matrix, which we call the *residual Hessian* of T.

[1] R is the set of indices of all the variables in $\mathbf{x}$.

Claim 6.2.1 (Factorization of $\det(H(T))$) *Let $\ell \in [\Delta]$ and $u_\ell \in \Sigma_\ell$ be chosen arbitrarily and $\overline{n}_{u_\ell} = \sum_{u'_\ell \in [s_{v_{\ell-1}}] \setminus \{u_\ell\}} n_{u'_\ell}$, which is equal to the number of variables in the siblings of Q_{u_ℓ}. Then,*

$$\det(H(T)) = \left(\prod_{(\mathbf{u},\mathbf{v}) \in S} \mathbf{x}_{(\mathbf{u},\mathbf{v})}^{n_{(\mathbf{u},\mathbf{v})}-2} \right) \prod_{\ell \in [\Delta]} \left(\prod_{\substack{u_\ell \in [s_{v_{\ell-1}}] : s_{v_{\ell-1}} \neq 1, \\ v_\ell \in [r_{u_\ell}]}} Q_{u_\ell}^{\overline{n}_{u_\ell}} \right) \times \det(H'(T)).$$

Remark 6.1 *By the notation $\prod_{\ell \in [\Delta]} \left(\prod_{\substack{u_\ell \in [s_{v_{\ell-1}}] : s_{v_{\ell-1}} \neq 1, \\ v_\ell \in [r_{u_\ell}]}} Q_{u_\ell} \right)$ we mean the product of all the $+$-rooted sub-ROFs of T whose parent product gate has fan-in at least 2.*

Proof: Observe that the multiplicity of $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ is $n_{(\mathbf{u},\mathbf{v})} - 2$ for every $(\mathbf{u},\mathbf{v}) \in S$. This is so because we are taking $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ common from the numerators of $n_{(\mathbf{u},\mathbf{v})}$ many rows and from the denominators of the row and the column labelled by $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$. Fix $\ell \in [\Delta], u_\ell \in [s_{v_{\ell-1}}]$ arbitrarily, such that $s_{v_{\ell-1}} \neq 1$, i.e., the $\times$ gate $v_{\ell-1}$ has at least two children. Then, it is easy to see that the multiplicity of Q_{u_ℓ} is equal to the number of rows of $H(T)$ from which Q_{u_ℓ} was taken out. Let $(\mathbf{u}',\mathbf{v}',k') \in R$ be arbitrary. Note that Q_{u_ℓ} comes out from the numerator of the entries of the $(\mathbf{u}',\mathbf{v}',k')$-th row if and only if $u'_1 = u_1, \ldots, v'_{\ell-1} = v_{\ell-1}$ and $u'_\ell \in [s_{v_{\ell-1}}] \setminus \{u_\ell\}, v'_\ell \in [r_{u'_\ell}], \ldots, u'_\Delta \in [s_{v'_{\Delta-1}}], v'_\Delta \in [r_{u'_\Delta}], k' \in [n_{(\mathbf{u}',\mathbf{v}')}]$. Then, clearly the multiplicity of Q_{u_ℓ} is equal to $\overline{n}_{u_\ell} = \sum_{u'_\ell \in [s_{v_{\ell-1}}] \setminus \{u_\ell\}} n_{u'_\ell}$, where $n_{u'_\ell} = |\mathrm{var}(Q_{u'_\ell})|$. □

The structure of $H'(T)$. Note that for $(\mathbf{u},\mathbf{v},k),(\mathbf{u}',\mathbf{v}',k') \in R$, the $((\mathbf{u},\mathbf{v},k),(\mathbf{u}',\mathbf{v}',k'))$-th entry of $H'(T)$ is one of the following.

1. If $(\mathbf{u},\mathbf{v},k) = (\mathbf{u}',\mathbf{v}',k')$ then the entry is 0.
2. If $(\mathbf{u},\mathbf{v}) = (\mathbf{u}',\mathbf{v}')$ and $k \neq k'$ then the entry is 1.
3. For every $i = 1, \ldots, \Delta$, let $(\mathbf{u},\mathbf{v})_{i-1} = (\mathbf{u}',\mathbf{v}')_{i-1}$.
   (a) If $u_i = u'_i$ and $v_i \neq v'_i$ then the entry is 0.
   (b) If $u_i \neq u'_i$ then it is equal to $\dfrac{\mathbf{x}_{(\mathbf{u}',\mathbf{v}')} \cdot \prod_{t \in [i+1,\Delta]} \prod_{\hat{u}'_t \in [s_{v'_{t-1}}] \setminus \{u'_t\}} Q_{\hat{u}'_t}}{Q_{u'_i}}$.

We call $\det(H'(T))$ as the spurious term of $\det(H(T))$. Note that $\det(H'(T))$ is a rational function in $\mathbf{x}$ over $\mathbb{F}$. In the following observation, we analyse the denominator of $\det(H'(T))$.

Observation 6.1 *The denominator of* $\det(H'(T))$ *is equal to* $d_T := \prod_{\ell\in[\Delta]} \left(\prod_{\substack{u_\ell\in[s_{v_{\ell-1}}]:s_{v_{\ell-1}}\neq 1,\\ v_\ell\in[r_{u_\ell}]}} Q_{u_\ell} \right)$.

Proof: Let $\ell \in [\Delta], u_\ell \in [s_{v_{\ell-1}}]$ be picked arbitrarily, such that $s_{v_{\ell-1}} \neq 1$. Let

$$(\mathbf{u},\mathbf{v})_{\ell-1} := (u_1, v_1, \cdots, u_{\ell-1}, v_{\ell-1}),$$

where for every $i \in [\ell-1], u_i \in [s_{v_i-1}], v_i \in [r_{u_i}]$ (recall that $s_{v_0} = m$). Then, it follows from the structure of $H'(T)$ that for $(\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{k}),(\mathbf{u}',\mathbf{v}',k') \in R$, the denominator of the $((\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{k}),(\mathbf{u}',\mathbf{v}',k'))$-th entry is Q_{u_ℓ} if and only if $(\hat{\mathbf{u}},\hat{\mathbf{v}})_{\ell-1} = (\mathbf{u},\mathbf{v})_{\ell-1} = (\mathbf{u}',\mathbf{v}')_{\ell-1}, \hat{u}_\ell \neq u_\ell$ and $u'_\ell = u_\ell$. Let

$$V_{u_\ell} := \{(\mathbf{u}',\mathbf{v}',k') \in R : (\mathbf{u}',\mathbf{v}')_{\ell-1} = (\mathbf{u},\mathbf{v})_{\ell-1}, u'_\ell = u_\ell\}$$

and

$$W_{u_\ell} := \left\{(\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{k}) \in R : (\hat{\mathbf{u}},\hat{\mathbf{v}})_{\ell-1} = (\mathbf{u},\mathbf{v})_{\ell-1}, \hat{u}_\ell \neq u_\ell\right\}.$$

It follows from the structure of $H'(T)$ that for any $(\mathbf{u}',\mathbf{v}',k') \in V_{u_\ell}$, all the entries of the $(\mathbf{u}',\mathbf{v}',k')$-th column of $H'(T)$ restricted to W_{u_ℓ} are the same. Pick $(\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{k}) \in W_{u_\ell}$ arbitrarily and subtract the $(\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{k})$-th row of $H'(T)$ from the $(\mathbf{u}'',\mathbf{v}'',k'')$-th row of $H'(T)$ for every $(\mathbf{u}'',\mathbf{v}'',k'') \in W_{u_\ell} \setminus \{(\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{k})\}$. After doing this, note that Q_{u_ℓ} appears in the denominator of the non-zero entries of exactly one row in $H'(T)$. Since this is true for every $\ell \in [\Delta], u_\ell \in [s_{v_{\ell-1}}]$, such that $s_{v_{\ell-1}} \neq 1$ and since every such Q_{u_ℓ} is irreducible (Observation 2.7), we get that the multiplicity of Q_{u_ℓ} in the denominator of $\det(H'(T))$ is equal to 1. □

Claim 6.2.1 and Observation 6.1 imply the following.

Claim 6.2.2 *For* $\ell \in [\Delta]$, *let* $\overline{n}_{u_\ell} = \sum_{u'_\ell \in [s_{v_{\ell-1}}]\setminus\{u_\ell\}} n_{u'_\ell}$. *Then,*

$$\det(H(T)) = \left(\prod_{(\mathbf{u},\mathbf{v})\in S} \mathbf{x}_{(\mathbf{u},\mathbf{v})}^{n_{(\mathbf{u},\mathbf{v})}-2} \right) \prod_{\ell\in[\Delta]} \left(\prod_{\substack{u_\ell\in[s_{v_{\ell-1}}]:s_{v_{\ell-1}}\neq 1,\\ v_\ell\in[r_{u_\ell}]}} Q_{u_\ell}^{\overline{n}_{u_\ell}-1} \right) \times g_T, \tag{6.1}$$

where $g_T = d_T \cdot \det(H'(T))$ *and* d_T *is the denominator of* $H'(T)$ *defined in Observation 6.1.*

This implies the following useful result.

Claim 6.2.3 (Factors of $\det(H(T))$) *Let $\mathbb{F}$ be an arbitrary field and $T = Q_1 \cdots Q_m$, where for every $u \in [m], Q_u$ is a $+$-rooted extended canonical ROF having product-depth Δ.*

1. *Let $(\mathbf{u}, \mathbf{v}) \in S$. If either $n_{(\mathbf{u},\mathbf{v})} \geq 3$ or there exists $\ell \in [\Delta]$, such that $s_{v_{\ell-1}} \geq 2$ and Q_{u_ℓ} computes $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$, where $n_{(\mathbf{u},\mathbf{v})} \leq 2$ then every $x \in \mathbf{x}_{(\mathbf{u},\mathbf{v})}$ is a factor of $\det(H(T))$.*

2. *Let $\ell \in [\Delta], s_{v_{\ell-1}} \neq 1$ and $u_\ell \in [s_{v_{\ell-1}}]$ be such that the polynomial computed by Q_{u_ℓ} is not a monomial. Let $\overline{n}_{u_\ell} = \sum_{u'_\ell \in [s_{v_{\ell-1}}] \setminus \{u_\ell\}} n_{u'_\ell}$. Then, the multiplicity of Q_{u_ℓ} in $\det(H(T))$ is at least $\overline{n}_{u_\ell} - 1$.*

Proof: Let $(\mathbf{u}, \mathbf{v}) \in S$ be arbitrary. If $n_{(\mathbf{u},\mathbf{v})} \geq 3$ then it follows from Claim 6.2.2 that every $x \in \mathbf{x}_{(\mathbf{u},\mathbf{v})}$ is a factor of $\det(H(T))$. Suppose there exists $\ell \in [\Delta]$, such that $s_{v_{\ell-1}} \geq 2$ and Q_{u_ℓ} computes $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$, such that $n_{(\mathbf{u},\mathbf{v})} \leq 2$. It is clear from the extended canonical structure of T that there does not exist $u'_\ell \in [s_{v_{\ell-1}}] \setminus \{u_\ell\}$, such that $Q_{u'_\ell}$ computes a monomial. Thus, for every $u'_\ell \in [s_{v_{\ell-1}}] \setminus \{u_\ell\}, \deg(Q_{u'_\ell}) \geq 2$. Let $n_{(\mathbf{u},\mathbf{v})} = 2$ then Claim 6.2.2 implies that every $x \in \mathbf{x}_{(\mathbf{u},\mathbf{v})}$ is a factor of $\det(H(T))$. This is so because the multiplicity of Q_{u_ℓ} in the middle factor of Equation (6.1) is at least 1.

Now, suppose $n_{(\mathbf{u},\mathbf{v})} = 1$. We know Q_{u_ℓ} computes $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$, which is now a variable. Note that if $\overline{n}_{u_\ell} \geq 3$ then Claim 6.2.2 implies that $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ is a factor of $\det(H(T))$, as before. Now suppose $\overline{n}_{u_\ell} = 2$. In this case, note that the degree of $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ in the factors of $\det(H(T))$ other that g_T is equal to zero. However, we show that $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ is a factor of g_T. It is not difficult to see from the structure of $H'(T)$ given above that if Q_{u_ℓ} appears in the denominator of any entry of $H'(T)$ then it also appears in the numerator of the same entry as Q_{u_ℓ} computes a monomial $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$. Since $s_{v_{\ell-1}} \neq 1$, the multiplicity of Q_{u_ℓ} in the denominator d_T of $\det(H'(T))$ computed in Observation 6.1 is equal to 1. Thus, Q_{u_ℓ} should be a factor of g_T, which is the numerator of $\det(H'(T))$. Thus, the variable $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ is a factor of $\det(H(T))$.

The second point of the claim follows immediately from Claim 6.2.2.

□

Corollary 6.1 *Let $T = Q_1 \cdots Q_m$ for some $m \geq 2$, where for every $u \in [m], Q_u$ is a $+$-rooted canonical ROF. If T computes a polynomial of degree at least 3 then there exists $u \in [m]$, such that Q_u is a factor of $\det(H(T))$.*

Proof: Let $m \geq 3$. Then, it follows from Claim 6.2.2 that for every $u \in [m], Q_u$ is a factor of $\det(H(T))$. Now, suppose $m = 2$. As T computes a polynomial of degree at least 3, there exists $u \in [2]$, such that Q_u computes a polynomial of degree at least 2, which implies that $|\text{var}(Q_u)| \geq 2$. Then, $Q_{u'}$ is a factor of $\det(H(T))$, where $u' \in [2] \setminus \{u\}$. □

Note that $g_T \in \mathbb{F}[\mathbf{x}]$. Now, our goal is to analyse g_T. In Section 6.5, we show that if $char(\mathbb{F}) \geq n$ or $char(\mathbb{F}) = 0, g_T$ is not equal to 0, which implies that $\det(H(T)) \neq 0$ over $\mathbb{F}$.

Now, we simplify $H'(T)$ by applying the following elementary row and column operations on it: For every $(\mathbf{u}, \mathbf{v}) \in S$ and for every $k \in [2, n_{(\mathbf{u},\mathbf{v})}]$, subtract the $(\mathbf{u}, \mathbf{v}, 1)$-th row from the $(\mathbf{u}, \mathbf{v}, k)$-th row and the $(\mathbf{u}, \mathbf{v}, 1)$-th column from the $(\mathbf{u}, \mathbf{v}, k)$-th column from $H'(T)$. This simplification would be very helpful in analysing $\det(H'(T))$. Observe that these elementary operations do not change the value of $\det(H'(T))$. Let $u_1, u_1' \in [m], u_1 \neq u_1'$. Note that before applying these elementary operations on $H'(T)$, all the entries in the sub-matrix of $H'(T)$, whose rows are indexed by R_{u_1} and columns by $\{(\mathbf{u}', \mathbf{v}', k'), k' \in [n_{(\mathbf{u}',\mathbf{v}')}]\}$ for some $(\mathbf{u}', \mathbf{v}') \in S_{u_1'}$ are same, i.e., $\left(\frac{\mathbf{x}_{(\mathbf{u}',\mathbf{v}')} \cdot \prod_{t\in[2,\Delta]} \prod_{\hat{u}_t' \in [s_{v_{t-1}'}] \setminus \{u_t'\}} Q_{\hat{u}_t'}}{Q_{u_1'}}\right)$. This implies the following.

Observation 6.2 *Let $u_1, u_1' \in [m], u_1 \neq u_1', (\mathbf{u}', \mathbf{v}', k') \in R_{u_1'}$ (recall $R_{u_1'}$). After applying the above mentioned elementary operations on $H'(T)$, the $(\mathbf{u}', \mathbf{v}', k')$-th column restricted to R_{u_1} is non-zero if and only if $k' = 1$. Further, the $(\mathbf{u}, \mathbf{v}, k)$-th entry in the $(\mathbf{u}', \mathbf{v}', 1)$-th column restricted to R_{u_1} is equal to $\left(\frac{\mathbf{x}_{(\mathbf{u}',\mathbf{v}')} \cdot \prod_{t\in[2,\Delta]} \prod_{\hat{u}_t' \in [s_{v_{t-1}'}] \setminus \{u_t'\}} Q_{\hat{u}_t'}}{Q_{u_1'}}\right)$ if $k = 1$ and 0 otherwise.*

For the sake of reader's convenience, we present here the matrix $H'(T)$, where $T = Q_1 \cdots Q_m$ and every Q_u is a $+$-rooted canonical ROF of product-depth 1. This view would also be helpful in Section 6.4 where we give the complete description of the spurious term of the Hessian determinant of a product-depth 2 canonical ROF. For every $u \in [m]$, let Q_u be given by the following equation

$$Q_u = \mathbf{x}_{u,1} + \cdots + \mathbf{x}_{u,r_u} + \alpha_u, \tag{6.2}$$

where for every $v \in [r_u]$, $\mathbf{x}_{u,v}$ is a monomial in $\mathbf{x}$-variables[1], $|\mathbf{x}_{u,v}| = n_{(u,v)}$, such that for distinct $v, v' \in [r_u], \mathbf{x}_{u,v}$ and $\mathbf{x}_{u,v'}$ are variable disjoint and $\alpha_u \in \mathbb{F}$. For $u \in [m]$, $\mathbf{x}_u := \cup_{v\in[r_u]} \mathbf{x}_{u,v}, |\mathbf{x}_u| := n_u$. It is easy to see that $H'(T)$ looks as follows.

$$H'(T) = \begin{array}{c} \\ \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_m \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_m \end{array} \\ \begin{pmatrix} B_1 & F_{1,2} & \cdots & F_{1,m} \\ F_{2,1} & B_2 & \cdots & F_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ F_{m,1} & F_{m,2} & \cdots & B_m \end{pmatrix} \end{array}, \tag{6.3}$$

[1]We will also treat $\mathbf{x}_{u,v}$ as a set of variable and the usage should be clear from the context.

where for distinct $u_1, u_1' \in [m], F_{u_1,u_1'}$ is an $n_{u_1} \times n_{u_1'}$ size sub-matrix of $H'(T)$, whose rows and columns are labelled by $\mathbf{x}_{u_1}$ and $\mathbf{x}_{u_1'}$ respectively, and for $u \in [m]$, B_u is the $n_u \times n_u$ size sub-matrix of $H'(T)$, whose rows and columns are labelled by $\mathbf{x}_u$. It follows from the structure of $H'(T)$ (after applying the elementary operations) that B_u looks as

$$B_u = \begin{bmatrix} B_{u,1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & B_{u_2} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & B_{u,r_u-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & B_{u,r_u} \end{bmatrix}_{n_u \times n_u} \tag{6.4}$$

, where for every $v \in [r_u], B_{u,v}$ is the following matrix.

$$B_{u,v} = \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & -2 & -1 & \cdots & -1 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & -1 & -1 & \cdots & -2 \end{pmatrix}_{n_{(u,v)} \times n_{(u,v)}}$$

Then, it follows from Observation 6.2 that $F_{u_1,u_1'}$ looks as

$$F_{u_1,u_1'} = \begin{array}{c} \\ x_{u_1,1,1} \\ \vdots \\ x_{u_1,1,n_{(u_1,1)}} \\ x_{u_1,2,1} \\ \vdots \\ x_{u_1,2,n_{(u_1,2)}} \\ \vdots \\ x_{u_1,r_{u_1},1} \\ \vdots \\ x_{u_1,r_{u_1},n_{(u_1,r_{u_1})}} \end{array} \begin{array}{c} \begin{array}{cccccccc} x_{u_1',1,1} & \cdots & x_{u_1',1,n_{(u_1',1)}} & x_{u_1',2,1} & \cdots & x_{u_1',r_{u_1'},1} & \cdots & x_{u_1',r_{u_1'},n_{(u_1',r_{u_1'})}} \end{array} \\ \begin{pmatrix} \frac{\mathbf{x}_{u_1',1}}{Q_{u_1'}} & \cdots & 0 & \frac{\mathbf{x}_{u_1',2}}{Q_{u_1'}} & \cdots & \frac{\mathbf{x}_{u_1',r_{u_1'}}}{Q_{u_1'}} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \frac{\mathbf{x}_{u_1',1}}{Q_{u_1'}} & \cdots & 0 & \frac{\mathbf{x}_{u_1',2}}{Q_{u_1'}} & \cdots & \frac{\mathbf{x}_{u_1',r_{u_1'}}}{Q_{u_1'}} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\mathbf{x}_{u_1',1}}{Q_{u_1'}} & \cdots & 0 & \frac{\mathbf{x}_{u_1',2}}{Q_{u_1'}} & \cdots & \frac{\mathbf{x}_{u_1',r_{u_1'}}}{Q_{u_1'}} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \end{array} \tag{6.5}$$

Recall the objective of the remainder of this chapter is to analyse $\det(H(T))$ and in particular

$\det(H'(T))$. In the following observation, we give the description of $\det(H(T))$ where T is a product of distinct variables. The proof of this immediately follows from the above discussion and it is easy to see that in this case, $\det(H'(T))$ is an integer.

Observation 6.3 (Hessian determinant of a multilinear monomial) *Let $T = x_1 \cdots x_m, m \geq 2$. Then,*

$$\det(H(T)) = (-1)^{(m-1)} \cdot (m-1) \prod_{i \in [m]} x_i^{m-2}.$$

Remark 6.2 (Hessian determinant of a general monomial) *Suppose $T = x_1^{e_1} \cdots x_n^{e_n}$, where for every $i \in [n], e_i \geq 1$. Then, it is also not difficult to show that*

$$\det(H(T)) = (-1)^{n-1} e_1 \cdots e_n \cdot (e_1 + e_2 + \cdots + e_n - 1) \times (x_1^{e_1 \cdot n-2} \cdots x_n^{e_n \cdot n-2}).$$

6.3 The Laplace expansion

We start with the following theorem, which gives a useful description of the determinant of a matrix. We will use this to understand $\det(H(T))$.

Theorem 6.1 (Laplace expansion of the determinant) [Jan08] *Let $\mathbb{F}$ be a field, $n \in \mathbb{N}$, $A \in M(n, \mathbb{F})$, whose rows and columns are indexed by the ordered tuple $(1, \ldots, n)$, $D = \det(A)$ and $1 \leq r \leq n$. Then, for every $1 \leq i_1 < \cdots < i_r \leq n$,*

$$D = \sum_{1 \leq j_1 < \cdots < j_r \leq n} (-1)^{\sum_{\ell \in [r]} (i_\ell + j_\ell)} \cdot D(i_1, i_2, \ldots, i_r \,|\, j_1, j_2, \ldots, j_r) \cdot \overline{D}(i_1, i_2, \ldots, i_r \,|\, j_1, j_2, \ldots, j_r), \tag{6.6}$$

where $D(i_1, i_2, \ldots, i_r \,|\, j_1, j_2, \ldots, j_r)$ is the order-r minor lying in the intersection of the i_1-th, $\ldots, i_r$-th rows and j_1-th, $\ldots, j_r$-th columns of A and $\overline{D}(i_1, i_2, \ldots, i_r \,|\, j_1, j_2, \ldots, j_r)$ is the order-$(n-r)$ minor lying in the intersection of remaining $(n-r)$ rows and $(n-r)$ columns of A.

We call Equation (6.6) as the *Laplace expansion* of D along $\{i_1, \ldots, i_r\}$. We would use this to analyse $\det(H'(T))$. We assume here that the set of children of every gate in T is ordered. Then, for $\ell \in [\Delta], u_\ell \in \Sigma_\ell$ and $v_\ell \in \prod_\ell$, the sets $[u_\ell - 1]$ and $[v_\ell - 1]$ consist of all the indexes of all the $+$-rooted and $\times$-rooted siblings of u_ℓ and v_ℓ, whose orders are less than u_ℓ and v_ℓ respectively. Recall that $R = \{(\mathbf{u}, \mathbf{v}, k) : (\mathbf{u}, \mathbf{v}) \in S, k \in [n_{(\mathbf{u},\mathbf{v})}]\}$ is the set of the indices of variables in C. We first label the rows and columns of $H'(T)$ with increasing natural numbers

using the map $\mu : R \to \mathbb{N}$, where for every $(\mathbf{u}, \mathbf{v}, k) \in R$,

$$\mu((\mathbf{u}, \mathbf{v}, k)) := \sum_{\ell \in [\Delta]} \left(\sum_{\hat{u}_\ell \in [u_\ell - 1]} n_{\hat{u}_\ell} + \sum_{\hat{v}_\ell \in [v_\ell - 1]} n_{\hat{v}_\ell} \right) + k,$$

where for every $\ell \in [\Delta]$, u_ℓ, v_ℓ are the coordinates of $(\mathbf{u}, \mathbf{v})$, $n_{\hat{u}_\ell}$ and $n_{\hat{v}_\ell}$ are the number of variables in the $+$-rooted ROF $Q_{\hat{u}_\ell}$ and $\times$-rooted ROF $T_{\hat{v}_\ell}$ respectively. The map μ imposes the order $\prec$ on R in the following way: let $(\mathbf{u}, \mathbf{v}, k), (\mathbf{u}', \mathbf{v}', k') \in R$. Then, $(\mathbf{u}', \mathbf{v}', k') \prec (\mathbf{u}, \mathbf{v}, k)$ if and only if $\mu((\mathbf{u}', \mathbf{v}', k')) < \mu\Big((\mathbf{u}, \mathbf{v}, k)\Big)$. Observe that for every $\ell \in [\Delta]$, $n_{u_\ell} = \sum_{v_\ell \in [r_{u_\ell}]} n_{v_\ell}$ and $n_{v_\ell} = \sum_{u_{\ell+1} \in [s_{v_\ell}]} n_{u_{\ell+1}}$. It is easy to prove the following.

Observation 6.4 *$\prec$ imposes the lexicographic ordering on R.*

We will see later in this section how the map μ becomes instrumental in the simplification of the Laplace expansion of $\det(H'(T))$.

From now onwards, we would always treat every subset of R as an ordered set with respect to $\prec$. Let $D = \det(H'(T))$. Fix $u_1 \in [m]$ arbitrarily. Recall the set R_{u_1} from Section 6.1. Then, $|R_{u_1}| = n_{u_1}$. Theorem 6.1 implies that the Laplace expansion of D along R_{u_1} is given as follows.

$$D = \sum_{C \subseteq R, |C| = n_{u_1}} sgn(R_{u_1}) \cdot sgn(C) \cdot D(R_{u_1}|C) \cdot D(\overline{R}_{u_1}|\overline{C}), \tag{6.7}$$

where $sgn(R_{u_1}) = (-1)^{\sum_{(\mathbf{u},\mathbf{v},k) \in R_{u_1}} \mu((\mathbf{u},\mathbf{v},k))}$, for any $C \subseteq R, |C| = n_{u_1}$, $sgn(C) = (-1)^{\sum_{(\mathbf{u},\mathbf{v},k) \in C} \mu((\mathbf{u},\mathbf{v},k))}$, $D(R_{u_1}|C)$ is the order-n_{u_1} minor lying in the intersection of rows and columns of $H'(T)$ labelled by R_{u_1} and C respectively and $D(\overline{R}_{u_1}|\overline{C})$ is the order-$(n - n_{u_1})$ minor lying in the intersection of the rows and columns of $H'(T)$ labelled by $\overline{R}_{u_1} := R \setminus R_{u_1}$ and $\overline{C} := R \setminus C$ respectively. The structure of $H'(T)$ and Observation 6.2 imply the following.

Observation 6.5 *Let $C \subseteq R$ be an arbitrary set, such that $|C| = n_{u_1}$. If C satisfies one of the following conditions then $D(R_{u_1}|C) \cdot D(\overline{R}_{u_1}|\overline{C}) = 0$.*

1. $|C \cap R_{u_1}| \leq n_{u_1} - 2$.
2. $|C \cap R_{u_1}| = n_{u_1} - 1$ *and* $(\mathbf{u}', \mathbf{v}', k') \in C \setminus R_{u_1}$, *such that* $k' \neq 1$.
3. $|C \cap R_{u_1}| = n_{u_1} - 1$ *and* $(\mathbf{u}, \mathbf{v}, k) \in R_{u_1} \setminus C$, *such that* $k \neq 1$.

Proof:

1. Suppose $|C \cap R_{u_1}| \leq n_{u_1} - 2$. Let $H(R_{u_1}|C)$ be the sub-matrix of $H'(T)$, whose rows and columns are labelled by R_{u_1} and C respectively. Then, $\det(H(R_{u_1}|C)) = D(R_{u_1}|C)$ and at least two columns in $H(R_{u_1}|C)$ are the columns of $H'(T)$ labelled by tuples in $\overline{R}_{u_1}$ and restricted to R_{u_1}. It follows from Observation 6.2 that such columns are $\mathbb{F}(\mathbf{x})$-linearly dependent. This implies $D(R_{u_1}|C) = 0$.

2. Suppose $(\mathbf{u}', \mathbf{v}', k') \in C \setminus R_{u_1}$, such that $k' \neq 1$. Observation 6.2 implies $D(R_{u_1}|C) = 0$.

3. Suppose $(\mathbf{u}, \mathbf{v}, k) \in R_{u_1} \setminus C$, such that $k \neq 1$. Then, it is easy to see from Observation 6.2 that the column $(\mathbf{u}, \mathbf{v}, k)$ restricted to the rows labelled by $\overline{R}_{u_1}$ is zero. This implies that $D(\overline{R}_{u_1}|\overline{C}) = 0$.

□

For $u_1 \in [m]$, recall that $A_{u_1} = [m] \setminus \{u_1\}$. Let $u_1' \in A_{u_1}$ and

$$\mathscr{C}_{u_1,u_1'} = \left\{ C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} := (R_{u_1} \setminus \{(\mathbf{u},\mathbf{v},1)\}) \cup \{(\mathbf{u}',\mathbf{v}',1)\} \ : \ (\mathbf{u},\mathbf{v}) \in S_{u_1}, (\mathbf{u}',\mathbf{v}') \in S_{u_1'} \right\}.$$

Note that for every $C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} \in \mathscr{C}_{u_1,u_1'}$, $|C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}| = n_{u_1}$. Every $C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ is ordered by $\prec$. Then, Observation 6.5 implies that Equation (6.7) can be written as follows

$$\begin{aligned} D = D(R_{u_1}|R_{u_1}) \cdot D(\overline{R}_{u_1}|\overline{R}_{u_1}) + \sum_{u_1' \in A_{u_1}} \Bigg(\sum_{C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} \in \mathscr{C}_{u_1,u_1'}} sgn(R_{u_1}) \cdot \\ sgn(C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) \cdot D(R_{u_1}|C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) \cdot D(\overline{R}_{u_1}|\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) \Bigg), \end{aligned} \tag{6.8}$$

where $\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} := R \setminus C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ and is ordered by $\prec$. Let $(\mathbf{u},\mathbf{v}) \in S_{u_1}, (\mathbf{u}',\mathbf{v}') \in S_{u_1'}$,

$$b_{(\mathbf{u},\mathbf{v},1)} := \sum_{\hat{v}_1 \in [v_1 - 1]} n_{\hat{v}_1} + \sum_{\ell \in [2,\Delta]} \left(\sum_{\hat{u}_\ell \in [u_\ell - 1]} n_{\hat{u}_\ell} + \sum_{\hat{v}_\ell \in [v_\ell - 1]} n_{\hat{v}_\ell} \right) + 1$$

and

$$b_{(\mathbf{u}',\mathbf{v}',1)} := \sum_{\hat{v}_1' \in [v_1' - 1]} n_{\hat{v}_1'} + \sum_{\ell \in [2,\Delta]} \left(\sum_{\hat{u}_\ell' \in [u_\ell' - 1]} n_{\hat{u}_\ell'} + \sum_{\hat{v}_\ell' \in [v_\ell' - 1]} n_{\hat{v}_\ell'} \right) + 1.$$

Note that $b_{(\mathbf{u},\mathbf{v},1)} = \mu((\mathbf{u},\mathbf{v},1)) - \sum_{\hat{u}_1 \in [u_1-1]} n_{\hat{u}_1}$ and $b_{(\mathbf{u}',\mathbf{v}',1)} = \mu((\mathbf{u}',\mathbf{v}',1)) - \sum_{\hat{u}'_1 \in [u'_1-1]} n_{\hat{u}'_1}$. Also, notice that $b_{(\mathbf{u},\mathbf{v},1)}$ and $b_{(\mathbf{u}',\mathbf{v}',1)}$ are the positions of $(\mathbf{u},\mathbf{v},1)$ and $(\mathbf{u}',\mathbf{v}',1)$ in the ordered sets R_{u_1} and $R_{u'_1}$. Now, we swap a few tuples in each of $C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ and $\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ to bring $(\mathbf{u}',\mathbf{v}',1)$ and $(\mathbf{u},\mathbf{v},1)$ to $b_{(\mathbf{u},\mathbf{v},1)}$-th and $b_{(\mathbf{u}',\mathbf{v}',1)}$-th positions in $C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ and $\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ respectively. This would be helpful in the simplifying Equation (6.8) and is shown in the following two cases.

1. $u_1 < u'_1$: Observe that in this case, the rightmost tuple in the ordered set $C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ is $(\mathbf{u}',\mathbf{v}',1)$. It is easy to verify that the number of right to left swaps required to bring $(\mathbf{u}',\mathbf{v}',1)$ to $b_{(\mathbf{u},\mathbf{v},1)}$-th position in $C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ is equal to $b_1 := n_{u_1} - b_{(\mathbf{u},\mathbf{v},1)}$. Let $C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ be the resulting set. Similarly, in $\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$, the leftmost tuple is $(\mathbf{u},\mathbf{v},1)$. Here, the number of left to right swaps required to bring $(\mathbf{u},\mathbf{v},1)$ at the $b_{(\mathbf{u}',\mathbf{v}',1)}$-th position in $\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ is equal to $b_2 := n_{u_1+1} + \cdots + n_{u'_1-1} + b_{(\mathbf{u}',\mathbf{v}',1)} - 1$. Let $\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ be the resulting set. Note that $C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ and $\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ are no longer ordered by $\prec$. Then, it is easy to show that:

$$\begin{aligned} D(R_{u_1}|C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) &= (-1)^{b_1} D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}), \\ D(\overline{R}_{u_1}|\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) &= (-1)^{b_2} D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}). \end{aligned} \tag{6.9}$$

2. $u_1 > u'_1$: In this case the leftmost tuple in the ordered set $C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ is $(\mathbf{u}',\mathbf{v}',1)$. Similarly, the rightmost tuple in $\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ is $(\mathbf{u},\mathbf{v},1)$. It is easy to verify that the number of left to right swaps required to bring $(\mathbf{u}',\mathbf{v}',1)$ at the $b_{(\mathbf{u},\mathbf{v},1)}$-th position in $C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ is equal to $b_3 := b_{(\mathbf{u},\mathbf{v},1)} - 1$. Similarly, observe that the number of right to left swaps performed on $\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ to bring $(\mathbf{u},\mathbf{v},1)$ at the $b_{(\mathbf{u}',\mathbf{v}',1)}$-th position in $\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ is equal to $b_4 = n_{u_1-1} + \cdots + n_{u'_1} - b_{(\mathbf{u}',\mathbf{v}',1)}$. In this case, we get

$$\begin{aligned} D(R_{u_1}|C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) &= (-1)^{b_3} D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}), \\ D(\overline{R}_{u_1}|\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) &= (-1)^{b_4} D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}). \end{aligned} \tag{6.10}$$

The values of $sgn(R_{u_1})$, $sgn(C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}), b_1, b_2, b_3$ and b_4 imply the following.

Observation 6.6 *Let $u_1, u'_1 \in [m]$ be such that $u_1 \neq u'_1$ and $C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} \in \mathscr{C}_{u_1,u'_1}$. Then,*

$$sgn(R_{u_1}) \cdot sgn(C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) \cdot (-1)^{b_1} \cdot (-1)^{b_2} = sgn(R_{u_1}) \cdot sgn(C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) \cdot (-1)^{b_3} \cdot (-1)^{b_4} = -1.$$

Then, Equations (6.8), (6.9), (6.10) and the above observation imply the following.

$$D = D(R_{u_1}|R_{u_1})D(\overline{R}_{u_1}|\overline{R}_{u_1}) - \sum_{\substack{u'_1 \in A_{u_1} \\ C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} \in \mathscr{C}_{u_1,u'_1}}} D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}). \quad (6.11)$$

Remark 6.3 *1. Let $u_1 \in [m]$.*

(a) $D(R_{u_1}|R_{u_1})$ is the determinant of $H'(Q_{u_1})$, which is the residual Hessian of Q_{u_1}. Since $H'(Q_{u_1})$ is a block diagonal matrix where the diagonal blocks are $H'(T_{v_1}), v_1 \in [r_{u_1}]$, $D(R_{u_1}|R_{u_1}) = \prod_{v_1 \in [r_{u_1}]} \det(H'(T_{v_1}))$. As for every $v_1 \in [r_{u_1}]$, the product-depth of T_{v_1} is one less that the product-depth of T, we say that $\det(H'(T_{v_1}))$ is a 'product-depth $(\Delta - 1)$' instance of D.

(b) Let T_1 be obtained from T by removing the sub-ROF rooted at u_1. Then, $D(\overline{R}_{u_1}|\overline{R}_{u_1}) = \det(H'(T_1))$. Thus, we say that $D(\overline{R}_{u_1}|\overline{R}_{u_1})$ is a 'product-depth Δ and top fan-in $(m-1)$' instance of D.

2. Let $u_1, u'_1 \in [m], u_1 \neq u'_1, (\mathbf{u},\mathbf{v}) \in S_{u_1}, (\mathbf{u}',\mathbf{v}') \in S_{u'_1}$. Let $H_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'),\overline{R}_{u_1}}$ be obtained from the sub-matrix of $H'(T)$, whose rows and columns are labelled by $\overline{R}_{u_1}$ by replacing the $(\mathbf{u}',\mathbf{v}',1)$-th column with the $(\mathbf{u},\mathbf{v},1)$-th column of $H'(T)$ confined to $\overline{R}_{u_1}$. Then, $D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \det(H_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'),\overline{R}_{u_1}})$. Similarly, let $H_{(\mathbf{u}',\mathbf{v}'),(\mathbf{u},\mathbf{v}),R_{u_1}}$ be obtained from the sub-matrix of $H'(T)$, whose rows and columns are labelled by R_{u_1} by replacing the $(\mathbf{u},\mathbf{v},1)$-th column with the $(\mathbf{u}',\mathbf{v}',1)$-th column of $H'(T)$ confined to R_{u_1}. Then, $D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \det(H_{(\mathbf{u}',\mathbf{v}'),(\mathbf{u},\mathbf{v}),R_{u_1}})$. Observe that $H_{(\mathbf{u}',\mathbf{v}'),(\mathbf{u},\mathbf{v}),R_{u_1}}$ is a smaller instance of $H_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'),\overline{R}_{u_1}}$. Due to this, we say that $D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ is a 'smaller instance' of $D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$.

