Black Box Absolute Reconstruction for Sums of Powers of Linear Forms

Pascal Koiran
Univ Lyon, EnsL, UCBL, CNRS, LIP, F-69342, LYON Cedex 07, France

Subhayan Saha
Univ Lyon, EnsL, UCBL, CNRS, LIP, F-69342, LYON Cedex 07, France

Abstract

We study the decomposition of multivariate polynomials as sums of powers of linear forms. We give a randomized algorithm for the following problem: If a homogeneous polynomial $f \in K[x_1, ..., x_n]$ (where $K \subseteq \mathbb{C}$) of degree d is given as a blackbox, decide whether it can be written as a linear combination of d-th powers of linearly independent complex linear forms. The main novel features of the algorithm are:

- For $d = 3$, we improve by a factor of n on the running time from the algorithm in [21]. The price to be paid for this improvement is that the algorithm now has two-sided error.
- For $d > 3$, we provide the first randomized blackbox algorithm for this problem that runs in time poly(n, d) (in an algebraic model where only arithmetic operations and equality tests are allowed). Previous algorithms for this problem [17] as well as most of the existing reconstruction algorithms for other classes appeal to a polynomial factorization subroutine. This requires extraction of complex polynomial roots at unit cost and in standard models such as the unit-cost RAM or the Turing machine this approach does not yield polynomial time algorithms.
- For $d > 3$, when f has rational coefficients (i.e. $K = \mathbb{Q}$), the running time of the blackbox algorithm is polynomial in n, d and the maximal bit size of any coefficient of f. This yields the first algorithm for this problem over $\mathbb{C}$ with polynomial running time in the bit model of computation.

These results are true even when we replace $\mathbb{C}$ by $\mathbb{R}$. We view the problem as a tensor decomposition problem and use linear algebraic methods such as checking the simultaneous diagonalisability of the slices of a tensor. The number of such slices is exponential in d. But surprisingly, we show that after a random change of variables, computing just 3 special slices is enough. We also show that our approach can be extended to the computation of the actual decomposition. In forthcoming work we plan to extend these results to overcomplete decompositions, i.e., decompositions in more than n powers of linear forms.

2012 ACM Subject Classification Theory of computation → Algebraic complexity theory; Computing methodologies → Algebraic algorithms; Computing methodologies → Linear algebra algorithms

Keywords and phrases reconstruction algorithms, tensor decomposition, sums of powers of linear forms, simultaneous diagonalisation, algebraic algorithm, black box

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2022.24

Related Version *Full Version*: `https://arxiv.org/abs/2110.05305`

1 Introduction

Lower bounds and polynomial identity testing are two fundamental problems about arithmetic circuits. In this paper we consider another fundamental problem: arithmetic circuit reconstruction. For an input polynomial f, typically given by a black box, the goal is to find the smallest circuit computing f within some class $\mathcal{C}$ of arithmetic circuits. This problem can be divided in two subproblems: a decision problem (can f be computed by a circuit of size s from the class $\mathcal{C}$?) and the reconstruction problem proper (the actual construction of the smallest circuit for f). In this paper we are interested in *absolute reconstruction*,

namely, in the case where $\mathcal{C}$ is a class of circuits over the field of complex numbers. The name is borrowed from *absolute factorization*, a well-studied problem in computer algebra (see e.g. [9, 10, 11, 24]). Most of the existing reconstruction algorithms appeal to a polynomial factorization subroutine, see e.g. [12, 13, 16, 17, 18, 19, 26]. This typically yields polynomial time algorithms over finite fields or the field of rational numbers. However, in standard models of computation such as the unit-cost RAM or the Turing machine this approach does *not* yield polynomial time algorithms for absolute reconstruction. This is true even for the decision version of this problem. In the Turing machine model, the difficulty is as follows. We are given an input polynomial f, say with rational coefficients, and want to decide if there is a small circuit $C \in \mathcal{C}$ for f, where C may have complex coefficients. After applying a polynomial factorization subroutine, a reconstruction algorithm will manipulate polynomials with coefficients in a field extension of $\mathbb{Q}$. If this extension is of exponential degree, the remainder of the algorithm will not run in polynomial time. This point is explained in more detail in [21] on the example of a reconstruction algorithm due to Neeraj Kayal [17]. One way out of this difficulty is to work in a model where polynomial roots can be extracted at unit cost, as suggested in a footnote of [14]. We will work instead in more standard models, namely, the Turing machine model or the unit-cost RAM over $\mathbb{C}$ with arithmetic operations only (an appropriate formalization is provided by the Blum-Shub-Smale model of computation [5, 6]). Before presenting our results, we present the class of circuits studied in this paper.

1.1 Sums of powers of linear forms

Let $f(x_1, \ldots, x_n)$ be a homogeneous polynomial of degree d. In this paper we study decompositions of the type:

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{r} l_i(x_1, \ldots, x_n)^d \tag{1}$$

where the l_i are linear forms. Such a decomposition is sometimes called a Waring decomposition, or a symmetric tensor decomposition. The smallest possible value of r is the symmetric tensor rank of f, and it is NP-hard to compute already for $d = 3$ [25]. One can nevertheless obtain polynomial time algorithms by restricting to a constant value of r [3]. In this paper we assume instead that the linear forms l_i are linearly independent (hence $r \leq n$). This setting was already studied by Kayal [17]. It turns out that such a decomposition is unique when it exists, up to a permutation of the l_i and multiplications by d-th roots of unity. This follows for instance from Kruskal's uniqueness theorem. For a more elementary proof, see [17, Corollary 5.1] and [21, Section 3.1].

Under this assumption of linear independence, the case $r = n$ is of particular interest. In this case, f is *equivalent* to the sum of d-th powers polynomial

$$P_d(x) = x_1^d + x_2^d + \cdots + x_n^d \tag{2}$$

in the sense that $f(x) = P_d(Ax)$ where A is invertible. A test of equivalence to P_d was provided in [17]. The resulting algorithm provably runs in polynomial time over the field of rational numbers, but this is not the case over $\mathbb{C}$ due to the appeal to polynomial factorization. The first equivalence test to P_d running in polynomial time over the field of complex numbers was given in [21] for $d = 3$. We will extend this result to arbitrary degree in this paper. In the general case $r \leq n$ we can first compute the number of essential variables of f [8, 17]. Then we can do a change of variables to obtain a polynomial depending only on its first r variables [17, Theorem 4.1], and conclude with a test of equivalence to P_r (see [21, Proposition 44] for details).

Equivalence and reconstruction algorithms over $\mathbb{Q}$ are number-theoretic in nature in the sense that their behavior is highly sensitive to number-theoretic properties of the coefficients of the input polynomial. This point is clearly illustrated by an example from [21]:

▶ **Example 1.** Consider the rational polynomial

$$f(x_1, x_2) = (x_1 + \sqrt{2}x_2)^3 + (x_1 - \sqrt{2}x_2)^3 = 2x_1^3 + 12x_1x_2^2.$$

This polynomial is equivalent to $P_3(x_1, x_2) = x_1^3 + x_2^3$ over $\mathbb{R}$ and $\mathbb{C}$ but not over $\mathbb{Q}$.

By contrast, equivalence and reconstruction algorithms over $\mathbb{R}$ and $\mathbb{C}$ are of a more geometric nature.

1.2 Connection to Tensor Decomposition

Using the relation between tensors and polynomials, we can see that a homogeneous degree-d polynomial $f \in \mathbb{K}[x_1, ..., x_n]$ can be written as a sum of d-th powers of linear forms over $\mathbb{K}$ if and only if there exist $v_i \in \mathbb{K}$ such that the corresponding symmetric tensor T_f can be decomposed as $T_f = \sum_i v_i^{\otimes d}$. This is often referred to as the tensor decomposition problem for the given tensor T.

Most tensor decompositions algorithms are numerical such as the ALS method [22] (which lacks a good complexity analysis), tensor power iteration [1] (for orthogonal tensor decomposition) or Jennrich's algorithm [15, 23] for ordinary tensors. A self-contained bit-complexity analysis of Jennrich's algorithm can be found in [4]. Unlike the above algorithm from [21], these numerical algorithms do not provide any decision procedure. The algebraic algorithm from [21] seems closest in spirit to Jennrich's: they both rely on simultaneous diagonalization and on linear independence assumptions on the vectors involved in the tensor decomposition. Algorithms for symmetric tensor decomposition can be found in the algebraic literature, see e.g. [2, 7]. These two papers do not provide any complexity analysis.

1.3 Results and methods

Our main contributions are as follows. Recall that P_d is the sum of d-th powers polynomials (2), and let us assume that the input $f \in \mathbb{C}[x_1, \ldots, x_n]$ is a homogeneous polynomial of degree d.

- **(i)** For $d = 3$, we improve by a factor of n on the running time of the test of equivalence to P_3 from [21]. The price to be paid for this improvement is that the algorithm now has two-sided error. The algorithm for the $d = 3$ case and a simpler analysis can be found in the full paper [20].
- **(ii)** For $d > 3$, we provide the first blackbox algorithm for equivalence to P_d with running time polynomial in n and d (more specifically, $O(n^2 d)$ calls to the blackbox and $O(n^2 d \log^2(d) \log\log(d) + n^{\omega+1})$ arithmetic operations) where ω is the exponent of matrix multiplication, in an algebraic model where only arithmetic operations and equality tests are allowed (i.e., computation of polynomial roots is *not allowed*).
- **(iii)** For $d > 3$, when f has rational coefficients this blackbox algorithm runs in polynomial time in the bit model of computation. More precisely, the running time is polynomial in n, d and the maximal bit size of any coefficient of f. This yields the first test of equivalence to P_d over $\mathbb{C}$ with polynomial running time in the bit model of computation.

