Analysis of Multivariate Encryption Schemes: Application to Dob

Morten Øygarden[1], Patrick Felke[2], and Håvard Raddum[1]

[1] Simula UiB
[2] University of Applied Sciences Emden-Leer
{morten.oygarden,haavardr}@simula.no,
patrick.felke@hs-emden-leer.de,

Abstract. In this paper, we study the effect of two modifications to multivariate public key encryption schemes: internal perturbation (*ip*), and Q_+. Focusing on the *Dob encryption scheme*, a construction utilising these modifications, we accurately predict the number of degree fall polynomials produced in a Gröbner basis attack, up to and including degree five. The predictions remain accurate even when fixing variables. Based on this new theory we design a novel attack on the Dob encryption scheme, which breaks Dob using the parameters suggested by its designers.

While our work primarily focuses on the Dob encryption scheme, we also believe that the presented techniques will be of particular interest to the analysis of other big–field schemes.

1 Introduction

Public key cryptography has played a vital role in securing services on the internet that we take for granted today. The security of schemes based on integer factorization and the discrete logarithm problem (DLP) is now well understood, and the related encryption algorithms have served us well over several decades.

In [23] it was shown that quantum computers can solve both integer factorization and DLP in polynomial time. While large scale quantum computers that break the actual implementations of secure internet communication have yet to be built, progress is being made in constructing them. This has led the community for cryptographic research to look for new public key primitives that are based on mathematical problems believed to be hard even for quantum computers, so called *post–quantum cryptography.*

In 2016 NIST launched a project aimed at standardizing post–quantum public key primitives. A call for proposals was made and many candidate schemes were proposed. The candidates are based on a variety of problems, including the shortest vector problem for lattices, the problem of decoding a random linear code, or the problem of solving a system of multivariate quadratic equations over a finite field (the MQ problem).

The first encryption scheme based on the MQ problem, named C^*, was proposed in [19] and was broken by Patarin in [21]. Since then, much work has gone into designing new central maps, as well as modifications that can enhance the security of existing ones. Several multivariate schemes have been proposed following C^*, for instance [22, 5, 25, 26]. While some of the schemes for digital signatures based on the MQ problem seem to be secure, it has been much harder to construct encryption schemes that are both efficient and secure. The papers [15, 20, 27, 24, 1], all present attacks on MQ-based public key encryption schemes, and as of now we are only aware of a few (e.g., [8, 29]) that remain unbroken.

In [18] a new kind of central mapping is introduced, which can be used to construct both encryption and signature schemes. The novel feature of the central mapping is that it has a high degree over an extension field, while still being easy to invert. The encryption variant proposed in [18] is called Dob and uses two types of modifications to its basic construction.

Our Contribution

The initial part of our work provides a theoretical analysis of (combinations of) two modifications for multivariate cryptosystems. The Q_+–modification was (to the best of our knowledge) first proposed in [18], while the second, internal perturbation (*ip*), has been in use for earlier schemes [11, 7, 8]. More specifically, we develop the tools for computing the dimension of the ideal associated with these modifications, at different degrees. This in turn provides key insights into the complexity of algebraic attacks based on Gröbner basis techniques.

As an application, we focus on the Dob encryption scheme proposed in [18]. We are able to deduce formulas that predict the exact number of first fall polynomials for degrees 3,4 and 5. These formulas furthermore capture how the number of degree fall polynomials changes as an attacker fixes variables, which also allows for the analysis of hybrid methods (see e.g., [3]).

Finally, the newfound understanding allow us to develop a novel attack on the Dob encryption scheme. Through analyzing and manipulating smaller, projected polynomial systems, we are able to extract and isolate a basis of the secret modifiers, breaking the scheme. While the details of the attack have been worked out for the Dob encryption scheme, we believe the techniques themselves could be further generalised to include different central maps and modifications.

Organisation

The paper is organized as follows. In Section 2 we recall the relation between $\mathbb{F}_2^d$ and $\mathbb{F}_{2^d}$, as well as the necessary background for solving multivariate systems over $\mathbb{F}_2$. In Section 3 we develop the general theory that explores the effectiveness of the modifications Q_+ and *ip* . Section 4 introduces the Dob scheme, and we deduce formulas that predict the number of degree fall polynomials for this construction. Experimental data verifying the accuracy of these formulas are presented in Section 5. In Section 6 we develop the novel attack on the Dob encryption scheme, using the information learned from the previous sections. Finally, sections 7 and 8 discuss and conclude the work.

Table of definitions

Throughout the paper we will use the notation in Table 1. We list it here for easy reference.

Term	Meaning
$B(n)$	$B(n) = \mathbb{F}_2[x_1, \dots, x_n]/\langle x_1^2 + x_1, \dots, x_n^2 + x_n \rangle$
$\overline{B}(n)$	$\overline{B}(n) = \mathbb{F}_2[x_1, \dots, x_n]/\langle x_1^2, \dots, x_n^2 \rangle$
$\overline{B}(n)_\nu$	The set of homogeneous polynomials of degree ν in n variables.
$\langle \mathcal{R} \rangle$	The ideal associated with the set of polynomials $\mathcal{R}$.
$\langle \mathcal{R} \rangle_\nu$	The ν–th degree part of a graded ideal $\langle \mathcal{R} \rangle$.
$\dim_\nu(\langle \mathcal{R} \rangle)$	The dimension of $\langle \mathcal{R} \rangle_\nu$ as an $\mathbb{F}_2$–vector space.
$\mathcal{P}^h$	A set of homogeneous quadratic polynomials over $\overline{B}(n)_2$
$\mathrm{Syz}(\mathcal{P}^h)_\nu$	The grade ν part of the (first) syzygy module of $\mathcal{P}^h$. (See section 2.1)
$\mathcal{T}(\mathcal{P}^h)_\nu$	The grade ν part of the trivial syzygy module of $\mathcal{P}^h$. (See section 2.1)
$\mathcal{S}(\mathcal{P}^h)_\nu$	$\mathcal{S}(\mathcal{P})_\nu = \mathrm{Syz}(\mathcal{P})_\nu / \mathcal{T}(\mathcal{P}^h)_\nu$.
Q_+, q_i, t	The Q_+ modifier, with $q_1, \dots, q_t$ added quadratic polynomials.
(ip), v_i, k	The *internal perturbation* modifier with $v_1, \dots, v_k$ linear forms.
$N_\nu^{(\alpha,\beta)}$	Estimate of the number of degree fall polynomials at degree ν.

Table 1: Notation used in the paper

2 Preliminaries

Multivariate big–field encryption schemes are defined using the field $\mathbb{F}_{q^d}$ and the d-dimensional vector space over the base field, $\mathbb{F}_q^d$. In practical implementations, $q = 2$ is very often used, and we restrict ourselves to only consider this case in the paper.

2.1 Polynomial System Solving

A standard technique used in the cryptanalysis of multivariate schemes, is to compute a Gröbner basis associated with the ideal $\langle p_i + y_i \rangle_{1 \le i \le m}$, for a fixed ciphertext $y_1, \dots, y_m$ (see for example [6] for more information on Gröbner bases). As we are interested in an encryption system, we can reasonably expect a unique solution in the boolean polynomial ring $B(n)$. In this setting the solution can be read directly from a Gröbner basis of any order.

One of the most efficient algorithms for computing Gröbner bases is F$_4$ [14]. In the usual setting, the algorithm proceeds in a step–wise manner; each step has an associated degree, D, where all the polynomial pairs of of degree D are reduced simultaneously using linear algebra. The degree associated with the most time consuming step is known as the *solving degree*, D_{solv}, and time complexity of F_4 can be estimated to be:

$$\text{Complexity}_{\text{GB}} = \mathcal{O}\Bigg(\bigg(\sum_{i=0}^{D_{solv}} \binom{d}{i} \bigg)^{\omega} \Bigg), \tag{1}$$

where $2 \leq \omega \leq 3$ denotes the linear algebra constant. Determining D_{solv} is in general difficult, but there is an important class of polynomial systems that is well understood. Recall that a homogeneous polynomial system, $\mathcal{F}^h = (f_1^h, \ldots, f_m^h) \in \overline{B}(n)^m$, is said to be *semi–regular* if the following holds; for all $1 \leq i \leq m$ and any $g \in \overline{B}(n)$ satisfying

$$g f_i^h \in \langle f_1^h, \ldots, f_{i-1}^h \rangle \text{ and } \deg(g f_i) < D_{reg}, \tag{2}$$

then $g \in \langle f_1^h, \ldots, f_i^h \rangle$. Here D_{reg} is the *degree of regularity* as defined in [2], (for $i = 1$ the ideal generated by $\emptyset$ is the 0–ideal). We will also need a weaker version of this definition, where we say that $\mathcal{F}^h$ is D_0–semi–regular, if the same condition holds, but for $D_0 < D_{reg}$ in place of D_{reg} in eq. (2). An inhomogeneous system $\mathcal{F}$ is said to be (D_0–)semi–regular if its upper homogeneous part is. For a quadratic, semi–regular system $\mathcal{F}$ over $\overline{B}(n)$, the Hilbert series of $\overline{B}(n)/\mathcal{F}$ is written as (Corollary 7 in [2]):

$$T_{m,n}(z) = \frac{(1+z)^n}{(1+z^2)^m}, \tag{3}$$

and the degree of regularity can be computed explicitly as the degree of the first non–positive term in this series. Determining whether a given polynomial system is semi–regular may, in general, be as hard as computing a Gröbner basis for it. Nevertheless, experiments seem to suggest that randomly generated polynomial systems behave as semi–regular sequences with a high probability [2], and the degree of regularity can in practice be used as the solving degree in eq. (1). We will denote the degree of regularity for a semi–regular sequence of m polynomials in n variables as $D_{reg}(m, n)$. On the other hand, it is well known that many big–field multivariate schemes are not semi–regular (e.g., [15][5]). In these cases the *first fall degree* is often used to estimate the solving degree ([9][20]). The first fall degree, according to [9], will be defined in definition 2, but before that we recall the definition of a *Macaulay matrix* associated to a polynomial system.

Definition 1. *Let $\mathcal{P}$ be an (inhomogeneous) polynomial system in $B(n)$, of degree two. An (inhomogeneous)* Macaulay matrix *of $\mathcal{P}$ at degree D, $M_D(\mathcal{P})$, is a matrix with entries in $\mathbb{F}_2$, such that:*

1. *The columns are indexed by the monomials of degree $\leq D$ in $B(n)$.*
2. *The rows are indexed by the possible combinations $x^\alpha p_i$, where $1 \leq i \leq n$ and $x^\alpha \in B(n)$ is a monomial of degree $\leq D - 2$. The entries in one row corresponds to the coefficients of the associated polynomial.*

Similarly, we define the homogeneous Macaulay matrix *of $\mathcal{P}$ at degree D, $\overline{M}_D(\mathcal{P})$, by considering $\mathcal{P}^h \in \overline{B}(n)$, only including monomials of degree D in the columns, and rows associated to combinations $x^\alpha p_i^h$, $deg(x^\alpha) = D - 2$.*

Syzygies and Degree Fall Polynomials. Let $\mathcal{P}^h = (p_1^h, \dots, p_m^h) \in \overline{B}(n)_2^m$ denote a homogeneous quadratic polynomial system. The set $\mathcal{P}^h$ induces a map:

$$\begin{array}{rll} \psi^{\mathcal{P}^h} : & \overline{B}(n)^m & \longrightarrow \overline{B}(n) \\ & (b_1, \dots, b_m) & \longmapsto \sum_{i=1}^m b_i p_i^h, \end{array} \tag{4}$$

which in turn splits into graded maps $\psi_{\nu-2}^{\mathcal{P}^h} : \overline{B}(n)_{\nu-2}^m \longrightarrow \overline{B}(n)_\nu$. The $\overline{B}(n)$–module $\mathrm{Syz}(\mathcal{P}^h)_\nu = \mathrm{Ker}(\psi_{\nu-2}^{\mathcal{P}^h})$ is known as the *ν–th grade of the (first) syzygy module of $\mathcal{P}^h$*. When $\nu = 4$, $\mathrm{Syz}(\mathcal{P}^h)_4$ will contain the *Koszul Syzygies*, which are generated by $(0, ..., 0, p_j^h, 0, ..., 0, p_i^h, 0, ..., 0)$ (p_j^h is in position i and p_i^h is in position j), and the *field syzygies*, which are generated by $(0, ..., 0, p_i^h, 0, ..., 0)$ (p_i^h in position i). These syzygies correspond to the cancellations $p_j^h p_i^h + p_i^h p_j^h = 0$ and $(p_i^h)^2 = 0$. As they are always present, and not dependent of the structure of $\mathcal{P}^h$, they are sometimes referred to as the *trivial syzygies*. More generally, we will define the submodule $\mathcal{T}(\mathcal{P}^h)_\nu \subseteq \mathrm{Syz}(\mathcal{P}^h)_\nu$ to be the ν–th graded component of the module generated by the Koszul and field syzygies, and denote $\mathcal{S}(\mathcal{P})_\nu = \mathrm{Syz}(\mathcal{P}^h)_\nu / \mathcal{T}(\mathcal{P}^h)_\nu$.

Definition 2. *The* first fall degree *associated with the quadratic polynomial system $\mathcal{P}$ is the natural number*

$$D_{ff} = min\{\ D \geq 2 \mid \mathcal{S}(\mathcal{P})_D \neq 0\ \}.$$

Representations over base and extension fields For any fixed isomorphism $\mathbb{F}_2^d \simeq \mathbb{F}_{2^d}$, there is a one–to–one correspondence between d polynomials in $B(d)$ and a univariate polynomial in $\mathbb{F}_{2^d}[X]/\langle X^{2^d} + X\rangle$ (see 9.2.2.2 in [4] for more details). For an integer j, let $w_2(j)$ denote the number of nonzero coefficients in the binary expansion of j. For a univariate polynomial $H(X)$, we define $\max_{w_2}(H)$ as the maximal $w_2(j)$ where j is the degree of a term occurring in H. Let $P(X)$ be the univariate representation of the public key of a multivariate scheme, and suppose there exists a polynomial $H(X)$ such that

$$\max_{w_2}(H(X)P(X)) < \max_{w_2}(H(X)) + \max_{w_2}(P(X)). \tag{5}$$

Then the multivariate polynomials corresponding to the product $H(X)P(X)$ will yield degree fall polynomials from (multivariate) degree $\max_{w_2}(H) + \max_{w_2}(P)$ down to degree $\max_{w_2}(HP)$.

It was mentioned in [15] that the presence of polynomials satisfying eq. (5) was the reason for Gröbner basis algorithms to perform exceptionally well on HFE–systems. Constructing particular polynomials that satisfy eq. (5) has also been a central component in the security analyzes found in [9] and [20].

3 Estimating the Number of Degree Fall Polynomials

We start by introducing a general setting, motivated by the Dob encryption scheme which we will focus on later. Let $\mathcal{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ be a system of m quadratic polynomials over $B(n)$. Furthermore, consider the following two modifiers[3]:

1. The *internal perturbation* (*ip*) modification chooses k linear combinations $v_1, \ldots, v_k$, and adds a random quadratic polynomial in the v_i's to each polynomial in $\mathcal{F}$.
2. The Q_+ modifier selects t quadratic polynomials $q_1, \ldots, q_t$, and adds a random linear combination of them to each polynomial in $\mathcal{F}$.

Let H_{ip} be the random quadratic polynomials in $v_1, \ldots, v_k$ and H_{Q_+} the random linear combinations of $q_1, \ldots, q_t$. A modification of the system $\mathcal{F}$ can then be written as

$$\begin{aligned} \mathcal{P} : \mathbb{F}_2^n &\longrightarrow \mathbb{F}_2^m \\ x &\longmapsto \mathcal{F}(x) + H_{ip}(x) + H_{Q_+}(x). \end{aligned} \tag{6}$$

The problem we will be concerned with in this section is the following: given full knowledge of the degree fall polynomials of the system $\mathcal{F}$, what can we say about the degree fall polynomials of the system $\mathcal{P}$?

3.1 The Big Picture

Let $\mathcal{F}^h$ and $\mathcal{P}^h$ denote the homogeneous parts of the systems $\mathcal{F}$ and $\mathcal{P}$ respectively, and consider them over $\overline{B}(n)$. For a positive integer $\alpha \leq k$, we define V^α to be the homogeneous ideal in $\overline{B}(n)$ that is generated by all possible combinations of α linear forms from the *ip* modification, i.e.:

$$V^\alpha = \langle (v_{i_1} v_{i_2} \cdots v_{i_\alpha})^h \mid 1 \leq i_1 < i_2 < \ldots < i_\alpha \leq k \rangle. \tag{7}$$

In other words, V^α is the product ideal $\overbrace{V^1 \cdot V^1 \cdot \ldots \cdot V^1}^{\alpha}$. Similarly, for the quadratic polynomials associated with the Q_+ modifier we define Q^β for a positive integer $\beta \leq t$ to be the product ideal:

$$Q^\beta = \langle (q_{i_1} q_{i_2} \cdots q_{i_\beta})^h \mid 1 \leq i_1 < i_2 < \ldots < i_\beta \leq t \rangle. \tag{8}$$

Finally, for $0 \leq \alpha \leq k$ and $0 \leq \beta \leq t$, we define the ideal of different combinations of the modifiers, $M^{(\alpha,\beta)} = \langle V^\alpha, Q^\beta \rangle$, along with the boundary cases $M^{(\alpha,0)} = V^\alpha$, $M^{(0,\beta)} = Q^\beta$ and $M^{(0,0)} = \langle 1 \rangle$.