6.4 The Hessian determinant of a product-depth 2 ROF

In the following claim, we give the complete description of g_T, where $T = Q_1 \cdots Q_m$ and for every $u \in [m], Q_u$ is a $+$-rooted extended canonical ROF of product-depth 1.

Claim 6.4.1 (The spurious term of $\det(H(T))$) *Let $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}$ and $\mathbb{F}$ be a field. Let $T = Q_1 \cdots Q_m, m \geq 2$, where for every $u \in [m], Q_u$ is a $+$-rooted extended canonical ROF of product-depth 1 as given in Equation (6.2) and there exists a $u \in [m]$, such that Q_u*

computes a polynomial of degree at least 2. Let $g_T := \prod_{u\in[m]} Q_u \cdot \det(H'(T))$. *Then,*

$$g_T = \prod_{u\in[m]} (-1)^{n_u - r_u} \left(\sum_{M\in\mathscr{P}([m])} \sum_{u\in M, v_u\in[r_u]} \beta_{M,\mathbf{v}_M} \cdot \prod_{u\in M} \mathbf{x}_{u,v_u} \right),$$

where $n_u = |\mathrm{var}(Q_u)|, r_u$ *is the number of non-constant children of the topmost* $+$*-gate of* Q_u, $\mathscr{P}([m])$ *is the power set of* $[m], \mathbf{v}_M = (v_u)_{u\in M}$, *and for* $M = \emptyset$ *or* $|M| = 1$,

$$\beta_{M,\mathbf{v}_M} = \prod_{u\in[m]} \prod_{v\in[r_u]} (n_{(u,v)} - 1) \prod_{\hat{u}\in[m]\setminus M} \alpha_{\hat{u}} \tag{6.12}$$

and for M *satisfying* $|M| \geq 2$, $\beta_{M,\mathbf{v}_M} = (-1)^{|M|-1} \cdot \beta'_{M,\mathbf{v}_M}$, *where*

$$\beta'_{M,\mathbf{v}_M} = \prod_{\hat{u}\in[m]\setminus M} \left(\alpha_{\hat{u}} \prod_{\hat{v}\in[r_{\hat{u}}]} (n_{(\hat{u},\hat{v})} - 1) \right) \prod_{\substack{u\in M \\ \hat{v}\in[r_u]\setminus\{v_u\}}} (n_{(u,\hat{v})} - 1) \left(\sum_{u\in M} n_{(u,v_u)} - 1 \right). \tag{6.13}$$

In particular, if $char(\mathbb{F}) = 0$ *or* $char(\mathbb{F}) \geq n$ *then* $g_T \neq 0$.

Consider the following claim, which is helpful in understanding the minor $D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_1),(u'_1,v'_1)})$ used in Equation (6.11). This claim will be used in the proof of Claim 6.4.1. Recall that for $u_1 \in [m], A_{u_1} = [m] \setminus \{u_1\}$.

Claim 6.4.2 (Coefficients of useful monomials) *Let* $n \in \mathbb{N}, \mathbf{x} = \{x_1, \ldots, x_n\}$ *and* $\mathbb{F}$ *be a field. Let* $m \geq 2, u_1 \in [m], u'_1 \in A_{u_1}$, $v_1 \in [r_{u_1}], v'_1 \in [r_{u'_1}]$ *and* $A_{u_1,u'_1} = [m] \setminus \{u_1, u'_1\}$. *Let* $M' \in \mathscr{P}(A_{u_1,u'_1}), \hat{v}''_1 \in [r_{\hat{u}''_1}]$ *be fixed arbitrarily for every* $\hat{u}''_1 \in M'$. *Then, the coefficient of the monomial* $\mathbf{x}_{u_1,v_1} \prod_{\hat{u}''_1\in M'} \mathbf{x}_{\hat{u}''_1,\hat{v}''_1}$ *in* $Q_{u_1} \prod_{u''_1\in A_{u_1,u'_1}} Q_{u''_1} \cdot D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_1),(u'_1,v'_1)})$ *is equal to*

$$\prod_{u\in A_{u_1}} (-1)^{(n_u - r_u) + |M'|} \prod_{\substack{u\in A_{u_1,u'_1}\setminus M', \\ v\in[r_u]}} (n_{(u,v)} - 1)\alpha_u \prod_{\substack{\hat{u}''_1\in M', \\ \hat{v}_1\in[r_{\hat{u}''_1}]\setminus\{\hat{v}''_1\}}} (n_{\hat{u}''_1,\hat{v}_1} - 1) \prod_{\hat{v}'_1\in[r_{u'_1}]\setminus\{v'_1\}} (n_{(u'_1,\hat{v}'_1)} - 1) n_{(u'_1,v'_1)}.$$

Let q *be an arbitrary monomial of* $Q_{u_1} \prod_{u''_1\in A_{u_1,u'_1}} Q_{u''_1} \cdot D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_1),(u'_1,v'_1)}), u''_1 \in A_{u_1,u'_1}$ *and* $(\mathbf{u}'', \mathbf{v}'') \in S_{u''_1}$. *Then,* q *contains* $\mathbf{x}_{u_1,v_1}$ *and the degree of* $\mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}$ *in* q *is at most 1.*

These two claims complement each other in the following way: To prove Claim 6.4.1, we need a 'top fan-in $m-1$ instance' of this claim as well as Claim 6.4.2. Further, to prove

Claim 6.4.2, we need a 'top fan-in $m-2$ instance' of Claim 6.4.1. Thus, their proofs are by mutual induction on each other. Consider the following useful observation, which immediately follows from Observation 6.2. Recall from notations that S is the set of all the paths in T, starting from the top-most layer of $+$ gates. When the product depth of T is 2, note that every element of S looks like (u,v), where $u \in [m], v \in [r_u]$. Further, recall that in this case, $R = \{(u,v,k) : (u,v) \in S, k \in [n_{(u,v)}]\}$.

Observation 6.7 *Let $u_1 \in [m]$ be arbitrary, and $\mathbf{j}$ be a 0-1 vector, whose entries are labelled by R, such that if $k = 1$ then the $(u,v,1)$-th entry is 1 otherwise it is 0. Let $\overline{R}_{u_1} = R \setminus R_{u_1}$ and $H_{(u',v'),\mathbf{j},\overline{R}_{u_1}}$ be obtained from the sub-matrix of $H'(T)$, whose rows and columns are indexed by $\overline{R}_{u_1}$ by replacing the $(u,v,1)$-th column with $\mathbf{j}$ restricted to $\overline{R}_{u_1}$. Then,*

$$D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_1),(u'_1,v'_1)}) = \frac{\mathbf{x}_{u_1,v_1}}{Q_{u_1}} \det(H_{(u',v'),\mathbf{j},\overline{R}_{u_1}}).$$

Further, every monomial in $Q_{u_1} \prod_{u''_1 \in A_{u_1,u'_1}} Q_{u''_1} \cdot D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_1),(u'_1,v'_1)})$ contains $\mathbf{x}_{u_1,v_1}$.

Remark. We would see a generalization of the above observation in Section 6.5.

Before going to the proofs of these claims, we mention the following observations, which follow immediately from the matrices given in Equation (6.4) and Equation (6.5). These would be used in the proofs of Claim 6.4.1 and Claim 6.4.2.

Observation 6.8 *Let $u \in [m]$. Then, $\det(B_u) = (-1)^{n_u - r_u} \prod_{v \in [r_u]} (n_{(u,v)} - 1)$, where B_u is given in Equation (6.4).*

Observation 6.9 *Let $u_1, u'_1 \in [m]$ be distinct and $v_1 \in [r_{u_1}], v'_1 \in [r_{u'_1}]$. Let $\widehat{B}_{u_1}$ be obtained by replacing the $(u_1,v_1,1)$-th column of B_{u_1} with the $(u'_1,v'_1,1)$-th column of F_{u_1,u'_1}, which is given in Equation (6.5). Then,*

$$\det(\widehat{B}_{u_1}) = (-1)^{n_{u_1} - r_{u_1}} \prod_{\hat{v}_1 \in [r_{u_1}] \setminus \{v_1\}} (n_{(u_1,\hat{v}_1)} - 1) \cdot n_{(u_1,v_1)} \cdot \frac{\mathbf{x}_{u'_1,v'_1}}{Q_{u'_1}}.$$

6.4.0.1 Proof of Claim 6.4.1

We begin by recalling the Laplace expansion of $D = \det(H'(T))$ given by Equation (6.11). As noted before, every element in S looks like (u_1,v_1), where $u_1 \in [m], v_1 \in [r_{u_1}]$. We prove this

claim by induction on m. Suppose $m = 2$. In this case, we set $u_1 = 1, u_1' = 2$ and thus Equation (6.11) becomes

$$D = D(R_1 \,|\, R_1) \cdot D(\overline{R}_1 \,|\, \overline{R}_1) - \left(\sum_{v_1 \in [r_1], v_2 \in [r_2]} D(R_1 \,|\, C'_{(1,v_1),(2,v_2)}) \cdot D(\overline{R}_1 \,|\, \overline{C}'_{(1,v_1),(2,v_2)}) \right). \quad (6.14)$$

Note that $D(R_1 \,|\, R_1) = \det(B_1)$ and $D(\overline{R}_1 \,|\, \overline{R}_1) = \det(B_2)$, ($B_u$ is defined in Equation (6.4) for $u \in [2]$). It follows from Observation 6.8 that

$$D(R_1 \,|\, R_1) = (-1)^{n_1 - r_1} \prod_{v_1 \in [r_1]} (n_{(1,v_1)} - 1) \quad \text{and} \quad D(\overline{R}_1 \,|\, \overline{R}_1) = (-1)^{n_2 - r_2} \prod_{v_2 \in [r_2]} (n_{(2,v_2)} - 1).$$

Further, it follows from Observation 6.9 that

$$D(R_1 \,|\, C'_{(1,v_1),(2,v_2)}) = (-1)^{n_1 - r_1} \prod_{\hat{v}_1 \in [r_1] \setminus \{v_1\}} (n_{(1,\hat{v}_1)} - 1) \cdot n_{(1,v_1)} \cdot \frac{\mathbf{x}_{2,v_2}}{Q_2} \quad \text{and}$$

$$D(\overline{R}_1 \,|\, \overline{C}'_{(1,v_1),(2,v_2)}) = (-1)^{n_2 - r_2} \prod_{\hat{v}_2 \in [r_2] \setminus \{v_2\}} (n_{(2,\hat{v}_2)} - 1) \cdot n_{(2,v_2)} \cdot \frac{\mathbf{x}_{1,v_1}}{Q_1}.$$

Let $g_T = Q_1 \cdot Q_2 \cdot D$. Then, on putting all these equations together, we get

$$g_T = \prod_{u \in [2]} (-1)^{n_u - r_u} \left(\prod_{\substack{u \in [2], \\ v_u \in [r_u]}} (n_{(u,v_u)} - 1) Q_1 \cdot Q_2 - \left(\sum_{\substack{v_1 \in [r_1], \\ v_2 \in [r_2]}} \prod_{\substack{u \in [2], \\ \hat{v}_u \in [r_u] \setminus \{v_u\}}} (n_{(u,\hat{v}_u)} - 1) \cdot n_{(u,v_u)} \cdot \mathbf{x}_{u,v_u} \right) \right).$$

It is easy to see from the above equation that for $M \in \{\emptyset, \{1\}, \{2\}\}$, the coefficient of $\prod_{u \in M} \mathbf{x}_{u,v_u}$ in g_T is equal to

$$\beta_{M,\mathbf{v}_M} = \prod_{u \in [2]} (-1)^{n_u - r_u} \prod_{v_1 \in [r_1]} (n_{(1,v_1)} - 1) \prod_{v_2 \in [r_2]} (n_{(2,v_2)} - 1) \cdot \prod_{u' \in [m] \setminus M} \alpha_{u'}$$

and for $M = \{1, 2\}$, the coefficient of $\mathbf{x}_{1,v_1} \mathbf{x}_{2,v_2}$ for $v_1 \in [r_1], v_2 \in [r_2]$ is equal to

$$\prod_{u \in [2]} (-1)^{n_u - r_u} (-1)^{|M| - 1} \left(\prod_{u \in M} \prod_{\hat{v}_u \in [r_u] \setminus \{v_u\}} (n_{(u,\hat{v}_u)} - 1) (\sum_{u \in M} n_{(u,v_u)} - 1) \right).$$

Thus, the base case holds. Now, suppose that $m \geq 3$ and the lemma holds for $(m-1)$. Recall Equation (6.11). As seen before, $D(R_{u_1} \mid R_{u_1}) = \det(B_{u_1})$ (B_{u_1} is defined in Equation (6.4) and we know $\det(B_{u_1})$ from Observation 6.8). Note that $D(\overline{R}_{u_1} \mid \overline{R}_{u_1})$ is the restriction of $H'(T)$ to the sub-matrix whose rows and columns are indexed by variables in $(m-1)$ sets $\mathbf{x}_{u'_1}, u'_1 \in A_{u_1}$. Thus, from induction hypothesis of this claim, we also know the value of $D(\overline{R}_{u_1} \mid \overline{R}_{u_1})$. If we can figure out $D(R_{u_1} \mid C'_{(u_1,v_1),(u'_1,v'_1)})$ and $D(\overline{R}_{u'_1} \mid \overline{C}'_{(u_1,v_1),(u'_1,v'_1)})$ for every $C_{(u_1,v_1),(u'_1,v'_1)} \in \mathscr{C}_{u_1,u'_1}$, we would be done. As $D(R_{u_1} \mid C'_{(u_1,v_1),(u'_1,v'_1)})$ is the minor lying in the intersection of the rows and columns of $H'(T)$ in the row block R_{u_1} and the column block $C'_{(u_1,v_1),(u'_1,v'_1)}$ respectively, Observation 6.9 implies

$$D(R_{u_1} \mid C'_{(u_1,v_1),(u'_1,v'_1)}) = (-1)^{n_{u_1}-r_{u_1}} \prod_{\hat{v}_1 \in [r_{u_1}] \setminus \{v_1\}} (n_{(u_1,\hat{v}_1)} - 1) \cdot n_{(u_1,v_1)} \cdot \frac{\mathbf{x}_{u'_1,v'_1}}{Q_{u'_1}}.$$

Let $M \in \mathscr{P}([m])$, $v_u \in [r_u]$ for $u \in M$, and $\mathbf{v} = (v_u)_{u \in M}$. It is easy to verify that the statement of the claim holds true for M satisfying either $M = \emptyset$ or $|M| = 1$. This is so because in these cases, the monomial $\prod_{u \in M} \mathbf{x}_{u,v_u}$ is only present in $Q_{u_1} Q_{u'_1} D(R_{u_1} | R_{u_1}) \cdot D(\overline{R}_{u_1} | \overline{R}_{u_1})$ and not in $Q_{u_1} Q_{u'_1} D(R_{u_1} | C'_{(u_1,v_1),(u'_1,v'_1)}) \cdot D(\overline{R}_{u_1} | \overline{C}'_{(u_1,v_1),(u'_1,v'_1)})$ for any $u'_1 \in A_{u_1}$.

Let $M = \{u_1, u'_1\}$ and $\mathbf{x}_{M,\mathbf{v}} = \prod_{u \in M} \mathbf{x}_{u,v_u}$. Observation 6.8 and the induction hypothesis of this claim applied on $D(\overline{R}_{u_1} | \overline{R}_{u_1})$ imply that the coefficient of $\mathbf{x}_{M,\mathbf{v}}$ in $T \cdot D(R_{u_1} \mid R_{u_1}) \cdot D(\overline{R}_{u_1} \mid \overline{R}_{u_1})$ is equal to

$$\prod_{u \in [m]} (-1)^{n_u - r_u} \prod_{\hat{v}_1 \in [r_{u_1}]} (n_{(u_1,\hat{v}_1)} - 1) \prod_{\hat{v}'_1 \in [r_{u'_1}]} (n_{(u'_1,\hat{v}'_1)} - 1) \prod_{u \in [m] \setminus M} \alpha_u \prod_{v \in [r_u]} (n_{(u,v)} - 1), \tag{6.15}$$

where $T = Q_1 \cdots Q_m$. It follows from Observation 6.9 and Claim 6.4.2 that the coefficient of the monomial $\mathbf{x}_{M,\mathbf{v}}$ in $T \cdot D(R_{u_1} \mid C'_{(u_1,v_{u_1}),(u'_1,v_{u'_1})}) \cdot D(\overline{R}_{u_1} \mid \overline{C}'_{(u_1,v_{u_1}),(u'_1,v_{u'_1})})$ is equal to

$$\prod_{u \in [m]} (-1)^{n_u - r_u} \prod_{\hat{v}_1 \in [r_{u_1}] \setminus \{v_{u_1}\}} (n_{u_1,\hat{v}_1} - 1) n_{(u_1,v_{u_1})} \prod_{\hat{v}'_1 \in [r_{u'_1}] \setminus \{v_{u'_1}\}} (n_{(u'_1,\hat{v}'_1)} - 1) n_{(u'_1,v_{u'_1})} \prod_{u \in [m] \setminus M} \alpha_u \prod_{v \in [r_u]} (n_{(u,v)} - 1) \tag{6.16}$$

and the coefficient of this monomial in other terms is equal to 0. On subtracting Equation (6.16) from Equation (6.15), we get the desired result for $|M| = 2$.

Let $|M| > 2$ and $M_{u_1} = M \setminus \{u_1\}$. Then, $|M_{u_1}| \geq 2$. It follows from Observation 6.8 and the induction hypothesis of this claim applied on $D(\overline{R}_{u_1} | \overline{R}_{u_1})$ that the coefficient of $\mathbf{x}_{M,\mathbf{v}}$ in

$T \cdot D(R_{u_1} \,|\, R_{u_1}) \cdot D(\overline{R}_{u_1} \,|\, \overline{R}_{u_1})$ is equal to

$$\prod_{u\in[m]} (-1)^{n_u - r_u}(-1)^{|M|-2} \prod_{\hat{v}_1 \in [r_{u_1}]} (n_{(u_1,\hat{v}_1)} - 1) \left(\prod_{\substack{u\in[m]\setminus M \\ v\in[r_u]}} (n_{(u,v)} - 1)\alpha_u \right)$$
$$\left(\prod_{\substack{\hat{u}'_1 \in M_{u_1} \\ \hat{v}'_1 \in [r_{\hat{u}'_1}]\setminus\{v_{\hat{u}'_1}\}}} (n_{(\hat{u}'_1,\hat{v}'_1)} - 1)\Big(\sum_{\hat{u}'_1 \in M_{u_1}} n_{(\hat{u}'_1, v_{\hat{u}'_1})} - 1 \Big) \right).$$

Here we have used the fact that $[m] \setminus M = ([m] \setminus \{u_1\}) \setminus M_{u_1}$. Note that $\mathbf{x}_{M,\mathbf{v}}$ is present in $T \cdot D(R_{u_1} \,|\, C'_{(u_1,v_1),(u'_1,v'_1)}) \cdot D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_1),(u'_1,v'_1)})$ only if $u'_1 \in M_{u_1}$, $v_1 = v_{u_1}$ and $v'_1 = v_{u'_1}$. This is so because every monomial in $Q_{u'_1} \cdot D(R_{u_1} | C'_{(u_1,v_1),(u'_1,v'_1)})$ contains $\mathbf{x}_{u'_1,v'_1}$ and Claim 6.4.2 implies that every monomial in $\prod_{u\in[m]\setminus\{u'_1\}} Q_u \cdot D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_1),(u'_1,v'_1)})$ contains $\mathbf{x}_{u_1,v_1}$. Let $u'_1 \in M_{u_1}$ and $M' = M_{u_1} \setminus \{u'_1\}$. Hence, $|M'| = |M| - 2$. Observation 6.9 applied on $Q_{u'_1} \cdot D(R_{u_1} | C'_{(u_1,v_{u_1}),(u'_1,v_{u'_1})})$ and Claim 6.4.2 applied on $\prod_{u\in[m]\setminus\{u'_1\}} Q_u \cdot D(\overline{R}_{u_1} | \overline{C}'_{(u_1,v_{u_1}),(u'_1,v_{u'_1})})$ imply that the coefficient of $\mathbf{x}_{M,\mathbf{v}}$ in $T \cdot D(R_{u_1} \,|\, C'_{(u_1,v_{u_1}),(u'_1,v_{u'_1})}) \cdot D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_{u_1}),(u'_1,v_{u'_1})})$ for $u'_1 \in M_{u_1}$ is equal to

$$\prod_{u\in[m]} (-1)^{n_u - r_u}(-1)^{|M'|} \prod_{\hat{v}_1 \in [r_{u_1}]\setminus\{v_{u_1}\}} (n_{(u_1,\hat{v}_1)} - 1) n_{(u_1,v_{u_1})} \left(\prod_{\substack{u\in[m]\setminus M \\ v\in[r_u]}} (n_{(u,v)} - 1)\alpha_u \right)$$
$$\left(\prod_{\substack{\hat{u}'_1 \in M_{u_1} \\ \hat{v}'_1 \in [r_{\hat{u}'_1}]\setminus\{v_{\hat{u}'_1}\}}} (n_{(\hat{u}'_1,\hat{v}'_1)} - 1) n_{(u'_1, v_{u'_1})} \right).$$

Here, we have used the fact that $[m] \setminus M = A_{u_1,u'_1} \setminus M'$, where $A_{u_1,u'_1} = [m] \setminus \{u_1, u'_1\}$. Thus, on substituting $|M'| = |M| - 2$ in the above equation, adding the coefficients of $\mathbf{x}_{M,\mathbf{v}}$ in $T \cdot D(R_{u_1} \,|\, C'_{(u_1,v_{u_1}),(u'_1,v_{u'_1})}) \cdot D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_{u_1}),(u'_1,v_{u'_1})})$ for every $u'_1 \in M_{u_1}$ and subtracting it from the coefficient of this monomial in $T \cdot D(R_{u_1} \,|\, R_{u_1}) \cdot D(\overline{R}_{u_1} \,|\, \overline{R}_{u_1})$, we get the desired result. This completes the proof of Claim 6.4.1. ■

6.4.0.2 Proof of Claim 6.4.2

We prove this by induction on $|A_{u_1}|$. Let $m = 2, |A_{u_1}| = 1, u_1 = 1, u_1' = 2$ and $v_1 \in [r_1], v_2 \in [r_2]$. Then, $A_{u_1,u_1'} = \emptyset$ and hence $M' = \emptyset$. Observation 6.9 implies that

$$Q_1 \cdot D(\overline{R}_1 \,|\, \overline{C}'_{(1,v_1),(2,v_2)}) = (-1)^{n_2 - r_2} \prod_{\hat{v}_2 \in [r_2] \setminus \{v_2\}} (n_{(2,\hat{v}_2)} - 1) n_{(2,v_2)} \mathbf{x}_{1,v_1}.$$

Thus, the base case holds.

Suppose $|A_{u_1}| \geq 2$. Let H be the sub-matrix of $H'(T)$, whose rows and columns are indexed by $\overline{R}_{u_1}$ and $\overline{C}'_{(u_1,v_1),(u_1',v_1')}$. Then, $\det(H) = D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_1),(u_1',v_1')})$. To prove the result for A_{u_1}, we would look at the Laplace's expansion of $\det(H)$ corresponding to the set of rows indexed by the ordered set $R_{u_1'} = \{(u_1', v_1', k') : v_1' \in [r_{u_1'}], k' \in [n_{(u_1',v_1')}]\}$. Consider the ordered set E, which is obtained from $R_{u_1'}$ by replacing $(u_1, v_1, 1)$ with $(u_1', v_1', 1)$. Then, E looks as

$$E = \left((u_1', 1, 1), \ldots, (u_1', v_1' - 1, n_{(u_1',v_1'-1)}), (u_1, v_1, 1), \ldots, (u_1', r_{u_1'}, n_{(u_1',r_{u_1'})}) \right).$$

Let $\overline{E} = \overline{R}_{u_1,u_1'} = \overline{R}_{u_1} \setminus R_{u_1'}$. Let $u_1'' \in A_{u_1,u_1'}$, where $A_{u_1,u_1'} = [m] \setminus \{u_1, u_1'\}$, $v_1'' \in [r_{u_1''}]$, $E_{(u_1'',v_1'')}$ be obtained from the ordered set E by replacing $(u_1, v_1, 1)$ with $(u_1'', v_1'', 1)$ and $\overline{E}_{(u_1'',v_1'')}$ be obtained from the ordered set $\overline{E}$ by replacing $(u_1'', v_1'', 1)$ with $(u_1, v_1, 1)$. Then, it is easy to verify that the structure of H, observation analogous to Observation 6.5 and the Laplace's expansion of $\det(H)$ imply

$$D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_1),(u_1',v_1')}) = D(R_{u_1'} \,|\, E) \cdot D(\overline{R}_{u_1,u_1'} \,|\, \overline{E}) - \left(\sum_{\substack{u_1'' \in A_{u_1,u_1'}, \\ v_1'' \in [r_{u_1''}]}} D(R_{u_1'} \,|\, E_{(u_1'',v_1'')}) \cdot D(\overline{R}_{u_1,u_1'} \,|\, \overline{E}_{(u_1'',v_1'')}) \right). \tag{6.17}$$

It follows from Observation 6.9 that

$$Q_{u_1} \cdot D(R_{u_1'} \,|\, E) = (-1)^{n_{u_1'} - r_{u_1'}} \prod_{\hat{v}_1' \in [r_{u_1'}] \setminus \{v_1'\}} (n_{(u_1',\hat{v}_1')} - 1) \cdot n_{(u_1',v_1')} \cdot \mathbf{x}_{u_1,v_1} \tag{6.18}$$

and for $u_1'' \in A_{u_1,u_1'}, v_1'' \in [r_{u_1''}]$,

$$Q_{u_1''} \cdot D(R_{u_1'} \,|\, E_{(u_1'',v_1'')}) = (-1)^{n_{u_1'} - r_{u_1'}} \prod_{\hat{v}_1' \in [r_{u_1'}] \setminus \{v_1'\}} (n_{(u_1',\hat{v}_1')} - 1) \cdot n_{(u_1',v_1')} \cdot \mathbf{x}_{u_1'',v_1''}. \tag{6.19}$$

Notice that $D(\overline{R}_{u_1,u_1'} \mid \overline{E})$ is the minor of $H'(T)$, whose rows and columns are indexed by $\overline{R}_{u_1,u_1'}$. Thus, it is a smaller instance of the determinant mentioned in the statement of Claim 6.4.1 as it is the product of $(m-2)$ preprocessed Hessians $H'(T_{u_1''}), u_1'' \in A_{u_1,u_1'}$. Thus, by the induction hypothesis of Claim 6.4.1, we can describe $D(\overline{R}_{u_1,u_1'} \mid \overline{E})$. Notice that $D(\overline{R}_{u_1,u_1'} \mid \overline{E}_{(u_1'',v_1'')})$ is a smaller instance of $D(\overline{R}_{u_1} \mid \overline{C}'_{(u_1,v_1),(u_1',v_1')})$ and hence we use the induction hypothesis of this claim here. Let $M' \in \mathscr{P}(A_{u_1,u_1'})$ be chosen arbitrarily. We prove the result in the following three cases.

1. If $M' = \emptyset$ then the coefficient of $\mathbf{x}_{u_1,v_1}$ in $\prod_{u\in[m]\setminus\{u_1'\}} Q_u \cdot D(R_{u_1'} \mid E_{(u_1'',v_1'')}) \cdot D(\overline{R}_{u_1,u_1'} \mid \overline{E}_{(u_1'',v_1'')})$ is equal to 0 for every $u_1'' \in A_{u_1,u_1'}, v_1'' \in [r_{u_1''}]$ and it follows from the induction hypothesis of Claim 6.4.1 the coefficient of $\mathbf{x}_{u_1,v_1}$ in $\prod_{u\in[m]\setminus\{u_1'\}} Q_u \cdot D(R_{u_1'} \mid E) \cdot D(\overline{R}_{u_1,u_1'} \mid \overline{E})$ is equal to

$$\prod_{u\in A_{u_1}} (-1)^{n_u - r_u} \prod_{\hat{v}_1'\in[r_{u_1'}]\setminus\{v_1'\}} (n_{(u_1',\hat{v}_1')} - 1) n_{(u_1',v_1')} \prod_{u_1''\in A_{u_1,u_1'}} (\alpha_{u_1''} \prod_{v_1''\in[r_{u_1''}]} (n_{(u_1'',v_1'')} - 1)).$$

   Thus, the claim holds in this case.

2. Let $M' = \{u_1''\}$ and $p = \mathbf{x}_{u_1,v_1} \cdot \mathbf{x}_{u_1'',v_1''}$. It follows from Observation 6.9 and the induction hypothesis of Claim 6.4.1 that the coefficient of p in $\prod_{u\in[m]\setminus\{u_1'\}} Q_u \cdot D(R_{u_1'}|E) \cdot D(\overline{R}_{u_1,u_1'}|\overline{E})$ is equal to

$$\prod_{u\in A_{u_1}} (-1)^{n_u - r_u} \prod_{\hat{v}_1'\in[r_{u_1'}]\setminus[v_1']} (n_{(u_1',\hat{v}_1')} - 1) n_{(u_1',v_1')} \prod_{u\in A_{u_1,u_1',u_1''}} (n_{(u,v_u)} - 1)\alpha_u \prod_{\hat{v}_1''\in[r_{u_1''}]} (n_{(u_1'',\hat{v}_1'')} - 1),$$

   where $A_{u_1,u_1',u_1''} = [m] \setminus \{u_1, u_1', u_1''\}$. Observation 6.9 applied on $Q_{u_1''} \cdot D(R_{u_1'}|E_{(u_1'',v_1'')})$ and the induction hypothesis of this claim applied on $\prod_{u\in[m]\setminus\{u_1',u_1''\}} Q_u \cdot D(\overline{R}_{u_1,u_1'}|\overline{E}_{(u_1'',v_1'')})$ implies that the coefficient of p in $\prod_{u\in[m]\setminus\{u_1'\}} Q_u \cdot D(R_{u_1'}|E_{(u_1'',v_1'')}) \cdot D(\overline{R}_{u_1,u_1'}|\overline{E}_{(u_1'',v_1'')})$ is

$$\prod_{u\in A_{u_1}} (-1)^{n_u - r_u} \prod_{\hat{v}_1'\in[r_{u_1'}]\setminus[v_1']} (n_{(u_1',\hat{v}_1')} - 1) n_{(u_1',v_1')} \prod_{u\in A_{u_1,u_1',u_1''}} (n_{(u,v_u)} - 1)\alpha_u \prod_{\hat{v}_1''\in[r_{u_1''}]\setminus\{v_1''\}} (n_{(u_1'',\hat{v}_1'')} - 1) n_{(u_1'',v_1'')}$$

   and the coefficient of this monomial in the other terms in Equation (6.17) is equal to 0. Now, the result directly follows by plugging in these coefficients of p Equation (6.17).

3. Suppose $|M'| \geq 2$. Let $v_1'' \in [r_{u_1''}]$ for $u_1'' \in M'$. By applying Observation 6.9 on $Q_{u_1} \cdot D(R_{u_1'} \mid E)$ and induction hypothesis of Claim 6.4.1 on $\prod_{\hat{u}_1''\in A_{u_1,u_1'}} Q_{\hat{u}_1''} \cdot D(\overline{R}_{u_1,u_1'} \mid \overline{E})$, the coef-

ficient of the monomial $p := \mathbf{x}_{u_1,v_1} \prod_{\hat{u}_1'' \in M'} \mathbf{x}_{\hat{u}_1'',\hat{v}_1''}$ in $Q_{u_1} \cdot \prod_{\hat{u}_1'' \in A_{u_1,u_1'}} Q_{\hat{u}_1''} \cdot D(R_{u_1'} \,|\, E) \cdot D(\overline{R}_{u_1,u_1'} \,|\, \overline{E})$ is equal to

$$\prod_{u \in A_{u_1}} (-1)^{n_u - r_u} \cdot (-1)^{|M'|-1} \prod_{\hat{v}_1' \in [r_{u_1'}] \setminus \{v_1'\}} (n_{(u_1',\hat{v}_1')} - 1) n_{(u_1',v_1')} \prod_{\substack{u \in A_{u_1,u_1'} \setminus M', \\ v \in [r_u]}} (n_{(u,v)} - 1) \cdot \alpha_u$$

$$\times \prod_{\substack{\hat{u}_1'' \in M', \\ \hat{v}_1 \in [r_{\hat{u}_1''}] \setminus \{\hat{v}_1''\}}} (n_{(\hat{u}_1'',\hat{v}_1)} - 1) \left(\sum_{\hat{u}_1'' \in M'} n_{(\hat{u}_1'',\hat{v}_1'')} - 1 \right).$$

Note that this monomial is in $Q_{u_1} \cdot \prod_{\hat{u}_1'' \in A_{u_1,u_1'}} Q_{\hat{u}_1''} \cdot D(R_{u_1'} \,|\, E_{(u_1'',\hat{v}_1'')}) \cdot D(\overline{R}_{u_1,u_1'} \,|\, \overline{E}_{(u_1'',\hat{v}_1'')})$ only if $u_1'' \in M'$ and $\hat{v}_1'' = v_1''$. This is so because every monomial in $Q_{u_1} \cdot D(R_{u_1'} | E_{(u_1'',\hat{v}_1'')})$ contains $\mathbf{x}_{(u_1'',\hat{v}_1'')}$. Thus, for $u_1'' \in M'$, by Observation 6.9 and induction hypothesis of this claim, the coefficient of p in $Q_{u_1} \cdot \prod_{\hat{u}_1'' \in A_{u_1,u_1'}} Q''_{\hat{u}_1} \cdot D(R_{u_1'} \,|\, E_{(u_1'',v_1'')}) \cdot D(\overline{R}_{u_1,u_1'} \,|\, \overline{E}_{(u_1'',v_1'')})$ is equal to

$$\prod_{u \in A_{u_1}} (-1)^{n_u - r_u} \cdot (-1)^{|M'|-1} \prod_{\hat{v}_1' \in [r_{u_1'}] \setminus \{v_1'\}} (n_{(u_1',\hat{v}_1')} - 1) n_{(u_1',v_1')} \prod_{\substack{u \in A_{u_1,u_1'} \setminus M', \\ v \in [r_u]}} (n_{(u,v)} - 1) \cdot \alpha_u$$

$$\times \left(\prod_{\substack{\hat{u}_1'' \in M', \\ \hat{v}_1 \in [r_{\hat{u}_1''}] \setminus \{\hat{v}_1''\}}} (n_{(u_1'',\hat{v}_1)} - 1) \cdot n_{(u_1'',v_1'')} \right).$$

This implies that the coefficient of p in Equation (6.17) is equal to

$$\prod_{u \in A_{u_1}} (-1)^{n_u - r_u} \cdot (-1)^{|M'|} \cdot \left(\prod_{\substack{u \in A_{u_1,u_1'} \setminus M', \\ v \in [r_u]}} (n_{(u,v)} - 1) \cdot \alpha_u \right) \times$$

$$\left(\prod_{\substack{\hat{u}_1'' \in M', \\ \hat{v}_1 \in [r_{\hat{u}_1''}] \setminus \{\hat{v}_1''\}}} (n_{(\hat{u}_1'',\hat{v}_1)} - 1) \prod_{\hat{v}_1' \in [r_{u_1'}] \setminus \{v_1'\}} (n_{(u_1',\hat{v}_1')} - 1) \cdot n_{(u_1',v_1')} \right).$$

Let $u_1'' \in A_{u_1,u_1'}, v_1'' \in [r_{u_1''}]$ be arbitrarily chosen. Then, we know that $D(\overline{R}_{u_1,u_1'} \,|\, \overline{E}_{(u_1'',v_1'')})$

is a smaller instance of $D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_1),(u'_1,v'_1)})$. Thus, by the induction hypothesis of this claim, every monomial q' in $\prod_{u\in[m]\setminus\{u'_1,u''_1\}} Q_u \cdot D(\overline{R}_{u_1,u'_1}|\overline{E}_{(u''_1,v''_1)})$ contains $\mathbf{x}_{u_1,v_1}$ and for every $(\mathbf{u}'',\mathbf{v}'') \in S_{u''_1}$, the degree of $\mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}$ in q' is at most 1. Let $\widehat{q}$ be an arbitrary monomial of $D(\overline{R}_{u_1,u'_1} \,|\, \overline{E})$. Then, it follows from Claim 6.4.1 that the degree of $\mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}$ in $\widehat{q}$ is at most 1. This along with Equations (6.18), (6.19) and (6.17) implies that if q is an arbitrary monomial in $Q_{u_1} \prod_{u''_1\in A_{u_1,u'_1}} Q_{u''_1} \cdot D(\overline{R}_{u_1} \,|\, \overline{C}'_{(u_1,v_1),(u'_1,v'_1)})$ and $(\mathbf{u}'',\mathbf{v}'') \in S_{u''_1}$ then q contains $\mathbf{x}_{u_1,v_1}$ and the degree of $\mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}$ in q is at most 1. This completes the proof of Claim 6.4.2. ■

6.5 The Hessian determinant of a general ROF

This section contains the major fraction of this chapter. We start this section by giving some useful notations, which is in continuation with the notations given in Section 6.1.