As outlined in Section 1.1, these results have application to decomposition into sums of powers of linearly independent linear forms over $\mathbb{C}$. Namely, we can decide whether the input polynomial admits such a decomposition, and if it does we can compute the number of terms

r in such a decomposition. The resulting algorithm runs in polynomial time in the algebraic model of computation, as in item (ii) above; when the input has rational coefficients it runs in polynomial time in the bit model of computation, as in (iii) (refer to Appendix B in the full paper [20] for a detailed complexity analysis). This is the first algorithm with these properties. It can be viewed as an algebraic, high order, black box version of Jennrich's algorithm.

Using the relation to tensor decomposition problem mentioned in Section 1.2, if an order d-tensor $T \in K^{n\times\cdots\times n}$ is given as a blackbox, we give an algorithm that runs in time $\mathrm{poly}(n,d)$ to check if there exist linearly independent vectors $v_i \in \mathbb{K}^n$ such that $T = \sum_{i=1}^{t} \alpha_i v_i^{\otimes d}$ for some $t \leq n$. Note here that $K \subseteq \mathbb{C}$ and $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. This can be found in more detail in Section 4 of the full paper [20].

Finally, in Section 5 of the full paper [20], we show that our linear algebraic approach can be extended to the computation of the actual decomposition in a model where we allow the computation of polynomial roots. We therefore obtain an alternative to the algorithm from [17] for this problem. That algorithm relies on multivariate polynomial factorization, whereas our algorithm relies on matrix diagonalization.

Real versus complex field. For $\mathbb{K} = \mathbb{R}$ and even degree there is obviously a difference between sums of d-th powers of linear forms and linear combinations of d-th powers. In this paper we wish to allow arbitrary linear combinations. For this reason, in the treatment of the high order case ($d > 3$) we are not interested in equivalence to P_d only. Instead, we would like to know whether the input is equivalent to some polynomial of the form $\sum_{i=1}^{n} \alpha_i x_i^d$ with $\alpha_i \neq 0$ for all i. We denote by $\mathcal{P}_d$ this class of polynomials (one could even assume that $\alpha_i = \pm 1$ for all i). At first reading, there is no harm in assuming that $\mathbb{K} = \mathbb{C}$. In this case, one can assume without loss of generality that $\alpha_i = 1$ for all i. For $\mathbb{K} = \mathbb{R}$, having to deal with the whole of $\mathcal{P}_d$ slightly complicates notations, but the proofs are not significantly more complicated than for $\mathbb{K} = \mathbb{C}$. For this reason, in all of our results we give a unified treatment of the two cases $\mathbb{K} = \mathbb{C}$ and $\mathbb{K} = \mathbb{R}$.

Methods. We associate to a homogeneous polynomial of degree d the (unique) symmetric tensor T of order d such that

$$f(x_1, \ldots, x_n) = \sum_{i_1,\ldots,i_d=1}^{n} T_{i_1\ldots i_d} x_{i_1} x_{i_2} \ldots x_{i_d}.$$

We recall that T is said to be symmetric if it is invariant under all $d!$ permutations of its indices. A *slice* of T (or by abuse of language, a slice of f) is a matrix of size n obtained by fixing the values of $d-2$ indices. In Section 2.1 we give a characterization of equivalence to P_d:

▶ **Theorem 2.** *A degree d homogeneous polynomial $f \in \mathbb{C}[x_1, \ldots, x_n]$ is equivalent to $P_d = \sum_{i=1}^{n} x_i^d$ if and only if its slices span a nonsingular matrix space and the slices are simultaneously diagonalizable by congruence, i.e., there exists an invertible matrix $Q \in M_n(\mathbb{C})$ such that for every slice S of f, the matrix $Q^T S Q$ is diagonal.*

This characterization is satisfactory from a purely structural point of view, but not from an algorithmic point of view because a tensor of order d has $\frac{d(d-1)}{2} n^{d-2}$ slices. The tensors encountered in this paper are all symmetric since they originate from homogeneous polynomials. Taking the symmetry constraints into consideration reduces the number of distinct slices to $\binom{n+d-3}{d-2}$ at most: this is the number of multisets of size $d-2$ in a set of n elements,

or equivalently the number of monomials of degree $d-2$ in the variables $x_1, \ldots, x_n$. This number remains much too large to reach our goal of a complexity polynomial in n and d. This problem has a surprisingly simple solution: our equivalence algorithm (Algorithm 1 below) needs to work with 3 slices only! This is true already for $d = 3$, and is the reason why we can save a factor of n compared to the algorithm of [21]. More detail on the degree 3 case can be found in Section 2 of the full paper [20].

Algorithm 1 Randomized algorithm to check polynomial equivalence to an element of $\mathcal{P}_d$.

Input: A degree-d homogeneous polynomial f
Let $R \in M_n(\mathbb{K})$ be a matrix such that its entries r_{ij} are picked uniformly and independently at random from a finite set S, and set $h(x) = f(Rx)$.
Let $\{T_{i_1 \ldots i_{d-2}}\}_{i_1 \ldots i_{d-2} \in [n]}$ be the slices of h.
We compute the slices $T_{\bar{1}}$,$T_{\bar{2}}$,$T_{\bar{3}}$, where $T_{\bar{i}}$ refers to the slice $T_{i \ldots i}$.
if $T_{\bar{1}}$ *is singular* **then**
  reject
else
  compute $T'_{\bar{1}} = (T_{\bar{1}})^{-1}$
  if $T'_{\bar{1}}T_{\bar{2}}$ *and* $T'_{\bar{1}}T_{\bar{3}}$ *commute and* $T'_{\bar{1}}T_{\bar{2}}$ *is diagonalisable over* $\mathbb{K}$ **then**
    accept
  else
    reject
  end
end

More precisely, the following test forms the heart of the algorithm: check that $T_{\bar{1}}^{-1}T_{\bar{2}}$ is diagonalizable, and commutes with $T_{\bar{1}}^{-1}T_{\bar{3}}$ where $T_{\bar{i}}$ refers to the slice $T_{i \ldots i}$. It may be surprising at first sight that we can work with 3 slices only of a tensor with $\binom{n+d-3}{d-2}$ slices. To give some plausibility to this claim, note that $T_{\bar{1}}, T_{\bar{2}}, T_{\bar{3}}$ are not slices of the input f, but slices of the polynomial $h(x) = f(Rx)$ obtained by a random change of variables. As a result, each slice of h contains some information on *all* of the slices of f.

Our algorithms are therefore quite simple but their analysis is quite non-trivial. In fact, analysing the case of "negative" inputs, i.e. input polynomials that are not equivalent to any polynomial in $\mathcal{P}_d$, forms the bulk of this paper. For $d > 3$, the notion of "weak-singularity" of matrices (Definition 14) has been introduced which along with the notions of "commutativity property" and "diagonalisability property" helps us to give us another equivalent criterion for testing equivalence to a polynomial in $\mathcal{P}_d$ in Theorem 15. Finally, the crucial part of the proof (for $d > 3$, and already for $d = 3$) is to show that testing commutativity of two matrices and diagonalisability of one matrix is enough for testing these properties for any "symmetric family of symmetric matrices" (refer to Definition 19) with high probability.

Note here that an arbitrary slice of the polynomial is hard to compute, when the polynomial is given as blackbox (because that requires computing arbitrary degree-d partial derivatives using the blackbox). Hence, this particular choice of slices is crucial because they can be computed in polynomial time.

1.4 Notations

We work in a field $\mathbb{K}$ which may be the field of real numbers or the field of complex numbers. Some of our intermediate results apply to other fields as well. We denote by $\mathbb{K}[x_1, \ldots, x_n]_d$ the space of homogeneous polynomials of degree d in n variables with coefficients in $\mathbb{K}$. A

homogeneous polynomial of degree d is also called a *degree-d form*. We denote by P_d the polynomial $\sum_{i=1}^{n} x_i^d$, and we say that a degree d form $f(x_1, \ldots, x_n)$ is equivalent to a sum of d-th powers if it is equivalent to P_d, i.e., if $f(x) = P_d(Ax)$ for some invertible matrix A. More generally, we denote by $\mathcal{P}_d$ the set of polynomials of the form $\sum_{i=1}^{n} \alpha_i x_i^d$ with $\alpha_i \neq 0$ for all i. As explained in Section 1.3, for $\mathbb{K} = \mathbb{R}$ we are not only interested in equivalence to P_d: we would like to know whether the input is equivalent to one of the elements of $\mathcal{P}_d$.

We denote by $M_n(\mathbb{K})$ the space of square matrices of size n with entries from $\mathbb{K}$. We denote by ω a feasible exponent for matrix multiplication, i.e., we assume that two matrices of $M_n(\mathbb{K})$ can be multiplied with $O(n^\omega)$ arithmetic operations in $\mathbb{K}$.

Throughout the paper, we will choose the entries r_{ij} of a matrix R independently and uniformly at random from a finite set $S \subset \mathbb{K}$. When we calculate the probability of some event E over the random choice of the r_{ij}, by abuse of notation instead of $\Pr_{r_{11},\ldots,r_{nn} \in S}[E]$ we simply write $\Pr_{R \in S}[E]$.

2 Equivalence to a linear combination of d-th powers

We can associate to a symmetric tensor T of order d the homogeneous polynomial

$$f(x_1, ..., x_n) = \sum_{i_1,\ldots,i_d \in [n]} T_{i_1 \ldots i_d} x_{i_1} \ldots x_{i_d}.$$

This correspondence is bijective, and the symmetric tensor associated to a homogeneous polynomial f can be obtained from the relation

$$T_{i_1 \ldots i_d} = \frac{1}{d!} \frac{\partial^d f}{\partial x_{i_1} \ldots \partial x_{i_d}}.$$

The $(i_1, ..., i_{d-2})$-th slice of T is the symmetric matrix $T_{i_1 \ldots i_{d-2}}$ with entries $(T_{i_1 \ldots i_{d-2}})_{i_{d-1}, i_d} = T_{i_1 \ldots i_d}$. Recall from Section 1.4 that we denote by $\mathcal{P}_d$ the set of polynomials of the form $\sum_{i=1}^{n} \alpha_i x_i^d$ with $\alpha_i \neq 0$ for all $i \in [n]$. In this section we analyze Algorithm 1. Recall from the introduction that this is a polynomial time algorithm for checking whether an input degree d form in n variables f is equivalent to some polynomial in $\mathcal{P}_d$. This means that $f(x) = P_d(Ax)$ for some $P_d \in \mathcal{P}_d$ and some invertible matrix A.