The following result is an important first step to understand how the degree fall polynomials in $\mathcal{F}$ behave when modifiers are introduced to the scheme.

[3] The authors of [18] named these two modifiers $\oplus$ and " + ". Note that in earlier literature (c.f. [28]), the " + " modification refers to a different modification than what is described in [18], and the $\oplus$ modification has been called *internal perturbation* (*ip*). (To the best of our knowledge, the " + " modification from [18] has not been used in earlier work). To avoid any confusion, we have chosen to stick with the name (*ip*) and use Q_+ for [18]'s "+"

Lemma 1. *Let $\mathcal{P}^h$, $\mathcal{F}^h$, $M^{(2,1)}$ be defined as above, and $\psi^{\mathcal{P}^h}$ be as defined in eq.* (4)*. Then $\langle \psi^{\mathcal{P}^h}(\mathcal{S}(\mathcal{F}))\rangle$ and $\langle \psi^{\mathcal{P}^h}(Syz(\mathcal{F}^h))\rangle$ are homogeneous subideals of $\langle \mathcal{P}^h\rangle \cap M^{(2,1)}$.*

Proof. We show the statement for $\langle \psi^{\mathcal{P}^h}(\mathrm{Syz}(\mathcal{F}^h))\rangle$; the case of $\langle \psi^{\mathcal{P}^h}(\mathcal{S}(\mathcal{F}))\rangle$ is similar. First note that $\psi^{\mathcal{P}^h}(\mathrm{Syz}(\mathcal{F}^h))$ is a group, as it is the image of a group under a group homomorphism. Secondly, for any element $\mathbf{a} = (a_1, \ldots, a_m) \in \mathrm{Syz}(\mathcal{F}^h)$, and any $r \in \overline{B}(n)$, we have $r\psi^{\mathcal{P}^h}(\mathbf{a}) = \psi^{\mathcal{P}^h}((ra_1, \ldots, ra_m))$, where also $(ra_1, \ldots, ra_m) \in \mathrm{Syz}(\mathcal{F}^h)$. It follows that $\psi^{\mathcal{P}^h}(\mathrm{Syz}(\mathcal{F}^h))$ is indeed an ideal.

The inclusion $\langle \psi^{\mathcal{P}^h}(\mathrm{Syz}(\mathcal{F}^h))\rangle \subseteq \langle \mathcal{P}^h\rangle$ follows directly from the definition of $\psi^{\mathcal{P}^h}$. For the other inclusion we note that, by construction, we can write $p_i^h = f_i^h + \sum_{j=1}^{t} b_{i,j} q_j^h + \sum_{j,l=0}^{k} c_{i,j,l}(v_j v_l)^h$, for all $1 \leq i \leq m$ and for suitable constants $b_{i,j}, c_{i,j,l} \in \mathbb{F}_2$, where f_i^h, p_i^h are the polynomials of $\mathcal{F}^h$ and $\mathcal{P}^h$ respectively. When $\mathbf{a} \in \mathrm{Syz}(\mathcal{F}^h)$, the f_i^h–parts in $\psi^{\mathcal{P}^h}(\mathbf{a})$ will vanish, and we are left with a polynomial that can be generated from the elements of V^2 and Q^1. Hence we also have $\langle \psi^{\mathcal{P}^h}(\mathrm{Syz}(\mathcal{F}^h))\rangle \subseteq M^{(2,1)}$.

In particular, there is the following chain of ideals

$$\langle \psi^{\mathcal{P}^h}(\mathcal{S}(\mathcal{F}))\rangle \subseteq \langle \psi^{\mathcal{P}^h}(\mathrm{Syz}(\mathcal{F}^h))\rangle \subseteq \langle \mathcal{P}^h\rangle \cap M^{(2,1)} \subseteq M^{(2,1)}. \tag{9}$$

We now allow ourselves to be slightly informal, in order to see how this all relates in practice to the cases we are interested in. At each degree ν, the dimension $\dim_\nu(M^{(2,1)})$ of $M_\nu^{(2,1)}$ as a vector space over $\mathbb{F}_2$ can be seen as a measure of how much information the modifiers can hide. An interesting case from an attackers point of view is when $\langle \psi^{\mathcal{P}^h}(\mathcal{S}(\mathcal{F}))\rangle_{\nu_0}$ has the maximal dimension $\dim_{\nu_0}(\langle \psi^{\mathcal{P}^h}(\mathcal{S}(\mathcal{F}))\rangle) = \dim_{\nu_0}(M^{(2,1)})$, for a relatively small ν_0. While 'excess' polynomials in $\langle \psi^{\mathcal{P}^h}(\mathcal{S}(\mathcal{F}))\rangle_{\nu_0}$ will sum to 0 in $\overline{B}(n)$, there is a chance that the corresponding inhomogeneous polynomials will result in degree fall polynomials when treated over $B(n)$. In particular, this yields an upper bound $D_{ff} \leq \nu_0$ on the first fall degree. We can do even better in practice.

Note that $(M^{(2,1)}\langle \mathcal{P}^h\rangle)_\nu$ will be a subspace of (the row space of) the Macaulay matrix $\overline{M}_\nu(\mathcal{P})$. As this matrix can be constructed by an attacker, we should count the possible combinations of polynomials from both $(M^{(2,1)}\langle \mathcal{P}^h\rangle)$ and the image of $\psi^{\mathcal{P}^h}(\mathcal{S}(\mathcal{F}))$. Some caution is warranted when counting these combinations. For instance, $\psi^{\mathcal{P}^h}(ms) \in M^{(2,1)}\langle \mathcal{P}^h\rangle$ for any $m \in M^{(2,1)}$ and $s \in \mathcal{S}(\mathcal{F})$, so we need to be careful in order to not count the same elements twice. For now we will keep up with our informal theme and denote '$M^{(2,1)}\langle \mathcal{P}^h\rangle$ modulo these collisions' by $\mathcal{P}_{M^{(2,1)}}$. We will deal with it more properly when computing its dimension in section 3.3. We also show later, in appendix A, that $\langle \psi^{\mathcal{P}^h}(\mathcal{T}(\mathcal{F}^h))\rangle \subseteq M^{(2,1)}\langle \mathcal{P}^h\rangle$, which is why we will focus on $\langle \psi^{\mathcal{P}^h}(\mathcal{S}(\mathcal{F}))\rangle$ (as opposed to $\langle \psi^{\mathcal{P}^h}(\mathrm{Syz}(\mathcal{F}^h))\rangle$).

We now have everything needed to discuss estimates of the number of degree fall polynomials at different degrees. We start by assuming that none of the degree fall polynomials we get from $\mathcal{S}(\mathcal{F})$ (under $\psi^{\mathcal{P}^h}$) can be reduced by lower–degree Macaulay matrices of $\mathcal{P}$. This allows us to directly use $\dim_\nu(\mathcal{S}(\mathcal{F}))$. We furthermore add $\dim_\nu(\mathcal{P}_{M^{(2,1)}})$, and subtract by $\dim_\nu(M^{(2,1)})$. This yields the expression for our first estimate of degree fall polynomials, $N_\nu^{(0,0)}$, at degree ν:

$$N_\nu^{(0,0)} = \dim_\nu(\mathcal{S}(\mathcal{F})) + \dim_\nu(\mathcal{P}_{M^{(2,1)}}) - \dim_\nu(M^{(2,1)}). \tag{10}$$

In a sense, $N_\nu^{(0,0)}$ can be thought of as estimating the number of degree fall polynomials, as an effect of 'over saturating' $M_\nu^{(2,1)}$. When $N_\nu^{(0,0)}$ is a positive number, this is the number of degree fall polynomials we expect to find (based on this effect); if $N_\nu^{(0,0)}$ is negative, there is no such over saturation, and we do not expect any degree fall polynomials at degree ν. The benefits of having the expression in eq. (10) is that the study of the relatively complex polynomial system $\mathcal{P}^h$ can be broken down to studying three simpler systems. The dimensions of $M^{(2,1)}$ and $\mathcal{P}_{M^{(2,1)}}$ can, in particular, be further studied under the assumptions that the modifiers form a semi–regular system. In addition to being a reasonable assumption as the modifiers are randomly chosen, this is also the ideal situation for the legitimate user, as this maximizes the dimension of $M^{(2,1)}$. Indeed, the study of $M^{(2,1)}$ and $\mathcal{P}_{M^{(2,1)}}$ will be continued in the following subsections. Before that, we will generalize the ideas presented so far, arriving at several expressions that can be used to estimate the number of degree fall polynomials.

Generalised Estimates of Degree Fall Polynomials. Let $M^{(\alpha,\beta)}\mathrm{Syz}(\mathcal{F})$ denote the module $\{ms \mid m \in M^{(\alpha,\beta)}, s \in \mathrm{Syz}(\mathcal{F})\}$ (which is well–defined since $\mathrm{Syz}(\mathcal{F})$ is a $\overline{B}(n)$–module), and define

$$\mathcal{S}(\mathcal{F})_{M^{(\alpha,\beta)}} := [M^{(\alpha,\beta)}\mathrm{Syz}(\mathcal{F})]/\mathcal{T}(\mathcal{F}).$$

Instead of considering *all* the syzygies $\mathcal{S}(\mathcal{F})$, we can start with submodules of the form $\mathcal{S}(\mathcal{F})_{M^{(\alpha,\beta)}}$. The benefit is that the ideal we need to 'over saturate' will now be $M^{(\alpha,\beta)}M^{(2,1)}$. In section 5 we will see several examples where this yields a better estimate than $N_\nu^{(0,0)}$. Following through with this idea, along with the same considerations discussed prior to eq. (10), we arrive at the following estimate for $\alpha, \beta \geq 0$:

$$\begin{aligned} N_\nu^{(\alpha,\beta)} &= \dim_\nu(\mathcal{S}(\mathcal{F})_{M^{(\alpha,\beta)}}) - \dim_\nu(M^{(\alpha,\beta)}M^{(2,1)}) \\ &\quad + \dim_\nu(\mathcal{P}^h_{M^{(\alpha,\beta)}M^{(2,1)}}). \end{aligned} \tag{11}$$

Recalling the convention that $M^{(0,0)} = \langle 1 \rangle$, this is indeed a generalisation of eq. (10).

We now have several different estimates for degree fall polynomials, varying with the choice of α, β. Any of these may be dominating, depending on the parameters of the scheme. The general estimate at degree ν is then taken to be their maximum:

$$N_\nu = \max\{0, N_\nu^{(\alpha,\beta)} \mid 0 \leq \alpha \leq k \text{ and } 0 \leq \beta \leq t\}. \tag{12}$$

Note in particular that if $N_\nu = 0$, then all our estimates are non–positive, and we do not expect any degree fall polynomials at this degree.

Consider now the main assumptions underlying these estimates. Firstly, recall that we assumed that none of the degree fall polynomials that can be made from $\psi^{\mathcal{P}}(\mathcal{S}(\mathcal{F})_{M^{(\alpha,\beta)}})$ will be reduced to 0 when solving the system $\mathcal{P}$. Secondly, the formulas implicitly assume that all the polynomials in $M^{(\alpha,\beta)}M^{(2,1)}$ need to be reduced before we can observe degree fall polynomials. The third assumption, concerning $\mathcal{P}^h_{M^{(\alpha,\beta)}M^{(2,1)}}$, will be specified in section 3.3.

Finally, we stress that the aim of this section has been to investigate one of the aspects that can lead to a system exhibiting degree fall polynomials. The estimates presented should not be used without care to derive arguments about lower bounds on the first fall degree. Nevertheless, we find that in practice these estimates and their assumptions seem to be reasonable. With the exception of a slight deviation in only two cases (see Section 4.3), the estimates lead to formulas that are able to describe all our experiments for the Dob encryption scheme that will be investigated in Section 4.

3.2 Dimension of the Modifiers

The estimate given in eq. (11) requires knowledge of the dimension of (products of) the ideals $M^{(\alpha,\beta)}$. These will in turn depend on the chosen modifications V^α and Q^β. In this section we collect various results that will be needed to determine these dimensions. We start with the following elementary properties.

Lemma 2. *Consider $M^{(\alpha,\beta)} = (V^\alpha + Q^\beta)$, and positive integers $\alpha_0, \alpha, \beta_0, \beta, \nu$. Then the following holds:*

(i) $V^{\alpha_0}V^\alpha = V^{\alpha_0+\alpha}$ *and* $Q^{\beta_0}Q^\beta = Q^{\beta_0+\beta}$.
(ii) $V^{\alpha_0}Q^{\beta_0} \subseteq V^\alpha Q^\beta$ *if* $\alpha \le \alpha_0$ *and* $\beta \le \beta_0$.
(iii) $M^{(\alpha_0,\beta_0)}M^{(\alpha,\beta)} = M^{(\alpha_0+\alpha,\beta_0+\beta)} + V^{\alpha_0}Q^\beta + V^\alpha Q^{\beta_0}$.
(iv) $dim_\nu(M^{(\alpha,\beta)}) = dim_\nu(Q^\beta) + dim_\nu(V^\alpha) - dim_\nu(Q^\beta \cap V^\alpha)$.
(v)
$$dim_\nu(M^{(\alpha_0,\beta_0)}M^{(\alpha,\beta)}) = dim_\nu(M^{(\alpha_0+\alpha,\beta_0+\beta)}) + dim_\nu(V^{\alpha_0}Q^\beta)$$
$$+ dim_\nu(V^\alpha Q^{\beta_0}) - dim_\nu(M^{(\alpha_0+\alpha,\beta_0+\beta)} \cap V^{\alpha_0}Q^\beta)$$
$$- dim_\nu(M^{(\alpha_0+\alpha,\beta_0+\beta)} \cap V^\alpha Q^{\beta_0}) - dim_\nu(V^{\alpha_0}Q^\beta \cap V^\alpha Q^{\beta_0})$$
$$+ dim_\nu(M^{(\alpha_0+\alpha,\beta_0+\beta)} \cap V^{\alpha_0}Q^\beta \cap V^\alpha Q^{\beta_0}).$$

Proof. Properties (i) – (iv) follow from the appropriate definitions in a straightforward manner; we give a brief sketch of property (v) here. From property (iii) we know that $M^{(\alpha_0,\beta_0)}M^{(\alpha,\beta)}$ can be written as the sum of the three ideals $M^{(\alpha_0+\alpha,\beta_0+\beta)}$, $V^{\alpha_0}Q^\beta$ and $V^\alpha Q^{\beta_0}$. We start by summing the dimension of each of these three ideals individually. Any polynomial belonging to exactly two of these subideals is now counted twice, which is why we subtract by the combinations intersecting two of these ideals. Lastly, a polynomial belonging to all three of the subideals will, at this point, have been counted thrice, and then subtracted thrice. Hence, we add the dimension of intersecting all three subideals.

The dimension $\dim_\nu(V^\alpha)$ can be further inspected using the following result.

Lemma 3. *Suppose that $v_1, \ldots, v_k$ are k linearly independent linear forms in $\overline{B}(n)$. Then*

$$dim_\nu(V^\alpha) = \sum_{\substack{i \geq \alpha, j \geq 0 \\ i+j=\nu}} \binom{k}{i} \binom{n-k}{j} \tag{13}$$

holds under the conventions that $\binom{a}{b} = 0$ if $b > a$, and $\binom{a}{0} = 1$.

Proof. As $v_1, \ldots, v_k$ are linearly independent, we can choose $n-k$ linear forms of $\overline{B}(n)$, $w_{k+1}, \ldots, w_n$, that constitute a change of variables

$$\overline{B}(n) \simeq \overline{B}' = \mathbb{F}_2[v_1, \ldots, v_k, w_{k+1}, \ldots w_n]/\langle v_1^2, \ldots, w_n^2 \rangle.$$

For any monomial $\gamma \in \overline{B}'$, we will define $\deg_v(\gamma)$ as its degree in the $v_1, \ldots, v_k$-variables, and $\deg_w(\gamma)$ as its degree in the variables $w_{k+1}, \ldots, w_n$. The elements of V^α of (total) degree ν, is now generated (in $\overline{B}'$ as an $\mathbb{F}_2$–vector space) by all monomials γ such that $\deg_v(\gamma) \geq \alpha$ and $\deg_v(\gamma) + \deg_w(\gamma) = \nu$. The number of all such monomials are counted in eq. (13).

