

# ON FUNCTIONAL DECOMPOSITION OF MULTIVARIATE POLYNOMIALS WITH DIFFERENTIATION AND HOMOGENIZATION\*

Shangwei ZHAO · Ruyong FENG · Xiao-Shan GAO

DOI: 10.1007/s11424-012-1144-8

Received: 1 June 2011

©The Editorial Office of JSSC & Springer-Verlag Berlin Heidelberg 2012

**Abstract** This paper gives a theoretical analysis for the algorithms to compute functional decomposition for multivariate polynomials based on differentiation and homogenization which were proposed by Ye, Dai, and Lam (1999) and were developed by Faugère, Perret (2006, 2008, 2009). The authors show that a degree proper functional decomposition for a set of randomly decomposable quartic homogeneous polynomials can be computed using the algorithm with high probability. This solves a conjecture proposed by Ye, Dai, and Lam (1999). The authors also propose a conjecture which asserts that the decomposition for a set of polynomials can be computed from that of its homogenization and show that the conjecture is valid with high probability for quartic polynomials. Finally, the authors prove that the right decomposition factors for a set of polynomials can be computed from its right decomposition factor space.

**Key words** Cryptosystem analysis, functional decomposition, homogeneous polynomials, multivariate polynomial, right factor space.

## 1 Introduction

A public key cryptosystem often relies on a hard mathematical problem. One of the hard mathematical problems used in cryptosystems is the functional decomposition problem (FDP) for multivariate polynomials<sup>[1]</sup>. The general FDP for multivariate polynomials has been proved to be NP-hard by Dickerson<sup>[2]</sup>. Based on this fact, Patarin and Goubin<sup>[3]</sup> proposed the  $2R$  scheme which is based on the difficulty of decomposing a set of quartic polynomials. Let  $K$  be a finite field with  $q$  elements. In the original design, the private key consists of

- 1) Three linear bijections  $r, s, t: K^n \rightarrow K^n$ .
- 2) Two quadratic polynomial mappings  $\psi, \phi: K^n \rightarrow K^n$ .

The public key consists of

- 1) The field  $K$  and  $n$ .

---

Shangwei ZHAO

*Beijing Electronic Science and Technology Institute, Beijing 100070, China.*

Ruyong FENG · Xiao-Shan GAO

*KLMM, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.*

\*This research is partially supported by a National Key Basic Research Project of China under Grant No. 2011CB302400 and by a Grant from NSFC with Nos 60821002 and 10901156.

◊ *This paper was recommended for publication by Editor Ziming LI.*

2) The composition of polynomial mapping  $\pi = t \circ \psi \circ s \circ \phi \circ r$ , which is a set of polynomials of degree four.

In the encryption system, the quadratic polynomials are chosen from the given S-boxes, which can be inverted easily. Given the composition of two quadratic polynomials, if we know the private key, then we can obtain the plaintext. Otherwise, it is difficult to invert the polynomials of degree four directly. So, attack on the 2R scheme is reduced to the functional decomposition of quartic polynomials.

Efficient algorithms for several special forms of FDP are known. Polynomial-time algorithms are proposed for univariate decomposition of multivariate polynomials and multivariate decomposition of univariate polynomials<sup>[4–6]</sup>. Efficient algorithms for a kind of monomial decompositions of rational functions are proposed in [7], which is further extended to a complete decomposition algorithm for rational parametrization of ruled-surfaces and other cases<sup>[8–10]</sup>.

Ye, Dai, and Lam proposed an efficient algorithm for decomposing a set of  $n$  polynomials of degree four into two sets of quadratic polynomials<sup>[11]</sup>. The key idea of computing the FDP is to differentiate  $f$  to obtain a set of cubic polynomials and try to recover the right decomposition factors from these cubic polynomials. The idea of differentiation introduced in [11] is a very powerful technique in tackling FDP of multivariate polynomials. In a series of papers<sup>[12–14]</sup>, Faugère and Perret made significant contributions to this problem by integrating the idea of differentiation and fast Gröbner basis computation. These algorithms are always running in polynomial-time, but may fail in certain cases. Faugère and Perret proved that the algorithms are valid for semi-regular multivariate polynomial sets. They also proved the correctness of the algorithm for homogeneous polynomials when the polynomials have (2,2) decomposition and (3,2) decomposition under a genericity assumption<sup>[15]</sup>. Note that their result on the (2,2) decomposition is quite similar to the result presented in Section 4 of this paper, but the corresponding part of our work was independently finished and used a different method. As a consequence, the current known schemes based on FDP of quartic homogeneous multivariate polynomials are considered broken.

As far as we know, the method based on differentiation and homogenization is the only efficient approach to tackle some of the general FDP. But, these algorithms make strong assumptions on the input polynomial sets and these assumptions are expected to be valid and can be removed. This paper focuses on the theoretical analysis of the decomposition algorithm based on differentiation and homogenization.

We prove that the algorithm can be used to compute a degree proper decomposition for a set of randomly decomposable quartic homogeneous polynomials with probability one when the base field is of characteristic zero, and with probability close to one when the base field is a sufficiently large finite field. This solves a conjecture proposed by Ye, Dai, and Lam<sup>[11]</sup>. We also propose a conjecture such that the decomposition for a set of polynomials can be computed from that of its homogenization, and when the polynomials are quartic we prove that the conjecture holds with probability one when the base field is of characteristic zero, and with probability close to one when the base field is a sufficiently large finite field. Finally, we prove that the right decomposition factors for a set of polynomials can be computed from its right decomposition factor space.

The rest of this paper is organized as follows. In Section 2, we give the main result. In Sections 3, 4 and 5, we prove our results for the three major steps of the algorithm. In Section 6, the algorithm is given and its complexity is analyzed. In Section 7, we conclude the paper by proposing two open problems.

## 2 Problem and Main Result

In this section, we will present the problem and give the main results of the paper.

Let  $K$  be a field and  $\mathcal{R} = K[x_1, \dots, x_n]$  the polynomial ring in indeterminates  $x_1, \dots, x_n$  over  $K$ . For natural numbers  $u$  and  $m$ , the functional composition of two sets of multivariate polynomials  $g = (g_1, \dots, g_u) \in K[x_1, \dots, x_m]^u$  and  $h = (h_1, \dots, h_m) \in \mathcal{R}^m$  is a set of polynomials in  $\mathcal{R}^u$ :

$$(f_1, \dots, f_u) = (g_1(h_1, \dots, h_m), \dots, g_u(h_1, \dots, h_m)), \tag{1}$$

that is,

$$f = g \circ h.$$

We call  $g$  and  $h$  the left and right decomposition factors of  $f$ , respectively. The decomposition is called nontrivial if both  $g$  and  $h$  contain nonlinear polynomials.

The functional decomposition problem (FDP) of multivariate polynomials is the inverse of the above functional composition procedure. That is, given a set of  $u$  polynomials  $f = (f_1, \dots, f_u) \in \mathcal{R}^u$  and a positive number  $m$ , to find  $g = (g_1, \dots, g_u) \in K[x_1, \dots, x_m]^u$  and  $h = (h_1, \dots, h_m) \in \mathcal{R}^m$  such that  $f = g \circ h$ .

It is shown that  $f$  always has a nontrivial decomposition when  $m > n$ , which is easy to construct<sup>[2]</sup>. Then we assume that  $1 \leq m \leq n$ . Moreover, note that in cryptosystems, the field  $K$  is usually finite and we usually consider the case that  $m = n$ . So in the following paper, assume that  $m = n$ .

A basic idea of the differentiation approach is to compute the linear space generated by the right factors of  $f$  from the linear space generated by certain differentiations of the polynomials in  $f$ . For a polynomial sequence  $f = (f_1, \dots, f_u) \in \mathcal{R}^u$  with a decomposition like (1), let

$$R_{(f,h)} = \text{span}_K\{h_1, \dots, h_n\}$$

be the linear space generated by  $h_i$  over  $K$ , called a right factor space of  $f$ .

Another idea of the approach is to use homogenization. More precisely, we first compute a decomposition for the homogenization of  $f$  and then try to recover a decomposition of  $f$  from this decomposition. Let  $d_f = \max(d_{f_i})$ ,  $d_g = \max(d_{g_i})$ ,  $d_h = \max(d_{h_i})$ . The homogenizations of  $f, g, h$  are, respectively, defined as follows<sup>[11,13]</sup>:

$$\begin{aligned} f^* &= \left( x_0^{d_f}, x_0^{d_f} f_1\left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0}\right), \dots, x_0^{d_f} f_u\left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0}\right) \right), \\ g^* &= \left( x_0^{d_g}, x_0^{d_g} g_1\left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0}\right), \dots, x_0^{d_g} g_u\left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0}\right) \right), \\ h^* &= \left( x_0^{d_h}, x_0^{d_h} h_1\left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0}\right), \dots, x_0^{d_h} h_n\left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0}\right) \right). \end{aligned}$$

Then the approach proposed in [11–14] can be divided into three major steps which will be explained later.

### Algorithm FDPMP

- 1) Compute a right factor space  $R_{(f^*,h^*)}$  for the homogenization  $f^*$  of  $f$ .
- 2) Compute a right factor space  $R_{(f,h)}$  from  $R_{(f^*,h^*)}$ .
- 3) Compute an FDP for  $f$  from  $R_{(f,h)}$ .

We will show that there exists a complete polynomial time algorithm for Step 3, while for Steps 1, 2, there exist probabilistic algorithms in certain cases. We will discuss Steps 1, 2, 3 in the next three sections.

A decomposition  $f = g \circ h$  satisfying the following condition

$$d_f = d_g \cdot d_h \quad (2)$$

is called a degree proper decomposition, where  $d_f$ ,  $d_g$ , and  $d_h$  are the degrees of  $f$ ,  $g$ , and  $h$ , respectively. All decompositions in this paper are assumed to be degree proper unless mentioned otherwise.