We prove the surprising fact that even when the input polynomial has degree d, checking commutativity of 2 matrices and the diagonalisability of one matrix is enough to check equivalence to a polynomial of $\mathcal{P}_d$. Let $\{T_{i_1,\ldots,i_{d-2}}\}_{i_1,\ldots,i_{d-2} \in [n]}$ be the slices of $h(x) = f(Rx)$. Recall that $T_{\bar{i}}$ denotes the slice $T_{i \ldots i}$. Algorithm 1 checks if $T'_{\bar{1}} T_{\bar{2}}$ commutes with $T'_{\bar{1}} T_{\bar{3}}$ and if $T'_{\bar{1}} T_{\bar{3}}$ is diagonalisable.

Interestingly though, arbitrary slices of a degree-d polynomial are hard to compute. These particular slices are special because they can be computed using a small number of calls to the blackbox and in small number of arithmetic operations (due to the fact that they are essentially repeated partial derivatives with respect to a single variable). Hence, they help us give a polynomial time algorithm. More precisely, we show that if the polynomial is given as a blackbox, the algorithm requires only $O(n^2 d)$ calls to the blackbox and $O(n^2 M(d) \log d + n^{\omega+1})$ many arithmetic operations. We do a detailed complexity analysis of this algorithm in Appendix B in the full paper [20].

The remainder of this section is devoted to a correctness proof for Algorithm 1, including an analysis of the probability of error. Our main result about this algorithm is as follows:

▶ **Theorem 3.** *If an input $f \in \mathbb{F}[x_1, ..., x_n]_d$ is not equivalent to some polynomial $P_d \in \mathcal{P}_d$, then f is rejected by the algorithm with high probability over the choice of the random matrix R. More precisely, if the entries $r_{i,j}$ of R are chosen uniformly and independently at random*

from a finite set $S \subseteq \mathbb{K}$, then the input will be rejected with probability $\geq 1 - \frac{2(d-2)}{|S|}$. Conversely, if f is equivalent to some polynomial $P_d \in \mathcal{P}_d$, then f will be accepted with high probability over the choice of the random matrix R. More precisely, if the entries $r_{i,j}$ are chosen uniformly and independently at random from a finite set $S \subseteq \mathbb{K}$, then the input will be accepted with probability $\geq 1 - \frac{n(d-1)}{|S|}$.

In Section 2.2, we give a proof of the second part of theorem, i.e., we analyze the behavior of Algorithm 1 on the polynomials that can be decomposed as a sum of dth powers of linearly independent linear forms (which we refer to as the positive inputs). We give a characterization of positive inputs in terms of the subspace of matrices spanned by their slices. Namely, we show that these subspaces must not be "weakly singular", and must satisfy the commutativity and diagonalizability properties. If a polynomial is not equivalent to some polynomial in $\mathcal{P}_d$, this can happen in several ways depending on which property fails. We analyze the failure of commutativity in Section 2.3.1 and failure of diagonalisability in Section 2.3.2. Then we collect everything together and prove the first part of the theorem in Section 2.3.3.

2.1 Characterisation of equivalence to $\mathcal{P}_d$

First, we show how the slices of a degree-d form evolve under a linear change of variables. Towards the proof of Theorem 2 we need some results from [21], which we recall in this section. We also give a converse of this in Theorem 8. First, let us recall how the slices of a polynomial evolve under a linear change of variables.

▶ **Theorem 4.** *Let g be a degree-d form with slices $\{S_{i_1 \dots i_{d-2}}\}_{i_1,\dots,i_{d-2} \in [n]}$ and let $f(x) = g(Ax)$. Then the slices $T_{i_1 \dots i_{d-2}}$ of f, are given by $T_{i_1 \dots i_{d-2}} = A^T D_{i_1 \dots i_{d-2}} A$ where $D_{i_1 \dots i_{d-2}} = \sum_{j_1 \dots j_{d-2} \in [n]} a_{j_1 i_1} \dots a_{j_{d-2} i_{d-2}} S_{j_1 \dots j_{d-2}}$ and $a_{i,j}$ are the entries of A.*
If $g = \sum_{i=1}^n \alpha_i x_i^d$, we have $D_{i_1 \dots i_{d-2}} = diag(\alpha_1(\prod_{m=1}^{d-2} a_{1,i_m}), \dots, \alpha_n(\prod_{m=1}^{d-2} a_{n,i_m}))$.

Proof. By definition of the slices of a polynomial,

$$S_{i_1 \dots i_{d-2}} = \frac{1}{d!} H_{\frac{\partial^{d-2} g}{\partial x_{i_1} \dots \partial x_{i_{d-2}}}}(x) \text{ and } T_{i_1 \dots i_{d-2}} = \frac{1}{d!} H_{\frac{\partial^{d-2} f}{\partial x_{i_1} \dots \partial x_{i_{d-2}}}}(x)$$

where $H_f(x)$ is the Hessian matrix of f at point x. Since $f(x) = g(Ax)$, by differentiating d times, we get that $\frac{\partial^d f}{\partial x_{i_1} \dots \partial x_{i_d}}(x) = \sum_{j_1 \dots j_d \in [n]} a_{j_1 i_1} \dots a_{j_d i_d} \frac{\partial^d g}{\partial x_{j_1} \dots \partial x_{j_d}}(Ax)$. Putting these equations in matrix form, and using the fact that $\frac{\partial^d g}{\partial x_{j_1} \dots \partial x_{j_d}}(Ax) = \frac{\partial^d g}{\partial x_{j_1} \dots \partial x_{j_d}}(x)$ we get the desired result. ◀

Instead of diagonalisation by congruence, it is convenient to work with the more familiar notion of diagonalisation by similarity, where an invertible matrix A acts by $S \mapsto A^{-1}SA$ instead of $A^T SA$. We collect the necessary material in the remainder of this section (and we refer to diagonalisation by similarity simply as *diagonalisation*).

The two following properties play a fundamental role throughout the paper.

▶ **Definition 5.** *Let $\mathcal{V}$ be a non-singular space of matrices.*

- *We say that $\mathcal{V}$ satisfies the **Commutativity Property** if there exists an invertible matrix $A \in \mathcal{V}$ such that $A^{-1}\mathcal{V}$ is a commuting subspace, i.e., $PQ = QP$ for any two matrices $P, Q \in A^{-1}\mathcal{V}$.*
- *We say that $\mathcal{V}$ satisfies the **Diagonalisability Property** if there exists an invertible matrix $B \in \mathcal{V}$ such that all the matrices in the space $B^{-1}\mathcal{V}$ are diagonalisable.*

The next result can be found in [21, Section 2.2].

▶ **Theorem 6.** *Let $\mathcal{V}$ be a non-singular subspace of matrices of $M_n(\mathbb{K})$. The following properties are equivalent.*

- *$\mathcal{V}$ satisfies the commutativity property.*
- *For all non-singular matrices $A \in \mathcal{V}$, $A^{-1}\mathcal{V}$ is a commuting subspace.*

▶ Remark 7. Let $\mathcal{V}$ be a non-singular subspace of matrices which satisfies the commutativity and diagonalisability properties. There exists an invertible matrix $B \in \mathcal{V}$ and an invertible matrix R which diagonalizes simultaneously all of $B^{-1}\mathcal{V}$ (i.e., $R^{-1}MR$ is diagonal for all $M \in B^{-1}\mathcal{V}$).

Proof. Pick an invertible matrix $B \in \mathcal{V}$ such that $\mathcal{W} = B^{-1}\mathcal{V}$ is a space of diagonalizable matrices. By Theorem 6, $\mathcal{W}$ is a commuting subspace. It is well known that a finite collection of matrices is simultaneously diagonalisable if and only if they commute, and each matrix in the collection is diagonalisable. We conclude by applying this result to a basis of $\mathcal{W}$ (any matrix R which diagonalises a basis will diagonalise all of $\mathcal{W}$). ◀

We now give an analogue of Theorem 6 for the diagonalisability property.

▶ **Theorem 8.** *Let $\mathcal{V}$ be a non-singular subspace of matrices which satisfies the commutativity property. The following properties are equivalent:*

- *$\mathcal{V}$ satisfies the diagonalisability property.*
- *For all non-singular matrices $A \in \mathcal{V}$, the matrices in $A^{-1}\mathcal{V}$ are simultaneously diagonalisable.*

Proof. Suppose that $\mathcal{V}$ satisfies the diagonalisability property. By the previous remark, we already know that there exists *some* invertible matrix $B \in \mathcal{V}$ such that the matrices in $B^{-1}\mathcal{V}$ are simultaneously diagonalisable by an invertible matrix R. We need to establish the same property for an arbitrary invertible matrix $A \in \mathcal{V}$. For any $M \in \mathcal{V}$, $A^{-1}M = (B^{-1}A)^{-1}(B^{-1}M)$. Hence $A^{-1}M$ is diagonalised by R since this matrix diagonalises both matrices $B^{-1}A$ and $B^{-1}M$. Since R is independent of the choice of $M \in \mathcal{V}$, we have shown that the matrices in $A^{-1}\mathcal{V}$ are simultaneously diagonalisable. ◀

The importance of the commutativity and diagonalisability properties stems from the fact that they provide a characterization of simultaneous diagonalisation by congruence, which in turn provides a characterization of equivalence to some polynomial in $\mathcal{P}_d$:

▶ **Theorem 9.** *Let $A_1, ..., A_k \in M_n(\mathbb{K})$ and assume that the subspace $\mathcal{V}$ spanned by these matrices is non-singular. There are diagonal matrices Λ_i and a non-singular matrix $R \in M_n(\mathbb{K})$ such that $A_i = R\Lambda_i R^T$ for all $i \in [k]$ if and only if $\mathcal{V}$ satisfies the Commutativity property and the Diagonalisability property.*

For a proof, see [21, Section 2.2] for $\mathbb{K} = \mathbb{C}$ and [21, Section 2.3] for $\mathbb{K} = \mathbb{R}$.