Lemma 4. *Let $q_1^h, \ldots, q_t^h$ be a D_0–semi–regular system of homogneous quadratic polynomials over $\overline{B}(n)$. Then, for any $2 \leq \nu < D_0$, we have*

$$dim_\nu(Q^1) = \binom{n}{\nu} - [z^\nu]T_{t,n}(z),$$

where $[z^\nu]T_{t,n}(z)$ denotes the coefficient of the monomial z^ν in the expansion of the series $T_{t,n}(z)$, as given in eq. (3).

Proof. By assumption, the series $T_{t,n}(z)$ coincides with the Hilbert series of $\overline{B}(n)/Q^1$, for the terms with degree $2 \leq \nu < D_0$. From the additive property of the Hilbert function, we have that $\dim_\nu(Q^1) = \dim_\nu(\overline{B}(n)) - [z^\nu]T_{t,n}(z)$, and it is well–known that $\dim_\nu(\overline{B}(n)) = \binom{n}{\nu}$.

Lemma 5. *Suppose that $(v_1, \ldots, v_k, q_1, \ldots, q_t)$ is D_0–semi–regular, and consider $1 \leq \alpha \leq k$ and $1 \leq \beta \leq t$. Then*

$$(V^\alpha \cap Q^\beta)_\nu = (V^\alpha Q^\beta)_\nu,$$

holds for all $\nu < D_0$.

Proof. (Sketch) The product of any pair of ideals is contained in their intersection. For the other direction, consider a non–trivial element $e \in (V^\alpha \cap Q^\beta)_\nu$. Then, for some polynomials f_i, g_j, we can write $e = \sum f_i q_{i_1}^h \cdots q_{i_\beta}^h \in Q_\nu^\beta$, and $e = \sum g_j v_{j_1} \cdots v_{j_\alpha} \in V_\nu^\alpha$, which yields the syzygy

$$\sum f_i(q_{i_1}^h \cdots q_{i_\beta}^h) + \sum g_j(v_{j_1} \cdots v_{j_\alpha})^h = 0.$$

By assumption, all syzygies of degree $< D_0$ of $(v_1, \ldots, v_k, q_1^h, \ldots, q_t^h)$ will be generated by the field and Koszul syzygies of the v_i– and q_j^h–polynomials. It follows that (after possibly reducing by syzygies generated by only $q_1^h, \ldots, q_t^h$) we have $f_i \in V^\alpha$. Similarly, we have $g_j \in Q^\beta$. In particular, $e \in V^\alpha Q^\beta$.

A general characterisation of the ideal $V^\alpha Q^\beta$ is trickier. We are content with discussing some special cases of its dimension, which will be of interest to us.

Example 1 *Suppose that $(v_1, \ldots, v_k, q_1, \ldots, q_t)$ is D_0–semi–regular, and let $1 \leq \alpha \leq k$ and $1 \leq \beta \leq t$.*

(a) The generators of $V^\alpha Q^\beta$ are of degree $\alpha + 2\beta$, hence $dim_\nu(V^\alpha Q^\beta) = 0$ for all $\nu < \alpha + 2\beta$. (This also holds without the D_0–semi–regularity assumption).

(b) Suppose furthermore that $D_0 > \alpha + 2\beta + 1$. Then $dim_{(\alpha+2\beta+1)}(V^\alpha Q^\beta) = \binom{t}{\beta} dim_{\alpha+1}(V^\alpha)$. To see this, note that $\langle V^\alpha Q^\beta \rangle_{\alpha+2\beta+1}$ is generated by elements of the form $v_{l_1} \ldots v_{l_\alpha} q_{c_1} \ldots q_{c_\beta} x_r$, where $1 \leq l_1 < \ldots < l_\alpha \leq k$, $1 \leq c_1 < \ldots < c_\beta \leq t$ and $1 \leq r \leq n$. The semi–regularity assumption assures that there will be no cancellations (save for the ones already accounted for in $dim_{\alpha+1}(V^\alpha)$).

(c) Suppose furthermore that $D_0 > \alpha + 2\beta + 2$, then $dim_{(\alpha+2\beta+2)}(V^\alpha Q^\beta) = \binom{t}{\beta} dim_{\alpha+2}(V^\alpha) - \binom{k}{\alpha}\left[\binom{t}{\beta} t - \binom{t}{\beta+1}\right]$. The reasoning is similar to (b), with the difference that $dim_{\alpha+2}(V^\alpha)$ will now include the polynomials of the form $q_c^h(v_{l_1} \ldots v_{l_\alpha})^h$. There are $\binom{k}{\alpha}\left[\binom{t}{\beta} t - \binom{t}{\beta+1}\right]$ combinations of these that will reduce to 0 over $\overline{B}(n)$ (when multiplied with the combinations $q_{c_1}^h \ldots q_{c_\beta}^h$).

3.3 Dimension of $\mathcal{P}_{M^{(\alpha,\beta)}M^{(2,1)}}$

As noted in section 3.1, we want $\mathcal{P}_{M^{(\alpha,\beta)}M^{(2,1)}}$ to be $M^{(\alpha,\beta)}M^{(2,1)}\langle \mathcal{P}^h \rangle$, modulo the polynomials of the form $\psi^{\mathcal{P}^h}(ms)$, for $ms \in \mathcal{S}(\mathcal{F})_{M^{(\alpha,\beta)}M^{(2,1)}}$. Computing the dimension of $(M^{(\alpha,\beta)}M^{(2,1)}\langle \mathcal{P}^h \rangle)_\nu$ directly might be difficult, seeing that $\mathcal{P}^h$ depends on $M^{(2,1)}$. To tackle this, we start with the assumption that the cancellations in $M^{(\alpha,\beta)}M^{(2,1)}\langle \mathcal{P}^h \rangle$ are only generated by the 'generic' cancellations, and cancellations coming from the underlying structure, depending on $\mathcal{F}$. By 'generic' cancellations we mean those generated by the Koszul– or field syzygies in either the p_i^h– or m_j–polynomials. The assumption furthermore implies that the second type of cancellations will lie in the image of $\psi^{\mathcal{P}^h}(\mathcal{S}(\mathcal{F})_{M^{(\alpha,\beta)}M^{(2,1)}})$. Let $\mathcal{G}_{SR}$ be a system of homogeneous quadratic polynomials, of the same size and number of variables as $\mathcal{P}^h$, such that $\{V^1, Q^1, \mathcal{G}_{SR}\}$ forms a semi–regular system. With the assumption outlined above, we have

$$\dim_\nu(\mathcal{P}_{M^{(\alpha,\beta)}M^{(2,1)}}) = \dim_\nu(M^{(\alpha,\beta)}M^{(2,1)}\mathcal{G}_{SR}) - \dim_\nu(\mathcal{S}(\mathcal{F})_{M^{(\alpha,\beta)}M^{(2,1)}}). \quad (14)$$

Indeed, any would–be cancellations that are over–counted in the term $\dim_\nu(M^{(\alpha,\beta)}M^{(2,1)}\mathcal{G}_{SR})$ would be subtracted in $-\dim_\nu(\mathcal{S}(\mathcal{F})_{M^{(\alpha,\beta)}M^{(2,1)}})$.

$\mathcal{S}(\mathcal{F})_{M^{(\alpha,\beta)}M^{(2,1)}}$ requires knowledge of the underlying central map, $\mathcal{F}$, and will be dealt with in the next section. Computing the dimensions of the product ideal $M^{(\alpha,\beta)}M^{(2,1)}\mathcal{G}_{SR}$ has many similarities with the work that was done in the previous subsection. In particular, the dimension at degree ν is zero if the degrees of all of its generators are $> \nu$. We conclude with the following short example, which covers the other cases that will be the most relevant to us.

Example 2 *Let $\mathcal{G}_{SR}$ be a system of d homogeneous quadratic polynomials over $\overline{B}(n)$, such that $\{V^1, Q^1, \mathcal{G}_{SR}\}$ forms a semi–regular system. Then*

$$dim_\nu(M^{(2,1)}\mathcal{G}_{SR}) = n\big[dim_{\nu-2}(Q^1) + dim_{\nu-2}(V^2)\big],$$

holds for $\nu = 4, 5$.

4 Number of Degree Fall Polynomials in the Dob Encryption scheme

There are several ways to construct a central map $\mathcal{F} : \mathbb{F}_2^d \to \mathbb{F}_2^d$. For big–field schemes, the idea is to fix an isomorphism $\phi : \mathbb{F}_2^d \to \mathbb{F}_{2^d}$ between the vector space over the base field and the extension field, and choose two random invertible $d \times d$-matrices over $\mathbb{F}_2$, called S and T. $\mathcal{F}$ is then constructed as the composition $\mathcal{F} = S \circ \phi^{-1} \circ F \circ \phi \circ T$, where $F(X) \in \mathbb{F}_{2^d}[X]$, $\max_{w_2}(F) = 2$, and such that $F(X) = Y$ is easy to solve for any given Y. In particular, this ensures that $\mathcal{F}$ is a system of d quadratic polynomials, and ciphertexts can easily be decrypted with the knowledge of the secret S, T and F. There are two main ways in the literature to construct F with these properties:

1. $F(X) = X^e$, where $w_2(e) = 2$. This is the case for C^* [19].
2. $F(X) = \sum_{i=0}^{t} c_i X^{e_i}$, where we have $w_2(e_i) \leq 2$ for all i, and each e_i is bounded by a relatively small constant b. This is used in HFE [22].

Indeed, both C^* and HFE have been suggested with the *ip*–modification, known as PMI an ipHFE, respectively [7, 11]. These schemes were broken in [16, 13], by specialised attacks recovering the kernel of the linear forms of the *ip*–modification. Nevertheless, a later version of the C^* variant, PMI+ [8], also added the “ + ” modification in order to thwart this attack, and remains unbroken. We note that ipHFE, PMI and PMI+ all fits into the framework presented in section 3, and the techniques presented here can be used to understand their resistance against algebraic attacks (recall that the “+” modification does not increase the security versus algebraic attacks). A comprehensive study of these schemes are beyond the scope of this work, as we focus on a newer construction that utilizes both the *ip*– and Q_+–modification.

4.1 The Dob Encryption Scheme

The *Two–Face* family, introduced in [18], presents a third way to construct a function $F(X)$. Writing $Y = F(X)$, we get the polynomial equation

$$E_1(X, Y) = Y + F(X) = 0.$$

When F has the Two–Face property, it can be transformed into a different polynomial $E_2(X,Y) = 0$, which has low degree in X and have 2–weight at most 2 for all exponents in X. The degree of E_2 in Y can be arbitrary. Given Y, it is then easy to compute an X that satisfies $E_2(X,Y) = 0$, or equivalently, $Y = F(X)$.

For a concrete instantiation, the authors of [18] suggest the polynomial

$$F(X) = X^{2^m+1} + X^3 + X, \tag{15}$$

where $d = 2m-1$. Dobbertin showed in [12] that F is a permutation polynomial. In [18], based on the results of [12], it is further pointed out that

$$E_2(X,Y) = X^9 + X^6Y + X^5 + X^4Y + X^3(Y^{2^m} + Y^2) + XY^2 + Y^3 = 0$$

holds for any pair $Y = F(X)$. Note that F itself has high degree in X, but the highest exponent of X found in E_2 is 9 and all exponents have 2–weight at most 2.

The public key $\mathcal{F}$ associated with eq. (15) under the composition described at the beginning of section 4 is called *nude Dob*, and was observed in [18] to be weak. More precisely, experiments show that the associated multivariate system has solving degree three. Indeed, in appendix D we will show that this is the case for any d.

The (full) Dob encryption scheme is made by extending nude Dob with the two modifications, Q_+ and ip, as described at the beginning of section 3. The public key is the d quadratic polynomials $\mathcal{P}$, constructed according to eq. (6). The secret key consists of S, T, H_{ip} and H_{Q_+}. The plaintext space of the scheme is $\mathbb{F}_2^d$ and encryption is done by evaluating $y = \mathcal{P}(x)$, producing the ciphertext y.

To decrypt, the receiver of a ciphertext y guesses on the values of $v_i(x)$ and $q_j(x)$ for all $1 \leq i \leq k$ and $1 \leq j \leq t$, and computes the corresponding values of the polynomials in H_{ip} and H_{Q_+}. These values are added to y, removing the effect of the modifiers when the guess is correct. The resulting value y' is then the ciphertext of the nude Dob. This can be decrypted by first multiplying y' with S^{-1}, resulting in Y from the central mapping, which is then inverted using E_2 and multiplied with T^{-1} to recover the candidate plaintext x_0. The initial guess is then verified by checking if all $v_i(x_0)$ and $q_j(x_0)$ indeed evaluate to the guessed values.

In order for decryption to have an acceptable time complexity, the size of the modifications, k and t, can not be too large. To decrypt a ciphertext one must on the average do 2^{k+t-1} inversions of $\mathcal{P}$ before the correct plaintext is found. In [18] it is suggested to use $k = t = 6$ for 80–bit security.

For the remainder of this work, we let $\mathcal{F}$ and $\mathcal{P}$ denote the public keys of nude Dob and the (full) Dob encryption scheme, respectively.

4.2 Syzygies of the Unmodified Dob Scheme

The goal of this subsection is to estimate the dimension of $\mathcal{S}(\mathcal{F})_\nu$, for $\nu = 3, 4, 5$. We start by inspecting F (eq. (15)) over the extension field $\mathbb{F}_{2^d}[X]/\langle X^{2^d} + X\rangle$.

Note that $\max_{w_2}(F) = 2$, and consider the following polynomials:

$$G_1 = XF \qquad \text{and} \qquad G_2 = (X^{2^m} + X^2)F. \tag{16}$$

One finds that G_1 and G_2 are both products of F and a polynomial of 2–weight one, but the resulting polynomials still have $\max_{w_2}(G_i) = 2$. They are then examples of polynomials satisfying eq. (5) from section 2.1, and will correspond to $2d$ degree fall polynomials at degree three, down to quadratic polynomials. They form all the syzygies we expect at degree three, hence we set

$$\dim_3(\mathcal{S}(\mathcal{F})) = 2d. \tag{17}$$

Recall that it was noted in [18] that experiments of nude Dob had a solving degree of three, though the authors did not provide a proof that this is always the case. The presence of G_1 and G_2 ensures that the first fall degree of nude Dob is three. A complete proof that the solution of nude Dob can be found by only considering polynomials of degree three is a little more involved, and is included in appendix D.

Things get more complicated for dimensions $\nu > 3$. While we expect the two polynomials G_1 and G_2 to generate a significant part of the syzygies, we also expect there to be other generators, as well as cancellations to keep track of. Due to the complexity of fully characterizing the higher degree parts of $\mathcal{S}(\mathcal{F})$, we instead found an expression for its dimension at degrees $\nu = 4, 5$ experimentally. The experimental setup is further described at the end of this subsection. Note that the formulas we present in this subsection will be a multiple of d. This strongly suggests that all the syzygies of the system come from its extension field structure. These relations could then, in principle, be written out analytically as was the case for $\nu = 3$. In particular, this makes it reasonable to expect the formulas to continue to hold for larger values of d (i.e., beyond our experimental capabilities).

In the subsequent formulas we introduce the following notation, which will be useful to us later. Whenever counting the syzygies that can be generated from syzygies of lower degree, we will multiply by n (the number of variables in an associated multivariate system), as opposed to d. For instance, let $(g_{i,1} \ldots, g_{i,d})$, $1 \le i \le d$ denote the d multivariate syzygies associated with G_1. Then $x_j(g_{i,1} \ldots, g_{i,d})$, $1 \le j \le n$ are syzygies at $\nu = 4$, and we will count all of these as[4] nd. For the Dob encryption scheme we of course have $n = d$, so this distinction may seem unnecessary at the moment, but later, in section 5, we will also consider the case $n < d$ as an attacker may fix certain variables.

For $\nu = 4$, we find the following expression:

$$\dim_4(\mathcal{S}(\mathcal{F})) = (2n - 1)d, \tag{18}$$

where we note that the term $2nd$ has been generated by G_1 and G_2, as described above.

[4] Not all of these will be linearly independent in $\mathcal{S}(\mathcal{F})$. For example, the d syzygies associated with $(X^{2^m} + X^2)G_1$ will correspond to syzygies in $\mathcal{T}(\mathcal{F}^h)$. This does not really matter, as the expressions eq. (18) and eq. (19) corrects for this.

For $\nu = 5$, we have

$$\dim_5(\mathcal{S}(\mathcal{F})) = \left(2\binom{n}{2} - n - 2d - 20\right)d. \tag{19}$$

Once more, some of these terms can be understood from the syzygies of lower degrees. The contribution from the polynomials G_1 and G_2 from $\nu = 3$ will now be the $2\binom{n}{2}d$ term. The term '$-d$' from $\nu = 4$ will now cause the '$-nd$' term.

Experimental Setup. The experiments used to test eq. (18) and eq. (19) have been done as follows. The public polynomials of nude Dob are first generated, and we consider their upper homogeneous part, $\mathcal{F}^h$, over $\overline{B}(d)$. $\mathrm{Dim}_\nu(\mathcal{S}(\mathcal{F}))$ is computed as the dimension of the kernel of the homogeneous Macaulay matrix $\overline{M}_\nu(\mathcal{F}^h)$, minus $\dim_\nu(\mathcal{T}(\mathcal{F}^h))$. For $\nu = 4, 5$ we tested all odd d, $25 \leq d \leq 41$, all matching the values predicted by eq. (18) and eq. (19).

