University of Louisville
ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

8-2019

A study of big field multivariate cryptography.

Ryann Cartor
University of Louisville

Follow this and additional works at: https://ir.library.louisville.edu/etd

 Part of the Other Mathematics Commons

Recommended Citation
Cartor, Ryann, "A study of big field multivariate cryptography." (2019). *Electronic Theses and Dissertations.* Paper 3294.
https://doi.org/10.18297/etd/3294

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.

A STUDY OF BIG FIELD MULTIVARIATE CRYPTOGRAPHY

By

Ryann Cartor
B.S., Bellarmine University, 2014
M.A., University of Louisville, 2016

A Dissertation
Submitted to the Faculty of the
College of Arts and Sciences of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy
in
Applied and Industrial Mathematics

Department of Mathematics
University of Louisville
Louisville, Kentucky

August 2019

A STUDY OF BIG FIELD MULTIVARIATE CRYPTOGRAPHY

Submitted by

Ryann Cartor

A Dissertation Approved on

May 17, 2019

by the Following Dissertation Committee:

Dr. Daniel Smith-Tone,
Dissertation Director

Dr. David Brown

Dr. Hamid Kulosman

Dr. Jinjia Li

Dr. Steven Seif

DEDICATION

To the Super Six (Mom, Dad, Max, Austin, and Allison).

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Daniel Smith-Tone for all of his help, advice, and inspiration. I would also like to thank my committee members for their time and dedication. I am incredibly grateful for the entire University of Louisville Math Department for supporting me throughout this journey and making the Natural Sciences building feel like a second home. I am especially grateful for my cohort members, Katie and Trevor. We arrived, survived, and thrived together.

I would also like to thank my family, for loving me and supporting me on the good days and on the crazy days. And a special thank you to Austin, for all that you do.

ABSTRACT

A STUDY OF BIG FIELD MULTIVARIATE CRYPTOGRAPHY

Ryann Cartor

May 17, 2019

As the world grapples with the possibility of widespread quantum computing, the cryptosystems of the day need to be up to date. Multivariate Public Key Cryptography is a leading option for security in a post quantum society. One goal of this work is to classify the security of multivariate schemes, especially C^* variants. We begin by introducing Multivariate Public Key Cryptography and will then discuss different multivariate schemes and the main types of attacks that have been proven effective against multivariate schemes. Once we have developed an appropriate background, we analyze security of different schemes against particular attacks. Specifically, we will analyze differential security of HFEv^- and PFLASH schemes. We then introduce a variant of C^* that may be used as an encryption scheme, not just as a signature scheme. Finally, we will analyze the security and efficiency of a (n, d, s, a, p, t) scheme in general. This allows for individuals to generally discuss security and performance of any C^* variant.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
AN INTRODUCTION TO CRYPTOGRAPHY

In order to illustrate the basic concepts of cryptography, we will introduce characters Alice, Bob, and Charlie. Alice and Bob want to have a private conversation, while Charlie wants to spy on them. In order to keep the contents of their conversation secret, Alice and Bob may turn to cryptography to exchange encrypted messages. Alice and Bob have many choices when deciding upon a cipher code. One of the first things they must decide is if they will use a private key or a public key to encrypt. In private key cryptography, both parties have enough information to encrypt and decrypt. Because of this, the way that messages are encrypted is kept private.

Private Key Toy Example: Alice comes up with the secret encryption code $C \equiv P + 3 \mod 26$. This would mean $X \mapsto A, Y \mapsto B, Z \mapsto C, A \mapsto D$, etc. Through secret channels, Alice sends Bob her cryptosystem. Bob encrypts his message using Alice's cryptosystem and sends back the message "KHOOR!" Charlie would not know the meaning of this encrypted message, but Alice uses her encryption/decryption key to know that $P \equiv C - 3 \mod 26$, and that this message says "Hello!" The secrecy of the encryption key is of utmost importance because if Charlie were to gain access to the key, he would automatically be able to decrypt messages.

Diffie-Hellman Key Exchange The purpose of the Diffie-Hellman Key Exchange is to create a shared secret between two parties. Alice and Bob will want to create a private key, but we consider that they do not have a way of sharing information through secure channels, meaning Charlie can spy on any correspondence between the two. The way to create this shared secret is to first decide on a prime number p and a base element $g \in \mathbb{F}_p \cong \mathbb{Z}_p$. This information cannot be secretly transmitted, so Charlie has access to the values of p and g. Now Alice will choose some integer $k_A < p$ and computes $g^{k_a} \mod p := a$, and Bob will choose $k_B < p$ and computes $g^{k_a} \mod p := b$. Alice and Bob will keep k_A and k_B private, but will send the values a and b to the other party. Now Alice and Bob can let the value $g^{k_A k_B} \in \mathbb{F}_p$ be their key.

Bob will compute this value by raising the value a he received from Alice to the power of his secret number k_B.

$$a^{k_B} \mod p = (g^{k_A})^{k_B} \mod p = g^{k_A k_B} \mod p$$

Alice will, respectively, raise b to the power k_A. Now both parties have the same shared secret, while Charlie does not have enough information to find the key. This is the basis of the Diffie-Hellman Problem, which is if you are given $g, a, b \in \mathbb{F}_p$ to then find $g^{k_A k_B}$. Clearly if an individual can solve the discrete log problem, the individual could also solve the Diffie-Hellman problem. The Discrete Log Problem is GI-hard, and it is conjectured that the difficulty of the Diffie-Hellman Problem is equivalent to that of the Discrete Log Problem. Problems that are GI-hard tend to be sub-exponential in complexity, but still difficult for classical computers.

1.1 Public Key Cryptography

Public-key cryptography centers around the idea of "one-way functions." These are functions that if you are given an input, it is easy to compute an output.

But, if you are given only an output, it is unreasonably difficult to invert the function without additional information.

The Diffie-Hellman Key Exchange creates a shared secret where the key and the shared secret are kept private. Public key schemes may be used for encryption in order to create a shared secret, where the encryption key is public. In practice, one individual may create a random string of numbers, use a public key to encrypt that random string, and then send the encrypted random message to the other party who may then decrypt the cipher-text to understand the original string. The random string is thus the two parties' shared secret. Using our characters, we say that Alice creates one way functions, which she uses to create her public key. Both Bob and Charlie can see the public equations. Bob encrypts his message, it looks like gibberish to Charlie, but Alice uses secret information only she knows to decrypt Bob's message

Signature Verification We can also use cryptosystems to guarantee the authenticity of a message. To describe how a scheme may be used for signature verification, we will again say that Alice and Bob would like to communicate, but now we will say that Charlie may try to send Bob information disguised as Alice. To defend against this, Alice and Bob will agree upon a signature scheme for their correspondence. Alice will create a public key P, and Alice and Bob will agree that Alice's signature will be the preimage of some certificate y. Alice created the public key, so she is able to compute $P^{-1}(y) := x$. She will write her message, sign it as x, and send it to Bob. Bob will receive her message, compute $P(x) = y$, and he will know the message came from Alice. If Charlie only knows that Alice is supposed to sign with the preimage of y, he will not know how to compute that using Alice's public key.

Hash Functions Hash functions are another way to provide security against message tampering. A hash function is a function, h, that will in general have the following properties:

- Collision free, meaning that if messages $x_1 \neq x_2$, then $h(x_1) \neq h(x_2)$.
- The same message x will always result in the same hash value $h(x)$.
- Small changes in the message should result in major changes in the hash value. So if x_1 and x_2 are similar, $h(x_1)$ and $h(x_2)$ are not.
- It is infeasible to guess x given $h(x)$
- Messages can be of variable length, while all hash values will have a fixed length.

We can see that generally speaking, a hash function is a function that pseudo-randomly assigns a message to a hash value, and that will always follow that assignment once it has been made. Well known hash functions include SHA functions (0 through 3), RIPEMD-160, BLAKE2, Whirlpool, and more.

To illustrate how hash functions may be used, consider the following scenario. Bob would like to send Alice a message. After Alice receives the message, she wants to make sure the message she is reading is actually what Bob sent, and that Charlie has not tampered with or altered the message in any way. In order to protect against outside tampering, Alice and Bob may choose to use a hash function.

The hash function is used along with the Bob's public key. Bob will send Alice $(x, v) = (x, P^{-}1(h(x)))$, and then Alice can check that P(v)=h(x). In order for Charlie to trick Alice, he would have had to find values x and v such that $P(v) = h(x)$. But, only Bob has access to the information necessary to invert P, and it would be infeasible for Charlie to guess a message x' such that $h(x) = h(x')$.

1.1.1 RSA

The public key cryptosystem RSA (named after authors Rivest, Shamir, and Adleman) was introduced in 1978 and is still widely used today. The security of this scheme is based on the difficulty of factoring large numbers into their prime factors. It is an open question whether breaking RSA has the same difficulty as the factoring problem.

We will consider the hypothetical situation that Alice and Bob would like to secretly converse in order to demonstrate how RSA works. Alice decides to create a public key to use for encryption. To do this, she first chooses two prime numbers, p and q, and computes $n = pq$. Alice will then choose some large integer d such that $\gcd(d, \varphi(n)) = 1$, where $\varphi(n) = (p-1)(q-1)$. Finally, she will find some e such that $e \cdot d \equiv 1 \pmod{\varphi(n)}$. Clearly, this means $e \cdot d = k\varphi(n) + 1$ for some k.

The reason behind computing integers e and d is to create a multiplicative inverse for decryption. It is clear that $\mathbb{Z}_n$ has zero divisors when n is not prime, but we can consider the unit group (the set of all invertible elements) of $\mathbb{Z}_n$. We will denote the unit group of $\mathbb{Z}_n$ as U_n. For primes p and q, where $n = p \times q$, $U_n \cong U_p \times U_q$, where U_p has size $p-1$, and U_q has size $q-1$. Thus, U_n has size $\varphi(n) = (p-1) \times (q-1)$. So for any $x \in U_n$, $x^{\varphi(n)} = (1 \pmod{n})$.

If Bob wants to send Alice the message M (represented as an integer between 0 and $n-1$) he will compute $C \equiv M^e \pmod{n}$. Once Alice receives the cipher text C, she computes

$$C^d = (M^e)^d = M^{k\varphi(n)+1} = (M^{\varphi(n)})^k M$$

$$C^d \equiv M \pmod{n}$$

and she then understands the original plain text. Adversary Charlie only sees C, n, and e, and is unable to find any multiplicative inverse d without knowing the prime factors of n.

1.2 Post Quantum Cryptography

In 1994 Peter Shor developed a polynomial-time algorithm that computes the prime factorization of a number on quantum computers. This algorithm will render RSA and similar schemes useless when quantum computing becomes prevelant. Post-quantum cryptography focuses on schemes where there is no clear quantum advantage.

1.3 Multivariate Cryptography

A specific subset of post-quantum cryptography is multivariate cryptography. Multivariate cryptosystems are composed of systems of quadratic equations, and the security of these schemes is based on the MQ-problem. The MQ-problem is the problem of solving systems of quadratic equations over a field. This problem is known to be NP-hard, which suggests it will be difficult even for quantum computers. The "one-way functions" used to construct multivariate schemes are created by function composition. Composing multiple easily invertible maps results in a function that is difficult to invert without knowing each of the individual functions.

CHAPTER 2
BIG FIELD MULTIVARIATE SCHEMES

2.1 Introduction to Big Field Schemes

Many multivariate schemes make use of the fact that given a finite field $\mathbb{F}_q$ and a degree n extension $\mathbb{K}$, then $\mathbb{K}$ is an $\mathbb{F}_q$-algebra. Schemes utilizing this structure are known as "big field" schemes.