The next lemma uses Theorem 4 to reveal some crucial properties about the subspace spanned by the slices of any degree-d form which is equivalent to some $g \in \mathcal{P}_d$.

▶ **Lemma 10.** *Let $f(x_1, ..., x_n)$ and $g(x_1, ..., x_n)$ be two forms of degree d such that $f(x) = g(Ax)$ for some non-singular matrix A.*

1. *If $\mathcal{U}$ and $\mathcal{V}$ denote the subspaces of $M_n(\mathbb{K})$ spanned respectively by the slices of f and g, we have $\mathcal{U} = A^T\mathcal{V}A$.*
2. *$\mathcal{V}$ is non-singular iff $\mathcal{U}$ is non-singular.*
3. *In particular, for $g \in \mathcal{P}_d$ the subspace $\mathcal{V}$ is the space of diagonal matrices and $\mathcal{U}$ is a non-singular subspace, i.e., it is not made of singular matrices only.*

Proof. Theorem 4 shows that $\mathcal{U} \subseteq A^T\mathcal{V}A$. Now since, $g(x) = f(A^{-1}x)$, same argument shows that $\mathcal{V} \subseteq A^{-T}\mathcal{U}A^{-1}$. This gives us that $\mathcal{U} = A^T\mathcal{V}A$.

For the second part of the lemma, let us assume that $\mathcal{V}$ is non-singular and $M_{\mathcal{U}}$ be an arbitrary matrix in $\mathcal{U}$. Using the previous part of the lemma, we know that there exists $M_{\mathcal{V}} \in \mathcal{V}$ such that $M_{\mathcal{U}} = A^T M_{\mathcal{V}} A$. Since $\mathcal{V}$ is non-singular, $\det(M_{\mathcal{V}}) \neq 0$. Taking determinant on both sides, we get that $\det(M_{\mathcal{U}}) = \det(A)^2\det(M_{\mathcal{V}}) \neq 0$ (since A is invertible, $\det(A) \neq 0$). For the converse, assume that $\mathcal{U}$ is non-singular. Following a similar proof, it can be shown that $\det(M_{\mathcal{U}}) \neq 0$.

For the third part of the lemma, let $\{S_{i_1\ldots i_{d-2}}\}_{i_1\ldots i_{d-2}\in[n]}$ be the slices of g. If $g = \sum_{i\in[n]} \alpha_i x_i^d$, such that $\alpha_i \neq 0$ for all i, $S_{\bar{i}}$ has α_i in the (i,i)-th position and 0 everywhere. Also, $S_{i_1,\ldots,i_{d-2}} = 0$, when the i_k's are not equal. Hence, $\mathcal{V}$ is the space of all diagonal matrices. Hence $\mathcal{V}$ is a non-singular space. Using the previous part of the lemma, we get that $\mathcal{U}$ is a non-singular space as well. ◀

The next lemma is effectively a converse of the second part of Lemma 10. It shows that if the slices of f are diagonal matrices, then the fact that they effectively originate from a symmetric tensor forces them to be extremely special.

▶ **Lemma 11.** *Let $f \in \mathbb{K}[x_1,\ldots,x_n]_d$ be a degree-d form. If the slices of f are diagonal matrices, then $f = \sum_{i\in[n]} \alpha_i x_i^d$ for some $\alpha_1,\ldots,\alpha_n \in \mathbb{K}$.*

Proof. Let $T_{i_1,\ldots,i_{d-2}}$ be the slices of f. Let $I = \{(i_{\sigma(1)},\ldots,i_{\sigma(d)}) | \sigma \in S_d\}$. Now since they are slices of a polynomial, we know that

$$(T_{i_1\ldots i_{d-2}})_{i_{d-1},i_d} = (T_{i_{\sigma(1)},\ldots,i_{\sigma(d-2)}})_{i_{\sigma(d-1)},i_{\sigma(d)}}. \tag{3}$$

We want to show that $T_{i_1,\ldots,i_d} \neq 0$ only if $i_1 = i_2 = \ldots = i_d$. Using (3), it is sufficient to show that $(T_{i_1,\ldots,i_{d-2}})_{i_{d-1},i_d} \neq 0$ only if $i_{d-1} = i_d$. This is true since $T_{i_1,\ldots,i_{d-2}}$ are diagonal matrices. This gives us that $f = \sum_{i\in[n]} \alpha_i x_i^d$. ◀

Now we are finally ready to prove a theorem that characterizes exactly the set of degree-d homogeneous polynomials which are equivalent to some $g \in \mathcal{P}_d$. This already appears as Theorem 2 in the introduction. We restate it now for the reader's convenience.

▶ **Theorem 12.** *A degree d form $f \in \mathbb{K}[x_1,\ldots,x_n]$ is equivalent to some polynomial $P_d \in \mathcal{P}_d$ if and only if its slices $\{T_{i_1,\ldots,i_{d-2}}\}_{i_1,\ldots,i_{d-2}\in[n]}$ span a non-singular matrix space and the slices are simultaneously diagonalisable by congruence, i.e., there exists an invertible matrix $Q \in M_n(\mathbb{K})$ such that the matrices $Q^T T_{i_1\ldots i_{d-2}} Q$ are diagonal for all $i_1,\ldots,i_{d-2} \in [n]$.*

Proof. Let $\mathcal{U}$ be the space spanned by $\{T_{i_1,\ldots,i_{d-2}}\}_{i_1,\ldots,i_{d-2}\in[n]}$. If f is equivalent to P_d, Theorem 4 shows that the slices of f are simultaneously diagonalisable by congruence and Lemma 10 shows that $\mathcal{U}$ is non-singular.

Let us show the converse. Since the slices $\{T_{i_1,\ldots,i_{d-2}}\}_{i_1,\ldots,i_{d-2}\in[n]}$ are simultaneously diagonalisable, there are diagonal matrices $\Lambda_{i_1\ldots i_{d-2}}$ and a non-singular matrix $R \in M_n(\mathbb{K})$ such that $T_{i_1\ldots i_{d-2}} = R\Lambda_{i_1\ldots i_{d-2}}R^T$ for all $i_1,\ldots,i_{d-2} \in [n]$. So now we consider $g(x) = f(R^{-T}x)$. Let $\{S_{i_1,\ldots,i_{d-2}}\}_{i_1,\ldots,i_{d-2}\in[n]}$ be the slices of g. Using Theorem 4, we get that $S_{i_1\ldots i_{d-2}} = (R^{-1})(\sum_{j_1\ldots j_{d-2}\in[n]} r_{j_1 i_1}\ldots r_{j_{d-2}i_{d-2}} R\Lambda_{j_1\ldots j_{d-2}}R^T)R^{-T} = \sum_{j_1\ldots j_{d-2}\in[n]} r_{j_1 i_1}\ldots r_{j_{d-2}i_{d-2}}\Lambda_{j_1\ldots j_{d-2}}$. This implies that $S_{j_1\ldots j_{d-2}}$ are also diagonal matrices. By Lemma 11, $g = \sum_{i\in[n]} \alpha_i x_i^d$. It therefore remains to be shown that $\alpha_i \neq 0$, for all $i \in [n]$. Let $\mathcal{V}$ be the subspace spanned by the slices of g and the slices of f span a non-singular matrix space $\mathcal{U}$. Since, $\mathcal{U}$ is a non-singular subspace of matrices, using part (2) of Lemma 10, we get that $\mathcal{V}$ is a non-singular subspace of matrices.

But if some α_i vanishes, for all $A \in \mathcal{V}$, $A_{\bar{i}} = 0$. Hence $\mathcal{V}$ is a singular subspace, which is a contradiction. This gives us that $g = \sum_{i=1}^{n} \alpha_i x_i^d$ where $\alpha_i \neq 0$ for all i. Hence, $g \in \mathcal{P}_d$ and f is equivalent to g. ◀

▶ **Theorem 13.** *Let $f \in \mathbb{K}[x_1, ..., x_n]$ be a degree-d form. f is equivalent to some polynomial $P_d \in \mathcal{P}_d$ iff the subspace $\mathcal{V}$ spanned by its slices $\{T_{i_1,...,i_{d-2}}\}_{i_1,...,i_{d-2}\in[n]}$ is a non-singular subspace and $\mathcal{V}$ satisfies the Commutativity Property and the Diagonalisability Property.*

Proof. This follows from Theorem 12 and Theorem 9 for $k = n^{d-2}$ to get the result. ◀

Notice here that the notion of weak-singularity is entirely dependent on the generating set of matrices. So it is more of a property of the generating set. But by abuse of language, we will call the span of the matrices weakly singular. To put it in contrast, usually the notion of singularity is a property of the subspace spanned by the matrices (irrespective of the generating set). It can be further observed that for all $n \geq 2$ and $d \geq 4$, non-singular families of matrices can be easily constructed which are weakly singular!

▶ **Definition 14** (Weak singularity)**.** *Let $\mathcal{V}$ be the space spanned by matrices $\{S_{i_1,...,i_{d-2}}\}_{i_1,...,i_{d-2}\in[n]}$. $\mathcal{V}$ is weakly singular if for all $\alpha = (\alpha_1, ..., \alpha_n)$, $det(\sum_{i_1,...,i_{d-2}\in[n]}(\prod_{k\in[d-2]} \alpha_{i_k})S_{i_1...i_{d-2}}) = 0$.*

Notice here that the notion of weak-singularity is entirely dependent on the generating set of matrices. So it is more of a property of the generating set. But by abuse of language, we will call the span of the matrices weakly singular. To put it in contrast, usually the notion of singularity is a property of the subspace spanned by the matrices (irrespective of the generating set).

