A Complete Cryptanalysis of the Post-Quantum Multivariate Signature Scheme Himq-3

Jintai Ding[1], Zheng Zhang[1](✉), Joshua Deaton[1], and Lih-Chung Wang[2]

[1] Department of Mathematical Science, University of Cincinnati, Cincinnati, USA
jintai.ding@gmail.com, {zhang2zh,deatonju}@mail.uc.edu
[2] Department of Applied Mathematics, National Dong Hwa University, Shoufeng, Taiwan
lcwang@gms.ndhu.edu.tw

Abstract. In 2017 Kyung-Ah Shim et al. proposed a multivariate signature scheme called Himq-3 which is a submission to National Institute of Standards and Technology (NIST) standardization process of post-quantum cryptosystems. The Himq-3 signature scheme can be classified into the oil vinegar signature scheme family. Similar to the rainbow signature scheme, the Himq-3 signature scheme uses a multilayer structure to shorten the signature size. Moreover the signing process is very fast due to a special system called L-inveritble cycle system that is used to invert the central map. In this paper, we provide a complete cryptanalysis to the Himq-3 signature scheme. We describe a new attack method called the singularity attack. This attack is based on the observation that the variables in the L-invertible cycle system are not allowed to be zero in a valid signature. For the completeness, we show step by step how variables and layers can be separated so that signature forgery can be performed. We claim that the complexity of our attack is much lower than the proposed security level.

Keywords: Post-quantum cryptography · Multivariate public key cryptography · Cryptanalysis · Oil vinegar signature scheme

1 Introduction

1.1 Background

The ability to authenticate digital messages has always been an important building block for any free, secure, and digital society. In 1976, Whitfield Diffie and Martin Hellman did a major contribution to construct a mathematical framework, known as digital signature scheme, in this direction. The digital signature algorithm (DSA), the RSA digital signature algorithm, and the elliptic curve digital signature algorithm were the only signature schemes that were allowed under the guidelines of the National Institute of Standards and Technology (NIST)'s

W. Meng et al. (Eds.): ICICS 2020, LNCS 12282, pp. 422–440, 2020.
https://doi.org/10.1007/978-3-030-61078-4_24

up to 2013. However, in 1999 Peter Shor showed that these number theory based signature schemes are weak to sufficiently powerful quantum computers [18]. This indicates a significant need to prepare the current communication system for a post-quantum world. Due to the rapid development of quantum computers, NIST believes that it is prudent to begin developing standards for post-quantum cryptography. The call for proposals started in December 2016. NIST expects to perform multiple rounds of evaluation over a period of three to five years.

1.2 Multivariate Public Key Cryptography

Multivariate Public Key Cryptography (MPKC) is one of the candidates that are believed to have the potential to resist quantum attacks [4]. The security of MPKC depends on the difficulty of solving a system of multivariate quadratic polynomials over a finite field. A breakthrough in MPKC was proposed by Matsumoto and Imai in 1988 [14]. Instead of working over the vector space k^n for a finite field k, they looked to a degree n extension of k in which an invertible map can be constructed. Unfortunately, this scheme was broken by Patarin using the linearization equation attack [15]. However, inspired by this attack, Patarin proposed the oil vinegar signature scheme [16]. The oil vinegar signature scheme can be classified into three groups: Balanced oil vinegar [16] (Patarin 1997), Unbalanced oil vinegar (UOV) [12] (Kipnis et al. 1999) and Rainbow [7], a multilayer signature scheme with unbalanced oil vinegar in each layer (Ding and Schmidt 2005). The balanced oil vinegar scheme was broken by Kipnis and Shamir [13] using the idea of invariant subspaces. The unbalanced oil vinegar scheme remains unbroken since its publication nearly 20 years ago. However, the main drawback of UOV is its large key size and signature size. Rainbow is considered to be one of the most promised post-quantum cryptography signature schemes. Its multilayer structure, in which oil variables from previous layer are reused as vinegar variables in next layer, reduces the key size and signature size. Detailed security analysis of rainbow signature scheme is presented in [5]. There are several other signature schemes that are closely related to rainbow such as TRMC, TTS, etc. More about those schemes and their security analysis can be found in [5]. The lifted unbalanced oil vinegar proposed by Ward et al. is another modification of UOV [2] which achieves small key size by restricting all the coefficients of public keys to be binary. In 2019, Ding et al. designed a new attack, the subfield differential attack on LUOV, which drops the complexity of solving LUOV blew the NIST security strength for non-prime extension case [9]. Both rainbow and LUOV have passed into the second round for the NIST post-quantum standardization project. There are also new secure multivariate encryption schemes [6,19].

1.3 The Himq-3 Scheme and the Singularity Attack

The Himq-3 signature scheme proposed by Kyung-Ah Shim et al. in 2017 is a round 1 candidate of NIST post quantum standardization. It can be viewed as a variant of multilayer UOV. Himq-3 attempts to be more efficient than rainbow.

A crucial component of the central map of Himq-3 is a system called L-invertible cycle system [8]. The function of this L-invertible cycle system is to make the central map invertible. Moreover it appears that this system works very efficiently. The authors claim that it is more efficient to solve the L-invertible cycle system than a system of linear equations by a Gaussian elimination [17]. However, the L-invertible cycle system also restricts the values to certain variables. The idea of our singularity attack is based on such restriction. We claim that if enough signatures can be collected, we can construct a system of linear equations of monomials in which the solutions will leak partial information about the private key.

1.4 Our Contributions

The main result of this paper is a complete attack on a NIST round 1 candidate: the Himq-3 signature scheme. This new attack method is called the singularity attack. This attack is simple and straightforward. It does not involve polynomial solving algorithms such as F4/F5 or XL algorithm. Neither do we need the rank attacks (Minrank/Highrank attacks). The most complicated algorithm in our attack is just Gaussian elimination. We will show that it is impossible for the Himq-3 signature scheme and its variant Himq-3F to fulfill the proposed security level under the singularity attack. We notice that the variables which play a very important role in inverting the central map cannot be equal to zero in honest signing process. Hence, the public key of the scheme cannot be treated as a random multivariate quadratic system. There are some structures in the public key that we can explore. We will first show that if enough signatures are obtained, we can figure out how those variables are transformed by the private key. Next, we will undo the effect of the private keys by separating the variables and extracting the layers so that the public key can be turned in to the form where forgeries can be made. We will discuss the complexity of our attack for each proposed set of parameters, and the experimental results will be provided. Moreover, we will give a toy example in the appendix to clarify the first step of our attack.

2 HIMQ-3 Signature Scheme

2.1 Preliminary

General Construction of Bipolar MPKC Signature Scheme. We first describe the general construction of a Bipolar MPKC signature scheme. Let $\mathbb{F}_q$ be a finite field of order q. The main idea for the construction of MPKC signature schemes is to construct a polynomial map $\mathcal{F} : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^m$, called the central map, defined by $\mathcal{F} = (\mathcal{F}^{(1)}, \cdots, \mathcal{F}^{(m)})$ of m equations in n variables such that it is easy to find pre-images for a given vector. To hide the ability to find pre-images and thus construct a public key from $\mathcal{F}$, one uses two invertible affine maps $\mathcal{S} : \mathbb{F}_q^m \rightarrow \mathbb{F}_q^m$, and $\mathcal{T} : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^n$. The public key is the composition

$\mathcal{P} = \mathcal{S} \circ \mathcal{F} \circ \mathcal{T}$. The private keys are the invertible affine maps $\mathcal{S}$, $\mathcal{T}$ and the central map $\mathcal{F}$ individually. The signing process for a document is as follows:

$$\mathbb{F}_q^m \xrightarrow{\mathcal{S}^{-1}} \mathbb{F}_q^m \xrightarrow{\mathcal{F}^{-1}} \mathbb{F}_q^n \xrightarrow{\mathcal{T}^{-1}} \mathbb{F}_q^n.$$

To verify the signature, one goes through the other direction by the public key $\mathcal{P}$:

$$\mathbb{F}_q^m \xleftarrow{\mathcal{P}} \mathbb{F}_q^n.$$

L-Invertible Cycle System. The Himq-3 scheme contains a system of quadratic equations called L-invertible cycle system. This system makes it possible for the Himq-3 scheme to invert its central map.