By a homogeneous decomposition  $f = g \circ h$ , we mean that each component of  $f$ ,  $g$ , and  $h$  are homogeneous of the same degree  $d_f$ ,  $d_g$ , and  $d_h$ , respectively. It is clear that a homogeneous decomposition is always degree proper.

Degree proper decompositions can be found by solving a set of nonlinear algebraic equations. But, general equation solving algorithms are usually of exponential complexity. In this paper, we will show that the scheme FDPMP can be developed into a polynomial time decomposition algorithm for certain degree proper decompositions with high probability for random homogeneous polynomials. Here, a set of polynomials  $f$  is called random or randomly decomposable if  $f = g \circ h$  and  $g, h$  are random polynomials.

Before presenting the main results of the paper, we first make some explanation on what the probability means in this paper.

**Remark 1** Let  $N = \binom{n+4}{4} - 1$ . Denote the  $N$ -dimension projective space over  $K$  as  $\mathbb{P}^N(K)$ . Fixed a monomial ranking, the coefficients of homogeneous polynomials with degree 4 can be regarded as the points in  $\mathbb{P}^N(K)$ . Let  $V$  be the set of the points in  $\mathbb{P}^N(K)$  corresponding to homogeneous polynomials with a degree proper decomposition. Then  $V$  is a variety in  $\mathbb{P}^N(K)$ . We will say a statement holds true with probability one if there is a nonempty open subset  $U$  of  $V$  such that the statement is true for all polynomials corresponding to the points in  $U$ . Here we consider the Zariski topology. Also note that a key step of the proof is to use concrete examples to show that the open set is actually nonempty.

We now give the main results of the paper.

**Theorem 1** *Let  $f \in \mathcal{R}^n$  be a set of quartic homogeneous polynomials, each polynomial is of the same degree, for  $n \geq 5$ , we have a polynomial time probabilistic algorithm to find a degree proper decomposition  $f = g \circ h$  for  $g, h \in \mathcal{R}^n$ . For a randomly decomposable  $f$ , the algorithm will give correct result with probability one when  $K$  is of characteristic zero, and with probability close to one when  $K = F_q$  and  $q$  is a sufficiently large number.*

In order to find the decomposition for a set of polynomials from that of its homogenization, we propose the following conjecture.

**Conjecture 1** For all homogeneous decompositions of  $f^* = G \circ H$ , we have  $x_0^{d_H} \in \text{span}_K\{H_0, H_1, \dots, H_n\}$ , where  $f^*$  is the homogenization of  $f$ .

We prove that Conjecture 1 holds with high probability when  $f, G, H$  are of degree 4, 2, 2, respectively and when  $d_G = 2$  and  $d_H > 2$ . Then we have the following theorem.

**Theorem 2** *Let  $f \in \mathcal{R}^n$  be a set of polynomials with degree less than or equal to four, and at least one polynomial has degree four. If we know one of the homogeneous decomposition of  $f^*$ , then we can recover a degree proper decomposition of  $f$  with probability one when  $K$  is of characteristic zero, and with probability close to one when  $K = F_q$  and  $q$  is a sufficiently large number, where  $f^*$  is the homogenization of  $f$ .*

The main idea to prove the above results is to consider the generic FDP. A generic polynomial of degree  $d$  in  $\mathcal{R}$  is of the form  $\sum u_{i_1 \dots i_n} x_1^{i_1} \cdots x_n^{i_n}$  ( $i_1 + \cdots + i_n \leq d$ ), where the  $u_{i_1 \dots i_n}$  are

algebraically independent over  $K$ . An FDP  $f = g \circ h$  is called a generic decomposition if  $g$  and  $h$  are generic polynomials of degrees greater than one.

We will show that if  $f = g \circ h$  is a generic FDP for two quadratic polynomials  $g$  and  $h$ , then we can compute  $g$  and  $h$  with a polynomial number of arithmetic operations in the coefficients fields of  $g$  and  $h$ . Furthermore, when the coefficients of  $g$  and  $h$  specialize to concrete values in the base field  $K$ , the algorithm still works with probability close to one.

### 3 Compute an FDP from a Right Factor Space

In this section, we will show how to compute a decomposition for  $f$  from its right factor space efficiently. We discuss this problem first, because the result in this section will be used in Section 5. Also, among the three steps of the Algorithm FDPMP, this is the only step that has a complete solution.

We first prove several basic properties for  $R_{(f,h)}$ .

**Lemma 1** *Two equivalent decompositions of  $f$  have the same right factor space.*

*Proof* Suppose that  $f$  has two equivalent decompositions  $g \circ h = g' \circ h'$ . By the definition of equivalent decompositions, there exists a nonsingular matrix  $A \in GL_n(K)$  such that  $h' = h \cdot A$ . Therefore,  $\text{span}_K\{h_1, \dots, h_n\} = \text{span}_K\{h'_1, \dots, h'_n\}$ . ■

The following result shows that the FDP of a set of polynomials can be reduced to the FDP of several single polynomials. Denote the set of all right factor spaces of  $F$  by  $SR_F$ .

**Lemma 2** *If  $f = (f_1, \dots, f_u) \in \mathcal{R}^u$ , then*

$$SR_f = \bigcap_{i=1}^u SR_{f_i}. \tag{3}$$

*Proof* It is clear that  $SR_f \subseteq \bigcap_{i=1}^u SR_{f_i}$ . Assume that  $W \in \bigcap_{i=1}^u SR_{f_i}$  and  $h_1, \dots, h_m$  be a basis of  $W$ . Then there are  $g_i \in K[x_1, \dots, x_m]$  such that  $f_i = g_i(h_1, \dots, h_m)$ . Hence,  $W \in SR_f$ . ■

Since computing the intersection of two linear spaces is easy, we may reduce the FDP of  $f$  to the FDP of a single polynomial  $f_i$ .

The approaches in [11–12] are based on the idea of right factor space. But, the power of this idea is not fully explained in previous work. For instance, it is assumed that the rank of  $R_{(f,h)}$  is  $n$  in [13]. It is clear that this condition is not necessarily correct since  $h$  can be a set of arbitrary polynomials. For instance, if  $h = (\sum_{i=1}^n x_i^2, x_2^2, \dots, x_2^2)$  then the rank of  $R_{(f,h)}$  is always two for any decomposition  $f = g \circ h$ .

The following result shows that we can recover a right decomposition factor for  $f$  from  $R_{(f,h)}$  under any condition.

**Theorem 3** *Let  $B = \{b_1, \dots, b_k\}$  be a basis of the right factor space  $R_{(f,h)} = \text{span}_K\{h_1, \dots, h_n\}$ . If  $\dim(R_{(f,h)}) = k = n$ , then  $B$  is a right decomposition factor of  $f$ . If  $\dim(R_{(f,h)}) = k < n$ , then  $(b_1, \dots, b_k, b_1, \dots, b_1)$  is a right decomposition factor of  $f$ .*

*Proof* Firstly, assume that  $\dim(R_{(f,h)}) = n$ . Since  $\{h_1, \dots, h_n\} \in R_{(f,h)}$  and  $B$  is a basis of  $R_{(f,h)}$ , each  $h_i$  can be expressed as a linear combination of  $\{b_1, \dots, b_n\}$ , that is, there exists an invertible matrix  $P \in GL_n(K)$  such that  $(h_1, \dots, h_n) = (b_1, \dots, b_n) \cdot P$ . Then  $f = g \circ h = g(X \cdot P) \circ (h \cdot P^{-1}) = g(X \cdot P) \circ (b_1, \dots, b_n)$ , where  $X = (x_1, \dots, x_n)$ . Therefore,  $B$  is also a right decomposition factor of  $f$ .

Secondly, let  $\dim(R_{(f,h)}) = k < n$ . For the decomposition of  $f = g \circ h$ , since  $\{h_1, \dots, h_n\} \in$

$R_{(f,h)}$  and  $B$  is a basis of  $R_{(f,h)}$ ,  $h_i = \sum_{j=1}^k a_{i,j} b_j$ .

Therefore, we have

$$(h_1, \dots, h_n) = (b_1, \dots, b_k) \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1k} & \cdots & a_{nk} \end{pmatrix},$$

and  $(a_{ij})_{k \times n}$  contains a nonsingular  $k \times k$  submatrix, or else  $\dim(\text{span}_K\{h_1, \dots, h_n\}) < k$ , a contradiction.

Suppose that

$$\det \begin{pmatrix} a_{11} & \cdots & a_{k1} \\ \vdots & \ddots & \vdots \\ a_{1k} & \cdots & a_{kk} \end{pmatrix} \neq 0.$$

Then  $(h_1, \dots, h_n) = (b_1, \dots, b_k, h_{k+1}, \dots, h_n)A$ , where

$$A = \begin{pmatrix} a_{11} & \cdots & a_{k1} & & \\ \vdots & \ddots & \vdots & & \\ a_{1k} & \cdots & a_{kk} & & \\ & & & I_{n-k} & \end{pmatrix}$$

is an  $n \times n$  invertible matrix. Moreover, let

$$B = \begin{pmatrix} & -a_{k+1,1} + 1 & -a_{k+2,1} + 1 & \cdots & -a_{n,1} + 1 \\ & -a_{k+1,2} & -a_{k+2,2} & \cdots & -a_{n,2} \\ I_k & \vdots & \vdots & \ddots & \vdots \\ & -a_{k+1,k} & -a_{k+2,k} & \cdots & -a_{n,k} \\ & & & & I_{n-k} \end{pmatrix}$$

be an  $n \times n$  invertible matrix. It is easy to see that

$$(b_1, \dots, b_k, b_1, \dots, b_1) = (b_1, \dots, b_k, h_{k+1}, \dots, h_n)B.$$

Hence,

$$(h_1, \dots, h_n) = (b_1, \dots, b_k, b_1, \dots, b_1)B^{-1}A.$$

Since  $B^{-1}A$  is nonsingular, there exists a  $g''$  such that  $f = g'' \circ (b_1, \dots, b_k, b_1, \dots, b_1)$  which is an equivalent form of  $f = g \circ h$ . So we can choose  $(b_1, \dots, b_k, b_1, \dots, b_1)$  as a right decomposition factor of  $f$ . ■

Note that the last  $n - k$  elements  $b_1$  in the right factor can be replaced with any  $b_i$  in Theorem 3.