By choosing a vector space isomorphism $\phi : \mathbb{F}_q \to \mathbb{K}$, we are guaranteed an equivalence between systems F and f, where F is a set of n polynomials in n variables over $\mathbb{F}_q$, and f is a univariate polynomial of the form

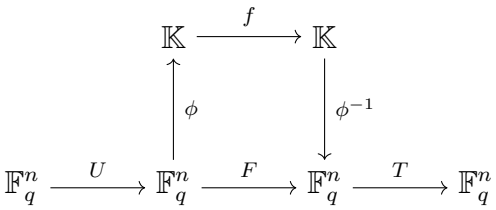
$$f(x) = \sum_{0 \leq i \leq j < n} \alpha_{ij} x^{q^i + q^j}, \quad x \in \mathbb{K}.$$

This equivalence is given by $F = \phi^{-1} \circ f \circ \phi$

To hide the structure of an easily invertible map F, the standard technique is to apply a morphism of polynomials, essentially choosing random linear maps that mix the input and output spaces of the central map. Formally, we define these morphisms as follows.

Definition 1. *A* polynomial morphism *between two systems of polynomials is a pair of affine maps (T, U) such that $G = T \circ F \circ U$. If both T and U are invertible, then the morphism is said to be an isomorphism and F and G are said to be isomorphic.*

Thus, for big field schemes, the construction of a public key can be summarized with the following diagram.



2.2 C* Scheme

Matsumoto and Imai introduced the first massively multivariate cryptosystem now known as C^* at Eurocrypt '88, [1]. The C^* scheme is a big field construction in which the central map $f : \mathbb{K} \to \mathbb{K}$ is the $\mathbb{F}_q$-quadratic monomial map $f(x) = x^{q^\theta+1}$. We call this function $\mathbb{F}_q$-quadratic because it is the product of two $\mathbb{F}_q$-linear terms (x^{q^θ} and x). In order to guarantee that this function has an inverse, it is required that $\gcd(q^{\theta+1}, q^n - 1) = 1$. The central map f is hidden by a polynomial morphism. In this case, the affine maps T and U both map from $\mathbb{F}_q^n$ to $\mathbb{F}_q^n$. Thus the public key is given by $P = T \circ \phi^{-1} \circ f \circ \phi \circ U$.

Encryption of a plaintext $x \in \mathbb{F}_q^n$ is accomplished by evaluating the public polynomials P at x. Decryption is accomplished by inverting each of the three component maps individually. The inversion of $v = f(u)$ is performed by solving $h(q^\theta + 1) = 1 \bmod (q^n - 1)$, and calculating $u = v^h$. The original intention for the C^* scheme was encryption, but it is also useful for digital signatures.

2.2.1 C* Toy Example

Creating the public key In order to complete a toy example, we must first choose our private information. We will let $q = 2$, which will make our base field $\mathbb{F}_2 = GF(2)$. We can choose $\mathbb{K}$ to be a degree $n = 5$ extension of $\mathbb{F}_q$, defined as $\mathbb{K} = \mathbb{F}/\langle x^5 + x^2 + 1\rangle$. We will choose our central map f to be $f(x) = x^{2^3+1}$. To create our public key, we will consider the input vector $\overline{x} = [x_1, x_2, x_3, x_4, x_5]^\top \in \mathbb{F}_2^5$.

We must also define our maps U and T, which for this example will be defined below:

$$U = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \end{bmatrix} \qquad T = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}$$

First we compose our input vector with our map U, which gives:

$$U\overline{x} = \begin{bmatrix} x_1 + x_2 + x_5 \\ x_1 + x_2 + x_4 \\ x_2 + x_3 + x_4 \\ x_1 + x_4 + x_5 \\ x_1 + x_2 + x_4 + x_5 \end{bmatrix}$$

We can then write $U\overline{x}$ as an element of the extension field by composing $U\overline{x}$ with our vector space isomporphism ϕ. Let β be a primitive element of the extension field. Then, we can define:

$$\phi \circ U(\overline{x}) = (x_1+x_2+x_5)+(x_1+x_2+x_4)\beta+(x_2+x_3+x_4)\beta^2+(x_1+x_4+x_5)\beta^4+(x_1+x_2+x_4+x_5)\beta^5$$

Now that we have mixed our input values with our affine map U and found the extension field representation of this information, we can evaluate our central map f at this element.

$$
\begin{aligned}
f(\phi(U(\overline{x}))) =& \Big(\phi(U(x))\Big)^{2^3+1} \\
=& \Big(\phi(Ux)\Big)^{2^3} \Big(\phi(U(x))\Big) \\
=& \Big((x_1+x_2+x_5)+(x_1+x_2+x_4)\beta+(x_2+x_3+x_4)\beta^2+(x_1+x_4+x_5)\beta^4 \\
&+(x_1+x_2+x_4+x_5)\beta^5\Big)^{2^3}\Big((x_1+x_2+x_5)+(x_1+x_2+x_4)\beta \\
&+(x_2+x_3+x_4)\beta^2+(x_1+x_4+x_5)\beta^4+(x_1+x_2+x_4+x_5)\beta^5\Big) \\
=& \Big((x_1+x_2+x_5)+(x_1+x_2+x_4)\beta^{2^3}+(x_2+x_3+x_4)\beta^{2\cdot 2^3}+(x_1+x_4+x_5)\beta^{4\cdot 2^3} \\
&+(x_1+x_2+x_4+x_5)\beta^{5\cdot 2^3}\Big)\Big((x_1+x_2+x_5)+(x_1+x_2+x_4)\beta \\
&+(x_2+x_3+x_4)\beta^2+(x_1+x_4+x_5)\beta^4+(x_1+x_2+x_4+x_5)\beta^5\Big) \\
=& (x_4x_5)+\beta(x_1x_3+x_2x_4)+\beta^6(x_1x_3)+\beta^7(x_1x_2+x_4x_5)+\beta^8(x_3x_4+x_5) \\
&+\beta^9(x_3x_4+x_2x_5)+\beta^{10}(x_3x_8)+\beta^{11}(x_2x_3)+\beta^{15}(x_2x_4)+\beta^{16}(x_2x_5+x_3x_5) \\
&+\beta^{17}(x_2)+\beta^{18}x_3+\beta^{20}x_1+\beta^{22}(x_1x_4)+\beta^{24}(x_2x_3)+\beta^{28}(x_1x_4)+\beta^{29}(x_1x_5) \\
&+\beta^{30}(x_1x_2+x_4+x_1x_5)
\end{aligned}
$$

The next step will be to find a base field representation of this element in the extension field. In order to use ϕ^{-1}, we require all of the β terms to be expressed with degrees less than 5. Because β is a primitive element, then we know that $\beta^5+\beta^2+1=0$ (from our definition of $\mathbb{K}$). This gives us the relationship that $\beta^5=\beta^2+1$. We can use this information to find other relations, for example, $\beta^6=\beta^3+\beta$. We can also use field equations to simplify the expressions. Because $\mathbb{F}_2=GF(2)$ we know that $x=-x, 2x=0$, and $x^j=x$ for each $x\in\mathbb{F}$. Using this information we are able to rewrite the output of f shown above to get the following.

$$
\begin{aligned}
f\Big(\phi(U(x))\Big) =& x_4x_5 + x_3x_4 + x_5 + x_2x_3 + x_2x_4 + x_2x_5 + x_2 + x_3 + x_1x_4 + x_1x_5 \\
&+ \beta(x_3x_4 + x_3x_5 + x_2 + x_3 + x_1x_4 + x_1x_2 + x_4 + x_1x_5) \\
&+ \beta^2(x_1x_2 + x_4x_5 + x_3x_4 + x_5 + x_2x_4 + x_1) \\
&+ \beta^3(x_1x_3 + x_5 + x_2x_4 + x_3x_5 + x_1 + x_2x_3 + x_1x_5) \\
&+ \beta^4(x_4x_5 + x_3x_4 + x_2x_4 + x_2 + x_2x_3 + x_4 + x_1x_5)
\end{aligned}
$$

The quantity $\phi^{-1}(f(\phi(U\overline{x})))$ is

$$
\begin{bmatrix}
x_4x_5 + x_3x_4 + x_5 + x_2x_3 + x_2x_4 + x_2x_5 + x_2 + x_3 + x_1x_4 + x_1x_5 \\
x_3x_4 + x_3x_5 + x_2 + x_3 + x_1x_4 + x_1x_2 + x_4 + x_1x_5 \\
x_1x_2 + x_4x_5 + x_3x_4 + x_5 + x_2x_4 + x_1 \\
x_1x_3 + x_5 + x_2x_4 + x_3x_5 + x_1 + x_2x_3 + x_1x_5 \\
x_4x_5 + x_3x_4 + x_2x_4 + x_2 + x_2x_3 + x_4 + x_1x_5
\end{bmatrix}
$$

We compose this with our map T to get the public key:

$$
P(\overline{x}) = \begin{bmatrix}
x_1x_3 + x_2x_3 + x_2x_4 + x_1x_5 + x_3x_5 + x_1 + x_5 \\
x_2x_3 + x_2x_4 + x_3x_4 + x_1x_5 + x_4x_5 + x_2 + x_4 \\
x_2x_3 + x_1x_4 + x_3x_4 + x_3x_5 + x_1 + x_3 + x_5 \\
x_1x_2 + x_1x_3 + x_1x_4 + x_1x_5 + x_4x_5 + x_1 + x_3 + x_5 \\
x_1x_3 + x_2x_3 + x_1x_4 + x_2x_4 + x_1x_5 + x_2x_5 + x_3x_5 + x_1 + x_3 + x_4
\end{bmatrix}
$$

To encrypt To encrypt a vector $\overline{x}$, you will evaluate P at $\overline{x}$. For a specific example, lets say that $\overline{x} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \end{bmatrix}^\top$. Then $P(\overline{x}) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^\top$.

To decrypt Let $P(\overline{x}) = \overline{y}$. Then, given $\overline{y}$, you will compute

$$U^{-1} \circ \phi^{-1} \circ f^{-1} \circ \phi \circ T^{-1}\overline{y}$$

to get $\overline{x}$.

2.2.2 Break of C^*

C^* was broken by Patarin in 1995 in [2] using linearization equations. The goal of the attack is to use the known quadratic equations to discover linear relationships between the plaintext and cipher text variables.

The first step of the attack is to denote $v = u^{q^\theta+1}$. Notice, $v, u \in \mathbb{K}$, so we can consider $u = \phi(x_1, \ldots, x_n)$ and $v = \phi(y_1, \ldots, y_n)$. Once this relationship is established, the next step is to raise both sides to the $q^\theta - 1$ power, which results in $v^{q^\theta-1} = u^{q^{2\theta}-1}$. The final step in creating the linearization equations is then to multiply both sides of the equation by uv, giving us $uv^{q^\theta} = u^{q^{2\theta}}v$. This equation is $\mathbb{F}_q$-linear in both plain text and cipher text variables, which renders the scheme no longer secure.

2.2.3 PFLASH and Other C^* Variants

After the break of C^*, modifiers were introduced in [3] in the hopes of regaining security. One such modifier is the minus modifier, which eliminates r equations from the public key. Revisiting the example from Section 2.2.1, if we were to apply a minus modifier, we may only publish 4 of the 5 public key equations (in this case, r would be equal to 1). C^* schemes that have a minus modifier applied are called C^{*-} schemes. One example of a C^{*-} scheme is SFLASH. SFLASH is a particular parameterization of a C^{*-} scheme which was thought to be secure and was recommended by the NESSIE consortium for smart card use. SFLASH was broken in [4] by employing the discrete differential, which will be discussed in Section 3.2.

Another modifier introduced is the projection modifier. The idea of projection is to fix the value of $d - n$ input variables to change the simplicity of the central map. PFLASH, see [5], is a specific parameterization of a projected C^{*-} scheme, which is used as a digital signature primitive. If we consider the projection

modifier $\pi_d : \mathbb{F}_n^d \to \mathbb{F}_q^d$ and the minus modifier $\pi_r : \mathbb{F}_q^d \to \mathbb{F}_q^{d-r}$, then the public key of PFLASH is given by $P = \pi_r \circ T \circ \phi^{-1} \circ f \circ \phi \circ U \circ \pi_d$.