Let $\mathbb{F}_q$ be a finite field with 2^k elements and l be an odd positive integer. The L-invertible cycle product system $\mathcal{Q}$ over $\mathbb{F}_q$ is defined by:

$$\mathcal{Q} : \alpha_1 x_1 x_2 = \beta_1, \alpha_2 x_2 x_3 = \beta_2, \cdots, \alpha_l x_l x_1 = \beta_l,$$

where α_i and β_i are nonzero elements in $\mathbb{F}_q$. We can rewrite the system $\mathcal{Q}$ in the form:

$$x_1 x_2 = \gamma_1, \cdots, x_l x_1 = \gamma_l,$$

where $\gamma_i = \beta_i / \alpha_i$.

Remark 1. Given an L-invertible cycle system $\mathcal{Q}$ as above, the solution of the system can be found as follows:

Let $A = \gamma_1 \gamma_2 \cdots \gamma_l$ and $B = \gamma_2 \gamma_4 \cdots \gamma_{l-1}$. We see that $x_1 = \frac{\sqrt{A}}{B}$, and $x_i = \gamma_{i-1}/x_{i-1}$ for $i = 2, \cdots, l-1$, and $x_l = \gamma_l / x_1$.

Remark 2. We call the variables in the L-invertible cycle system the *cycle variables* and a quadratic product of cycle variables are called *cycle product*. An important observation is that in any solution, the value of the cycle variables must be nonzero.

2.2 Description of the Himq-3 Scheme

The Himq-3 signature scheme can be classified as a new variant of UOV scheme, which shares the layer structure with the rainbow signature scheme [7]. Namely, one solves oil variables in previous layer, and plug in the solutions to next layer to solve new oil variables. We will now describe the particulars of the Himq-3 central map.

Let us denote the finite field by $\mathbb{F}_q$ of order $q = 2^k$. Let v, o_1, o_2, o_3 be positive integers where o_1 and o_2 are odd, we need the conditions that $v \geq o_1 + 1$ and $o_1 \geq o_2 \geq o_3$. Further, let the number of equations $m = o_1 + o_2 + o_3$ and the number of variables $n = v + m$. The Himq-3 central map contains n variables in the following four types.

Variables	Name
$x_1, \cdots, x_v$	v variables
$x_{v+1}, \cdots, x_{v+o_1}$	o_1 variables
$x_{v+o_1+1}, \cdots, x_{v+o_1+o_2}$	o_2 variables
$x_{v+o_1+o_2+1}, \cdots, x_{v+o_1+o_2+o_3}$	o_3 variables

Define $\mathbf{x} = (x_1, \cdots, x_n)$. The central map $\mathcal{F} = (\mathcal{F}^{(1)}, \cdots, \mathcal{F}^{(m)})$ of the Himq-3 signature scheme is defined by three layers:

First Layer. The first layer contains polynomials

$$\mathcal{F}^{(i)}(\mathbf{X}) = \Phi_i(\mathbf{X}) + \delta_i x_{v+i} x_{v+i+1}$$

for $i = 1, \cdots, o_1 - 1$ and

$$\mathcal{F}^{(o_1)}(\mathbf{X}) = \Phi_{o_1}(\mathbf{X}) + \delta_{o_1} x_{v+o_1} x_{v+1}$$

in which δ_i is a nonzero constant in $\mathbb{F}_q$. The term $\Phi_i(\mathbf{X})$ is a quadratic polynomial in v variables $(x_1, \cdots, x_v)$ defined by

$$\Phi_i(\mathbf{X}) = \sum_{j=1}^{v} \alpha_{i,j} x_j x_{1+(i+j-1)(\text{mod } v)}$$

where $\alpha_{i,j}$ is a nonzero element in $\mathbb{F}_q$. Each polynomial of the first layer consists of a quadratic polynomial Φ_i only in v variable in the front and a cycle product in o_1 variables in the end. To invert the first layer, one randomly assigns values to v variables, which in turn makes the first layer into a L-invertible cycle system in o_1 variables. If the constant terms are nonzero, the system can be easily solved by Remark 1. Otherwise, randomly assign values to v variables again and repeat the process.

Second Layer. The polynomials

$$\mathcal{F}^{(o_1+i)}(\mathbf{X}) = \Psi_i(\mathbf{X}) + \delta_{o_1+i} x_{v+o_1+i} x_{v+o_1+i+1}$$

for $i = 1, \cdots, o_2 - 1$, and

$$\mathcal{F}^{(o_1+o_2)}(\mathbf{X}) = \Psi_{o_2}(\mathbf{X}) + \delta_{o_1+o_2} x_{v+o_1+o_2} x_{v+o_1+1}$$

form the second layer in which δ_i is a nonzero constant in $\mathbb{F}_q$. The term $\Psi_i(\mathbf{X})$ is a quadratic polynomial in v and o_1 variables $(x_1, \cdots, x_{v+o_1})$ defined by

$$\Psi_i(\mathbf{X}) = \sum_{j=1}^{v} \alpha'_{i,j} x_j x_{v+(i+j-1)(\text{mod } o_1)}$$

where $\alpha'_{i,j}$ is a nonzero element in $\mathbb{F}_q$. Similar to the first layer, each polynomial of the second layer is formed by a quadratic polynomial Ψ_i in v and o_1 variables in the front and a cycle product in o_2 variables in the end. To invert the second layer, one plugs the values assigned to v variables and the solutions to o_1 variables from previous layer into Ψ_i, then the second layer becomes a L-invertible cycle system in which o_2 variables can be solved provided that the constant terms are nonzero.