**Corollary 1** *Corresponding to a given right factor space  $R_{(f,h)}$ ,  $f$  has a unique decomposition under the relation of equivalence.*

Restricted to decomposition of quartic polynomials considered in Theorem 2, the following result can be proved easily.

**Theorem 4** *Use the same assumption as Theorem 2. If  $R_{(f,h)}$  is known, we can compute  $g$  with  $O(n^{3\omega})$  arithmetic operations in the field  $K$ , where  $2 \leq \omega < 3$ .*

*Proof* Suppose  $R_{(f,h)} = \text{span}_K(h_1, \dots, h_k)$  is known. Then a right decomposition factor of  $f$  is also known by Theorem 3. To find  $g$ , we may simply by solving a system of linear equations with the coefficients of  $g$  as indeterminates. Note that  $g$  has  $nC_n^2 = O(n^3)$  coefficients. Then we need  $O((n^3)^\omega) = O(n^{3\omega})$  arithmetic operations in  $K$  to find  $g$ , where  $\omega$  is the matrix exponent<sup>[16]</sup> to measure the complexity of solving linear equations. ■

### 4 Decomposition of a Set of Homogenous Polynomials

In this section, we consider the decomposition of  $f$  when each polynomial of it is homogeneous of the same degree. More precisely, we will consider the following problem: “Let  $f$  be a set of quartic homogeneous polynomials. Find a decomposition  $f = g \circ h$  where  $g, h$  are sets of quadratic homogeneous polynomials.”

We may consider the problem in two steps. First, we compute the following linear space over  $K$

$$\tilde{V}_f = \text{span}_K \left\{ \frac{\partial f_i}{\partial x_j} : 1 \leq i \leq u, 1 \leq j \leq n \right\}. \tag{4}$$

Since  $f = g \circ h$  and  $g$  consists of quadratic polynomials, it is clear that  $\tilde{V}_f$  is contained in the following linear space.

$$V_{f,h} = \text{span}_K \{x_i h_j : 1 \leq i, j \leq n\}. \tag{5}$$

The following example shows that  $\tilde{V}_f$  could be a proper subset of  $V_{f,h}$ .

**Example 1** Let  $f = (xy^2z, x^2y^2 + xy^2z, xy^2z + y^2z^2)$ ,  $g = (xz, x^2 + xz, xz + z^2)$ ,  $h = (xy, y^2, yz)$ . It is easy to check that  $f = g \circ h$ . We have  $\tilde{V}_f = \text{span}_K \{xyz, y^2z, yz^2, xy^2, x^2y\}$  and  $V_{f,h} = \text{span}_K \{xyz, y^2z, yz^2, xy^2, x^2y, y^3\}$ .  $\tilde{V}_f$  is a proper subset of  $V_{f,h}$ . Later in this section, we will see that  $h$  cannot be recovered from its corresponding  $\tilde{V}_f$  in this example.

The idea of the algorithm is to compute  $\tilde{V}_f$  first, then try to recover  $V_{f,h}$  from  $\tilde{V}_f$ , and finally compute  $R_{(f,h)}$  from  $V_{f,h}$ . We will analyze the above procedure in the following two subsections. The problem is divided into two cases:  $u = n$  or  $u < n$ .

#### 4.1 The Case When $u = n$

We divide the procedure into two steps: to compute  $V_{f,h}$  from  $\tilde{V}_f$  and to recover  $R_{(f,h)}$  from  $V_{f,h}$ .

##### A. Compute $V_{f,h}$ from $\tilde{V}_f$

When  $u = n$ ,  $\tilde{V}_f$  is generated by  $n^2$  cubic polynomials, and  $\dim(\tilde{V}_f) \leq \dim(V_{f,h}) \leq n^2$ . In the next theorem, we will show that the probability for  $\tilde{V}_f = V_{f,h}$  is close to one under some conditions. The idea of the proof is to find a nonsingular matrix  $A$  in some indeterminates such that if a set of specialization of these indeterminates does not vanish  $|A|$  then  $\tilde{V}_f = V_{f,h}$ .

The following lemma will be used.

**Lemma 3**<sup>[17]</sup> For  $f \in F_q[x_1, \dots, x_n]$  with degree  $\deg(f) = d \geq 0$ . Then the equation  $f(x_1, \dots, x_n) = 0$  has at most  $dq^{n-1}$  solutions in  $F_q^n$ .

**Theorem 5** For randomly chosen  $g$  and  $h$ , let  $f = g \circ h$ . Then

- 1)  $\tilde{V}_f = V_{f,h}$  with probability one when the field  $K$  is of characteristic zero.
- 2)  $\tilde{V}_f = V_{f,h}$  with probability close to one when  $K = F_q$  and  $q$  is sufficiently large.

*Proof* Assume that

$$f_i = \sum_{1 \leq k, l \leq n} a_{k,l}^{(i)} h_k h_l, \quad 1 \leq i \leq n,$$

where  $a_{k,l}^{(i)} = a_{l,k}^{(i)}$  for  $1 \leq k, l \leq n$ , and

$$h_i = \sum_{1 \leq k \leq l \leq n} b_{k,i}^{(i)} x_k x_l, \quad 1 \leq i \leq n.$$

Then

$$\frac{\partial f_i}{\partial x_j} = \sum_{1 \leq k, l \leq n} a_{k,l}^{(i)} \left( h_k \frac{\partial h_l}{\partial x_j} + h_l \frac{\partial h_k}{\partial x_j} \right).$$

Let

$$U_i = \left( \frac{\partial f_1}{\partial x_i}, \frac{\partial f_2}{\partial x_i}, \dots, \frac{\partial f_n}{\partial x_i} \right), V_i = (x_i h_1, x_i h_2, \dots, x_i h_n), \text{ for } i = 1, 2, \dots, n.$$

Let  $U = (U_1, U_2, \dots, U_n)^T$  and  $V = (V_1, V_2, \dots, V_n)^T$ . Each  $\frac{\partial f_i}{\partial x_j}$  can be represented by a linear combination of  $\{x_k h_l, 1 \leq k, l \leq n\}$  over  $K$  and the coefficients are expressions in  $a_{k,l}^{(i)}, b_{k,l}^{(i)}$ . So, there exists an  $n^2 \times n^2$  matrix  $A$  such that  $U = A \cdot V$ , where the elements of  $A$  are polynomials in  $a_{k,l}^{(i)}, b_{k,l}^{(i)}$ . We will prove the  $\det(A) \neq 0$ . We make the following substitutions in  $A$ :  $a_{k,l}^{(i)} = (k+l)^i$  and

$$b_{k,l}^{(i)} = \begin{cases} 1, & \text{if } i = k = l; \\ 0, & \text{or else;} \end{cases}$$

and denote the new matrix by  $\bar{A}$ . After making these substitutions, one has  $f_i = \sum_{k,l} (k+l)^i x_k^2 x_l^2$  and  $h_i = x_i^2$  for  $1 \leq k, l, i \leq n$ . Now, we have

$$\frac{\partial f_i}{\partial x_s} = 4 \sum_{k=1}^n (s+k)^i x_s x_k^2, \quad \text{for } i, s = 1, 2, \dots, n,$$

which implies that for all  $s = 1, 2, \dots, n$ ,

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_s} \\ \frac{\partial f_2}{\partial x_s} \\ \vdots \\ \frac{\partial f_n}{\partial x_s} \end{pmatrix} = \begin{pmatrix} 4(1+s) & 4(2+s) & \cdots & 4(n+s) \\ 4(1+s)^2 & 4(2+s)^2 & \cdots & 4(n+s)^2 \\ \vdots & \vdots & \ddots & \vdots \\ 4(1+s)^n & 4(2+s)^n & \cdots & 4(n+s)^n \end{pmatrix} \cdot \begin{pmatrix} x_s x_1^2 \\ x_s x_2^2 \\ \vdots \\ x_s x_n^2 \end{pmatrix}. \tag{6}$$

Therefore,  $\det(\bar{A})$  is the products of a constant and  $n$  Vandermonde determinants, which is nonzero. Hence  $\det(A) \neq 0$ . One can easily see that the total degree of  $\det(A)$  in  $a_{k,l}^{(i)}, b_{k,l}^{(i)}$  equals  $2n^2$ .

When  $g$  and  $h$  specialize to concrete polynomials in  $\mathcal{R}^n$ , if  $A$  is invertible then each element of  $V_{f,h}$  can be represented by a linear combination of the elements of  $\tilde{V}_f$ . So,  $V_{f,h} = \tilde{V}_f$ . When  $K$  is of characteristic zero,  $\det(A) \neq 0$  with probability one in the sense explained in Remark 1. When  $K = F_q$ , by Lemma 3,  $\det(A) \neq 0$  with probability at least  $\frac{q-d}{q} = \frac{q-2n^2}{q}$  which is close



to one when  $q$  is sufficiently large, where  $d$  is the degree of the equation. These conclude the theorem. ■

When  $\tilde{V}_f \neq V_{f,h}$ , Ye, et al.<sup>[11]</sup> proposed a heuristic method to enlarge  $\tilde{V}_f$ , but there is no theoretical guarantee that the enlarged  $\tilde{V}_f$  is equal to  $V_{f,h}$ .