▶ **Theorem 15.** *Let $f \in \mathbb{K}[x_1, ..., x_n]$ be a degree-d form. f is equivalent to some polynomial $P_d \in \mathcal{P}_d$ iff the subspace $\mathcal{V}$ spanned by its slices $\{T_{i_1,...,i_{d-2}}\}_{i_1,...,i_{d-2}\in[n]}$ is not a weakly singular subspace, satisfies the Commutativity Property and the Diagonalisability Property.*

Proof. First we show that if $f = P_d(Ax)$ such that $P_d \in \mathcal{P}_d$ i.e. $P_d(x) = \sum_{i=1}^{n} \alpha_i x_i^d$ where $\alpha_i \neq 0$ for all $i \in [n]$ and A is invertible, then $\mathcal{V}$ is not a weakly singular subspace, satisfies the commutativity property and the diagonalisability property.
Let $\{S_{i_1...i_{d-2}}\}_{i_1,...,i_{d-2}\in[n]}$ be the slices of P_d. Then $S_{\bar{i}} = \alpha_i \text{diag}(e_i)$ where e_i is the i-th standard basis vector, and all other slices are 0. From Theorem 4,

$$T_{i_1...i_{d-2}} = A^T D_{i_1...i_{d-2}} A = A^T (\sum_{k\in[n]} a_{ki_1}...a_{ki_{d-2}} S_{\bar{k}})A.$$

Now we define

$$\begin{aligned} T(\bar{\beta}) &= \sum_{i_1,...,i_{d-2}\in[n]} (\prod_{k\in[d-2]} \beta_{i_k}) T_{i_1...i_{d-2}} \\ &= \sum_{i_1,...,i_{d-2}\in[n]} (\prod_{k\in[d-2]} \beta_{i_k}) A^T (\text{diag}(\alpha_1(\prod_{m\in[d-2]} a_{1i_m}), ..., \alpha_n(\prod_{m\in[d-2]} a_{ni_m})))A \\ &= A^T \text{diag}(\alpha_1(\sum_{i_1,...,i_{d-2}\in[n]} (\prod_{k\in[d-2]} \beta_{i_k} a_{1i_k})), ..., \alpha_n(\sum_{i_1,...,i_{d-2}\in[n]} (\prod_{k\in[d-2]} \beta_{i_k} a_{ni_k})))A. \end{aligned}$$

Taking determinant on both sides, $\det(T)(\bar{\beta}) = \det(A)^2 \prod_{m=1}^{n} T_m(\bar{\beta})$ where $T_m(\bar{\beta}) = \alpha_m(\sum_{i_1,...,i_{d-2}\in[n]}(\prod_{k\in[d-2]} \beta_{i_k} a_{mi_k}))$. Since, A is invertible, none of its rows are all 0. Hence for all $m_0 \in [n]$, there exists $j_0 \in [n]$, such that $a_{m_0 j_0} \neq 0$. Then $\text{coeff}_{\beta_{j_0}^{d-2}}(T_{m_0}) = a_{m_0 j_0}^{d-2} \neq 0$. Hence $T_{m_0} \not\equiv 0$ for all $m_0 \in [n]$ which implies that $\det(T) \not\equiv 0$. Therefore, there exists $\bar{\beta^0}$ such that $\det(T)(\bar{\beta^0}) \neq 0$. This proves that $\det(\sum_{i_1,...,i_{d-2}\in[n]}(\prod_{k\in[d-2]} \beta_{i_k}) T_{i_1...i_{d-2}}) \not\equiv 0$.

Hence, $\mathcal{V} = \text{span}\{T_{i_1\ldots i_{d-2}}\}_{i_1,\ldots,i_{d-2}\in[n]}$ is not weakly singular. Theorem 13 gives us that the subspace spanned by the slices $\mathcal{V}$ satisfies the commutativity property and the diagonalisability property.

For the converse, if $\mathcal{V}$ is not a weakly singular subspace, then it is a non-singular subspace as well. And it satisfies the commutativity property and the diagonalisability property. By Theorem 13, we get that f is equivalent to some polynomial in $\mathcal{P}_d$. ◀

2.2 Analysis for positive inputs

In this section we analyze the behavior of Algorithm 1 on inputs that are equivalent to some polynomial in $\mathcal{P}_d$ (which we refer to as the positive inputs). We recall here again that by $T_{\bar{1}}$, we denote the slice $T_{11\ldots1}$.

▶ **Lemma 16.** *Let $f \in \mathbb{K}[x_1,\ldots,x_n]_d$ with slices $\{S_{i_1,\ldots,i_{d-2}}\}_{i_1,\ldots,i_{d-2}\in[n]}$, such that the subspace $\mathcal{V}$ spanned by the slices is not weakly singular. Let $h(x) = f(Rx)$ where the entries $r_{i,j}$ are chosen uniformly and independently at random from a finite set $S \subseteq \mathbb{K}$. Let $\{T_{i_1,\ldots,i_{d-2}}\}_{i_1,\ldots,i_{d-2}\in[n]}$ be the slices of h. Then $Pr_{R\in S}[T_{\bar{1}}$ is invertible$] \geq 1 - \frac{n(d-1)}{|S|}$.*

Proof. We can obtain the slices $T_{i_1\ldots i_{d-2}}$ of h from the slices $S_{i_1\ldots i_{d-2}}$ of f using Theorem 4. Namely, we have $T_{i_1\ldots i_{d-2}} = R^T D_{i_1\ldots i_{d-2}} R$ where $D_{i_1\ldots i_{d-2}} = \sum_{j_1\ldots j_{d-2}\in[n]}(\prod_{m\in[d-2]} r_{j_m,i_m})S_{j_1\ldots j_{d-2}}$. Therefore $T_{\bar{1}}$ is invertible iff R and $D_{\bar{1}}$ are invertible. Applying Schwartz-Zippel lemma to $\det(R)$ shows that R is singular with probability at most $\frac{n}{|S|}$. We will show that $D_{\bar{1}}$ is singular with probability at most $\frac{n(d-2)}{|S|}$. The lemma then follows from the union bound. Matrix $D_{\bar{1}}$ is not invertible iff $\det(D_{\bar{1}}) = 0$. Since, $D_{\bar{1}} = \sum_{j_1\ldots j_{d-2}\in[n]}(\prod_{m\in[d-2]} r_{j_m,1})S_{j_1\ldots j_{d-2}}$, $\det(D_{\bar{1}}) \in \mathbb{K}[r_{1,1},\ldots,r_{n,1}]$ and $\deg(\det(D_{\bar{1}})) \leq n(d-2)$. Since, $\mathcal{V}$ is not weakly singular, there exists some choice of $\alpha = (\alpha_1,\ldots,\alpha_n)$, such that $S = \sum_{i_1,\ldots,i_{d-2}\in[n]}(\prod_{m\in[d-2]} \alpha_{i_m})S_{i_1\ldots i_{d-2}}$ is invertible. Hence, $\det(S) \neq 0$. This gives us that $\det(D_{\bar{1}})(\alpha) \neq 0$. which gives us that $\det(D_{\bar{1}}) \not\equiv 0$. From the Schwartz-Zippel lemma, it follows that $\Pr_{R\in S}[\det(D_{\bar{1}}) = 0] \leq \frac{n(d-2)}{|S|}$. ◀

Recall here from Section 1.4, we define by $\mathcal{P}_d$, the set of all polynomials of the form $\sum_{i=1}^n \alpha_i x_i^d$ such that $0 \neq \alpha_i \in \mathbb{K}$ for all $i \in [n]$.

▶ **Lemma 17.** *Given $A \in M_n(\mathbb{K})$, let $\{T_{i_1,\ldots,i_{d-2}}\}_{i_1,\ldots,i_{d-2}\in[n]}$ be the slices of $h(x) = P_d(Ax)$ where $P_d \in \mathcal{P}_d$. If $T_{\bar{1}}$ is invertible, define $T'_{\bar{1}} = (T_{\bar{1}})^{-1}$. Then $T'_{\bar{1}}T_{\bar{2}}$ commutes with $T'_{\bar{1}}T_{\bar{3}}$ and $T'_{\bar{1}}T_{\bar{2}}$ is diagonalisable.*

Proof. Let $P_d = \sum_{i=1}^n \alpha_i x_i^d$ where $\alpha_i \neq 0$. By Theorem 4,

$$T_{i_1\ldots i_{d-2}} = A^T(\text{diag}(\alpha_1(\prod_{m=1}^{d-2} a_{1,i_m}),\ldots,\alpha_n(\prod_{m=1}^{d-2} a_{n,i_m})))A = A^T D_{i_1\ldots i_{d-2}} A.$$

If $T_{\bar{1}}$ is invertible, the same is true of A and $D_{\bar{1}}$. The inverse $(D_{\bar{1}})^{-1}$ is diagonal like $D_{\bar{1}}$, hence $(D_{\bar{1}})^{-1}D_{\bar{2}}$ and $(D_{\bar{1}})^{-1}D_{\bar{3}}$ are both diagonal as well and must therefore commute. Now, $T'_{\bar{1}}T_{\bar{2}}T'_{\bar{1}}T_{\bar{3}} = A^{-1}((D_{\bar{1}})^{-1}D_{\bar{2}}(D_{\bar{1}})^{-1}D_{\bar{3}})A = A^{-1}((D_{\bar{1}})^{-1}D_{\bar{3}}(D_{\bar{1}})^{-1}D_{\bar{2}})A = T'_{\bar{1}}T_{\bar{3}}T'_{\bar{1}}T_{\bar{2}}$.

Finally, $T'_{\bar{1}}T_{\bar{2}} = A^{-1}((D_{\bar{1}})^{-1}D_{\bar{2}})A$ so this matrix is diagonalisable. ◀

We are now in a position to prove the easier half of Theorem 3.