Third Layer. The third layer is composed of the polynomials

$$\mathcal{F}^{(o_1+o_2+i)}(\mathbf{X}) = \sum_{v+1 \le l \le j \le v+o_1} \beta_{l,j}^{(i)} x_l x_j + \Theta_i(\mathbf{X}) + \Theta'_i(\mathbf{X}) + \epsilon_i x_{o_1+o_2+i}$$

for $i = 1, \cdots, o_3$, in which $\beta_{l,j}^{(i)}$ and ϵ_i are elements in $\mathbb{F}_q$. The polynomials Θ_i and Θ'_i are quadratic polynomials in variables $(x_1, \cdots, x_n)$ defined by

$$\Theta_i(\mathbf{X}) = \sum_{j=1}^{v+o_1} \gamma_{i,j} x_j x_{v+o_1+(i+j-1)(\mathrm{mod}\ o_2)},$$

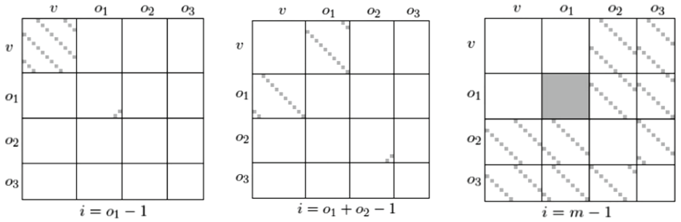
and

$$\Theta'_i(\mathbf{X}) = \sum_{j=1}^{v+o_1+o_2} \gamma'_{i,j} x_j x_{v+o_1+o_2+(i+j-1)(\mathrm{mod}\ o_3)}$$

where $\gamma_{i,j}$ and $\gamma'_{i,j}$ are nonzero elements in $\mathbb{F}_q$. We notice also that the o_3 variables are never multiplied together by themselves like oil variables in a UOV scheme. In addition they only appear in the polynomials of third layer, which makes the scheme under the threat of the highrank attack [3]. The third layer can be turned into a linear system in o_3 variables only once the random values assigned to v variables and solutions to o_1 and o_2 variables from the first and second layers respectively are plugged in. Hence, o_3 variables can be simply solved by a Gaussian elimination.

Remark 3. The design rationale of the individual $\Phi_i, \Psi_i, \Theta_i, \Theta'_i$ is to increase the rank of the symmetric matrices associate to the polynomials so that they achieve maximum amount of rank for the variables they involve. The purpose of such design is to prevent the scheme from the minrank attack [11].

We borrow from the authors of Himq-3 the graphs of symmetric matrices associated to the quadratic part of central map polynomials [17].



$i = o_1 - 1$ $i = o_1 + o_2 - 1$ $i = m - 1$

2.3 The Proposed Parameters

The authors of Himq-3 proposed the following sets of parameters for three different levels of security.

Security level	$\|\mathbb{F}_q\|$	v	o_1	o_2	o_3
128-bit	2^8	36	15	15	15
192-bit	2^8	56	25	25	25
256-bit	2^8	84	33	33	32

The Himq-3 signature scheme is claimed to be secure against all known attacks for these three levels of security according to the security analysis provided in [17]. We will show that the Himq-3 signature scheme meets none of these three security levels against our singularity attack. The complexities of our attack on Himq-3 with the last two sets of parameters are even very far away from the target level of security.

2.4 Compared to Rainbow Signature Scheme

A significant difference between rainbow and Himq-3 is the way to invert the central map. Rainbow uses the unbalanced oil vinegar structure, to be more specific, in each layer one solves new oil variables by Gaussian elimination given the random values assigned to vinegar variables and the solutions to oil variables from previous layers as new vinegar variables. Different from the rainbow signature scheme, the Himq-3 signature scheme uses the L-invertible cycle system to invert the first and second layers, and Gaussian elimination is only performed in the last layer. Due to this reason, the authors claim that the times of signing and verification of Himq-3 are respectively 3.1 times and 1.3 times faster than those of rainbow at the 128-bit level of security [17]. In addition, the sparse polynomials of the central map make the secrete key relatively small. The authors also claim that the secrete key size of Himq-3 is only 11.5% of that of rainbow. However, the L-invertible cycle system does not only speed up the signing process, but also puts restriction on certain variables. As we metioned in earlier, the cycle variables in the L-invertible cycle system cannot be equal to zero for any validly made signatures. One can see that the o_1 and o_2 variables are the cycle variables in the central map. Therefore, these nonzero variables give away the randomness of Himq-3. The comparison of key sizes, signature size and performance between Himq-3 and rainbow can be found in [17].

2.5 Himq-3 Variant: Himq-3F

Himq-3F is a generalization of Himq-3. Himq-3F fully fills the $v \times v$ parts in the first layer and $v \times o_1$ parts in the second layer. In addition, it shares the third layer with Himq-3. However, the quadratic product of the cycle variables in the central map of Himq-3F remains unchanged. Hence, the way to invert the central map of Himq-3F is essentially the same as Himq-3.

3 The Singularity Attack

3.1 Notations and Definitions

The central map of Himq-3

$$\mathcal{F} = (\mathcal{F}^{(1)}(x_1, \cdots, x_n), \cdots, \mathcal{F}^{(m)}(x_1, \cdots, x_n))$$

is defined in the same way as in Sect. 2. Let $\mathcal{S} : \mathbb{F}_q^m \to \mathbb{F}_q^m$ and $\mathcal{T} : \mathbb{F}_q^n \to \mathbb{F}_q^n$ be two invertible affine linear maps such that the public key is in the form:

$$\mathcal{P} = \mathcal{S} \circ \mathcal{F} \circ \mathcal{T} = (\mathcal{P}^{(1)}(x_1, \cdots, x_n), \cdots, \mathcal{P}^{(m)}(x_1, \cdots, x_n)).$$

Let $\mathbf{Q}_i$ be the symmetric matrix associated to the quadratic part of $\mathcal{F}^{(i)}$ for $i = 1, \cdots, m$. The matrix $\mathbf{P}_i$ denotes the symmetric matrix associate to the quadratic part of public key polynomials $\mathcal{P}^{(i)}$ for $i = 1, \cdots, m$. Let $\mathbf{S}$ and $\mathbf{T}$ be the matrix representations of $\mathcal{S}$ and $\mathcal{T}$ respectively. Next, We define $\mathbf{Q}'_i = \mathbf{T}^t \mathbf{Q}_i \mathbf{T}$ for $1 \le i \le m$, and $\mathcal{F}'_i = \mathbf{X}^t \mathbf{Q}'_i \mathbf{X}$ for $1 \le i \le m$.

We further define some subspaces of $\mathbb{F}_q^n$ as follows:

$V = \{\mathbf{X} \in \mathbb{F}_q^n : x_{v+1} = \cdots = x_n = 0\}$,
$O_1 = \{\mathbf{X} \in \mathbb{F}_q^n : x_1 = \cdots = x_v = x_{v+o_1+1} = \cdots = x_n = 0\}$,
$O_2 = \{\mathbf{X} \in \mathbb{F}_q^n : x_1 = \cdots = x_{v+o_1} = x_{v+o_1+o_2+1} = \cdots = x_n = 0\}$,
$O_3 = \{\mathbf{X} \in \mathbb{F}_q^n : x_1 = x_2 = \cdots = x_{v+o_1+o_2} = 0\}$,
$VO_1O_2 = \{\mathbf{X} \in \mathbb{F}_q^n : x_{v+o_1+o_2+1} = \cdots = x_n = 0\}$, and
$VO_1 = \{\mathbf{X} \in \mathbb{F}_q^n : x_{v+o_1+1} = \cdots = x_n = 0\}$.

3.2 General Idea of the Attack

The key observation is that the cycle variables cannot be equal to zero when evaluated at a honestly generated signature. In addition, this fact does not change under the change of basis $\mathcal{T}$. In other words, even if $\mathcal{T}$ is applied to mix the variables, the positions in the L-invertible cycle system part in the polynomials $\mathcal{F}^{(i)}$ for $1 \le i \le o_1 + o_2$ still cannot be equal to zero no matter what linear combinations of variables are plugged in. Since the scheme is constructed over a finite field with 2^k elements, it is a basic knowledge that if we raise any nonzero element a in the field to the power of $2^k - 1$, then $a^{2^k-1} = 1$. For this reason, if we evaluate the transformed cycle variables at the signatures under the effect of $\mathcal{T}$, and then raise their powers to $2^k - 1$, we will obtain some equations in variables of $\mathcal{T}$. Thus, if we have access to enough signatures, we will obtain enough equations. If the system of equations can be solved, we will get partial information about the private key $\mathcal{T}$ which immediately gives us the transformed cycles variables. The next step is to use those transformed cycle variables to further separate the layers and other variables. This can be accomplished easily by basic linear algebra. The Himq-3F keeps the L-invertible cycle system in the first and second layer of its central map. The same restriction applies to these variables in the L-invertible cycle system in Himq-3F. Hence, the singularity attack works for the Himq-3F as well.