**B. Recover  $R_{(f,h)}$  from  $V_{f,h}$**

In this subsection, we assume that the space  $V_{f,h}$  is known and show how to recover  $R_{(f,h)}$  from  $V_{f,h}$ . Given a vector space  $V \subseteq K[x_1, \dots, x_n]$  and a set  $S \subseteq K[x_1, \dots, x_n]$ , we define  $(V : S) = \{h | \forall s \in S, sh \in V\}$ .

By the definition of  $V_{f,h}$ ,  $x_i h_j \in V_{f,h}$  for all  $i, j$ . Hence,  $h_j \in (V_{f,h} : x_i)$ , and then  $R_{(f,h)} \subseteq (V_{f,h} : x_i)$ , for all  $i$ . So we have

$$R_{(f,h)} \subseteq \cap_i (V_{f,h} : x_i) = (V_{f,h} : L),$$

where  $L$  is the linear space generated by the variables  $x_1, \dots, x_n$ .

Note that  $R_{(f,h)} \subseteq (\tilde{V}_f : L)$  does not always hold. In Example 1,  $(\tilde{V}_f : L) = \{yz, xy\}$  while  $R_{(f,h)} = \{yz, xy, y^2\}$ .  $(\tilde{V}_f : L)$  is a proper subset of  $R_{(f,h)}$ . However, by Theorem 5, in the general case,  $R_{(f,h)} \subseteq (V_{f,h} : L) = (\tilde{V}_f : L)$  with probability one when  $K$  is of characteristic zero and close to one when  $K = F_q$  and  $q$  is sufficiently large.

One may ask that whether  $R_{(f,h)} = (V_{f,h} : L)$ . It is not always true as shown by the following example.

**Example 2** Let  $f = (x^2y^2, x^4 + y^4)$ ,  $g = (xy, x^2 + y^2)$  and  $h = (x^2, y^2)$ .  $(V_{f,h} : L) = \text{span}_K\{xy, x^2, y^2\}$ .  $R_{(f,h)}$  is a proper subset of  $(V_{f,h} : L)$ .

Ye, et al.<sup>[11]</sup> proposed a conjecture which suggests that for random  $R_{(f,h)}$ , the two spaces are equal with probability close to one no matter whether  $\dim(R_{(f,h)}) = n$  or  $\dim(R_{(f,h)}) < n$ . The conjecture is as follows:

**Conjecture Y<sup>[11]</sup>** Let  $W$  be a linear space of dimension  $\leq n$  consisting of quadratic forms in  $n$  variables  $x_1, \dots, x_n$ , and  $L$  be the linear space generated by  $x_1, \dots, x_n$ ,  $V = \sum_{1 \leq i \leq n} x_i W$ .

For randomly chosen  $W$ , the probability  $\rho$  that  $(V : L) = W$  is very close to one when  $n > 2$ .

It is one of the theoretical foundations of the differentiation approach. Ye, et al.<sup>[11]</sup> did not prove it and just gave a justification with some heuristic arguments. The work of Faugère and Perret is also based on this basic fact. When the number of basis for  $(V_{f,h} : L)$  equals  $n$ , they regarded  $(V_{f,h} : L)$  as  $R_{(f,h)}$  in their algorithm<sup>[12–13]</sup>.

We will give a proof of the conjecture when  $n \geq 5$ . Actually, we will extend the conjecture into a more general case that  $W$  and  $L$  are linear spaces consisting of homogeneous polynomials with higher degree and give a proof for this extension of the conjecture. The assumption  $n \geq 5$  is not a strict limitation since in practical usages,  $n$  is much larger than five. The number  $q$  is always large in  $2R$  or  $2R^-$  scheme<sup>[12–13]</sup>. Before proving the conjecture, we need a technical lemma. Let  $P = (p_1, \dots, p_n) \in \mathbb{N}^n$ . In the following, we will always use  $X^P$  to denote the monomial  $x_1^{p_1} \dots x_n^{p_n}$  and  $M(d', x_1, \dots, x_n)$  to denote the set of all monomials in  $x_1, \dots, x_n$  with degree  $d'$ .

**Lemma 4** Assume that  $h_i = \sum_{|P|=d} a_P^{(i)} X^P \in K[a_P^{(i)}][x_1, \dots, x_n]$  are homogeneous polynomials in  $x_1, \dots, x_n$  with degree  $d$ , where  $i = 1, 2, \dots, n + 1$  and  $a_P^{(i)} \in K$ . Assume that  $d' < d$  and  $n \geq 2d$ . Then if  $\{mh_i | m \in M(d', x_1, \dots, x_n), i = 1, 2, \dots, n + 1\}$  are linearly dependent over  $K$ , then  $(a_P^{(i)})$  will vanish a set of nonzero polynomials with total degree at most  $n \binom{n+d'-1}{d'}$ .

*Proof* Let us consider  $a_P^{(i)}$  as indeterminates for a moment. Assume that  $H = \sum c_{m,i} mh_i$

where  $c_{m,i}$  are indeterminates. Regarding  $H$  as a polynomial in  $x_1, \dots, x_n$ , one can see that  $H$  is a polynomial with  $\binom{n+d+d'-1}{d+d'}$  monomials whose coefficients are polynomials in  $c_{m,i}, a_P^{(i)}$ . Setting  $H = 0$ , one can get a system of the equations as follows:  $A\vec{c} = 0$ , where  $A$  is a  $\binom{n+d+d'-1}{d+d'}$  by  $n\binom{n+d'-1}{d'}$  matrix with entries linearly in the  $a_P^{(i)}$ , and  $\vec{c} = (c_{m_1,1}, \dots, c_{m_j,i}, \dots)^T$ . By the computation, one can show that  $\binom{n+d+d'-1}{d+d'} > n\binom{n+d'-1}{d'}$ . Hence,  $A$  is of full rank if and only if  $\{mh_i | m \in M(d', x_1, \dots, x_n), i = 1, 2, \dots, n+1\}$ , are linearly independent. To prove  $A$  is of full rank, one only need to prove this for a specialization of  $A$ . Since  $n \geq 2d$ , let  $h_1 = x_1^d, h_2 = x_2^d, \dots, h_n = x_n^d, h_{n+1} = x_1x_2 \cdots x_d + x_{d+1}x_{d+2} \cdots x_{2d}$ . It leads to a specialization of the matrix  $A$ . Denote this specialization by  $\bar{A}$ . We claim that  $\bar{A}$  is of full rank, which is equivalent to claim that the polynomials  $mx_j^d, m(x_1 \cdots x_d + x_{d+1} \cdots x_{2d}), m \in M(d', x_1, \dots, x_n), j = 1, 2, \dots, n$ , are linearly independent. Assume that

$$\bar{H} = \sum_{m,i} \bar{c}_{m,i} mx_i^d + \sum_m \bar{b}_m m(x_1 \cdots x_d + x_{d+1} \cdots x_{2d}) = 0,$$

where  $\bar{c}_{m,i}, \bar{b}_m \in K$ . For convenience, denote  $\partial x_1^{q_1} \cdots \partial x_n^{q_n}$  by  $\partial m$  where  $m = x_1^{q_1} \cdots x_n^{q_n}$ . When  $d' < d$  and  $n \geq 2d$ , one can see that

$$\begin{aligned} \frac{\partial^{d+d'}(\bar{H})}{\partial mx_i^d} &= \begin{cases} * \bar{c}_{m,i}, & \forall m' \in M, m'x_1 \cdots x_d \neq mx_i^d \text{ and } m'x_{d+1} \cdots x_{2d} \neq mx_i^d; \\ * \bar{c}_{m,i} + * \bar{b}_{m'}, & \exists m' \in M \text{ s.t. } m'x_1 \cdots x_d = mx_i^d \text{ or } m'x_{d+1} \cdots x_{2d} = mx_i^d; \end{cases} \\ \frac{\partial^{d+d'}(\bar{H})}{\partial mh} &= \begin{cases} * \bar{b}_m, & h = x_1 \cdots x_d \text{ and } \forall m' \in M \forall i, mx_1 \cdots x_d \neq m'x_i^d; \\ * \bar{c}_{m',i} + * \bar{b}_m, & h = x_{d+1} \cdots x_{2d} \text{ and } \exists m' \in M \exists i \text{ s.t. } mx_1 \cdots x_d = m'x_i^d; \end{cases} \end{aligned}$$

where  $*$  denote positive integers. Since  $\frac{\partial^{d+d'}(\bar{H})}{\partial \bar{m}} = 0$  for all monomials  $\bar{m}$ , the claim is proved. Therefore,  $A$  is of full rank. Now, consider the  $a_P^{(i)}$  as the elements in  $K$ . If  $\{mh_i | m \in M(d', x_1, \dots, x_n), i = 1, 2, \dots, n+1\}$  are linearly dependent, which implies that  $A\vec{c} = 0$  has a nontrivial solution, then  $(a_P^{(i)})$  must vanish the determinants of all  $n\binom{n+d'-1}{d'}$  by  $n\binom{n+d'-1}{d'}$  submatrices of  $A$ . This completes the proof. ▀

Let  $h = (h_1, h_2, \dots, h_n)$ , where the  $h_i$  are homogeneous polynomials with the same degree in  $K[x_1, x_2, \dots, x_n]$  and let  $d_h$  be the degree of  $h_i$ . Denote

$$U(h, d') = \text{span}_K \{mh_i | m \in M(d', x_1, \dots, x_n), i = 1, 2, \dots, n\}.$$

Let  $W = \text{span}_K \{h_1, h_2, \dots, h_n\}$ . Then we have

**Theorem 6** *For randomly chosen  $h_1, h_2, \dots, h_n$ , if  $d' < d_h$  and  $n > 2d_h$ , then the probability  $\rho$  that  $(U(h, d') : x_i^{d'}) = W$  is one when the field  $K$  is of characteristic zero and close to one when  $K = F_q$  with  $q$  sufficiently large, where  $i = 1, 2, \dots, n$ .*

*Proof* It suffices to prove the theorem for the case  $i = 1$ . The other cases can be proved by the similar way.