4.3 Degree Fall Polynomials of the (modified) Dob Scheme

We now have all the tools needed to write out explicit formulas for (variants of) the estimates $N_\nu^{(\alpha,\beta)}$, $\nu \leq 5$, for the Dob scheme. The approach for the formulas is as follows. Equation (11) is used as a foundation, and $\dim_\nu(\mathcal{S}(\mathcal{F}))$ is given according to section 4.2. For the dimension of the modifiers, and $\mathcal{P}_{M^{(\alpha,\beta)}M^{(2,1)}}$, we will combine the results discussed in section 3.2 and section 3.3. In particular, we will assume that the chosen modifying polynomials $\{v_1, \ldots, v_k, q_1, \ldots, q_t\}$ form a $(\nu + 1)$–semi–regular system. The dimensions that are not covered by combining the results discussed so far, will be commented on separately. For the convenience of the reader, the non–trivial dimensions have been marked with an overbrace in the equations. The exceptions are eq. (24) and eq. (25), which are covered in greater depth in appendix B. Recall also our convention that $\binom{a}{b} = 0$, if $b > a$, and $\binom{a}{0} = 1$.

$\boldsymbol{\nu = 3}$. At this degree we only consider $N^{(0,0)}$.

$$N_3^{(0,0)} = \overbrace{2d}^{\dim_3(\mathcal{S}(\mathcal{F}))} - \overbrace{\left((n-k)\binom{k}{2} + \binom{k}{3}\right)}^{\dim_3(V^2)} - \overbrace{nt}^{\dim_3(Q^1)} . \tag{20}$$

$\boldsymbol{\nu = 4}$.

$$\begin{aligned} N_4^{(0,0)} = {} & \overbrace{(2n-1)d}^{\dim_4(\mathcal{S}(\mathcal{F}))} + \overbrace{d\left(t + \binom{k}{2}\right)}^{\dim_4(\mathcal{P}_{M^{(2,1)}})} - \overbrace{\left(t\binom{n}{2} - \binom{t}{2} - t\right)}^{\dim_4(Q^1)} \\ & - \overbrace{\left(\binom{k}{2}\binom{n-k}{2} + \binom{k}{3}(n-k) + \binom{k}{4}\right)}^{\dim_4(V^2)} + \overbrace{t\binom{k}{2}}^{\dim_4(Q^1 \cap V^2)} . \end{aligned} \tag{21}$$

At $\nu = 4$, we also consider the estimate $N_4^{(1,0)}$, i.e., multiplying everything with the k linear forms from the *ip*–modifier. In particular, this means that $(\mathcal{S}(\mathcal{F})_{M^{(1,0)}})_4$ is spanned by the combinations $v_j^h(g_{i,1}\dots,g_{i,d})$, $1 \le j \le k$ and $1 \le i \le 2d$, where we recall that $(g_{i,1}\dots,g_{i,d})$ denote the $2d$ multivariate syzygies associated with G_1 and G_2 (eq. (16))

$$\begin{aligned} N_4^{(1,0)} = {} & \overbrace{2kd}^{\dim_4\left(\mathcal{S}(\mathcal{F})_{M^{(1,0)}}\right)} - \overbrace{\left(\binom{k}{3}(n-k)+\binom{k}{4}\right)}^{\dim_4(V^3)} \\ & - \overbrace{t\left(k(n-k)+\binom{k}{2}\right)}^{\dim_4(Q^1V^1)}. \end{aligned} \tag{22}$$

***ν* = 5.** At degree 5, $\mathcal{S}(\mathcal{F})_{M^{(2,1)}}$ (in eq. (14)) is no longer trivial. Indeed, it will now consist of the possible combinations $v_{j_1}^h v_{j_2}^h(g_{i,1}\dots,g_{i,d})$ and $q_j^h(g_{i,1}\dots,g_{i,d})$.

$$\begin{aligned} N_5^{(0,0)} = {} & \overbrace{\left(2\binom{n}{2}-n-2d-20\right)d}^{\dim_5(\mathcal{S}(\mathcal{F}))} - \overbrace{\left(t\binom{n}{3}-n\binom{t}{2}-tn\right)}^{\dim_5(Q^1)} \\ & - \overbrace{\left(\binom{k}{2}\binom{n-k}{3}+\binom{k}{3}\binom{n-k}{2}+\binom{k}{4}(n-k)+\binom{k}{5}\right)}^{\dim_5(V^2)} \\ & + \overbrace{t\left(\binom{k}{2}(n-k)+\binom{k}{3}\right)}^{\dim_5(Q^1\cap V^2)} \\ & + \overbrace{ntd + d\left(\binom{k}{2}(n-k)+\binom{k}{3}\right) - 2dt - 2d\binom{k}{2}}^{\dim_5\left(\mathcal{P}_{M^{(2,1)}}\right)}. \end{aligned} \tag{23}$$

As mentioned above, it is a bit more involved to derive $N_5^{(1,1)}$ and $N_5^{(2,1)}$, and we will refer to appendix B for more details. It would also appear that our assumptions are slightly off for these two estimates, as our experiments consistently yield $4d$ more degree fall polynomials than we are able to predict (see remark 3 for more details). We present the experimentally adjusted versions in Equations (24) and (25):

$$\begin{aligned} N_5^{(1,1)} = {} & d\left(k(2n-k-2)+t(2+k)+\binom{k}{3}+4\right) - \binom{t}{2}n - \binom{k}{3}\binom{n-k}{2} \\ & - \binom{k}{5} - \binom{k}{4}(n-k) - t\left(k\binom{n-k}{2}+\binom{k}{2}(n-k)-kt\right). \end{aligned} \tag{24}$$

$$\begin{aligned} N_5^{(2,1)} = 2d\Big(\binom{k}{2} + t + 2\Big) - \Big(\binom{k}{4}(n-k) + \binom{k}{5}\Big) \\ - t\Big(\binom{k}{2}(n-k) + \binom{k}{3}\Big) - \binom{t}{2}n. \end{aligned} \tag{25}$$

5 Experimental Results on Degree Fall Polynomials

In the previous section we developed the theory on how to estimate the number of first fall polynomials, ending up with several formulas. This section is focused on the accuracy of these formulas, and how they can be used by an attacker. Note that since we are interested in the unique structure of the Dob encryption scheme, we will always assume that 'generic' degree fall polynomials do not interfere. More specifically, when inspecting a system of d polynomials in n variables at degree ν, we assume that d and n is chosen such that $D_{reg}(d,n) > \nu$.

5.1 Fixing Variables

The formulas separate d, the size of the field extension, and n, the number of variables. While the Dob encryption scheme uses $d = n$, an attacker can easily create an overdetermined system with $n < d$ by fixing some variables. This approach, known as the hybrid method, can be viewed as a trade–off between exhaustive search and Gröbner basis techniques, and its benefits are well–known for semi–regular sequences [3]. From eqs. (20) to (25), we find that for the relevant choices of parameters (d, t, k), a greater difference between n and d can increase the number of degree fall polynomials. This means that a hybrid method will have a more intricate effect on a Dob system, than what we would expect from random systems. To a certain extent, an attacker can "tune" the number of degree fall polynomials, by choosing the amount of variables to fix. Of course, if the intent is to find a solution of the polynomial system through a Gröbner basis, this comes at the added cost of solving the system 2^r times, where r is the number of fixed variables, but in section 6 we will present a different attack that circumvents this exponential factor.

Finally, one could ask whether it is reasonable to expect eqs. (20) to (25) to be accurate after fixing a certain number of variables. It is, for instance, possible that different degree fall polynomials will cancel out, as certain variables are fixed. However, if past experience with the hybrid method is any indicator, such cancellations are very rare, and we see no reason that the extension field structure increases the probability for such cancellations to happen. As we will see in section 5.3 this is supported by the experiments we have run; the formulas remain precise, even as n is varied.

5.2 Using the Degree Fall Formulas

We briefly recall how the formulas found in section 4.3 relate to the public polynomials of a Dob encryption scheme. Let $\mathcal{P}$ be the polynomial system associated with a Dob scheme of fixed parameters (d, n, t, k) (where n is as described in section 5.1). We expect the non–trivial dimension (i.e., the dimension of the part that is not generated by $\mathcal{T}(\mathcal{F})$) of the kernel of $\overline{M}_\nu(\mathcal{P})$ to be given by the maximal of the formulas $N_\nu^{(\alpha,\beta)}$, for $\nu = 3, 4, 5$.

If a step–wise algorithm such as F_4 is used, we expect the formulas to predict the number of degree falls polynomials, but *only* at the first fall degree. Suppose, for instance, that $N_3 = 0$, but $N_4 > 0$. Then this algorithm runs a second step at degree 4, using the newly found degree fall polynomials. This means that there are effectively more available polynomials in the system when (if) a step of degree 5 is performed, and in this case we do not expect the formulas we have for N_5 to be accurate.

Note in particular that if all the formulas we have are non–positive, an attacker is likely required to go up to step degree ≥ 6 in order to observe first fall polynomials.

5.3 Experimental Results

We have run a number of experiments with the Dob system of varying parameters (d, n, t, k). A subset of them is presented in table 2, and the rest can be found in appendix G. Gröbner bases of the systems were found using the F_4 algorithm implemented in the computational algebra system Magma. The script used for the experiments is available at [17].

In table 2 (and appendix G) we use the following notation. 'D_{ff}' is the experimentally found first fall degree. 'N (predicted)' is the number of first fall polynomials as predicted by the equations in section 4.3. 'N (Magma)' is the number of first fall polynomials read from the verbose output of Magma, written as 'degree : {# degree fall polynomials at this degree}'. The solving degree D_{solv} was found experimentally by Magma. This has been measured as the degree where the most time consuming step of the algorithm took place. In the instances where the algorithm did not run to completion due to memory constraints, we give D_{solv} as $\geq X$, where X is the degree of the step where termination occurred. The degree of regularity for semi–regular systems of the same size, $D_{reg}(d, n)$, is also given. 'Step Degrees' lists the degrees of the steps that are being performed by F_4 up until linear relations are found. Once a sufficient number of linear relations are found, Magma restarts F_4 with the original system, as well as these linear relations. This restart typically needs a few rounds before the entire basis is found, but its impact on the running time of the algorithm is negligible, which is why we have chosen to exclude it when listing the step degrees. For convenience, the step where first fall polynomials are found is marked in blue and the solving step marked in red. Purple is used to mark the steps where these two coincide.

Table 2: Degree fall polynomials for Dob encryption schemes of various parameters.

d	n	t	k	D_{ff}	N	N	D_{solv}	Step
		(+)	*(ip)*		(predicted)	(Magma)	($D_{reg}(d,n)$)	Degrees
53	53	0	0	3	$N_3^{(0,0)}$: 106	2:106	3 (9)	2,3,3
53	53	0	3	4	$N_4^{(0,0)}$: 1999	3:1999	4 (9)	2,3,4,4
53	53	3	0	4	$N_4^{(0,0)}$: 1596	3:1596	4 (9)	2,3,4,4
59	29	0	7	4	$N_4^{(1,0)}$: 21	3:21	5 (5)	2,3,4,4,5
37	25	2	3	4	$N_4^{(0,0)}$: 692	3:692	4 (5)	2,3,4,4
31	29	0	8	5	$N_5^{(1,1)}$: 478	4:478	5 (6)	2,3,4,5,5,5
31	30	0	8	5	$N_5^{(2,1)}$: 264	4:264	5 (6)	2,3,4,5,5,5,4
39	37	1	7	5	$N_5^{(2,1)}$: 136	4:136	≥ 6 (7)	2,3,4,5,5,5,6...
57	38	4	6	5	$N_5^{(1,1)}$: 2086	4:2086	≥ 6 (6)	2,3,4,5,5,6. . .
57	37	4	6	5	$N_5^{(1,1)}$: 2847	4:2847	5 (6)	2,3,4,5,5
129	50	6	6	5	$N_5^{(0,0)}$: 64024	4:64024	≥ 5 (6)	2,3,4,5,5...

A first observation is that in all experiments we find that 'N (predicted)' matches 'N (Magma)'. We also find that fixing variables affects the cross–over point between the formulas $N_\nu^{(\alpha,\beta)}$, as for instance seen in the rows 6 and 7. We note that $N_\nu^{(0,0)}$ tend to be dominant when $n << d$, and that $N_5^{(2,1)}$ only seems to have an impact when k is large and t is small.

For the majority of cases we observe that $D_{ff} = D_{solv}$ or $D_{solv} + 1$, but one should be careful in drawing any conclusions from this, seeing that our experiments are in practice limited to computations of $D < 6$. The relation between n and D_{solv} is also noteworthy. For instance, in row 9 we have $d = 57$ and $n = 38$; D_{ff} is 5, but $D_{solv} \geq 6$. In row 10 we fix one more variable, $n = 37$ (while keeping everything else as before), and find $D_{solv} = 5$.

Impact on Known Attacks. The solving degree of big field schemes are often estimated using the first fall degree. In cases where $D_{solv} > D_{ff}$, we observed instances where it is beneficial for an attacker to fix (a few) variables in order to lower the D_{solv} for each guess. Without a better understanding of D_{solv} and how it is affected by fixing variables, it seems that the approximation $D_{ff} \approx D_{solv}$ is conservative, yet reasonable, when estimating the complexity of direct/hybrid attacks against Dob system.

Another attack that may greatly benefit from the detailed formulas for degree fall polynomials obtained in section 3, is an adapted version of the distinguishing attack that was proposed for HFEv- (Section 5 in [10]). An attacker fixes random linear forms, and distinguishes between the cases where (some of) the fixed linear forms are in the span of $(v_1, \ldots, v_k)$, and when none of them are, by the use of Gröbner basis techniques. Indeed, if *one* of the fixed linear forms are in this span, the number of degree fall polynomials will be the same as for a system with $k-1$

ip linear forms. Hence, a distinguisher based on the formulas presented here will work even without a drop in first fall degree, making the attack more versatile.

The deeper understanding for how the modifiers work allows for an even more efficient attack on the Dob scheme, which we now present.

6 A New Attack on the Dob Encryption Scheme

In the previous two sections we have studied how degree fall polynomials can occur in the Dob scheme, and have verified the accuracy of our resulting formulas through experiments. In this section we will show how all these insights can be combined to a novel attack. In section 6.1, we shall see that adding an extra polynomial to the system can leak information about the modification polynomials. We will see how this information can be used to retrieve (linear combinations of) the secret *ip* linear forms, and the homogeneous quadratic part of the Q_+ modification, in sections 6.2 and 6.3. We investigate how Gröbner basis algorithms perform with this extra information in section 6.4, and finally discuss the complexity of the entire attack in section 6.5.

6.1 Adding an Extra Polynomial

In section 3.1 we discussed how products of the modifiers and public polynomials affect the number of degree fall polynomials, through $\mathcal{P}_{M^{(2,1)}}$. One would also expect a similar effect to take place when adding a random polynomial to the system.

Consider a set of parameters for the Dob scheme, where the number of first fall polynomials is determined by $N_\nu^{(0,0)}$, for some $\nu > 3$. Let $\mathcal{P}$ be the public key of this scheme, and consider a randomly chosen homogeneous polynomial p_R of degree $\nu - 2$. As it is unlikely that the randomly chosen p_R has any distinct interference with $\mathcal{P}$, we expect $(\langle p_R\rangle \cap M^{(2,1)})_\nu$ to be generated by the possible combinations $p_R q_i^h$, and $p_R(v_j v_l)^h$. Furthermore, since the generators of $\mathcal{S}(\mathcal{F})$ have degree at least 3, we do not expect any collision between $\psi^{\mathcal{P}^h}(\mathcal{S}(\mathcal{F}))$ and $\langle p_R\rangle$ at degree ν (cf. section 3.3). From these considerations, we estimate the number of degree fall polynomials for the system $\{\mathcal{P}, p_R\}$ at degree ν to be:

$$N_\nu(\{\mathcal{P}, p_R\}) = N_\nu^{(0,0)}(\mathcal{P}) + t + \binom{k}{2}. \tag{26}$$

We ran a few experiments that confirm this intuition, the details are given in table 3. First, we confirmed that the degree fall polynomials of $\mathcal{P}$ were indeed given by $N_\nu^{(0,0)}(\mathcal{P})$, before applying Magma's implementation of the F_4 algorithm on the system $\{\mathcal{P}, p_R\}$. Recall also our convention that $\binom{0}{2} = 0$ when applying eq. (26).