3.3 Finding the Cycle Variables

Suppose that the private key $(\mathcal{F}, \mathcal{T}, \mathcal{S})$ has been generated with its corresponding public key $\mathcal{P} = \mathcal{S} \circ \mathcal{F} \circ \mathcal{T}$. The private key $\mathcal{T}$ can be expressed as an invertible matrix $(a_{ij})_{1\leq i,j\leq n}$ and a vector $\mathbf{b} = (b_1, \cdots, b_n)$ so that for any $(x_1, \cdots, x_n) \in \mathbb{F}_q^n$, we have that

$$\mathcal{T}((x_1,\cdots,x_n)) = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_{1i}x_i + b_1 \\ \sum_{i=1}^n a_{2i}x_i + b_2 \\ \vdots \\ \sum_{i=1}^n a_{ni}x_i + b_n \end{bmatrix}.$$

Our goal is to find how the private key $\mathcal{T}$ transforms the cycle variables used in the L-invertible cycle system (up to a multiplication by a non-zero constant). Namely, we want to find the transformed cycle variables in the form of linear combinations of $\gamma_j\,(\sum_{i=1}^n a_{ji}x_i + b_j)$ for $v+1 \leq j \leq v+o_1+o_2$, and for some nonzero constant $\gamma_j \in \mathbb{F}_q$. Let us denote a signature by $\sigma = (\sigma_1, \cdots, \sigma_n)$, then for $v+1 \leq j \leq v+o_1+o_2$ we have that $\sum_{i=1}^n a_{ji}\sigma_i + b_j \neq 0$ because a cycle variable cannot be zero when evaluated at a signature by the signing process as described above. Since $\mathbb{F}_q$ is a finite field with $q = 2^k$ elements, the nonzero elements of $\mathbb{F}_q$ form a multiplicative group $\mathbb{F}_q^*$. So for any $\gamma_j \in \mathbb{F}_q^*$ and for any signature σ, we obtain that

$$1 = \left(\sum_{i=1}^n \gamma_j a_{ji}\sigma_i + \gamma_j b_j\right)^{2^k-1} = \prod_{h=1}^k \left(\sum_{i=1}^n \gamma_j a_{ji}\sigma_i + \gamma_j b_j\right)^{2^{k-h}}$$

As we are working in characteristic two we have that

$$\prod_{h=1}^k \left(\sum_{i=1}^n \gamma_j a_{ji}\sigma_i + \gamma_j b_j\right)^{2^{k-h}} = \prod_{h=1}^k \left(\sum_{i=1}^n (\gamma_j a_{ji}\sigma_i)^{2^{k-h}} + (\gamma_j b_j)^{2^{k-h}}\right).$$

Since the vector $\mathbf{b}$ is randomly chosen, we first consider the main case when $b_j \neq 0$. The case in which $b_j = 0$ can be solved analogously. Now we can set $\gamma_j = b_j^{-1}$ to obtain

$$\prod_{h=1}^k \left(\sum_{i=1}^n (b_j^{-1} a_{ji}\sigma_i)^{2^{k-h}} + 1\right) = 1.$$

Let $\tilde{a}_{ji} = b_j^{-1}a_{ij}$ and perform the above product, we get

$$\tilde{a}_{j1}^{2^k-1}\sigma_1^{2^k-1} + \tilde{a}_{j1}^{2^k-2}\tilde{a}_{j2}\sigma_1^{2^k-2}\sigma_2 + \cdots + \tilde{a}_{jn}\sigma_n + 1 = 1.$$

If we treat the individual monomials of the $\tilde{a}_{ij}$'s as individual variables, we obtain a homogeneous linear equation with $(n+1)^k - 1$ terms. We get another homogeneous linear equation if we use a different signature. Hence by collecting around $(n+1)^k-1$ signatures we can build a linear system.

For $v+1 \leq j \leq v+o_1+o_2$, we list the monomials of $\tilde{a}_{ij}$ in the order: $\tilde{a}_{j1}^{2^k-1}, \tilde{a}_{j1}^{2^k-2}\tilde{a}_{j2}, \cdots, \tilde{a}_{jn}$. Moreover, for each signature $\sigma_i = (\sigma_{i,1}, \cdots, \sigma_{i,n})$, the corresponding coefficients are: $\sigma_{i,1}^{2^k-1}, \sigma_{i,1}^{2^k-2}\sigma_{i,2}, \cdots, \sigma_{i,n}$. A matrix can be simply constructed by having these corresponding coefficients as a row for each signature we use. Therefore the size of this matrix is $(n+1)^k-1$ by $(n+1)^k-1$ if we use $(n+1)^k-1$ signatures. If follows that we obtain a homogeneous linear system: $\mathbf{Ax}=\mathbf{0}$, where $\mathbf{A}$ is the matrix whose rows are $(\sigma_{i,1}^{2^k-1}, \sigma_{i,1}^{2^k-2}\sigma_{i,2}, \cdots, \sigma_{i,n})$ for each signature σ_i, and the vector $\mathbf{x} = (\tilde{a}_{j1}^{2^k-1}, \tilde{a}_{j1}^{2^k-2}\tilde{a}_{j2}, \cdots, \tilde{a}_{jn})^t$.

Remark 4. Assume that $b_j \neq 0$, for $v+1 \leq j \leq v+o_1+o_2$, the vector $\tilde{\mathbf{a}}_j = (\tilde{a}_{j1}^{2^k-1}, \tilde{a}_{j1}^{2^k-2}\tilde{a}_{j2}, \cdots, \tilde{a}_{jn})^t$ is contained in the kernel of $\mathbf{A}$. Moreover, it is obvious that they are linearly independent. It follows that $\text{Rank}(\mathbf{A}) \leq (n+1)^k-1-(o_1+o_2)$. In fact, according to our experiments, with overwhelming probability, $\text{Rank}(\mathbf{A}) = (n+1)^k-1-(o_1+o_2)$.