Assume that  $h_i = \sum_{|P|=d_h} a_P^{(i)} X^P \in K[x_1, \dots, x_n]$ , where the  $a_P^{(i)} \in K$ . Denote  $\bar{U} = \{H \in U(h, d') | x_1^{d'} | H\}$ . For  $\sum_i G_i h_i \in \bar{U}$ , let

$$G_i = \tilde{G}_{0,i} x_1^{d'} + \tilde{G}_{1,i} x_1^{d'-1} + \cdots + \tilde{G}_{d'-2,i} x_1^2 + \tilde{G}_{d'-1,i} x_1 + \tilde{G}_{d',i}$$

and

$$h_i = \tilde{h}_{0,i} x_1^{d_h} + \tilde{h}_{1,i} x_1^{d_h-1} + \cdots + \tilde{h}_{d_h-2,i} x_1^2 + \tilde{h}_{d_h-1,i} x_1 + \tilde{h}_{d_h,i},$$

where  $\tilde{G}_{0,i}, \tilde{G}_{1,i}, \dots, \tilde{G}_{d',i}$  are homogeneous polynomials in  $x_2, \dots, x_n$  with degrees  $0, 1, \dots, d'$ , respectively and  $\tilde{h}_{0,i}, \tilde{h}_{1,i}, \dots, \tilde{h}_{d_h,i}$  are homogeneous polynomials in  $x_2, \dots, x_n$  with degrees  $0, 1, \dots, d_h$ , respectively. Since  $\sum_i G_i h_i \equiv 0 \pmod{x_1^{d'}}$ , we have

$$\begin{aligned} \sum_i G_i h_i &\equiv \sum_i \left( x_1^{d'-1} \left( \tilde{G}_{1,i} \tilde{h}_{d_h,i} + \tilde{G}_{2,i} \tilde{h}_{d_h-1,i} + \dots + \tilde{G}_{d',i} \tilde{h}_{d_h-d'+1,i} \right) \right. \\ &\quad \left. + x_1^{d'-2} \left( \tilde{G}_{2,i} \tilde{h}_{d_h,i} + \tilde{G}_{3,i} \tilde{h}_{d_h-1,i} + \dots + \tilde{G}_{d',i} \tilde{h}_{d_h-d'+2,i} \right) + \dots \right. \\ &\quad \left. + x_1^2 \left( \tilde{G}_{d'-2,i} \tilde{h}_{d_h,i} + \tilde{G}_{d'-1,i} \tilde{h}_{d_h-1,i} + \tilde{G}_{d',i} \tilde{h}_{d_h-2,i} \right) \right. \\ &\quad \left. + x_1 \left( \tilde{G}_{d'-1,i} \tilde{h}_{d_h,i} + \tilde{G}_{d',i} \tilde{h}_{d_h-1,i} \right) + \tilde{G}_{d',i} \tilde{h}_{d_h,i} \right) \\ &\equiv 0 \pmod{x_1^{d'}}. \end{aligned}$$

Therefore,

$$\sum_i \left( \tilde{G}_{1,i} \tilde{h}_{d_h,i} + \tilde{G}_{2,i} \tilde{h}_{d_h-1,i} + \dots + \tilde{G}_{d',i} \tilde{h}_{d_h-d'+1,i} \right) = 0, \tag{7}$$

$$\sum_i \left( \tilde{G}_{2,i} \tilde{h}_{d_h,i} + \tilde{G}_{3,i} \tilde{h}_{d_h-1,i} + \dots + \tilde{G}_{d',i} \tilde{h}_{d_h-d'+2,i} \right) = 0, \tag{8}$$

$$\begin{aligned} &\vdots \\ \sum_i \left( \tilde{G}_{d'-2,i} \tilde{h}_{d_h,i} + \tilde{G}_{d'-1,i} \tilde{h}_{d_h-1,i} + \tilde{G}_{d',i} \tilde{h}_{d_h-2,i} \right) &= 0, \tag{9} \end{aligned}$$

$$\sum_i \left( \tilde{G}_{d'-1,i} \tilde{h}_{d_h,i} + \tilde{G}_{d',i} \tilde{h}_{d_h-1,i} \right) = 0, \tag{10}$$

$$\sum_i \tilde{G}_{d',i} \tilde{h}_{d_h,i} = 0. \tag{11}$$

Assume that for each  $1 \leq k \leq d'$ ,  $\{m \tilde{h}_{d_h,i} | m \in M(k, x_2, \dots, x_n), i = 2, \dots, n\}$  are linearly independent. Then by the equalities (7)–(11), one has  $\tilde{G}_{j,i} = 0$  for  $j = 1, 2, \dots, d'$  and  $i = 1, 2, \dots, n$ . Therefore  $\bar{U} = \{\sum_i \tilde{G}_{0,i} h_i\} \subseteq W$ . Note that  $(U(h, d') : x_1^{d'}) = \bar{U}$ . Hence  $(U(h, d') : x_1^{d'}) = W$ . By Lemma 4, the  $a_p^{(i)}$  such that for some  $k \leq d'$ ,  $\{m h_i | m \in M(k, x_2, \dots, x_n), i = 1, 2, \dots, n\}$  are linearly dependent are the zeroes of some polynomials with degree at most  $(n-1) \binom{n+d+k-2}{d+k} (\leq (n-1) \binom{n+d+d'-2}{d+d'} \triangleq N)$ .

Hence, when  $K$  is of characteristic zero, the probability that  $(U(h, d') : x_1^{d'}) = W$  is one; when  $K = F_q$ , by Lemma 3 the probability that  $(U(h, d') : x_1^{d'}) = W$  is at least  $\frac{q-N}{q}$  which is close to one when  $q$  is sufficiently large<sup>[17]</sup>. ■

**Remark 2** In general, when  $K$  is algebraically closed, Theorem 6 does not hold for sufficiently large integer  $d'$ . For randomly chosen  $h_1, \dots, h_n$ , the set of zeroes of  $\{h_1, \dots, h_n\}$  in  $\mathbb{P}(K)^{n-1}$  is empty, where  $\mathbb{P}(K)^{n-1}$  is  $n-1$  dimension projective space over  $K$ . Then by the Projective Weak Nullstellensatz Theorem<sup>[18]</sup>, there is some integer  $r$  such that  $\langle x_1, \dots, x_n \rangle^r \subseteq \langle U(h, r-d_h) \rangle$ . Let  $d' = r-d_h$ . Then  $M(r, x_1, \dots, x_n) \subseteq U(h, d')$ , which implies that  $M(d_h, x_1, \dots, x_n) \subseteq (U(h, d') : x_1^{d'})$ . However, in general,  $W \neq \text{span}_K(M(d_h, x_1, \dots, x_n))$ .

**Corollary 2** Conjecture Y is correct over  $K$  when  $n \geq 5$ , where  $K$  is of characteristic zero or is a finite field consisting of a sufficiently large number of elements.

As a consequence of Theorem 5 and Corollary 2, we have the following result.

**Theorem 7** *If  $f$  is a randomly decomposable and  $n \geq 5$ , then  $(\tilde{V}_f : L) = R_{(f,h)}$  with probability one when  $K$  is of characteristic zero and with probability close to one when  $q$  is sufficiently large where  $K = F_q$ .*

Therefore, we can recover  $R_{(f,h)}$  from  $\tilde{V}_f$  directly with high probability if the FDP of  $f$  is randomly chosen.

Faugère and Perret assumed that  $\tilde{V}_f = V_{f,h}$  in their papers, since they assumed that the decomposition is random, the dimension of  $R_{(f,h)}$  spanned by  $h_1, \dots, h_n$  is  $n$ , and  $\dim(\tilde{V}_f) \geq \dim(V_{f,h})^{[12]}$ .

**Theorem 8** *Under the same assumptions as Theorem 7. If  $\tilde{V}_f$  is known, we can compute  $R_{(f,h)}$  with complexity  $O(n^{3\omega})$  arithmetic operations in  $K$  with probability one when  $K$  is of characteristic zero and with probability close to one when  $q$  is sufficiently large when  $K = F_q$ .*

*Proof* It suffices to randomly choose a linear polynomial  $l$  in  $x_1, x_2, \dots, x_n$  and compute  $(\tilde{V}_f : l)$  to obtain  $R_{(f,h)}$ . Without loss of generality, assume that  $l = x_1 + c_2x_2 + \dots + c_nx_n$ . Let  $y_1 = l, y_i = x_i, i = 2, 3, \dots, n$ , that is,

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & c_2 & \cdots & c_n \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \tag{12}$$

Then there exists a matrix  $M_l \in GL_K(n)$  such that

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = M_l \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \tag{13}$$

where

$$M_l = \begin{pmatrix} 1 & -c_2 & \cdots & -c_n \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}. \tag{14}$$

Denote it by  $X = M_l \cdot Y$ .