With all this in mind, assume for the moment that $d = n$, and consider a homogeneous Macaulay matrix of $\{\mathcal{P}^h, p_R\}$ at degree ν, $\overline{M}_\nu(\{\mathcal{P}^h, p_R\})$. Any

Table 3: First fall polynomials of Dob encryption schemes with an added, randomly chosen polynomial p_R.

d	n	$\deg(p_R)$	t	k	D_{ff}	N	N
			(Q_+)	(ip)		(predicted)	(Magma)
31	29	2	2	2	4	$N_4 : 705$	3:705
45	30	2	6	0	4	$N_4 : 342$	3:342
75	39	3	6	6	5	$N_5 : 4695$	4:4695
39	37	3	6	0	5	$N_5 : 9036$	4:9036

element in the (left) kernel of this matrix can in general be written as:

$$h_R p_R + \sum_{i=1}^{d} h_i p_i^h = 0, \tag{27}$$

for some homogeneous quadratic polynomials $h_i \in \overline{B}(d)_{\nu-2}$, $1 \leq i \leq d$, and $h_R \in \overline{B}(d)_2$. From the discussion above, we expect that the only way p_R contributes to these kernel elements is through the trivial syzygies, multiplications with p_i^h or p_R, and through multiplying with the generators of $M^{(2,1)}$. It follows that any polynomial h_R, from eq. (27), will be in the span of[5]

$$\mathcal{H} := \{p_1^h, \ldots, p_d^h, p_R, q_1^h, \ldots, q_t^h, (v_1 v_2)^h, \ldots, (v_{k-1} v_k)^h\}. \tag{28}$$

Hence, given enough kernel elements of $\overline{M}_\nu(\{\mathcal{P}^h, p_R\})$, a set of generators of Span($\mathcal{H}$) can be found. In the next subsection we will generalise this observation to the case where a number of variables are fixed, i.e. $n < d$.

6.2 Gluing Polynomials

Let W_η denote a non-empty subset of r variables, i.e. $W_\eta = \{x_{\eta_1}, \ldots, x_{\eta_r}\}$ for integers $1 \leq \eta_1 < \ldots < \eta_r \leq d$. For $n = d - r$, there is a natural projection map associated to W_η, $\pi_{W_\eta} : B(d) \to B(d)/W_\eta \simeq B(n)$, that fixes the variables in W_η to 0. For any polynomial system $\mathcal{R}$ over $B(d)$, we will also write $\pi_{W_\eta}(\mathcal{R})$ to mean the system consisting of all polynomials in $\mathcal{R}$ under π_{W_η}. Suppose now that the number of first fall polynomials of a Dob system $\mathcal{P}$ is given by $N_\nu^{(0,0)}$, after fixing r variables to 0, i.e., $n = d - r$. Let W_η be the set of variables we fix. Following a similar line of reasoning as in section 6.1, we find that $\pi_{W_\eta}(h_R)$ from a kernel element of the Macaulay matrix associated with $\pi_{W_\eta}(\{P^h, p_R\})$ will no longer be in the span of $\mathcal{H}$, but rather lie in the span of $\pi_{W_\eta}(\mathcal{H})$. To ease notation, we will write $\mathcal{H}_\eta = \pi_{W_\eta}(\mathcal{H})$. A natural question is whether we can

[5] If p_R has degree ≥ 3, then the syzygy $p_R^2 + p_R = 0$ will be of degree $> \nu$. In this case p_R will not be among the generators of $\mathcal{H}$. We shall see later, in Remark (2), that the effect of p_R can also be removed in the degree 2 case, but at an added cost to the run time.

recover $\mathcal{H}$, by using different variable sets $W_1, \ldots, W_\rho$, and finding generators for the associated polynomial sets $\mathcal{H}_1, \ldots, \mathcal{H}_\rho$. We answer this question positively in this subsection.

Let $\widetilde{W_\eta} := \{x_1, \ldots, x_d\} \backslash W_\eta$ denote the complement of W_η, and note that $\mathcal{H}_\eta$ only contains information about the set of monomials $A(W_\eta) := \{x_i x_j \mid x_i, x_j \in \widetilde{W_\eta}\}$. In order to guarantee that the family $\mathcal{H}_1, \ldots, \mathcal{H}_\rho$ can give complete information about $\mathcal{H}$ we need to ensure that for any choice of $1 \leq i \leq j \leq d$, we have $x_i, x_j \in \widetilde{W_\eta}$ for at least one $1 \leq \eta \leq \rho$. In other words, the sets $\widetilde{W_1}, \ldots, \widetilde{W_\rho}$ must cover all possible quadratic monomials.

In practice, both d and the size r of the variable sets will be determined by the chosen Dob parameters[6]. This naturally leads to the following problem:

Definition 3 (The (Quadratic) (r,d)–Covering Problem). *For integers $1 < r < d-1$, find the smallest number ρ of variable sets, each of size r, such that*

$$A(W_1) \cup \ldots \cup A(W_\rho) = \{x_i x_j \mid 1 \leq i < j \leq d\}.$$

In Appendix E we present a constructive solution to this problem, which provides a good upper bound for ρ that is sufficient for our use case. The upper bound is given by the following lemma

Lemma 6. *The (Quadratic) (r,d)–Covering Problem is upper bounded by*

$$\rho \leq \binom{\left\lceil \frac{d}{\lfloor (d-r)/2 \rfloor} \right\rceil}{2}.$$

We illustrate the strategy for recovering $\mathcal{H}$ in the simple case when $d = 3r$. In this particular case, the method above yields $\rho = 3$, where W_1, W_2 and W_3 are pairwise, disjoint variable sets. We may write the following matrix:

$$\begin{array}{c|cccccc|}
 & W_1 * W_1 & W_1 * W_2 & W_1 * W_3 & W_2 * W_2 & W_2 * W_3 & W_3 * W_3 \\
H_1 & 0 & 0 & 0 & * & * & * \\
H_2 & * & 0 & * & 0 & 0 & * \\
H_3 & * & * & 0 & * & 0 & 0
\end{array}$$

Here $W_i * W_j$, $i, j \in \{1, 2, 3\}$, is understood as a list of the monomials $x_a x_b$ where $x_a \in W_i$ and $x_b \in W_j$ (under any fixed ordering and $a \neq b$), and we write H_l to mean the rows associated with a fixed set of generators for $\mathcal{H}_l$. A 0 in the matrix means that the respective submatrix is the zero matrix, whereas $*$ denotes that the submatrix may take non-zero values. By construction, if the submatrix whose rows are H_l, and columns are $W_i * W_j$, is denoted by $*$, then it forms a set of generators for $\mathcal{H}$ restricted to the monomials in $W_i * W_j$. In

[6] We will see later that the gluing also requires some overlap between the variable sets, but this is not a problem for the parameters we are interested in.

particular, the submatrix with columns $W_3 * W_3$ and rows H_1 spans the same row-space as the submatrix with columns $W_3 * W_3$ and rows H_2. We will use this observation to construct a new matrix, denoted $H_1 \cap_{W_3} H_2$, that combine the useful information from H_1 and H_2 in the following procedure.

1. Since $\{p_1^h, \dots, p_d^h, p_R\}$ are known, we start by finding $t+\binom{k}{2}$ vectors in the row space of H_2 that are linearly independent of $\pi_{W_2}(\{p_1^h, \dots, p_d^h, p_R\})$. Denote the set of these vectors Y_2.
2. If $|W_3 * W_3| >> d + t + \binom{k}{2} + 1$, then for each vector $y_i \in Y_2$, we can expect a unique vector z_i in the row space of H_1, such that $y_i + z_i$ is 0 along the columns associated with $W_3 * W_3$. Find such an z_i for each $y_i \in Y_2$ through Gaussian elimination.
3. We now have $t + \binom{k}{2}$ pairs (y_i, z_i) that are used to define the $(t + \binom{k}{2}) \times \binom{d}{2}$ matrix $(H_1 \cap_{W_3} H_2)$ over $\mathbb{F}_2$ in the following manner. For each row index i_0 and column index j_0, we define the entry at $[i_0, j_0]$ to be

$$(H_1 \cap_{W_3} H_2)[i_0, j_0] = \begin{cases} y_{i_0}[j_0], & \text{if } j_0 \text{ is associated with a monomial in } W_3 * W_3 \\ y_{i_0}[j_0] + z_{i_0}[j_0], & \text{otherwise.} \end{cases}$$

The above procedure uses the common information found in the columns of $W_3 * W_3$ to combine vectors from H_1 and H_2. We may think of this as "gluing" polynomials along $W_3 * W_3$, hence the name of the technique. Now consider the following matrix.

$$\begin{array}{c|cccccc|} & W_1 * W_1 & W_1 * W_2 & W_1 * W_3 & W_2 * W_2 & W_2 * W_3 & W_3 * W_3 \\ (H_1 \cap_{W_3} H_2) & * & 0 & * & * & * & * \\ H_3 & * & * & 0 & * & 0 & 0 \end{array}$$

Note in particular that the polynomials associated with $(H_1 \cap_{W_3} H_2)$ forms a set of generators for $\pi_{W_1 * W_2}(\mathcal{H})$. In order to recover the information of the monomials in $W_1 * W_2$, we need only glue the vectors of $(H_1 \cap_{W_3} H_2)$, with combinations from the row space of H_3, using the same procedure as described above. Since both $(H_1 \cap_{W_3} H_2)$ and H_3 may take non–zero values at $W_1 * W_1$ and $W_2 * W_2$, we expect the gluing to result in $t + \binom{k}{2}$ unique polynomials if $|(W_1 * W_1) \cup (W_2 * W_2)| >> d + t + \binom{k}{2} + 1$. By construction, all of the resulting $t + \binom{k}{2}$ polynomials associated with $(H_1 \cap_{W_3} H_2) \cap_{W_1} H_3$ will be in the span of $\langle p_1^h, \dots, p_d^h, p_R, q_1^h, \dots, q_t^h, \dots (v_i v_j)^h \dots \rangle$, but none of them in the span of $\langle p_1^h, \dots, p_d^h, p_R \rangle$. Hence we define $\mathcal{G}$ to be the set consisting of the polynomials $\{p_1^h, \dots, p_d^h, p_R\}$, as well as the polynomials associated with $(H_1 \cap_{W_3} H_2) \cap_{W_1} H_3$, and note that $\mathcal{G}$ is, by construction, a system of polynomials that are linearly equivalent to $\mathcal{H}$.

As a proof of concept, we implemented retrieving $\mathcal{G}$ from a toy example of the Dob scheme, with $d = 45$, $t = 6$ and $k = 0$, using the method described above. The interested reader can find more details in appendix C, example 3.

The General Case In the case of a general family of variable sets $W_1, \dots, W_\rho$, we will not be able to set up the straightforward matrices that was shown above. The gluing process can still be done in a similar, iterative manner. For instance, the submatrix associated with $\mathcal{H}_\eta$ will have 0 for each monomial $x_i x_j$ where x_i or $x_j \in W_\eta$, and $*$ otherwise. As above, we expect to be able to glue $\mathcal{H}_\eta$ with $\mathcal{H}_\psi$ if the number of their common $*$–monomials exceeds $d + t + \binom{k}{2} + 1$.

6.3 Retrieving the Linear Forms from *ip*

Suppose now that a set of generators $\mathcal{G}$ for Span($\mathcal{H}$) has been found, as described in section 6.2. The goal is to recover k linear forms that are generators for $\langle v_1, \dots, v_k \rangle$. In order to simplify our arguments we will assume $k \geq 5$. The cases $2 \leq k \leq 4$ will be discussed in Remark 1.

Consider the kernel of the homogeneous Macaulay matrix $\overline{M}_3(\mathcal{G})$. From the definition of $\mathcal{H}$ (*eq.* (28)), we find that Span($\mathcal{H}$) contains all the homogeneous nude Dob–polynomials $f_1^h, \dots, f_d^h$, as well as all the combinations $(v_i v_j)^h$, $1 \leq i < j \leq k$. Each polynomial $(v_i v_j)^h$ generates the two kernel elements $v_i(v_i v_j)^h$ and $v_j(v_i v_j)^h$ (which are trivial when working over $\overline{B}(d)$). The nude Dob–polynomials will generate the $2d$ kernel elements associated with the degree fall polynomials discussed in section 4.2. We would like to separate these two types of kernel elements. To this end, we suggest constructing a smaller system, $\mathcal{G}'$, by removing three polynomials from $\mathcal{G}$, that are in the span of $\{p_1^h, \dots, p_d^h\}$. Indeed, the idea is that this will work as a self–imposed minus modifier, which will remove the effect of the Dob–polynomials of $\mathcal{G}$ at degree 3.

On the other hand, some kernel elements generated by combinations of the $(v_i v_j)^h$–elements can still be observed for $\mathcal{G}'$ at degree 3. More specifically, suppose $\mathcal{G}'$ was created from $\mathcal{G}$ by removing p_1^h, p_2^h and p_3^h. Then Span($\mathcal{G}'$) may not necessarily contain $(v_1 v_j)^h$ itself, for any $2 \leq j \leq k$, but it will contain the combination $(v_1 v_j)^h + b_{1,j} p_1^h + b_{2,j} p_2^h + b_{3,j} p_3^h$, for some $b_{1,j}, b_{2,j}, b_{3,j} \in \mathbb{F}_2$. By considering these equations for all j, and eliminating p_1^h, p_2^h and p_3^h, we find that Span($\mathcal{G}'$) will contain a polynomial $z_1 = \sum_{j=2}^{k} a_j (v_1 v_j)^h$, where $a_2, \dots, a_k \in \mathbb{F}_2$ are not all 0, using the assumption that $k \geq 5$. The polynomial $v_1 z_1$ will subsequently be reduced to 0 over $\overline{B}(d)$. Similarly, we are guaranteed to find polynomials $z_2, \dots, z_k$. We assume that these are the only contributors to the kernel. In particular, this means that each kernel element of $\overline{M}_3(\mathcal{G}')$ can be written as $\sum l_i g_i = 0$, with $g_i \in \mathcal{G}'$, and each l_i a linear form in Span($\{v_1, \dots, v_k\}$). It follows that an attacker can retrieve a basis $v_1^*, \dots, v_k^*$ of $\langle v_1, \dots, v_k \rangle$, by determining k linearly independent l_i's from these kernel elements.

Remark 1. In the text above, we suggest removing $a = 3$ polynomials from $\mathcal{G}$, and assumed $k \geq 5$. We note that removing $a = 2$ polynomials is the smallest number needed for the prediction[7] $2d - ad$ to be non–positive, but this setting

[7] This formula follows a similar line of reasoning as in section 4.3, but with the minus modifier instead of Q_+. Cf. also [20] for a study on how the minus modifier affects degree fall polynomials for a somewhat related scheme.

could lead to some complications if the prediction turns out to be slightly off. Hence, $a \geq 3$ seems to be reasonable in order to exclude any interference from the Dob–structure at degree 3. Subsequently, we assume $k \geq a+2$ in order to guarantee the existence of the polynomials z_i.

When $k = 4$, an attacker might choose $a = 2$, as described above. Some experiments also seem to suggest that we can still find enough generators for $\langle v_1, v_2, v_3, v_4 \rangle$, even when choosing $a = 3$ (even though our arguments does not hold in this case). Hence we do not expect $k = 4$ to pose any real challenge for an attacker. Lastly, if $k = 2, 3$, then $\binom{k}{2}$ is so small that an attacker can skip this step of the attack altogether and simply guess the values for the $v_i v_j$–combinations directly in the step described in the next section.

The retrieval of $\mathcal{G}$ and $v_1^*, \dots, v_k^*$, as described in this subsection, has been implemented and verified on the toy example with parameters $d = 63$, $t = 1$ and $k = 4$. This is further described in example 4, in appendix C.

6.4 Solving the Extended Dob System

Assume now that an attacker has followed the steps described in the previous subsections, and has recovered a system $\mathcal{G}$ (section 6.2), as well as a basis $\{v_1^*, \dots, v_k^*\}$ that generates $\langle v_1, \dots, v_k \rangle$ (section 6.3). Now fix a set of generators $q_1^*, \dots, q_k^*$ for the polynomials that are in $\mathrm{Span}(\mathcal{G})$, but not in

$$\mathrm{Span}(\{p_1^h, \dots, p_d^h, p_R, (v_i^* v_j^*)^h \mid 1 \leq i < j \leq k\}).$$

With all this information, we consider the associated *extended Dob system*, $\mathcal{P}_E$, defined by:

$$\mathcal{P}_E := \{p_1, \dots, p_d, p_R, q_1^*, \dots, q_t^*, v_1^*, \dots, v_k^*\}. \tag{29}$$

For any given ciphertext, an attacker with access to an extended Dob system can guess constant values for the polynomials $p_R, q_1^*, \dots, q_t^*, v_1^*, \dots, v_k^*$, and check the guess by finding a Gröbner basis for $\mathcal{P}_E$.

Remark 2. It might be in the interest of an attacker to find a system $\mathcal{P}_E$ that does not depend on the random element p_R. If this is the case, one can choose a second random element p'_R, and construct a second system $\mathcal{P}'_E = \{p_1, \dots, p_d, p'_R, q_1'^*, \dots, q_t'^*, v_1'^*, \dots, v_k'^*\}$. A third system, $\mathcal{P}''_E = \{p_1, \dots, p_d, q_1''^*, \dots, q_t''^*, v_1''^*, \dots, v_k''^*\}$, independent of the random elements p_R and p'_R, can now be found by determining generators for $\mathrm{Span}(\mathcal{P}_E) \cap \mathrm{Span}(\mathcal{P}'_E)$. Solving the latter system $\mathcal{P}''_E$ could be easier, as one does not have to guess values for the random element. As a result, this could be a beneficial trade–off if the attack is not dominated by finding extended Dob systems. By abuse of notation, we will also call this latter system $\mathcal{P}''_E$ an extended Dob system, as long as we are careful about the factor 2 in the complexity estimates.