6.5.1 Notations

1. For every $(\mathbf{u},\mathbf{v}) \in S$, let $b(\mathbf{u},\mathbf{v})$ be defined as the number of multiplication gates with fan-in at least 2, in the path $(\mathbf{u},\mathbf{v})$. Recall $n_{(\mathbf{u},\mathbf{v})} = |\mathbf{x}_{(\mathbf{u},\mathbf{v})}|$. If $n_{(\mathbf{u},\mathbf{v})} \geq 2$ then we define $a(\mathbf{u},\mathbf{v}) = b(\mathbf{u},\mathbf{v})$ otherwise $a(\mathbf{u},\mathbf{v}) = b(\mathbf{u},\mathbf{v}) + 1$.

2. Let $(\mathbf{u},\mathbf{v}) = (u_1, v_1, \ldots, u_\Delta, v_\Delta) \in S$. Suppose there exists $i \in [\Delta]$, such that for every $j < i, r_{u_j} \neq 1, s_{v_j} \neq 1, r_{u_i} \neq 1$ and for every $k \in [i+1,\Delta], r_{u_k} = 1$ and $s_{v_{k-1}} = 1$. Then,

$$W_{(\mathbf{u},\mathbf{v})} := \Big\{(\mathbf{u}',\mathbf{v}') := (u_1, v_1, \cdots, u_{i-1}, v_{i-1}, u_i, v'_i, u'_{i+1}, \ldots, v'_\Delta) \,:\, v'_i \in [r_{u_i}] \setminus \{v_i\},$$
$$\text{for every } k \in [i+1,\Delta], s_{v'_{k-1}} = 1, r_{u'_k} = 1\Big\}.$$

   Suppose $(\mathbf{u}',\mathbf{v}') \in W_{(\mathbf{u},\mathbf{v})}$. Then, it is easy to see that the nodes v_i in $(\mathbf{u},\mathbf{v})$ and v'_i in $(\mathbf{u}',\mathbf{v}')$ compute monomials.

3. For every $\ell \in [\Delta]$,

   (a) Let $\Sigma_{\ell,1} := \{u_\ell \in \Sigma_\ell \,:\, \text{there exists } v_\ell \in [r_{u_\ell}], n_{v_\ell} = 1\}$, where n_{v_ℓ} is the number of variables in the $\times$-rooted ROF T_{v_ℓ}. Since T is canonical, it follows immediately that for every $u_\ell \in \Sigma_{\ell,1}$, there exists a unique $v_\ell \in [r_{u_\ell}]$, such that $n_{v_\ell} = 1$. Let $\overline{\Sigma}_{\ell,1} := \Sigma_\ell \setminus \Sigma_{\ell,1}$.

   (b) Let $M \subseteq [m]$. Then, $\Sigma_\ell^M \subseteq \Sigma_\ell$ and $\prod_\ell^M \subseteq \prod_\ell$ are such that for every $u_\ell \in \Sigma_\ell^M, v_\ell \in \prod_\ell^M$, there exist $u, u' \in M$, such that the $+$ gate u_ℓ is present in the sub-ROF Q_u

and the $\times$ gate v_ℓ is present in the sub-ROF $Q_{u'}$. Further, $\widehat{\Sigma}_\ell^M \subseteq \Sigma_\ell^M$ is such that for every $u_\ell \in \widehat{\Sigma}_\ell^M$, the parent $\times$ gate of u_ℓ has fan-in at least two.

(c)

$$\mathscr{V}_\ell = \left\{\mathfrak{v}_\ell := (v_{u_\ell})_{u_\ell \in \Sigma_\ell} : \forall u_\ell \in \Sigma_{\ell,1}, v_{u_\ell} \in [r_{u_\ell}],\ \text{s.t. } n_{v_{u_\ell}} = 1, \forall u_\ell \in \overline{\Sigma}_{\ell,1}, v_{u_\ell} \in [r_{u_\ell}]\right\}.$$

In other words, for every $\ell \in [\Delta], \mathfrak{v}_\ell$ is a tuple, whose entries are labelled by the $+$ gates in Σ_ℓ and for every $u_\ell \in \Sigma_\ell$, the u_ℓ-th entry of $\mathfrak{v}_\ell$ is exactly one non-constant child of the $+$ gate u_ℓ. Further, for $u_\ell \in \Sigma_{\ell,1}$, if $v_{u_\ell} \in [r_{u_\ell}]$ is such that $n_{u_{v_\ell}} = 1$ then the u_ℓ-th coordinate of $\mathfrak{v}_\ell$ is equal to v_{u_ℓ}. Since T is canonical, such a v_{u_ℓ} is unique.

4. Let $\mathscr{V} = \mathscr{V}_1 \times \cdots \times \mathscr{V}_\Delta$. Then, an element $\mathfrak{v} \in \mathscr{V}$ looks as $\mathfrak{v} = (\mathfrak{v}_\ell)_{\ell \in [\Delta]}$, where for every $\ell \in [\Delta], \mathfrak{v}_\ell \in \mathscr{V}_\ell$.
   The following sets would be used to define the *nice monomials* in g_T.

5. Let $u \in [m]$ and $\mathfrak{v} \in \mathscr{V}$ be fixed arbitrarily. Then,

   (a) $S_{\mathfrak{v},u,0} = \left\{(\mathbf{u},\mathbf{v}) := (u_\ell, v_{u_\ell})_{\ell \in [\Delta]} : u_1 = u, \text{ and for } \ell \in [2,\Delta], u_\ell \in [s_{v_{u_{\ell-1}}}]\right\}$. Note here that for every $\ell \in [\Delta]$, the coordinate v_{u_ℓ} in $(\mathbf{u},\mathbf{v})$ is fixed by the tuple $\mathfrak{v}$. Suppose $u \in [m]$ is such that one of its children computes a variable. Then, $|S_{\mathfrak{v},u,0}| = 1$.

   (b) For $i \in [\Delta]$,

$$\begin{aligned} S_{\mathfrak{v},u,i} = \Big\{&(\mathbf{u},\mathbf{v}) = (u_1,\ldots,v_i,u_{i+1},v_{u_{i+1}},\ldots,u_\Delta,v_{u_\Delta}) : u_1 = u, v_1 \in [r_{u_1}], \\ &\forall j \in [2,i-1], v_j \in [r_{u_j}], u_j \in [s_{v_{j-1}}], u_i \in [s_{v_{i-1}}], v_i \in [r_{u_i}] \setminus \{v_{u_i}\}, \\ &u_{i+1} \in [s_{v_i}], \forall k \in [i+2,\Delta], u_k \in [s_{v_{u_{k-1}}}]\Big\}. \end{aligned}$$

   Note that for any $(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u,i}$ and for $j \in [1,i-1]$, the coordinate v_j of $(\mathbf{u},\mathbf{v})$ is picked arbitrarily from $[r_{u_j}]$, v_i is picked arbitrarily from $[r_{u_i}] \setminus \{v_{u_i}\}$, where v_{u_i} is fixed by $\mathfrak{v}$ and for $k \in [i+1,\Delta], v_{u_k}$ is also fixed by $\mathfrak{v}$[1].

6.5.2 Technical lemmas.

Lemma 6.1 (Description of nice monomials) *Let $\Delta, n \in \mathbb{N}, \Delta \geq 2, \mathbb{F}$ be a field such that either $\mathrm{char}(\mathbb{F}) = 0$ or $\mathrm{char}(\mathbb{F}) \geq n$, $T = Q_1 \cdots Q_m$, where for every $u \in [m], Q_u$ is a $+$-rooted extended canonical ROF of product-depth Δ and there exists $u \in [m]$, such that Q_u*

[1] We emphasize here again that it is very important that for any $+$ gate u_i, $[r_{u_i}]$ does not contain the label of its constant child, if there is any.

computes a polynomial of degree at least 2 and $|\mathrm{var}(T)| = n$. *Let* d_T *be the denominator of* $\det(H'(T))$ *given in Observation 6.1 and* $g_T := d_T \cdot \det(H'(T))$. *Then,* $g_T = g_{T_1} + g_{T_2}$, *where* $g_{T_1} = \sum_{\mathfrak{v} \in \mathscr{V}} (-1)^b \beta_{\mathfrak{v}} \cdot p_{\mathfrak{v}}$, $b = \sum_{\ell \in [\Delta]} (s(\prod_\ell) - r(\Sigma_\ell)) + (m-1)$, $p_{\mathfrak{v}} = \prod_{u \in [m]} \prod_{i \in [0, \Delta-1]} \prod_{(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u,i}} \mathbf{x}_{(\mathbf{u},\mathbf{v})}^{a(\mathbf{u},\mathbf{v})-i}$ *and* $\beta_{\mathfrak{v}} = \beta_{\mathfrak{v},0} \prod_{u \in [m]} \prod_{i \in [\Delta-1]} \beta_{\mathfrak{v},u,i}$, *where*

$$\beta_{\mathfrak{v},0} = \left(\sum_{\substack{u \in [m] \\ (\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u,0}}} n_{(\mathbf{u},\mathbf{v})} - 1 \right) \prod_{\substack{u \in [m] \\ (\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u,0}}} \prod_{(\hat{\mathbf{u}},\hat{\mathbf{v}}) \in W_{(\mathbf{u},\mathbf{v})}} (n_{(\hat{\mathbf{u}},\hat{\mathbf{v}})} - 1)$$

and for every $u \in [m], i \in [\Delta - 1]$,

$$\beta_{\mathfrak{v},u,i} = \prod_{(\mathbf{u}',\mathbf{v}')_i \in B_{\mathfrak{v},u,i}} \left(\left(\sum_{\substack{(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u,i}: \\ (\mathbf{u},\mathbf{v})_i = (\mathbf{u}',\mathbf{v}')_i}} n_{(\mathbf{u},\mathbf{v})} - 1 \right) \prod_{\substack{(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u,i}, \\ (\mathbf{u},\mathbf{v})_i = (\mathbf{u}',\mathbf{v}')_i}} \prod_{(\hat{\mathbf{u}},\hat{\mathbf{v}}) \in W_{(\mathbf{u},\mathbf{v})}} (n_{(\hat{\mathbf{u}},\hat{\mathbf{v}})} - 1) \right),$$

where $B_{\mathfrak{v},u,i} = \{(\mathbf{u}',\mathbf{v}')_i = (u_1', v_1', \ldots, u_i', v_i') : (\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u,i}\}$. *Further,* $g_{T_2} \in \mathbb{F}[\mathbf{x}]$ *and* g_{T_1} *and* g_{T_2} *are monomial disjoint.*

The proof of Lemma 6.1 is dependent on the next lemma, where we understand the minor $D(\overline{R}_{u_1} | \overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ given in Equation (6.11). Let $u_1 \in [m]$. Then, recall $A_{u_1} = [m] \setminus \{u_1\}$. Further, for $\ell \in [\Delta], M \subseteq [m]$, recall the definition of the set $\widehat{\Sigma}_\ell^M$.

Lemma 6.2 *Let* $n \in \mathbb{N}$, $\mathbb{F}$ *be a field such that either* $\mathrm{char}(\mathbb{F}) = 0$ *or* $\mathrm{char}(\mathbb{F}) \geq n$, $u_1, u_1' \in [m], u_1 \neq u_1'$ *and* $A_{u_1,u_1'} = [m] \setminus \{u_1, u_1'\}$, *where* m *is the fan-in of the top-most gate of* T *and* $n = |\mathrm{var}(T)|$, *where* T *is a product-rooted extended canonical ROF considered in Lemma 6.1. Let* $(\mathbf{u},\mathbf{v}) \in S_{u_1}, (\mathbf{u}',\mathbf{v}') \in S_{u_1'}$.

1. *The denominator of* $D(\overline{R}_{u_1} | \overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ *is equal to*

$$\bar{d}_{((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))} := \frac{Q_{u_1} \prod_{\widehat{u}_1 \in A_{u_1,u_1'}} Q_{\widehat{u}_1} \prod_{\ell \in [2,\Delta]} \left(\prod_{\widehat{u}_\ell' \in \widehat{\Sigma}_\ell^{A_{u_1}}} Q_{\widehat{u}_\ell'} \right)}{\prod_{\ell \in [2,\Delta]: s_{v'_{\ell-1}} \neq 1} Q_{u_\ell'}},$$

where $u_\ell', \ell \in [2, \Delta]$ *correspond to* $(\mathbf{u}',\mathbf{v}')$. *Then,* $\widetilde{D}(\overline{R}_{u_1} | \overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) := D(\overline{R}_{u_1} | \overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) \cdot \bar{d}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ *is a polynomial in* $\mathbb{F}[\mathbf{x}]$.

2. *Let* $\mathfrak{v} \in \mathscr{V}, (\mathbf{u}', \mathbf{v}') \in S_{\mathfrak{v},u_1',0}, (\mathbf{u}, \mathbf{v}) \in S_{\mathfrak{v},u_1,0}$ *and*

$$q_{\mathfrak{v},u_1'} := \prod_{(\hat{\mathbf{u}},\hat{\mathbf{v}})\in S_{\mathfrak{v},u_1,0}} \mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})} \times \prod_{j\in[a(\mathbf{u}',\mathbf{v}')-1]} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u_1',0}:\\ (\hat{\mathbf{u}}',\hat{\mathbf{v}}')_j=(\mathbf{u}',\mathbf{v}')_j, \hat{u}'_{j+1}\neq u'_{j+1}}} \mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-j} \right)$$
$$\times \prod_{\substack{u_1''\in A_{u_1,u_1'},\\ (\mathbf{u}'',\mathbf{v}'')\in S_{\mathfrak{v},u_1'',0}}} \mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}^{a(\mathbf{u}'',\mathbf{v}'')} \times \prod_{\substack{i\in[\Delta-1],\ (\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},\widehat{u}_1',i}\\ \widehat{u}_1'\in A_{u_1}}} \mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-i}.$$

Then, the coefficient of $q_{\mathfrak{v},u_1'}$ *in* $\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ *is equal to*

$$(-1)^c\, n_{(\mathbf{u}',\mathbf{v}')} \prod_{\widehat{u}_1'\in A_{u_1}} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},\widehat{u}_1',0}\\ (\mathbf{u}'',\mathbf{v}'')\in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}} (n_{(\mathbf{u}'',\mathbf{v}'')}-1) \prod_{i\in[\Delta-1]} \beta_{\mathfrak{v},\widehat{u}_1',i} \right),$$

where $c = \sum_{\ell\in[\Delta]} \left(s(\prod_\ell^{A_{u_1}}) - r(\Sigma_\ell^{A_{u_1}}) \right) + |A_{u_1}| - 1$ *and* $\beta_{\mathfrak{v},\widehat{u}_1',i}$ *is defined in Lemma 6.1.*

3. *Let* $(\mathbf{u}', \mathbf{v}') \in S_{u_1'}$ *be picked arbitrarily and* q *be an arbitrary monomial of* $\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$. *Let* $(\hat{\mathbf{u}}', \hat{\mathbf{v}}') \in S_{u_1'}$ *be such that there exists* $i \in [\Delta]$, *such that either* $(\mathbf{u}', \mathbf{v}')_{i-1} = (\hat{\mathbf{u}}', \hat{\mathbf{v}}')_{i-1}$, $u'_i = \hat{u}'_i$ *and* $v'_i \neq \hat{v}'_i$ *or* $(\mathbf{u}', \mathbf{v}')_i = (\hat{\mathbf{u}}', \hat{\mathbf{v}}')_i$ *and* $u'_{i+1} \neq \hat{u}'_{i+1}$. *Then,* $\deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}} q \leq (a(\hat{\mathbf{u}}', \hat{\mathbf{v}}') - i)$[1]. *Let* $u_1'' \in A_{u_1,u_1'}, (\mathbf{u}'', \mathbf{v}'') \in S_{u_1''}$. *Then,* $\deg_{(\mathbf{u}'',\mathbf{v}'')} q \leq a(\mathbf{u}'', \mathbf{v}'')$.

Remark 6.4 1. *Note that if* $\mathbb{F}$ *is a finite field with* $char(\mathbb{F}) \geq n$ *or* $char(\mathbb{F}) = 0$ *then for every* $\mathfrak{v} \in \mathscr{V}$, *the coefficients of* $p_{\mathfrak{v}}$ *in* g_T *and* $q_{\mathfrak{v},u_1'}$ *in* $\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ *is non-zero. This is so because* $n = \sum_{(\mathbf{u},\mathbf{v})\in S} n_{(\mathbf{u},\mathbf{v})}$.

2. *Let* $j \in [a(\mathbf{u}', \mathbf{v}') - 1], (\hat{\mathbf{u}}', \hat{\mathbf{v}}') \in S_{\mathfrak{v},u_1',0}$, *such that* $(\hat{\mathbf{u}}', \hat{\mathbf{v}}')_j = (\mathbf{u}', \mathbf{v}')_j$ *and* $\hat{u}'_{j+1} \neq u'_{j+1}$. *Observe that this immediately implies* $a(\hat{\mathbf{u}}', \hat{\mathbf{v}}') \geq j$. *Similarly, for* $i \in [\Delta - 1], \hat{u}_1' \in A_{u_1}, (\hat{\mathbf{u}}', \hat{\mathbf{v}}') \in S_{\mathfrak{v},\widehat{u}_1',i}, a(\hat{\mathbf{u}}', \hat{\mathbf{v}}') \geq i$. *This implies that* $p_{\mathfrak{v}}$ *in Lemma 6.1 and* $q_{\mathfrak{v},u_1'}$ *in Lemma 6.2 are monomials in* $\mathbf{x}$ *variables.*

3. *The notation* $\prod_{\ell\in[2,\Delta]:\, s_{v'_{\ell-1}}\neq 1} Q_{u'_\ell}$ *means the product of all +-rooted sub-ROFs except* $Q_{u_1'}$ *on the path* $(\mathbf{u}', \mathbf{v}')$ *whose parent product gates have fan-in at least 2.*

[1] $\deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}} q$ means the degree of $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ in q.

Lemma 6.1 and Lemma 6.2 complement each other. We call the monomials of g_T as the nice monomials in g_T. Before proceeding with the proofs, we want to draw the attention of the reader that in the product-depth 2 case (Claim 6.4.1), the nice monomials in g_T are the monomials of the kind $\prod_{u\in M}\mathbf{x}_{u,v_u}$, where $M=[m]$ and for $u\in M$, if Q_u contains a variable then $v_u\in[r_u]$ is such that $n_{(u,v_u)}=1$ otherwise $v_u\in[r_u]$ is arbitrary.

6.5.3 Proof of Lemma 6.1

We have already seen the proof of this for $\Delta=1$ in Claim 6.4.1. We now prove this lemma for a fixed $\Delta\geq 2$. Let $D=\det(H'(T))$ and $u_1\in[m]$ be fixed arbitrarily. We first give the high level overview of the proof.

Proof idea. Recall $g_T=d_T\cdot\det(H'(T))$. We look at the Laplace's expansion of g_T after clearing the denominators in Equation (6.11). We aim to find the coefficients of the set of monomials $p_{\mathfrak{v}},\mathfrak{v}\in\mathscr{V}$ in g_T. In Equation (6.11), we have a positive part and a negative part. We find the coefficient of a fixed $p_{\mathfrak{v}}$ in the positive part using the induction on Lemma 6.1. We want to mention here that in the proof, we use induction at two level: One at the top fan-in of T for a fixed product-depth Δ and other at the product-depth of the underlying $\times$-rooted ROF. After that, we find the coefficient of $p_{\mathfrak{v}}$ in the negative part of the Laplace's equation. An immediate problem is that there are more than one of negative 'terms' in the Laplace's expansion of g_T, captured by the set $\mathscr{C}_{u_1,u_1'}$ for $u_1'\in A_{u_1}$ and it is not clear which 'term' contains the monomial $p_{\mathfrak{v}}$. Using the structure of $p_{\mathfrak{v}}$, we are able to describe all the negative terms in the Laplace's expansion of g_T, in which the coefficient of $p_{\mathfrak{v}}$ is non-zero (see Claim 6.5.1). Then, we use Lemma 6.2 to compute the coefficient of $p_{\mathfrak{v}}$ in the negative part and by subtracting it from the coefficient of $p_{\mathfrak{v}}$ in the positive part, we get its coefficient in g_T.

Now, we start the proof. We start with normalizing Equation (6.11) by removing the denominators of every term involved in this equation. We first calculate the denominators of $D(R_{u_1}|R_{u_1})\cdot D(\overline{R}_{u_1}|\overline{R}_{u_1})$.

1. As $D(R_{u_1}|R_{u_1})=\prod_{v_1\in[r_{u_1}]}\det(H'(T_{v_1}))$, it follows from the first point of Remark 6.3 given after Equation (6.11) that by invoking Observation 6.1 on $\det(H'(T_{v_1}))$, which is a product-depth $(\Delta-1)$ instance of $\det(H'(T))$, for every $v_1\in[r_{u_1}]$, we get that the

denominator of $D(R_{u_1}|R_{u_1})$ is equal to

$$d_{u_1} := \prod_{v_1 \in [r_{u_1}]} \prod_{\ell \in [2,\Delta]} \left(\prod_{\substack{u_\ell \in [s_{v_{\ell-1}}]: s_{v_{\ell-1}} \neq 1, \\ v_\ell \in [r_{u_\ell}]}} Q_{u_\ell} \right).$$

2. As noted in Remark 1 given after Equation (6.11), $D(\overline{R}_{u_1}|\overline{R}_{u_1})$ is a product-depth Δ and top fan-in $(m-1)$ instance of $\det(H'(T))$. Thus, by changing $[m]$ to A_{u_1} in the definition of d_T given in Observation 6.1, we get that the denominator of $D(\overline{R}_{u_1}|\overline{R}_{u_1})$ is equal to

$$\bar{d}_{u_1} := \prod_{u'_1 \in A_{u_1}} Q_{u'_1} \prod_{\ell \in [2,\Delta]} \left(\prod_{\substack{u'_\ell \in [s_{v'_{\ell-1}}]: s_{v'_{\ell-1}} \neq 1, \\ v'_\ell \in [r_{u'_\ell}]}} Q_{u'_\ell} \right).$$

This implies that the denominator of $D(R_{u_1}|R_{u_1}) \cdot D(\overline{R}_{u_1}|\overline{R}_{u_1})$ is equal to $\frac{d_T}{Q_{u_1}}$. Now, we calculate the denominator of $D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) \cdot D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$, where $(\mathbf{u},\mathbf{v}) \in S_{u_1}$ and $(\mathbf{u}',\mathbf{v}') \in S_{u'_1}$ are arbitrary.

1. As noted in Remark 2 given after Equation (6.11), $D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ is a smaller instance of $D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$. Now, if we set $u_1 = u'_1, A_{u_1} = \{u_1\}, A_{u_1,u'_1} = \emptyset$ and change $s_{v'_{\ell-1}}$ to $s_{v_{\ell-1}}$, u'_ℓ to u_ℓ for $\ell \in [2,\Delta]$ in Lemma 6.2, we get that the denominator of $D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ is equal to

$$d_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} := \frac{Q_{u'_1} \prod_{\ell \in [2,\Delta]} \prod_{\widehat{u}_\ell \in \widehat{\Sigma}_\ell^{\{u_1\}}} Q_{\widehat{u}_\ell}}{\prod_{\ell \in [2,\Delta]: s_{v_{\ell-1}} \neq 1} Q_{u_\ell}}.$$

2. The denominator of $D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ is given in Lemma 6.2.

Then, it is not difficult to see that the denominator of $D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) \cdot D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ is equal to $\left(\frac{d_T}{\prod_{\ell \in [2,\Delta]: s_{v_{\ell-1}} \neq 1} Q_{u_\ell} \times \prod_{\ell \in [2,\Delta]: s_{v'_{\ell-1}} \neq 1} Q_{u'_\ell}} \right)$. Let

$$\widetilde{D}(R_{u_1}|R_{u_1}) = d_{u_1} \cdot D(R_{u_1}|R_{u_1}),\ \widetilde{D}(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = d_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} \cdot D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}),$$

$$\widetilde{D}(\overline{R}_{u_1}|\overline{R}_{u_1}) = \bar{d}_{u_1} \cdot D(\overline{R}_{u_1}|\overline{R}_{u_1}),\ \widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \bar{d}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} \cdot D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}),$$

where $\bar{d}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ is defined in Lemma 6.2 and $g_T = d_T \cdot D$. Then, Observation 6.1 implies that Equation (6.11) can be rewritten as

$$\begin{aligned} g_T =& Q_{u_1} \cdot \widetilde{D}(R_{u_1}|R_{u_1})\widetilde{D}(\overline{R}_{u_1}|\overline{R}_{u_1}) - \sum_{\substack{u'_1 \in A_{u_1},\\ C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} \in \mathscr{C}_{u_1,u'_1}}} \prod_{\ell \in [2,\Delta]:\, s_{v_{\ell-1}} \neq 1} Q_{u_\ell} \\ & \times \prod_{\ell \in [2,\Delta]:\, s_{v'_{\ell-1}} \neq 1} Q_{u'_\ell} \times \widetilde{D}(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}). \end{aligned} \tag{6.20}$$

For simplicity, we use the terminology $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20) to refer to the polynomial $\prod_{\ell \in [2,\Delta]:\, s_{v_{\ell-1}} \neq 1} Q_{u_\ell} \times \prod_{\ell \in [2,\Delta]:\, s_{v'_{\ell-1}} \neq 1} Q_{u'_\ell} \times \widetilde{D}(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$. Let $(\mathbf{u},\mathbf{v}) \in S_{u_1}$ and $(\mathbf{u}',\mathbf{v}') \in S_{u'_1}$, such that $u_1 \neq u'_1$. Let $H_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'),R_{u_1}}$ and $H_{(\mathbf{u}',\mathbf{v}'),(\mathbf{u},\mathbf{v}),\overline{R}_{u_1}}$ be the $n_{u_1} \times n_{u_1}$ and $(n-n_{u_1}) \times (n-n_{u_1})$ size sub-matrices of $H'(T)$, whose rows are indexed by R_{u_1} and $\overline{R}_{u_1}$ and columns are indexed by $C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ and $\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ respectively, such that $D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \det(H_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'),R_{u_1}})$ and $D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \det(H_{(\mathbf{u}',\mathbf{v}'),(\mathbf{u},\mathbf{v}),\overline{R}_{u_1}})$. It follows from Observation 6.2 that the only non-zero entry of the $(\mathbf{u}',\mathbf{v}',1)$-th column of $H_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'),R_{u_1}}$ is $\left(\frac{\mathbf{x}_{(\mathbf{u}',\mathbf{v}')} \prod_{\ell \in [2,\Delta]} \prod_{\hat{u}'_\ell \in [s_{v'_{\ell-1}}] \setminus \{u'_\ell\}} Q_{\hat{u}'_\ell}}{Q_{u'_1}}\right)$, where $u'_\ell, \ell \in [\Delta]$ correspond to $(\mathbf{u}',\mathbf{v}')$ and the only non-zero entry of the $(\mathbf{u},\mathbf{v},1)$-th column of $H_{(\mathbf{u}',\mathbf{v}'),(\mathbf{u},\mathbf{v}),\overline{R}_{u_1}}$ is $\left(\frac{\mathbf{x}_{(\mathbf{u},\mathbf{v})} \prod_{\ell \in [2,\Delta]} \prod_{\hat{u}_\ell \in [s_{v_{\ell-1}}] \setminus \{u_\ell\}} Q_{\hat{u}_\ell}}{Q_{u_1}}\right)$, where $u_\ell, \ell \in [\Delta]$ correspond to $(\mathbf{u},\mathbf{v})$. This along with Observation 6.2 implies the following. Recall from the notations that R is the set of indices of variables.

Observation 6.10 *Let* $\mathbf{j}$ *be the 0-1 column vector, whose entries are labelled by* $(\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{k}) \in R$*, such that if* $\hat{k} = 1$ *then the* $(\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{k})$*-entry of* $\mathbf{j}$ *is 1, otherwise 0. Let* $H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1}}$ *and* $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}}$ *be obtained by replacing the columns labelled by* $(\mathbf{u}',\mathbf{v}',1)$ *and* $(\mathbf{u},\mathbf{v},1)$ *in* $H_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'),R_{u_1}}$ *and* $H_{(\mathbf{u}',\mathbf{v}'),(\mathbf{u},\mathbf{v}),\overline{R}_{u_1}}$ *with* $\mathbf{j}$ *restricted to* R_{u_1} *and* $\overline{R}_{u_1}$ *respectively. Then,*

$$D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \left(\frac{\mathbf{x}_{(\mathbf{u}',\mathbf{v}')} \prod_{\ell \in [2,\Delta]} \prod_{\hat{u}'_\ell \in [s_{v'_{\ell-1}}] \setminus \{u'_\ell\}} Q_{\hat{u}'_\ell}}{Q_{u'_1}}\right) \cdot \det(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1}})$$

and

$$D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \left(\frac{\mathbf{x}_{(\mathbf{u},\mathbf{v})}\prod\limits_{\ell\in[2,\Delta]\hat{u}_\ell\in[s_{v_{\ell-1}}]\setminus\{u_\ell\}} Q_{\hat{u}_\ell}}{Q_{u_1}}\right)\cdot \det(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}}).$$

Recall that the denominators of $D(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ *and* $D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ *are* $d_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ *and* $\bar{d}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ *respectively. Then, the denominators of* $\det(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1}})$ *and* $\det(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u'_1}})$ *are* $\frac{d_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}}{Q_{u'_1}}$ *and* $\frac{\bar{d}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}}{Q_{u_1}}$*, respectively. Let* $\widetilde{\det}(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1}}) = \det(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1}}) \times \frac{d_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}}{Q_{u'_1}}$ *and* $\widetilde{\det}(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u'_1}}) = \det(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u'_1}}) \times \frac{\bar{d}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}}{Q_{u_1}}$*. Then, after substituting the values of* $d_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ *and* $\bar{d}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ *in the above two equations, we get*

$$\prod_{\ell\in[2,\Delta]:\, s_{v_{\ell-1}}\neq 1} Q_{u_\ell}\cdot \widetilde{D}(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \mathbf{x}_{(\mathbf{u}',\mathbf{v}')}\prod_{\ell\in[2,\Delta]}\prod_{\hat{u}'_\ell\in[s_{v'_{\ell-1}}]:\, s_{v'_{\ell-1}}\neq 1} Q_{\hat{u}'_\ell}\times\widetilde{\det}(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1}}) \tag{6.21}$$

and

$$\prod_{\ell\in[2,\Delta]:\, s_{v'_{\ell-1}}\neq 1} Q_{u'_\ell}\cdot \widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \mathbf{x}_{(\mathbf{u},\mathbf{v})}\prod_{\ell\in[2,\Delta]}\prod_{\hat{u}_\ell\in[s_{v_{\ell-1}}]:\, s_{v_{\ell-1}}\neq 1} Q_{\hat{u}_\ell}\times\widetilde{\det}(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}}). \tag{6.22}$$

Remark 6.5 *Let* $(\hat{\mathbf{u}},\hat{\mathbf{v}})\in S$*. Then, the notation* $\prod\limits_{\ell\in[2,\Delta]\hat{u}_\ell\in[s_{\hat{v}_{\ell-1}}]:\, s_{\hat{v}_{\ell-1}}\neq 1} Q_{\hat{u}_\ell}$ *means the product of the children of all the* $\times$ *gates on the path* $(\hat{\mathbf{u}},\hat{\mathbf{v}})$ *except* $\hat{v}_\Delta$*, such that each of these* $\times$ *gates has fan-in at least 2.*

Proof: It is shown in Claim 6.5.2 that the denominators of $\det(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1}})$ and $\det(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u'_1}})$ are $\frac{d_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}}{Q_{u'_1}}$ and $\frac{\bar{d}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}}{Q_{u_1}}$, respectively. The remaining details are easy to verify. □

It is easy to note the following from the structures of $H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1}}$ and $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}}$.

Observation 6.11 *Let* $u_1,u'_1\in[m], u_1\neq u'_1, p_1,p_2$ *be arbitrary monomials of* $\widetilde{\det}(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1}})$ *and* $\widetilde{\det}(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u'_1}})$ *respectively and* $(\hat{\mathbf{u}},\hat{\mathbf{v}})\in S_{u_1}$ *and* $(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{u'_1}$*. Then,* p_1 *and* p_2 *do not contain* $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ *and* $\mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}$ *respectively.*

We now calculate the coefficient of the monomial $p_{\mathfrak{v}}$ in g_T, where $\mathfrak{v}\in\mathscr{V}$. Recall that $p_{\mathfrak{v}} = \prod\limits_{u\in[m]}\prod\limits_{i\in[0,\Delta-1]}\prod\limits_{(\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u,i}}\mathbf{x}_{(\mathbf{u},\mathbf{v})}^{a(\mathbf{u},\mathbf{v})-i}$, where $S_{\mathfrak{v},u,i}$ and $a(\mathbf{u},\mathbf{v})$ are given in Points 5 and 1 of the second part of the notations. The following claim is very helpful as it shows which all terms in the negative part of Equation (6.20) contain $p_{\mathfrak{v}}$.

Claim 6.5.1 *Let $u_1 \in [m], \mathfrak{v} \in \mathscr{V}, u_1' \in A_{u_1}, (\mathbf{u},\mathbf{v}) \in S_{u_1}$ and $(\mathbf{u}',\mathbf{v}') \in S_{u_1'}$. Then, the monomial $p_{\mathfrak{v}}$ is in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20) if and only if $(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u_1,0}$ and $(\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u_1',0}$.*

We first complete the proof of Lemma 6.1 assuming the above claim, whose proof is given in Section 6.5.4. Let $\mathfrak{v} \in \mathscr{V}$ be fixed arbitrarily. We now use induction on m of this lemma for a fixed product-depth Δ to prove that the coefficient of $p_{\mathfrak{v}}$ in g_T is $\beta_{\mathfrak{v}}$. As m is the fan-in of the top multiplication gate in T, $m \geq 2$. We analyse the coefficient of $p_{\mathfrak{v}}$ in the following two cases. Recall the definition of $\Sigma_{\ell,1}$ for some $\ell \in [\Delta]$ from Point 3 of the second part of the notations.

Case 1: $|\{u_1, u_1'\} \cap \Sigma_{1,1}| \geq 1$. Without loss of generality, let $u_1 \in \Sigma_{1,1}$. Then, we know that there exists a unique $v_1 \in [r_{u_1}]$, such that $n_{v_1} = 1$. Let $(\mathbf{u},\mathbf{v}) \in S$, be such that the first two entries of $(\mathbf{u},\mathbf{v})$ from the left are u_1 and v_1 and for every $\ell \in [2,\Delta], u_\ell = v_\ell = 1$. It is important to note that $S_{\mathfrak{v},u_1,0} = \{(\mathbf{u},\mathbf{v})\}$, $a(\mathbf{u},\mathbf{v}) = 1$ and $n_{(\mathbf{u},\mathbf{v})} = |\mathbf{x}_{(\mathbf{u},\mathbf{v})}| = 1$. Let H_1' be the sub-matrix of $H'(T)$, whose rows and columns are labelled by R_{u_1}. Then, observe that all the entries of the column of H_1' labelled by the variable $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ are zero. As $D(R_{u_1}|R_{u_1}) = \det(H_1')$, we get that $D(R_{u_1}|R_{u_1}) = 0$. Then, Equation (6.20) looks as

$$g_T = - \sum_{\substack{u_1' \in A_{u_1}, \\ C_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} \in \mathscr{C}_{u_1,u_1'}}} \left(\prod_{\ell \in [2,\Delta]:\, s_{v_{\ell-1}} \neq 1} Q_{u_\ell} \times \prod_{\ell \in [2,\Delta]:\, s_{v'_{\ell-1}} \neq 1} Q_{u'_\ell} \right. \\ \left. \times \widetilde{D}(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) \widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) \right), \tag{6.23}$$

Let $(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u_1,0}, (\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u_1',0}$ be picked arbitrarily. We first give a factorization of $p_{\mathfrak{v}}$ in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20), which would be helpful in figuring out its coefficient in this term. Recall from Lemma 6.2 that

$$q_{\mathfrak{v},u_1'} = \prod_{(\hat{\mathbf{u}},\hat{\mathbf{v}}) \in S_{\mathfrak{v},u_1,0}} \mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})} \times \prod_{j \in [a(\mathbf{u}',\mathbf{v}')-1]} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}') \in S_{\mathfrak{v},u_1',0}: \\ (\hat{\mathbf{u}}',\hat{\mathbf{v}}')_j = (\mathbf{u}',\mathbf{v}')_j, \hat{u}'_{j+1} \neq u'_{j+1}}} \mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-j} \right) \\ \times \prod_{\substack{u_1'' \in A_{u_1,u_1'}, \\ (\mathbf{u}'',\mathbf{v}'') \in S_{\mathfrak{v},u_1'',0}}} \mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}^{a(\mathbf{u}'',\mathbf{v}'')} \times \prod_{\substack{i \in [\Delta-1], \\ \widehat{u}_1' \in A_{u_1}}} \prod_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}') \in S_{\mathfrak{v},\widehat{u}_1',i}} \mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-i}.$$

Let $q_{\mathfrak{v},u_1}$ be obtained from the monomial $q_{\mathfrak{v},u_1'}$ by making the following changes: change $(\mathbf{u}',\mathbf{v}')$

to $(\mathbf{u},\mathbf{v})$, set $A_{u_1,u_1'}=\emptyset, u_1'=u_1, u_1=u_1'$, and $A_{u_1}=\{u_1\}$. Then, $q_{\mathfrak{v},u_1}$ looks as

$$q_{\mathfrak{v},u_1} := \prod_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u_1',0}} \mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')} \times \prod_{j\in[a(\mathbf{u},\mathbf{v})-1]} \left(\prod_{\substack{(\hat{\mathbf{u}},\hat{\mathbf{v}})\in S_{\mathfrak{v},u_1,0}:\\ (\hat{\mathbf{u}},\hat{\mathbf{v}})_j=(\mathbf{u},\mathbf{v})_j, \hat{u}_{j+1}\neq u_{j+1}}} \mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}^{a(\hat{\mathbf{u}},\hat{\mathbf{v}})-j} \right) \prod_{i\in[\Delta-1]} \prod_{(\hat{\mathbf{u}},\hat{\mathbf{v}})\in S_{\mathfrak{v},u_1,i}} \mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}^{a(\hat{\mathbf{u}},\hat{\mathbf{v}})-i}. \tag{6.24}$$

Let

$$p_{u_1} = \mathbf{x}_{(\mathbf{u},\mathbf{v})}^{a(\mathbf{u},\mathbf{v})-1} \cdot \prod_{j\in[a(\mathbf{u},\mathbf{v})-1]} \left(\prod_{\substack{(\hat{\mathbf{u}},\hat{\mathbf{v}})\in S_{\mathfrak{v},u_1,0}:\\ (\hat{\mathbf{u}},\hat{\mathbf{v}})_j=(\mathbf{u},\mathbf{v})_j, \hat{u}_{j+1}\neq u_{j+1}}} \mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}^{j-1} \right),$$

$$p_{u_1'} = \mathbf{x}_{(\mathbf{u}',\mathbf{v}')}^{a(\mathbf{u}',\mathbf{v}')-1} \cdot \prod_{j\in[a(\mathbf{u}',\mathbf{v}')-1]} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u_1',0}:\\ (\hat{\mathbf{u}}',\hat{\mathbf{v}}')_j=(\mathbf{u}',\mathbf{v}')_j, \hat{u}'_{j+1}\neq u'_{j+1}}} \mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{j-1} \right).$$

Then, it is easy to observe that

$$p_{\mathfrak{v}} = q_{\mathfrak{v},u_1} \cdot q_{\mathfrak{v},u_1'} \cdot p_{u_1} \cdot p_{u_1'}. \tag{6.25}$$

Further, observe that p_{u_1} and $p_{u_1'}$ are contributed by $\prod_{\ell\in[2,\Delta]: s_{v_{\ell-1}}\neq 1} Q_{u_\ell}$ and $\prod_{\ell\in[2,\Delta]: s_{v'_{\ell-1}}\neq 1} Q_{u'_\ell}$ respectively. The following observation argues that the only factorization of $p_{\mathfrak{v}}$ in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20) is given by Equation (6.25).