To solve the linear system, we first perform a Gaussian elimination on this matrix $\mathbf{A}$, and turn the linear system into a reduced echelon form $\mathbf{A}'\mathbf{x}=\mathbf{0}$. We start at the bottom of $\mathbf{A}'$. If $\mathbf{A}$ has rank $(n+1)^k-1-(o_1+o_2)$, then in the last nonzero row of $\mathbf{A}'$, most entries will equal to zero and the nonzero entries will only appear in the last o_1+o_2+1 columns in variables $\tilde{a}_{jn}^{o_1+o_2+1}, \tilde{a}_{jn}^{o_1+o_2}, \tilde{a}_{jn}^{o_1+o_2-1}, \cdots, \tilde{a}_{jn}$. Hence, converting this back into a polynomial means we have a univariate polynomial equation which we can thus solve by the Berlekamp's algorithm. One can see that if $2^k-1 \geq o_1+o_2+1$, we will obtain a univariate polynomial. Solving the univariate polynomial allows us to get our possibilities for $\tilde{a}_{jn}$ (as the above equation will be true for any of the $\tilde{a}_{ji}$'s, $v+1 \leq j \leq v+o_1+o_2$, we will return all of these values). We then move up the matrix to the first time that $\tilde{a}_{j(n-1)}$ appears only with powers of itself and $\tilde{a}_{jn}$. As we already know what $\tilde{a}_{jn}$ can be, this is also a univariate polynomial equation. For each of our possible solutions to $\tilde{a}_{jn}$, we plug in and get the possible solutions to $\tilde{a}_{j(n-1)}$. Continue this process until we collect all the $\tilde{a}_{ji}$ for which $b_j \neq 0$. On the other hand, to avoid the inequality $2^k-1 \geq o_1+o_2+1$, the size of the field is then forced to be small, which will reduce the complexity of other attacks such as a direct attack or a min/high rank attack [17].

Remark 5. The process is essentially the same as for the case $b_j = 0$ except that we then guess the last available $\tilde{a}_{ji}$ to be non-zero hence enabling us to set $\gamma_j = \tilde{a}_{ji}^{-1}$ for that particular $\tilde{a}_{ji}$. Repeat until all of the $\tilde{a}_{ji}$ are found, which generally is after the first few guesses. Since there are less variables in this case, the resulting matrix is of smaller size than the previous matrix. A toy exam is provided in the Appendix to demonstrate this step.

The collection of $\tilde{a}_{ji}$ that we found actually tells us the transformed cycle variables. Let us denote by

$$x'_j = \begin{cases} \sum_{i=1}^{n} \tilde{a}_{ji} x_i + 1 \text{ if } b_j \neq 0 \\ \sum_{i=1}^{n} \tilde{a}_{ji} x_i \text{ if } b_j = 0 \end{cases}$$

for $v + 1 \leq j \leq v + o_1 + o_2$, the transformed cycle variables under the effect of $\mathcal{T}$. Next we will use these variables to further separate the layers.

3.4 Extract the Second Layer

Let us recall how the polynomials of the central map are defined in these three different layers. Each first layer polynomial contains Φ_i where only v variables times one of themselves. Moreover, quadratic terms of a v variable multiplied by an o_3 variable appear in every third layer polynomial. In addition, in a second layer polynomial, every quadratic term contains a cycle variable (either an o_1 or an o_2 variable) as a factor. Thus, if we set the cycle variables equal to 0, the quadratic terms in the polynomials of the second layer will vanish but not those from first and third layers. Since we found the transformed cycle variables $\{x'_{v+1}, \cdots, x'_{v+o_1+o_2}\}$, we will use them to extract the second layer.

Setting the transformed cycle variables equal to zero can be accomplished by constructing the quotient ring

$$\mathbb{F}_q[x_1, \cdots, x_n]/\langle x'_{v+1}, \cdots, x'_{v+o_1+o_2} \rangle.$$

Let ϕ be the natural homomorphism:

$$\phi : \mathbb{F}_q[x_1, \cdots, x_n] \longrightarrow \mathbb{F}_q[x_1, \cdots, x_n]/\langle x'_{v+1}, \cdots, x'_{v+o_1+o_2} \rangle.$$

Consider the polynomials $\phi(\mathcal{P}^{(i)}) = \tilde{\mathcal{P}}^{(i)}$ for $i = 1, \cdots, m$. The quadratic terms in the second layer polynomials will vanish in this quotient ring, while the quadratic terms in the first and third layer polynomials will not. Let us construct a matrix $\mathbf{M}_1$ whose rows are formed by the coefficients of quadratic terms of each $\tilde{\mathcal{P}}^{(i)}$ for $i = 1, \cdots, m$. The matrix $\mathbf{M}_1$ cannot be of full rank because the polynomials $\tilde{\mathcal{P}}^{(1)} \cdots, \tilde{\mathcal{P}}^{(m)}$ do not contain any quadratic terms from the second layer polynomials, which already vanish in the quotient ring. If we apply a Gaussian elimination on this matrix $\mathbf{M}_1$, the bottom o_2 rows will all be zero, and they represent the quadratic part of the second layer polynomials in the quotient ring. By applying the same Gaussian elimination over the public keys, we can get o_2 linear combinations of the polynomials of the second layer by themselves, namely, o_2 linear combinations of $\mathcal{F}'_i$ (equivalently o_2 linear combinations of $\mathbf{Q}'_i$) for $o_1 + 1 \leq i \leq o_1 + o_2$ are found. Let $\bar{\mathcal{F}}_i$ be those o_2 linear combinations of $\mathcal{F}'_{o_1+1}, \cdots, \mathcal{F}'_{o_1+o_2}$ for $i = o_1 + 1, \cdots, o_1 + o_2$. Let us denote by $\bar{\mathbf{Q}}_i$ the symmetric matrices associated to the quadratic part of $\bar{\mathcal{F}}_i$ for $i = o_1 + 1, \cdots, o_1 + o_2$. The structure of those polynomials is not visible yet since there is a change of basis $\mathcal{T}$ still acting on them. Having the second layer extracted will enable us to further separate the variables.

3.5 Distinguish o_1 Variables from o_2 Variables

The variables $\{x'_{v+1}, \cdots, x'_{v+o_1+o_2}\}$ we obtained in Sect. 4.3 can either be a transformed o_1 variable or a transformed o_2 variable under the change of basis of $\mathcal{T}$. We will use the second layer that we extracted to distinguish which type of cycle variable they act under the effect of $\mathcal{T}$. Observe that the quadratic terms in a second layer polynomial in the central map are either a product of an o_1 variable multiplied by an v variable or a product of an o_2 variable multiplied by another o_2 variable. Hence, we can set all the variables $x'_{v+1}, \cdots, x'_{v+o_1+o_2}$ equal to zero except one. If the one left is a transformed o_1 variable under the effect of $\mathcal{T}$, then quadratic part of $\mathcal{F}'_i$ will not vanish for $o_1+1 \leq i \leq o_1+o_2$. If it is a transformed o_2 variable under the effect of change of basis, then the quadratic part of $\mathcal{F}'_i$ will vanish for $o_1+1 \leq i \leq o_1+o_2$. As we already obtained o_2 linear combinations $\bar{\mathcal{F}}_i$ of $\mathcal{F}'_{o_1+1}, \cdots, \mathcal{F}'_{o_1+o_2}$ in Sect. 3.4, we can construct the quotient rings one by one, and check if the quadratic part of $\bar{\mathcal{F}}_i$ for $i = o_1+1, \cdots, o_1+o_2$ vanishes or not in the quotient rings. It follows that we will immediately know which x'_j is a transformed o_1 variable and which one is a transformed o_2 variable under the effect of $\mathcal{T}$.

3.6 Getting the Linear Combinations of First and Second Layers

In the central map, o_3 variables only appear in the third layer, and they are multiplied by v, o_1 and o_2 variables. Hence, we may use o_3 variables to get rid of the third layer. It is obvious that the space O_3 is contained in the kernel of $\mathbf{Q}_i$ for $o_1+1 \leq i \leq o_1+o_2$. So it follows that $\mathcal{T}^{-1}(O_3)$ can be found by taking intersections of ker $\bar{\mathbf{Q}}_i$ for $o_1+1 \leq i \leq o_1+o_2$. In addition, for $i = o_1+1, \cdots, o_1+o_2$, the image of $\mathbf{Q}_i$ is contained in the space VO_1O_2. Therefore, $\mathcal{T}^{-1}(VO_1O_2)$ can be obtained by collecting the images of $\bar{\mathbf{Q}}_i$ for $o_1+1 \leq i \leq o_1+o_2$. Note that we may not get the full space $\mathcal{T}^{-1}(VO_1O_2)$ in general, we provide an analysis for the probability of getting the full space in the Appendix. One will see that for the proposed parameters, the space can be obtained with overwhelming probability.