It is important to note that there is only a polynomial morphism between the central map and the public key, as opposed to the two being isomorphic. The morphism of polynomials problem is known to be NP-hard, as seen in [6], which gives hope that the information lost to the public key may secure the scheme.

Verification of a signature is accomplished by evaluating the public polynomials at the given signature. Signing is done by finding preimages of each of the private maps. To find a preimage of $\pi_r \circ T\phi^{-1}$, randomly append r values to the message, then apply T^{-1} and ϕ. Once f is inverted, an element in the preimage of $\phi \circ U$ and in the image of π_d is selected as the signature.

2.3 HFE

Hidden Field Equation (HFE) scheme (introduced in [7]) is a generalization of the C^* construction where the monomial map is replaced by a more general polynomial with a degree bound D. Given the base field $\mathbb{F}_q$ and the degree n extension $\mathbb{K}$, we choose a quadratic polynomial $f : \mathbb{K} \to \mathbb{K}$ of degree bound D. Thus f has the form:

$$f(x) = \sum_{\substack{i \leq j \\ q^i + q^j \leq D}} \alpha_{i,j} x^{q^i + q^j} + \sum_{\substack{i \\ q^i \leq D}} \beta_i x^{q^i} + \gamma,$$

where $\alpha_{i,j}, \beta_i, \gamma \in \mathbb{K}$. The public key is then constructed via the isomorphism:

$$P = T \circ \phi^{-1} \circ f \circ \phi \circ U.$$

Inversion is accomplished by first taking a ciphertext $y = P(x)$, computing $v = T^{-1}(y)$, solving $v = f(u)$ for u via the Berlekamp algorithm, see [8], and then recovering $x = U^{-1}(u)$.

The HFE scheme was designed to be used as an encryption or a signature scheme. To generate a signature (or to decrypt), one computes, successively, $v = T^{-1}y$, $u = f^{-1}(v)$ and $x = U^{-1}u$. The vector x is the signature (or the plaintext). For verification (or encryption), one simply evaluates the public polynomials, P, at x. If $P(x)$ which is equal to $T \circ f \circ U(x)$ is equal to y, the signature is authenticated (or the ciphertext is y).

HFE was presented in the hopes of creating a scheme that is protected against Patarin's linearization equations. The scheme is safe from that attack, but HFE is vulnerable against MinRank and differential attacks, which we will discuss in a later chapter.

2.3.1 HFEv and Other Variants

The break of HFE lead to the application of modifiers in the hopes of constructing a secure adaptation of the scheme. Once again a minus modifier was proposed, which lead to the creation of HFE$^-$. This scheme seemed to hold promise, but a new key recovery attack for HFE$^-$ was proposed in [9].

When using HFE as a signature scheme, another possible modification is to add vinegar variables. The vinegar modifier adds extra variables $\widetilde{x_1}, \widetilde{x_2}, \ldots, \widetilde{x_v}$ into the public key, that can be assigned random values upon inversion. The effect of adding vinegar variables is that new quadratic terms, formed from both products of vinegar variables and HFE variables and products among vinegar variables, increase the rank of the public key. The central map of the HFEv scheme has the form

$$f(\mathbf{x}) = \sum_{\substack{i \leq j \\ q^i + q^j \leq D}} \alpha_{i,j} \mathbf{x}^{q^i + q^j} + \sum_{\substack{i \\ q^i \leq D}} \beta_i(\widetilde{x_1}, \widetilde{x_2}, \ldots, \widetilde{x_v}) \mathbf{x}^{q^i} + \gamma(\widetilde{x_1}, \widetilde{x_2}, \ldots, \widetilde{x_v})$$

where $\mathbf{x} \in \mathbb{K}$, $\alpha_{i,j} \in \mathbb{K}$, $\beta_i : \mathbb{F}_q^v \to \mathbb{K}$ is linear, and $\gamma : \mathbb{F}_q^v \to \mathbb{K}$ is quadratic.

The vector valued functions of HFEv map from $\mathbb{F}_q^{n+v} \to \mathbb{F}_q^n$, unline the vector valued functions of HFE which map from $\mathbb{F}_q^n \to \mathbb{F}_q^n$. But, as shown in [10–12], it can be useful to express these functions over the extension field $\mathbb{K}$.

One way to create the representation over $\mathbb{K}$ is by augmenting an additional $n - v$ elements to the input of f. We can consider the input variables $\hat{y}$ (vinegar variables) and $\hat{x}$ (HFE variables) and $\hat{f}$, which is now a bivariate function over $\mathbb{K}$. We may now write f in the following form:

$$f(x,y) = \sum_{\substack{i \leq j \\ q^i + q^j \leq D}} \alpha_{i,j} x^{q^i + q^j} + \sum_{\substack{i \\ q^i \leq D}} \beta_i x^{q^i} y^{q^j} + \sum_{0 \leq i \leq j < n} \gamma_{ij} y^{q^i + q^j}$$

Here we see an obvious distinction among the types of monomials.

The HFEv^- scheme applies a minus modifier to an HFEv scheme. Just as before, the minus modifier removes r of the public equations. This alteration is designed to destroy some of the information of the big field operations latent in the public key.

2.3.2 HFEv Toy Example

Creation of the public key For the purposes of this toy example, we will only consider the central map f, and we will let U and T be identity maps. We will choose $q = 2$, $\mathbb{F}_q = GF(2)$, $\mathbb{K} = \mathbb{F}_2 / < x^4 + x + 1 >, n = 4$, and $v = 3$. We will define our central map to be $f = x^{q^1+q^2} + xy + y^{q^1+q^2}$, and consider β a primitive element of the extension field. Using our information we compute:

$$\begin{aligned}
\hat{f}\begin{pmatrix}\hat{x}\\ \hat{y}\end{pmatrix} =&(x_1+\beta x_2+\beta^2x_3+\beta^3x_4)^{q^3}+(x_1+\beta x_2+\beta^2x_3+\beta^3x_4)(y_1+\beta y_2+\beta^2y_3)\\
&+(y_1+\beta y_2+\beta^2y_3)^{q^3}\\
=&x_1^3+\beta x_1^2x_2+\beta^2x_1x_2^2+\beta^3x_2^3+\beta^2x_1^2x_3+\beta^4x_2^2x_3+\beta^4x_1x_3^2+\beta^5x_2x_3^2+\beta^6x_3^3\\
&+\beta^3x_1^2x_4+\beta^5x_2^2x_4+\beta^7x_3^2x_4+\beta^6x_1x_4^2+\beta^7x_2x_4^2+\beta^8x_3x_4^2+\beta^9x_4^3+y_1^3\\
&+\beta y_1^2y_2+\beta^2y_1y_2^2+\beta^3y_2^3+\beta^2y_1^2y_3+\beta^4y_2^2y_3+\beta^4y_1y_3^2+\beta^5y_2y_3^2+\beta^6y_3^3\\
&+x_1y_1+\beta x_2y_1+\beta^2x_3y_1+\beta^3x_4y_1+\beta x_1y_2+\beta^2x_2y_2+\beta^3x_3y_2+\beta^4x_4y_2\\
&+\beta^2x_1y_3+\beta^3x_2y_3+\beta^4x_3y_3+\beta^5x_4y_3\\
=&(x_1+y_1+x_1y_1+x_3x_4+x_2x_4+x_3x_4+x_2x_3+x_1x_3+y_2y_3+y_1y_3+x_4y_2+x_3y_3)\\
&+\beta(x_1x_2+x_2y_1+x_1y_2+x_2x_4+x_4y_3+x_3x_4+x_2x_4+x_4+x_1x_3+x_4y_2+x_3y_3)\\
&+\beta^2(x_1x_2+x_1x_3+y_1y_2+y_1y_3+x_3y_1+x_2y_2+x_1y_3+x_2x_3+x_2x_4+y_2y_3\\
&+x_4y_3+x_3+x_1x_4+y_3+x_3x_4)+\beta^3(x_2+x_1x_4+y_2+x_4y_1+x_3y_2+x_2y_3+x_3\\
&+x_1x_4+y_3+x_3x_4+x_2x_4+x_4)
\end{aligned}$$

We will now denote the vinegar variables y_1, y_2, and y_3 as x_5, x_6, and x_7, respectively. This gives us our public key $P(\overline{x})$:

$$\begin{bmatrix}
x_1+x_5+x_1x_5+x_3x_4+x_2x_4+x_3x_4+x_2x_3+x_1x_3+x_6x_7+x_5x_7+x_4x_6+x_3x_7\\
x_1x_2+x_2x_5+x_1x_6+x_2x_4+x_4x_7+x_3x_4+x_2x_4+x_4+x_1x_3+x_4x_6+x_3x_7\\
x_1x_2+x_1x_3+x_5x_6+x_5x_7+x_3x_5+x_2x_6+x_1x_7+x_2x_3+x_2x_4+x_6x_7+x_4x_7+x_3+x_1x_4+x_7+x_3x_4\\
x_2+x_1x_4+x_6+x_4x_5+x_3x_6+x_2x_7+x_3+x_1x_4+x_7+x_3x_4+x_2x_4+x_4
\end{bmatrix}$$

Signature Creation Alice wants to send Bob an important message, but Bob needs a way to ensure the message he receives is from Alice. Alice creates her public key P, and they agree upon a signature. Bob states he wants Alice's signature to be x where x is the preimage of $y = \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix}$

Alice will then apply random values to the vinegar variables x_5, x_6, x_7. She will then use her information regarding the creation of the public key to invert f, which is now quadratic in $\widehat{x}$.

Alice finds that $x = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$ is a preimage of $y = \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix}$, and signs her message accordingly.

Signature Verification Bob plugs $x = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$ in to P and verifies that $P(x) = \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix} = y$.

CHAPTER 3
ATTACKS

There are three major classes of attacks that have proved effective against big field schemes. These classes of attacks include differential techniques, MinRank attacks, and algebraic attacks. All of these attacks are related in some way to the Q-rank of a scheme.

Definition 2. *Given a central map f, we choose a matrix representation $\mathbf{F}$ of the central map f such that*

$$f(X) = \begin{bmatrix} X & X^q & \dots & X^{q^{d-1}} \end{bmatrix} \mathbf{F} \begin{bmatrix} X & X^q & \dots & X^{q^{d-1}} \end{bmatrix}^\top$$

*We call the rank of the quadratic form $\mathbf{F}$, the Q-*rank *of f*

The MinRank key recovery attack has a complexity directly dependent on the Q-rank of the central map. The differential symmetry attack is relevant when the Q-rank of the central map is minimal in the relevant algebra. The direct algebraic attack has a complexity dependent on the degree of regularity of the public key which is usually a linear function of the Q-rank. We review each of these techniques. The subsequent chapters will focus heavily on analyzing the security of different multivariate schemes against these attacks

3.1 Direct Algebraic Attack

The most straightforward attack of a multivariate cryptosystem is to try to directly invert the public key via Gröbner bases. In this section we will define

Gröbner bases and outline the process of such an attack.

3.1.1 Gröbner Basis

Definition 3. *Let $G = \{g_1, \ldots, g_\ell\} \subset \mathbb{F}[X] = \mathbb{F}[X_1, \ldots, X_m]$ be a finite set of polynomials in m variables over a field $\mathbb{F}$; and let I be the ideal of $\mathbb{F}[X]$ that they generate. We say that G is a* Gröbner basis *for the ideal I if every nonzero $f \in I$ has a leading term that is divisible by the leading term of at least one of the g_i. (Page 74, [13])*

Informally, a Gröbner basis is a special collection of functions that generate an ideal, I. It is important to note that this differs from our normal idea of a basis in that the linear combinations of the functions to form elements of our ideal I are not unique.

Many resources have been devoted to creating algorithms to efficiently find Gröbner bases. Buchberger created the first algorithm to compute the Gröbner basis of a set of polynomials in [14]. Later, Faugére created the F4 algorithm, published in [15], which greatly improves the efficiency of finding a Gröbner basis.