Observation 6.12 *Let $u_1'\in A_{u_1}, (\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u_1,0}$ and $(\mathbf{u}',\mathbf{v}')\in S_{\mathfrak{v},u_1',0}$ be arbitrary. Let $p_{\mathfrak{v}} = h_1\cdot h_2\cdot h_3\cdot h_4$, where h_1,h_2,h_3,h_4 are monomials of $f_1 := \prod_{\ell\in[2,\Delta]: s_{v_{\ell-1}}\neq 1} Q_{u_\ell}$, $f_2 := \prod_{\ell\in[2,\Delta]: s_{v'_{\ell-1}}\neq 1} Q_{u'_\ell}$, $f_3 := \widetilde{D}(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ and $f_4 := \widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ respectively. Then, $h_1 = p_{u_1}, h_2 = p_{u_1'}, h_3 = q_{\mathfrak{v},u_1}$ and $h_4 = q_{\mathfrak{v},u_1'}$, where $p_{u_1}, p_{u_1'}, q_{\mathfrak{v},u_1}$ and $q_{\mathfrak{v},u_1'}$ are the monomials considered in Equation (6.25).*

Proof: Recall that $p_{\mathfrak{v}} = \prod_{u\in[m]}\prod_{i\in[0,\Delta-1]}\prod_{(\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u,i}} \mathbf{x}_{(\mathbf{u},\mathbf{v})}^{a(\mathbf{u},\mathbf{v})-i}$. First observe that for arbitrary $i_1,i_2\in[\Delta-1], (\hat{\mathbf{u}},\hat{\mathbf{v}})\in S_{\mathfrak{v},u_1,i_1}$ and $(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u_1',i_2}, \mathbf{x}_{(\mathbf{u},\mathbf{v})}$ and $\mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}$ are not present in any monomial of f_1 and f_2. Let $(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u_1',0}, j\in[a(\mathbf{u}',\mathbf{v}')-1]$, such that $(\hat{\mathbf{u}}',\hat{\mathbf{v}}')_j = (\mathbf{u}',\mathbf{v}')_j$ and $\hat{u}'_{j+1}\neq u'_{j+1}$.

Then, it follows from Lemma 6.2 that $\deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}} h_4 \leq a(\hat{\mathbf{u}}',\hat{\mathbf{v}}') - j$. Since $\widetilde{D}(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ is a smaller instance of $\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$, we can show that if $(\hat{\mathbf{u}},\hat{\mathbf{v}}) \in S_{\mathfrak{v},u_1,0}, j' \in [a(\mathbf{u},\mathbf{v})-1]$, such that $(\hat{\mathbf{u}},\hat{\mathbf{v}})_{j'} = (\mathbf{u},\mathbf{v})_{j'}$ and $\hat{u}_{j'+1} \neq u_{j'+1}$ then $\deg_{\mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}} h_3 \leq a(\hat{\mathbf{u}},\hat{\mathbf{v}}) - j'$. Notice that $\deg_{\mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}} f_3 = a(\hat{\mathbf{u}},\hat{\mathbf{v}}) - j'$ and $\deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}} f_4 = a(\hat{\mathbf{u}}',\hat{\mathbf{v}}') - j$, $\deg_{\mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}} f_1 = j' - 1, \deg_{\mathbf{x}_{(\mathbf{u},\mathbf{v})}} f_1 = a(\mathbf{u},\mathbf{v}) - 1$, $\deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}} f_2 = j-1$ and $\deg_{\mathbf{x}_{(\mathbf{u}',\mathbf{v}')}} f_2 = a(\mathbf{u}',\mathbf{v}')-1$. This is so because $q_{\mathfrak{v},u_1}, q_{\mathfrak{v},u'_1}, p_{u_1}$ and $p_{u'_1}$ are monomials of f_3, f_4, f_1, and f_2 respectively. Then, it follows from Observation 6.14 that $\deg_{\mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}} f_3 = \deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}} f_4 = 1$. Further, Equations (6.21) and (6.22) imply that $\deg_{\mathbf{x}_{(\mathbf{u}',\mathbf{v}')}} f_3 = \deg_{\mathbf{x}_{(\mathbf{u},\mathbf{v})}} f_4 = 1$. Now, it is easy to see that this immediately implies that $h_1 = p_{u_1}$ and $h_2 = p_{u'_1}$. This implies that h_3 and h_4 should contain $\prod_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u'_1,0}} \mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ and $\prod_{(\hat{\mathbf{u}},\hat{\mathbf{v}})\in S_{\mathfrak{v},u_1,0}} \mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}$ respectively. Let $i_1, i_2 \in [\Delta-1], (\hat{\mathbf{u}},\hat{\mathbf{v}}) \in S_{\mathfrak{v},u_1,i_1}$ and $(\hat{\mathbf{u}}',\hat{\mathbf{v}}') \in S_{\mathfrak{v},u'_1,i_2}$ be arbitrary. Then, it is not difficult to see from Equations (6.21) and (6.22) that $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ and $\mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}$ are not present in f_3 and f_4 respectively. This is so because $\widetilde{\det}(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1}})$ and $\widetilde{\det}(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}})$ given in Equations (6.21) and (6.22) do not contain $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ and $\mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}$ respectively. It follows from this discussion that $h_3 = q_{\mathfrak{v},u_1}$ and $h_4 = q_{\mathfrak{v},u'_1}$. □

The above observation and Equation (6.25) imply that the coefficient of $p_{\mathfrak{v}}$ in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term is the product of coefficients of $q_{\mathfrak{v},u_1}$ and $q_{\mathfrak{v},u'_1}$ in $\widetilde{D}(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ and $\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ respectively. Now, we find the coefficient of $p_{\mathfrak{v}}$ in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20) by induction on m.

Base Case: $m = 2$. Let $[m] = \{u_1, u'_1\}$. Pick $(\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u'_1,0}$ arbitrarily. Observe that in this case, $\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \widetilde{D}(R_{u'_1}|C'_{(\mathbf{u}',\mathbf{v}'),(\mathbf{u},\mathbf{v})})$, which implies that $\widetilde{D}(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ and $\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ are similar to each other. Let $q_{\mathfrak{v},u'_1}$ be as defined in Lemma 6.2 (note that in the base case $A_{u_1,u'_1} = \emptyset$). From Lemma 6.2, the coefficient of $q_{\mathfrak{v},u'_1}$ in $\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ is

$$(-1)^{c_1} \cdot n_{(\mathbf{u}',\mathbf{v}')} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u'_1,0} \\ (\mathbf{u}'',\mathbf{v}'')\in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}} (n_{(\mathbf{u}'',\mathbf{v}'')} - 1) \right) \prod_{i\in[\Delta-1]} \beta_{\mathfrak{v},u'_1,i},$$

where $c_1 = \sum_{\ell\in[\Delta]} \left(s(\prod_\ell^{A_{u_1}}) - r(\Sigma_\ell^{A_{u_1}}) \right)$ and $A_{u_1} = \{u'_1\}$. Recall how $q_{\mathfrak{v},u_1}$ was obtained from $q_{\mathfrak{v},u'_1}$. This implies the following.

Observation 6.13 *The coefficient of the monomial $q_{\mathfrak{v},u_1}$ in $\widetilde{D}(R_{u_1}|C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ is equal to*

$$(-1)^{c_2}\, n_{(\mathbf{u},\mathbf{v})} \left(\prod_{\substack{(\hat{\mathbf{u}},\hat{\mathbf{v}})\in S_{\mathfrak{v},u_1,0}\\ (\mathbf{u}'',\mathbf{v}'')\in W_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}}} (n_{(\mathbf{u}'',\mathbf{v}'')}-1) \right) \prod_{i\in[\Delta-1]} \beta_{\mathfrak{v},u_1,i},$$

where $c_2 = \sum_{\ell\in[\Delta]} \Big(s(\prod_\ell^{\{u_1\}}) - r(\Sigma_\ell^{\{u_1\}}) \Big)$.

Then, the coefficient of $p_{\mathfrak{v}}$ in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20) is equal to

$$(-1)^{b-1}\, n_{(\mathbf{u},\mathbf{v})} \cdot n_{(\mathbf{u}',\mathbf{v}')} \prod_{u\in[m]} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u,0}\\ (\mathbf{u}'',\mathbf{v}'')\in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}} (n_{(\mathbf{u}'',\mathbf{v}'')}-1) \prod_{i\in[\Delta-1]} \beta_{\mathfrak{v},u,i} \right). \tag{6.26}$$

where $b = \sum_{\ell\in[\Delta]} (s(\prod_\ell) - r(\Sigma_\ell)) + (m-1)$. Since $S_{\mathfrak{v},u_1,0} = \{(\mathbf{u},\mathbf{v})\}$, Claim 6.5.1 and Equation (6.23) imply that the coefficient of $p_{\mathfrak{v}}$ in g_T is equal to

$$(-1)^{b}\, n_{(\mathbf{u},\mathbf{v})} \left(\sum_{(\mathbf{u}',\mathbf{v}')\in S_{\mathfrak{v},u_1',0}} n_{(\mathbf{u}',\mathbf{v}')} \right) \prod_{u\in[m]} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u,0}\\ (\mathbf{u}'',\mathbf{v}'')\in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}} (n_{(\mathbf{u}'',\mathbf{v}'')}-1) \prod_{i\in[\Delta-1]} \beta_{\mathfrak{v},u,i} \right). \tag{6.27}$$

Using the facts that $n_{(\mathbf{u},\mathbf{v})} = 1$ and $S_{\mathfrak{v},u_1,0} = \{(\mathbf{u},\mathbf{v})\}$, the above equation can be rewritten as

$$(-1)^{b} \left(\sum_{\substack{(\mathbf{u}',\mathbf{v}')\in S_{\mathfrak{v},u_1',0},\\ (\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u_1,0}}} n_{(\mathbf{u}',\mathbf{v}')} + n_{(\mathbf{u},\mathbf{v})} - 1 \right) \prod_{u\in[m]} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u,0}\\ (\mathbf{u}'',\mathbf{v}'')\in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}} (n_{(\mathbf{u}'',\mathbf{v}'')}-1) \prod_{i\in[\Delta-1]} \beta_{\mathfrak{v},u,i} \right). \tag{6.28}$$

This proves the base case.

Induction step. Let $m \geq 3$ and assume the statement holds true for $m-1$. Let $(\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u_1',0}$ and recall $S_{\mathfrak{v},u_1,0} = \{(\mathbf{u},\mathbf{v})\}$. Let $q_{\mathfrak{v},u_1}$ and $q_{\mathfrak{v},u_1'}$ be the monomials defined in Equation (6.24) and Lemma 6.2 respectively. Then, an argument similar to the one used in Observation 6.12 implies that the factorization of $p_{\mathfrak{v}}$ in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20) given in Equation (6.25) is unique. It is easy to verify from Equation (6.25), Observation 6.13, Lemma

6.2 and Claim 6.5.1 that the coefficient of $p_{\mathfrak{v}}$ in g_T the same as given by Equation (6.27). As $n_{(\mathbf{u},\mathbf{v})} = 1$, this coefficient is equal to the one present in Equation (6.28). This proves the induction step.

Case 2: $\{u_1, u_1'\} \cap \Sigma_{1,1} = \emptyset$. In this case, for every $(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u_1,0}, (\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u_1',0}, n_{(\mathbf{u},\mathbf{v})} \geq 2$ and $n_{(\mathbf{u}',\mathbf{v}')} \geq 2$. Recall Equation (6.20). We want to find the coefficient of $p_{\mathfrak{v}}$ in g_T. As noted above, Equation (6.26) gives the coefficient of $p_{\mathfrak{v}}$ in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20), where $(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u_1,0}, (\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u_1',0}$. Further, Claim 6.5.1 implies that the coefficient of $p_{\mathfrak{v}}$ in the negative part of Equation (6.20) is

$$(-1)^{b-1} \left(\sum_{(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u_1,0}} n_{(\mathbf{u},\mathbf{v})} \right) \left(\sum_{u_1' \in A_{u_1}} \sum_{(\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u_1',0}} n_{(\mathbf{u}',\mathbf{v}')} \right) \times \prod_{u \in [m]} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}') \in S_{\mathfrak{v},u,0} \\ (\mathbf{u}'',\mathbf{v}'') \in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}} (n_{(\mathbf{u}'',\mathbf{v}'')} - 1) \prod_{i \in [\Delta-1]} \beta_{\mathfrak{v},u,i} \right). \tag{6.29}$$

In this case, we mainly figure out the coefficient of $p_{\mathfrak{v}}$ in the positive part of Equation (6.20) and then put the things together. We compute the required coefficient by induction on m.

Base case: $m = 2$. Let $[m] = \{u_1, u_1'\}$. In this case, $A_{u_1,u_1'} = \emptyset$. Note that $D(\overline{R}_{u_1}|\overline{R}_{u_1}) = D(R_{u_1'}|R_{u_1'})$, which implies that the two factors $D(\overline{R}_{u_1}|\overline{R}_{u_1}), D(R_{u_1}|R_{u_1})$ of the positive part of Equation (6.20) are similar and it is easy to see that the positive part of Equation (6.20) looks as $Q_{u_1} \cdot \widetilde{D}(R_{u_1}|R_{u_1}) \cdot Q_{u_1'} \cdot \widetilde{D}(R_{u_1'}|R_{u_1'})$. Let

$$p_1 = \prod_{i \in [0,\Delta-1]} \prod_{(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u_1,i}} \mathbf{x}_{(\mathbf{u},\mathbf{v})}^{a(\mathbf{u},\mathbf{v})-i} \text{ and } p_2 = \prod_{i \in [0,\Delta-1]} \prod_{(\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u_1',i}} \mathbf{x}_{(\mathbf{u}',\mathbf{v}')}^{a(\mathbf{u}',\mathbf{v}')-i}.$$

Then, note that $p_{\mathfrak{v}} = p_1 \cdot p_2$. Observe that $Q_{u_1} \cdot \widetilde{D}(R_{u_1}|R_{u_1})$ and $Q_{u_1'} \cdot \widetilde{D}(R_{u_1'}|R_{u_1'})$ are variable disjoint. This implies that the above factorization of $p_{\mathfrak{v}}$ in $Q_{u_1} \cdot \widetilde{D}(R_{u_1}|R_{u_1}) \cdot Q_{u_1'} \cdot \widetilde{D}(R_{u_1'}|R_{u_1'})$ is unique. We first calculate the coefficient of p_1 in $Q_{u_1} \cdot \widetilde{D}(R_{u_1}|R_{u_1})$ and since $Q_{u_1} \cdot \widetilde{D}(R_{u_1}|R_{u_1})$ and $Q_{u_1'} \cdot \widetilde{D}(R_{u_1'}|R_{u_1'})$ are similar to each other, we also get the coefficient of p_2 in $Q_{u_1'} \cdot \widetilde{D}(R_{u_1'}|R_{u_1'})$. We noted in Remark 1 given after Equation (6.11) that $D(R_{u_1}|R_{u_1}) = \prod_{\hat{v}_1 \in [r_{u_1}]} \det(H'(T_{\hat{v}_1}))$ and each $\det(H'(T_{\hat{v}_1}))$ is a product-depth $(\Delta-1)$ instance of $\det(H'(T))$. Fix $\hat{v}_1 \in [r_{u_1}]$ arbitrarily. By the induction hypothesis of Observation 6.1 on the product-depth of $T_{\hat{v}_1}$, we get that the

denominator of $\det(H'(T_{\hat{v}_1}))$ is equal to

$$d_{T_{\hat{v}_1}} := \prod_{\ell\in[2,\Delta]} \left(\prod_{\substack{\hat{u}_\ell\in[s_{\hat{v}_{\ell-1}}]:s_{\hat{v}_{\ell-1}}\neq 1,\\ \hat{v}_\ell\in[r_{\hat{u}_\ell}]}} Q_{\hat{u}_\ell} \right).$$

Let $\widetilde{\det}(H'(T_{\hat{v}_1})) = d_{T_{\hat{v}_1}} \cdot \det(H'(T_{\hat{v}_1}))$. Then, clearly we get $\widetilde{D}(R_{u_1}|R_{u_1}) = \prod_{\hat{v}_1\in[r_{u_1}]} \widetilde{\det}(H'(T_{\hat{v}_1}))$. Let $v_1 = v_{u_1}$, where v_{u_1} is the coordinate of $\mathfrak{v}$ labelled by u_1, where $\mathfrak{v}$ is used to define the monomial $p_{\mathfrak{v}}$. Then,

$$\widetilde{D}(R_{u_1}|R_{u_1}) = \prod_{\hat{v}_1\in[r_{u_1}]\setminus\{v_1\}} \widetilde{\det}(H'(T_{\hat{v}_1})) \times \widetilde{\det}(H'(T_{v_1})). \tag{6.30}$$

We want to find the coefficient of p_1 in $Q_{u_1}\cdot\widetilde{D}(R_{u_1}|R_{u_1})$. Suppose p_1' is defined as

$$p_1' = \prod_{(\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u_1,0}} \mathbf{x}_{(\mathbf{u},\mathbf{v})}^{a(\mathbf{u},\mathbf{v})-1} \prod_{i\in[\Delta-1]} \left(\prod_{(\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u_1,i}} \mathbf{x}_{(\mathbf{u},\mathbf{v})}^{a(\mathbf{u},\mathbf{v})-i} \right).$$

Since $\widetilde{\det}(H'(T_{\hat{v}_1}))$ is a product-depth $\Delta-1$ instance of g_T for every $\hat{v}_1\in[r_{u_1}]$, it follows from Equation (6.30) and Observation 6.14 that for every $(\hat{\mathbf{u}},\hat{\mathbf{v}})\in S_{\mathfrak{v},u_1,0}$, the degree of $\mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}$ in any monomial of $\widetilde{D}(R_{u_1}|R_{u_1})$ is at most $a(\hat{\mathbf{u}},\hat{\mathbf{v}})-1$. Thus, it is easy to see that $p_1 = \prod_{(\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u_1,0}} \mathbf{x}_{(\mathbf{u},\mathbf{v})}\cdot p_1'$ is the only factorization of p_1 in $Q_{u_1}\cdot\widetilde{D}(R_{u_1}|R_{u_1})$. Thus, the coefficient of p_1 in $Q_{u_1}\cdot\widetilde{D}(R_{u_1}|R_{u_1})$ is equal to the coefficient of p_1' in $\widetilde{D}(R_{u_1}|R_{u_1})$.

It is not difficult to see from the definition of $S_{\mathfrak{v},u_1,0}$ that $\prod_{(\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u_1,0}} \mathbf{x}_{(\mathbf{u},\mathbf{v})}$ is contributed by Q_{u_1} and we would show that p_1' is contributed by $\widetilde{D}(R_{u_1}|R_{u_1})$. Let $\hat{v}_1\in[r_{u_1}]$ and $\hat{u}_2\in[s_{\hat{v}_1}]$ be fixed arbitrarily. Let

$$S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,0} := \left\{ (\mathbf{u}'',\mathbf{v}'') = (u_1,\hat{v}_1,u_2'',v_{u_2''},\ldots,u_\Delta'',v_{u_\Delta''}) \,:\, u_2''=\hat{u}_2, \ell\in[3,\Delta], u_\ell''\in[s_{v_{\ell-1}''}] \right\}, \tag{6.31}$$

where, for every $\ell \in [2,\Delta]$, $v_{u''_\ell}$ is fixed by $\mathfrak{v}$. For $i \in [\Delta-2]$,

$$S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i} := \Big\{ (\mathbf{u}'',\mathbf{v}'') = (u_1,\hat{v}_1,u''_2,\ldots,v''_{i+1},u''_{i+2},v_{u''_{i+2}},\ldots,u''_\Delta,v_{u''_\Delta}) : u''_2 = \hat{u}_2, v''_2 \in [r_{u''_2}] \\ \forall j \in [3,i], u''_j \in [s_{v''_{j-1}}], v''_j \in [r_{u''_j}], u''_{i+1} \in [s_{v''_i}], v''_{i+1} \in [r_{u''_{i+1}}] \setminus \{v_{u''_{i+1}}\}, \\ u''_{i+2} \in [s_{v''_{i+1}}], \forall k \in [i+3,\Delta], u''_k \in [s_{v_{u''_{k-1}}}] \Big\}, \tag{6.32}$$

where $v_{u''_{i+1}},\ldots,v_{u''_{\Delta-1}}$ are fixed by $\mathfrak{v}$. As $v_1 = v_{u_1}$, it is easy to note that

$$S_{\mathfrak{v},u_1,0} = \bigcup_{u_2 \in [s_{v_1}]} S_{\mathfrak{v},u_1,v_1,u_2,0} \text{ and } S_{\mathfrak{v},u_1,1} = \bigcup_{\hat{v}_1 \in [r_{u_1}]\setminus\{v_1\}} \bigcup_{\hat{u}_2 \in [s_{\hat{v}_1}]} S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,0}, \tag{6.33}$$

Further, for every $i \in [2,\Delta-1]$,

$$S_{\mathfrak{v},u_1,i} = \bigcup_{\hat{v}_1 \in [r_{u_1}]} \bigcup_{\hat{u}_2 \in [s_{\hat{v}_1}]} S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i-1}. \tag{6.34}$$

For $\hat{v}_1 \in [r_{u_1}]$, let

$$p_{1,\hat{v}_1} := \prod_{i\in[0,\Delta-2]} \prod_{\hat{u}_2 \in [s_{\hat{v}_1}]} \prod_{(\mathbf{u}'',\mathbf{v}'') \in S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i}} \mathbf{X}_{(\mathbf{u}'',\mathbf{v}'')}^{(a(\mathbf{u}'',\mathbf{v}'')-1)-i}.$$

This along with Equations (6.33) and (6.34) implies that the monomial p'_1 defined above can be written as $p'_1 = \prod_{\hat{v}_1 \in [r_{u_1}]} p_{1,\hat{v}_1}$. To figure out the coefficient of $p_{1,\hat{v}_1}$, we first show that it is a 'smaller instance' of $p_{\mathfrak{v}}$. Recall the definition of $\Sigma_\ell^{\{u_1\}}$ from Point 3 of the second part of the notations. Let $\Sigma_\ell^{\{(u_1,\hat{v}_1)\}} \subseteq \Sigma_\ell^{\{u_1\}}$ such that every $u_\ell \in \Sigma_\ell^{\{u_1\}}$ lies in the sub-ROF, whose parent is $\hat{v}_1$ and grandparent is u_1. Let $\mathfrak{v}^{\{(u_1,\hat{v}_1)\}} := (\mathfrak{v}_\ell^{\{(u_1,\hat{v}_1)\}})_{\ell\in[2,\Delta]}$, where $\mathfrak{v}_\ell^{\{(u_1,\hat{v}_1)\}} = (v_{u_\ell})_{u_\ell \in \Sigma_\ell^{\{(u_1,\hat{v}_1)\}}}$ and for every $u_\ell \in \Sigma_{\ell,1}^{\{(u_1,\hat{v}_1)\}}, v_{u_\ell} \in [r_{u_\ell}]$ satisfies $n_{v_{u_\ell}} = 1$, where such v_{u_ℓ} is unique as T is canonical and for $u_\ell \in \Sigma_\ell^{\{(u_1,\hat{v}_1)\}} \setminus \Sigma_{\ell,1}^{\{(u_1,\hat{v}_1)\}}, v_{u_\ell} \in [r_{u_\ell}]$ is arbitrary. Then, note that $p_{\mathfrak{v}^{\{(u_1,\hat{v}_1)\}}}$ in $\widetilde{\det}(H'(T_{\hat{v}_1}))$ is similar to $p_{\mathfrak{v}}$ in g_T. Since the product-depth of $T_{\hat{v}_1}$ is one less than the product-depth of T, the value of $a(\mathbf{u}'',\mathbf{v}'')$ also reduces by 1 for every $(\mathbf{u}'',\mathbf{v}'') \in S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i}$. Thus,

$$p_{\mathfrak{v}^{\{(u_1,\hat{v}_1)\}}} = \prod_{i\in[0,\Delta-2]} \prod_{\hat{u}_2 \in [s_{\hat{v}_1}]} \prod_{(\mathbf{u}'',\mathbf{v}'') \in S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i}} \mathbf{X}_{(\mathbf{u}'',\mathbf{v}'')}^{(a(\mathbf{u}'',\mathbf{v}'')-1)-i}.$$

Hence, $p_{\mathfrak{v}^{\{(u_1,\hat{v}_1)\}}} = p_{1,\hat{v}_1}$. Since, $\widetilde{\det}(H'(T_{\hat{v}_1}))$ is a product-depth $(\Delta-1)$ instance of g_T, it follows from the induction hypothesis of Lemma 6.1 on the product-depth $(\Delta-1)$, that the coefficient of $p_{1,\hat{v}_1}$ in $\widetilde{\det}(H'(T_{\hat{v}_1}))$ is equal to $(-1)^{c_{\hat{v}_1}}\beta_{\mathfrak{v},u_1,\hat{v}_1,0}\prod_{\hat{u}_2\in[s_{\hat{v}_1}]}\prod_{i\in[2,\Delta-1]}\beta_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i}$, where $c_{\hat{v}_1} = \sum_{\ell\in[2,\Delta]}\Big(s(\prod_\ell^{\{(u_1,\hat{v}_1)\}}) - r(\Sigma_\ell^{\{(u_1,\hat{v}_1)\}}) + s_{\hat{v}_1} - 1\Big)$,

$$\beta_{\mathfrak{v},u_1,\hat{v}_1,0} = \left(\sum_{\substack{\hat{u}_2\in[s_{\hat{v}_1}],\\(\mathbf{u}'',\mathbf{v}'')\in S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,0}}} n_{(\mathbf{u}'',\mathbf{v}'')} - 1\right)\left(\prod_{\substack{\hat{u}_2\in[s_{\hat{v}_1}],\\(\mathbf{u}'',\mathbf{v}'')\in S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,0},\\(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in W_{(\mathbf{u}'',\mathbf{v}'')}}} (n_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')} - 1)\right) \tag{6.35}$$

and for $\hat{u}_2 \in [s_{\hat{v}_1}]$ and $i \in [2,\Delta-1]$,

$$\begin{aligned}\beta_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i} = &\prod_{(\hat{\mathbf{u}},\hat{\mathbf{v}})_i\in B_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i}}\left(\sum_{\substack{(\mathbf{u}'',\mathbf{v}'')\in S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i-1}\\(\mathbf{u}'',\mathbf{v}'')_i=(\hat{\mathbf{u}},\hat{\mathbf{v}})_i}} n_{(\mathbf{u}'',\mathbf{v}'')} - 1\right)\\ &\times\prod_{(\hat{\mathbf{u}},\hat{\mathbf{v}})_i\in B_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i}}\left(\prod_{\substack{(\mathbf{u}'',\mathbf{v}'')\in S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i-1}\\(\mathbf{u}'',\mathbf{v}'')_i=(\hat{\mathbf{u}},\hat{\mathbf{v}})_i,}}\prod_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in W_{(\mathbf{u}'',\mathbf{v}'')}}(n_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')} - 1)\right),\end{aligned} \tag{6.36}$$

where $B_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i} = \{(\hat{\mathbf{u}},\hat{\mathbf{v}})_i = (\hat{u}_1,\ldots,\hat{v}_i) : (\hat{\mathbf{u}},\hat{\mathbf{v}}) \in S_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i-1}\}$. It follows from Equations (6.33), (6.34), (6.35) and (6.36) that

$$\prod_{\hat{v}_1\in[r_{u_1}]\setminus\{v_1\}}\beta_{\mathfrak{v},u_1,\hat{v}_1,0} = \beta_{\mathfrak{v},u_1,1} \tag{6.37}$$

and for $i \in [2,\Delta-1]$,

$$\prod_{\hat{v}_1\in[r_{u_1}]}\prod_{\hat{u}_2\in[s_{\hat{v}_1}]}\beta_{\mathfrak{v},u_1,\hat{v}_1,\hat{u}_2,i} = \beta_{\mathfrak{v},u_1,i}. \tag{6.38}$$

This along with Equation (6.33) implies that, the coefficient of p'_1 in $\prod_{\hat{v}_1\in[r_{u_1}]}\widetilde{\det}(H'(T_{\hat{v}_1}))$ and

hence the coefficient of p_1 in $Q_{u_1} \cdot D(R_{u_1}|R_{u_1})$ is equal to

$$(-1)^{c_2} \left(\sum_{(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u_1,0}} n_{(\mathbf{u},\mathbf{v})} - 1 \right) \prod_{\substack{(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u_1,0}, \\ (\hat{\mathbf{u}},\hat{\mathbf{v}}) \in W_{(\mathbf{u},\mathbf{v})}}} (n_{(\hat{\mathbf{u}},\hat{\mathbf{v}})} - 1) \prod_{i \in [\Delta-1]} \beta_{\mathfrak{v},u_1,i},$$

where $c_2 = \sum_{\ell \in [\Delta]} \left(s(\prod_\ell^{\{u_1\}}) - r(\sum_\ell^{\{u_1\}}) \right)$. Similarly, the coefficient of p_2 in $Q_{u_1'} \cdot \widetilde{D}(R_{u_1'}|R_{u_1'})$ is equal to

$$(-1)^{c_1} \left(\sum_{(\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u_1',0}} n_{(\mathbf{u}',\mathbf{v}')} - 1 \right) \prod_{\substack{(\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u_1',0} \\ (\hat{\mathbf{u}}',\hat{\mathbf{v}}') \in W_{(\mathbf{u}',\mathbf{v}')}}} (n_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')} - 1) \prod_{i \in [\Delta-1]} \beta_{\mathfrak{v},u_1',i},$$

where $c_1 = \sum_{\ell \in [\Delta]} \left(s(\prod_\ell^{\{u_1'\}}) - r(\sum_\ell^{\{u_1'\}}) \right)$. As $p_{\mathfrak{v}} = p_1 \cdot p_2$, these two equations imply that the coefficient of p_1 in the positive part of Equation (6.20) is equal to

$$(-1)^{c_1+c_2} \prod_{u \in [m]} \left(\left(\sum_{(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u,0}} n_{(\mathbf{u},\mathbf{v})} - 1 \right) \prod_{\substack{(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u,0} \\ (\hat{\mathbf{u}},\hat{\mathbf{v}}) \in W_{(\mathbf{u},\mathbf{v})}}} (n_{(\hat{\mathbf{u}},\hat{\mathbf{v}})} - 1) \prod_{i \in [\Delta-1]} \beta_{\mathfrak{v},u,i} \right)$$

Note that $c_1 + c_2 = b - 1$, where $b = \sum_{\ell \in [\Delta]} (s(\prod_\ell) - r(\sum_\ell)) + 1$. Then, Equation (6.20) implies that on subtracting Equation (6.29) from the above equation, we get that the coefficient of $p_{\mathfrak{v}}$ in g_T is equal to $(-1)^b \beta_{\mathfrak{v},0} \cdot \prod_{i \in [\Delta-1], u \in [m]} \beta_{\mathfrak{v},u,i}$, where $b = \sum_{\ell \in [\Delta]} (s(\prod_\ell) - r(\sum_\ell)) + 1$. This proves the base case.

Inductive step. Let $m \geq 3$ and assume that the induction hypothesis holds for $m-1$. Then, we want to find the coefficient of $p_{\mathfrak{v}}$ in the positive part of Equation (6.20). Let $p_1 = \prod_{i \in [0,\Delta-1]} \prod_{(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u_1,i}} \mathbf{x}_{(\mathbf{u},\mathbf{v})}^{a(\mathbf{u},\mathbf{v})-i}$ and $p_2 = \prod_{u_1' \in A_{u_1}} \prod_{i \in [\Delta-1]} \prod_{(\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u_1',i}} \mathbf{x}_{(\mathbf{u}',\mathbf{v}')}^{a(\mathbf{u}',\mathbf{v}')-i}$. Then, note that $p_{\mathfrak{v}} = p_1 \cdot p_2$. Observe that $Q_{u_1} \cdot \widetilde{D}(R_{u_1}|R_{u_1})$ and $\widetilde{D}(\overline{R}_{u_1}|\overline{R}_{u_1})$ are variable disjoint. This implies that the above factorization of $p_{\mathfrak{v}}$ in $Q_{u_1} \cdot \widetilde{D}(R_{u_1}|R_{u_1}) \cdot \widetilde{D}(\overline{R}_{u_1}|\overline{R}_{u_1})$ is unique. We have already computed the coefficient of p_1 in $Q_{u_1} \cdot \widetilde{D}(R_{u_1}|R_{u_1})$ in the base case. Note that p_2 can be obtained by replacing $[m]$ with A_{u_1} in the definition of $p_{\mathfrak{v}}$. It follows from Remark 1 given after Equation (6.11), $\widetilde{D}(\overline{R}_{u_1}|\overline{R}_{u_1})$ is a product-depth Δ and top fan-in $(m-1)$ instance of $\det(H'(T))$, thus

the induction hypothesis of Lemma 6.1 implies that the coefficient of p_2 in $\widetilde{D}(\overline{R}_{u_1}|\overline{R}_{u_1})$ is equal to

$$(-1)^{c_3}\left(\sum_{\substack{u'_1\in A_{u_1}\\(\mathbf{u}',\mathbf{v}')\in S_{\mathfrak{v},u'_1,0}}} n_{(\mathbf{u}',\mathbf{v}')}-1\right)\prod_{\substack{u'_1\in A_{u_1}\\(\mathbf{u}',\mathbf{v}')\in S_{\mathfrak{v},u'_1,0}}}\prod_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in W_{(\mathbf{u}',\mathbf{v}')}}(n_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}-1)\prod_{\substack{i\in[\Delta-1],\\u'_1\in A_{u_1}}}\beta_{\mathfrak{v},u'_1,i},$$

where $c_3=\sum_{\ell\in[\Delta]}\left(s(\prod_\ell^{A_{u_1}})-r(\sum_\ell^{A_{u_1}})\right)+|A_{u_1}|-1$. On multiplying the coefficients of p_1 and p_2 together, we get that the coefficient of $p_{\mathfrak{v}}$ in the positive part of Equation (6.20) is equal to

$$(-1)^{c_2+c_3}\left(\sum_{(\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u_1,0}} n_{(\mathbf{u},\mathbf{v})}-1\right)\left(\sum_{\substack{u'_1\in A_{u_1}\\(\mathbf{u}',\mathbf{v}')\in S_{\mathfrak{v},u'_1,0}}} n_{(\mathbf{u}',\mathbf{v}')}-1\right)\times$$
$$\prod_{\substack{u\in[m]\\(\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u,0}}}\prod_{(\hat{\mathbf{u}},\hat{\mathbf{v}})\in W_{(\mathbf{u},\mathbf{v})}}(n_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}-1)\times\prod_{\substack{i\in[\Delta-1],\\u\in[m]}}\beta_{\mathfrak{v},u,i}.$$

As $c_2+c_3=b-1$, it follows from Equation (6.20) that on subtracting the coefficient given in Equation (6.29) from the above equation, we get the coefficient of $p_{\mathfrak{v}}$ in g_T, is equal to $(-1)^b\beta_{\mathfrak{v}}$. This completes the inductive step.

It is clear that g_{T_1} and g_{T_2} are monomial disjoint. This completes the proof of Lemma 6.1. ■

6.5.4 Proof of Claim 6.5.1

Let $p_{\mathfrak{v}}$ be present in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20) where $(\mathbf{u},\mathbf{v})\in S_{u_1}$, $(\mathbf{u}',\mathbf{v}')\in S_{u'_1}$ and $u'_1\in A_{u_1}$. Then, Equations (6.21) and (6.22) of Observation 6.10 imply that every monomial of this term contains $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ and $\mathbf{x}_{(\mathbf{u}',\mathbf{v}')}$ and hence $\mathbf{x}_{(\mathbf{u},\mathbf{v})},\mathbf{x}_{(\mathbf{u}',\mathbf{v}')}$ are present in $p_{\mathfrak{v}}$. This implies that there exist $i,j\in[0,\Delta-1]$, such that $(\mathbf{u}',\mathbf{v}')\in S_{\mathfrak{v},u'_1,i}$ and $(\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u_1,j}$. We want to show that $i=j=0$.