For all  $f \in \mathcal{R}$ , define  $M_l(f) = f|_{X=M_l \cdot Y}$ ,  $M_l^{-1}(g) = g|_{Y=M_l^{-1} \cdot X}$ , where  $g \in K[y_1, \dots, y_n]$ . Then  $M_l^{-1}M_l(f) = f$ , and  $M_l(f_1f_2) = M_l(f_1)M_l(f_2)$ . So  $M_l(l) = l|_{X=M_l \cdot Y} = y_1$ . Let  $M_l(\tilde{V}_f) = \{p|_{X=M_l \cdot Y} : \text{for all } p \in \tilde{V}_f\}$ .

Then we have  $r \in (\tilde{V}_f : l) \Leftrightarrow rl \in \tilde{V}_f \Leftrightarrow M_l(rl) \in M_l(\tilde{V}_f) \Leftrightarrow M_l(r)M_l(l) \in M_l(\tilde{V}_f) \Leftrightarrow M_l(r) \in (M_l(\tilde{V}_f) : M_l(l)) \Leftrightarrow M_l(r) \in (M_l(\tilde{V}_f) : y_1) \Leftrightarrow r \in M_l^{-1}(M_l(\tilde{V}_f) : y_1)$ , that is,  $(\tilde{V}_f : l) = M_l^{-1}(M_l(\tilde{V}_f) : y_1)$ .

So in order to compute  $(\tilde{V}_f : l)$ , we can first transform the polynomials in  $\tilde{V}_f$  by a nonsingular coordinate substitution  $X = M_l \cdot Y$  to obtain  $M_l(\tilde{V}_f)$ , and then compute  $(M_l(\tilde{V}_f) : y_1)$ . Finally, transform  $(M_l(\tilde{V}_f) : y_1)$  to  $(\tilde{V}_f : l)$  by the inverse transformation  $Y = M_l^{-1} \cdot X$ . The main arithmetic complexity relies on the computation of  $(M_l(\tilde{V}_f) : y_1)$ .

We construct a matrix  $S$  to represent the polynomials of  $M_l(\tilde{V}_f)$  in a basis of monomials of degree three. Each row of  $S$  corresponds to the coefficients of each polynomial of  $M_l(\tilde{V}_f)$  with respect to the monomials of degree three. Suppose that the monomials are sorted so that the last  $n(n + 1)/2$  columns of  $S$  correspond to monomials which can be divided by  $y_1$ . Then perform linear elimination to  $S$ , we can obtain polynomials which can be divided by  $y_1$ , denoted by  $t_i, i = 1, 2, \dots, k$ , if  $n \geq 5$ , then  $k \leq n^{[14]}$ . Then  $(M_l(\tilde{V}_f) : y_1) = \{t_i/y_1, i = 1, 2, \dots, k\}$ . Note that  $S$  is an  $n^2 \times C_{n+2}^3$  matrix. Then we need  $O((n^3)^\omega) = O(n^{3\omega})$  arithmetic operations to compute  $(M_l(\tilde{V}_f) : y_1)$ . The whole arithmetic complexity of computing  $(\tilde{V}_f : l)$  is also  $O(n^{3\omega})$ . ■

4.2 The Case when  $u < n$

We now consider the case of  $u < n$ . In this case, Faugère and Perret extended  $\tilde{V}_f$  and  $V_{f,h}$  to new linear spaces  $\tilde{V}_{fd}$  and  $V_{fd}$ :

$$\tilde{V}_{fd} = \text{span}_K \left\{ m \frac{\partial f_i}{\partial x_j} : m \in M(d), 1 \leq i \leq u, 1 \leq j \leq n \right\}, \tag{15}$$

$$V_{fd} = \text{span}_K \{ m' h_j : m' \in M(d + 1), 1 \leq i, j \leq n \}, \tag{16}$$

where  $M(d)$  represents the set of monomials of degree  $d$ . It is obvious that  $\tilde{V}_{fd} \subseteq V_{fd}$ . Faugère and Perret<sup>[12–13]</sup> required  $\dim(\tilde{V}_{fd}) \geq \dim(V_{fd})$  by choosing a proper integer  $d$ , which means  $\tilde{V}_{fd} = V_{fd}$ .

Assume that  $V_{fd}$  is known, and we try to recover  $R_{(f,h)}$  from  $V_{fd}$ . By the definition of  $V_{fd}$ ,  $mh_j \in V_{fd}$  for all  $m \in M(d+1)$  and  $j$ . Hence,  $h_j \in (V_{fd} : x_i^{d+1})$ , and then  $R_{(f,h)} \subseteq (V_{fd} : x_i^{d+1})$ , for all  $i$ . Hence,  $R_{(f,h)} \subseteq \cap_i (V_{fd} : x_i^{d+1})$ . The approach in [12–14] makes use of this property, and recovers  $R_{(f,h)}$  by computing the quotient  $(V_{fd} : x_i^{d+1})$  for some  $i$ . Faugère and Perret<sup>[12–14]</sup> chose that  $i = n$ .

In the case that  $d_g = d_h = 2$ , Theorem 6 fails if  $u < n$ .

However, in the general case, if the degrees of  $g$  and  $h$  are more than 2, then from Theorem 6, we can obtain  $R_{(f,h)}$  by computing the quotient  $(V_{fd} : x_i^{d+1})$  when  $d + 1 < d_h$ .

From the above discussion, we can see that the results listed above provide a theoretical guarantee for the previous work [11–14] in certain sense.

5 Recover the Decomposition of  $f$  from  $f^*$

In this section, we study the relationship between the FDPs of a set of polynomials  $f$  and that of its homogenization  $f^*$ . We will show that with high probability, we can recover a decomposition for  $f$  from a decomposition of  $f^*$ .

For a general FDP  $f = g \circ h$ , the following result gives the connection between the FDP of  $f$  and the FDP of its homogenization  $f^*$ .

**Lemma 5** *If  $f = g \circ h$ , then  $x_0^{d_g d_h - d_f} f^* = g^* \circ h^*$ , where  $d_g, d_h, d_f$  are the degrees of  $g, h, f$ , respectively.*

*Proof* If  $f = g \circ h$ , we have  $d_g \cdot d_h \geq d_f$ . Hence,

$$f_i \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right) = g_i \left( h_1 \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right), \dots, h_n \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right) \right).$$

Then

$$\begin{aligned}
 g^* \circ h^* &= \left( x_0^{d_g d_h}, x_0^{d_g d_h} g_1 \left( h_1 \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right), \dots, x_0^{d_g d_h} h_n \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right) \right), \right. \\
 &\quad \left. \dots, x_0^{d_g d_h} g_u \left( h_1 \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right), \dots, h_n \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right) \right) \right) \\
 &= \left( x_0^{d_g d_h}, x_0^{d_g d_h} f_1 \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right), \dots, x_0^{d_g d_h} f_u \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right) \right) \\
 &= x_0^{d_g d_h - d_f} \left( x_0^{d_f}, x_0^{d_f} f_1 \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right), \dots, x_0^{d_f} f_u \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right) \right) \\
 &= x_0^{d_g d_h - d_f} f^*.
 \end{aligned}$$

Then the proof is completed. █

As a consequence, we have  $(f \circ g)^* = f^* \circ g^*$  if  $d_f \cdot d_g = d_h^{[11,13]}$ .

The following result gives a necessary and sufficient condition for  $f$  to have an FDP in terms of its homogenization  $f^*$ .

**Theorem 9** *Let  $f = (f_1, \dots, f_u) \in \mathcal{R}^u$ . Then,  $f$  has a decomposition if and only if there exist natural numbers  $s, t$  such that  $x_0^s f^* = g' \circ h'$  is a homogeneous decomposition and  $x_0^t \in \text{span}_K \{h'_0, \dots, h'_n\}$ .*

*Proof* If  $f$  has a decomposition  $f = g \circ h$ , let  $s = x_0^{d_g d_h - d_f}, g' = g^*, h' = h^*, t = d_h$  in Lemma 5. Then the conclusion holds. █

We now prove the other direction. If there are natural numbers  $s, t$  such that  $x_0^s f^* = g' \circ h'$  is a homogeneous decomposition and  $x_0^t \in \text{span}_K \{h'_0, \dots, h'_n\}$ , then  $\text{deg}(h') = t, \text{deg}(g') = \frac{s+d_f}{t}$ , and we can choose  $g', h'$  such that  $x_0^s f^*$  has the following homogeneous decomposition form by Theorem 3:

$$\begin{aligned}
 x_0^s f^* &= \left( x_0^{s+d_f}, x_0^{s+d_f} f_1 \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right), \dots, x_0^{s+d_f} f_u \left( \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right) \right) \\
 &= \left( x_0^{\frac{s+d_f}{t}}, g'_1, \dots, g'_u \right) \circ \left( x_0^t, h'_1, \dots, h'_n \right)
 \end{aligned}$$

and  $\text{deg}(g'_i) = \frac{s+d_f}{t}, \text{deg}(h'_i) = t$ . Let  $x_0 = 1$ . We have

$$\begin{aligned}
 f &= (f_1, \dots, f_u) \\
 &= (g'_1(1, x_1, \dots, x_n), \dots, g'_u(1, x_1, \dots, x_n)) \circ (h'_1(1, x_1, \dots, x_n), \dots, h'_n(1, x_1, \dots, x_n)),
 \end{aligned}$$

which is a decomposition of  $f$ . █

As a consequence of Lemma 5 and Theorem 9, we have

**Corollary 3** *Let  $f = (f_1, \dots, f_u) \in \mathcal{R}^u$ . Then,  $f$  has a degree proper decomposition if and only if there is a natural number  $t$  such that  $f^*$  has a homogeneous decomposition  $f^* = g' \circ h'$  and  $x_0^t \in \text{span}_K \{h'_0, \dots, h'_n\}$ .*

In order to use the idea of homogenization, we need to solve the following problem.

**Problem 1** For all homogeneous decompositions of  $f^* = G \circ H$ , whether  $x_0^{d_H} \in \text{span}_K \{H_0, \dots, H_n\}$ .