▶ **Theorem 18.** *If an input $f \in \mathbb{K}[x_1,\ldots,x_n]_d$ is equivalent to some polynomial $P_d \in \mathcal{P}_d$ then f will be accepted by Algorithm 1 with high probability over the choice of the random matrix R. More precisely, if the entries $r_{i,j}$ are chosen uniformly and independently at random from a finite set $S \subseteq \mathbb{K}$, then the input will be accepted with probability $\geq (1 - \frac{n(d-1)}{|S|})$.*

Proof. We start by assuming that $f = P_d(Bx)$ for some $P_d \in \mathcal{P}_d$ where B is an invertible matrix. By Theorem 15, we know that the subspace spanned by the slices of f is not weakly singular. We can therefore apply Lemma 16, the first slice $T_{\bar{1}}$ of $h(x) = f(Rx)$ is invertible with probability at least $1 - \frac{n(d-1)}{|S|}$. Moreover if $T_{\bar{1}}$ is invertible, Lemma 17 shows that, f will always be accepted. (We can apply this lemma to h since $h = P_d(RBx)$). ◀

2.3 Analysis of negative inputs

In this section, we analyse the behaviour of Algorithm 1 on the inputs that are not equivalent to any polynomial in $\mathcal{P}_d$ (which we refer to as the negative inputs). The main goal is to show that the algorithm rejects negative inputs with high probability.

2.3.1 Failure of commutativity

▶ **Definition 19.** *Let $\{S_{i_1,\dots,i_d}\}_{i_1,\dots,i_d\in[n]}$ be a family of matrices. We say that the matrices form a symmetric family of symmetric matrices if each matrix in the family is symmetric and for all permutations $\sigma \in S_d$, $S_{i_1,\dots,i_d} = S_{i_{\sigma(1)}\dots i_{\sigma(d)}}$.*

In the next lemma, we show that if a symmetric family of symmetric matrices (this family has size n^d) is not a commuting family, then two linear combinations of these matrices formed by picking just $2n$ elements at random also do not commute with high probability.

▶ **Lemma 20** (General commutativity lemma)**.** *Let $\{S_{i_1,\dots,i_d}\}_{i_1,\dots,i_d\in[n]}$ be a symmetric family of symmetric matrices in $M_n(\mathbb{K})$ that do not form a commuting family. Pick $\alpha = \{\alpha_1, \dots, \alpha_n\}$ and $\alpha' = \{\alpha'_1, \dots, \alpha'_n\}$ uniformly and independently at random from a finite set $S \subset \mathbb{K}$. We define*

$$M_\alpha = \sum_{i_1,\dots,i_d\in[n]} \Big(\prod_{m\in[d]} \alpha_{i_m}\Big) S_{i_1,\dots,i_d} \text{ and } M_{\alpha'} = \sum_{j_1,\dots,j_d\in[n]} \Big(\prod_{m\in[d]} \alpha'_{j_m}\Big) S_{j_1,\dots,j_d}.$$

Then, $Pr_{\alpha,\alpha'\in S}\Big[M_\alpha, M_{\alpha'} \text{ don't commute}\Big] \geq \Big(1 - \frac{2d}{|S|}\Big)$.

Proof. We want to bound the probability of error, i.e $\Pr_{\alpha,\alpha'\in S}\Big[M_\alpha M_{\alpha'} - M_{\alpha'} M_\alpha \neq 0\Big]$. The expression $M_\alpha M_{\alpha'} - M_{\alpha'} M_\alpha$ can be written as $\sum_{\substack{i_1,\dots,i_d\in[n]\\ j_1,\dots,j_d\in[n]}} (\prod_{m\in[d]} \alpha_{i_m}\alpha'_{j_m})(S_{i_1\dots i_d}S_{j_1\dots j_d} - S_{j_1\dots j_d}S_{i_1\dots i_d})$. For a fixed $r, s \in [n]$, we define the polynomial

$$P^{r,s}_{\text{comm}}(\alpha, \alpha') = \sum_{i_1,\dots,i_d,j_1,\dots,j_d\in[n]} \Big(\prod_{m\in[d]} \alpha_{i_m}\alpha'_{j_m}\Big) m^{r,s}_{i_1\dots i_d j_1\dots j_d}$$

where $m^{r,s}_{i_1\dots i_d j_1\dots j_d} = (S_{i_1\dots i_d}S_{j_1\dots j_d} - S_{j_1\dots j_d}S_{i_1\dots i_d})_{r,s}$.

First note that by construction M_α commutes with $M_{\alpha'}$ if and only if for all $r, s \in [n]$ such that $P^{r,s}_{\text{comm}}(\alpha, \alpha') = 0$. Since, $\{S_{i_1,\dots,i_d}\}$ is not a commuting family, there exists $i^0_1, \dots, i^0_d, j^0_1, \dots, j^0_d \in [n]$, such that $S_{i^0_1\dots i^0_d}S_{j^0_1\dots j^0_d} - S_{j^0_1\dots j^0_d}S_{i^0_1\dots i^0_d} \neq 0$. Hence, there exists some entry (r_0, s_0) such that $(S_{i^0_1\dots i^0_d}S_{j^0_1\dots j^0_d} - S_{j^0_1\dots j^0_d}S_{i^0_1\dots i^0_d})_{r_0,s_0} \neq 0$.

Now we claim that $P^{r_0,s_0}_{\text{comm}}(\alpha, \alpha') \not\equiv 0$. It is enough to show that the coefficient of $\alpha_{i^0_1}\dots\alpha_{i^0_d}\alpha'_{j^0_1}\dots\alpha'_{j^0_d}$ in $P^{r_0,s_0}_{\text{comm}}(\alpha, \alpha')$ is non-zero. Let $I_0 = \{(i^0_{\sigma(1)}, \dots, i^0_{\sigma(d)}) | \sigma \in S_d\}$ and let $J_0 = \{(j^0_{\sigma(1)}, \dots, j^0_{\sigma(d)}) | \sigma \in S_d\}$. Then $\text{coeff}_{\alpha_{i^0_1}\dots\alpha_{i^0_d}\alpha'_{j^0_1}\dots\alpha'_{j^0_d}}(P^{r_0,s_0}_{\text{comm}}) = \sum_{\bar{i}\in I_0, \bar{j}\in J_0} m^{r_0 s_0}_{\bar{i}\bar{j}}$. The matrices $S_{i_1\dots i_d}$ form a symmetric family in the sense of Definition 19. Therefore, for all $\bar{i} \in I_0, \bar{j} \in J_0$, $m^{r_0 s_0}_{\bar{i}\bar{j}}$ are equal. This gives us that $\text{coeff}_{\alpha_{i^0_1}\dots\alpha_{i^0_d}\alpha'_{j^0_1}\dots\alpha'_{j^0_d}}(P^{r_0,s_0}_{\text{comm}}) =$

$|I_0||J_0|(m^{r_0,s_0}_{i_1^0\ldots i_d^0 j_1^0\ldots j_d^0}) \neq 0$. Hence $P^{r_0,s_0}_{\text{comm}} \not\equiv 0$ and $\deg(P^{r_0,s_0}_{\text{comm}}) \leq 2d$ and using Schwartz-Zippel lemma, we get that, $\Pr_{\alpha,\alpha' \in S}[P^{r_0,s_0}_{\text{comm}}(\alpha,\alpha') \neq 0] \geq 1 - \frac{2d}{|S|}$. Putting $r = r_0, s = s_0$, this gives us that $\Pr_{\alpha,\alpha' \in S}\Big[M_\alpha, M_{\alpha'} \text{ don't commute}\Big] \geq \Big(1 - \frac{2d}{|S|}\Big)$. ◀

The next result relies on the above lemma. Theorem 21 gives us a way to analyze the case when the slices of the input polynomial fail to satisfy the commutativity property (recall that this property is relevant due to Theorem 15).

▶ **Theorem 21.** *Let $f \in \mathbb{K}[x_1,\ldots,x_n]_d$ be a degree d form such that the subspace of matrices $\mathcal{V}$ spanned by its slices is not weakly singular and does not satisfy the commutativity property. Let $h(x) = f(Rx)$ where the entries $(r_{i,j})$ of R are chosen uniformly and independently at random from a finite set $S \subset \mathbb{K}$. Let $\{T_{i_1\ldots i_{d-2}}\}_{i_1,\ldots,i_{d-2}\in[n]}$ be the slices of h. If $T_{\bar{1}}$ is invertible, define $T'_{\bar{1}} = (T_{\bar{1}})^{-1}$. Then $Pr[T_{\bar{1}}$ is invertible and $T'_{\bar{1}}T_{\bar{2}}, T'_{\bar{1}}T_{\bar{3}}$ commute $] \leq \frac{2(d-2)}{|S|}$.*

Proof. Let $\{S_{i_1\ldots i_{d-2}}\}_{i_1,\ldots,i_{d-2}\in[n]}$ be the slices of f. By Theorem 4, we know that

$$T_{i_1\ldots i_{d-2}} = R^T\Big(\sum_{j_1\ldots j_{d-2}\in[n]}\Big(\prod_{m\in[d-2]} r_{j_m,i_m}S_{j_1\ldots j_{d-2}}\Big)\Big)R.$$

Let us define $D_{i_1\ldots i_{d-2}} = \sum_{j_1\ldots j_{d-2}\in[n]}(\prod_{m\in[d-2]} r_{j_m,i_m})S_{j_1\ldots j_{d-2}}$. Then we have for all $i \in \{2,\ldots,n\}$:

$$T'_{\bar{1}}T_{\bar{i}} = R^{-1}(D_{\bar{1}})^{-1}(R)^{-T}R^T D_{\bar{i}}R = R^{-1}\Big(\sum_{j_1\ldots j_{d-2}\in[n]}\Big(\prod_{m\in[d-2]} r_{j_m,i}\Big)(D_{\bar{1}})^{-1}S_{j_1\ldots j_{d-2}}\Big)R. \quad (4)$$

So, if $T_{\bar{1}}$ is invertible, $T'_{\bar{1}}T_{\bar{2}}$ commutes with $T'_{\bar{1}}T_{\bar{3}}$ iff $(D_{\bar{1}})^{-1}D_{\bar{2}}$ commutes with $(D_{\bar{1}})^{-1}D_{\bar{3}}$.