Suppose $i\neq 0$. Then, $i\in[\Delta-1]$ and $(\mathbf{u}',\mathbf{v}')\in S_{\mathfrak{v},u'_1,i}$. Let $k\in[0,i-1]$ be such that $v'_1=v_{u'_1},\ldots,v'_k=v_{u'_k}$ but $v'_{k+1}\neq v_{u'_{k+1}}$, where $v'_1,\ldots,v'_{k+1}$ correspond to $(\mathbf{u}',\mathbf{v}')$ and $v_{u'_1},\ldots v_{u'_{k+1}}$ are fixed by $\mathfrak{v}$. Let $(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u'_1,0}$ be such that for every $\ell\in[k]$, $\hat{u}'_\ell=u'_\ell,\hat{v}'_\ell=v'_\ell$ and $\hat{u}'_{k+1}=u'_{k+1}$. Then, $\hat{v}'_{k+1}=v_{u'_{k+1}}$. It follows from the definition of $p_{\mathfrak{v}}$ that $\deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}p_{\mathfrak{v}}$ should be equal to $a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')$. As $p_{\mathfrak{v}}$ is in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20), it is easy

to see from Equations (6.20), (6.21) and Point 3 of Lemma 6.2 that in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20), $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ is contributed only by

$$h := \prod_{\ell\in[2,\Delta]\hat{u}'_\ell\in[s_{v'_{\ell-1}}]:\, s_{v'_{\ell-1}}\neq 1} Q_{\hat{u}'_\ell}\cdot \widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$$

in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term. This is so because it is easy to see that in Equation (6.21), $\widetilde{\det}(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1}})$ does not contain a monomial divisible by $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$. Let p and q be arbitrary monomials of the polynomials $\prod_{\ell\in[2,\Delta]\hat{u}'_\ell\in[s_{v'_{\ell-1}}]:\, s_{v'_{\ell-1}}\neq 1} Q_{\hat{u}'_\ell}$ and $\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ respectively. Then, observe that $\deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}} p \leq k$, as $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ can only be contributed by $Q_{u'_\ell}, \ell\in[2,k+1]$. It follows from Lemma 6.2 that $\deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}} q \leq (a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-(k+1))$. This implies the degree of $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ in any monomial of h and hence in any monomial of the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20) is at most $(a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-1)$. Thus, $p_{\mathfrak{v}}$ is not present in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20), which is a contradiction. Hence, $i=0$. Similarly, we can show that $j=0$.

It follows from the factorization of $p_{\mathfrak{v}}$ given in Equation (6.25) and Lemma 6.2 that if $(\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u_1,0}, (\mathbf{u}',\mathbf{v}')\in S_{\mathfrak{v},u'_1,0}$ then $p_{\mathfrak{v}}$ is in the $((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))$-th term of Equation (6.20). This proves the converse and completes the proof. ■

6.5.5 Proof of Lemma 6.2

Recall that the objective of this lemma is to understand $D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$. As noted in Observation 6.10,

$$D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \mathbf{x}_{(\mathbf{u},\mathbf{v})}\prod_{\ell\in[2,\Delta]}\prod_{\hat{u}_\ell\in[s_{v_{\ell-1}}]\setminus\{u_\ell\}} Q_{\hat{u}_\ell}\times\frac{\det(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}})}{Q_{u_1}}, \tag{6.39}$$

where $u_\ell, v_\ell, \ell\in[2,\Delta]$ correspond to $(\mathbf{u},\mathbf{v})$ and $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}}$ is defined in Observation 6.10. Let $D_1 = \det(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}})$. Consider the following claim.

Claim 6.5.2 *Let $u_1,u'_1\in[m], u_1\neq u'_1$ and $\bar{d}_{((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))}$ be as given in Lemma 6.2. Then, the denominator of D_1 is $\bar{d}'_{((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))} := \frac{\bar{d}_{((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))}}{Q_{u_1}}$. Let $\mathfrak{v}\in\mathscr{V}, (\mathbf{u}',\mathbf{v}')\in S_{\mathfrak{v},u'_1,0}, (\mathbf{u},\mathbf{v})\in S_{\mathfrak{v},u_1,0}$*

and

$$q'_{\mathfrak{v},u'_1} := \prod_{j\in[a(\mathbf{u}',\mathbf{v}')-1]} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u'_1,0},\\ (\hat{\mathbf{u}}',\hat{\mathbf{v}}')_j=(\mathbf{u}',\mathbf{v}')_j,\\ \hat{u}'_{j+1}\neq u'_{j+1}}} \mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-j} \right) \times \prod_{u''_1\in A_{u_1,u'_1}} \prod_{(\mathbf{u}'',\mathbf{v}'')\in S_{\mathfrak{v},u''_1,0}} \mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}^{a(\mathbf{u}'',\mathbf{v}'')}$$
$$\times \prod_{i\in[\Delta-1],\widehat{u}'_1\in A_{u_1}} \prod_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},\widehat{u}'_1,i}} \mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-i}.$$

Then, the coefficient of $q'_{\mathfrak{v},u'_1}$ *in* $\widetilde{D}_1 := D_1 \times \bar{d}'_{((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))}$ *is equal to*

$$(-1)^c\, n_{(\mathbf{u}',\mathbf{v}')} \prod_{\widehat{u}'_1\in A_{u_1}} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},\widehat{u}'_1,0}\\ (\mathbf{u}'',\mathbf{v}'')\in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}} (n_{(\mathbf{u}'',\mathbf{v}'')}-1) \prod_{i\in[\Delta-1]} \beta_{\mathfrak{v},\widehat{u}'_1,i} \right),$$

where $c = \sum_{\ell\in[\Delta]} \left(s(\prod_\ell^{A_{u_1}}) - r(\Sigma_\ell^{A_{u_1}}) \right) + |A_{u_1}| - 1$ *and* $\beta_{\mathfrak{v},\widehat{u}'_1,i}$ *is defined in Lemma 6.1. Let* $(\mathbf{u}',\mathbf{v}') \in S_{u'_1}$ *be picked arbitrarily and* q_1 *be an arbitrary monomial of* $\widetilde{D}_1$. *Let* $(\hat{\mathbf{u}}',\hat{\mathbf{v}}') \in S_{u'_1}$ *be such that there exists* $i \in [\Delta]$, *such that either* $(\mathbf{u}',\mathbf{v}')_{i-1} = (\hat{\mathbf{u}}',\hat{\mathbf{v}}')_{i-1}$, $u'_i = \hat{u}'_i$ *and* $v'_i \neq \hat{v}'_i$ *or* $(\mathbf{u}',\mathbf{v}')_i = (\hat{\mathbf{u}}',\hat{\mathbf{v}}')_i$ *and* $u'_{i+1} \neq \hat{u}'_{i+1}$. *Then,* $\deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}} q_1 \leq (a(\hat{\mathbf{u}}',\hat{\mathbf{v}}') - i)$. *Let* $u''_1 \in A_{u_1,u'_1}, (\mathbf{u}'',\mathbf{v}'') \in S_{u''_1}$. *Then,* $\deg_{(\mathbf{u}'',\mathbf{v}'')} q \leq a(\mathbf{u}'',\mathbf{v}'')$.

We first complete the proof of Lemma 6.2 assuming Claim 6.5.2, whose proof is given in Section 6.5.6. Equation (6.39) and Claim 6.5.2 imply that the denominator of $D(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ is equal to $\bar{d}_{((\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}'))}$. On clearing the denominator of Equation (6.39), we get

$$\widetilde{D}(\overline{R}_{u_1}|\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}) = \mathbf{x}_{(\mathbf{u},\mathbf{v})} \prod_{\ell\in[2,\Delta]} \prod_{\hat{u}_\ell\in[s_{v_{\ell-1}}]\setminus\{u_\ell\}} Q_{\hat{u}_\ell} \times \widetilde{\det}(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}}). \tag{6.40}$$

Recall $(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u_1,0}$. Note that $q_{\mathfrak{v},u'_1} = q'_{\mathfrak{v},u'_1} \times \prod_{(\hat{\mathbf{u}},\hat{\mathbf{v}})\in S_{\mathfrak{v},u_1,0}} \mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}$. Observe that the monomial $\prod_{(\hat{\mathbf{u}},\hat{\mathbf{v}})\in S_{\mathfrak{v},u_1,0}} \mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}$ is present in $\mathbf{x}_{(\mathbf{u},\mathbf{v})} \prod_{\ell\in[2,\Delta]} \prod_{\hat{u}_\ell\in[s_{v_{\ell-1}}]\setminus\{u_\ell\}} Q_{\hat{u}_\ell}$. Since $\prod_{\ell\in[2,\Delta]} \prod_{\hat{u}_\ell\in[s_{v_{\ell-1}}]\setminus\{u_\ell\}} Q_{\hat{u}_\ell}$ and $\widetilde{\det}(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}})$ are variable disjoint, the factorization of $q_{\mathfrak{v},u'_1}$ given in the beginning of this paragraph is unique. Thus, the coefficient of $q'_{\mathfrak{v},u'_1}$ in $\widetilde{D}_1$ in Claim 6.5.2 is equal to the coefficient of $q_{\mathfrak{v},u'_1}$ in $\widetilde{D}(\overline{R}_{u_1}|\overline{C}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')})$ in Lemma 6.2. This completes the proof of Lemma 6.2. ∎

6.5.6 Proof of Claim 6.5.2

We prove this claim by induction on the product-depth Δ. Let $\Delta = 1$ and $D_1 = \det(H_{(u',v'),\mathbf{j},\overline{R}_{u_1}})$, where $H_{(u',v'),\mathbf{j},\overline{R}_{u_1}}$ is the matrix defined in Observation 6.7. It follows from Observation 6.7 and Claim 6.4.2 that the denominator of D_1 is equal to $\prod_{u_1'' \in A_{u_1,u_1'}} Q_{u_1''}$. Further, the monomial $q'_{\mathfrak{v},u_1'}$ in this case looks as $\prod_{u_1'' \in A_{u_1,u_1'}} \mathbf{x}_{u_1'',v_{u_1''}}$, where for every $u_1'' \in A_{u_1,u_1'}, v_{u_1''} \in [r_{u_1''}]$ is picked in the following way: if there exists $v_1'' \in [r_{u_1''}]$, such that $n_{(u_1'',v_1'')} = 1$ then $v_{u_1''} = v_1''$ otherwise $v_{u_1''}$ is picked arbitrarily. Then, the coefficient of this monomial in $\widetilde{D}_1 := \prod_{u_1'' \in A_{u_1,u_1'}} Q_{u_1''} \cdot D_1$ is given by Claim 6.4.2, which is equal to

$$\prod_{u \in A_{u_1}} (-1)^{(n_u - r_u)+|A_{u_1}|-1} \prod_{\substack{\hat{u}_1'' \in A_{u_1}, \\ \hat{v}_1 \in [r_{\hat{u}_1''}] \setminus \{\hat{v}_1''\}}} (n_{\hat{u}_1'',\hat{v}_1} - 1) n_{(u_1',v_1')}.$$

It is not difficult to see that this coefficient is equal to the coefficient of $q'_{\mathfrak{v},u_1'}$ given in the statement of the instance of this claim for $\Delta = 1$. Further, if $u_1'' \in A_{u_1,u_1'}, (\mathbf{u}'', \mathbf{v}'') \in S_{u_1''}$ then Claim 6.4.2 implies that the degree of $\mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}$ in any monomial of $\widetilde{D}_1$ is at most 1. This proves the base case.

Now, suppose this claim holds for product-depth $\Delta - 1$. We understand D_1 by looking at its Laplace's expansion. Recall the definition of $R_{u_1'}$ from notations. Then, $|R_{u_1'}| = n_{u_1'}$, where $n_{u_1'}$ is the number of variables appearing in the $+$-rooted sub-ROF $Q_{u_1'}$ and we assume $R_{u_1'}$ is ordered by $\prec$ defined in Section 6.3. Let $\overline{R}_{u_1,u_1'} := \overline{R}_{u_1} \setminus R_{u_1'}$ also be ordered by $\prec$. Recall from Observation 6.10 that the matrix $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}}$ is obtained from the sub-matrix of $H'(T)$ whose rows and columns are labelled with $\overline{R}_{u_1}$ by replacing the $(\mathbf{u}, \mathbf{v}, 1)$-th column with the vector $\mathbf{j}$ confined to $\overline{R}_{u_1}$ and not with the $(\mathbf{u}', \mathbf{v}', 1)$-th column of $H'(T)$ confined to $\overline{R}_{u_1}$. Then, Theorem 6.1 implies that

$$D_1 = \sum_{C \subseteq \overline{R}_{u_1}, |C| = n_{u_1'}} sgn(R_{u_1'}) \cdot sgn(C) \cdot D(R_{u_1'}|C) \cdot D(\overline{R}_{u_1,u_1'}|\overline{C}).$$

Let E be the set of tuples labelling the columns of $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}}$, where the first coordinates of these tuples are u_1' and E be ordered by $\prec$. We want to mention that we have used here E instead of $R_{u_1'}$ (although both sets are same) to emphasise on the fact that the column of $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}}$ labelled by $(\mathbf{u}, \mathbf{v}, 1)$ is $\mathbf{j}$ confined to $\overline{R}_{u_1}$. Let $\overline{E} = \overline{R}_{u_1} \setminus E$. Observe that $\overline{E} = \overline{R}_{u_1,u_1'}$.

Let $(\mathbf{u}'', \mathbf{v}'') \in S_{u_1''}$, where $u_1'' \in A_{u_1,u_1'}$ and $E_{(\mathbf{u}'',\mathbf{v}'')}$ be obtained from E by replacing $(\mathbf{u}, \mathbf{v}, 1)$

with $(\mathbf{u}'', \mathbf{v}'', 1)$, such that $E_{(\mathbf{u}'',\mathbf{v}'')}$ as an ordered set it is same as E with the change that the position of $(\mathbf{u}'', \mathbf{v}'', 1)$ in $E_{(\mathbf{u}'',\mathbf{v}'')}$ and the position of $(\mathbf{u}, \mathbf{v}, 1)$ in E are same. Let $\overline{E}_{(\mathbf{u}'',\mathbf{v}'')}$ be obtained from the ordered set $\overline{R}_{u_1,u_1'}$ by replacing the tuple $(\mathbf{u}'', \mathbf{v}'', 1)$ in $\overline{R}_{u_1,u_1'}$ with $(\mathbf{u}, \mathbf{v}, 1)$. Then, note that $E_{(\mathbf{u}'',\mathbf{v}'')}$ and $\overline{E}_{(\mathbf{u}'',\mathbf{v}'')}$ are similar to the sets $C'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ and $\overline{C}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$ mentioned in Section 6.3 respectively and neither $E_{(\mathbf{u}'',\mathbf{v}'')}$ nor $\overline{E}_{(\mathbf{u}'',\mathbf{v}'')}$ is ordered by $\prec$. Then, it is easy to show that the arguments similar to those used in Section 6.3 to converge to Equation (6.11) imply the following.

$$D_1 = D(R_{u_1'}|E) \cdot D(\overline{R}_{u_1,u_1'}|\overline{E}) - \left(\sum_{\substack{u_1'' \in A_{u_1,u_1'}, \\ (\mathbf{u}'',\mathbf{v}'') \in S_{u_1''}}} D(R_{u_1'}|E_{(\mathbf{u}'',\mathbf{v}'')}) \cdot D(\overline{R}_{u_1,u_1'}|\overline{E}_{(\mathbf{u}'',\mathbf{v}'')}) \right), \quad (6.41)$$

where

1. $D(R_{u_1'}|E_{(\mathbf{u}'',\mathbf{v}'')})$ is the determinant of the matrix obtained from the sub-matrix of $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}}$, whose rows and columns are labelled by $R_{u_1'}$ and E respectively, by replacing the column labelled by $(\mathbf{u}, \mathbf{v}, 1)$ with its $(\mathbf{u}'', \mathbf{v}'', 1)$-th column restricted to $R_{u_1'}$.

2. $D(\overline{R}_{u_1,u_1'}|\overline{E}_{(\mathbf{u}'',\mathbf{v}'')})$ is the determinant of the matrix obtained from the sub-matrix of $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},\overline{R}_{u_1}}$, whose rows and columns are labelled by $\overline{R}_{u_1,u_1'}$ and $\overline{E}$ respectively, by replacing the column labelled by $(\mathbf{u}'', \mathbf{v}'', 1)$ with the vector $\mathbf{j}$ confined to $\overline{R}_{u_1,u_1'}$.

It is easy to note from the structure of $H'(T)$ given in Section 6.2 that the following equation holds.

$$D(R_{u_1'}|E_{(\mathbf{u}'',\mathbf{v}'')}) = \mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')} \cdot \prod_{\ell \in [2,\Delta]} \prod_{\hat{u}_\ell'' \in [s_{v''_{\ell-1}}] \setminus \{u_\ell''\}} Q_{\hat{u}_\ell''} \times \frac{\det(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u_1'}})}{Q_{u_1''}},$$

where $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u_1'}}$ is obtained from the sub-matrix of $H'(T)$, whose rows and columns are labelled by $R_{u_1'}$ by replacing the $(\mathbf{u}, \mathbf{v}, 1)$-th column with the vector $\mathbf{j}$ confined to $R_{u_1'}$. Then, observe that $\det(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u_1'}}) = D(R_{u_1'}|E)$. Thus,

$$D(R_{u_1'}|E_{(\mathbf{u}'',\mathbf{v}'')}) = \mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')} \cdot \prod_{\ell \in [2,\Delta]} \prod_{\hat{u}_\ell'' \in [s_{v''_{\ell-1}}] \setminus \{u_\ell''\}} Q_{\hat{u}_\ell''} \times \frac{D(R_{u_1'}|E)}{Q_{u_1''}}.$$

Then, Equation (6.41) can be re-written as

$$D_1 = D(R_{u'_1}|E)\left(D(\overline{R}_{u_1,u'_1}|\overline{E}) - \left(\sum_{\substack{u''_1\in A_{u_1,u'_1},\\(\mathbf{u}'',\mathbf{v}'')\in S_{u''_1}}} \mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')} \prod_{\substack{\ell\in[2,\Delta],\\ \hat{u}''_\ell\in[s_{v''_{\ell-1}}]\setminus\{u''_\ell\}}} Q_{\hat{u}''_\ell} \cdot \frac{D(\overline{R}_{u_1,u'_1}|\overline{E}_{(\mathbf{u}'',\mathbf{v}'')})}{Q_{u''_1}}\right)\right). \tag{6.42}$$

Now, we prove Claim 6.5.2 in three parts.

Part 1. The denominator of D_1. We first calculate the denominators of the minors in the R.H.S of Equation (6.42) and then put the things together.

1. Denominator of $D(R_{u'_1}|E)$: As noted above, $\det(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u'_1}}) = D(R_{u'_1}|E)$. Then, it is easy to show that

$$D(R_{u'_1}|E) = \prod_{\hat{v}'_1\in[r_{u'_1}]\setminus\{v'_1\}} \det(H'(T_{\hat{v}'_1})) \times \det(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u'_1},v'_1}), \tag{6.43}$$

   where $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u'_1},v'_1}$ is the sub-matrix of $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u'_1}}$, whose rows and columns are labelled by the set $\{(\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{k}) \in R_{u'_1} : \hat{v}_1 = v'_1\}$. Let $D_2 = \det(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u'_1},v'_1})$. Note that D_2 is a product-depth $(\Delta-1)$ instance of the determinant studied in Claim 6.5.2. Thus, on replacing $[2,\Delta]$ by $[3,\Delta]$, A_{u_1,u'_1} by $[s_{v'_1}]\setminus\{u'_2\}$ and A_{u_1} by $[s_{v'_1}]$ in the definition of $\bar{d}'_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')}$, we get the denominator of D_2. So, from the induction hypothesis of Claim 6.5.2 for product-depth $(\Delta-1)$ along with these changes, we get that the denominator of D_2 is equal to

$$\frac{\prod_{\widehat{u}_2\in[s_{v'_1}]\setminus\{u'_2\}} Q_{\widehat{u}_2} \prod_{\ell\in[3,\Delta]} \prod_{\widehat{u}'_\ell\in\widehat{\Sigma}_\ell^{[s_{v'_1}]}} Q_{\widehat{u}'_\ell}}{\prod_{\ell\in[3,\Delta]:s_{v'_{\ell-1}}\neq 1} Q_{u'_\ell}},$$

   where $u'_\ell, \ell \in [3,\Delta]$ correspond to $(\mathbf{u}',\mathbf{v}')$. Observe that for every $\hat{v}'_1 \in [r_{u'_1}]\setminus\{v'_1\}$, $\det(H'(T_{\hat{v}'_1}))$ is a product-depth $(\Delta-1)$ instance of the determinant studied in Lemma 6.1 and can be seen by making the following changes in Lemma 6.1: $[\Delta]$ is replaced by $[2,\Delta]$, u_ℓ, v_ℓ are replaced with $\hat{u}'_\ell$ and $\hat{v}_{\ell'}$ respectively. Then, by Observation 6.1 we get

that the denominator of $\prod_{\hat{v}'_1 \in [r_{u'_1}] \setminus \{v'_1\}} \det(H'(T_{\hat{v}'_1}))$ is equal to

$$\prod_{\hat{v}'_1 \in [r_{u'_1}] \setminus \{v'_1\}} \left(\prod_{\ell \in [2,\Delta]} \prod_{\hat{u}'_\ell \in [s_{\hat{v}'_{\ell-1}}]: s_{\hat{v}'_{\ell-1}} \neq 1, \hat{v}'_\ell \in [r_{\hat{u}'_\ell}]} Q_{\hat{u}'_\ell} \right).$$

On putting the things together, observe that the denominator of $D(R_{u'_1}|E)$ is equal to

$$d_1 := \frac{\prod_{\hat{v}_1 \in [r_{u'_1}]} \left(\prod_{\ell \in [2,\Delta] \hat{u}_\ell \in [s_{\hat{v}_{\ell-1}}]: s_{\hat{v}_{\ell-1}} \neq 1, \hat{v}_\ell \in [r_{\hat{u}_\ell}]} Q_{\hat{u}_\ell} \right)}{\prod_{\ell \in [2,\Delta]: s_{v'_{\ell-1}} \neq 1} Q_{u'_\ell}}.$$

2. Denominator of $D(\overline{R}_{u_1,u'_1}|\overline{E})$: As noted above $\overline{E} = \overline{R}_{u_1,u'_1}$, which immediately implies that $D(\overline{R}_{u_1,u'_1}|\overline{E})$ is a product-depth Δ and top fan-in $(m-2)$ instance of the determinant computed in Lemma 6.1. Thus, by using the induction hypothesis of Lemma 6.1, we get that the denominator of $D(\overline{R}_{u_1,u'_1}|\overline{E})$ is equal to

$$d_2 = \prod_{u''_1 \in A_{u_1,u'_1}, v''_1 \in [r_{u''_1}]} Q_{u''_1} \prod_{\ell \in [2,\Delta]} \left(\prod_{u''_\ell \in [s_{v''_{\ell-1}}]: s_{v''_{\ell-1}} \neq 1, v''_\ell \in [r_{u''_\ell}]} Q_{u''_\ell} \right).$$

3. Denominator of $D(\overline{R}_{u_1,u'_1}|\overline{E}_{(\mathbf{u}'',\mathbf{v}'')})$: Observe that it is the smaller instance of the determinant studied in Claim 6.5.2 and by the induction hypothesis of this claim, we get that the denominator of $D(\overline{R}_{u_1,u'_1}|\overline{E}_{(\mathbf{u}'',\mathbf{v}'')})$ is equal to

$$d_3 := \frac{\prod_{\widehat{u}_1 \in A_{u_1,u'_1,u''_1}} Q_{\widehat{u}_1} \prod_{\ell \in [2,\Delta]} \prod_{\widehat{u}''_\ell \in \widehat{\Sigma}_\ell^{A_{u_1,u'_1}}} Q_{\widehat{u}''_\ell}}{\prod_{\ell \in [2,\Delta]: s_{v''_{\ell-1}} \neq 1} Q_{u''_\ell}},$$

where $A_{u_1,u'_1,u''_1} = [m] \setminus \{u_1, u'_1, u''_1\}$ and $u''_\ell, \ell \in [2,\Delta]$ correspond to $(\mathbf{u}'', \mathbf{v}'')$.

Now, on putting these things together, it is not difficult to see that Equation (6.42) implies

that the denominator of D_1 is equal to

$$\bar{d}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} := \frac{\prod\limits_{\hat{u}_1 \in A_{u_1,u_1'}} Q_{\hat{u}_1'} \cdot \prod\limits_{\ell \in [2,\Delta]} \prod\limits_{\hat{u}_\ell' \in \widehat{\Sigma}_\ell^{A_{u_1}}} Q_{\hat{u}_\ell'}}{\prod\limits_{\ell \in [2,\Delta]: s_{v'_{\ell-1}} \neq 1} Q_{u_\ell'}}.$$

This completes the proof of this part. Then, on normalising the denominator of D_1, Equation (6.42) can be rewritten as

$$\widetilde{D}_1 = \widetilde{D}(R_{u_1'}|E)\left(\widetilde{D}(\overline{R}_{u_1,u_1'}|\overline{E}) - \left(\sum_{\substack{u \in A_{u_1,u_1'}, \\ (\mathbf{u}'',\mathbf{v}'') \in S, u_1'' = u}} \mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')} \prod_{\substack{\ell \in [2,\Delta], \\ \hat{u}_\ell'' \in [s_{v''_{\ell-1}}]: \\ s_{v''_{\ell-1}} \neq 1}} Q_{\hat{u}_\ell''} \cdot \widetilde{D}(\overline{R}_{u_1,u_1'}|\overline{E}_{(\mathbf{u}'',\mathbf{v}'')})\right)\right), \tag{6.44}$$

where $\widetilde{D}_1 = \bar{d}_{(\mathbf{u},\mathbf{v}),(\mathbf{u}',\mathbf{v}')} \cdot D_1$, $\widetilde{D}(R_{u_1'}|E) = d_1 \cdot D(R_{u_1'}|E)$, $\widetilde{D}(\overline{R}_{u_1,u_1'}|\overline{E}) = d_2 \cdot D(\overline{R}_{u_1,u_1'}|\overline{E})$ and $\widetilde{D}(\overline{R}_{u_1,u_1''}|\overline{E}_{(\mathbf{u}'',\mathbf{v}'')}) = d_3 \cdot D(\overline{R}_{u_1,u_1''}|\overline{E}_{(\mathbf{u}'',\mathbf{v}'')})$.

Part 2. Coefficient of $q'_{\mathfrak{v},u_1'}$: Let $(\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u_1',0}$ and $(\mathbf{u},\mathbf{v}) \in S_{\mathfrak{v},u_1,0}$. We use induction on $|A_{u_1}|$ to compute the coefficient of $q'_{\mathfrak{v},u_1'}$ in $\widetilde{D}_1$.

Base Case. $|A_{u_1}| = 1$: Let $A_{u_1} = \{u_1'\}$. In this case $A_{u_1,u_1'} = \emptyset$. It is easy to see from Equation (6.44) that as $A_{u_1,u_1'} = \emptyset$ in this case, we get

$$\widetilde{D}_1 = \widetilde{D}(R_{u_1'}|E). \tag{6.45}$$

Equation (6.43) would be helpful to understand this. Note that on clearing the denominators in this equation, we get the following.

$$\widetilde{D}(R_{u_1'}|E) = \prod_{\hat{v}_1' \in [r_{u_1'}] \setminus \{v_1'\}} \widetilde{\det}(H'(T_{\hat{v}_1'})) \times \widetilde{\det}(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1'},v_1'}). \tag{6.46}$$

Let $\widetilde{D}_2 = \widetilde{\det}(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1'},v_1'})$. Let $\mathfrak{v} \in \mathscr{V}$ be picked arbitrarily. Then, we want to find the

coefficient of $q'_{\mathfrak{v},u'_1}$ in $\widetilde{D}(R_{u'_1}|E)$. For $\hat{v}'_1 \in [r_{u'_1}] \setminus \{v'_1\}$, let

$$p_{1,\hat{v}'_1} = \prod_{\substack{i\in[0,\Delta-2],\\ \hat{u}'_2\in[s_{\hat{v}'_1}]}} \prod_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u'_1,\hat{v}'_1,\hat{u}'_2,i}} \mathbf{X}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{(a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-1)-i},$$

$$q_2 = \left(\prod_{\substack{j\in[2,a(\mathbf{u}',\mathbf{v}')-1]:\\ (\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u'_1,v'_1,u'_2,0},\\ (\hat{\mathbf{u}}',\hat{\mathbf{v}}')_j=(\mathbf{u}',\mathbf{v}')_j,\\ \hat{u}'_{j+1}\neq u'_{j+1}}} \mathbf{X}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-j} \right) \times \left(\prod_{\substack{\hat{u}'_2\in[s_{\hat{v}'_1}]\setminus\{u'_2\},\\ (\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u'_1,v'_1,\hat{u}'_2,0}}} \mathbf{X}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-1} \right) \prod_{\substack{i\in[2,\Delta-1],\\ \hat{u}'_2\in[s'_{v_1}],\\ (\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u'_1,v'_1,\hat{u}'_2,i-1}}} \mathbf{X}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-i},$$

where for any $\hat{v}'_1 \in [r_{u'_1}]$, $S_{\mathfrak{v},u'_1,\hat{v}'_1,\hat{u}'_2,i}$ is defined similar to the one defined in Equation (6.32) and in the definition of q_2, u'_2 is the third coordinate of $(\mathbf{u}',\mathbf{v}')$ from the left. Then, the variants of Equation (6.33) and Equation (6.34) by replacing u_1, v_1 with u'_1, v'_1 respectively imply that $q'_{\mathfrak{v},u'_1} = q_1 \cdot q_2$, where $q_1 = \prod_{\hat{v}'_1\in[r_{u'_1}]\setminus\{v'_1\}} p_{1,\hat{v}'_1}$. This is so because $v'_1 = v_{u'_1}$ is fixed by $\mathfrak{v}$ since $(\mathbf{u}',\mathbf{v}') \in S_{\mathfrak{v},u'_1,0}$. It follows from the discussion similar to the one given after Equation (6.34) that $p_{1,\hat{v}'_1}$ is present in $\widetilde{\det}(H'(T_{\hat{v}'_1}))$. As $\widetilde{\det}(H'(T_{\hat{v}'_1}))$ is a product-depth $\Delta-1$ instance of g_T studied in Lemma 6.1, we get the coefficient of $p_{1,\hat{v}'_1}$ in $\widetilde{\det}(H'(T_{\hat{v}'_1}))$ from the induction hypothesis of Lemma 6.1 on the product-depth is equal to

$$(-1)^{c_{\hat{v}'_1}} \beta_{\mathfrak{v},u'_1,\hat{v}'_1,0} \prod_{i\in[2,\Delta-1]} \prod_{\hat{u}'_2\in[s_{\hat{v}'_1}]} \beta_{\mathfrak{v},u'_1,\hat{v}'_1,\hat{u}'_2,i},$$

where $c_{\hat{v}'_1} = \sum_{\ell\in[2,\Delta]} \left(s(\prod_\ell^{\{(u'_1,\hat{v}'_1)\}}) - r(\Sigma_\ell^{\{(u'_1,\hat{v}'_1)\}}) \right) + s_{\hat{v}'_1} - 1$ and the coefficients $\beta_{\mathfrak{v},u'_1,\hat{v}'_1,0}$ and $\beta_{\mathfrak{v},u'_1,\hat{v}'_1,\hat{u}'_2,i}$, $i \in [2,\Delta-1]$ are similar to those defined in Equations (6.35) and (6.36) respectively. Thus, we get that the coefficient of q_1 in $\prod_{\hat{v}'_1\in[r_{u'_1}]\setminus\{v'_1\}} \widetilde{\det}(H'(T_{\hat{v}'_1}))$ is equal to

$$(-1)^{\sum_{\hat{v}'_1\in[r_{u'_1}]\setminus\{v'_1\}} c_{\hat{v}'_1}} \prod_{\substack{\hat{v}'_1\in[r_{u'_1}]\setminus\{v'_1\},\\ \hat{u}'_2\in[s_{\hat{v}'_1}]}} \beta_{\mathfrak{v},u'_1,\hat{v}'_1,\hat{u}'_2,0} \times \prod_{i\in[2,\Delta-1]} \prod_{\substack{\hat{v}'_1\in[r_{u'_1}]\setminus\{v'_1\},\\ \hat{u}'_2\in[s_{\hat{v}'_1}]}} \beta_{\mathfrak{v},u'_1,\hat{v}'_1,\hat{u}'_2,i}.$$

Now, we argue that q_2 is present in $\widetilde{D}_2$. Observe that if we make the following changes to $q'_{\mathfrak{v},u'_1}$ and use Equation (6.33), it becomes q_2: replace $[\Delta-1]$ with $[2,\Delta-1]$, $[a(\mathbf{u}',\mathbf{v}')-1]$

with $[2, a(\hat{\mathbf{u}}', \hat{\mathbf{v}}') - 1], \hat{u}_1'$ with $\hat{u}_2'$, for $i \geq 1$, $S_{\mathfrak{v},u_1',i}$ with $S_{\mathfrak{v},u_1',v_1',u_2',i-1}$, set $A_{u_1} = [s_{v_1'}], A_{u_1,u_1'} = [s_{v_1'}] \setminus \{u_2'\}$. As noted before, $\widetilde{D}_2$ is a product-depth $(\Delta - 1)$ instance of the determinant studied in Claim 6.5.2, we get that q_2 is present in $\widetilde{D}_2$. Thus, by the induction hypothesis of Claim 6.5.2 on the product-depth $\Delta - 1$ and applying the above mentioned changes, the coefficient of q_2 in $\widetilde{D}_2$ is equal to

$$(-1)^{c_2}\, n_{(\mathbf{u}',\mathbf{v}')} \prod_{\hat{u}_2' \in [s_{v_1'}]} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}') \in S_{\mathfrak{v},u_1',v_1',\hat{u}_2',0}, \\ (\mathbf{u}'',\mathbf{v}'') \in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}} (n_{(\mathbf{u}'',\mathbf{v}'')} - 1) \prod_{i \in [2,\Delta-1]} \beta_{\mathfrak{v},u_1',v_1',\hat{u}_2',i} \right),$$

where $c_2 = \sum_{\ell \in [2,\Delta]} \left(s(\prod_\ell^{[s_{v_1'}]}) - r(\Sigma_\ell^{[s_{v_1'}]}) \right) + s_{v_1'} - 1$. It is easy to verify that $\sum_{\ell \in [2,\Delta]} \left(s(\prod_\ell^{[s_{v_1'}]}) - r(\Sigma_\ell^{[s_{v_1'}]}) \right)$ $= \sum_{\ell \in [2,\Delta]} \left(s(\prod_\ell^{\{(u_1',v_1')\}}) - r(\Sigma_\ell^{\{(u_1',v_1')\}}) \right)$. Thus, $c_2 = \sum_{\ell \in [2,\Delta]} \left(s(\prod_\ell^{\{(u_1',v_1')\}}) - r(\Sigma_\ell^{\{(u_1',v_1')\}}) \right) + s_{v_1'} - 1$. As $(\mathbf{u}', \mathbf{v}') \in S_{\mathfrak{v},u_1',0}$, observe that $S_{\mathfrak{v},u_1',0} = \bigcup_{\hat{u}_2' \in [s_{\hat{v}_1'}] \setminus \{u_2'\}} S_{\mathfrak{v},u_1',v_1',\hat{u}_2',0} \bigcup S_{\mathfrak{v},u_1',v_1',u_2',0}$, which implies the following

$$\prod_{\hat{u}_2' \in [s_{v_1'}]} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}') \in S_{\mathfrak{v},u_1',v_1',\hat{u}_2',0}, \\ (\mathbf{u}'',\mathbf{v}'') \in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}} (n_{(\mathbf{u}'',\mathbf{v}'')} - 1) \right) = \prod_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}') \in S_{\mathfrak{v},u_1',0}} \prod_{(\mathbf{u}'',\mathbf{v}'') \in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}} (n_{(\mathbf{u}'',\mathbf{v}'')} - 1).$$

Note that for $\hat{v}_1' \in [r_{u_1'}] \setminus \{v_1'\}$, $\widetilde{\det}(H'(T_{\hat{v}_1'}))$ and $\widetilde{\det}(H_{(\mathbf{u},\mathbf{v}),\mathbf{j},R_{u_1'},v_1'})$ given in Equation (6.46) are variable disjoint. This implies that the factorization of $q_{\mathfrak{v},u_1'}$ as $q'_{\mathfrak{v},u_1'} = q_1 \cdot q_2$ in $\widetilde{D}(R_{u_1'}|E)$ is unique and thus the coefficient of $q'_{\mathfrak{v},u_1'}$ in $\widetilde{D}(R_{u_1'}|E)$ is equal to the coefficient of $q_1 \cdot q_2$. On multiplying together the coefficients of q_1 and q_2 and using the variants of Equations (6.37) and (6.38) by changing u_1 to u_1' and v_1 to v_1' , we get that the coefficient of $q_{\mathfrak{v},u_1'}$ in $\widetilde{D}_1$ is equal to

$$(-1)^c\, n_{(\mathbf{u}',\mathbf{v}')} \prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}') \in S_{\mathfrak{v},u_1',0}, \\ (\mathbf{u}'',\mathbf{v}'') \in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}} (n_{(\mathbf{u}'',\mathbf{v}'')} - 1) \times \prod_{i \in [\Delta-1]} \beta_{\mathfrak{v},u_1',i},$$

where $c = \prod_{\ell \in [\Delta]} \left(s(\prod_\ell^{\{u_1'\}}) - r(\Sigma_\ell^{\{u_1'\}}) \right)$. This proves the base case.