Then we propose Conjecture 1, which means that the problem has a positive answer.

If the conjecture is true, we may conclude that to compute a degree proper decomposition of  $f$  is equivalent to compute a homogeneous decomposition of  $f^*$ . Therefore, we can obtain a

right factor space  $R_{(f,h)}$  of  $f$  from  $R_{(f^*,h^*)}$  in the same way with the method in the proofs of Theorem 9.

First, we show that Conjecture 1 is valid in a very special case.

**Theorem 10** *Conjecture 1 has a positive answer in the field of complex numbers if the degrees of  $f^*$ ,  $G$  and  $H$  are 4, 2, 2 respectively and  $n = 2$ .*

*Proof* In the field  $K = \mathbb{C}$ , if  $G$  is nondegenerate, we can assume that  $G$  has the following standard form  $G = x_0^2 + x_1^2 + x_2^2$  by nonsingular linear substitution<sup>[19]</sup> (If  $G$  is degenerate, then we can assume that  $G = x_0^2 + x_1^2$  or  $G = x_0^2$ , it is easy to see that the Conjecture holds in either case).

Firstly, we claim that we can assume  $H_0 = x_0^2 + c_0$ ,  $H_1 = b_1x_0 + c_1$ , and  $H_2 = b_2x_0 + c_2$ , where  $c_i$  are quadratic homogeneous polynomials and  $b_i$  are linear homogeneous polynomials in variables  $x_1$  and  $x_2$ . Since we consider the decomposition over the field of complex numbers, we may assume that  $H_k = a_kx_0^2 + G_k$  ( $k = 0, 1, 2$ ), where  $G_k$  does not contain  $x_0^2$ . Since  $x_0^4 = H_0^2 + H_1^2 + H_2^2$ ,  $a_0^2 + a_1^2 + a_2^2 = 1$ . Without loss of generality, we may assume  $a_0^2 + a_1^2 \neq 0$ . Let  $H'_0 = \frac{a_1H_1}{\sqrt{a_0^2+a_1^2}} + \frac{a_0H_0}{\sqrt{a_0^2+a_1^2}}$  and  $H'_1 = \frac{a_0H_1}{\sqrt{a_0^2+a_1^2}} - \frac{a_1H_0}{\sqrt{a_0^2+a_1^2}}$ . We have

$$H_0^2 + H_1^2 = (H'_0)^2 + (H'_1)^2$$

and  $H'_1$  does not contain the term  $x_0^2$ . Repeat the above procedure one more time, we obtain three new polynomials  $H''_0, H''_1, H''_2$  such that  $H''_1$  and  $H''_2$  do not contain  $x_0^2$ . Since  $x_0^4 = H''_0^2 + H''_1^2 + H''_2^2$ , we have  $H''_0 = x_0^2 + b_0x_0 + c_0$ . Comparing the coefficients of  $x_0^3$ , we have  $b_0 = 0$ . Thus, the claim is proved.

Since  $x_0^4 = H_0^2 + H_1^2 + H_2^2$ , we have  $-c_0(c_0 + 2x_0^2) = H_1^2 + H_2^2 = (H_1 + iH_2)(H_1 - iH_2)$ . We will discuss it in the following two cases.

1) When  $c_0 + 2x_0^2$  is irreducible, then there exist constants  $\alpha, \beta \in K$  such that  $H_1 + iH_2 = \alpha(c_0 + 2x_0^2)$ ,  $H_1 - iH_2 = \beta c_0$ , or  $H_1 - iH_2 = \alpha(c_0 + 2x_0^2)$ ,  $H_1 + iH_2 = \beta c_0$ . In either case, we have  $x_0^2 \in \text{span}_K\{H_1, H_2\}$ .

2) When  $c_0 + 2x_0^2$  is reducible, then there exists a linear polynomial  $p$  in variables  $x_1, x_2$  such that  $c_0 + 2x_0^2 = (\sqrt{2}x_0 + p)(\sqrt{2}x_0 - p)$ , where  $c_0 = -p^2$ .

If  $H_1 + iH_2$  has a factor  $\sqrt{2}x_0 + p$  or  $\sqrt{2}x_0 - p$ , without loss of generality, assume  $\sqrt{2}x_0 + p$  is a factor of  $H_1 + iH_2$ , then there exist constants  $\alpha, \beta \in K$  such that  $H_1 + iH_2 = \alpha p(\sqrt{2}x_0 + p)$  and  $H_1 - iH_2 = \beta p(\sqrt{2}x_0 - p)$ . Then  $p^2 \in \text{span}_K\{H_1, H_2\}$ . Since  $H_0 = x_0^2 + c_0 = x_0^2 - p^2$ , then  $p^2 \in \text{span}_K\{H_0, H_1, H_2\}$ .

If  $c_0 + 2x_0^2$  is a factor of  $H_1 + iH_2$ , then the same as the case 1),  $x_0^2 \in \text{span}_K\{H_1, H_2\}$ .

The above discussion shows that  $x_0^2 \in \text{span}_K\{H_0, H_1, H_2\}$ . ■

We now show that Conjecture 1 is valid with high probability in more general cases.

**Lemma 6** *If  $f^* = G \circ H$  is homogeneous decomposition and the degrees of  $f, G, H$  are 4, 2, 2, respectively, then  $x_0^2 \in \text{span}\{H_0, \dots, H_n\}$  with probability one when  $K$  is of character zero, and with probability close to one when  $K = F_q$  and  $q$  is sufficiently large. Moreover, it is also correct when  $d_G = 2, d_H > 2$ .*

*Proof* Since the first element of  $f^*$  is  $x_0^4$ , by the definition in Section 4, we know that  $x_0^3 \in \tilde{V}_{f^*} \subseteq V_{f^*,H}$ , and hence  $x_0^2 \in (V_{f^*,H} : x_0)$ . By Theorem 6,  $(V_{f^*,H} : x_0) = \text{span}\{H_0, H_1, \dots, H_n\}$  with probability one when  $K$  is of characteristic zero, and with probability close to one when  $K = F_q$  and  $q$  is a sufficiently large number. As a consequence, we have  $x_0^2 \in \text{span}\{H_0, H_1, \dots, H_n\}$  with probability one when  $K$  is of characteristic zero, and with probability close to one when  $K = F_q$  and  $q$  is a sufficiently large number. It also holds for the case when  $d_G = 2, d_H > 2$  because of the same reason. ■



By Corollary 3 and Lemma 6, we can prove Theorem 2.

For a set of non-homogeneous polynomials  $f$ , we then can compute a degree proper decomposition of its homogenization  $f^*$  and then recover a decomposition of  $f$  from that of  $f^*$ . Unfortunately, the method based on differentiation dose not work for the computation of functional decomposition of  $f^*$  directly. One of key steps of the method presented in Section 4 is that one could compute  $V_{f,h}$  from  $\tilde{V}_f$ . In general, as proved in the following theorem,  $f^*$  does not satisfy the condition of Theorem 5.

**Lemma 7** *Let  $h_1, \dots, h_n$  be polynomials in  $\mathcal{R}$  and  $(\tilde{h}_1, \dots, \tilde{h}_n) = (h_1, \dots, h_n)A$ , where  $A$  is an invertible matrix with order  $n$  over  $K$ . If the polynomials  $x_i h_j, i, j = 1, 2, \dots, n$  are linearly independent over  $K$ , then the polynomials  $x_i \tilde{h}_j, i, j = 1, 2, \dots, n$  are also linearly independent over  $K$ .*

*Proof* Assume that there are  $n^2$  elements  $b_{i,j}$  in  $K$  such that  $\sum_{i,j} b_{i,j} x_i \tilde{h}_j = 0$ . Let  $A = (a_{i,j})$ .

Then we have

$$\sum_{i,j=1}^n b_{i,j} x_i \sum_{k=1}^n a_{k,j} h_k = \sum_{i,k=1}^n \left( \sum_{j=1}^n a_{k,j} b_{i,j} \right) x_i h_k = 0.$$

Therefore  $\sum_{j=1}^n a_{k,j} b_{i,j} = 0$  for all  $i, k = 1, 2, \dots, n$ . Since  $A$  is invertible,  $b_{i,j} = 0$  for all  $i, j = 1, 2, \dots, n$ . ■

**Theorem 11** *Let  $f^* = (x_0^4, f_1^*, \dots, f_n^*)$  where the  $f_i^*$  are quartic homogeneous polynomials. Assume that  $f^* = G \circ H$  is a homogeneous decomposition of  $f^*$ . Then  $\tilde{V}_{f^*} \subsetneq V_{f^*,H}$  with high probability.*

*Proof* Obviously,  $\dim(\tilde{V}_{f^*}) < (n+1)^2$ . Then it suffices to prove that  $\dim(V_{f^*,H}) = (n+1)^2$  with high probability. Assume that  $H = (H_0, H_1, \dots, H_n)$ . By Lemma 6,  $x_0^2 \in \text{span}_K\{H_0, H_1, \dots, H_n\}$  with probability one when  $K$  is of character zero, and with probability close to one when  $K = F_q, q$  is sufficiently large. Now assume that  $x_0^2 \in \text{span}\{H_0, H_1, \dots, H_n\}$ . Then  $x_0^2$  can be extended into a basis of  $\text{span}_K\{H_0, H_1, \dots, H_n\}$ , say  $\tilde{H}_0 = x_0^2, \tilde{H}_1, \dots, \tilde{H}_n$ . Then we have

$$f^* = \tilde{G} \circ (\tilde{H}_0, \tilde{H}_1, \dots, \tilde{H}_n). \tag{17}$$

By Lemma 7, it suffices to prove that  $V_{f^*,\tilde{H}} = (n+1)^2$  with high probability. Furthermore, it suffices to prove that the polynomials  $x_i x_0^2, x_i h_j$  are linearly independent over  $K(b_{k,l}^{(j)})$  where  $h_j = \sum_{1 \leq k \leq l \leq n} b_{k,l}^{(j)} x_k x_l$  are generic polynomials.