In order to get a better understanding of solving extended Dob systems, we introduce the following modification for multivariate schemes.

Definition 4. *For a polynomial system $\mathcal{P}'$, we define the modification $\mathcal{L}_+$ by choosing l_0 linear forms, and appending linear combinations of them to each polynomial in $\mathcal{P}'$.*

Consider an extended Dob system, $\mathcal{P}_E$, where all coefficients have been guessed correctly. Since q_i^* does not contain any information about the linear part of the q_i–polynomials, it follows that $\text{Span}(\mathcal{P}_E)$ will contain a Dob system that is only modified with the $\mathcal{L}_+$–modification, where $l_0 = t$. Moreover, this Dob system has d equations and $d - k$ variables[8]. The problem of estimating the complexity of finding a solution to $\mathcal{P}_E$, can then be reduced to that of estimating the complexity of finding a Gröbner basis for Dob with the $\mathcal{L}_+$–modification. While a thorough analysis of this $\mathcal{L}_+$–modification is beyond the scope of this work, we point out a couple of immediate properties.

Firstly, seeing that the first fall degree only depends on the upper homogeneous part of a polynomial system, it is unaffected by the $\mathcal{L}_+$–modification. In particular, we expect $2d$ degree fall polynomials at degree 3, as in the case for nude Dob (section 4.2). Secondly, if running an algorithm such as F_4, a second batch of degree fall polynomials will emerge at the first step of degree 4. To see this, note that Dob with the $\mathcal{L}_+$–modification can be written over the quotient ring $\mathbb{F}_{2^d}[X]/\langle X^{2^d} + X\rangle$ as

$$F_{\mathcal{L}_+}(X) = X(X^{2^m} + X^2) + L(X) + C_E, \tag{30}$$

where C_E is a constant in $\mathbb{F}_{2^d}$, and $L(X) = \sum_{i=1}^{m} c_i X^{2^i}$, with $c_i \in \mathbb{F}_{2^d}$, is a polynomial of binary weight one. $XF_{\mathcal{L}_+}$ is one of the combinations that induce degree fall polynomials at degree 3, and $X^4XF_{\mathcal{L}_+}$ will correspond to cubic[9] (multivariate) polynomials found at the second step of degree 3. Upon running a subsequent step at degree 4, the polynomial $L(X)X^4XF_{\mathcal{L}_+}$ will correspond to d multivariate cubic polynomials, and would hence be counted as degree fall polynomials.

We ran a few experiments for extended Dob systems, $\mathcal{P}_E$, the results of which can be found in appendix F.

6.5 Complexity of the Attack

The attack proposed in this section have two main parts. The first step is to construct an extended Dob system, $\mathcal{P}_E$. In the second step, an attacker solves this system for a particular ciphertext. Suppose an attacker fixes $d - n$ variables in order to find ρ polynomial systems $\mathcal{H}_1, \ldots, \mathcal{H}_\rho$ from the kernel elements of Macaulay matrices of degree $D_0 \geq 3$. The gluing operations, determining the linear forms $v_1^*, \ldots, v_k^*$, and the quadratic forms $q_1^*, \ldots, q_t^*$ only involve Macaulay

[8] Here we implicitly assume that k variables have been eliminated by the linear forms v_i^*.

[9] For nude Dob, the polynomial X^5F can be used to create linear polynomials (eq. (34)). The crucial difference is that in this case, the linear term X can be cancelled out at degree 3, whereas this is not possible for a general $L(X)$.

matrices of degree at most three. Hence, we expect the first step to be dominated by recovering generators for the polynomial systems $\mathcal{H}_i$. While the optimal choice of attack parameters may depend on the parameters of the Dob encryption scheme, as a rule of thumb it seems best to first minimize D_0, then n, and lastly ρ. In practice, minimizing n involves choosing the smallest n such that $D_{reg}(d, n) > D_0$, for a fixed d. As a result of this minimization, the Macaulay matrix at degree D_0 is close to being square, and hence we estimate the complexity of step one as $\rho\binom{n}{D_0}^{\omega}$, where ω is the linear algebra constant. Recall from remark 2 that this process is performed twice if the attacker wishes to remove the effect of p_R from $\mathcal{P}_E$; let $\delta = 1$ denote if this is the case, and $\delta = 0$ otherwise. It follows that the total attack complexity can be estimated as

$$\mathcal{C}_{\text{Attack}} = \max\left\{2^{\delta}\rho\binom{n}{D_0}^{\omega}, \mathcal{C}_{\mathcal{P}_E,\delta} \;\middle|\; \delta \in \{0,1\}\right\}, \tag{31}$$

where $\mathcal{C}_{\mathcal{P}_E,\delta}$ denotes the complexity of finding a solution for $\mathcal{P}_E$ (with or without p_R, depending on δ). While we do not have a general estimate for the complexity this second step, we discuss how to estimate it in the case of the 80–bit secure parameter set proposed in Section 2.4 of [18], in the following.

Security of the Suggested Parameters. Let $d = 129$, and $t = k = 6$ for the Dob encryption scheme. Using equations (3) and (21) we find that it is not possible to choose an n such that $N_4^{(0,0)}$ is positive, and $D_{reg}(129, n) > 4$. For degree 5, we find that $n = 50$ is the smallest number such that $N_5^{(0,0)}$ is positive, and $D_{reg}(129, 50) > 5$. Indeed, for this choice of parameters, we get:

$$N_5^{(0,0)}(129, 50, 6, 6) = 64024,$$

which is exactly the number of degree fall polynomials observed in the last row of table 2. For this choice of parameters, ρ is upper bounded by 15, due to lemma 6. In this case we can do even better, and use $\rho = 11$, as described in appendix E. Choosing $\delta = 1$, we find that the first step requires about $22\binom{50}{5}^{\omega}$ operations. For step two, we note from table 4 in appendix F that the extended Dob system with modifications $t = k = 6$ has a solving degree of 4 in all the experiments we can run. Conjecturing that this behaviour extends to $d = 129$, we estimate the complexity of step two to be $\mathcal{C}_{\mathcal{P}_E,1} = 2^{12}\binom{123}{4}^{\omega}$, where the factor 2^{12} is the cost of finding the correct constants for $q_1^*, \ldots, q_6^*$ and $v_1^*, \ldots, v_6^*$. We have also used $123 = 129 - 6$ as the number of variables in this system, seeing that 6 variables are eliminated by the linear forms v_i^*.

Using $\omega = 2.4$ step one requires about 2^{55} operations, and step two is estimated at 2^{67}. Using Strassen's algorithm with $\omega = 2.8$ (a rather pessimistic choice for an attacker as it assumes that it is not possible to take advantage of the sparse matrix structure of the systems), the numbers are 2^{63} for step one, and 2^{77} for step two. Either option leads to a time complexity below the proposed 80–bit security.

7 The Security of the Dobbertin Permutation for Cryptographic Use

A natural question to ask is whether it is possible to find parameters for an efficient and secure version of the Dob encryption scheme. As our attack can be split in two phases, one could either try to make either of these infeasible for an attacker. We have seen that the modifications of the Dob encryption scheme is not as effective as initially hoped in hiding the degree fall polynomials of nude Dob. Furthermore, an attacker has a lot of flexibility in fixing variables, and gluing together polynomials that reveals information about the secret modifications. Even if secure parameters could be found for degree five, there is always the question of how the number of degree fall polynomials grows for larger degrees, i.e., determining N_ν for $\nu > 5$. For these reasons it seems likely that a significant increase to t, k, and/or d is needed, which would in turn have a large negative impact on decryption time and/or public key size.

Another idea could be to make solving the extended Dob system (phase two of the attack) infeasible. We note for instance that if the suggested parameters (see section 6.5) had instead used $t = 12$ and $k = 0$, then the extended system would not have been susceptible to a straightforward hybrid attack, since the computations would likely go up to at least degree five for each guess (see table 4 and the surrounding discussion in appendix F). We do, however, stress that essentially basing the security on the $\mathcal{L}_+$ modification (definition 4) seems like a risky endeavour: an attacker is still able to learn a lot of information about the structure of the system from its degree fall polynomials. This extra information could potentially be exploited in a more sophisticated attack.

On the other hand, the analysis presented in this work may not prove much of a threat to the use of the Dob permutation in signature schemes. The authors of [18] suggested the minus modification for a Dob signature scheme. While there is reason to believe that this modification has similar characteristics to the Q_+ modification (we note that the behaviour of the Q_+ modification is somewhat reminiscent to what was analysed in [20], though the central maps differ), the key difference is that signing time does not depend exponentially on the number of polynomials removed. For instance, in [18] a version of the Dob signature scheme is suggested using $d = 257$, and removing 129 polynomials for 128–bit security. It seems unlikely that our techniques will be successful when such a large number of modifications are in place, even when degrees > 5 are taken into account.

Lastly, we note that the analysis presented here has solely been focused on the Dobbertin permutation, and hence the security of the generalisations discussed in [18], i.e., the families 'Pat', 'Mac' and 'Super Two–Face', remains an open question.

8 Conclusions

We have presented an analysis of the effectiveness the Q_+ and *ip* modifications against algebraic attacks. The theory was then applied to the Dob encryption scheme, along with a novel attack on this construction. Not only does the attack break the suggested parameter set, its flexibility and effectiveness allows us to conclude that the Dobbertin permutation seems unsuited for use in encryption schemes.

There are several directions where the ideas presented here may inspire future work. Firstly, the modifications are treated as ideals, whose dimensions can be examined. If different types of modifications, such as minus and vinegar, can be included in this framework, it could lead to a deeper understanding of the security of an even larger subclass of big–field schemes. Secondly, the attack introduces new tools for the cryptanalysis of multivariate schemes. The gluing technique allows an attacker to collect useful information after fixing a number of variables. As there is no need for correct guesses, the exponential factor usually associated with hybrid methods is avoided. Furthermore, the technique does not rely on heuristic assumptions on the relation between the first fall and solving degrees.

In light of this, we believe that security analyses of big–field multivariate schemes ought not only focus on the first fall degree directly, but also how this degree changes when fixing variables. Cryptographers wishing to design encryption schemes by adding limited modification to an otherwise weak polynomial system should be particularly aware of the effect presented in this work.

Acknowledgements

Morten Øygarden has been funded by The Research Council of Norway through the project "qsIoT: Quantum safe cryptography for the Internet of Things". The authors would like to thank Carlos Cid for useful discussions on this work.

References

1. D. Apon, D. Moody, R. Perlner, D. Smith-Tone, and J. Verbel. Combinatorial rank attacks against the rectangular simple matrix encryption scheme. In *International Conference on Post-Quantum Cryptography*, pages 307–322. Springer, 2020.
2. M. Bardet, J.-C. Faugère, and B. Salvy. Complexity of Gröbner basis computation for Semi-regular Overdetermined sequences over $\mathbb{F}_2$ with solutions in $\mathbb{F}_2$. 2003. [Research Report] RR-5049, INRIA, inria-00071534.
3. L. Bettale, J.-C. Faugère, and L. Perret. Hybrid approach for solving multivariate systems over finite fields. *Journal of Mathematical Cryptology*, 3(3):177–197, 2009.
4. C. Carlet. Vectorial boolean functions for cryptography. In Y. Crama and P. L. Hammer, editors, *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, pages 398–469. Cambridge University Press, 2010.
5. R. Cartor and D. Smith-Tone. EFLASH: A New Multivariate Encryption Scheme. In C. Cid and M. Jacobson Jr., editors, *Selected Areas in Cryptography – SAC 2018*, volume 11349 of *Lecture Notes in Computer Science*, pages 281–299. Springer International Publishing, 2019.
6. D. A. Cox, J. Little, and D. O'shea. *Using algebraic geometry*, volume 185. Springer Science & Business Media, 2006.

7. J. Ding. A new variant of the Matsumoto-Imai cryptosystem through perturbation. In *International Workshop on Public Key Cryptography*, pages 305–318. Springer, 2004.
8. J. Ding and J. E. Gower. Inoculating multivariate schemes against differential attacks. In *International Workshop on Public Key Cryptography*, pages 290–301. Springer, 2006.
9. J. Ding and T. J. Hodges. Inverting HFE systems is quasi-polynomial for all fields. In *Annual Cryptology Conference*, pages 724–742. Springer, 2011.
10. J. Ding, R. Perlner, A. Petzoldt, and D. Smith-Tone. Improved cryptanalysis of HFEv- via projection. In *International Conference on Post-Quantum Cryptography*, pages 375–395. Springer, 2018.
11. J. Ding and D. Schmidt. Cryptanalysis of HFEv and internal perturbation of HFE. In *International Workshop on Public Key Cryptography*, pages 288–301. Springer, 2005.
12. H. Dobbertin. Almost perfect nonlinear power functions on gf (2/sup n/): the welch case. *IEEE Transactions on Information Theory*, 45(4):1271–1275, 1999.
13. V. Dubois, L. Granboulan, and J. Stern. Cryptanalysis of HFE with internal perturbation. In *International Workshop on Public Key Cryptography*, pages 249–265. Springer, 2007.
14. J. C. Faugère. A new efficient algorithm for computing Gröbner bases (F4). *Journal of pure and applied algebra*, 139(1-3):61–88, 1999.
15. J.-C. Faugère and A. Joux. Algebraic cryptanalysis of hidden field equation (HFE) cryptosystems using Gröbner bases. In *Annual International Cryptology Conference*, pages 44–60. Springer, 2003.
16. P.-A. Fouque, L. Granboulan, and J. Stern. Differential cryptanalysis for multivariate schemes. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 341–353. Springer, 2005.
17. `https://github.com/Simula-UiB/Attack-On-The-Dob-Encryption-Scheme`.
18. G. Macario-Rat and J. Patarin. Two-face: New public key multivariate schemes. In *International Conference on Cryptology in Africa*, pages 252–265. Springer, 2018.
19. T. Matsumoto and H. Imai. Public quadratic polynomial-tuples for efficient signature-verification and message-encryption. In D. Barstow, W. Brauer, P. Brinch Hansen, D. Gries, D. Luckham, C. Moler, A. Pnueli, G. Seegmüller, J. Stoer, N. Wirth, and C. G. Günther, editors, *Advances in Cryptology — EUROCRYPT '88*, pages 419–453, Berlin, Heidelberg, 1988. Springer Berlin Heidelberg.
20. M. Øygarden, P. Felke, H. Raddum, and C. Cid. Cryptanalysis of the multivariate encryption scheme EFLASH. In *Cryptographers' Track at the RSA Conference*, pages 85–105. Springer, 2020.
21. J. Patarin. Cryptanalysis of the Matsumoto and Imai public key scheme of Eurocrypt'88. In *Annual International Cryptology Conference*, pages 248–261. Springer, 1995.
22. J. Patarin. Hidden fields equations (HFE) and isomorphisms of polynomials (IP): Two new families of asymmetric algorithms. In *International Conference on the Theory and Applications of Cryptographic Techniques*, pages 33–48. Springer, 1996.
23. P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. Ieee, 1994.
24. D. Smith-Tone and J. Verbel. A rank attack against extension field cancellation. In *International Conference on Post-Quantum Cryptography*, pages 381–401. Springer, 2020.

25. A. Szepieniec, J. Ding, and B. Preneel. Extension field cancellation: A new central trapdoor for multivariate quadratic systems. In *Post-Quantum Cryptography*, pages 182–196. Springer, 2016.
26. C. Tao, H. Xiang, A. Petzoldt, and J. Ding. Simple matrix–a multivariate public key cryptosystem (MPKC) for encryption. *Finite Fields and Their Applications*, 35:352–368, 2015.
27. Y. Wang, Y. Ikematsu, D. H. Duong, and T. Takagi. The secure parameters and efficient decryption algorithm for multivariate public key cryptosystem EFC. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 102(9):1028–1036, 2019.
28. C. Wolf and B. Preneel. Taxonomy of public key schemes based on the problem of multivariate quadratic equations. Cryptology ePrint Archive, Report 2005/077, 2005. `https://eprint.iacr.org/2005/077`.
29. T. Yasuda, Y. Wang, and T. Takagi. Multivariate encryption schemes based on polynomial equations over real numbers. In *International Conference on Post-Quantum Cryptography*, pages 402–421. Springer, 2020.