Induction step. Let $|A_{u_1}| \geq 2$. Recall Equation (6.44). Let $q'_{\mathfrak{v},u'_1}$ be as given in the claim. Let

$$q_1 = \prod_{j\in[a(\mathbf{u}',\mathbf{v}')-1]} \left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u'_1,0},\\ (\hat{\mathbf{u}}',\hat{\mathbf{v}}')_j=(\mathbf{u}',\mathbf{v}')_j,\\ \hat{u}'_{j+1}\neq u'_{j+1}}} \mathbf{X}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-j} \right) \times \prod_{i\in[\Delta-1]} \prod_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},u'_1,i}} \mathbf{X}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}^{a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-i}$$

and

$$q_2 = \prod_{i\in[0,\Delta-1]} \left(\prod_{u''_1\in A_{u_1,u'_1}} \prod_{(\mathbf{u}'',\mathbf{v}'')\in S_{\mathfrak{v},u''_1,i}} \mathbf{X}_{(\mathbf{u}'',\mathbf{v}'')}^{a(\mathbf{u}'',\mathbf{v}'')-i} \right).$$

Then, note that $q'_{\mathfrak{v},u'_1} = q_1 \cdot q_2$. It follows from Equation (6.44) that q_1 is contributed by $\widetilde{D}(R_{u'_1}|E)$ and q_2 is contributed by the other factor of $\widetilde{D}_1$ in the R.H.S of Equation (6.44). As $\widetilde{D}(R_{u'_1}|E)$ and that other factor are variable disjoint, the above factorization of $q'_{\mathfrak{v},u'_1}$ is unique. We have already calculated the coefficient of q_1 in $\widetilde{D}(R_{u'_1}|E)$ in the base case. Observe that q_2 is obtained from $p_{\mathfrak{v}}$ on replacing $[m]$ with A_{u_1,u'_1} and hence is a smaller instance of $p_{\mathfrak{v}}$. Now, we calculate the coefficient of $q_1 \cdot q_2$ in the positive and the negative part of Equation (6.44). As already noted, $\widetilde{D}(\overline{R}_{u_1,u'_1}|\overline{E})$ is a product-depth Δ and top fan-in $(m-2)$ instance of the determinant studied in Lemma 6.1. Thus, it follows from the induction hypothesis of Lemma 6.1 that the coefficient of $q_1 \cdot q_2$ in the positive part of Equation (6.44) is equal to

$$(-1)^{c_1}\, n_{(\mathbf{u}',\mathbf{v}')} \left(\sum_{\substack{u''_1\in A_{u_1,u'_1},\\ (\mathbf{u}'',\mathbf{v}'')\in S_{\mathfrak{v},u''_1,0}}} n_{(\mathbf{u}'',\mathbf{v}'')} - 1 \right) \prod_{\hat{u}'_1\in A_{u_1}} \prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},\hat{u}'_1,0},\\ (\hat{\mathbf{u}},\hat{\mathbf{v}})\in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}} (n_{(\hat{\mathbf{u}},\hat{\mathbf{v}})} - 1) \prod_{\substack{i\in[\Delta-1],\\ \hat{u}'_1\in A_{u_1}}} \beta_{\mathfrak{v},\hat{u}'_1,i}, \quad (6.47)$$

where $c_1 = \sum_{\ell\in[\Delta]} \left(s(\prod_\ell^{A_{u_1}}) - r(\Sigma_\ell^{A_{u_1}}) \right) + |A_{u_1,u'_1}| - 1 = \sum_{\ell\in[\Delta]} \left(s(\prod_\ell^{A_{u_1}}) - r(\Sigma_\ell^{A_{u_1}}) \right) + |A_{u_1}| - 2$. Now, we compute the coefficient of q_2 in the negative part of Equation (6.44). It is easy to verify that a claim similar to Claim 6.5.1 implies that q_2 is present in the $(\mathbf{u}'',\mathbf{v}'')$-th term of Equation (6.44) if and only if $(\mathbf{u}'',\mathbf{v}'') \in S_{\mathfrak{v},u''_1,0}$ for some $u''_1 \in A_{u_1,u'_1}$. As seen before, $\widetilde{D}(\overline{R}_{u_1,u''_1}|\overline{E}_{(\mathbf{u}'',\mathbf{v}'')})$ is a smaller instance of the determinant studied in Claim 6.5.2. Thus, by the induction hypothesis of Claim 6.5.2, the base case and the variant of Claim 6.5.1 for Equation (6.44), the coefficient

of $q_{\mathfrak{v},u'_1}$ in the negative part of Equation (6.44) is equal to

$$(-1)^{c_1}\left(\sum_{\substack{u''_1\in A_{u_1,u'_1},\\(\mathbf{u}'',\mathbf{v}'')\in S_{\mathfrak{v},u''_1,0}}} n_{(\mathbf{u}',\mathbf{v}')}\cdot n_{(\mathbf{u}'',\mathbf{v}'')}\right)\prod_{\hat{u}'_1\in A_{u_1}}\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},\hat{u}'_1,0},\\(\hat{\mathbf{u}},\hat{\mathbf{v}})\in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}}(n_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}-1)\prod_{\substack{i\in[\Delta-1],\\ \hat{u}'_1\in A_{u_1}}}\beta_{\mathfrak{v},\hat{u}'_1,i}. \tag{6.48}$$

On subtracting Equation (6.48) from Equation (6.47), we get that the coefficient of $q_{\mathfrak{v},u'_1}$ in $\widetilde{D}_1$ is equal to

$$(-1)^c\, n_{(\mathbf{u}',\mathbf{v}')}\prod_{\hat{u}'_1\in A_{u_1}}\left(\prod_{\substack{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{\mathfrak{v},\hat{u}'_1,0}\\(\mathbf{u}'',\mathbf{v}'')\in W_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}}(n_{(\mathbf{u}'',\mathbf{v}'')}-1)\prod_{i\in[\Delta-1]}\beta_{\mathfrak{v},\hat{u}'_1,i}\right),$$

where $c=\sum_{\ell\in[\Delta]}\left(s(\prod_\ell^{A_{u_1}})-r(\Sigma_\ell^{A_{u_1}})\right)+|A_{u_1}|-1$. This proves the induction hypothesis.

Part 3. Other details: Let $(\mathbf{u}',\mathbf{v}')\in S_{u'_1}$ be an arbitrary tuple[1] and q_1 be an arbitrary monomial in $\widetilde{D}_1$. Let q_1 be an arbitrary monomial of $\widetilde{D}_1$. Let $(\hat{\mathbf{u}}',\hat{\mathbf{v}}')\in S_{u'_1}$ be such that there exists $i\in[\Delta]$, such that $(\mathbf{u}',\mathbf{v}')_{i-1}=(\hat{\mathbf{u}}',\hat{\mathbf{v}}')_{i-1}, u'_i=\hat{u}'_i$ and $v'_i\neq\hat{v}'_i$. It follows from Equation (6.44) that in $\widetilde{D}_1$, $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ is contributed only by $\widetilde{D}(R_{u'_1}|E)$. We give an upper bound on the degree of $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ in q_1 for $i=1$ and $i\geq 2$ separately. Before proceeding, we note the following observation, which is used in the proof of this part. Its proof is given in the subsequent section.

Observation 6.14 *Let $T=Q_1\cdots Q_m$, where for every $u\in[m], Q_u$ is a $+$-rooted extended canonical ROF of product-depth Δ. Let $(\mathbf{u},\mathbf{v})\in S$ be arbitrary. Then, the degree of $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ in g_T is at most $a(\mathbf{u},\mathbf{v})$.*

Let $i=1$. Then, Equations (6.44), (6.45) and (6.46) imply that in $\widetilde{D}_1$, $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ is contributed only by $\widetilde{\det}(H'(T_{\hat{v}'_1}))$ for some $\hat{v}'_1\in[r_{u'_1}]\setminus\{v'_1\}$. As $\widetilde{\det}(H'(T_{\hat{v}'_1}))$ is a product-depth $\Delta-1$ instance of g_T, it follows from Observation 6.14 applied on $\widetilde{\det}(H'(T_{\hat{v}'_1}))$ that in any monomial of $\widetilde{\det}(H'(T_{\hat{v}'_1}))$, the degree of $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ is at most $a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-1$ as the product-depth in this case has reduced by 1. Thus, $\deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}} q\leq a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-1$.

Suppose $i\geq 2$, then it follows from Equations (6.45) and (6.46) that in $\widetilde{D}_1$, only $\widetilde{D}_2$ contributes $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$. Let $q_1=q'_1\cdot q'_2$ be such that q'_1 and q'_2 are monomials in $\prod_{\hat{v}'_1\in[r_{u'_1}]\setminus\{v'_1\}}\widetilde{\det}(H'(T_{\hat{v}'_1}))$ and $\widetilde{D}_2$ respectively. Note that $\prod_{\hat{v}'_1\in[r_{u'_1}]\setminus\{v'_1\}}\widetilde{\det}(H'(T_{\hat{v}'_1}))$ does not contain $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$. As $\widetilde{D}_2$ is a

[1] We are not assuming here that $(\mathbf{u}',\mathbf{v}')\in S_{\mathfrak{v},u'_1,0}$

product-depth $\Delta-1$ instance of $\widetilde{D}_1$ and as the product-depth has reduced by 1, it follows from the induction hypothesis of this claim that $\deg_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')} q_2' \leq (a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-1)-(i-1)$. Since q_1' does not contain $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$, $\deg_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')} q_1 \leq a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-i$.

Let $(\hat{\mathbf{u}}',\hat{\mathbf{v}}') \in S_{u_1'}$ be such that there exists $i \in [\Delta]$, such that $(\mathbf{u}',\mathbf{v}')_i = (\hat{\mathbf{u}}',\hat{\mathbf{v}}')_i$ and $u_{i+1}' \neq \hat{u}_{i+1}'$. Then, it follows from Equations (6.44) and (6.46) that in $\widetilde{D}_1$, only $\widetilde{D}_2$ contributes $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$. Let $R_{[u_1',u_i']} := \{(\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{k}) \in R : (\hat{\mathbf{u}},\hat{\mathbf{v}})_{i-1} = (\mathbf{u}',\mathbf{v}')_{i-1}, u_i' = \hat{u}_i'\}$ and $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{[u_1',u_i']},v_i'}$ be the sub-matrix of $H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{u_1'},v_1'}$, whose rows and columns are indexed by $\{(\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{k}) \in R_{[u_1',u_i']} : \hat{v}_i = v_i'\}$. Let $\widetilde{D}_{i+1} = \widetilde{\det}(H_{(\mathbf{u}',\mathbf{v}'),\mathbf{j},R_{[u_1',u_i']},v_i'})$. Then, it is not difficult to see that $\widetilde{D}_{i+1}$ is a product-depth $(\Delta-i)$ instance of $\widetilde{D}_1$.

Now, we recursively expand $\widetilde{D}_1$ by using the Laplace's expansion till level i. For example, Equation (6.44) is the expansion of $\widetilde{D}_1$ till level 1 and in the next level, we expand $\widetilde{D}_2$ given in Equation (6.46) using its Laplace's expansion and so on. The last minor that gets expanded in this process is $\widetilde{D}_i$. Then, it is not difficult to show that by doing so, $\widetilde{D}_{i+1}$ is a factor of $\widetilde{D}_1$ and it is the only factor, which contains $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$. Thus, it is sufficient to upper bound the degree of $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ in an arbitrary monomial of $\widetilde{D}_{i+1}$.

Let q_{i+1} be an arbitrary monomial of $\widetilde{D}_{i+1}$. Then, it is easy to see that if $\mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})}$ appears in q_{i+1} for some $(\hat{\mathbf{u}},\hat{\mathbf{v}}) \in S$ then $(\hat{\mathbf{u}},\hat{\mathbf{v}})_i = (\mathbf{u}',\mathbf{v}')_i$. We know that $\widetilde{D}_{i+1}$ is a product-depth $(\Delta-i)$ instance of $\widetilde{D}_1$ and $(\hat{\mathbf{u}}',\hat{\mathbf{v}}')_i = (\mathbf{u}',\mathbf{v}')_i, \hat{u}_{i+1}' \neq u_{i+1}'$. Then, notice that finding the degree of $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ in q_{i+1} is similar to finding the degree of $\mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')} \in S_{u_1''}$ in an arbitrary monomial q_1 of $\widetilde{D}_1$ for some $u_1'' \in A_{u_1,u_1'}$ in the product-depth $(\Delta-i)$ set-up. Thus, by using the induction hypothesis of Claim 6.5.2 on the part of the claim which upper bounds the degree of $\mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}$ in q_1 for some $(\mathbf{u}'',\mathbf{v}'') \in S_{u_1''}$, where $u_1'' \in A_{u_1,u_1'}$, we get that the degree of $\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}$ in any monomial of $\widetilde{D}_{i+1}$ is at most $a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-i$. This implies $\deg_{\mathbf{x}_{(\hat{\mathbf{u}}',\hat{\mathbf{v}}')}}(q_1) \leq a(\hat{\mathbf{u}}',\hat{\mathbf{v}}')-i$.

Now, let $u_1'' \in A_{u_1,u_1'}$ and $(\mathbf{u}'',\mathbf{v}'') \in S_{u_1'}$ be arbitrary. It follows from Equation (6.44) that $\mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}$ is contributed by the factor of $\widetilde{D}_1$ other that $\widetilde{D}(R_{u_1'}|E)$ in the R.H.S of this equation. In this factor, notice that $\widetilde{D}(\overline{R}_{u_1,u_1'}|\overline{E})$ is a product-depth Δ and top fan-in $m-2$ instance of g_T and it follows from Observation 6.14 that the degree of $\mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}$ in any monomial of $\widetilde{D}(\overline{R}_{u_1,u_1'}|\overline{E})$ is at most $a(\mathbf{u}'',\mathbf{v}'')$. Further, let $\hat{u}_1 \in A_{u_1,u_1'}, (\hat{\mathbf{u}},\hat{\mathbf{v}}) \in S_{\hat{u}_1}$. Suppose that either $u_1'' \neq \hat{u}_1$ or there exists $i \in [\Delta]$, such that either $(\mathbf{u}'',\mathbf{v}'')_{i-1} = (\hat{\mathbf{u}},\hat{\mathbf{v}})_{i-1}, u_i'' = \hat{u}_i$ and $v_i'' \neq \hat{v}_i$ or $(\mathbf{u}'',\mathbf{v}'')_i = (\mathbf{u}',\mathbf{v}')_i$ and $u_{i+1}'' \neq \hat{u}_{i+1}$. Since $\widetilde{D}(\overline{R}_{u_1,\hat{u}_1}|\overline{E}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})})$ is a smaller top fan-in instance of $\widetilde{D}_1$ for product-depth Δ, by using the induction hypothesis of Claim 6.5.2, it is not difficult to show that the degree of $\mathbf{x}_{(\mathbf{u}'',\mathbf{v}'')}$ in any monomial of $\left(\mathbf{x}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})} \prod_{\substack{\ell\in[2,\Delta],\\ \hat{u}_\ell''\in[s_{\hat{v}_{\ell-1}}]:s_{\hat{v}_{\ell-1}}\neq 1}} Q_{\hat{u}_\ell''} \cdot \widetilde{D}(\overline{R}_{u_1,u_1'}|\overline{E}_{(\hat{\mathbf{u}},\hat{\mathbf{v}})})\right)$ is at most $a(\mathbf{u}'',\mathbf{v}'')$.

This complete the proof of Claim 6.5.2 ■

Proof of Observation 6.14

Consider $H(T)$ given in the beginning of Section 6.2 before the factors were taken out common from the numerators and denominators of its rows and columns. For every $k \in [n_{(\mathbf{u},\mathbf{v})}]$, take out $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ from the numerator of the $(\mathbf{u},\mathbf{v},k)$-th row of $H(T)$ and $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ from the denominator of the $(\mathbf{u},\mathbf{v},k)$-th row and $(\mathbf{u},\mathbf{v},k)$-th column of $H(T)$. Let $(\mathbf{u}',\mathbf{v}',k') \in R$ be arbitrary. Then, the structure of modified $H(T)$ implies that the numerator of an entry of the $(\mathbf{u}',\mathbf{v}',k')$-th row of $H(T)$ contains $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ if and only if there exists $\ell \in [\Delta]$, such that $(\mathbf{u},\mathbf{v})_{\ell-1} = (\mathbf{u}',\mathbf{v}')_{\ell-1}, u_\ell \neq u'_\ell$. This is so because Q_{u_ℓ} is present in the numerator of every entry of this row and Q_{u_ℓ} contains $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$. This immediately implies that the number of rows in the modified $H(T)$ containing $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ is equal to $\overline{n}_{u_\ell} = \sum_{\ell\in[\Delta]} \sum_{u'_\ell \in [s_{v_{\ell-1}}]\setminus\{u_\ell\}} n_{u'_\ell}$, where $n_{u'_\ell}$ is the number of variables present in the $+$-rooted sub-ROF $Q_{u'_\ell}$. Thus, the total degree of $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ in $\det(H(T))$ is at most $n_{(\mathbf{u},\mathbf{v})} - 2 + \sum_{\ell\in[\Delta]} \overline{n}_{u_\ell}$. Consider the factorization of $\det(H(T))$ given in Claim 6.2.2. Observe that there exists a monomial in $\prod_{(\mathbf{u},\mathbf{v})\in S} \mathbf{x}_{(\mathbf{u},\mathbf{v})}^{n_{(\mathbf{u},\mathbf{v})}-2} \prod_{\ell\in[\Delta]} \left(\prod_{\substack{u_\ell\in[s_{v_{\ell-1}}]:\\ s_{v_{\ell-1}}\neq 1, v_\ell\in[r_{u_\ell}]}} Q_{u_\ell}^{\overline{n}_{u_\ell}-1} \right)$, in which the degree of $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ in g_T is equal to $n_{(\mathbf{u},\mathbf{v})} - 2 + \sum_{\ell\in[\Delta]} (\overline{n}_{u_\ell} - 1)$. It is not difficult to verify that $|\{Q_{u'_\ell} : \ell \in [\Delta], u'_\ell \in [s_{v_{\ell-1}}] \setminus \{u_\ell\}\}| = a(\mathbf{u},\mathbf{v})$. This immediately implies that the degree of $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$ in g_T is at most $a(\mathbf{u},\mathbf{v})$. ■

6.5.7 About the essential variables in $\det(H(T))$

We first show that if T is a regular ROF, such that T computes a polynomial of degree at least 3 then all the variables appearing in T are essential for $\det(H(T))$ and then talk about essential variables in $\det(H(T))$ when T is a $\times$-rooted canonical ROF (without the regularity condition). Consider the following observation.

Observation 6.15 *Let $(\mathbf{u},\mathbf{v}) \in S$ be such that $n_{(\mathbf{u},\mathbf{v})} \geq 2$ and T be a $\times$-rooted ROF, such that $|\mathrm{var}(T)| = n$. Suppose p is an arbitrary monomial in $\det(H(T))$ and $\alpha \in \mathbb{F}$ is the coefficient of p in $\det(H(T))$. Then, for every $j,k \in [n_{(\mathbf{u},\mathbf{v})}], j \neq k$, $\deg_{x_j}(\alpha \cdot p) = \deg_{x_k}(\alpha \cdot p)$.*

Proof: Suppose this is not true and there exist distinct $j,k \in [n_{(\mathbf{u},\mathbf{v})}]$, such that $\deg_{x_j}(\alpha \cdot p) = e_1, \deg_{x_k}(\alpha \cdot p) = e_2$ and $e_1 > e_2$. Let $\mathrm{var}(T) = \{x_1,\ldots,x_n\}$ and $c \in \mathbb{F}^\times \setminus \{1\}$ be such that it is

not a root of unity[1]. Let A be an $n \times n$ diagonal matrix, whose rows and columns are indexed by $\{x_1, \ldots, x_n\}$, such that the x_j-th entry of A is c, x_k-th entry is c^{-1} and all the other diagonal entries are 1. As $\det(A) = 1$, Corollary 2.1 implies that

$$\det(H(T)) = \det(H(T))(A \cdot \mathbf{x}).$$

We get a monomial $c^{e_1 - e_2}\alpha p$ in $\det(H(T))$, which is not true as $c^{e_1-e_2} \neq 1$. Hence, $e_1 = e_2$. □
Now, we talk about the essential variables of a $\times$-rooted regular ROF.

Claim 6.5.3 *Let $n \in \mathbb{N}, \mathbb{F}$ be a field satisfying either $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) \geq n$, T be a $\times$-rooted regular ROF over $\mathbb{F}$ and $\mathbf{x} = \{x_1, \ldots, x_n\}$ be the set of variables appearing in T. If $n \geq 3$ then every variable in $\mathbf{x}$ is present in the Hessian determinant of T.*

Proof: Consider the extended canonical form of T which is $T = Q_1 \ldots Q_m$, where $m \geq 2$ and for every $u \in [m], Q_u$ is a $+$-rooted sub-ROF of T having product-depth equal to Δ, where $\Delta \in \mathbb{N}$. If $\Delta = 0$ then the claim immediately follows from Observation 6.3 since $n \geq 3$.

Suppose $x \in \mathbf{x}$ is an arbitrary variable. If x is connected to a $\times$ gate that computes a polynomial of degree at least 3 then Claim 6.2.2 implies that x is present in $\det(H(T))$. Suppose $x \in \mathbf{x}$ is connected to a $\times$ gate, which computes a degree 2 monomial. Since T is a regular ROF, there does not exist a $+$ gate in T, which has a variable child. Suppose $\Delta = 1$. Then, it is easy to verify from Claim 6.4.1 that as T is a regular ROF, there exists a monomial in g_T that contains x.

Suppose $\Delta \geq 2$. We show that there exists a $\mathfrak{v} \in \mathscr{V}$, such that the monomial $p_{\mathfrak{v}}$ contains x. Let $(\mathbf{u}, \mathbf{v}) \in S$ be such that $x \in \mathbf{x}_{(\mathbf{u},\mathbf{v})}$ and $n_{(\mathbf{u},\mathbf{v})} = 2$. Suppose $\ell \in [\Delta]$ is such that $r_{u_\ell} \geq 2$ and for every $i \in [\ell + 1, \Delta], r_{u_i} = s_{v_{i-1}} = 1$, where u_i, v_i are the coordinates of $(\mathbf{u}, \mathbf{v})$. Since T is a regular ROF, it is easy to see that there exists $\mathfrak{v} \in \mathscr{V}$, such that the entry in $\mathfrak{v}$ labelled by u_ℓ is equal to v_ℓ. Since for every $i \in [\ell + 1, \Delta], r_{u_i} = s_{v_{i-1}} = 1$ the entry in $\mathfrak{v}$ labelled by u_i is equal to v_i for every $i \in [\ell + 1, \Delta]$. Thus, $p_{\mathfrak{v}}$ contains $\mathbf{x}_{(\mathbf{u},\mathbf{v})}$. As either $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) \geq n$, Lemma 6.1 implies that the coefficient of $p_{\mathfrak{v}}$ in g_T is non-zero. Thus, $x \in \text{var}(\det(H(T)))$. □

This along with Observation 2.5 and Observation 2.6 help us show the following.

Claim 6.5.4 *Let $n \in \mathbb{N}$ and $\mathbb{F}$ be a field satisfying either $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) \geq n$. Let T be a $\times$-rooted regular ROF over $\mathbb{F}$, such that $|\text{var}(T)| = n$. If $n \geq 3$ then all the variables present in T are essential for the Hessian determinant of T.*

[1]One way to do this is to pick c from a finite subset of $\mathbb{F}$ size greater than $\deg(\det(H(T)))$. If $\mathbb{F}$ is small, we work with a large enough extension field $\mathbb{G}$ of $\mathbb{F}$ and consider $\det(H(T))$ over $\mathbb{G}$. Then, with high probability, c is not a root of unity.

Proof: Let $\mathbf{x} = \text{var}(T)$. As $n \geq 3$, Lemma 5.1 implies that $\det(H(T)) \neq 0$. Consider the following equation.

$$\sum_{x \in \mathbf{x}} \beta_x \frac{\partial \det(H(T))}{\partial x} = 0,$$

where for every $x \in \mathbf{x}, \beta_x \in \mathbb{F}$. Let $x \in \mathbf{x}$ be an arbitrary variable. Suppose x is a child of a $\times$ gate, which computes a polynomial of degree at least 3. We know from Claim 6.2.2 that x is a factor of $\det(H(T))$. Then, Observation 2.5 implies that $\beta_x = 0$.

Now, suppose x is connected to a $\times$ gate, which computes a degree 2 monomial, say $x \cdot x'$. Observation 6.15 implies that if p is an arbitrary monomial of $\det(H(T))$ then the degrees of x and x' in p are same. We know from Claim 6.5.4 that x appears in $\det(H(T))$, which implies that there exists $i \geq 1$, such that $(x \cdot x')^i$ has non-zero coefficient in $\det(H(T))$. Then, Observation 2.6 implies that $\beta_x = 0$. Since $\beta_x = 0$ for every $x \in \mathbf{x}$, Fact 2.11 implies that every variable in $\mathbf{x}$ is essential for $\det(H(T))$. □

Chapter 7

Conclusion

In this chapter, we summarize our main results and mention some open questions. The central theme of this thesis is the equivalence testing problem (in short, ET). An ET for a polynomial family $\{f_n\}_{n\in\mathbb{N}}$ (similarly, a circuit class $\mathscr{C}$) takes input black-box access to a $g \in \mathbb{F}[\mathbf{x}]$ and decides whether g is in the orbit of an $f \in \{f_n\}_{n\in\mathbb{N}}$ (respectively, a circuit in $\mathscr{C}$). If yes, it outputs an $f \in \{f_n\}_{n\in\mathbb{N}}$ and an $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ (respectively, a $\mathsf{C} \in \mathscr{C}$ and an $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$) such that $g = f(A\mathbf{x})$ (respectively, $g = \mathsf{C}(A\mathbf{x})$). In this thesis, we study equivalence tests for two polynomial families, namely the families of the Nisan-Wigderson design polynomials (in short, NW) and the determinant, and a circuit class, namely the class of regular read-once arithmetic formulas (in short, ROFs). In the process of designing an equivalence test for the family of NW, we also study some fundamental structural and algorithmic results questions to the symmetries of NW. An invertible matrix A is called a symmetry of NW if $\mathsf{NW} = \mathsf{NW}(A\mathbf{x})$. The structural questions for NW that we study are related to properties of characterization by symmetries and characterization by circuit identities, and the algorithmic questions include a circuit testing algorithm and a flip theorem for NW. These questions are important from the perspective of GCT and have been studied for the permanent. The content of this chapter is divided into three parts - the first one is devoted to the results on NW, the second one to the ET for the family of determinant, and the last one to the ET for the class of regular ROFs.

7.1 Structural and algorithmic results on NW

In ACT, many polynomial families like the families of permanent, determinant, iterated matrix multiplication (in short, IMM) polynomial, elementary symmetric polynomial, power symmetric polynomial, Nisan-Wigderson polynomial (NW) etc. have been used to prove lower bounds for various classes of arithmetic circuits. Unlike the other families mentioned here, not much

is known about the family of NW. We know that the family of NW is in VNP but it is not known whether it is in VP or VNP-complete. On the other hand, the families of determinant, IMM, power symmetric polynomial, elementary symmetric polynomial are in VP and the family of permanent is VNP-complete. Several other interesting results are also known for other polynomials. For instance, the permanent, the determinant, the IMM, the power symmetric polynomials are characterized by their symmetries whereas the elementary symmetric polynomial is not characterized by its symmetries. We also have efficient equivalence tests for all the polynomial families mentioned before except the family of NW. In addition to this, two efficient circuit testing algorithms and a flip theorem are also known for the permanent. In the first work of this thesis, we study characterization by symmetries, circuit testing, flip theorem and a special case of ET for the family of NW. We talk about these results below.

Structural results. The first result is related to the characterization by symmetries property (Definition 2.24). This property bypasses the natural proof barrier (see Section 1.3.1). This property is important from the viewpoint of GCT; GCT aims to exploit the characterization by symmetries properties of the permanent and the determinant to show that the permanent of an $n \times n$ symbolic matrix is not an affine projection of a poly(n)-size determinant. In this thesis, we show that NW is characterized by its symmetries over $\mathbb{C}$ but not over $\mathbb{R}$ and $\mathbb{Q}$. On the other hand, the permanent of an $n \times n$ symbolic matrix is characterized by its symmetries over fields having more than n elements [MM62, Gro12]. The main reason for this contrasting result for NW over different fields is as follows: The characterization by symmetries property of NW over $\mathbb{C}$ very crucially uses some symmetries of NW which are obtained from a d-th primitive root of unity, where d is a prime number used as a parameter in the definition of NW. We also show that in the absence of such symmetries over $\mathbb{R}$ and $\mathbb{Q}$, NW can not be characterized by its symmetries over these two fields. To show that NW is not characterized by its symmetries over $\mathbb{R}$ and $\mathbb{Q}$, we use the following two results about NW, which were studied in the author's master's thesis [Gup17]: The structure of the group of symmetries of NW (see Theorem 3.3) and a structural insight obtained from the analysis of the Lie algebra of NW (see Claim 3.1.2).

Apart from this, we also showed that NW is characterized by circuit identities (Definition 2.25) over any field. This property means there are poly(d) many polynomial indentities satisfied by NW, where every polynomial in each of these identities is computable by a poly(d) size arithmetic circuit, and an $f \in \mathbb{F}[\mathbf{x}]$ satisfies these identities if and only if $f = \alpha \cdot \mathsf{NW}$ for some $\alpha \in \mathbb{F}$. The proof of the property that NW is characterized by circuit identities also uses some symmetries of NW. This property implies two algorithmic results for NW namely, a circuit testing algorithm and a flip theorem.

Algorithmic results. Our first algorithmic result is a randomized polynomial time circuit testing algorithm for NW over almost any field. It takes input black-box access to a circuit C and checks whether C computes NW. We know that the family of NW is in VNP but is not known to be in VP. In the absence of a proof that $\mathsf{VP} = \mathsf{VNP}$, it becomes a natural problem to test if given an arithmetic circuit C, whether C computes NW. Circuit testing for NW is a special case of ET for NW and it is also required in an ET for NW. We exploit the characterization by circuit identities property of NW to design a circuit testing algorithm for NW. Two efficient circuit testing algorithms are also known for the permanent [Lip89, Mul10].

The second result is a flip theorem for NW. It says that if NW is not computable by arithmetic circuits of size at most s then in randomized polynomial time, we can compute a list of poly(s) many 'witness points' over the underlying field against all arithmetic circuits of size at most s. This means that for every circuit C of size at most s, there exists a witness point $\mathbf{a}$ in this list such that $\mathsf{NW}(\mathbf{a}) \neq \mathsf{C}(\mathbf{a})$. Using the characterization by circuit identities property of NW, we also show that a polynomial time black-box PIT algorithm for arithmetic circuits of size-$10s$ implies that we can compute the list of above mentioned poly(s) many witness points in polynomial time. A flip theorem is also known for the permanent [Mul10, Mul11b].

The third result is related to equivalence test for the family of NW. We give a randomized polynomial time reduction from general ET for NW to block permuted ET for NW, which determines if there exists an invertible block-permuted matrix (recall the definition from Section 1.3.1) A such that the input polynomial f satisfies $f = \mathsf{NW}(A\mathbf{x})$. We give a randomized polynomial time algorithm for a special case of block-permuted ET for NW, which we call block-diagonal permutation scaling ET (in short, BD-PS ET). A BD-PS ET determines whether there exists a block-diagonal permutation matrix A (recall the definition from Section 1.3.1) and an invertible scaling matrix B such that $f = \mathsf{NW}(AB\mathbf{x})$. This algorithm works over the field of real numbers and over finite fields satisfying $d \nmid (|\mathbb{F}| - 1)$, where d is a prime number used as a parameter in the definition of NW. The BD-PS ET crucially uses symmetries of NW. On the other hand, a complete randomized polynomial time ET is known for the permanent [Kay12].

Future work. Our contributions draw more parallels between the permanent and NW: It was known that both of these polynomials are in VNP. Our work implies that both of these are characterized by its symmetries over $\mathbb{C}$, both have randomized polynomial time circuit testing algorithm, and flip theorems hold for both of these. However, certain aspects of the family of NW are still not clear. We note some interesting open questions about NW below.

1. **Complexity of zero-testing for** NW: The zero-testing for NW is the following algorithmic problem: Suppose d is a prime number. Given a point $\mathbf{a} \in \{0,1\}^{d^2}$, determine whether $\mathsf{NW}(\mathbf{a}) = 0$. This is an interesting problem and has been well-studied in the Boolean complexity theory. It is known as Andreev's problem in the Boolean complexity literature (see [Joh86]). It is easy to see that the zero-testing problem for NW is in NP - if $\mathsf{NW}(\mathbf{a}) = 1$ then we can give a set $\{(i, l_i) : i, l_i \in \mathbb{F}_d\}$ as a certificate, where for every $i \in \mathbb{F}_d, a_{i,l_i} = 1$ and $\prod_{i \in \mathbb{F}_d} x_{i,l_i}$ is a monomial of NW. Is the zero testing for NW NP-complete? This question was stated in [Joh86] and has remained unresolved till date. Recently, [Pot19] gave an $\mathsf{AC}^0[\oplus]$ lower bound for Andreev's problem. On the other hand, the zero-testing problem for the permanent is in P - it is related to checking whether a bipartite graph has a perfect matching.

2. **Complexity of** NW: We know that the family of NW is in VNP. Is the family of NW in VP or is it VNP-complete? Complexities of several other useful polynomial families in ACT like the families of permanent, determinant, IMM, elementary symmetric polynomials etc. are well-studied. If the family of NW is in VP then the zero-testing for NW can be solved efficiently. If this family turns out to be VNP-complete then it would also be good as it will enrich the list of VNP-complete polynomials. Unlike NP-complete functions, we do not have many examples of VNP-complete families. A list of VNP-complete families is given in [Bür00] and most of these polynomials have graph theoretic definition. If NW is VNP-complete then we will also have a VNP-complete family having an algebraic definition. Apart from this, the complexity of NW is also interesting from the viewpoint of GCT because some properties like characterization by symmetries, which plays an important role in GCT, have been studied for NW. If the family of NW turns out to be VNP-complete then it can be a substitute for the permanent in GCT.

3. **A full ET for** NW: We saw a special case of ET for NW, which we called the block-diagonal permutation scaling (BD-PS) ET for NW. This ET crucially used symmetries and other structural insights of NW. We saw in Section 3.2.3 of Chapter 3 that this ET is a special case of the block-permuted equivalence test (in short, BP ET) for NW. We also gave a randomized polynomial time reduction from general ET for NW to its BP ET over almost every field. Thus, a BP ET for NW would imply a full ET for NW. Can we design an efficient BP ET for NW? The ideas used in designing a BD-PS ET for NW can play a crucial role in its BP ET.

7.2 DET over finite fields and $\mathbb{Q}$

In the second work, we study equivalence testing problem for the determinant (in short, DET). This is an important problem from the perspective of GCT. As DET deals with checking if a given polynomial is in the orbit of the determinant, it is a natural first step in the direction of understanding whether permanent is in the orbit closure of a polynomial size determinant. Recall from Section 1.1.2.2 that showing this would separate the complexities of permanent and determinant. A randomized polynomial time DET over $\mathbb{C}$ was given by Kayal [Kay12]. A randomized polynomial time DET over a finite field $\mathbb{F}_q$ was given in [KNS19], where if the input polynomial is equivalent to the determinant then the DET outputs a certificate matrix over a degree n extension field of $\mathbb{F}_q$. Before our work, DET over $\mathbb{Q}$ was not known.

DET over finite fields. In this work, we give a randomized polynomial time DET over finite fields satisfying mild conditions on the size and the characteristic. Our DET algorithm outputs a certificate matrix over the *base field* $\mathbb{F}_q$ and *not* over an extension field of $\mathbb{F}_q$.

DET over $\mathbb{Q}$. We give the *first* randomized DET over $\mathbb{Q}$ with oracle access to an integer factoring algorithm IntFact. If f is equivalent to the determinant of an $n \times n$ symbolic matrix, this DET outputs a certificate matrix over $\mathbb{Q}$. This DET algorithm runs in polynomial time if n is bounded. However, for unbounded n we have a DET over $\mathbb{Q}$, which runs in randomized polynomial time but outputs a certificate matrix over an *extension field* $\mathbb{L}$ of $\mathbb{Q}$ satisfying $[\mathbb{L} : \mathbb{Q}] \leq n$. This DET does not require oracle access to IntFact. We show that it is unlikely to get rid of oracle access to IntFact from the DET over $\mathbb{Q}$. In particular, we show that assuming the Generalized Riemann Hypothesis (GRH), there exists a randomized polynomial time reduction from factoring square-free integers to DET for quadratic forms (i.e., $n = 2$ case) over $\mathbb{Q}$. This shows that assuming GRH, DET for quadratic forms over $\mathbb{Q}$ and IntFact are randomized polynomial time reducible to each other and hence it is unlikely to get rid of IntFact oracle from DET over $\mathbb{Q}$.

Relation between FMAI and DET. Our main technical contribution is to give a randomized polynomial time reduction from DET to another problem called the *full matrix algebra isomorphism* (FMAI). This reduction holds over almost every field. FMAI is a well-studied problem in computer algebra. The FMAI problem takes input an $\mathbb{F}$-algebra $\mathscr{A} \subseteq M_{n^2}(\mathbb{F})$ and decides whether $\mathscr{A}$ is isomorphic to the full matrix algebra $M_n(\mathbb{F})$. If the answer is yes, it also outputs an $\mathbb{F}$-algebra isomorphism $\varphi : \mathscr{A} \to M_n(\mathbb{F})$. FMAI algorithms are known over finite fields and over $\mathbb{Q}$ (see Section 2.2.4).