Let E_1 be the event that $T_{\bar{1}}$ is invertible and $T'_{\bar{1}}T_{\bar{2}}$ commutes with $T'_{\bar{1}}T_{\bar{3}}$. Let E'_1 be the event that $D_{\bar{1}}$ is invertible and $(D_{\bar{1}})^{-1}D_{\bar{2}}$ commutes with $(D_{\bar{1}})^{-1}D_{\bar{3}}$. Let E_4 be the event that R is invertible. Then we have that $E_1 = E'_1 \cap E_4$.

Let E_2 be the event that $D_{\bar{1}}$ is invertible and $\{(D_{\bar{1}})^{-1}S_{i_1,\ldots,i_{d-2}}\}_{i_1,\ldots,i_{d-2}\in[n]}$ is not a commuting family. Since $\mathcal{V}$ does not satisfy the commutativity property, $(D_{\bar{1}})^{-1}\mathcal{V}$ is not a commuting subspace if $D_{\bar{1}}$ is invertible. Hence, the event that $D_{\bar{1}}$ is invertible is the same as the event E_2. This also implies that $E'_1 \subseteq E_2$. Setting $A_{i_1\ldots i_{d-2}} = (D_{\bar{1}})^{-1}S_{i_1\ldots i_{d-2}}, \alpha_i = r_{i,2}, \alpha'_i = r_{i,3}$ and then using Lemma 20, we can conclude that $\Pr_{R\in S}\Big[E'_1|E_2\Big] \leq \frac{2(d-2)}{|S|}$.

Note here that $D_{\bar{1}}$ depends only on the random variables $r_{i,1}$ for all $i \in [n]$ and therefore is independent of $r_{k,2}$ and $r_{l,3}$ for all $k,l \in [n]$, because we assume that the entries of R are all picked uniformly and independently at random.

Let E_3 be the event that $T_{\bar{1}}$ is invertible. Now we know that $T_{\bar{1}}$ is invertible iff R and $D_{\bar{1}}$ are invertible. Then, we have $E_3 = E_2 \cap E_4$ Hence, the probability of error can be bounded as follows: $\Pr_{R\in S}[E_1] = \Pr_{R\in S}[E'_1 \cap E_2 \cap E_4] \leq \Pr_{R\in S}[E'_1|E_2] \leq \frac{2(d-2)}{|S|}$. ◀

2.3.2 Failure of diagonalisability

Theorem 21 gives us a way to analyze the case when the slices of the input polynomial fail to satisfy the commutativity property. With the results in the present section we will be able to analyze the case where the commutativity property is satisfied, but the diagonalisability property fails (recall that these properties are relevant due to Theorem 15).

▶ **Proposition 22.** *Let $\mathcal{U} \subseteq M_n(\mathbb{K})$ be a commuting subspace of matrices. We define $\mathcal{M} := \Big\{M\Big|M \text{ is diagonalisable and } M \in \mathcal{U}\Big\}$. Then $\mathcal{M}$ is a linear subspace of $\mathcal{U}$. In particular, if there exists $A \in \mathcal{U}$ such that A is not diagonalisable then $\mathcal{M}$ is a proper linear subspace of $\mathcal{U}$.*

Proof. $\mathcal{M}$ is trivially closed under multiplication by scalars. Let $M, N \in \mathcal{M}$. These two matrices are diagonalisable by definition of $\mathcal{M}$, and they commute since $\mathcal{M} \subseteq \mathcal{U}$. Hence they are simultaneously diagonalisable. Thus $\mathcal{M}$ is closed under addition as well, which implies that it is a linear subspace of $\mathcal{U}$. ◀

▶ **Lemma 23.** *Let $\{A_{i_1\dots i_d}\}_{i_1,..,i_d\in[n]} \in M_n(\mathbb{K})$ be a commuting family of symmetric matrices. Let us assume that this family is symmetric in the sense of Definition 19 and there exists $i_1^0, ..., i_d^0 \in [n]$ such that $A_{i_1^0\dots i_d^0}$ is not diagonalisable. Let $S \subset \mathbb{K}$ be a finite set. Then $D = \sum_{i_1,\dots,i_d=1}^{n}(\prod_{m\in[d]}\alpha_{i_m})A_{i_1\dots i_d}$ is diagonalisable with probability at most $\frac{d}{|S|}$ when $\alpha_1, ..., \alpha_n$ are chosen uniformly and independently at random from S.*

Proof. We define $\mathcal{U} = \text{span}\{A_{i_1\dots i_d}\}_{i_1,..,i_d\in[n]}$. We also define the class of matrices $\mathcal{M} := \left\{M \middle| M \text{ is diagonalisable and } M \in \mathcal{U}\right\}$. So, we want to show that $\Pr_{\alpha\in S}\left[D \in \mathcal{M}\right] \leq \frac{d}{|S|}$.

Now using Proposition 22, and the hypothesis that there exists $A_{i_1^0\dots i_d^0} \in \mathcal{U} \setminus \mathcal{M}$, we get that $\mathcal{M}$ is a proper linear subspace of $\mathcal{U}$. So $\mathcal{M}$ is an intersection of hyperplanes. Since $A_{i_1^0\dots i_d^0} \notin \mathcal{M}$, there exists a linear form $l_{\mathcal{M}}(X) = \sum_{i,j\in[n]} a_{ij}X_{ij}$ corresponding to a hyperplane such that $l_{\mathcal{M}}(M) = 0$ for all $M \in \mathcal{M}$ and $l_{\mathcal{M}}(A_{i_1^0\dots i_d^0}) \neq 0$. We know that if D is diagonalisable, then $l_{\mathcal{M}}(D) \neq 0$. We compute the polynomial $l_{\mathcal{M}}(D)(\alpha) = \sum_{i_1,\dots,i_d\in[n]}(\prod_{m\in[d]}\alpha_{i_m})m_{i_1\dots i_d}$ where $m_{i_1\dots i_d} = (\sum_{k,l\in[n]} a_{kl}(A_{i_1\dots i_d})_{k,l})$. Now we claim that $l_{\mathcal{M}}(D) \not\equiv 0$. We show this by proving that the coefficient of $\alpha_{i_1^0}\dots\alpha_{i_d^0}$ in $l_{\mathcal{M}}(D)(\alpha)$ is not equal to 0. Let $I_0 = \{(i^0_{\sigma(1)}, ..., i^0_{\sigma(d)})|\sigma \in S_d\}$. Then $\text{coeff}_{\alpha_{i_1^0}\dots\alpha_{i_d^0}}(D) = \sum_{\bar{i}\in I_0} m_{\bar{i}}$. Since the matrices $A_{i_1\dots i_d}$ form a symmetric family, the $m_{\bar{i}}$ are equal for all $\bar{i} \in I_0$. Also, since, $l_{\mathcal{M}}(A_{i_1^0\dots i_d^0}) \neq 0$, we get that $\sum_{k,l\in[n]} a_{k,l}(A_{i_1^0\dots i_d^0})_{k,l} \neq 0$. This gives us that $m_{i_1^0\dots i_d^0} \neq 0$. Hence, we get that $\text{coeff}_{\alpha_{i_1^0}\dots\alpha_{i_d^0}}(D) = |I_0|m_{i_1^0\dots i_d^0} \neq 0$. Thus, $l_{\mathcal{M}}(D) \not\equiv 0$ and $\deg(l_{\mathcal{M}}(D)) \leq d$. From Schwartz-Zippel lemma, we have $\Pr_{\alpha\in S}\left[D \in \mathcal{M}\right] \leq \Pr_{\alpha\in S}[l_{\mathcal{M}}(D)(\alpha) = 0] \leq \frac{d}{|S|}$. ◀

The last result for this section is an analogue of the Theorem 21 for the diagonalisability property.

▶ **Theorem 24.** *Let $f \in \mathbb{K}[x_1, ..., x_n]_d$ be a degree-d form such that the subspace $\mathcal{V}$ spanned by its slices is a not weakly-singular subspace, satisfies the commutativity property, but does not satisfy the diagonalisability property. Let $h(x) = f(Rx)$ where the entries $r_{i,j}$ of R are chosen uniformly and independently at random from a finite set $S \subset \mathbb{K}$, Let $\{T_{i_1\dots i_{d-2}}\}_{i_1,\dots,i_{d-2}\in[n]}$ be the slices of h. If $T_{\bar{1}}$ is invertible, define $T'_{\bar{1}} = (T_{\bar{1}})^{-1}$. Then $Pr[T_{\bar{1}}$ is invertible and $T'_{\bar{1}}T_{\bar{2}}$ is diagonalisable $] \leq \frac{d-2}{|S|}$.*

Proof. As in the proof of Theorem 21, we use the expression for $T'_{\bar{1}}T_{\bar{2}}$ which we obtain from the definition of the slices i.e. $R^{-1}(\sum_{j_1\dots j_{d-2}\in[n]}(\prod_{m\in[d-2]} r_{j_m,2})(D_{\bar{1}})^{-1}S_{j_1\dots j_{d-2}})R$ where $D_{\bar{1}} = \sum_{j_1\dots j_{d-2}\in[n]}(\prod_{m\in[d-2]} r_{j_m,1})S_{j_1\dots j_{d-2}}$. So if $T_{\bar{1}}$ is invertible, $T'_{\bar{1}}T_{\bar{2}}$ is diagonalisable iff $M = (\sum_{j_1\dots j_{d-2}\in[n]}(\prod_{m\in[d-2]} r_{j_m,2})(D_{\bar{1}})^{-1}S_{j_1\dots j_{d-2}})$ is diagonalisable.

We denote by E_1 the event that $T_{\bar{1}}$ is invertible and $T'_{\bar{1}}T_{\bar{2}}$ is diagonalisable and by E'_1 the event that $D_{\bar{1}}$ is invertible and M is diagonalisable. Let E_4 be the event that R is invertible. Hence, $E_1 = E'_1 \cap E_4$.