The first step to compute the Gröbner basis of an ideal, is to decide on a term ordering. You will get different solutions depending on the chosen order, so it is important to be consistent. Common orderings include Lexicographical ordering, Degree-Lexicographical ordering, and Degree-Reverse-Lexicographical ordering.

The process of finding a Gröbner basis requires the use of the S-polynomial.

Definition 4. *The* S-polynomial *of two nonzero polynomials $f, g \in \mathbb{F}[X_1, \ldots, X_m]$ is*

$$S(f,g) = \frac{L}{lt(f)}f - \frac{L}{lt(g)}g$$

where L denotes the least common multiple of the leading terms of f and g. That is, the power product of lowest total degree that is divisible by both $lt(f)$ and $lt(g)$

Now we will consider the following process, which follows Buchberger's algorithm. Let $I \subset \mathbb{F}[X_1, \ldots, X_m]$ be the ideal generated by the set $F = \{f_1, f_2, \ldots, f_{\ell'}\}$. To create the Gröbner basis, one would reduce the S-polynomial $S(f_i, f_j)$ modulo F until a polynomial h_{ij} is obtained that either is 0 or has leading term that cannot be reduced (for all $1 \leq i < j \leq \ell'$). In the case that h_{ij} is nonzero, h_{ij} is added to the set F. This process continues, adding $f_{\ell'+1}, f_{\ell'+2}, \ldots$ to the set F, until you have a set $G = \{f_1, \ldots, f_\ell\}$ such that $S(f_i, f_j)$ reduces to 0 modulo G for all $1 \leq i < j \leq \ell$. This process will give you a Gröbner basis of I and will terminate in finitely many steps.

Definition 5. *A Gröbner basis $\{g_1, \ldots, g_\ell\}$ of an ideal $I \subset \mathbb{F}[X_1, \ldots, X_m]$ is said to be* minimal *if all of the g_i are monic and if the leading term of g_i does not divide the leading term of g_j for $i \neq j$, with $i, j = 1, \ldots, l$. (Page 77, [13])*

Once we have a minimal Gröbner basis, we are able to derive a reduced Gröbner basis. It is shown in [13] that every ideal of $\mathbb{F}[X_1, \ldots, X_m]$ has a unique reduced Gröbner basis.

Definition 6. *A Gröbner basis $\{g_1, \ldots, g_\ell\}$ of an ideal $I \subset \mathbb{F}[X_1, \ldots, X_m]$ is said to be* reduced *if all of the g_i are monic and if none of the terms of g_i is divisible by the leading term of g_j where $j \neq i$. (Page 78, [13]).*

Reduced Gröbner bases are unique for any given ideal and monomial ordering. Thus, two ideals are equal iff they have the same reduced Gröbner basis.

3.1.2 Attacks Using Gröbner Basis

As discussed in the previous subsection, the Gröbner basis of a set of polynomials is a set of functions that generate an ideal. We will consider the set of polynomials

$$p_1 = y_1, p_2 = y_2, \ldots, p_n = y_n$$

where p_i represents the i^{th} public polynomial and y_i represents the i^{th} coordinate of a ciphertext. If an adversary has intercepted a secret ciphertext $[y_1, y_2, \ldots, y_n]$ then the adversary may try to directly solve the set of polynomials using Gröbner basis techniques.

Let $F \subset \mathbb{F}[X_1, \ldots, X_n]$, and denote $V(F) = \{\mathbf{x} \in \mathbb{F}^n : f(\mathbf{x}) = 0, \forall f \in F\}$ as the variety of the set F. Recall the set of polynomials $p_1 = y_1, p_2 = y_2, \ldots, p_n = y_n$, consider the equivalent set $p_1 - y_1 = 0, p_2 - y_2 = 0, \ldots, p_n - y_n = 0$. We will use the following Lemma to establish a link between $p_1 - y_1 = 0, p_2 - y_2 = 0, \ldots, p_n - y_n = 0$ and the Gröbner basis it will produce.

Lemma 1. *Consider two sets of polynomials $F, G \subset \mathbb{F}[X_1, \ldots, X_n]$. If F and G generate the same ideal, then $V(F) = V(G)$.*

Proof. Let I be an ideal generated by F.

1. Show $V(F) \subseteq V(I)$

   Let $\mathbf{x} \in V(F)$. Thus, $f(\mathbf{x}) = 0$ for all $f \in F$. Consider $\hat{f} \in I$. Then, by the definition of I, $\hat{f} = \sum p_i f_i$, where $f_i \in F$. So we see, $\hat{f}(\mathbf{x}) = \sum p_i(\mathbf{x}) f_i(\mathbf{x}) = 0$, because $f_i(\mathbf{x}) = 0$ for all $f_i \in F$. Thus, for all $\mathbf{x} \in V(F)$ we know $\mathbf{x} \in V(I)$. Therefore, $V(F) \subseteq V(I)$

2. Show $V(I) \subseteq V(F)$

   Recall, $F \subseteq I$. Let $\mathbf{x} \in V(I)$. Thus $f(\mathbf{x}) = 0$ for all $f \in I$. Let $\hat{f} \in F$. Then, $\hat{f} \in I$. Thus, $\hat{f}(\mathbf{x}) = 0$. Therefore, $V(I) \subseteq V(F)$.

3. Consider G.

   Let G be a generating set of I. As shown above, $V(I) = V(F)$. By similar arguments, we can show $V(G) = V(I)$. Therefore $V(F) = V(G)$.

□

So if we start with the polynomial system $p_i - y_i = 0$ for all $1 \leq i \leq n$, then we can find a Gröbner basis for this system of polynomials. As we have found, the variety of the Gröbner basis will be the same as the variety of the original of public polynomials. This is advantageous for an adversary because it is known how to find the variety of a Gröbner basis. The first coordinate of $\mathbf{x} \in V(G)$ will be a root of the greatest common divisor of polynomials of the basis that depends only of the first variable. After substituting in the first coordinate of $\mathbf{x}$, the second coordinate will be a root of the greatest common divisor of the resulting polynomials that depends only on this second variable, and so on, until all of the coordinates of $\mathbf{x}$ have been found. Thus,

$$P(\mathbf{x}) - \mathbf{y} = 0 \Longrightarrow P(\mathbf{x}) = \mathbf{y}.$$

The complexity of solving such systems relies on the degree of regularity of the system, which can be defined as the smallest degree at which a nontrivial syzygy producing a degree fall is generated in the Gröbner basis algorithm.

3.2 Differential Techniques

Differential attacks make use of the discrete differential of a function. The discrete differential of a function $f : \mathbb{K} \to \mathbb{K}$ is the bivariate function

$$Df(a, x) = f(a + x) - f(a) - f(x) + f(0).$$

The discrete differential is similar to the derivative of a function in the sense that it depresses the degree of a function, but the discrete differential will also introduce another variable. For example, if you have a cubic function f, then the discrete differential of f, Df, will be a bi-quadratic function.

Pataran's linearization equations in [2] can be viewed as a differential attack. Notice that the discrete differential of the C^* monomial $f(x) = x^{q^\theta+1}$ is $Df(a, x) = ax^{q^\theta} + xa^{q^\theta}$. If $a = x$ then we would get $Df(x, x) = 2x^{q^\theta+1}$, which is equal to zero

in characteristic two. Now consider setting $v = f(u)$ and compute the following:

$$\begin{aligned} 0 = Df(v, f(u)) &= \big(f(u)\big)v^{q^\theta} + v\big(f(u)\big)^{q^\theta} \\ &= u^{q^\theta+1}v^{q^\theta} + vu^{q^{2\theta}+q^\theta} \\ &= u^{q^\theta}(uv^{q^\theta} + vu^{q^{2\theta}}) \end{aligned}$$

Notice, we now have a bilinear relationship between u (plain-text variables) and v (cipher-text variables) over $\mathbb{F}_q$.

Moving beyond the direct application of the differential, discrete differentials are the foundation of differential symmetry and differential invariant attacks. It is useful to note that while Df is a function over $\mathbb{F}_q^n$, we can define Df as a matrix on $\mathbb{K}$. Notice that we can express the i^{th} coordinate of Df, notated as $[Df(y, x)]_i$, in the following way:

$$[Df(y, x)]_i = y^\top \mathbf{Df}_i x$$

where $\mathbf{Df}$ is the matrix form of Df. If we consider $f(x)$ to be the C^* polynomial, then $\mathbf{Df}$ is the $n \times n$ matrix with 1's in the $(0, \theta)$ and $(\theta, 0)$ coordinates, and 0's everywhere else. This notation will help us in later analysis.

3.2.1 Differential Symmetry

Linear differential symmetry attacks attempt to find linear maps L that "factor through" the differential of the central map in an interesting way. Specifically, the goal is to find maps M satisfying

$$Df(Ma, x) + Df(a, Mx) = \Lambda_M Df(a, x) \tag{3.1}$$

If such a map can be found, it allows one to "remove" a minus modifier by discovering new linear combinations of the central maps that are linearly independent of the public key.

This technique was used in [4] to fatally attack SFLASH. We recall the discrete differential of a function f is defined as $Df(a,x) = f(a+x) - f(a) - f(x) + f(0)$. When $f(x) = x^{q^\theta+1}$ then $Df(a,x) = ax^{q\theta} + a^{q\theta}x$. If we symmetrically apply an element of the extension field to the inputs of the differential, we have $Df(\sigma a, x) = \sigma a x^{q\theta} + (\sigma a)^{q\theta}x$ and $Df(a,\sigma x) = a(\sigma x)^{q\theta} + a^{q\theta}x\sigma$, where $\sigma \in \mathbb{K}$. Thus for all $\sigma \in \mathbb{K}$, $Df(\sigma a, x) + Df(a, \sigma x) = (\sigma + \sigma^{q^\theta})Df(a,x)$.

Consider the public key of the C^* scheme given by $P = \hat{T} \circ \phi^{-1} \circ f\phi \circ \hat{U}$. For ease of notation (for this section only) we will let $T := \hat{T} \circ \phi^{-1}$ and $U := \phi \circ \hat{U}$, so we can denote $P = T \circ f \circ U$. We will now analyze the differential of the *public* equations P. This will be computed as $DP(a,x) = T \circ Df(U(a), U(x))$, where $DP : \mathbb{F}_q^n \to \mathbb{F}_q^n$.

Now consider symmetrically applying elements of $\mathbb{K}$ as we did before. We want to apply this relation to $U(a)$ and $U(x)$, so we define $M_\sigma(x) = \sigma U(x)$.

$$\begin{aligned} DP(M_\sigma(a), x) + DP(a, M_\sigma(x)) &= T \circ Df(\sigma U(a), U(x)) + T \circ Df(U(a), \sigma U(x)) \\ &= T \circ (\sigma + \sigma^{q^\theta})(Df(U(a), U(x)) \\ &= T \circ (\sigma + \sigma^{q^\theta})(T^{-1}(DP(a,x)) \end{aligned}$$

Now let DP_Π denote the differential of a C^{*-} scheme. We can consider $P_\Pi = \Pi \circ T \circ f \circ U$, where Π denotes the minus modifier.

$$\begin{aligned} DP_\Pi(M_\sigma(a), x) + DP_\Pi(a, M_\sigma(x)) &= \Pi \circ T \circ Df(\sigma U(a), U(x)) \\ &\quad + \Pi \circ T \circ Df(U(a), \sigma U(x)) \\ &= \Pi \circ T \circ (\sigma + \sigma^{q^\theta})(Df(U(a), U(x)) \\ &= \Pi \circ T \circ (\sigma + \sigma^{q^\theta})(T^{-1}(DP(a,x)) \end{aligned}$$

So now we have a relationship between the published C^{*-} key (left hand

side), and the breakable unmodified C^* public equations (right hand side). The left hand side is linear in the unknown coefficients M_σ, and the right and side is linear in the unknown coefficients of $T \circ (\sigma + \sigma^{q^\theta}) \circ T^{-1}$. To match our notation in Equation 3.1, $M = M_\sigma$, $\Lambda_M = T \circ (\sigma + \sigma^{q^\theta}) \circ T^{-1}$.