Our reduction from DET to FMAI exploits the rich structure of the Lie algebra of the determinant, denoted $\mathfrak{g}_{\mathsf{Det}_n}$. It is well-known that over a field $\mathbb{F}$ satisfying $char(\mathbb{F}) \nmid n$, $\mathfrak{g}_{\mathsf{Det}_n} = \mathscr{L}_{\text{row}} \oplus \mathscr{L}_{\text{col}}$, where $\mathscr{L}_{\text{row}} := Z_n \otimes I_n, \mathscr{L}_{\text{col}} := I_n \otimes Z_n$, I_n is the identity matrix and Z_n is the set of $n \times n$ traceless matrices over $\mathbb{F}$. Suppose $f = \mathsf{Det}(A\mathbf{x})$ for some $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ then the Lie algebra of f, denoted $\mathfrak{g}_f$, is a direct sum of $\mathscr{F}_{\text{row}} := A^{-1} \cdot \mathscr{L}_{\text{row}} \cdot A$ and $\mathscr{F}_{\text{col}} := A^{-1} \cdot \mathscr{L}_{\text{col}} \cdot A$. Suppose $f \in \mathbb{F}[\mathbf{x}]$ is the input of DET. In the first phase, the algorithm decomposes $\mathfrak{g}_f$ as follows: It computes $\mathfrak{g}_f$ and then computes a special set $\mathscr{P}$ of linear operators on $\mathfrak{g}_f$. Then, it computes irreducible invariant spaces (Definition 2.16) of $\mathscr{P}$. We show that if f is equivalent to Det_n then $\mathscr{F}_{\text{col}}$ and $\mathscr{F}_{\text{row}}$ are the only irreducible invariant subspaces of $\mathscr{P}$. This is how the algorithm gets hold of $\mathscr{F}_{\text{col}}$ and $\mathscr{F}_{\text{row}}$. The reason this decomposition is important is because the algebra $\mathscr{A}$ generated by $\mathscr{F}_{\text{col}}$ is isomorphic to $M_n(\mathbb{F})$. On invoking FMAI on $\mathscr{A}$, we get an $\mathbb{F}$-algebra isomorphism $\varphi : \mathscr{A} \to M_n(\mathbb{F})$. Using φ and the Skolem-Noether theorem (Theorem 2.1), we compute an $A \in \mathrm{GL}(n^2, \mathbb{F})$ such that $f = \mathsf{Det}_n(A\mathbf{x})$, provided f is equivalent to Det_n.

We also give a reduction from DET to FMAI over fields satisfying $char(\mathbb{F}) \nmid n$, which is efficient when the value of the parameter n is bounded. This reduction crucially uses the property that the determinant is characterized by its Lie algebra, i.e., if $f \in \mathbb{F}[\mathbf{x}]$ such that $\mathscr{L}_{\text{col}}$ is an $\mathbb{F}$-subspace of $\mathfrak{g}_f$ then f is a scalar multiple of the determinant. In a follow up work, [MNS20] gave a randomized polynomial time reduction from FMAI to DET for any n.

Future work. Now, we mention an open question in this direction.

1. **An efficient DET over $\mathbb{Q}$ with oracle access to IntFact**. We saw in the first part of Theorem 1.7 that if we insist that our DET algorithm outputs a certificate matrix over $\mathbb{Q}$, the DET algorithm takes oracle access to IntFact and is efficient only when the input parameter n is bounded. We also saw in Theorem 1.9 that it is unlikely to get rid of IntFact oracle from this variant of DET over $\mathbb{Q}$. Can we design a randomized polynomial time algorithm, that takes input black-box access to an n^2-variate polynomial f over $\mathbb{Q}$, has oracle access to IntFact, and determines whether f and Det_n are equivalent, outputs a certificate matrix over $\mathbb{Q}$ and runs in poly(n) time for every value of n?

7.3 An ET for regular ROFs

In the third work, we give the first randomized polynomial time ET for the class of regular ROFs. This ET takes oracle access to quadratic form equivalence (in short, QFE) and works over fields satisfying some mild restrictions on the size and the characteristic. ET for regular ROFs generalizes QFE over $\mathbb{C}$ and ET algorithms for two sub-classes of regular ROFs, namely

the class of sum-product polynomials and the class of ROANFs. Efficient ET algorithms for these two sub-classes were given recently in [MS21]. We gave a randomized polynomial time ET for the class of general ROFs in a follow-up work [GST22]. This ET also takes oracle access to QFE and works over fields satisfying some mild restrictions on the size and the characteristic. The ET for general ROFs is not a part of this thesis.

Hessian determinant of a regular ROF. The ET for the class of regular ROFs crucially uses some properties of the Hessian determinant of a regular ROF. We study these properties for the Hessian determinant of a canonical ROF (Definition 2.39). Since a regular ROF is canonical by definition, all these properties also hold for the Hessian determinant of a regular ROF. We list these properties below.

- **Non-zeroness of the Hessian determinant**. We show that if $\mathtt{C}$ is a canonical ROF then over any field $\mathbb{F}$ satisfying $char(\mathbb{F}) = 0$ or $char(\mathbb{F}) \geq |\text{var}(\mathtt{C})|$, the Hessian determinant of $\mathtt{C}$, denoted $\det(H_{\mathtt{C}})$, is non-zero. However, this may not be true over a finite field $\mathbb{F}$ satisfying $char(\mathbb{F}) < |\text{var}(\mathtt{C})|$. For example, if $\mathtt{C} = x_1x_2x_3$ then $\det(H_{\mathtt{C}}) = 0$ over $\mathbb{F}$ satisfying $char(\mathbb{F}) = 2$. We prove the non-zeroness of $\det(H_{\mathtt{C}})$ by analysing the structures and the coefficients of some *nice monomials* in $\det(H_{\mathtt{C}})$.

- **Essential variables of the Hessian determinant**. We show that if T is a $\times$-rooted regular ROF such that $\deg(T) \geq 3$ then every variable appearing in T is essential for $\det(H_T)$. This result is obtained by analysing the structures of nice monomials in $\det(H_T)$ and using the fact that T is regular. On the other hand, if $\deg(T) = 2$ then $\det(H_T) \in \mathbb{F}^{\times}$.

- **Factors of the Hessian determinant**. We study factors of the Hessian determinant of a $\times$-rooted canonical ROF T and show that if $\deg(T) \geq 3$, then there exists a child of the topmost $\times$ gate of T, which is also a factor of $\det(H_T)$.

Efficient ET for regular ROFs. The input of the ET algorithm is black-box access to an $f \in \mathbb{F}[\mathbf{x}]$ in the orbit of a $+$-rooted regular ROF. Using the polynomial factorization algorithm in [KT90], we can reduce the ET for a $\times$-rooted ROF to the ET for a $+$-rooted ROF. In the first phase, the ET algorithm computes an $A \in \text{GL}(|\mathbf{x}|, \mathbb{F})$ such that $f(A\mathbf{x})$ is the sum of variable disjoint $\times$-rooted terms. To accomplish this phase, the algorithm uses the knowledge of essential variables in the Hessian determinant of a regular ROF, a basic approach by Kayal (see Section 1.4.3) and a QFE over $\mathbb{F}$. In the next phase, the algorithm gets black-box access to every $\times$-rooted terms of $f(A\mathbf{x})$ from black-box of f, then factorizes these terms using the algorithm in [KT90] and then recurses on these terms. The knowledge of the factors of the

Hessian determinant of a regular ROF plays a crucial role in obtaining black-box access to a $\times$-rooted term of $f(A\mathbf{x})$ using just one black-box query to f. It is important that we use exactly one black-box query to f in this case, otherwise the running time of the algorithm can be exponential in $|\mathbf{x}|$ as the product-depth of f can be as large as $|\mathbf{x}|$. This is how the knowledge of essential variables and factors of the Hessian determinant of a regular ROF plays an important role in the ET for the class of regular ROFs.

Future work. Now we mention some open questions in this direction.

1. **Equivalence test for univariate-substituted ROFs**. A *univariate-substituted ROF* is an arithmetic formula, which is obtained from an ROF $\mathtt{C}$ by replacing every variable x in $\mathtt{C}$ with a univariate polynomial in x. The class of univariate-substituted ROFs is well-studied - polynomial time black-box PIT algorithm and reconstruction algorithm for univariate-substituted ROFs are known [SV14, MV18][1]. We have given a randomized polynomial time equivalence test for the class of general ROFs in [GST22]. As the class of univariate-substituted ROFs generalizes the class of ROFs, it is natural to ask the following: Can we design an efficient ET for the class of univariate-substituted ROFs?

   An ET for univariate-substituted ROFs is an important problem as it generalizes QFE over arbitrary fields not having characteristic equal to two: Let $\mathbb{F}$ be a field such that $char(\mathbb{F}) \neq 2$ and $g \in \mathbb{F}[\mathbf{x}]$ be a quadratic form not having redundant variables (Definition 2.32). It follows from the well-known classification results on quadratic forms that g is in the orbit of $h := \alpha_1 x_1^2 + \cdots + \alpha_n x_n^2$, where $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$. Thus, QFE over $\mathbb{F}$ boils down to checking if two given quadratic forms are equivalent to h. Note that h is a depth 2 univariate-substituted ROF. Thus, an ET for univariate-substituted ROF would generalize QFE over $\mathbb{F}$. An ET for univariate-substituted ROFs would also generalize the reconstruction algorithm for univariate-substituted ROFs studied in [SV14, MV18] and equivalence test for sum of univariates model considered in [GKP18]. We hope that the ideas used in the ET for general ROFs given in [GST22] can be helpful in designing an equivalence test for the class of univariate-substituted ROFs. We are also hopeful that the detailed analysis of the Hessian determinant of a canonical ROF given in Chapter 6 can be helpful in analysing the Hessian determinant of a univariate-substituted ROF, which can be an important component in the ET for univariate substituted ROFs.

2. **Equivalence test for other classes of circuits**. As mentioned in Section 1.1.4, PIT for orbits of various circuit classes like ROFs, sparse polynomials, bounded-width ROABPs

[1] A univariate-substituted ROF was called a *preprocessed ROF* in [SV14, MV18].

etc. have been studied recently in [MS21, ST21, BG21]. Like ROFs, do the classes of sparse polynomials and ROABPs also admit efficient equivalence tests?

A variant of equivalence testing for sparse polynomials, called *equivalence testing under shifts*, over rings has been recently studied in [CGS22]. Equivalence testing for sparse polynomials under shifts is the following algorithmic problem: Let $f \in R[\mathbf{x}]$ be a polynomial over a ring R. Given f, determine if there exists an $\mathbf{a} \in R^{|\mathbf{x}|}$ such that number of non-zero monomials in $f(\mathbf{x} + \mathbf{a})$ is less than the number of non-zero monomials in f. Recently, [CGS22] showed some hardness results for equivalence testing for sparse polynomials under shifts over integral domains, which are not fields.

3. **Reconstruction of random arithmetic formulas**. As mentioned in Section 1.3.3, an efficient ET for the class of ROFs implies an efficient algorithm to reconstruct random arithmetic formulas in the *high number of variables* setting, i.e., when the number of variables n is greater than the size s of the formula. Can we reconstruct random arithmetic formulas efficiently when n is much smaller than s?

Bibliography

[Aar16] Scott Aaronson. The p=?np. *Open problems in mathematics*, 2016. https://www.scottaaronson.com/papers/pnp.pdf. 3, 6, 8

[AB03] Manindra Agrawal and Somenath Biswas. Primality and identity testing via chinese remaindering. *J. ACM*, 50(4):429–443, jul 2003. 3, 9

[AFS+18] Matthew Anderson, Michael A. Forbes, Ramprasad Saptharishi, Amir Shpilka, and Ben Lee Volk. Identity Testing and Lower Bounds for Read-k Oblivious Algebraic Branching Programs. *TOCT*, 10(1):3:1–3:30, 2018. Conference version appeared in the proceedings of CCC 2016. 27

[AGKS14] Manindra Agrawal, Rohit Gurjar, Arpita Korwar, and Nitin Saxena. Hitting-sets for roabp and sum of set-multilinear circuits. *SIAM Journal on Computing*, 44, 06 2014. 234

[Agr05] Manindra Agrawal. Proving lower bounds via pseudo-random generators. In *Proceedings of the 25th International Conference on Foundations of Software Technology and Theoretical Computer Science*, FSTTCS '05, page 92–105, Berlin, Heidelberg, 2005. Springer-Verlag. 233

[AKS02] Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. Primes is in p. *Annals of Mathematics*, 160:781 – 793, 09 2002. 3, 9

[AM10] V. Arvind and Partha Mukhopadhyay. The ideal membership problem and polynomial identity testing. *Inf. Comput.*, 208(4):351–363, apr 2010. 233

[AMV15] Matthew Anderson, Dieter Melkebeek, and Ilya Volkovich. Derandomizing polynomial identity testing for multilinear constant-read formulae. *Computational Complexity*, 24:695–776, 2015. 233, 234

[Ang88] Dana Angluin. Queries and concept learning. *Machine Learning*, 2(4):319–342, apr 1988. 10

[Ara11] Manuel Araújo. Classification of quadratic forms. https://www.math.tecnico.ulisboa.pt/~ggranja/manuel.pdf, 2011. 14, 27, 70

[AS05] Manindra Agrawal and Nitin Saxena. Automorphisms of finite rings and applications to complexity of problems. In *23rd Annual Symposium on Theoretical Aspects of Computer Science, STACS 2005*, pages 1–17, 2005. 13, 14

[AS06] Manindra Agrawal and Nitin Saxena. Equivalence of f-algebras and cubic forms. In *23rd Annual Symposium on Theoretical Aspects of Computer Science, STACS 2006*, pages 115–126, 2006. 14

[ASSS16] Manindra Agrawal, Chandan Saha, Ramprasad Saptharishi, and Nitin Saxena. Jacobian Hits Circuits: Hitting Sets, Lower Bounds for Depth-D Occur-k Formulas and Depth-3 Transcendence Degree-k Circuits. *SIAM J. Comput.*, 45(4):1533–1562, 2016. Conference version appeared in the proceedings of STOC 2012. 233, 234

[AT85] V. S. Alagar and Mai Thanh. Fast polynomial decomposition algorithms. In Bob F. Caviness, editor, *EUROCAL '85*, pages 150–153, Berlin, Heidelberg, 1985. Springer Berlin Heidelberg. 13

[Ats06] Albert Atserias. Distinguishing SAT from polynomial-size circuits, through black-box queries. In *21st Annual IEEE Conference on Computational Complexity (CCC 2006), 16-20 July 2006, Prague, Czech Republic*, pages 88–95, 2006. 22

[AV08] Manindra Agrawal and V. Vinay. Arithmetic Circuits: A Chasm at Depth Four. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 67–75. IEEE Computer Society, 2008. 231, 233

[AW21] Josh Alman and Virginia Vassilevska Williams. A refined laser method and faster matrix multiplication. SODA '21, page 522–539, USA, 2021. Society for Industrial and Applied Mathematics. 2

[BÖ00] Peter Bürgisser. Cook's versus valiant's hypothesis. *Theor. Comput. Sci.*, 235(1):71–88, mar 2000. 5

[Bab16] László Babai. Graph isomorphism in quasipolynomial time [extended abstract]. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '16, page 684–697. Association for Computing Machinery, 2016. 1

[BBB+00] Amos Beimel, Francesco Bergadano, Nader H. Bshouty, Eyal Kushilevitz, and Stefano Varricchio. Learning functions represented as multiplicity automata. *J. ACM*, 47(3):506–530, may 2000. 235

[Ber70] Elwyn R Berlekamp. Factoring polynomials over large finite fields. *Mathematics of Computation*, 24:713–735, 1970. 61, 99

[Ber84] Stuart J. Berkowitz. On computing the determinant in small parallel time using a small number of processors. *Information Processing Letters*, 18(3):147–150, 1984. 2

[BFP15] Jérémy Berthomieu, Jean-Charles Faugère, and Ludovic Perret. Polynomial-time algorithms for quadratic isomorphism of polynomials: The regular case. *J. Complex.*, 31(4):590–616, 2015. 14

[BG21] Vishwas Bhargava and Sumanta Ghosh. Improved hitting set for orbit of roabps. In Mary Wootters and Laura Sanità, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2021, August 16-18, 2021, University of Washington, Seattle, Washington, USA (Virtual Conference)*, volume 207 of *LIPIcs*, pages 30:1–30:23. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. 10, 16, 27, 205

[BGKS21] Vishwas Bhargava, Ankit Garg, Neeraj Kayal, and Chandan Saha. Learning generalized depth-three arithmetic circuits in the non-degenerate case. *Electron. Colloquium Comput. Complex.*, page 155, 2021. 236

[BHH95] Nader H. Bshouty, Thomas R. Hancock, and Lisa Hellerstein. Learning arithmetic read-once formulas. *SIAM J. Comput.*, 24(4):706–735, 1995. Conference version appeared in the proceedings of STOC 1992. 27, 29

[BLS16] Nikhil Balaji, Nutan Limaye, and Srikanth Srinivasan. An almost cubic lower bound for ΣΠΣ circuits computing a polynomial in VP. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:143, 2016. 231

[BMS13] M. Beecken, J. Mittmann, and N. Saxena. Algebraic independence and blackbox identity testing. *Information and Computation*, 222:2–19, 2013. 38th International Colloquium on Automata, Languages and Programming (ICALP 2011). 233

[BOT88] Michael Ben-Or and Prasoon Tiwari. A deterministic algorithm for sparse multivariate polynomial interpolation. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, STOC '88, page 301–309, New York, NY, USA, 1988. Association for Computing Machinery. 233, 234

[BR90] László Babai and Lajos Rónyai. Computing irreducible representations of finite groups. *Mathematics of Computation*, 55(192):705–722, 1990. 71

[Bre76] R.P Brent. Multiple-precision zero-finding methods and the complexity of elementary function evaluation. *Analytic Computational Complexity*, pages 151–176, 1976. 93

[BS83] Walter Baur and Volker Strassen. The Complexity of Partial Derivatives. *Theor. Comput. Sci.*, 22:317–330, 1983. 7, 230

[BSV20] Vishwas Bhargava, Shubhangi Saraf, and Ilya Volkovich. Reconstruction of depth-4 multilinear circuits. In *Proceedings of the Thirty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '20, page 2144–2160, USA, 2020. Society for Industrial and Applied Mathematics. 235

[BSV21] Vishwas Bhargava, Shubhangi Saraf, and Ilya Volkovich. *Reconstruction Algorithms for Low-Rank Tensors and Depth-3 Multilinear Circuits*, page 809–822. Association for Computing Machinery, New York, NY, USA, 2021. 235

[Bür00] Peter Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*, volume 7 of *Algorithms and computation in mathematics*. Springer, 2000. 4, 5, 200

[BZ85] David R. Barton and Richard Zippel. Polynomial decomposition algorithms. *Journal of Symbolic Computation*, 1(2):159–168, 1985. 12, 13

[Car06] Enrico Carlini. Reducing the number of variables of a polynomial. In *Algebraic Geometry and Geometric Modeling*, pages 237–247. Springer Berlin Heidelberg, 2006. 53, 62

[CELS18] Suryajith Chillara, Christian Engels, Nutan Limaye, and Srikanth Srinivasan. A Near-Optimal Depth-Hierarchy Theorem for Small-Depth Multilinear Circuits. In

Mikkel Thorup, editor, *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 934–945. IEEE Computer Society, 2018. 232

[CGS22] Suryajith Chillara, Coral Grichener, and Amir Shpilka. On hardness of testing equivalence to sparse polynomials under shifts, 2022. https://arxiv.org/abs/2207.10588. 205

[CIK97] Alexander Chistov, Gábor Ivanyos, and Marek Karpinski. Polynomial time algorithms for modules over finite dimensional algebras. In *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation*, ISSAC '97, page 68–74, New York, NY, USA, 1997. Association for Computing Machinery. 106

[CKS18] Chi-Ning Chou, Mrinal Kumar, and Noam Solomon. Some closure results for polynomial factorization and applications, 2018. https://arxiv.org/abs/1803.05933. 233, 234

[CKSV20] Prerona Chatterjee, Mrinal Kumar, Adrian She, and Ben Lee Volk. A quadratic lower bound for algebraic branching programs. In *Proceedings of the 35th Computational Complexity Conference*, CCC '20, Dagstuhl, DEU, 2020. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. 20, 230

[CKW11] Xi Chen, Neeraj Kayal, and Avi Wigderson. Partial Derivatives in Arithmetic Complexity and Beyond. *Foundations and Trends in Theoretical Computer Science*, 6(1-2):1–138, 2011. 21, 50, 51, 232

[CM14] Suryajith Chillara and Partha Mukhopadhyay. Depth-4 lower bounds, determinantal complexity: A unified approach. In *31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014), STACS 2014, March 5-8, 2014, Lyon, France*, pages 239–250, 2014. 19

[CMM17] Sunil K. Chebolu, Dan McQuillan, and Ján Mináč. Witt's cancellation theorem seen as a cancellation. *Expositiones Mathematicae*, 35(3):300–314, 2017. 17

[Coh03] Joel S. Cohen. *Computer Algebra and Symbolic Computation*. CRC Press, 2003. 3, 12, 13

[Csa76] L. Csanky. Fast parallel matrix inversion algorithms. *SIAM Journal on Computing*, 5(4):618–623, 1976. 2

[CT65] James W. Cooley and John W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, 19:297–301, 1965. 3

[CU13] Henry Cohn and Christopher Umans. Fast matrix multiplication using coherent configurations. In *SODA*, 2013. 2

[CW90] Don Coppersmith and Shmuel Winograd. Matrix multiplication via arithmetic progressions. *J. Symb. Comput.*, 9(3):251–280, mar 1990. 2

[CZ81] David Geoffrey Cantor and Hans Zassenhaus. A new algorithm for factoring polynomials over finite fields. *Mathematics of Computation*, 36:587–592, 1981. 3

[de 97] W.A. de Graaf. Calculating the structure of a semisimple lie algebra. *Journal of Pure and Applied Algebra*, 117-118:319–329, 1997. 106

[Dic89] Matthew Thomas Dickerson. *The functional decomposition of polynomials*. PhD thesis, Cornell university, 1989. 12

[Dic93] Matthew T. Dickerson. General polynomial decomposition and the s-1-decomposition are np-hard. *International Journal of Foundations of Computer Science*, 04(02):147–156, 1993. 13

[DKSS08] Anindya De, Piyush P. Kurur, Chandan Saha, and Ramprasad Saptharishi. Fast integer multiplication using modular arithmetic. In Cynthia Dwork, editor, *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 499–506. ACM, 2008. 3

[DL78] Richard A. Demillo and Richard J. Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7(4):193–195, 1978. 9, 60

[DMPY12] Zeev Dvir, Guillaume Malod, Sylvain Perifel, and Amir Yehudayoff. Separating multilinear branching programs and formulas. In Howard J. Karloff and Toniann Pitassi, editors, *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 615–624. ACM, 2012. 232

[DS07] Zeev Dvir and Amir Shpilka. Locally decodable codes with two queries and polynomial identity testing for depth 3 circuits. *SIAM Journal on Computing*, 36(5):1404–1434, 2007. 233

[DS13] A. Davie and AJ Stothers. Improved bound for complexity of matrix multiplication. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 143, 04 2013. 2

[DSY10] Zeev Dvir, Amir Shpilka, and Amir Yehudayoff. Hardness-randomness tradeoffs for bounded depth arithmetic circuits. *SIAM Journal on Computing*, 39(4):1279–1293, 2010. 233

[Dvi08] Zeev Dvir. On the size of kakeya sets in finite fields. *Journal of the American Mathematical Society*, 22(4):1093–1097, jun 2008. 2

[Dvi12] Zeev Dvir. Incidence theorems and their applications. *Foundations and Trends® in Theoretical Computer Science*, 6, 08 2012. 2

[FGT16] Stephen Fenner, Rohit Gurjar, and Thomas Thierauf. Bipartite perfect matching is in quasi-nc. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '16, page 754–763, New York, NY, USA, 2016. Association for Computing Machinery. 9

[FK09] Lance Fortnow and Adam R. Klivans. Efficient learning algorithms yield circuit lower bounds. *J. Comput. Syst. Sci.*, 75(1):27–36, 2009. 9, 234

[FLMS15] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower Bounds for Depth-4 Formulas Computing Iterated Matrix Multiplication. *SIAM J. Comput.*, 44(5):1173–1201, 2015. Conference version appeared in the proceedings of STOC 2014. 231

[FP09a] Jean-Charles Faugère and Ludovic Perret. An efficient algorithm for decomposing multivariate polynomials and its applications to cryptography. *Journal of Symbolic Computation*, 44:1676–1689, 12 2009. 13

[FP09b] Jean-Charles Faugère and Ludovic Perret. High order derivatives and decomposition of multivariate polynomials. pages 207–214, 01 2009. 13

[FPS08] Lance Fortnow, Aduri Pavan, and Samik Sengupta. Proving SAT does not have small circuits with an application to the two queries problem. *J. Comput. Syst. Sci.*, 74(3):358–363, 2008. 22

[Fro97] Georg Frobenius. Ueber die darstellung der endlichen gruppen durch linearc substitutionen. *Sitzungber. der Berliner Akademie*, 7:994–1015, 1897. 20

[FS13] Michael A. Forbes and Amir Shpilka. Quasipolynomial-Time Identity Testing of Non-commutative and Read-Once Oblivious Algebraic Branching Programs. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 243–252, 2013. 234, 235

[FSS14] Michael A. Forbes, Ramprasad Saptharishi, and Amir Shpilka. Hitting sets for multilinear read-once algebraic branching programs, in any order. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14, page 867–875, New York, NY, USA, 2014. Association for Computing Machinery. 233, 234

[FSV17] Michael A. Forbes, Amir Shpilka, and Ben Lee Volk. Succinct hitting sets and barriers to proving algebraic circuits lower bounds. STOC 2017, page 653–664, New York, NY, USA, 2017. Association for Computing Machinery. 20

[FvzGP10] Jean-Charles Faugère, Joachim von zur Gathen, and Ludovic Perret. Decomposition of generic multivariate polynomials. In *Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation*, ISSAC '10, page 131–137, New York, NY, USA, 2010. Association for Computing Machinery. 13

[Fü09] Martin Fürer. Faster integer multiplication. *Proceedings of the Annual ACM Symposium on Theory of Computing*, 39, 01 2009. 3

[Ges16] Fulvio Gesmundo. Gemetric aspects of iterated matrix multiplication. *Journal of Algebra*, 461:42–64, 2016. 20

[GG13] Joachim von zur Gathen and Jrgen Gerhard. *Modern Computer Algebra*. Cambridge University Press, USA, 3rd edition, 2013. 3

[GGKS19] Ankit Garg, Nikhil Gupta, Neeraj Kayal, and Chandan Saha. Determinant equivalence test over finite fields and over Q. In Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi, editors, *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*, volume 132 of *LIPIcs*, pages 62:1–62:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. 24, 94

[GK86] S Goldwasser and J Kilian. Almost all primes can be quickly certified. In *Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing*, STOC

'86, page 316–329, New York, NY, USA, 1986. Association for Computing Machinery. 3

[GK98] Dima Grigoriev and Marek Karpinski. An Exponential Lower Bound for Depth 3 Arithmetic Circuits. In *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*, pages 577–582, 1998. 231

[GKKS14a] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the chasm at depth four. *J. ACM*, 61(6), dec 2014. 24

[GKKS14b] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the Chasm at Depth Four. *J. ACM*, 61(6):33:1–33:16, 2014. Conference version appeared in the proceedings of CCC 2013. 231

[GKKS16] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic circuits: A chasm at depth 3. *SIAM Journal on Computing*, 45(3):1064–1079, 2016. 231, 233

[GKL11] Ankit Gupta, Neeraj Kayal, and Satyanarayana V. Lokam. Efficient Reconstruction of Random Multilinear Formulas. In *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pages 778–787, 2011. 235

[GKP18] Ignacio García-Marco, Pascal Koiran, and Timothée Pecatte. Polynomial Equivalence Problems for Sum of Affine Powers. In Manuel Kauers, Alexey Ovchinnikov, and Éric Schost, editors, *Proceedings of the 2018 ACM on International Symposium on Symbolic and Algebraic Computation, ISSAC 2018, New York, NY, USA, July 16-19, 2018*, pages 303–310. ACM, 2018. 11, 18, 204

[GKQ13] Ankit Gupta, Neeraj Kayal, and Youming Qiao. Random Arithmetic Formulas Can Be Reconstructed Efficiently. In *Proceedings of the 28th Conference on Computational Complexity, CCC 2013, K.lo Alto, California, USA, 5-7 June, 2013*, pages 1–9, 2013. 28, 236

[GKS20] Ankit Garg, Neeraj Kayal, and Chandan Saha. Learning sums of powers of low-degree polynomials in the non-degenerate case. In Sandy Irani, editor, *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 889–899. IEEE, 2020. 236

[GKSS17] Joshua A. Grochow, Mrinal Kumar, Michael E. Saks, and Shubhangi Saraf. Towards an algebraic natural proofs barrier via polynomial identity testing. *CoRR*, abs/1701.01717, 2017. 20

[GQ21] Joshua A. Grochow and Youming Qiao. On the complexity of isomorphism problems for tensors, groups, and polynomials I: tensor isomorphism-completeness. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference, ITCS 2021, January 6-8, 2021, Virtual Conference*, volume 185 of *LIPIcs*, pages 31:1–31:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. 14

[GQT21] Joshua A. Grochow, Youming Qiao, and Gang Tang. Average-case algorithms for testing isomorphism of polynomials, algebras, and multilinear forms. In Markus Bläser and Benjamin Monmege, editors, *38th International Symposium on Theoretical Aspects of Computer Science, STACS 2021, March 16-19, 2021, Saarbrücken, Germany (Virtual Conference)*, volume 187 of *LIPIcs*, pages 38:1–38:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. 14

[GR00] Dima Grigoriev and Alexander A. Razborov. Exponential Lower Bounds for Depth 3 Arithmetic Circuits in Algebras of Functions over Finite Fields. *Appl. Algebra Eng. Commun. Comput.*, 10(6):465–487, 2000. Conference version appeared in the proceedings of FOCS 1998. 231

[Gro12] Joshua Abraham Grochow. *Symmetry and equivalence relations in classical and geometric complexity theory*. PhD thesis, Department of Computer Science, The University of Chicago, Chicago, Illinois, 2012. 6, 8, 20, 21, 50, 198

[GS19] Nikhil Gupta and Chandan Saha. On the symmetries of and equivalence test for design polynomials. In Peter Rossmanith, Pinar Heggernes, and Joost-Pieter Katoen, editors, *44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, August 26-30, 2019, Aachen, Germany*, volume 138 of *LIPIcs*, pages 53:1–53:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. 18, 72

[GST20] Nikhil Gupta, Chandan Saha, and Bhargav Thankey. A super-quadratic lower bound for depth four arithmetic circuits. In *Proceedings of the 35th Computational Complexity Conference*, CCC '20, Dagstuhl, DEU, 2020. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. 19, 231

[GST22] Nikhil Gupta, Chandan Saha, and Bhargav Thankey. Equivalence test for read-once arithmetic formulas. *Electronic Colloquium on Computational Complexity (ECCC)*, 2022. https://eccc.weizmann.ac.il/report/2022/099/. 11, 17, 18, 29, 38, 126, 203, 204

[GT20] Rohit Gurjar and Thomas Thierauf. Linear matroid intersection is in quasi-nc. *Comput. Complex.*, 29(2):9, 2020. 9

[Gup14] Ankit Gupta. Algebraic geometric techniques for depth-4 pit & sylvester-gallai conjectures for varieties. *Electron. Colloquium Comput. Complex.*, 21:130, 2014. 233

[Gup17] Nikhil Gupta. Towards a characterization of the symmetries of the nisan-wigderson polynomial family. Master's thesis, Indian Institute of Science, 2017. iv, 21, 23, 24, 31, 33, 73, 75, 76, 83, 84, 198

[Gut16] Larry Guth. *Polynomial Methods in Combinatorics*. University Lecture Series. American Mathematical Society, 2016. 2

[GV88] D. Yu. Grigor'ev and N. N. Vorobjov. Solving systems of polynomial inequalities in subexponential time. *J. Symb. Comput.*, 5(1–2):37–64, feb 1988. 13

[Hal03] Brian C Hall. *Lie Groups, Lie Algebras and Representations: An Elementary introduction*. Graduate Texts in Mathematics. Springer, 2003. 51

[Har70] R. A. Harshman. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis. *UCLA Working Papers in Phonetics*, 16:1–84, 1970. 11

[Hås90] Johan Håstad. Tensor Rank is NP-Complete. *J. Algorithms*, 11(4):644–654, 1990. 235

[HH91] Thomas R. Hancock and Lisa Hellerstein. Learning read-once formulas over fields and extended bases. In Manfred K. Warmuth and Leslie G. Valiant, editors, *Proceedings of the Fourth Annual Workshop on Computational Learning Theory, COLT 1991, Santa Cruz, California, USA, August 5-7, 1991*, pages 326–336. Morgan Kaufmann, 1991. 27, 29

[HS80] J. Heintz and C. P. Schnorr. Testing polynomials which are easy to compute (extended abstract). In *Proceedings of the Twelfth Annual ACM Symposium on Theory*

of Computing, STOC '80, page 262–272, New York, NY, USA, 1980. Association for Computing Machinery. 233

[Hüt16] Jesko Hüttenhain. The Stabilizer of Elementary Symmetric Polynomials. *CoRR*, abs/1607.08419, 2016. 21

[HvdH21] David Harvey and Joris van der Hoeven. Integer multiplication in time $O(n\log n)$. *Annals of Mathematics*, 193(2):563 – 617, 2021. 3

[HW99] Ming-Deh Huang and Yiu-Chung Wong. Solvability of systems of polynomial congruences modulo a large prime. *Comput. Complex.*, 8(3):227–257, dec 1999. 13

[HY11] Pavel Hrubeš and Amir Yehudayoff. Arithmetic complexity in ring extensions. *Theory of Computing*, 7:119–129, 01 2011. 4, 5

[Ier89] D. Ierardi. Quantifier elimination in the theory of an algebraically-closed field. In *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing*, page 138–147, New York, NY, USA, 1989. Association for Computing Machinery. 13

[IQ19] Gábor Ivanyos and Youming Qiao. Algorithms Based on *-Algebras, and Their Applications to Isomorphism of Polynomials with One Secret, Group Isomorphism, and Polynomial Identity Testing. *SIAM J. Comput.*, 48(3):926–963, 2019. Conference version appeared in the proceedings of SODA 2018. 14

[IRS12] Gábor Ivanyos, Lajos Rónyai, and Josef Schicho. Splitting full matrix algebras over algebraic number fields. *Journal of Algebra*, 354:211–223, 2012. 71

[Jan08] Milan Janjić. A proof of generalized laplace's expansion theorem. *Bull. Soc. Math. Banja Luka*, 2008. 151

[Joh86] David S. Johnson. The np-completeness column: An ongoing guide. *J. Algorithms*, 7(2):289–305, 1986. 200

[Kal85] K. Kalorkoti. A Lower Bound for the Formula Size of Rational Functions. *SIAM J. Comput.*, 14(3):678–687, 1985. 24, 230

[Kal87] E. Kaltofen. Single-factor hensel lifting and its application to the straight-line complexity of certain polynomials. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 443–452, New York, NY, USA, 1987. Association for Computing Machinery. 3

[Kal89] Erich Kaltofen. Factorization of polynomials given by straight-line programs. In *Randomness and Computation*, pages 375–412. JAI Press, 1989. 3

[Kay11] Neeraj Kayal. Efficient algorithms for some special cases of the polynomial equivalence problem. In Dana Randall, editor, *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 1409–1421. SIAM, 2011. 11, 15, 18, 36, 53, 62, 63

[Kay12] Neeraj Kayal. Affine projections of polynomials: extended abstract. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 643–662, 2012. iv, 6, 13, 17, 23, 24, 52, 61, 199, 201

[KI03] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '03, page 355–364, New York, NY, USA, 2003. Association for Computing Machinery. 233

[KL89] Dexter Kozen and Susan Landau. Polynomial decomposition algorithms. *Journal of Symbolic Computation*, 7(5):445–456, 1989. 13

[KLSS14] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. Super-polynomial lower bounds for depth-4 homogeneous arithmetic formulas. STOC '14, page 119–127, New York, NY, USA, 2014. Association for Computing Machinery. 231

[KLSS17] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas. *SIAM J. Comput.*, 46(1):307–335, 2017. Conference version appeared in the proceedings of FOCS 2014. 19, 231

[KLZ96] Dexter Kozen, Susan Landau, and Richard Zippel. Decomposition of algebraic functions. *J. Symb. Comput.*, 22:235–246, 01 1996. 13

[KMSV13] Zohar S. Karnin, Partha Mukhopadhyay, Amir Shpilka, and Ilya Volkovich. Deterministic identity testing of depth-4 multilinear circuits with bounded top fan-in. *SIAM Journal on Computing*, 42(6):2114–2131, 2013. 233

[KNS19] Neeraj Kayal, Vineet Nair, and Chandan Saha. Average-case linear matrix factorization and reconstruction of low width algebraic branching programs. *Comput. Complex.*, 28(4):749–828, 2019. iv, 17, 24, 201, 236

[KNST19] Neeraj Kayal, Vineet Nair, Chandan Saha, and Sébastien Tavenas. Reconstruction of full rank algebraic branching programs. *ACM Trans. Comput. Theory*, 11(1):2:1–2:56, 2019. Conference version appeared in the proceedings of CCC 2017. 12, 17, 20, 24, 53, 60, 61, 62, 63, 66

[Koe21] W. Koepf. *Computer Algebra: An Algorithm-Oriented Introduction*. Springer Undergraduate Texts in Mathematics and Technology. Springer International Publishing, 2021. 3

[Koi12] Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theor. Comput. Sci.*, 448:56–65, 2012. 231

[KS01] Adam R. Klivans and Daniel A. Spielman. Randomness efficient identity testing of multivariate polynomials. In *Proceedings on 33rd Annual ACM Symposium on Theory of Computing, July 6-8, 2001, Heraklion, Crete, Greece*, pages 216–223, 2001. 233, 234

[KS06] Adam R. Klivans and Amir Shpilka. Learning Restricted Models of Arithmetic Circuits. *Theory of Computing*, 2(10):185–206, 2006. 235

[KS07a] Zohar Karnin and Amir Shpilka. Black box polynomial identity testing of generalized depth-3 arithmetic circuits with bounded top fan-in. *Combinatorica*, 31:333–364, 01 2007. 233

[KS07b] Zohar Karnin and Amir Shpilka. Black box polynomial identity testing of generalized depth-3 arithmetic circuits with bounded top fan-in. *Combinatorica*, 31:333–364, 01 2007. 235

[KS07c] Neeraj Kayal and Nitin Saxena. Polynomial identity testing for depth 3 circuits. *Computational Complexity*, 16:115–138, 01 2007. 233