Let E_2 be the event that $D_{\bar{1}}$ is invertible and $\{(D_{\bar{1}})^{-1}S_{i_1\dots i_{d-2}}\}_{i_1,\dots,i_{d-2}\in[n]}$ is a commuting family and there exists $j_1\dots j_{d-2} \in [n]$ such that $(D_{\bar{1}})^{-1}S_{j_1\dots j_{d-2}}$ is not diagonalisable. Since $\mathcal{V}$ satisfies the commutativity property and does not satisfy the diagonalisability property, by Theorem 8, the event that $D_{\bar{1}}$ is invertible is the event same as E_2. It can also be observed that $E'_1 \subseteq E_2$.

Setting $A_{i_1 \ldots i_{d-2}} = (D_{\bar{1}})^{-1} S_{i_1 \ldots i_{d-2}}$ and setting $\alpha_i = r_{i,2}$ for all $i \in [n]$ and using Lemma 23, we get that $\Pr_{R \in S}\left[E_1' | E_2\right] \leq \frac{d-2}{|S|}$. Now we know that $T_{\bar{1}}$ is invertible iff R and $D_{\bar{1}}$ is invertible. Let E_3 be the event that $T_{\bar{1}}$ is invertible. Then, we have $E_3 = E_2 \cap E_4$. The probability of error can finally be bounded as follows: $\Pr_{R \in S}[E_1 \cap E_3] = \Pr_{R \in S}[E_1' \cap E_2 \cap E_4] \leq \Pr_{R \in S}[E_1' | E_2] \leq \frac{d-2}{|S|}$. ◀

2.3.3 Finishing the analysis for negative inputs

In this section we complete the proof of Theorem 3. The case of positive inputs was treated in Section 2.2. It therefore remains to prove the following result.

▶ **Theorem 25.** *If an input $f \in \mathbb{K}[x_1, ..., x_n]_d$ is not equivalent to some polynomial $P_d \in \mathcal{P}_d$, then f is rejected by the algorithm with high probability over the choice of the random matrix R. More precisely, if the entries $r_{i,j}$ are chosen uniformly and independently at random from a finite set $S \subseteq \mathbb{K}$, then the input will be rejected with probability $\geq (1 - \frac{2(d-2)}{|S|})$.*

Proof. Let $\{S_{i_1,\ldots,i_{d-2}}\}_{i_1,\ldots,i_{d-2} \in [n]}$ be the slices of f and $\mathcal{V} = \text{span}\{S_{i_1,\ldots,i_{d-2}}\}$. From Theorem 13 and Theorem 9, we know that if $f \not\sim P_d$, then there are three disjoint cases:

1. **Case 1:** $\mathcal{V}$ is a weakly singular subspace of matrices.
2. **Case 2:** $\mathcal{V}$ is not a weakly singular subspace and $\mathcal{V}$ does not satisfy the commutativity property.
3. **Case 3:** $\mathcal{V}$ is not a weakly singular subspace, $\mathcal{V}$ satisfies the commutativity property but does not satisfy the diagonalisability property.

Now we try to upper bound the probability of error in each case.
In case 1, $T_{\bar{1}} = R^T(\sum_{j_1 \ldots j_{d-2} \in [n]} r_{j_1,1} \ldots r_{j_{d-2},1} S_{j_1 \ldots j_{d-2}})R \in R^T \mathcal{V} R$ is always singular for any choice of $r_{j,1}$. So f is rejected with probability 1 in this case.
In case 2, we can upper bound the probability of error as follows:

$$\begin{aligned}
&\Pr_{R \in S}[f \text{ is accepted by the algorithm}] \\
&= \Pr_{R \in S}[T_{\bar{1}} \text{ is invertible, } T_{\bar{1}}' T_{\bar{2}}, T_{\bar{1}}' T_{\bar{3}} \text{ commute }, T_{\bar{1}}' T_{\bar{2}} \text{ is diagonalisable}] \\
&\leq \Pr_{R \in S}[T_{\bar{1}} \text{ is invertible, } T_{\bar{1}}' T_{\bar{2}}, T_{\bar{1}}' T_{\bar{3}} \text{ commute }].
\end{aligned}$$

Using Theorem 21, we get that this occurs with probability at most $\frac{2(d-2)}{|S|}$. In Case 3, we have the following upper bound on the probability of error,

$$\begin{aligned}
&\Pr_{R \in S}[f \text{ is accepted by the algorithm}] \\
&= \Pr_{R \in S}[T_{\bar{1}} \text{ is invertible, } T_{\bar{1}}' T_{\bar{2}}, T_{\bar{1}}' T_{\bar{3}} \text{ commute }, T_{\bar{1}}' T_{\bar{2}} \text{ is diagonalisable}] \\
&\leq \Pr_{R \in S}[T_{\bar{1}} \text{ is invertible, } T_{\bar{1}}' T_{\bar{2}} \text{ is diagonalisable}].
\end{aligned}$$

By Theorem 24, this occurs with probability $\leq \frac{d-2}{|S|}$. Therefore in all these three cases, the algorithm rejects f with probability at least $1 - \frac{2(d-2)}{|S|}$. ◀

References

1 Animashree Anandkumar, Rong Ge, Daniel Hsu, Sham M. Kakade, and Matus Telgarsky. Tensor decompositions for learning latent variable models. *Journal of Machine Learning Research*, 15(80):2773–2832, 2014. URL: `http://jmlr.org/papers/v15/anandkumar14b.html`.

2 Alessandra Bernardi, Alessandro Gimigliano, and Monica Ida. Computing symmetric rank for symmetric tensors. *Journal of Symbolic Computation*, 46(1):34–53, 2011.

3 Vishwas Bhargava, Shubhangi Saraf, and Ilya Volkovich. *Reconstruction Algorithms for Low-Rank Tensors and Depth-3 Multilinear Circuits*, pages 809–822. Association for Computing Machinery, New York, NY, USA, 2021. doi:10.1145/3406325.3451096.

4 Aditya Bhaskara, Moses Charikar, Ankur Moitra, and Aravindan Vijayaraghavan. Smoothed analysis of tensor decompositions. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 594–603, New York, NY, USA, 2014. Association for Computing Machinery. doi:10.1145/2591796.2591881.

5 L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and Real Computation*. Springer-Verlag, 1998.

6 L. Blum, M. Shub, and S. Smale. On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines. *Bulletin of the American Mathematical Society*, 21(1):1–46, July 1989.

7 Jérôme Brachat, Pierre Comon, Bernard Mourrain, and Elias Tsigaridas. Symmetric tensor decomposition. *Linear Algebra and its Applications*, 433(11-12):1851–1872, 2010.

8 Enrico Carlini. Reducing the number of variables of a polynomial. In *Algebraic geometry and geometric modeling*, Math. Vis., pages 237–247. Springer, Berlin, 2006. doi:10.1007/978-3-540-33275-6_15.

9 Guillaume Cheze and André Galligo. Four lectures on polynomial absolute factorization. In *Solving polynomial equations*, pages 339–392. Springer, 2005.

10 Guillaume Chèze and Grégoire Lecerf. Lifting and recombination techniques for absolute factorization. *Journal of Complexity*, 23(3):380–420, 2007.

11 Shuhong Gao. Factoring multivariate polynomials via partial differential equations. *Mathematics of computation*, 72(242):801–822, 2003.

12 Ignacio García-Marco, Pascal Koiran, and Timothée Pecatte. Reconstruction algorithms for sums of affine powers. In *Proc. International Symposium on Symbolic and Algebraic Computation (ISSAC)*, pages 317–324, 2017. doi:10.1145/3087604.3087605.

13 Ignacio García-Marco, Pascal Koiran, and Timothée Pecatte. Polynomial equivalence problems for sums of affine powers. In *Proc. International Symposium on Symbolic and Algebraic Computation (ISSAC)*, 2018.

14 Ankit Garg, Nikhil Gupta, Neeraj Kayal, and Chandan Saha. Determinant equivalence test over finite fields and over $\mathbb{Q}$. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 26, page 42, 2019.

15 Richard Harshman. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis. *UCLA working papers in phonetics*, 1970.

16 Zohar Karnin and Amir Shpilka. Reconstruction of generalized depth-3 arithmetic circuits with bounded top fan-in. In *24th Annual IEEE Conference on Computational Complexity (CCC)*, pages 274–285, 2009.

17 Neeraj Kayal. Efficient algorithms for some special cases of the polynomial equivalence problem. In *Symposium on Discrete Algorithms (SODA)*. Society for Industrial and Applied Mathematics, January 2011. URL: https://www.microsoft.com/en-us/research/publication/efficient-algorithms-for-some-special-cases-of-the- polynomial-equivalence-problem/.

18 Neeraj Kayal, Vineet Nair, Chandan Saha, and Sébastien Tavenas. Reconstruction of full rank algebraic branching programs. *ACM Transactions on Computation Theory (TOCT)*, 11(1):2, 2018.

19 Neeraj Kayal and Chandan Saha. Reconstruction of non-degenerate homogeneous depth three circuits. In *Proc. 51st Annual ACM Symposium on Theory of Computing (STOC)*, pages 413–424, 2019.

20 Pascal Koiran and Subhayan Saha. Black box absolute reconstruction for sums of powers of linear forms, 2021. doi:10.48550/arXiv.2110.05305.

21 Pascal Koiran and Mateusz Skomra. Derandomization and absolute reconstruction for sums of powers of linear forms. *Theoretical Computer Science*, 887:63–84, 2021. doi:10.1016/j.tcs.2021.07.005.

22 T. Kolda and B. Bader. Tensor decompositions and applications. *SIAM Rev.*, 51:455–500, 2009.

23 A. Moitra. *Algorithmic Aspects of Machine Learning*. Cambridge University Press, 2018.

24 Hani Shaker. Topology and factorization of polynomials. *Mathematica Scandinavica*, pages 51–59, 2009.

25 Yaroslav Shitov. How hard is the tensor rank? *arXiv preprint*, 2016. arXiv:1611.01559.

26 Amir Shpilka. Interpolation of depth-3 arithmetic circuits with two multiplication gates. *SIAM Journal on Computing*, 38(6):2130–2161, 2009.