Once the new linearly independent combinations of the central maps have been discovered, the adversary may mount other attacks (such as Patarin's linearization equations) onto the recovered representation of the scheme and break it. This attack lead to the implementation of using a projection modifier in conjunction with a minus modifier, leading to schemes such as PFLASH (parameters defined in [5]).

3.2.2 Differential Invariants

Differential invariants can also be used to weaken the security of certain schemes. Informally, a function has a differential invariant, $V \subseteq \mathbb{K}$, if the image of V under all the coordinates of the matrix form of Df lies in a fixed subspace with the same or smaller dimension size. More formally we define the following:

Definition 7. *Let $f : \mathbb{F}_q^n \to \mathbb{F}_q^m$ be a function. A* differential invariant *of f is a subspace $V \subseteq \mathbb{K}$ with the property that there is a subspace $W \subseteq \mathbb{K}$ such that $dim(W) \leq dim(V)$ and $\forall A \in Span_{\mathbb{F}_q}(Dfi)$ it holds that $AV \subseteq W$.*

The point of searching for a differential invariant is to try to create a linear relationship between plaintext and ciphertext variables. If we can find a differential invariant V, then we can establish a linear relationship using the discrete differential. This explanation follows much of the analysis done in [12].

Assume V is a differential invariant of some function f. Then we define $V^\perp$ to be the set of elements $x \in \mathbb{K}$ such that the dot product $\langle x, A \rangle = 0$ for all $v \in V$ and $A \in Span(Df_i)$. In other words, we can consider $V^\perp$ to be the set of elements

orthogonal to AV. Once we have found V and defined $V^\perp$, we may then choose linear functions $M : \mathbb{K} \to V$ and $M^\perp : \mathbb{K} \to V^\perp$.

Expressing the differential as a matrix we find that

$$[Df(M^\perp y, Mx)]_i = (M^\perp y)^\top (\mathbf{Df}_i(Mx)).$$

Notice that $M^\perp y \in V^\perp$ and $\mathbf{Df}_i Mx \in AV$. So for $1 \leq i \leq n$ and $\forall x, y \in \mathbb{K}$ we have

$$[Df(M^\perp y, Mx)]_i = (M^\perp y)^\top (\mathbf{Df}_i(Mx)) = 0$$

Our next goal will be to rewrite M and $M^\perp$. Consider Proposition 1 from [12] which states the following:

> If A, B are two $m \times n$ matrices, then $rank(A) = rank(B)$ if and only if there exist nonsingular matrices C, D such that $A = CBD$.

Now consider M and $M^\perp$. Without loss of generality, assume that $rank(M^\perp) \leq rank(M)$. If $rank(M^\perp) = rank(M)$, then by the above proposition, there exist nonsingular matrices S, T such that $M^\perp = SMT$. If $rank(M^\perp) < rank(M)$, then compose M with singular matrix Y so that $rank(M^\perp) = rank(YM)$. Applying the above proposition to $M^\perp$ and YM, we know there exist nonsingular matrices S, T such that $M^\perp = S(YM)T = S'MT$ where S' is singular. Restating our above result, for all $x, y \in \mathbb{K}$

$$Df(SMTy, MTx) = 0.$$

Thus we have found a linear relationship between plaintext and ciphertext variables using the differential invariant V.

3.3 MinRank

The first effective attack on HFE was presented in [10] and is now commonly called the Kipnis-Shamir (KS) attack. Their idea is to express the central

polynomial f as a matrix in quadratic form $\mathbf{F}$ (recall Definition 2) over $\mathbb{K}$. As the reader easily notices, the degree bound on f implies that $\mathbf{F}$ has only a small block of nonzero values and thus has low rank. We call the rank of this quadratic form the Q-rank of f.

The attack in [10] exploits this low Q-rank property of HFE by first finding a formula for the public key over the extension field. The next step of the attack is to compute the matrix forms of all of the Frobenius powers of this map, and the final goal is to find a low rank linear combination of these matrices with coefficients chosen from $\mathbb{K}$. The attack can be effective, but all of the algebra takes place in $\mathbb{K}$ which can be cumbersome.

The KS attack was significantly improved for determined or slightly over-determined schemes in [11], where the authors introduce minors modeling. The modeling of the low rank property in the KS attack requires structures defined over $\mathbb{K}$, whereas the authors of [11] noticed that a $\mathbb{K}$-linear combination of the *public* quadratic forms defined over $\mathbb{F}_q$ also has low rank. Thus one may construct a system of equations over the small field, resolve this system via Gröbner bases over the small field, and finally recover the variety over the big field. This requires the most intensive calculations to be performed over the base field, providing a significant advantage.

CHAPTER 4
ON THE DIFFERENTIAL SECURITY OF THE HFEv$^-$ PRIMITIVE

This chapter will analyze the security of HFEv^- schemes against differential attacks. Recall from Section 2.3.1 the definition of an HFEv^- scheme. HFEv^- is a big-field scheme where the central map is of the form

$$f(x,y) = \sum_{\substack{i\leq j \\ q^i+q^j\leq D}} \alpha_{i,j}x^{q^i+q^j} + \sum_{\substack{i \\ q^i\leq D}} \beta_i x^{q^i}y^{q^j} + \sum_{0\leq i\leq j<n} \gamma_{ij}y^{q^i+q^j} \tag{4.1}$$

and r of the public key equations are deleted. Much of this work has been completed by considering the security of an HFEv scheme, analyzed in [12], and then extending the analysis to the HFEv^- scheme.

4.1 Linear Symmetry

Recall from Section 3.2.1 that a general linear differential symmetry is a relation of the form

$$Df(Mx,a) + Df(x,Ma) = \Lambda_M Df(a,x).$$

where $M, \Lambda_M : \mathbb{K} \to \mathbb{K}$ are $\mathbb{F}_q-$linear maps.

While attacks similar to that of [10,16] exploited some multiplicative relation on central maps of schemes with some algebraic structure over the base field, it was shown in [17] that general linear differential symmetries based on more complex relations exist, in general. Therefore, when analyzing the potential threat, it becomes necessary to classify the possible linear differential symmetries. If we succeed

in characterizing parameters which provably eliminate nontrivial differential symmetric relations, we prove security against the entire class of differential symmetric attacks, even those utilizing relations not yet discovered. To this end, we evaluate the security of HFEv against such adversaries. We explicitly consider parameter restrictions that guarantee the existence of only trivial differential symmetries.

4.1.1 Linear Symmetry for HFEv

In our analysis, we will begin by considering the differential of our core map. The discrete differential is

$$D\hat{f}\Big(\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}, \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} \Big) = Df(a,b,x,y)$$

By the bilinearity of $D\hat{f}$ we see that Df is multi-affine, meaning Df is affine in each of its inputs when the remaining inputs are fixed. Evaluating this differential we obtain

$$\begin{aligned} Df(a,b,x,y) = &\sum_{\substack{0\le i\le j<n \\ q^i+q^j\le D}} \alpha_{i,j}(x^{q^i}a^{q^j} + x^{q^j}a^{q^i}) \\ &+ \sum_{\substack{0\le i\le j<n \\ q^i\le D}} \beta_{i,j}(x^{q^i}b^{q^j} + y^{q^j}a^{q^i}) \\ &+ \sum_{0\le i\le j<n} \gamma_{i,j}(y^{q^i}b^{q^j} + y^{q^j}b^{q^i}) \end{aligned} \tag{4.2}$$

noting that Df is a $\mathbb{K}$-bilinear form. For ease of computation, we will choose the following representation for $\mathbb{K}$:

$$x \mapsto \begin{bmatrix} x & x^q & x^{q^2} & \cdots & x^{q^{n-1}} \end{bmatrix}^\top$$

Similarly, we may map our oil-vinegar vector as

$$x \mapsto \begin{bmatrix} x & x^q & x^{q^2} & \cdots & x^{q^{n-1}}y & y^q & y^{q^2} & \cdots & y^{q^{n-1}} \end{bmatrix}^\top$$

and Df is thus represented by the $2n \times 2n$ matrix where the $(i,j)^{th}$ and $(j,i)^{th}$ entries in the upper left $n \times n$ block are the coefficients $\alpha_{i,j}$, and the $(i,j)^{th}$ entries in the upper right block and the $(j,i)^{th}$ entries in the lower left block are the coefficients $\beta_{i,j}$, while the $(i,j)^{th}$ and the $(j,i)^{th}$ entries in the lower right block are the coefficients $\gamma_{i,j}$. Note, that any $\mathbb{F}_q$-linear map $M : \mathbb{K} \to \mathbb{K}$ can be represented by $Mx = \sum_{i=0}^{n-1} m_i x$. Thus, as demonstrated in [12], under our representation we can model M as,

$$M = \begin{pmatrix} m_0 & m_1 & \cdots & m_{n-1} \\ m_{n-1}^q & m_0^q & \cdots & m_{n-2}^q \\ \vdots & \vdots & \ddots & \vdots \\ m_1^{q^{n-1}} & m_2^{q^{n-1}} & \cdots & m_0^{q^{n-1}} \end{pmatrix}.$$

However, when viewing an $\mathbb{F}_q$-linear map over our vector $\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}$, we may consider the $2n \times 2n$ matrix

$$\overline{M} = \begin{pmatrix} m_{00,0} & m_{00,1} & \cdots & m_{00,n-1} & m_{01,0} & m_{01,1} & \cdots & m_{01,n-1} \\ m_{00,n-1}^q & m_{00,0}^q & \cdots & m_{00,n-2}^q & m_{01,n-1}^q & m_{01,0}^q & \cdots & m_{01,n-2}^q \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{00,1}^{q^{n-1}} & m_{00,1}^{q^{n-1}} & \cdots & m_{00,0}^{q^{n-1}} & m_{01,1}^{q^{n-1}} & m_{01,2}^{q^{n-1}} & \cdots & m_{01,0}^{q^{n-1}} \\ m_{10,0} & m_{10,1} & \cdots & m_{10,n-1} & m_{11,0} & m_{11,1} & \cdots & m_{11,n-1} \\ m_{10,n-1}^q & m_{10,0}^q & \cdots & m_{10,n-2}^q & m_{11,n-1}^q & m_{11,0}^q & \cdots & m_{11,n-2}^q \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{10,1}^{q^{n-1}} & m_{10,1}^{q^{n-1}} & \cdots & m_{10,0}^{q^{n-1}} & m_{11,1}^{q^{n-1}} & m_{11,2}^{q^{n-1}} & \cdots & m_{11,0}^{q^{n-1}} \end{pmatrix}$$

For computational reference, we will label each row and column $modulo(n)$, i.e., each coordinate of the entry (i,j), will be represented by a residue class modulo n. If we assume that f is vulnerable to a differential attack, then there exists a

nontrivial linear mapping $\overline{M}$ such that the differential symmetry is satisfied. Computing such a symmetry inducing map requires the solution of $4n^2$ highly dependent, but random, equations in the $8n$ unknown coefficients of $\overline{M}$ and $\Lambda_{\overline{M}}$ over $\mathbb{K}$. Since trivial symmetries (such as multiplication by scalars) are exhibited by every map, we know that there exist nontrivial solutions. Even assuming unit time for $\mathbb{K}$-arithmetic operations, for realistic parameters this process is very inefficient; with the more realistic assumption of costly $\mathbb{K}$-arithmetic operations, this task is unsatisfactory in key generation.