[KS09] Neeraj Kayal and Shubhangi Saraf. Blackbox polynomial identity testing for depth 3 circuits. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 198–207, 2009. 233

[KS12] Neeraj Kayal and Chandan Saha. On the sum of square roots of polynomials and related problems. *ACM Trans. Comput. Theory*, 4(4), nov 2012. 3

[KS14a] Mrinal Kumar and Shubhangi Saraf. The limits of depth reduction for arithmetic formulas: it's all about the top fan-in. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 136–145, 2014. 19, 231

[KS14b] Mrinal Kumar and Shubhangi Saraf. Superpolynomial lower bounds for general homogeneous depth 4 arithmetic circuits. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part I*, pages 751–762, 2014. 231

[KS16a] Neeraj Kayal and Chandan Saha. Lower Bounds for Depth-Three Arithmetic Circuits with small bottom fanin. *Computational Complexity*, 25(2):419–454, 2016. Conference version appeared in the proceedings of CCC 2015. 19

[KS16b] Mrinal Kumar and Shubhangi Saraf. Arithmetic circuits with locally low algebraic rank. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 34:1–34:27, 2016. 19, 233

[KS16c] Mrinal Kumar and Shubhangi Saraf. Sums of products of polynomials in few variables: Lower bounds and polynomial identity testing. In *Proceedings of the 31st Conference on Computational Complexity*, CCC '16, Dagstuhl, DEU, 2016. 233

[KS17a] Mrinal Kumar and Ramprasad Saptharishi. An exponential lower bound for homogeneous depth-5 circuits over finite fields. In *32nd Computational Complexity Conference, CCC 2017, July 6-9, 2017, Riga, Latvia*, pages 31:1–31:30, 2017. 19, 232

[KS17b] Mrinal Kumar and Shubhangi Saraf. On the Power of Homogeneous Depth 4 Arithmetic Circuits. *SIAM J. Comput.*, 46(1):336–387, 2017. Conference version appeared in the proceedings of FOCS 2014. 231

[KS19] Neeraj Kayal and Chandan Saha. Reconstruction of non-degenerate homogeneous depth three circuits. In Moses Charikar and Edith Cohen, editors, *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 413–424. ACM, 2019. 236

[KS21a] Pascal Koiran and Subhayan Saha. Black Box Absolute Reconstruction for Sums of Powers of Linear Forms. *CoRR*, abs/2110.05305, 2021. 18

[KS21b] Pascal Koiran and Mateusz Skomra. Derandomization and absolute reconstruction for sums of powers of linear forms. *Theor. Comput. Sci.*, 887:63–84, 2021. 18

[KS22] Deepanshu Kush and Shubhangi Saraf. Improved low-depth set-multilinear circuit lower bounds. 2022. CCC '22. 19

[KSS14] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 146–153, 2014. 18, 19, 231

[KST16] Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. An Almost Cubic Lower Bound for Depth Three Arithmetic Circuits. In *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, pages 33:1–33:15, 2016. 19, 231

[KT90] Erich Kaltofen and Barry M. Trager. Computing with Polynomials Given By Black Boxes for Their Evaluations: Greatest Common Divisors, Factorization, Separation of Numerators and Denominators. *J. Symb. Comput.*, 9(3):301–320, 1990. Conference version appeared in the proceedings of FOCS 1988. 3, 29, 36, 37, 39, 40, 61, 203, 235

[Kum17] Mrinal Kumar. A Quadratic Lower Bound for Homogeneous Algebraic Branching Programs. In *Proceedings of the 32nd Computational Complexity Conference*, CCC '17, pages 19:1–19:16, 2017. 230

[KUW85] Richard Karp, Eli Upfal, and Avi Wigderson. Constructing a perfect matching is in random nc. volume 6, pages 22–32, 12 1985. 9

[Lam04] T. Y. Lam. *Introduction To Quadratic Forms Over Fields*. American Mathematical Society, 2004. 70

[LG12] François Le Gall. Faster algorithms for rectangular matrix multiplication. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 514–523, 2012. 2

[LG14] François Le Gall. Powers of tensors and fast matrix multiplication. In *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, ISSAC '14, page 296–303, New York, NY, USA, 2014. Association for Computing Machinery. 2

[Lip89] Richard J. Lipton. New directions in testing. In *Distributed Computing And Cryptography, Proceedings of a DIMACS Workshop, Princeton, New Jersey, USA, October 4-6, 1989*, pages 191–202, 1989. 8, 21, 199

[LLL82a] Arjen Lenstra, H. Lenstra, and Lovász László. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261, 12 1982. 3

[LLL82b] Arjen K Lenstra, Hendrik W Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4):515–534, 1982. 61, 99

[Lor08] Falko Lorenz. *Algebra Volumne 2: Fields with structures, Algebras and advanced topics*. Springer, 2008. 48

[LRA93] S. E. Leurgans, R. T. Ross, and R. B. Abel. A decomposition for three-way arrays. *SIAM Journal on Matrix Analysis and Applications*, 14(4):1064–1083, 1993. 11

[LST21] N. Limaye, S. Srinivasan, and S. Tavenas. Superpolynomial lower bounds against low-depth algebraic circuits. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 804–814, Los Alamitos, CA, USA, feb 2021. IEEE Computer Society. 231, 232, 233, 234

[LV03] Richard J. Lipton and Nisheeth K. Vishnoi. Deterministic identity testing for multivariate polynomials. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, January 12-14, 2003, Baltimore, Maryland, USA*, pages 756–760. ACM/SIAM, 2003. 233

[Lá79] Lovász László. On determinants, matchings and random algorithms. volume 79, pages 565–574, 01 1979. 9

[MM62] Marvin Marcus and Francis May. The permanent function. *Canadian Journal of Mathematics*, 14:177–189, 1962. 20, 198

[MNS20] Janaky Murthy, Vineet Nair, and Chandan Saha. Randomized Polynomial-Time Equivalence Between Determinant and Trace-IMM Equivalence Tests. In Javier Esparza and Daniel Král', editors, *45th International Symposium on Mathematical*

Foundations of Computer Science, MFCS 2020, August 24-28, 2020, Prague, Czech Republic, volume 170 of *LIPIcs*, pages 72:1–72:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. 17, 26, 202

[MS01] Ketan Mulmuley and Milind A. Sohoni. Geometric complexity theory I: an approach to the P vs. NP and related problems. *SIAM J. Comput.*, 31(2):496–526, 2001. 5

[MS21] Dori Medini and Amir Shpilka. Hitting sets and reconstruction for dense orbits in vp_{e} and ΣΠΣ circuits. In *36th Computational Complexity Conference, CCC 2021, July 20-23, 2021, Toronto, Ontario, Canada (Virtual Conference)*, volume 200 of *LIPIcs*, pages 19:1–19:27. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. v, 10, 16, 18, 27, 28, 203, 205

[Mul07] Ketan Mulmuley. On P vs. NP, Geometric Complexity Theory, and the Flip I: a high level view. *CoRR*, abs/0709.0748, 2007. 22

[Mul10] Ketan Mulmuley. Explicit proofs and the flip. *CoRR*, abs/1009.0246, 2010. 8, 21, 22, 199

[Mul11a] Ketan Mulmuley. Geometric complexity theory vi : The flip via positivity. 2011. http://ramakrishnadas.cs.uchicago.edu/gct6.pdf. 8

[Mul11b] Ketan Mulmuley. On P vs. NP and geometric complexity theory: Dedicated to Sri Ramakrishna. *J. ACM*, 58(2):5:1–5:26, 2011. 22, 199

[MV97] Meena Mahajan and V. Vinay. A combinatorial algorithm for the determinant. In *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '97, page 730–738, USA, 1997. Society for Industrial and Applied Mathematics. 24

[MV18] Daniel Minahan and Ilya Volkovich. Complete derandomization of identity testing and reconstruction of read-once formulas. *ACM Trans. Comput. Theory*, 10(3):10:1–10:11, 2018. Conference version appeared in the proceedings of CCC 2017. 27, 29, 204, 234, 235

[MVV87] Ketan Mulmuley, Umesh V. Vazirani, and Vijay V. Vazirani. Matching is as easy as matrix inversion. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 345–354, New York, NY, USA, 1987. Association for Computing Machinery. 9

[Nai19] Vineet Nair. *On Learning and Lower Bound Problems Related to the Iterated Matrix Multiplication Polynomial*. PhD thesis, Indian Institute of Science, Bangalore, 2019. 34, 95

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.*, 49(2):149–167, 1994. 19

[NW97] Noam Nisan and Avi Wigderson. Lower Bounds on Arithmetic Circuits Via Partial Derivatives. *Computational Complexity*, 6(3):217–234, 1997. Conference version appeared in the proceedings of FOCS 1995. 231

[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014. 2

[Pat96] Jacques Patarin. Hidden fields equations (HFE) and isomorphisms of polynomials (IP): two new families of asymmetric algorithms. In *Advances in Cryptology - EUROCRYPT '96, International Conference on the Theory and Application of Cryptographic Techniques, Saragossa, Spain, May 12-16, 1996, Proceeding*, pages 33–48, 1996. 14

[PG97] Jacques Patarin and Louis Goubin. Asymmetric cryptography with s-boxes. In *Proceedings of the First International Conference on Information and Communication Security*, ICICS '97, page 369–380, Berlin, Heidelberg, 1997. Springer-Verlag. 12

[Pip22] Nicholas Pippenger. A formula for the determinant, 2022. https://arxiv.org/abs/2206.00134. 2

[Pot19] Aditya Potukuchi. On the ac^0[oplus] complexity of andreev's problem. In Arkadev Chattopadhyay and Paul Gastin, editors, *39th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2019, December 11-13, 2019, Bombay, India*, volume 150 of *LIPIcs*, pages 25:1–25:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. 200

[Rab80] Michael O Rabin. Probabilistic algorithm for testing primality. *Journal of Number Theory*, 12(1):128–138, 1980. 3

[Raz06] Ran Raz. Separation of Multilinear Circuit and Formula Size. *Theory of Computing*, 2(6):121–135, 2006. Conference version appeared in the proceedings of FOCS 2004. 232

[Raz09] Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *J. ACM*, 56(2), apr 2009. 24, 232

[Raz10] Ran Raz. Elusive Functions and Lower Bounds for Arithmetic Circuits. *Theory of Computing*, 6(1):135–177, 2010. Conference version appeared in the proceedings of STOC 2008. 19, 232

[Raz13] Ran Raz. Tensor-Rank and Lower Bounds for Arithmetic Formulas. *J. ACM*, 60(6):40:1–40:15, 2013. Conference version appeared in the proceedings of STOC 2010. 232

[Ron87] L. Ronyai. Simple algebras are difficult. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 398–408, New York, NY, USA, 1987. Association for Computing Machinery. 35, 116, 117

[Rón90] Lajos Rónyai. Computing the Structure of Finite Algebras. *J. Symb. Comput.*, 9(3):355–373, 1990. 71

[RR97] Alexander A Razborov and Steven Rudich. Natural proofs. *Journal of Computer and System Sciences*, 55(1):24–35, 1997. 20

[RY09] Ran Raz and Amir Yehudayoff. Lower Bounds and Separations for Constant Depth Multilinear Circuits. *Computational Complexity*, 18(2):171–207, 2009. Conference version appeared in the proceedings of CCC 2008. 24, 232

[Sap15] Ramprasad Saptharishi. A selection of known lower bounds in arithmetic circuits, 2015. Github survey. 232

[Sax06] Nitin Saxena. *Morphisms of rings and applications to complexity.* PhD thesis, Indian Institute of Technology, Kanpur, 2006. 13, 14

[Sax08] Nitin Saxena. Diagonal circuit identity testing and lower bounds. In *Proceedings of the 35th International Colloquium on Automata, Languages and Programming - Volume Part I*, ICALP '08, page 60–71, Berlin, Heidelberg, 2008. Springer-Verlag. 233

[Sch80] Jacob T. Schwartz. Fast Probabilistic Algorithms for Verification of Polynomial Identities. *J. ACM*, 27(4):701–717, 1980. 9, 60, 81

[Ser73] Jean-Pierre Serre. *A course in arithmetic.* Springer, 1973. 14, 27, 70

[Sha92] Adi Shamir. Ip = pspace. *J. ACM*, 39(4):869–877, oct 1992. 2, 3, 9

[Shi16] Yaroslav Shitov. How hard is the tensor rank? *arXiv*, abs/1611.01559, 2016. 235

[Sho05] Victor Shoup. *A Computational Introduction to Number Theory and Algebra*. Cambridge University Press, USA, 2005. 3

[Shp09] Amir Shpilka. Interpolation of depth-3 arithmetic circuits with two multiplication gates. *SIAM J. Comput.*, 38(6):2130–2161, 2009. 235

[Sin16] Gaurav Sinha. Reconstruction of real depth-3 circuits with top fan-in 2. In *Proceedings of the 31st Conference on Computational Complexity*, CCC '16, Dagstuhl, DEU, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. 235

[Sin22] Gaurav Sinha. Efficient reconstruction of depth three arithmetic circuits with top fan-in two. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference, ITCS 2022, January 31 - February 3, 2022, Berkeley, CA, USA*, volume 215 of *LIPIcs*, pages 118:1–118:33. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. 235

[SS71] Arnold Schönhage and Volker Strassen. Schnelle multiplikation großer zahlen. *Computing*, 7:281–292, 1971. 2, 3

[SS77] R. Solovay and V. Strassen. A fast monte-carlo test for primality. *SIAM J. Comput.*, 6(1):84–85, mar 1977. 3

[SS97] Victor Shoup and Roman Smolensky. Lower Bounds for Polynomial Evaluation and Interpolation Problems. *Computational Complexity*, 6(4):301–311, 1997. Conference version appeared in the proceedings of FOCS 1991. 232

[SS11] Nitin Saxena and C. Seshadhri. An almost optimal rank bound for depth-3 identities. *SIAM Journal on Computing*, 40(1):200–224, 2011. 233

[SS12] Nitin Saxena and C. Seshadhri. Blackbox identity testing for bounded top-fanin depth-3 circuits: The field doesn't matter. *SIAM Journal on Computing*, 41(5):1285–1298, 2012. 233

[SS13] Nitin Saxena and C. Seshadhri. From sylvester-gallai configurations to rank bounds: Improved blackbox identity test for depth-3 circuits. *J. ACM*, 60(5), oct 2013. 233

[ST17] O. Svensson and J. Tarnawski. The matching problem in general graphs is in quasi-nc. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 696–707, Los Alamitos, CA, USA, October 2017. IEEE Computer Society. 9

[ST21] Chandan Saha and Bhargav Thankey. Hitting sets for orbits of circuit classes and polynomial families. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2021, August 16-18, 2021, University of Washington, Seattle, Washington, USA (Virtual Conference)*, volume 207 of *LIPIcs*, pages 50:1–50:26. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. 6, 9, 10, 16, 27, 205

[Str69] Volker Strassen. Gaussian elimination is not optimal. *Numer. Math.*, 13(4):354–356, aug 1969. 2

[Str73a] Volker Strassen. Die berechnungskomplexiät von elementarysymmetrischen funktionen und von iterpolationskoeffizienten. *Numerische Mathematik*, 20:238–251, 1973. 7

[Str73b] Volker Strassen. Die berechnungskomplexität von elementarsymmetrischen funktionen und von interpolationskoeffizienten. *Numerische Mathematik*, 20:238–251, 1973. 230

[Str73c] Volker Strassen. Vermeidung von divisionen. *Journal für die reine und angewandte Mathematik*, 264:184–202, 1973. 4

[SV10] Amir Shpilka and Ilya Volkovich. On the relation between polynomial identity testing and finding variable disjoint factors. In Samson Abramsky, Cyril Gavoille, Claude Kirchner, Friedhelm Meyer auf der Heide, and Paul G. Spirakis, editors, *Automata, Languages and Programming, 37th International Colloquium, ICALP 2010, Bordeaux, France, July 6-10, 2010, Proceedings, Part I*, volume 6198 of *Lecture Notes in Computer Science*, pages 408–419. Springer, 2010. 235

[SV14] Amir Shpilka and Ilya Volkovich. On Reconstruction and Testing of Read-Once Formulas. *Theory of Computing*, 10(18):465–514, 2014. Conference version appeared in the proceedings of STOC 2008. 27, 204, 235

[SV15] Amir Shpilka and Ilya Volkovich. Read-once polynomial identity testing. *Comput. Complex.*, 24(3):477–532, 2015. Conference versions appeared in the proceedings of STOC 2008 and APPROX-RANDOM 2009. 27

[SV17] Shubhangi Saraf and Ilya Volkovich. Black-box identity testing of depth-4 multilinear circuits. *Combinatorica*, 38, 12 2017. 233

[SW01] Amir Shpilka and Avi Wigderson. Depth-3 arithmetic circuits over fields of characteristic zero. *Computational Complexity*, 10(1):1–27, 2001. Conference version appeared in the proceedings of CCC 1999. 24, 231

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic Circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, 2010. 232, 234

[Tao13] Terence Tao. Algebraic combinatorial geometry: the polynomial method in arithmetic combinatorics, incidence combinatorics, and number theory. *EMS Surveys in Mathematical Sciences*, 1, 10 2013. 2

[Tav15] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. *Inf. Comput.*, 240:2–11, 2015. Conference version appeared in the proceedings of MFCS 2013. 231

[Thi98] Thomas Thierauf. The isomorphism problem for read-once branching programs and arithmetic circuits. *Chicago J. Theor. Comput. Sci.*, 1998, 1998. 13

[Val79] Leslie G. Valiant. Completeness Classes in Algebra. In *Proceedings of the 11h Annual ACM Symposium on Theory of Computing, April 30 - May 2, 1979, Atlanta, Georgia, USA*, pages 249–261, 1979. 4, 5

[Val82] L.G. Valiant. Reducibility by algebraic projections. *de L'Enseignement Mathematique: Logic and Algorithmic*, pages 365 – 380, 1982. 4

[Vol16] Ilya Volkovich. A Guide to Learning Arithmetic Circuits. In *Proceedings of the 29th Conference on Learning Theory, COLT 2016, New York, USA, June 23-26, 2016*, pages 1540–1561, 2016. 9, 234

[Vol17] Ilya Volkovich. On some computations on sparse polynomials. In Klaus Jansen, José D. P. Rolim, David Williamson, and Santosh S. Vempala, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Tech-*

niques, APPROX/RANDOM 2017, August 16-18, 2017, Berkeley, CA, USA, volume 81 of *LIPIcs*, pages 48:1–48:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. 235

[von90] Joachim von zur Gathen. Functional decomposition of polynomials: The wild case. *Journal of Symbolic Computation*, 10(5):437–452, 1990. 13

[VSBR83] Leslie G. Valiant, Sven Skyum, S. Berkowitz, and Charles Rackoff. Fast Parallel Computation of Polynomials Using Few Processors. *SIAM J. Comput.*, 12(4):641–644, 1983. 231

[vzG90] Joachim von zur Gathen. Functional decomposition ofpolynomials: the tame case. *Journal of Symbolic Computation*, 9(3):281–299, 1990. Computational algebraic complexity editorial. 13

[Wal13] Lars Ambrosius Wallenborn. Computing the hilbert symbol, quadratic form equivalence and integer factoring. Diploma thesis, Rheinischen Friedrich-Wilhelms-Universität Bonn, 2013. 71

[Wil12] Virginia Vassilevska Williams. Multiplying matrices faster than coppersmith-winograd. STOC '12, page 887–898, New York, NY, USA, 2012. Association for Computing Machinery. 2

[Wit37] Ernst Witt. Theorie der quadratischen Formen in beliebigen Körpern. *J. Reine Angew. Math.*, 176:31–44, 1937. 16, 27

[Yau16] Morris Yau. Almost cubic bound for depth three circuits in VP. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:187, 2016. 231

[Ye94] Yinyu Ye. Combining binary search and newton's method to compute real roots for a class of real functions. *J. Complexity*, 10(3):271–280, 1994. 93

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation, EUROSAM '79, An International Symposiumon Symbolic and Algebraic Computation, Marseille, France, June 1979, Proceedings*, pages 216–226, 1979. 9, 60, 81

Appendix A

A survey of results on lower bounds, PIT and reconstruction

In this appendix, we present a brief survey of the progress made in the three most important problems in algebraic complexity theory, namely lower bounds, polynomial identity testing (PIT) and reconstruction of arithmetic circuits. We have given an introduction to these three problems along with their connections with equivalence testing problem in Sections 1.1.3 - 1.1.5.

A.1 Lower bounds

Lower bounds for general circuits, formulas, ABPs. The best known lower bound for general arithmetic circuits is super-linear. [Str73b, BS83] showed that any arithmetic circuit computing the d-th power symmetric polynomial or the d-th elementary symmetric polynomial in n variables requires size $\Omega(n \log d)$. The situation is slightly better for arithmetic formulas, for which a quadratic lower bound is known [Kal85, CKSV20]. A quadratic lower bound on the size of a 'homogeneous' algebraic branching program (ABP) (see Definition 2.36) computing the n-th power symmetric polynomial in n variables was given in [Kum17]. Later, [CKSV20] showed that any 'layered' ABP computing the same polynomial should have size $\Omega(n^2)$.

As proving good lower bounds for general arithmetic circuits, formulas, or ABPs seem very difficult, it is natural to focus on the restricted classes of arithmetic circuits. One can hope that proving strong lower bounds for such restricted classes might also give us a handle on showing good lower bounds for general arithmetic circuits. One such natural restricted class is the class of low-depth circuits. It is known because of the *depth reduction* results in ACT that to prove super-polynomial lower bounds on the size of general arithmetic circuits, it is sufficient to prove 'strong enough' lower bounds on the size of low-depth arithmetic circuits.

Depth reduction results. [VSBR83] showed that if an n-variate degree d polynomial f is computed by an arithmetic ciruit of size s then f can also be computed by another arithmetic circuit of size $\text{poly}(s,d)$ and depth $O(\log d(\log s + \log d))$. Building on this, Agrawal and Vinay showed in [AV08] that if a degree d polynomial f is computed by an arithmetic circuit of size $2^{o(d+d\log\frac{n}{d})}$ then it can also be computed by a depth 4 circuit (i.e., $\Sigma\Pi\Sigma\Pi$ circuit) of size $2^{o(d+d\log\frac{n}{d})}$. This result was further refined in [Koi12, Tav15] and Tavenas showed that if f having degree $d = \text{poly}(n)$ is computed by an arithmetic circuit of size s then it can also be computed by a depth 4 circuit of size $2^{O(\sqrt{d\log(sd)\log n})}$, where the fan-in of every $\times$ gate in the resulting depth 4 circuit is upper bounded by $O(\sqrt{d})$. These results hold over any field and if f is homogeneous then the resulting depth 4 circuit is also homogeneous [1]. Further, [GKKS16, Tav15] showed that if a degree $d = \text{poly}(n)$ polynomial f is computed by a size-s arithmetic circuit, it can also be computed by a depth 3 circuit (i.e., $\Sigma\Pi\Sigma$ circuit) of size $2^{O(\sqrt{d\log s\log n})}$. This result holds only over the fields of characteristic zero and if f is homogeneous then the resulting depth 3 circuit need not be homogeneous.

These results say that to show that the permanent of an $n\times n$ symbolic matrix is not computed by arithmetic circuits of size $\text{poly}(n)$, it is sufficient to show that every homogeneous $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[O(\sqrt{n})]}$ circuit (i.e., a depth 4 circuit where the fan-in of every multiplication gate is $O(\sqrt{n})$) or a depth 3 circuit (over a field of characteristic zero) for the permanent needs size $2^{\omega(\sqrt{n}\log n)}$. This naturally motivates the study of lower bounds for homogeneous depth 4 and (non-homogeneous) depth 3 arithmetic circuits.

Lower bounds for constant-depth circuits before [LST21]. Depth 2 circuit (i.e., $\Sigma\Pi$ circuit) is the simplest low-depth arithmetic circuit. Any polynomial containing exponentially many monomials requires $\Sigma\Pi$ circuits of exponential size. Thus, the smallest interesting class from the viewpoint of lower bounds is that of $\Sigma\Pi\Sigma$ circuits. Nisan-Wigderson gave an exponential lower bound on the size of homogeneous $\Sigma\Pi\Sigma$ circuit over any field in [NW97]. Thereafter, an exponential lower bound for $\Sigma\Pi\Sigma$ circuits over finite fields was given in [GK98, GR00]. But, before [LST21], the best known lower bound for general $\Sigma\Pi\Sigma$ circuits over fields of characteristic zero was almost cubic [KST16, BLS16, Yau16], which improved a quadratic lower bound given in [SW01]. The situation was similar for the depth 4 circuits. A long line of research converged to a $2^{\Omega(\sqrt{n}\log n)}$ lower bound for homogeneous depth 4 circuits over any field [GKKS14b, KSS14, FLMS15, KS14a, KLSS14, KS14b, KLSS17, KS17b]. Before [LST21], the best known lower bound on the size of a general depth 4 circuit was super-quadratic [GST20].

[1] An arithmetic circuit C is said to be homogeneous is every node in C computes a homogeneous polynomial.

An exponential lower bound for homogeneous depth 5 circuits over small finite fields was given in [KS17a]. Before [LST21], the best known lower bound for arithmetic circuits having product-depth[1] $\Delta = O(\log n)$, was $O(\Delta n^{1+\frac{1}{\Delta}})$ [SS97, Raz10].

A breakthrough on constant-depth arithmetic circuits by [LST21]. Last year, Limaye, Srinivasan and Tavenas gave the first super-polynomial lower bound for unrestricted arithmetic circuits of constant-depth. This was a quantum leap in the status of lower bound for constant-depth arithmetic circuits. In particular, they showed that for $N \in \mathbb{N}$, $d = o(\log N)$, if C is an arithmetic circuit of product-depth Δ which computes $\mathrm{IMM}_{n,d}$, where $N = n^2 d$, then the size of C is greater than $N^{d^{\exp(-O(\Delta))}}$. Their result holds over any field having characteristic equal to 0 or greater than d.

Lower bounds for mutlilinear circuits. An arithmetic circuit C is said to be multilinear if every gate in C computes a multilinear polynomial[2]. Raz showed in [Raz13] that for $d = O(\frac{\log n}{\log\log n})$, if an n-variate degree d polynomial f is computed by a polynomial size arithmetic formula then f can also be computed by a polynomial size *set-multilinear circuit*. Thus, a super-polynomial lower bound on the size of set-multilinear circuits computing a low-degree polynomial implies a super-polynomial lower bound on the size of arithmetic formulas.

In [Raz09], Raz gave a super-polynomial lower bound for arithmetic formulas computing the determinant or the permanent. Thereafter, he showed in [Raz06] that there exists a multilinear polynomial f, which can be computed by a multilinear circuit of size $\mathrm{poly}(n)$ but every multilinear formula computing f requires size $n^{\Omega(\log n)}$, which implies a super-polynomial separation between multilinear circuits and multilinear formulas. A super-polynomial separation between multilinear formulas and *multilinear branching programs* was given in [DMPY12]. Raz and Yehudayoff in [RY09] gave a super-polynomial separation between multilinear circuits of depths Δ and $\Delta + 1$, where Δ is a constant. [CELS18] later improved this result and gave exponential separation.

We direct interested readers to [SY10, CKW11, Sap15] for a detailed exposure to lower bound results in ACT.

[1] The product-depth of an arithmetic circuit C is the maximum number of product gates on any path from an input gate to the output gate of C.

[2] A polynomial is said to be multilinear if the degree of every variable in this polynomial is at most 1.

A.2 Polynomial identity testing

Connection between lower bounds and PIT. It was shown in [KI03] that if a polynomial time algorithm exists for PIT over integers then either NEXP $\not\subseteq$ P/poly or the permanent is not computed by a polynomial size arithmetic circuit. [KI03] also showed a (partial) converse - if there exists an exponential time computable multilinear polynomial, which is not computed by polynomial size arithmetic circuits then PIT has a quasi-polynomial time deterministic algorithm. It was shown in [HS80, Agr05] that if a polynomial time deterministic black-box PIT algorithm exists then there is a polynomial f, whose coefficients can be computed in PSPACE and any arithmetic circuit for f has exponential size. Results analogous to [KI03] relating PIT and lower bounds for bounded depth arithmetic circuits are also known [DSY10, CKS18].

PIT for constant-depth circuits before [LST21]. A tight connection between PIT and lower bounds makes the task of coming up with a deterministic polynomial time PIT algorithm very challenging. Thus, researchers have focused on restricted models of computations like low-depth arithmetic circuits. Polynomial time black-box PIT algorithms are known for depth 2 circuits [BOT88, KS01, LV03]. Before [LST21], designing efficient sub-exponential PIT algorithm even for depth 3 circuits appeared very demanding. It follows from the depth reduction results that polynomial time black-box PIT for depth 4 circuits or depth 3 circuits imply a sub-exponential time deterministic black-box PIT algorithm for general arithmetic circuits (see [AV08, GKKS16]). Thus, PIT for special cases of depth 3 and depth 4 arithmetic circuits was studied. In [Sax08], a polynomial time white-box PIT algorithm for *depth 3 powering circuits* was given. A black-box PIT algorithm for the same model was given by [FSS14], which runs in quasi-polynomial time for arbitrary depth 3 powering circuit but runs in polynomial time if the fan-in of topmost gate (or the top fan-in) of such a circuit is a constant. PIT for depth 3 circuits with bounded top fan-in has also received a lot of attention and after lot of work, a polynomial time black-box PIT algorithm is known when the top fan-in is a constant [DS07, KS07a, KS07c, KS09, AM10, SS11, SS12, SS13]. A polynomial time black-box PIT algorithm was given in [ASSS16] for depth 3 circuits where the *transcendence degree* of the set of polynomials computed by product gates is a constant. PIT for some restricted classes of depth 4 circuits have also been studied. Polynomial time algorithm for black-box PIT is known for multilinear depth 4 circuits with constant top fan-in [AMV15, KMSV13, SV17]. Efficient PIT algorithms are also known for depth 4 circuits with other constraints (see [BMS13, Gup14, KS16c, KS16b]).

A breakthrough on PIT for constant-depth circuits [LST21]. The first sub-exponential time black-box PIT algorithm for constant-depth arithmetic circuits was given in [LST21], which appeared a hard nut to crack before this work. This PIT algorithm was achieved by combining the super-polynomial lower bound for constant-depth circuits given in [LST21] with a result in [CKS18].

PIT for constant-read arithmetic circuits/ABPs. PIT algorithms have also been studied for models of computations, where a variable is read constantly many times. Polynomial time deterministic black-box PIT algorithms are known for *read-once arithmetic formulas* (ROFs) [MV18], for *constant-depth constant-read multilinear arithmetic formulas* [AMV15] and for *constant-depth constant-occur formulas* [ASSS16]. Quasi-polynomial deterministic black-box PIT algorithms for *read-once oblivious algebraic branching programs* (ROABPs) with known variable ordering, for *multilinear unknown variable order ROABPs*, and for *unknown variable order ROABPs* were given in [FS13], [FSS14] and [AGKS14] respectively.

A.3 Arithmetic circuit reconstruction

Relation of circuit reconstruction with lower bounds and PIT. It was shown in [FK09] that a randomized polynomial time reconstruction algorithm for a circuit class $\mathscr{C}$ implies that there exists a polynomial f, whose evaluations on Boolean inputs can be computed in BPEXP but f can not be computed by polynomial size circuits from $\mathscr{C}$. After that, [Vol16] showed that if there exists a deterministic polynomial time reconstruction algorithm for $\mathscr{C}$ then there exists a multilinear polynomial f computable in exponential time such that any circuit from $\mathscr{C}$ that computes f requires exponential size. Observe that a deterministic polynomial time reconstruction algorithm for $\mathscr{C}$ immediately implies a deterministic black-box PIT algorithm for $\mathscr{C}$. In several cases, deterministic black-box PIT algorithms have led to the discovery of efficient reconstruction algorithms (see Chapter 5 of [SY10]). Thus, designing polynomial time reconstruction algorithms (randomized or deterministic) for general arithmetic circuits is a very challenging task. Similar to PIT and lower bounds, the natural next step in this situation is to study reconstruction of restricted classes of arithmetic circuits. We give a brief summary of two types of reconstruction algorithms here: *worst-case algorithms*, which work for all circuits in a circuit class $\mathscr{C}$ and *average-case algorithms*, which work for *almost all* circuits in $\mathscr{C}$.

Worst-case reconstruction algorithms. Deterministic polynomial time algorithms for reconstruction of $\Sigma\Pi$ circuits are known [BOT88, KS01]. It is easy to see from the depth reduction results that polynomial time reconstruction algorithms for depth 3 or depth 4 circuits would im-

mediately imply sub-exponential time reconstruction algorithms for general circuits. Thus, the next step to focus on sub-classes of depth 3 and depth 4 circuits. One such sub-class is $\Sigma\Pi\Sigma(k)$ circuits, where k is the top fan-in of a depth 3 circuit. When $k = 1$, the reconstruction problem is same as black-box polynomial factorisation problem, for which a randomized polynomial time algorithm is known [KT90]. A randomized $\text{poly}(n, |\mathbb{F}|)$ time (respectively, quasi-polynomial $(n, d, |\mathbb{F}|)$ time)[1] algorithm was given by Shpilka [Shp09] for multilinear $\Sigma\Pi\Sigma(2)$ (respectively, $\Sigma\Pi\Sigma(2)$) circuits[2]. This result was later derandomized and extended for constant value of k in [KS07b]. As the running time of these algorithms depend on $|\mathbb{F}|$, these algorithms are efficient only over small finite fields. The first polynomial time randomized algorithm for $\Sigma\Pi\Sigma(2)$ circuits over fields of characteristic zero was given in [Sin16]. Recently, [Sin22] gave a randomized reconstruction algorithm for $\Sigma\Pi\Sigma(2)$ circuits, where the running time is $\text{poly}(n, d, \log |\mathbb{F}|)$.

Apart from $\Sigma\Pi\Sigma(k)$, reconstuction algorithms for other sub-classes of $\Sigma\Pi\Sigma$ circuits have also been studied. [Hås90] and [Shi16] proved that reconstructing a smallest *set-multilinear* depth 3 circuit and a smallest depth 3 powering circuit is NP-hard respectively. Randomized polynomial time reconstruction algorithms for set-multilinear depth 3 circuits are known [BBB+00, KS06], where the outputs are read-once algebraic branching programs (ROABPs) and hence these algorithms are improper. [KS06] also gave a $\text{poly}(m, n, 2^d)$ algorithm to reconstruct a $\Sigma\Pi\Sigma$ circuit having m multiplication gates. Recently, [BSV21] gave polynomial time randomized algorithms for reconstructing set-multilinear $\Sigma\Pi\Sigma(k)$ circuits, depth 3 powering circuits with top fan-in k and multilinear $\Sigma\Pi\Sigma(k)$ circuits. In these algorithms, k is a constant. Their algorithms are deterministic over $\mathbb{R}$ and $\mathbb{C}$. Reconstruction algorithms are also studied for multilinear $\Sigma\Pi\Sigma\Pi(k)$ class, where k is the top fan-in. A polynomial time algorithm for $k = 1$ was given in [SV10]. A deterministic polynomial time reconstruction algorithm for $k = 2$ was given in [Vol17] and a randomized polynomial time reconstuction algorithm for *random* multilinear $\Sigma\Pi\Sigma\Pi(2)$ circuits was given in [GKL11]. Recently, a deterministic quasi-polynomial reconstruction algorithm is given for multilinear $\Sigma\Pi\Sigma\Pi(k)$ is given in [BSV20] when k is a constant. The running time of their algorithm depends on $|\mathbb{F}|$, hence the algorithm is efficient only over small fields. A polynomial time reconstruction algorithm for read-once arithmetic formulas was given in [MV18], which improved the results in [SV14]. A quasi-polynomial time deterministic algorithm for reconstructing ROABPs was given in [FS13], for which a randomized polynomial time algorithm was given by [KS06].

[1] Unless otherwise specified, n and d denote the number of variables and the degree of the underlying polynomial respectively.

[2] The algorithm is proper for multilinear $\Sigma\Pi\Sigma(2)$ circuits but in case of general $\Sigma\Pi\Sigma(2)$ circuits, the algorithm either outputs a $\Sigma\Pi\Sigma(2)$ circuit or a depth 3 circuit of quasi-polynomial size, depending upon the *rank* of the underlying polynomial.

Non-degenerate reconstruction algorithms. As reconstructing an arithmetic circuit in the worst-case is very challenging and as some of its instances are NP-hard, it is natural to consider average-case reconstruction algorithms. An average-case algorithm for a class $\mathscr{C}$ reconstructs circuits from $\mathscr{C}$ satisfying some *non-degeneracy conditions*. These conditions are satisfied with high probability by circuits chosen randomly from $\mathscr{C}$ according to some distribution and thus such algorithms work for *almost all* circuits in $\mathscr{C}$. As seen above, efficient (worst-case) reconstruction algorithms for depth 3 and multilinear depth 4 circuits are known only when the top fan-in is a constant. [KS19] gave a randomized polynomial time algorithm to reconstruct non-degenerate homogeneous depth 3 circuits. Their algorithm handles circuits with large top fain-in. In [BGKS21], a polynomial time randomized algorithm for reconstructing non-degenerate *generalized depth 3 circuits* was given. A randomized polynomial time reconstruction algorithm for non-degenerate *sum of low-degree polynomials*[1] was given in [GKS20]. A randomized algorithm for reconstructing *random* n-variate width w ABPs was given in [KNS19], where $n \geq 4w^2$. [GKQ13] gave a randomized polynomial time algorithm to reconstruct *random* arithmetic formulas in the *alternate normal form* (ANF)[2].

[1]Such polynomials have the following structure: $f = \alpha_1 Q_1^m + \cdots + \alpha_s Q_s^m$, where $\alpha_1, \cdots, \alpha_s \in \mathbb{F}$ and every Q_i is a homogeneous polynomial of degree t.

[2]An arithmetic formula is said to be in the alternating normal form if its underlying tree is a complete binary tree, where the leaves are labelled by affine forms and it has alternate layers of $+$ and $\times$ gate.