A Trivial Syzygies under $\psi^{\mathcal{P}^h}$

The image $\psi^{\mathcal{P}^h}(\mathcal{T}(\mathcal{F}^h))$, where $\mathcal{T}(\mathcal{F}^h)$ denotes the trivial syzygies, warrants some extra attention. Write $p_i^h = f_i^h + \sum_j a_{i,j} m_j$, where m_j denote the modifiers q_i^h and $(v_i v_l)^h$. Then the image of a Koszul syzygy is

$$\psi^{\mathcal{P}^h}((0,\dots,0,f_{i_0}^h,0\dots,0,f_{j_0}^h,0\dots,0) = f_{i_0}^h\Bigg(\sum_j a_{j_0,j} m_j\Bigg) + f_{j_0}^h\Bigg(\sum_j a_{i_0,j} m_j\Bigg).$$

Note that the same polynomial can be written as

$$\Bigg(\sum_j a_{j_0,j} m_j\Bigg) p_{i_0}^h + \Bigg(\sum_j a_{i_0,j} m_j\Bigg) p_{j_0}^h = f_{i_0}^h\Bigg(\sum_j a_{j_0,j} m_j\Bigg) + f_{j_0}^h\Bigg(\sum_j a_{i_0,j} m_j\Bigg).$$

A similar observation can be done for the field syzygies, which ensures that $\langle\psi^{\mathcal{P}^h}(\mathcal{T}(\mathcal{F}^h))\rangle \subseteq M^{(2,1)}\langle\mathcal{P}^h\rangle$.

B Deriving Formulas for Degree Fall Polynomials

$\mathbf{N_5^{(1,1)}}$: Let us start by examining $(\mathcal{S}(\mathcal{F})_{M^{(1,1)}})_5$. The polynomials involving the quadratic polynomials from Q_+, q_i^h, are easy to classify, as they would only appear as products with the $2d$ degree fall polynomials at $\nu = 3$ (from eq. (16)). The elements containing the *ip* linear forms are slightly more involved. At first glance, the $\nu = 3$ syzygies will generate $2d \cdot \dim_2(V^1)$, but we also need to take into consideration the cancellations appearing at $\nu = 4$ (which sums up to the $-d$ term in eq. (18)). Assuming that none of these cancellations can be factorized by a linear form in $\mathrm{Span}(v_1,\dots,v_k)$ (which is highly likely when $n >> k$), we will need to subtract by $-kd$ to account for these cancellations.

Turning our attention to the modifiers, we can combine (v) and (ii) from Lemma 2, to get

$$\dim_5(M^{(2,1)}M^{(1,1)}) = \dim_5(M^{(3,2)}) + \dim_5(V^1Q^1) - \dim_5(M^{(3,2)} \cap V^1Q^1).$$

Expecting that $(Q^2 \cap V^3)_5$ is empty, and using Lemma 2 (iv), we can further rewrite this as

$$\begin{aligned}\dim_5(M^{(2,1)}M^{(1,1)}) = \dim_5(Q^2) &+ \dim_5(V^3) + \dim_5(V^1Q^1)\\ &- \dim_5(Q^2 \cap V^1Q^1) - \dim_5(V^3 \cap V^1Q^1).\end{aligned}$$

Example 1 (c) covers $\dim_5(V^1Q^1)$, and we will deal with the intersections through ad hoc arguments. We expect $\langle Q^2 \cap V^1Q^1\rangle_5$ to be generated by the the possible combinations $q_iq_jv_l$, so we estimate its dimension to be $k\binom{t}{2}$. Similarly, $\langle V^3 \cap V^1Q^1\rangle_5$ is expected to be generated by the combinations $v_iv_jv_rq_l$, and its dimension will be counted by $t\binom{k}{3}$.

Lastly, we examine $\mathcal{P}_{M^{(1,1)}M^{(2,1)}}$. At degree 5 the only possible combinations are $v_iv_jv_rp_l$, and $v_iq_jp_l$, and we need not have to worry with intersections, as we did for $\mathcal{P}_{M^{(2,1)}}$. All this information sums up to the following:

$$\begin{aligned}\big(N_5^{(1,1)}\big)' = {}& \overbrace{d\bigg(2k(n-k)+2\binom{k}{2}+2t-k\bigg)}^{\dim_5(\mathcal{S}(\mathcal{F})_{M^{(1,1)}})} - \overbrace{\binom{t}{2}n}^{\dim_5(Q^2)}\\ & - \overbrace{\bigg(\binom{k}{3}\binom{n-k}{2}+\binom{k}{4}(n-k)+\binom{k}{5}\bigg)}^{\dim_5(V^3)}\\ & - \overbrace{t\bigg(k\binom{n-k}{2}+\binom{k}{2}(n-k)+\binom{k}{3}\bigg)+k\bigg(t^2-\binom{t}{2}\bigg)}^{\dim_5(Q^1V^1)}\\ & + \overbrace{\binom{t}{2}k}^{\dim_5(Q^2\cap V^1Q^1)} + \overbrace{\binom{k}{3}t}^{\dim_5(V^3\cap V^1Q^1)} + \overbrace{d\bigg(kt+\binom{k}{3}\bigg)}^{\dim_5(\mathcal{P}_{M^{(1,1)}M^{(2,1)}})}\quad .\end{aligned}\tag{32}$$

Remark 3. We have run tests for $\dim_5(\mathcal{S}(\mathcal{F})_{M^{(1,1)}})$, $\dim_5(M^{(2,1)}M^{(1,1)})$ and $\dim_5(\mathcal{P}_{M^{(1,1)}M^{(2,1)}})$, and separately they agree with what we have counted above. However, when running tests for $(N_5^{(1,1)})'$ as a whole, we find that the theoretical formula presented in eq. (32) consistently undershoots the number of degree fall polynomials by $4d$. For this reason, we adjust eq. (24) in the main part of the text by this value, i.e., $N_5^{(1,1)} = \big(N_5^{(1,1)}\big)' + 4d$.

$\mathbf{N_5^{(2,1)}}$: The degree five part of $\mathcal{S}(\mathcal{F})_{M^{(2,1)}}$ will be in the span of the degree fall polynomials at degree 3 (from G_1 and G_2 eq. (16)), multiplied with the modifiers q_i and v_jv_l. An application of Lemma 2 (iv) and (v) leads to

$$\dim_5(M^{(2,1)}M^{(2,1)}) = \dim_5(V^4) + \dim_5(Q^2) + \dim_5(V^2Q^1)$$

Example 1 (b) is used to compute $\dim_5(V^2Q^1)$, and we furthermore expect no polynomials of degree five in $\mathcal{P}_{M^{(2,1)}M^{(2,1)}}$. All this sums up to the following estimate:

$$\begin{aligned}\left(N_5^{(2,1)}\right)' = {}& \overbrace{2d\left(\binom{k}{2} + t\right)}^{\dim_5(\mathcal{S}(\mathcal{F})_{M^{(2,1)}})} - \overbrace{\left(\binom{k}{4}(n-k) + \binom{k}{5}\right)}^{\dim_5(V^4)} \\ & - \overbrace{t\left(\binom{k}{2}(n-k) + \binom{k}{3}\right)}^{\dim_5(Q^1V^2)} - \overbrace{\binom{t}{2}n}^{\dim_5(Q^2)} .\end{aligned} \tag{33}$$

Similarly to what was discussed in remark 3, we also find that the theoretically predicted $\left(N_5^{(2,1)}\right)'$ is off by $4d$ in experiments. Hence, we adjust for this in eq. (25) by setting $N_5^{(2,1)} = \left(N_5^{(2,1)}\right)' + 4d$.

C Experimental Examples

In order to test our attack strategy, we implemented and verified the following two toy examples in Magma. We checked that we do indeed find $t + \binom{k}{2}$ polynomials that are in $\mathrm{Span}(p_1^h, \ldots, p_d^h, p_R, q_1^h, \ldots, q_t^h, \ldots, (v_iv_j)^h, \ldots)$, but not in $\mathrm{Span}(p_1^h, \ldots, p_d^h, p_R)$. For the latter example we also verified that we retrieve k linear forms in $\mathrm{Span}(v_1, \ldots, v_k)$. The implementation is available at [17].

Example 3 *The first toy example is that of a Dob encryption scheme where $d = 45$, $t = 6$ and $k = 0$. Fixing no variables, $n = d$, we find that the equations eqs. (20) to (22) are negative, and hence we do not expect this system to have any degree fall polynomials at degrees ≤ 4. If we instead fix 15 variables, $n = 30$, we get $N_4 = N_4^{(0,0)} = 336$. If we, in addition, add a randomly chosen homogeneous quadratic polynomial p_R to the system, we get 342 degree fall polynomials at degree 4 (see eq. (26)).*

Following section 6.2, we split the variables into three disjoint sets: $W_1 = \{x_1, \ldots, x_{15}\}$, $W_2 = \{x_{16}, \ldots, x_{30}\}$ and $W_3 = \{x_{31}, \ldots, x_{45}\}$. Let $\mathcal{P}$ denote the public polynomials of the scheme, and for $i = 1, 2, 3$, compute the kernel of $\overline{M}_4(\pi_{W_i}(\{\mathcal{P}, p_R\}))$. Let $\mathcal{H}_i$ be the system of polynomials that gets multiplied with p_R in creating these kernels, and find a basis for it (which will be of dimension $d + t + 1 = 52$). The polynomial sets $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_3$ are now glued together as detailed in section 6.2. Note in particular that we do not expect any problems with the gluing, seeing that $|W_3 * W_3| = \binom{15}{2} = 105 > 52$.

Example 4 *The second toy example had parameters $d = 63$, $t = 1$ and $k = 4$. Fixing no variables, we find $N_4 = N_4^{(1,0)} = 25$. If we fix 21 variables, we find that $N_4^{(0,0)}$ is dominant, i.e., $N_4 = N_4^{(0,0)} = 445$. Adding a random quadratic polynomial yields 452 degree fall polynomials at degree 4 (see eq. (26)).*

As in the example above, we divide into three equal sets: $W_1 = \{x_1, \ldots, x_{21}\}$, $W_2 = \{x_{22}, \ldots, x_{42}\}$ and $W_3 = \{x_{43}, \ldots, x_{63}\}$, and followed the steps described in Sections 6.2 and 6.3.

D Nude Dob is Fully Broken at Degree 3

In [18] it is stated that experiments indicate that nude Dob has a solving degree 3. We will show that this is indeed the case. In the following, all computations are over either $B(d)$ or $\mathbb{F}_{2^d}[X]/\langle X^{2^d}+X\rangle$. Consider $F(X)+C=0$, where F is as defined in eq. (15), and $C \in \mathbb{F}_{2^d}$ a ciphertext we wish to solve for. Tedious hand calculation shows that

$$\left.\begin{array}{l} C^2\left((1+X^2)(XF)\right)^{2^m} + (C^2+X^4)F^{2^{m+1}} + X^{2^{m+1}}(C^2F^2+F^4) \\ +C^{2^m+2}F^{2^m} + C^{2^m+2}XF + \left((X^4+X^2)(XF)+(X^2+X)CF\right)^2 \end{array}\right\}(I)$$
$$= \left.\begin{array}{l} X^{16} + (C^{2^{m+1}} + C^{2^m+2} + C^4 + C^2)X^4 \\ +(C^2+1)X^8 + (C^{2^m+2}+C^4)X^2 + C^{2^m+3}X. \end{array}\right\}(II) \tag{34}$$

The polynomial (II) is linearized and of degree 16. Thus its zeros form subspace of dimension at most 4. It follows that (II) will correspond to a linear system $l_1(x_1,\ldots,x_d) = \ldots = l_d(x_1,\ldots,x_d) = 0$ of rank at least $d-4$, from which a plaintext from an intercepted ciphertext can be easily recovered.

It remains to show that polynomial (I) can be computed from the public key, using polynomials of degree at most 3. Recall from section 4.2 that XF correspond to degree fall polynomials down to degree two. Each such polynomial will correspond to a solution $a_{i,j}, \gamma_{i,j}, \beta_i, \delta \in \mathbb{F}_2$, for the equation

$$\begin{gathered}(a_{1,0}+a_{1,1}x_1+\ldots+a_{1,d}x_d)p_1+\ldots+(a_{d,0}+a_{d,1}x_1+\ldots+a_{d,d}x_d)p_d+ \\ \sum \gamma_{i,j}x_ix_j + \sum \beta_i x_i + \delta = 0.\end{gathered}$$

As described in section 4.2, we expect this solution space to be of dimension $2d$. Let $d_1,\ldots,d_{2d}$ be a basis of the degree fall polynomials derived in this step, i.e., a basis of the partial polynomials $\sum \gamma_{i,j}x_ix_j + \sum \beta_i x_i + \delta$ from this solution space. Since the only terms in (I) of 2–weight four are generated from (XF) and can be substituted by the above degree fall polynomials, we may find solutions $a'_{i,j}, \beta'_i, \delta' \in \mathbb{F}_2$ for the following system.

$$\begin{gathered}(a'_{1,0}+a'_{1,1}x_1+\ldots+a'_{1,d}x_d)p_1+\ldots+(a'_{d,0}+a'_{d,1}x_1+\ldots+a'_{d,d}x_d)p_d+ \\ (a'_{d+1,1}x_1+\cdots+a'_{d+1,d}x_d)d_1+\cdots+(a'_{3d,0}+a'_{3d,1}x_1+\cdots+a'_{3d,d}x_d)d_{2d}+ \\ \beta'_1x_1+\cdots+\beta'_dx_d+\delta'=0.\end{gathered}$$

In particular, the linear forms from (II) can be written $l_j = \sum \beta'_i x_i + \delta'$, where the β' and δ'–coefficient will be associated with solutions of this system.

Since all the systems described above only includes polynomials of degree at most three, finding a plaintext remains practical, even for $d = 129$. In practice one can also apply algorithms that can exploit degree fall polynomials, such as F_4. If this is the case, the polynomials associated with XF will be found in the first step of degree three, and the linear polynomials (II) will be found in the ensuing step of degree three.

E Proof of Lemma 6

By a slight abuse of notation we will consider $\widetilde{W_\eta}$ to include integers, by listing the index of the variables it contains. Recall the (r, d) covering problem, which can be stated as follows: for given d and $r < d-1$, find ρ subsets $\widetilde{W_\eta} \subset \{1, \ldots, d\}$ of size $d - r$, such that for any pair (i, j) where $1 \leq i < j \leq d$, $\{i, j\} \subset \widetilde{W_\eta}$ for at least one η.

Proof (of Lemma 6). Let $s = \lfloor (d-r)/2 \rfloor$. We divide $\{1, \ldots, d\}$ into blocks of size s:

$$C_b = \{(b-1)s + 1, \ldots, bs\}, \text{ for } 1 \leq b \leq \lfloor d/s \rfloor$$

.

Let the sets $\widetilde{W_\eta}$ for $1 \leq \eta \leq \binom{\lfloor d/s \rfloor}{2}$ be defined as the union of C_a and C_b, for all choices of $1 \leq a < b \leq \lfloor d/s \rfloor$. In the case $d - r$ is odd, we also add one arbitrary extra number to each set to make sure that each $\widetilde{W_\eta}$ contains exactly $d - r$ numbers.

Any $\{i, j\} \subset \{1, \ldots, s\lfloor d/s \rfloor\}$ will then be contained in at least one $\widetilde{W_\eta}$. If both i and j belong to the same block C_b, then all $\widetilde{W_\eta}$ involving C_b will contain $\{i, j\}$. If $i \in C_a$ and $j \in C_b$ for $a \neq b$, then the set $\widetilde{W_\eta} = C_a \cup C_b$ will contain $\{i, j\}$. Hence the $\binom{\lfloor d/s \rfloor}{2}$ sets constructed will cover all pairs from $\{1, \ldots, s\lfloor d/s \rfloor\}$.

If s divides d we are done. Otherwise, to cover all pairs of numbers in $\{1, \ldots, d\}$ it is sufficient to create $\lfloor d/s \rfloor$ new $\widetilde{W}$-sets consisting of
$\{s\lfloor d/s \rfloor + 1, \ldots, d\} \cup C_b \cup \{s - (d - s\lfloor d/s \rfloor) \text{ extra numbers}\}$,
where $1 \leq b \leq \lfloor d/s \rfloor$, and the extra numbers are arbitrary. The total number of sets will then be $\binom{\lceil d/s \rceil}{2}$, and replacing s with $\lfloor (d-r)/2 \rfloor$ we get Lemma 6.

For the particular case $d = 129, r = 79$ (which is used in Section 6.5) we get $\rho \leq 15$. Doing the exercise in practice we find that $\rho = 11$ is sufficient to solve the problem by extending the block C_5 to cover all numbers $101, \ldots, 129$, and modifying slightly the sets involving C_5.

F Experiments with Extended Dob Systems

In table 4 we have run some experiments on the extended Dob System, without the random polynomial q_R (see remark 2). We have chosen to fix $k = 6$, and vary $t = 3, 6, 10$. As noted in section 6.4, all the systems has $2d$ degree fall polynomials at degree 3. Furthermore, additional degree fall polynomials will be found at the first step of degree 4. This can be observed under “Step Degrees”, where the initial degrees are $2, 3, 3, 4, 4(\ldots)$. The exceptions are when $t = 3$, where a solution is found already at degree 3. Despite the low first fall degree, the solving degree seems to grow with t.