Having these two spaces allows us to perform a change of basis on the public key so that the variables will be placed in their own positions. Take the o_3 basis vectors of $\mathcal{T}^{-1}(O_3)$ and the $v+o_1+o_2$ basis vectors of $\mathcal{T}^{-1}(VO_1O_2)$, and perform a change of basis on $\mathbf{P}_i$ for $i = 1, \cdots, m$. We get new matrices $\mathbf{P}'_1, \cdots, \mathbf{P}'_m$. The quadratic terms of a v, o_1 and o_2 variable multiplied by an o_3 variable will be in their own submatrix.

The o_3 variables do not appear in the polynomials of the first and second layer at all, hence for a first or second layer polynomial, the submatrix in the top right/down left corner should vanish. On the other hand, the third layer polynomials contains quadratic monomials of vo_3, o_1o_3 and o_2o_3. Hence for a third layer polynomial, the submatrix in top right/down left corner will not vanish.

$$\left[\begin{array}{ccccc|c} \multicolumn{5}{c|}{VO_1O_2} & O_3 \\ * & * & * & * & * & \\ * & * & * & * & * & \\ * & * & * & * & * & \\ * & * & * & * & * & \\ * & * & * & * & * & \\ - & - & - & - & - & - \\ & & & & & \end{array}\right] \left[\begin{array}{ccccc|c} \multicolumn{5}{c|}{VO_1O_2} & O_3 \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ - & - & - & - & - & - \\ * & * & * & * & * & \end{array}\right]$$

We use a similar method stated in Subsect. 3.4 to get the linear combinations of polynomials of first and second layer. Let us construct a matrix $\mathbf{M}_2$ whose rows are formed by the entries in the top right (vo_1o_2 by o_3) submatrix of each $\mathbf{P}'_i$ for $i = 1, \cdots, m$. Then the matrix $\mathbf{M}_2$ cannot be full rank since there are $o_1 + o_2$ zero rows generated by the first and second layer polynomials which are mixed by $\mathcal{S}$ with other nonzero rows. Apply a Gaussian elimination on the matrix $\mathbf{M}_2$, the bottom $o_1 + o_2$ zero rows will represent the linear combinations of first and second layer polynomials. Apply the same Gaussian elimination over the public key, one obtains $o_1 + o_2$ linear combinations of polynomials $\mathcal{F}'_i$ (equivalently $o_1 + o_2$ linear combinations of $\mathbf{Q}'_i$) for $1 \leq i \leq o_1 + o_2$. Due to the change of basis map $\mathcal{T}$, the structure of those polynomials is not visible. For simplicity, let $\tilde{\mathbf{Q}}_1, \cdots, \tilde{\mathbf{Q}}_{o_1+o_2}$ denote these $o_1 + o_2$ linear combinations of matrices $\mathbf{Q}'_i$

3.7 Separate the First Layer Out

Let us consider the symmetric matrices associated to the linear combinations of polynomials of first and second layers in the central map (these symmetric matrices can be visualized by overlapping $\mathbf{Q}_i$ for $1 \leq i \leq o_1 + o_2$. See the pictures of these matrices in Sect. 2). The entries representing the cycle products of an o_2 variable multiplied by one of themselves are in different spots, and the submatrices of the o_2 by o_2 part are one off the full rank. Thus, if we take the images of these symmetric matrices and then take intersections of those image spaces, we can get rid of images produced by the entries in the o_2 by o_2 part and the space VO_1 can be obtained. It follows that the space $\mathcal{T}^{-1}(VO_1)$ can be found by taking the images of $\tilde{\mathbf{Q}}_i$ for $1 \leq i \leq o_1 + o_2$, then taking the intersections.

Now we use a similar method to extract the first layer as we did in Subsect. 3.6. We can perform a change of basis on the public key to turn the variables to their own positions since we have the space $\mathcal{T}^{-1}(VO_1)$, the exact transformed o_2 variables under the effect of $\mathcal{T}$, and the space $\mathcal{T}^{-1}(O_3)$. After performing a change of basis on $\mathbf{P}_i$ for $i = 1, \cdots, m$, the o_2 and o_3 variables will go to their own positions, but v and o_1 variables are still mixed together. We obtain the new matrices $\bar{\mathbf{P}}_i$ for $i = 1, \cdots, m$. Recall that in a first layer polynomial, there are quadratic terms of a v variable multiplied by a v variable, and an o_1 variable multiplied by another o_1 variable. Moreover, the second layer polynomials contain quadratic terms of a v variable multiplied by an o_1 variable and cycle products of an o_2 variable by an o_2 variable. Hence, for a first layer polynomial, the submatrix of o_2 by o_2 part will vanish. While for a second layer polynomial, the submatrix of o_2 by o_2 part will not vanish.

$$\left[\begin{array}{ccccc|cc} & & VO_1 & & & O_2 & O_3 \\ * & * & * & * & * & & \\ * & * & * & * & * & & \\ * & * & * & * & * & & \\ * & * & * & * & * & & \\ * & * & * & * & * & & \\ \hline & & & & & & \\ & & & & & & \\ & & & & & & \end{array}\right] \left[\begin{array}{ccccc|cc} & & VO_1 & & & O_2 & O_3 \\ * & * & * & * & * & & \\ * & * & * & * & * & & \\ * & * & * & * & * & & \\ * & * & * & * & * & & \\ * & * & * & * & * & & \\ \hline & & & & & & \\ & & & & & * \; * & \\ & & & & & * \; * & \\ & & & & & & \end{array}\right]$$

Let us construct a matrix $\mathbf{M}_3$ whose rows are formed by the entries of each o_2 by o_2 submatrix of $\bar{\mathbf{P}}_i$ for $i = 1, \cdots, m$. It follows that the matrix $\mathbf{M}_3$ cannot be of full rank because there are o_1 zero rows generated by the first layer polynomials which are mixed by $\mathcal{S}$ with other nonzero rows. Perform a Gaussian elimination on $\mathbf{M}_3$, the bottom o_1 zero rows represent the first layer polynomials. Let us apply the same Gaussian elimination on the public key, we can get o_1 linear combinations of first layer polynomials, namely o_1 linear combinations of $\mathcal{F}'_i$ (equivalently o_1 linear combinations of $\mathbf{Q}'_i$) for $1 \leq i \leq o_1$. Again, because of the change of basis map $\mathcal{T}$, no structure can be seen from those polynomials. Let $\bar{\mathbf{Q}}_1, \cdots, \bar{\mathbf{Q}}_{o_1}$ be the o_1 linear combinations of $\mathbf{Q}'_i$ for $i = 1, \cdots, o_1$.