Assume that

$$\sum_i t_{i,0} x_i x_0^2 + \sum_{i,j} t_{i,j} x_i h_j = 0.$$

Combine the same terms, and let the coefficients of each monomials equal to 0, then we can get a set of equations with  $t_{i,j}$  as variables and  $b_{k,l}^{(j)}$  as coefficients, that is,

$$MT = 0,$$

where  $M$  is an  $n^2 \times n^2$  matrix whose entries are polynomials in  $b_{k,l}^{(j)}$  and  $T$  is the vector  $(\dots, t_{i,0}, \dots, t_{i,j}, \dots)^T$ . We will show that  $\det(M) \neq 0$ . Consider the specialization of  $b_{k,l}^{(j)}$ . We specialize  $b_{k,l}^{(j)}$  to 1 for  $j = k = l$  and to 0 for other cases. Then  $h_j$  becomes  $x_j^2$  under this specialization. Assume that  $M$  becomes  $\tilde{M}$  under the above specialization. Since the



polynomials  $x_i x_j^2, i, j = 0, 1, \dots, n$  are linearly independent over  $K$ ,  $\det(\widetilde{M}) \neq 0$ . Therefore,  $\det(M) \neq 0$ , which implies the conclusion.  $\blacksquare$

## 6 Algorithm and Complexity

Let  $f \in \mathcal{R}^n$  be a set of quartic homogeneous polynomials, each polynomial is of the same degree. We now give the algorithm to find a degree proper decomposition of  $f$ . We prove that it is a polynomial time algorithm with high successful probability. Note that the algorithm is essentially the same as that given in [11]. Our main contribution is the analysis of the algorithm.

### Algorithm FDPMPH4

**Input** Let  $f \in \mathcal{R}^n$  be a set of quartic homogeneous polynomials, each polynomial is of the same degree.

**Output**  $g, h \in \mathcal{R}^n$  such that  $f = g \circ h$  is a degree proper decomposition of  $f$ . The algorithm may fail even if such a decomposition exists.

**Step 1** Compute  $\widetilde{V}_f := \text{span}_K \{ \frac{\partial f_i}{\partial x_j} : i, j = 0, 1, \dots, n \}$ .

**Step 2** Compute  $R_{(f,h)} := (\widetilde{V}_f : l)$  as stated in the proof of Theorem 8.

**Step 3** Perform a linear elimination to the generators of  $R_{(f,h)}$  to obtain a basis  $(h_1, \dots, h_k)$  of  $R_{(f,h)}$ . If  $k = n$ , then  $h = (h_1, \dots, h_n)$ ; if  $k < n$ , then  $h = (h_1, \dots, h_k, h_1, \dots, h_1)$ ; if  $k > n$ , then the algorithm terminates without finding a result.

**Step 4** Compute the coefficients of  $g$  by solving a system of linear equations as shown in Theorem 4.

The following result shows that Theorem 1 is valid.

**Theorem 12** *Algorithm FDPMPH4 needs  $O(n^{3\omega})$  arithmetic operations in the field  $K$ , where  $2 \leq \omega < 3$ . For a randomly decomposable  $f$ , the algorithm computes the decomposition with probability one when  $K$  is of characteristic zero, and with probability close to one when  $K = F_q \cdot q$  is a sufficiently large number.*

*Proof* The complexity of the whole algorithm depends on Steps 2 and 4, both of them cost  $O(n^{3\omega})$  arithmetic operations by Theorems 4 and 8. Then we have a polynomial time algorithm to find a degree proper decomposition  $f = g \circ h$  for  $g, h \in \mathcal{R}^n$  with probability one when  $K$  is of characteristic zero, and with probability close to one when  $K = F_q \cdot q$  is a sufficiently large number.  $\blacksquare$

## 7 Conclusion and Problems

In this paper, we give a theoretical analysis for the approaches of computing functional decomposition for multivariate polynomials based on differentiation and homogenization proposed in [11–14]. We show that a degree proper functional decomposition for a set of quartic homogeneous polynomials can be computed using the algorithm with high probability for randomly decomposable polynomials. We propose a conjecture asserting that the decomposition for a set of polynomials can be computed from the decomposition of its homogenization, and show that the conjecture is valid in several cases with high probability. Finally, we prove that the right decomposition factors for a set of polynomials can be computed from its right decomposition factor space.

Despite of the significant progresses, the general FDP for multivariate polynomials is widely open. Some of the basic problems related to FDP of multivariate polynomials are not resolved. We will give two basic open problems below.

The first problem is about the existence of an algorithm for FDP.

**Problem 2** Given  $f \in \mathcal{R}^n$ , to find an FDP for  $f$  is decidable or not.

Note that in a decomposition  $f = g \circ h$ , the degrees of  $g$  and  $h$  could be arbitrarily high. Consider the following two transformations:

$$\begin{aligned} T_1 &: (x_1, \dots, x_n) \Rightarrow (x_1 + p, x_2 \cdots, x_n), \\ T_2 &: (x_1, \dots, x_n) \Rightarrow (x_1 - p, x_2 \cdots, x_n), \end{aligned}$$

where  $p$  is a polynomial in  $x_2, \dots, x_n$  of any degree. Then  $T_1 \circ T_2 = (x_1, \dots, x_n)$ . For any decomposition  $f = g \circ h$ ,  $f = (g \circ T_1) \circ (T_2 \circ h)$  is also a decomposition of  $f$ . Therefore, one way to solve Problem 2 is to find the smallest possible degrees of  $g$  and  $h$  if a decomposition exists.

The second problem is about the computational complexity of FDP. In this aspect, even the simplest case is not resolved.

**Problem 3** Let  $f \in \mathcal{R}^n$  be a set of quartic polynomials. Estimate the complexity of computing a degree proper FDP of  $f$  over a finite field  $K = F_q$ . In particular, does there exist a polynomial-time algorithm for Boolean polynomials?

## References

- [1] J. Patarin and L. Goubin, Asymmetric cryptography with S-boxes, *Proceedings of ICICS'97*, Lecture Notes in Computer Science, Springer, **1334**, 1997.
- [2] M. Dickerson, The functional decomposition of polynomials, Ph.D Thesis, TR 89-1023, Department of Computer Science, Cornell University, Ithaca, NY, 1989.
- [3] J. Patarin and L. Goubin, Asymmetric cryptography with S-boxes (extended version), Available at <http://citeseer.ist.psu.edu/patarin97asymmetric.html>, 1997.
- [4] J. Gutierrez, R. Rubio, and J. von zur Gathen, Multivariate polynomial decomposition, *Algebra in Engineering, Communication and Computing*, 2003, **14**(1): 11–31.
- [5] J. von zur Gathen, Functional decomposition of polynomials: The tame case, *Journal of Symbolic Computation*, 1990, **9**: 281–299.
- [6] J. von zur Gathen, Functional decomposition of polynomials: The wild case, *Journal of Symbolic Computation*, 1990, **10**: 437–452.
- [7] E. W. Chionh, X. S. Gao, and L. Y. Shen, Inherently improper surface parametric supports, *Computer Aided Geometric Design*, 2006, **23**: 629–639.
- [8] J. Li, L. Shen, and X. S. Gao, Proper reparametrization of rational ruled surface, *Journal of Computer Science & Technology*, 2008, **23**(2): 290–297.
- [9] J. Li and X. S. Gao, The proper parametrization of a special class of rational parametric equations, *Journal of Systems Science & Complexity*, 2006, **19**(3): 331–339.
- [10] L. Shen, E. Chionh, X. S. Gao, and J. Li, Proper reparametrization for inherently improper unirational varieties, *Journal of Systems Science & Complexity*, 2011, **24**(2): 367–380.
- [11] D. F. Ye, K. Y. Lam, and Z. D. Dai, Cryptanalysis of “2R” Schemes, *Advances in Cryptology-CRYPTO 1999*, Lecture Notes in Computer Science, Springer, 1999, **1666**: 315–325.

- 
- [12] J. C. Faugère and L. Perret, Cryptanalysis of  $2R^-$  schemes, *Advances in Cryptology-CRYPTO 2006*, Lecture Notes in Computer Science, 2006, **4117**: 357–372.
  - [13] J. C. Faugère and L. Perret, An efficient algorithm for decomposing multivariate polynomials and its applications to cryptography, *Special Issue of JSC, “Gröbner Bases techniques in Coding Theory and Cryptography”*, 2008.
  - [14] J. C. Faugère and L. Perret, High order derivatives and decomposition of multivariate polynomials, 2009, *Proc. ISSAC*, ACM Press, 2009.
  - [15] J. C. Faugère, J. von zur Gathen, and L. Perret, Decomposition of generic multivariate polynomials, *ISSAC*, 2010.
  - [16] D. Coppersmith and S. Winograd, Matrix multiplication via arithmetic progressions, *Journal of Symbolic Computation*, 1990, **9**(3): 251–280.
  - [17] R. Lidl and H. Niederreiter, *Finite Fields*, Addison-Wesley Publishing Company, 1983.
  - [18] D. Cox, J. Little, and D. O’Shea, *Ideals, Varieties, and Algorithms*, 2nd ed., Springer-Verlag, 1996.
  - [19] T. Y. Lam, *The Algebraic Theory of Quadratic Forms*, Benjamin, 1973.
  - [20] A. Schinzel, *Polynomials with Special Regard to Reducibility*, Cambridge University Press, 2000.