To make the solution of such systems of equations more efficient, we derive the structure of the equations and develop a two step process for verifying trivial differential symmetric structure. The first step involves finding equations which only involve a subset of the variables. The existence of such equations is guaranteed by the degree bound of the HFE monomials. This information is then bootstrapped to eliminate many unknown coefficients of $\overline{M}$ resulting in a very small system of equations which can be solved explicitly. We remark here that this methodology also suggests a method for estimating the probability of the existence of a differential symmetry for the HFEv primitive. The existence of a nontrivial symmetry corresponds to systems for which the rank of the system of equations is less than $8n$. Under the heuristic that under row reduction these systems of equations behave like random $8n \times 8n$ matrices, we obtain a probability of roughly $1 - q^{-1}$ that the scheme has no nontrivial differential symmetry.We note that this heuristic is almost certainly false since trivial symmetries do exist. This quantity does represent a lower bound, however, and thus may offer support for larger base fields.

We begin by considering the entries of the matrix $\overline{M}^{\top} Df + Df\overline{M}$. The contribution of any monomial $\alpha_{i,j}x^{q^i+q^j}$ to the i^{th} row of $Df\overline{M}$ is given by

$$\left(\begin{array}{cccccccc} \alpha_{i,j}m^j_{00,-j} & \alpha_{i,j}m^j_{00,1-j} & \cdots & \alpha_{i,j}m^j_{00,-1-j} & \alpha_{i,j}m^j_{01,-j} & \alpha_{i,j}m^j_{01,1-j} & \cdots & \alpha_{i,j}m^j_{01,-1-j} \end{array} \right)$$

while the contribution to the j^{th} row is

$$\begin{pmatrix} \alpha_{i,j}m^i_{00,-i} & \alpha_{i,j}m^i_{00,1-i} & \cdots & \alpha_{i,j}m^i_{00,-1-i} & \alpha_{i,j}m^i_{01,-i} & \alpha_{i,j}m^i_{01,1-i} & \cdots & \alpha_{i,j}m^i_{01,-1-i} \end{pmatrix}$$

By symmetry, the i^{th} and and j^{th} columns of $\overline{M}^\top Df$ Df are the same as their respective rows.

It is clear that the rows and columns associated with coefficients of vinegar monomials as well as terms associated with mixing monomials may be represented similarly. However, it should be noted that those terms associated with mixing monomials will be multiplied by linear coefficients $m_{00,*}, m_{01,*}, m_{10,*}$, and $m_{11,*}$, while coeffcients associated with vinegar variables are multiplied only by the linear coeffcients $m_{10,*}$ and $m_{11,*}$.

The above patterns can be extended to characterize the contribution to the i^{th} row and j^{th} row of monomials of the form $\beta_{i,j}x^{q^i}y^{q^j}$ and $\gamma_{i,j}y^{q^i+q^j}$ as well. We note, however, that γ coefficients interact with entries from the lower block matrices while β coefficients interact with coefficients from all block matrices.

Now that we have characterized the left side of the central map described in Equation 4.1, we will consider the entries of $\Lambda_{\overline{M}}Df$. For every monomial of f, say $\alpha_{i',j'}x^{q^i+q^j}, \beta_{r,s}x^{q^r}y^{q^s}$, or $\gamma_{u,v}y^{q^s+q^v}$, under the mapping of $\Lambda_{\overline{M}}$ we have terms of the form: $l_\ell\alpha^{q^\ell}_{i,j}x^{q^{i+\ell}+q^{j+\ell}}$ $l_\ell\beta^{r+\ell}_{r,s}x^{q^{s+\ell}}y^{q^j}$, and $l_\ell\gamma^{q^\ell}_{u,v}y^{q^{u+\ell}+q^{v+\ell}}$. Clearly, this results in every nonzero entry, say (r,s), of our Df matrix being raised to the power of q^ℓ and shifted along a forty-five degree angle to entry $(r+\ell, s+\ell)$. Thus, for every monomial in f there are two possible nonzero entries in the i^{th} row, with possible overlap.

This discrete geometrical interpretation of the action of M and D on the coefficients of f is central to this analysis. A graphical representation of these relations is provided in Figure 4.1

As in [12], the possibility of a differential symmetry can be determined by

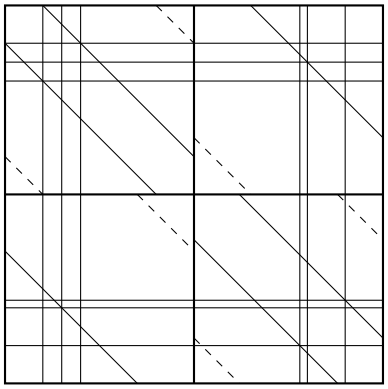


Figure 4.1: Graphical representation of the equation $M^\top Df + DfM = \Lambda_M Df$ for the HFEv polynomial $f(x) = \alpha_{i,j}x^{q^i+q^j} + \beta_{r,s}x^{q^r}y^{q^s} + \gamma_{u,v}y^{q^u+q^v}$. Horizontal and vertical lines represent nonzero entries in $M^\top Df + DfM$ while diagonal lines represent nonzero entries in $\Lambda_M Df$. We may consider this diagram as a genus 4 surface containing straight lines.

setting the matrix representation of $M^\top Df + DfM$ equal to the matrix $\Lambda_M Df$. We will demonstrate an algorithm, given some specific constraints, that will help provide secure keys to be generated automatically. Due to the structure of our M matrix, we need to work within each $m_{i,j}$ matrix independently. The following algorithm for $m_{0,0}$ extends very naturally to the other 3 matrices. For clarity, all m terms in the description below are $m_{0,0}$ terms.

Let $\alpha_{i,j}, \beta_{r,s}, \gamma_{u,v}$ represent the coefficients of our monomials in our core map. Consider the i^{th} row of $M^\top Df + DfM$. For all w not occurring as a power of q of our HFE or mixing monomials in f, or difference as a of powers of q in an exponent of a monomial in f plus i, the (i, w) entry is $\alpha_{i,j}m_{w-j}^{q^i} = 0$ (respectively $\beta_{i,j}m_{w-j}^{q^j}$). Consider the r^{th} row. For all w not occurring as an exponent of q in a vinegar monomial or as a difference of powers of q in an exponent of a monomial in f plus s, the $(r, w)^{th}$ entry is $\beta_{r,s}m_{k-s}^{q^s} = 0$. Hence, we can use those relations to look for non-zero entries of $m_{0,0}$.

After putting those relations into Algorithm 2 (listed in Appendix I), we can

generate a set of exponents that occur in the core map for every i and r. Each set provides a list of indices of all possible non-zero m's. For each index not occurring in any such set, the corresponding coefficient m must equal zero due to the fact that there must be a coordinate in the equation $M^\top Df + DfM = \Lambda_M Df$ setting a constant multiple of m to zero. Thus, the intersection of all sets generated produces a list of all possible non-zero entries for the sub-matrix $m_{0,0}$.

Once this list is obtained, the variables shown to have value zero are eliminated from the system of equations. After repeating a similar algorithm for each of the remaining three submatrices a significantly diminished system of equations is produced which is then solved explicitly. After running this algorithm with realistic values satisfying the above constraints and matching the parameter sizes of [53] along with using mild restrictions on the powers of the mixing and vinegar monomials, the only non-zero value obtained is m_0.

We note that it is possible that these restrictions, especially the restriction for these experiments on the number of monomials, place a lower bound on the number of vinegar variables required to achieve such a structure. On the other hand, with numerous small-scale experiments without parameter restrictions and using the full number of monomials we found that structurally the only nonzero value for the matrix $m_{0,0}$ is the m_0 term.

Since we have only a single non-zero term, our $m_{0,0}$ matrix is a diagonal matrix. A similar analysis for each of the remaining submatrices reveals the same structure. Thus we find that the only possible structure for M under these constraints satisfying a differential symmetry for HFEv is

$$\overline{M} = \left[\begin{array}{c|c} cI & dI \\ \hline dI & cI \end{array}\right]$$

Furthermore, we can prove by way of Theorem 2 from [18], that the coefficients $c, d \in \mathbb{F}_q$.

We note that this map induces a trivial differential symmetry. To see this, note that the (nonpartial) differential of any bivariate function is bilinear in its vector inputs. Thus

$$\begin{aligned} Dg(\overline{M}[a,b]^\top,[x,y]^\top) &= Dg([ca+db,da+cb]^\top,[x,y]^\top) \\ &= Dg([ca+db,cb+da]^\top,[x,y]^\top) \\ &= Dg(c[a,b]^\top,[x,y]^\top)+Dg(d[b,a]^\top,[x,y]^\top) \\ &= cDg(a,b,x,y)+dDg(b,a,x,y) \\ &= (c+d)Dg(a,b,x,y). \end{aligned} \tag{4.3}$$

Consequently, for the parameters provided by Algorithm 2 found in Appendix I HFEv provably has no nontrivial differential symmetric structure.

It should be noted that the restrictions provided on the powers of q of the monomials of our f does lower the entropy of our key space and likely raise the number of required vinegar variables to a level which is either unsafe or undesirable. However, there is still plenty of entropy with these restrictions and we obtain provable security against the differential symmetric attack. The restrictions provided are just a base line for this technique and our experiments with small scale examples indicate that even when we insist that every possible monomial satisfying the HFE degree bound is required to have a nonzero coefficient, the generalized algorithm still outputs only the trivial solution. Thus we can achieve provable security with minimal loss of entropy.

4.1.2 Linear Symmetry for HFEv^-

The algorithm extends naturally to HFEv^-. Every non-zero entry from the system generated by HFEv is also in the system generated by HFEv^-, but with a few additional entries, see Figure 4.2. We choose a basis in which an example minus projection is a polynomial of degree q^2. For every i^{th} row, we also know that for

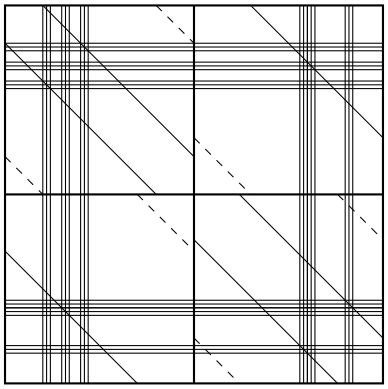


Figure 4.2: Graphical representation of the equation $M^\top Df + DfM = \Lambda_M Df$ for the HFEv with the minus modifier given by the projection $\pi(x) = x^{q^2} + \rho x^q + \tau x$. Horizontal and vertical lines represent nonzero entries in $M^\top Df + DfM$ while diagonal lines represent nonzero entries in $\Lambda_M Df$. We note that each triple of lines corresponds to a single monomial in the central map.

any w which is not a power of $\alpha + n$ or $\beta + n$ (where $n < 2$), the $(i, w)^{th}$ entry is $\alpha_{i,j} m_{w-j}^{q^j} = 0$. For the s^{th} row for all w not being a power of $\beta + n$ or $r + n$ where $n < 2$, the $(s, w)^{th}$ entry is $\beta_{r,s} m_{w-r}^{q^r} = 0$ A visualization is provided in Figure 4.2.

Again, we can use these relations, along with the relations described in the HFEv system, to create a list of sets of all non-zero areas on $m_{0,0}$ using Algorithm 3. Each of these sets contains indices which are possibly non-zero, thus entries not in that set are definitively equal to zero.

By taking the intersection of all the sets, you can find the final locations of non-zero entries for our sub matrix $m_{0,0}$. In doing so, with realistic values from [19], the only non-zero value obtained is m_0. This again gives us security against symmetrical attacks by having M being a block matrix consisting of diagonal matrices with an argument similar to [12].

4.2 Differential Invariants

$Df(SMTy, MTx) = 0$.

As discussed in Section 3.2.2, if we have a differential invariant V, then there exists a linear relationship between plain text and cipher text variables. To determine the effectiveness of a differential invariant attack, we evaluate the liklihood that a differential invariant exists.

Minimal Generators over Intermediate Subfield Consider the following statement about the structure of the coordinate ring of a subspace of an extension field over an intermediate extension.