Table 4: Step and Solving Degrees for Extended Dob Systems

d	t (Q_+)	k (ip)	D_{solv}	Step Degrees
55	3	6	3	2,3,3,3,3
67	3	6	3	2,3,3,3,3
55	6	6	4	2,3,3,4,4,3,3
61	6	6	4	2,3,3,4,4,3,4
65	6	6	4	2,3,3,4,4,4
67	6	6	4	2,3,3,4,4,4
55	10	6	5	2,3,3,4,4,5
67	10	6	≥ 5	2,3,3,4,4,5...

G Experiments with the Dob Encryption System

There is substantial freedom in the choice of parameters, d, n, t, k, that can be associated to a Dob encryption system (with some fixed variables). In light of this, we find that elaborate experimentation is necessary in order to gain confidence in the degree fall estimates presented in Equations (20) – (25) of section 4.3. We hope to make a stride towards such confidence by presenting various experiments in Tables 5 – 8 (in addition to what was presented in Table 2 of section 5). The setup of the tables is as described in section 5.3. An entry where none of our formulas predict a positive value is marked with a "-" under "N (predicted)". We also do need register the experimentally found number of degree fall polynomials, "N (Magma)", if the registered first fall degree is the same as $D_{reg}(d, n)$. Experiments that ran out of memory has been marked with an inequality in D_{solv}, and "..." under "Step Degrees". The last number in "Step Degrees" marks the step that ran out of memory. It is worth noting that in all the experiments we have run, the number of degree fall polynomials we predict, "N (predicted)", matches exactly the number of registered first fall polynomials. Furthermore, in the cases where "N (predicted)" is marked with "-", the experimental first fall degree is indeed ≥ 6.

Table 5: Dob encryption scheme for various parameters, $D_{ff} = 3, 4$.

d	n	t	k	D_{ff}	N	N	D_{solv}	Step
		(Q_+)	(ip)		(predicted)	(Magma)	$(D_{reg}(d,n))$	Degrees
53	53	0	5	4	$N_4^{(1,0)}$: 45	3:45	5 (9)	2,3,4,4,5,4
49	49	0	0	3	$N_3^{(0,0)}$: 98	2:98	3 (9)	2,3,3
29	29	3	0	4	$N_4^{(0,0)}$: 528	3:528	4 (6)	2,3,4,4
29	29	4	0	4	$N_4^{(0,0)}$: 155	3:155	4 (6)	2,3,4,4,4
33	32	4	0	4	$N_4^{(0,0)}$: 237	3:237	4 (6)	2,3,4,4,4
29	29	0	2	3	$N_3^{(0,0)}$: 31	2:31	3 (6)	2,3,3,3,3,3
29	29	0	3	4	$N_4^{(0,0)}$: 739	3:739	4 (6)	2,3,4,4
31	31	0	3	4	$N_4^{(0,0)}$: 822	3:822	4 (6)	2,3,4,4
31	30	0	3	4	$N_4^{(0,0)}$: 842	3:842	4 (6)	2,3,4,4
31	31	0	4	4	$N_4^{(1,0)}$: 139	3:139	4 (6)	2,3,4,4,4
29	29	0	4	4	$N_4^{(1,0)}$: 131	3:131	4 (6)	2,3,4,4,4
33	33	0	4	4	$N_4^{(1,0)}$: 147	3:147	4 (7)	2,3,4,4,4
35	35	0	4	4	$N_4^{(1,0)}$: 155	3:155	4 (7)	2,3,4,4,4
29	25	0	4	4	$N_4^{(0,0)}$: 250	3:250	4 (5)	2,3,4,4,4
31	31	0	5	4	$N_4^{(1,0)}$: 45	3:45	$\geq$ 5 (6)	2,3,4,4,5...
29	29	1	4	4	$N_4^{(1,0)}$: 25	3:25	5 (6)	2,3,4,4,5,4
35	26	0	6	4	$N_4^{(1,0)}$: 5	3:5	5 (5)	2,3,4,4,5
59	29	0	7	4	$N_4^{(1,0)}$: 21	3:21	5 (5)	2,3,4,4,5
31	29	2	2	4	$N_4^{(0,0)}$: 702	3:702	4 (6)	2,3,4,4,3
37	24	0	5	4	$N_4^{(0,0)}$: 204	3:204	4 (5)	2,3,4,4
37	25	0	5	4	$N_4^{(1,0)}$: 165	3:165	4 (5)	2,3,4,4,3
37	24	1	4	4	$N_4^{(0,0)}$: 508	3:508	4 (5)	2,3,4,4
37	25	1	4	4	$N_4^{(0,0)}$: 434	3:434	4 (5)	2,3,4,4
79	33	0	6	4	$N_4^{(0,0)}$: 500	3:500	4 (5)	2,3,4,4
83	34	0	6	4	$N_4^{(0,0)}$: 561	3:561	4 (5)	2,3,4,4

Table 6: Dob encryption scheme for various parameters, $D_{ff} \geq 5$.

d	n	t	k	D_{ff}	N	N	D_{solv}	Step
		(Q_+)	(ip)		(predicted)	(Magma)	$(D_{reg}(d,n))$	Degrees
53	53	0	8	≥ 6	-	-	≥ 6 (9)	2,3,4,5,6...
37	37	5	0	5	$N_5^{(0,0)}$: 12617	3:397, 4:12220	5 (7)	2,3,4,5,4,3
35	30	0	8	5	$N_5^{(1,1)}$: 1568	4:1568	5 (6)	2,3,4,5,5
35	34	0	8	5	$N_5^{(3)}$: 224	4:224	6 (7)	2,3,4,5,5,6
41	31	0	9	5	$N_5^{(3)}$: 218	4:218	5 (6)	2,3,4,5,5,5
35	35	1	6	5	$N_5^{(2)}$: 2714	4:2714	5 (7)	2,3,4,5,5,5
31	29	2	4	5	$N_5^{(1)}$: 5869	4:5869	5 (6)	2,3,4,5,5
31	31	0	6	5	$N_5^{(2)}$: 4407	4:4407	5 (6)	2,3,4,5,5
33	32	0	6	5	$N_5^{(2)}$: 4984	4:4984	5 (6)	2,3,4,5,5
33	32	0	7	5	$N_5^{(2)}$: 2596	4:2596	5 (6)	2,3,4,5,5
33	32	0	8	5	$N_5^{(3)}$: 244	4:244	6 (6)	2,3,4,5,5,5,6
33	32	0	9	6	-	-	6 (6)	2,3,4,5,6,6
33	31	0	8	5	$N_5^{(2,1)}$: 314	4:314	5 (6)	2,3,4,5,5,5
31	28	0	8	5	$N_5^{(1,1)}$: 1172	4:1172	5 (6)	2,3,4,5,5
33	28	0	9	6	-	-	6 (6)	2,3,4,5,6
35	28	0	6	5	$N_5^{(1,1)}$: 5964	3:49, 4:5915	5 (6)	2,3,4,5,4,5
37	35	6	0	5	$N_5^{(0,0)}$: 8048	4:8048	5 (7)	2,3,4,5,5
37	37	6	0	5	$N_5^{(0,0)}$: 6364	4:6364	5 (7)	2,3,4,5,5,3
39	37	6	0	5	$N_5^{(0,0)}$: 9030	4:9030	5 (7)	2,3,4,5,5,3
37	35	7	0	5	$N_5^{(0,0)}$: 2969	4:2969	5 (7)	2,3,4,5,5,5
37	35	8	0	6	-	4:4817 5:96104	6 (7)	2,3,4,5,6,5
39	38	6	0	5	$N_5^{(0,0)}$: 8136	4:8136	5 (7)	2,3,4,5,5,3
39	38	5	0	5	$N_5^{(0,0)}$: 14940	3:429, 4:14511	5 (7)	2,3,4,5,4,3
37	36	7	0	5	$N_5^{(0,0)}$: 1644	4:1644	5 (7)	2,3,4,5,5,5,3
39	38	2	4	5	$N_5^{(0,0)}$: 5458	4:5458	5 (7)	2,3,4,5,5
39	38	4	2	5	$N_5^{(0,0)}$: 16112	4:16112	5 (7)	2,3,4,5,5
37	36	2	4	5	$N_5^{(0,0)}$: 5578	4:5578	5 (7)	2,3,4,5,5
39	39	0	6	5	$N_5^{(1,1)}$: 6255	4:6255	5 (7)	2,3,4,5,5
37	37	0	6	5	$N_5^{(1,1)}$: 5769	4:5769	5 (7)	2,3,4,5,5
35	35	0	6	5	$N_5^{(1,1)}$: 5299	4:5299	5 (7)	2,3,4,5,5
33	33	0	6	5	$N_5^{(1,1)}$: 4845	4:4845	5 (7)	2,3,4,5,5
37	36	0	6	5	$N_5^{(1,1)}$: 5940	4:5940	5 (7)	2,3,4,5,5
39	35	0	6	5	$N_5^{(1,1)}$: 6883	4:6883	5 (7)	2,3,4,5,5
37	35	0	6	5	$N_5^{(1,1)}$: 6091	4:6091	5 (7)	2,3,4,5,5
37	34	0	6	5	$N_5^{(1,1)}$: 6222	4:6222	5 (7)	2,3,4,5,5
35	34	0	6	5	$N_5^{(1,1)}$: 5454	4:5454	5 (7)	2,3,4,5,5
35	33	0	6	5	$N_5^{(1,1)}$: 5589	4:5589	5 (7)	2,3,4,5,5
33	31	0	6	5	$N_5^{(1,1)}$: 5103	4:5103	5 (6)	2,3,4,5,5

Table 7: Dob encryption scheme for various parameters, $D_{ff} \geq 5$.

d	n	t	k	D_{ff}	N	N	D_{solv}	Step
		(Q_+)	(ip)		(predicted)	(Magma)	$(D_{reg}(d,n))$	Degrees
37	37	1	6	5	$N_5^{(1,1)}$: 2816	4:2816	5 (7)	2,3,4,5,5,5
39	38	1	6	5	$N_5^{(1,1)}$: 3304	4:3304	5 (7)	2,3,4,5,5,5
39	39	1	6	5	$N_5^{(1,1)}$: 2910	4:2910	5 (7)	2,3,4,5,5,5
37	36	1	6	5	$N_5^{(1,1)}$: 3182	4:3182	5 (7)	2,3,4,5,5,5
43	37	1	6	5	$N_5^{(1,1)}$: 5384	4:5384	5 (7)	2,3,4,5,5
43	39	1	6	5	$N_5^{(1,1)}$: 4718	4:4718	5 (7)	2,3,4,5,5,5
41	39	1	6	5	$N_5^{(1,1)}$: 3814	4:3814	5 (7)	2,3,4,5,5,5
41	38	1	6	5	$N_5^{(1,1)}$: 4184	4:4184	5 (7)	2,3,4,5,5,5
37	36	2	6	5	$N_5^{(1,1)}$: 400	4:400	$\geq$ 6 (7)	2,3,4,5,5,6...
37	30	2	6	5	$N_5^{(1,1)}$: 3100	4:3100	5 (6)	2,3,4,5,5
37	30	3	6	5	$N_5^{(1,1)}$: 1350	4:1350	5 (6)	2,3,4,5,5
37	31	2	6	5	$N_5^{(1,1)}$: 2730	4:2730	5 (6)	2,3,4,5,5
37	32	2	6	5	$N_5^{(1,1)}$: 2328	4:2328	5 (6)	2,3,4,5,5
41	35	2	6	5	$N_5^{(1,1)}$: 2578	4:2578	5 (6)	2,3,4,5,5,5
41	34	2	6	5	$N_5^{(1,1)}$: 3028	4:3028	5 (6)	2,3,4,5,5
41	34	3	6	5	$N_5^{(1,1)}$: 630	4:630	$\geq$ 6 (6)	2,3,4,5,5,6...
41	35	3	6	$\geq$ 6	-	-	$\geq$ 6 (6)	2,3,4,5,6...
47	33	3	6	5	$N_5^{(1,1)}$: 3603	4:3603	5 (6)	2,3,4,5,5
37	29	4	6	5	$N_5^{(1,1)}$: 231	4:231	5 (6)	2,3,4,5,5,5
43	32	5	6	6	-	-	6 (6)	2,3,4,5,6
47	32	5	6	5	$N_5^{(1,1)}$: 34	4:34	6 (6)	2,3,4,5,5,6
61	36	6	6	6	-	-	6 (6)	2,3,4,5,6
75	39	6	6	5	$N_5^{(0,0)}$: 4674	4:4674	5 (6)	2,3,4,5,5
33	31	0	7	5	$N_5^{(1,1)}$: 3009	4:3009	5 (6)	2,3,4,5,5
33	30	0	7	5	$N_5^{(1,1)}$: 3387	4:3387	5 (6)	2,3,4,5,5
37	33	0	7	5	$N_5^{(1,1)}$: 3900	4:3900	5 (6)	2,3,4,5,5
37	34	0	7	5	$N_5^{(1,1)}$: 3473	4:3473	5 (7)	2,3,4,5,5
37	35	0	7	5	$N_5^{(1,1)}$: 3011	4:3011	5 (7)	2,3,4,5,5,5
35	30	0	7	5	$N_5^{(1,1)}$: 4179	4:4179	5 (6)	2,3,4,5,5
35	33	0	7	5	$N_5^{(1,1)}$: 3024	4:3024	5 (7)	2,3,4,5,5
39	35	0	7	5	$N_5^{(1,1)}$: 3943	4:3943	5 (7)	2,3,4,5,5
39	36	0	7	5	$N_5^{(1,1)}$: 3474	4:3474	5 (7)	2,3,4,5,5,5
39	37	0	7	5	$N_5^{(1,1)}$: 2970	4:2970	5 (7)	2,3,4,5,5,5
39	35	0	8	5	$N_5^{(2,1)}$: 394	4:394	$\geq$ 6 (7)	2,3,4,5,5,5,6...
41	34	0	8	5	$N_5^{(1,1)}$: 1408	4:1408	5 (6)	2,3,4,5,5,5
49	36	0	8	5	$N_5^{(1,1)}$: 4060	4:4060	5 (6)	2,3,4,5,5
57	38	0	8	5	$N_5^{(1,1)}$: 7000	4:7000	5 (6)	2,3,4,5,5
55	37	0	8	5	$N_5^{(1,1)}$: 6638	4:6638	5 (6)	2,3,4,5,5
53	37	0	8	5	$N_5^{(1,1)}$: 5494	4:5494	5 (6)	2,3,4,5,5

Table 8: Dob encryption scheme for various parameters, $D_{ff} \geq 5$.

d	n	t	k	D_{ff}	N	N	D_{solv}	Step
		(Q_+)	(ip)		(predicted)	(Magma)	$(D_{reg}(d,n))$	Degrees
49	34	3	6	5	$N_5^{(1,1)}$: 3894	4:3894	5 (6)	2,3,4,5,5
51	35	3	6	5	$N_5^{(1,1)}$: 4195	4:4195	5 (6)	2,3,4,5,5
53	35	4	6	5	$N_5^{(1,1)}$: 2525	4:2525	5 (6)	2,3,4,5,5
51	35	4	6	5	$N_5^{(1,1)}$: 1669	4:1669	5 (6)	2,3,4,5,5,5
49	33	4	6	5	$N_5^{(1,1)}$: 2219	4:2219	5 (6)	2,3,4,5,5
57	36	4	6	5	$N_5^{(1,1)}$: 3564	4:3564	5 (6)	2,3,4,5,5
57	35	4	6	5	$N_5^{(1,1)}$: 4237	4:4237	5 (6)	2,3,4,5,5
65	37	6	6	5	$N_5^{(1,1)}$: 780	4:780	5 (6)	2,3,4,5,5,5
81	41	0	7	5	$N_5^{(0,0)}$: 25534	4:25534	5 (6)	2,3,4,5,5
81	42	0	7	5	$N_5^{(1,1)}$: 23613	4:23613	5 (6)	2,3,4,5,5
81	43	0	7	5	$N_5^{(1,1)}$: 23487	4:23487	5 (6)	2,3,4,5,5
83	41	0	7	5	$N_5^{(0,0)}$: 29450	4:29450	5 (6)	2,3,4,5,5
83	42	0	7	5	$N_5^{(0,0)}$: 24910	4:24910	5 (6)	2,3,4,5,5
83	43	0	7	5	$N_5^{(1,1)}$: 24643	4:24643	5 (6)	2,3,4,5,5
37	30	2	7	5	$N_5^{(1,1)}$: 1127	4:1127	5 (6)	2,3,4,5,5
41	31	3	7	5	$N_5^{(1,1)}$: 58	4:58	6 (6)	2,3,4,5,5,6
45	32	2	8	5	$N_5^{(1,1)}$: 88	4:88	6 (6)	2,3,4,5,5,6
89	43	3	5	5	$N_5^{(0,0)}$: 55986	2:424, 3:8514, 4:47048	5 (6)	2,3,4,5,3
93	44	2	6	5	$N_5^{(0,0)}$: 46246	4:46246	5 (6)	2,3,4,5,5