3.8 Getting Transformed V Space

Once the first layer is obtained, it is easy to get the space $\mathcal{T}^{-1}(V)$. In a linear combination of symmetric matrices $\mathbf{Q}_i$ for $i = 1, \cdots, o_1$, the entries representing the cycle products of an o_1 variable multiplied by another o_1 variable are in different spots, and the submatrix of the o_1 by o_1 part is one off full rank. Hence, taking the images of $\mathbf{Q}_i$ for $1 \leq i \leq o_1$ and then taking the intersections will yield the space V. It follows that $\mathcal{T}^{-1}(V)$ can be obtained by taking images of $\bar{\mathbf{Q}}_i$ for $1 \leq i \leq o_1$ and then taking the intersections.

3.9 Invert Change of Basis

We have extracted the first layer and the second layer from the public key, in other words, we have undone the work that the private key $\mathcal{S}$ does. Additionally, we now have all the information required to create a change of basis which will undo $\mathcal{T}$'s effect of hiding the cycle structure in the public key. We do not need the exact transformed v and transformed o_3 variables as they do not appear in the L-invertible cycle system. As long as these variables are mapped to a linear combination of themselves we will have no problem inverting the central map as done in the original scheme. Hence, having just the spaces of $\mathcal{T}^{-1}(V)$ and $\mathcal{T}^{-1}(O_3)$ is enough. However, the cycle variables must each be mapped to another cycle variable. That is, we must know exactly how $\mathcal{T}$ changed these variables, and also its affine part cannot be ignored. Fortunately, we have already found this up to a scalar multiple when we found $x'_{v+1}, \cdots, x'_{v+o_1+o_2}$.

3.10 Complexity

The most complicated step throughout the entire attack is to do a Gaussian elimination over the square matrix of size $(n+1)^k - 1$. The complexity of solving such linear system is $((n+1)^k - 1)^\omega$, where ω is called the complexity exponent of linear algebra [1]. The best published estimates to date gives $\omega \approx 2.3727$ [10, 20]. The complexity of singularity attack on Himq-3 for all three sets proposed parameters is stated in the table.

$\|\mathbb{F}_q\|$, v, o_1, o_2, o_3	Security level	# of Signatures	Complexity $\omega = 2.3727, \omega = 2$
2^8, 36, 15, 15, 15	128-bit	2^{51}	2^{120}, 2^{102}
2^8, 56, 25, 25, 25	192-bit	2^{56}	2^{134}, 2^{112}
2^8, 84, 33, 33, 32	256-bit	2^{60}	2^{143}, 2^{120}

It can be seen that the Himq-3 scheme does not meet the target levels of security. The complexities of our attack on Himq-3 with last two sets of proposed parameters are much lower than the target levels of security. It is obvious that the complexity of our attack is dominated by the size of the field. So we do not leave too much room for the authors of Himq-3 to save the scheme by choosing different parameters. The set of proposed parameters of Himq-3F for 128-bit level of security is $|\mathbb{F}_q| = 2^8$, $v = 36, o_1 = 13, o_2 = 17, o_3 = 15$. Thus, the complexity of the singularity attack on Himq-3F for this set of parameters is approximately 2^{121} if we use $\omega = 2.3727$ and 2^{102} if $\omega = 2$. So Himq-3F does not meet the claimed level of security.

4 Experimental Results

We ran our attack 100 times with Magma of version V2-24 on three sets of parameters and record the times it took to obtain part of $\mathcal{T}$ including evaluating the signatures. Our hardware is a workstation of Intel Core i7-9700, 8 Core, 12 MB Cache, 3.0 Ghz.

v, o_1, o_2, o_3	Field	Find cycle variables	Find $\mathcal{T}^{-1}(VO_1O_2)$	Time in seconds
$7, 3, 3, 2$	$q = 2^3$	100	100	7.651
$9, 3, 3, 2$	$q = 2^3$	100	86	20.770
$11, 5, 5, 4$	$q = 2^3$	100	100	302.843
$13, 3, 3, 2$	$q = 2^3$	100	0	115.252

5 Conclusion

We presented a complete cryptanalysis of a NIST round 1 submission Himq-3. This attack method may also be applied to other cryptosystems in which there are some restrictions on its variables. So our singularity attack is a warning for cryptographers not to restrict the variables used in design of central map from being zero. According to our complexity analysis and experimental results, Himq-3 and its variant Himq-3F can be defeated with overwhelming probability at much lower costs than the target security levels. However, our attack method does not apply to the rainbow scheme since there is no restriction on any variables in the scheme.

Acknowledgments. J. Ding, Z. Zhang and J. Deaton would like to thank partial support of NSF (Grant: $\#CNS - 1814221$). J. Ding would like to thank NIST, and the TAFT Research Center for many years' support. Finally, we are grateful for the comments of the referees helping us improve the quality of this paper.

A Toy Example

We provide a toy example to clarify the step 3.3. In this example, we choose $k = 3$, thus our field is the finite field of 2^3 elements. The finite field will be represented by $\{0, 1, w, w^2, \cdots, w^6\}$, where w is a generator in the multiplicative group of the finite field. Let $n = 2$. For the sake of clarity. We use a linear map instead of a affine map. Our linear map $\mathcal{T}$ is randomly chosen to be the matrix

$$\begin{bmatrix} w^2 & w^2 \\ w^3 & w \end{bmatrix}.$$

Suppose we obtain a set of signatures (x_1, x_2):

$$\begin{aligned}
&(w, w^5), (w^5, w), (w^2, 1), (w^6, w^5), (0, w^2), (w^5, w^3), (1, w^6), (0, w^5),\\
&(0, w^2), (1, 0), (w^5, w^6), (0, w), (w^5, w^3), (1, w), (w^5, 0), (w^6, 1), (w^6, w^3),\\
&(w, w^4), (w^2, w^5), (w^3, w), (1, w^6), (w, 1), (w^2, w), (w^2, w), (w^4, w), (w^4, 1), (w^4, w^2).
\end{aligned}$$

We first construct a generic polynomial $g = a_1x_1 + a_2x_2$. We assume that this polynomial is never equal to zero. Hence, in this finite field, $g^{2^3-1} = (a_1x_1 + a_2x_2)^{2^3-1} = 1$. We can rewrite this equation as: $(a_1x_1 + a_2x_2)^{2^3-1} = (a_1x_1 + a_2x_2)^{2^{3-1}}(a_1x_1 + a_2x_2)^{2^{3-2}}(a_1x_1 + a_2x_2)^{2^{3-3}} = 1$. Since this is a field of characteristic 2, the equations turns out to be

$$((a_1x_1)^{2^{3-1}} + (a_2x_2)^{2^{3-1}})((a_1x_1)^{2^{3-2}} + (a_2x_2)^{2^{3-2}})((a_1x_1)^{2^{3-3}} + (a_2x_2)^{2^{3-3}}) = 1.$$

Multiply the product out, we have

$$a_1^7x_1^7 + a_1^6a_2x_1^6x_2 + a_1^5a_2^2x_1^5x_2^2 + a_1^4a_2^3x_1^4x_2^3 + a_1^3a_2^4x_1^3x_2^4 + a_1^2a_2^5x_1^2x_2^5 + a_1a_2^6x_1x_2^6 + a_2^7x_2^7 + 1 = 0.$$