Lemma 2. *Let $\mathbb{L}/\mathbb{K}/\mathbb{F}_q$ be a tower of finite extensions with $|\mathbb{L} : \mathbb{K}| = m$ and $|\mathbb{K} : \mathbb{F}_q| = n$. Let V be an $\mathbb{F}_q$-subspace of $\mathbb{L}$. Then $I(V)$ has m multivariate generators over $\mathbb{K}$ of the form*

$$\mathscr{M}_V^{(k)}(x_0, \ldots, x_{m-1}) = \sum_{\substack{0 \le i < n \\ 0 \le j < m}} \alpha_{ijk} x_j^{q^i}.$$

This Lemma is proven in [20], and will prove insightful during the following analysis. We note that the minimal polynomials studied in [12] correspond to the special case of the above lemma in which $m = 1$. Given our characterization from Section 2.3.1 of the central map of HFEv^- as a bivariate polynomial over $\mathbb{K}$, we are primarily interested in the $m = 2$ case of Lemma 2.

4.2.1 Invariant Analysis of HFEv

As in [12], we consider $Df(SMTa, MTx)$, where T is nonsingular, S is a possibly singular map which sends V into $V^\perp$ and $M : \mathbb{K} \to \mathbb{K}$ s a projection onto V. Without loss of generality we'll assume that M projects onto V. Then MT is another projection onto V. SMT is a projection onto $V^\perp$. An important distinction is that for this case, the a and x above are actually two dimensional vectors over $\mathbb{K}$.Thus $dim(V) + dim(V^\perp) \ge n$.

Theorem 1. *Let $\mathbb{K}$ be a degree n extension of the finite field $\mathbb{F}_q$. Let f be an HFEv central map. With high probability, f has no nontrivial differential invariant structure.*

Proof. We will denote the quantity $MT[x,y]^\top$ as $[\hat{x},\hat{y}]$. Suppose we have

$$f(x,y) = \sum_{\substack{0\le i\le j<n\\ q^i+q^j\le D}} \alpha_{ij}x^{q^i+q^j} + \sum_{\substack{0\le i\le j<n\\ q^i\le D}} \beta_{ij}x^{q^i}y^{q^j} + \sum_{0\le i\le j<n} \gamma_{ij}y^{q^i+q^j}.$$

Applying the differential with respect to the vector $[x,y]^\top$, we obtain Equation 4.2. Substituting $SMT[a,b]^\top$ and $MT[x,y]^\top$ into Equation 4.2 we derive:

$$Df(S[\hat{a},\hat{b}]^\top,\hat{x},\hat{y}) = Df(S_{11}\hat{a}+S_{12}\hat{b}, S_{21}\hat{a}+S_{22}\hat{b},\hat{x},\hat{y}).$$

For notational convenience, let $\widetilde{a} = S_{11}\hat{a}+S_{12}\hat{b}$ and $\widetilde{b} = S_{21}\hat{a}+S_{22}\hat{b}$. Plugging these values into the previous equation we get

$$\begin{aligned} Df(\widetilde{a},\widetilde{b},\hat{x},\hat{y}) = &\sum_{\substack{0\le i\le j<n\\ q^i+q^j\le D}} \alpha_{i,j}(\widetilde{a}^{q^i}\hat{x}^{q^j} + \widetilde{a}^{q^j}\hat{x}^{q^i})\\ &+ \sum_{\substack{0\le i\le j<n\\ q^i\le D}} \beta_{i,j}(\widetilde{a}^{q^i}\hat{y}^{q^j} + \hat{x}^{q^j}\widetilde{b}^{q^i})\\ &+ \sum_{0\le i\le j<n} \gamma_{i,j}(\widetilde{b}^{q^i}\hat{y}^{q^j} + \widetilde{b}^{q^j}\hat{y}^{q^i}) \end{aligned} \tag{4.4}$$

In contrast to the situation with HFE, these monomials are not necessarily independent. By Lemma 2, the generators of $I(V)$ have the form

$$\sum_{0\le i<n} r_{ij}x^{q^i} + \sum_{0\le i<n} s_{ij}y^{q^j} \text{ for } j\in\{1,2\},$$

where $r_{ij}, s_{ij}\in\mathbb{K}$. Clearly these expressions evaluate to zero on $(\hat{x},\hat{y})$. Evaluating 4.4 modulo $I(V)$ (only on the variables $\hat{x},\hat{y}$), we obtain:

$$\begin{aligned} Df(\widetilde{a},\widetilde{b},\hat{x},\hat{y}) = &\sum_{\substack{0\le i<n\\ 0\le j<d_x}} \left[\alpha'_{ij}(\widetilde{a})^{q^i} + \beta'_{ij}(\widetilde{b})^{q^i}\right]\hat{x}^{q^j}\\ &+ \sum_{\substack{0\le i<n\\ 0\le j<d_y}} \left[\gamma'_{ij}(\widetilde{a})^{q^i} + \delta'_{ij}(\widetilde{b})^{q^i}\right]\hat{y}^{q^j}, \end{aligned} \tag{4.5}$$

where d_x and d_y are the largest occurring powers of $\hat{x}$ and $\hat{y}$, respectively. After the reduction modulo $I(V)$, the remaining monomials $\hat{x}, \ldots, \hat{x}^{q^{d_x}}$ and $\hat{y}, \ldots, \hat{y}^{q^{d_y}}$ are independent. Thus, for $Df(\widetilde{a}, \widetilde{b}, \hat{x}, \hat{y}) = 0$, each polynomial expression multiplied by a single $\hat{x}^{q^j}$ or $\hat{y}^{q^j}$ must be identically zero, that is to say that for all $0 \leq j \leq d_x$

$$\sum_{0 \leq i < n} \left[\alpha'_{ij} (\widetilde{a})^{q^i} + \beta'_{ij} (\widetilde{b})^{q^i} \right] = 0 \tag{4.6}$$

and for all $0 \leq j \leq d_y$

$$\sum_{0 \leq i < n} \left[\gamma'_{ij} (\widetilde{a})^{q^i} + \delta'_{ij} (\widetilde{b})^{q^i} \right] = 0 \tag{4.7}$$

The left hand sides of 4.7 and 4.6 are $\mathbb{F}$-linear functions in $S[\hat{a}, \hat{b}]^\top$. Thus we can express each such equality over $\mathbb{F}$ as

$$LS\left[\hat{a}_0, \ldots, \hat{a}_{n-1}, \hat{b}_0, \ldots, \hat{b}_{n-1}\right]^\top = 0$$

where L is an $n \times 2n$ matrix with entries in $\mathbb{F}$. We note specifically that the coefficients of L depend on V and the choices of coefficients in the central map f. For randomly chosen coefficients retaining the HFEv structure, we expect an L derived from an equation of the form 4.7 or 4.6 to have high rank with very high probability, more than $1 - q^{-n}$. Thus the dimension of the intersections of the null spaces of each L is zero with probability at least $1 - 2q^{-n}$.

Clearly, the condition for these equations to be satisfied is that S sends V to the intersection of the null spaces of each such L. Thus S is with high probability the zero map on V and so $V^\perp = \{0\}$. This generates a contradiction, however, since $2n \leq dim(V) + dim(V^\perp) < 2n$. Thus, with probability greater than $1 - 2q^{-n}$, f has no nontrivial differential invariant structure.

□

4.2.2 Invariant Analysis of HFEv^-

The situation for HFEv^- is quite similar, but the probabilities are slightly different. Specifically one must note that since the condition of being a differential invariant is a condition on the span of the public differential forms, under projection this condition is weaker and easier to satisfy. For specificity, we consider the removal of a single public equation, though, critically, a very similar though notationally messy analysis is easy to derive in the general case.

We may model the removal of a single equation, which can be considered a corank 1 projection, as a projection of the form $\pi(x) = x^q + x$ applied after the central map. We will show this by proving the more general statement below. This proof was originally published in [9].

Claim 1. *Let $\Pi \circ T$ be a corank a linear transformation on $\mathbb{F}_q^n$. There exists both a nonsingular linear transformation S and a degree q^a polynomial π such that*

$$\Pi \circ T = S \circ \phi^{-1} \circ \pi \circ \phi$$

Proof. Let V be the kernel of $\Pi \circ T$ and let $\pi = \mathscr{M}_V = \Pi_{v \in V}(x - v)$. We may call π the "minimal polynomial" of V because this is the polynomial of minimal degree such that every element of V is a root. Note that $|V| = q^a$. This is because each $v \in V$ will have 0's in the first $n - a$ coordinates, while the remaining a terms can be any value in $\mathbb{F}_q$. Thus we see from the definition of π that $\mathscr{M}_V(x)$ has degree q^a and will be of the form

$$x^{q^a} + c_{a-1}x^{q^{a-1}} + \cdots + c_1 x^q + c_0 x$$

where $c_i \in \mathbb{K}$.

Now let $B_v = \{b_{n-a}, b_{n-a+1}, \ldots, b_{n-1}\}$ be a basis for V. We can extend B_v into some vector $B = \{b_0, \ldots, b_{n-a-1}, b_{n-a}, b_{n-a+1}, \ldots, b_{n-1}\}$ such that B is a basis for $\mathbb{F}_q^n$.

Let M be the matrix that maps the standard basis to B. We can see that the last a columns of the matrix representations of $M^{-1}(\Pi \circ T)M$ and $M^{-1}(\phi \circ \pi \circ \phi)M$ will be zero.

Observe that there exist invertible matrices A and A', corresponding to row operations, such that both $AM^{-1}(\Pi \circ T)M$ and $A'M^{-1}(\phi \circ \pi \circ \phi)M$ are in reduced echelon form; that is:

$$AM^{-1}(\Pi \circ T)M = \left[\begin{array}{c|c} \mathrm{I} & 0 \\ \hline 0 & 0 \end{array}\right] = A'M^{-1}(\phi \circ \pi \circ \phi)M$$

Solving for $\Pi \circ T$, we obtain

$$\Pi \circ T = MA^{-1}A'M^{-1}(\phi^{-1} \circ \pi \circ \phi)$$

Let $S = MA^{-1}A'M^{-1}$ and the proof is complete. □

So to remove one equation, we consider Π to be a corank 1 projection, $a = 1$, $|V| = q^1$, and $\pi = \Pi_{v \in V}(x - v) = x^q + x$. We have

$$\Pi \circ T \circ \phi^{-1} \circ f \circ \phi \circ U = S \circ \phi^{-1} \circ (\pi f) \circ \phi \circ U$$

4.3 Degree of Regularity, Q-Rank, and Parameters

Further considerations for the security of HFEv^- are the degree of regularity, a quantity closely connected to the complexity of algebraic attacks, and the Q-rank of the public key. A careful analysis of each of these quantities reveals that they support the security of HFEv^- against an algebraic attack such as [21] and against the Kipnis-Shamir methodology and its improvements, see [10, 11].

In [22], it is shown that an upper bound for the Q-rank of an HFE component, the number of removed equations, and the Q-rank of the vinegar component. For Gui-96(96,5,6,6), here $q = 2$, $n = 96$, $D = 5$, $v = 6$, and $r = 6$, this quantity is

roughly 15. Furthermore, in [19], experimental evidence in the form of analysis of toy variants is provided indicating that this estimate is tight. Thus the complexity of a Kipnis-Shamir style attack is roughly $O(n^3 q^{15n})$.

Also in [22], a formula for an upper bound on the degree of regularity for HFEv^- systems is derived. Given the parameters of Gui-96(96,5,6,6), the degree of regularity is expected to be 9. Further, experiments are provided in [19], supporting the tightness of this approximation formula for toy schemes with n as large as 38. With this degree of regularity the expected complexity of inverting the system via Gröbner basis techniques is given by

$$\binom{96-6+9}{9}^{\omega} \approx 2^{93}$$

where $2 \leq \omega < 3$ is the linear algebra constant. We note than an error in the approximation of the degree of regularity can easily change this estimate by a factor of a few thousand. Still, it seems clear that each of these avenues of attack is unviable.