We view the products of a_i as variables, and x_i as coefficients. If we evaluate these coefficients at the signatures, we get $(n+1)^k = 27$ vectors which will be the rows of the matrix. We apply echelon form on this matrix and then remove the zero rows. The new matrix is:

$$\begin{bmatrix} 1\,0\,0\,0\,0\,0 & 0 & 0 & 1 \\ 0\,1\,0\,0\,0\,0 & w^5 & 0 & w^4 \\ 0\,0\,1\,0\,0\,0 & w^2 & 0 & w^6 \\ 0\,0\,0\,1\,0\,0 & w^4 & 0 & w^5 \\ 0\,0\,0\,0\,1\,0 & w^3 & 0 & w \\ 0\,0\,0\,0\,0\,1 & w^6 & 0 & w^2 \\ 0\,0\,0\,0\,0\,0 & 0 & 1 & 1 \end{bmatrix}$$

Our next goal is to turn this matrix back to polynomials. Recall the order of the monomials, we get 7 multivariate polynomials:

$$\begin{aligned} &a_1^7 + 1 \\ &a_1^6 a_2 + w^5 a_1 a_2^6 + w^4 \\ &a_1^5 a_2^2 + w^2 a_1 a_2^6 + w^6 \\ &a_1^4 a_2^3 + w^4 a_1 a_2^6 + w^5 \\ &a_1^3 a_2^4 + w^3 a_1 a_2^6 + w \\ &a_1^2 a_2^5 + w^6 a_1 a_2^6 + w^2 \\ &a_2^7 + 1 \end{aligned}$$

The first and last polynomials do not help, they are trivial. Remember that we are not looking for the original values for a_i, we only need solutions for a_i up to unit multiple. Therefore, we can set $a_1 = 1$, and if we pick the second polynomial, we then get a univariate polynomial $w^5 a_2^6 + a_2 + w^4$. The roots are $a_2 = 1$ and $a_2 = w^5$.

Let us check our solution with the linear map $\mathcal{T} = \begin{bmatrix} w^2 & w^2 \\ w^3 & w \end{bmatrix}$. It is clear that $a_1 = 1$ and $a_2 = 1$ are unit multiples of $a_1 = w^2$ and $a_2 = w^2$. Now if we check the second row, The original values are:

$$\begin{aligned} a_1 &= w^3 \\ a_2 &= w \end{aligned}$$

If we multiply the inverse of w^3 by w, we get w^{-2} which is exactly equal to w^5 in the finite field of 2^3 elements.

B Getting Transformed VO_1O_2 Space

We know that there are o_1 column vectors in the $v \times o_1$ part of each symmetric matrix $\mathbf{Q}_i$ for $i = o_1 + 1, \cdots, o_1 + o_2$. So we have $o_1 o_2$ such vectors. Assume that these $o_1 o_2$ vectors do not span the entire V space. Let us take $v-1$ vectors and look at the span of these $v-1$ vectors. Therefore, the probability of the next vector being in the span of these $v-1$ vector is $\frac{q^{v-1}-1}{q^v} \approx \frac{1}{q}$. There are $o_1 o_2 - (v-1)$ vectors to check, so the probability of failing to fill the entire space is $1/q^{o_1 o_2 - (v-1)}$. Thus we can conclude that if $o_1 o_2$ is larger enough than v, we can always get the full space. All the sets of proposed parameters satisfy this condition, so we do not need to worry about this case at all.

References

1. Albrecht, M.R., Bard, G.V., Pernet, C.: Efficient dense gaussian elimination over the finite field with two elements. arXiv preprint arXiv:1111.6549 (2011)
2. Beullens, W., Szepieniec, A., Vercauteren, F., Preneel, B.: LUOV: signature scheme proposal for NIST PQC project (2017)
3. Coppersmith, D., Stern, J., Vaudenay, S.: Attacks on the birational permutation signature schemes. In: Stinson, D.R. (ed.) CRYPTO 1993. LNCS, vol. 773, pp. 435–443. Springer, Heidelberg (1994). https://doi.org/10.1007/3-540-48329-2_37
4. Ding, J., Petzoldt, A.: Current state of multivariate cryptography. IEEE Secur. Priv. **15**(4), 28–36 (2017). https://doi.org/10.1109/MSP.2017.3151328
5. Ding, J., Gower, J., Schmidt, D.: Multivariate public key cryptosystems. In: Jajodia, S. (ed.) Advances in Information Security. Springer (2006)
6. Ding, J., Petzoldt, A., Wang, L.: The cubic simple matrix encryption scheme. In: Mosca, M. (ed.) PQCrypto 2014. LNCS, vol. 8772, pp. 76–87. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-11659-4_5
7. Ding, J., Schmidt, D.: Rainbow, a new multivariable polynomial signature scheme. In: Ioannidis, J., Keromytis, A., Yung, M. (eds.) ACNS 2005. LNCS, vol. 3531, pp. 164–175. Springer, Heidelberg (2005). https://doi.org/10.1007/11496137_12
8. Ding, J., Wolf, C., Yang, B.-Y.: l-invertible cycles for $\mathcal{M}$ultivariate $\mathcal{Q}$uadratic ($\mathcal{MQ}$) public key cryptography. In: Okamoto, T., Wang, X. (eds.) PKC 2007. LNCS, vol. 4450, pp. 266–281. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-71677-8_18
9. Ding, J., Zhang, Z., Deaton, J., Schmidt, K., Vishakha, F.: New attacks on lifted unbalanced oil vinegar. In: The 2nd NIST PQC Standardization Conference (2019)
10. Dumas, J.G., Pernet, C.: Computational linear algebra over finite fields. arXiv preprint arXiv:1204.3735 (2012)
11. Goubin, L., Courtois, N.T.: Cryptanalysis of the TTM cryptosystem. In: Okamoto, T. (ed.) ASIACRYPT 2000. LNCS, vol. 1976, pp. 44–57. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-44448-3_4
12. Kipnis, A., Patarin, J., Goubin, L.: Unbalanced oil and vinegar signature schemes. In: Stern, J. (ed.) EUROCRYPT 1999. LNCS, vol. 1592, pp. 206–222. Springer, Heidelberg (1999). https://doi.org/10.1007/3-540-48910-X_15
13. Kipnis, A., Shamir, A.: Cryptanalysis of the oil and vinegar signature scheme. In: Krawczyk, H. (ed.) CRYPTO 1998. LNCS, vol. 1462, pp. 257–266. Springer, Heidelberg (1998). https://doi.org/10.1007/BFb0055733
14. Matsumoto, T., Imai, H.: Public quadratic polynomial-tuples for efficient signature-verification and message-encryption. In: Barstow, D., et al. (eds.) EUROCRYPT 1988. LNCS, vol. 330, pp. 419–453. Springer, Heidelberg (1988). https://doi.org/10.1007/3-540-45961-8_39

15. Patarin, J.: Cryptanalysis of the Matsumoto and Imai public key scheme of eurocrypt'88. In: Coppersmith, D. (ed.) CRYPTO 1995. LNCS, vol. 963, pp. 248–261. Springer, Heidelberg (1995). https://doi.org/10.1007/3-540-44750-4_20
16. Patarin, J.: The oil and vinegar algorithm for signatures. In: Dagstuhl Workshop on Cryptography 1997 (1997)
17. Shim, P., Kim: HIMQ-3: a high speed signature scheme based on multivariate quadratic equations (2017)
18. Shor, P.W.: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM Rev. **41**(2), 303–332 (1999)
19. Tao, C., Diene, A., Tang, S., Ding, J.: Simple matrix scheme for encryption. In: Gaborit, P. (ed.) PQCrypto 2013. LNCS, vol. 7932, pp. 231–242. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38616-9_16
20. Williams, V.V.: Breaking the Coppersmith-Winograd barrier (2011)