Still another attack vector is to put the entropy of the key space to the test with techniques such as those mentioned in [23] for deriving equivalence classes of keys. With our most restrictive instance of the key verification algorithm in Section 4.1.2, we have a key space consisting of roughly q^{13n} central maps, roughly q^{6n} of which can be seen as equivalent keys as in [23]. Thus provable security against the differential adversary can be achieved with a key space of size far beyond the reach of the "guess-then-IP" strategy.

CHAPTER 5
AN UPDATED SECURITY ANALYSIS OF PFLASH

SFLASH, was thought to be secure until its break in [4] (described in Section 3.2.1). This scheme was attacked by exploiting its differential symmetry. The authors of [24] present a way to resist the attack on SFLASH through projection, which produces a scheme we now call PFLASH. PFLASH is still a very fast signature scheme and is amenable to low-power environments without sacrificing side-channel resistance. This projected C^{*-} system is shown to resist differential cryptanalysis for restricted parameters, that is, when the degree is bounded by $q^{n/2-d}$, in [17] and is fully specified with paractical parameters in [5].

Since the design of PFLASH there have been a number of cryptanalytic developments in the big field venue. The development of differential invariant attacks in [25] and their further application in [26] are examples of advancement in this active area. Furthermore, the improved efficiency of the Kipnis-Shamir (KS) attack of [10] presented in [11] is directly impactful to PFLASH, as one can consider PFLASH as a possibly high degree but still low rank version of HFE$^-$.

We expand and update the analysis in [17] and [5] proving resistance to differential and rank techniques for the vast majority of parameters, and we verify that the provably secure key spaces are less limiting than previous works suggest. This improvement is directly impactful, providing further assurance that attacks based on equivalent keys cannot weaken PFLASH.

The degree bound restriction in [17] reduces the dimension of possible private keys by a factor of more than two. Our updated differential analysis verifies the

security of the scheme when the central map has no degree bound, and thus assures us that very little entropy is lost in the key space when restricting to parameters that are provably secure against differential adversaries.

In [5], an argument for the resistance of PFLASH to the technique of [11, Section 8.2] when PFLASH is considered as a low degree projected HFE$^-$ scheme is provided. We make this assessment more robust by also considering the possibility of an adversary attempting to remove the projection modifier from PFLASH considering it to be a higher rank HFE$^-$ scheme. Whereas in the former case, the attack is impossible, in the latter case, the algebraic structure allows the possibility that the attack can succeed; however, the complexity of the attack is directly computed and shown to be infeasible.

5.1 Updated Differential Analysis of Projected Primitive

As discussed in [17], we may assume that the projection mapping is tied to f and consider differential symmetries of $f \circ \pi$ where π is chosen in a basis such that $deg(\pi) = q^d$. Clearly, if $f \circ \pi$ has a differential symmetry then the equation $Df(Ma, \pi x) + Df(\pi a, Mx) = \Lambda_M Df(\pi a, \pi x)$ is satisfied for some M. We can express this relation with matrix multiplication, namely

$$a^\top(\Pi^\top \mathbf{Df} M)x + a^\top(M^\top \mathbf{Df}\Pi)x = \Lambda_M[a^\top(\Pi^\top \mathbf{Df}\Pi)x],$$

where $\mathbf{Df}$ is the matrix representing Df as a bilinear form over $\mathbb{K}$, having one in the $(0, \theta)$ and $(\theta, 0)$ coordinates and zero elsewhere, where $\Pi x = \sum_{i=0}^{d} \beta_i x^{q^i}$ and where $Mx = \sum_{i=0}^{n-1} m_i x^{q^i}$.

Examining this equation, we see that $a^\top(\Pi^\top \mathbf{Df} M)x + a^\top(M^\top \mathbf{Df}\Pi)x$ will have nonzero entries restricted to certain coordinates depending only on d and θ, see Figure 5.1. Similarly, the right hand side of the equation, $\Pi^\top \mathbf{Df}\Pi$, has a structure dependent upon d and θ, see Figure 5.2. Notice, the graphs may look

different depending on the choice of θ and d.

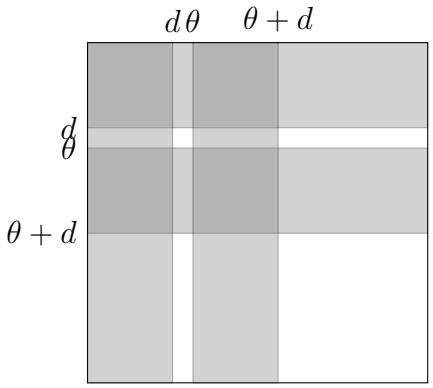


Figure 5.1: The shape of the matrix representation over $\mathbb{K}$ of $Df(Ma, \pi x) + Df(\pi a, Mx)$. Shaded regions correspond to possibly nonzero values.

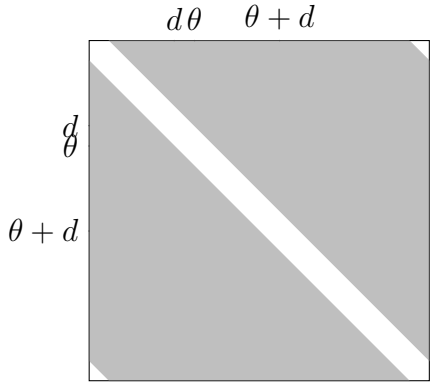


Figure 5.2: The shape of the matrix representation of $\Lambda_M Df(\pi a, \pi x)$ over $\mathbb{K}$. Shaded regions correspond to possibly nonzero values.

The strategy for finding conditions on π, M and Λ_M for the existence of such a symmetry is then to find coordinates in which one side of this matrix equation is zero while the other side involves only a single unknown coefficient of M or Λ_M. While this system of equations is nonlinear in the coefficients of π, it is linear in both the unknown coefficients of M and those of Λ_M.

The system contains many more equations than variables, but certainly generates a positive dimensional ideal. The reason is that for any fixed π, $M = a\pi$ for any $a \in \mathbb{F}_q$ generates a solution. On the other hand, for a fixed π and a fixed θ, the above system becomes linear with the number of nonzero equations depending on both d and θ. Even in the best case, the number of equations is far larger than the number of variables. Since the coefficients of π are the only source of randomness for this system of linear equations, the great number of equations are not independent in a probabilistic sense. Therefore, probabilistic arguments are difficult, though

extensive experiments show that the solution space is generally one dimensional.

Luckily, we can do better by bootstrapping the result of [17]. Specifically, we examine the case when $\theta > \frac{n}{2}$.

Lemma 3. *$f(x^{q^\rho}) = f(x)^{q^\rho}$ when $f(x) = x^{q^\theta+1}$*

Proof. $f(x^{q^\rho}) = (x^{q^\rho})^{q^\theta+1} = x^{(q^\theta+1)q^\rho} = \left(x^{q^\theta+1}\right)^{q^\rho} = f(x)^{q^\rho}$

□

Consider the special case of Lemma 3 when $\rho = -\theta$. After applying this map to the output of **Df**, the nonzero terms, originally in the $(\theta, 0)$ and $(0, \theta)$ coordinates, are transported to the $(0, -\theta)$ and $(-\theta, 0)$ coordinates, respectively. This observation leads to the following theorem, revealing that most parameters of PFLASH are provably secure against a differential adversary.

Theorem 1. *Let $f(x) = x^{q^\theta+1}$ be a C^* map, and let M and $\pi x := \sum_{i=0}^{d} x^{q^i}$ be linear. Suppose that f satisfies the symmetric relation:*

$$Df(Ma, \pi x) + Df(\pi a, Mx) = \Lambda_M Df(\pi a, \pi x).$$

If $d < min\{\frac{n}{2} - \theta, |n - 3\theta|, \theta - 1\}$, or if $d < \{\theta - \frac{n}{2}, |2n - 3\theta|, n - \theta - 1\}$, then $M = M_\sigma \circ \pi$ for some $\sigma \in k$.

Proof. *Assume $Df(Ma, \pi x) + Df(\pi a, Mx) = \Lambda_M Df(\pi a, \pi x)$ holds true. Then, we have two cases.*

1.) $\theta < \frac{n}{2}$

By [17, Theorem 3], we are done.

2.) $\theta > \frac{n}{2}$

Let $\widetilde{f}(x) = f(x)^{q^{-\theta}} = f\left(x^{q^{-\theta}}\right)$

We have,

$$Df(Ma, \pi x) + Df(\pi a, Mx) = \Lambda_M Df(\pi a, \pi x)$$

$$[Df(Ma, \pi x) + Df(\pi a, Mx)]^{q^{-\theta}} = [\Lambda_M Df(\pi a, \pi x)]^{q^{-\theta}}$$

$$[Df(Ma, \pi x) + Df(\pi a, Mx)]^{q^{-\theta}} = L_\theta^{-1} \Lambda_M Df(\pi a, \pi x)$$

Let L_θ represent the map that raises terms to the θ^{th} power. We can use the definition of the discrete differential to expand the left hand side of the equation. By linearity, we can distribute the exponent $q^{-\theta}$ to each term. After applying our lemma we get the following,

$$\tilde{f}(Ma+\pi x)+\tilde{f}(Ma)+\tilde{f}(\pi x)+\tilde{f}(\pi a+Mx)+\tilde{f}(\pi a)+\tilde{f}(Mx) = L_\theta^{-1}\Lambda_M Df(\pi a, \pi x)$$

By adding $0 = 2\tilde{f}(0)$ to the left and applying $I = L_\theta L_\theta^{-1}$ to the right we get,

$$D\tilde{f}(Ma, \pi x) + D\tilde{f}(\pi a, Mx) = L_\theta^{-1}\Lambda_M (L_\theta L_\theta^{-1}) Df(\pi a, \pi x)$$

And by the lemma we have,

$$D\tilde{f}(Ma, \pi x) + D\tilde{f}(\pi a, Mx) = L_\theta^{-1}\Lambda_M L_\theta D\tilde{f}(\pi a, \pi x)$$

We now have a relation on $\widetilde{f}(x)$ where $-\theta + d < \frac{n}{2}$. Now we can apply [17, Theorem 3] to conclude that $M = M_\sigma \circ \pi$ for some $\sigma \in k$.

□

We note that the existence of a differential symmetry on $f \circ \pi$ implies a solution of the equation in Theorem 1 as well as the commutativity of M_σ and π. Since the commutativity of M_σ and π requires that π is L-linear, where $\mathbb{F}_q \subseteq L \subseteq k$ and $\sigma \in L$, for any nontrivial differential symmetry to exist, $(d, n) > 1$. Thus, there is a most desirable value of d from an efficiency and security standpoint: $d = 1$.

Let us specifically consider this most desired value $d = 1$. Then the only restriction on θ for provable differential security is

$$\theta \in \left(2, \frac{n-1}{3}\right) \cup \left(\frac{n+1}{3}, \frac{n}{2} - 1\right) \cup \left(\frac{n}{2} + 1, \frac{2n-1}{3}\right) \cup \left(\frac{2n+1}{3}, n-2\right).$$

Furthermore, since $\theta = \frac{n}{2}$ always produces a many-to-one map in any characteristic, the restriction to provably secure parameters for PFLASH eliminates at most four possible values for θ for all extension degrees n.

5.2 Extension to PFLASH

We now generalize the analysis of the previous section in application to PFLASH. First we derive a heuristic argument for bootstrapping the provable security of the composition $f \circ \pi$ to statistical security for the projected primitive. We then clarify the resistance of PFLASH to analysis as an HFE$^-$ scheme. Finally, we derive security bounds for various PFLASH parameters.

5.2.1 Differential Analysis

As previously mentioned, proof that differential symmetries do not exist for the central map of a scheme verifies that a differential adversary cannot recover a full rank key. Such a proof does not, however, verify that a differential adversary cannot find a symmetry revealing the extension field multiplicative structure and directly attack the scheme.

To illustrate this principal, imagine a high degree variant of HFE in which the central map has the form $f(x) = x^{q^\theta+1} + \pi_2(Q(x))$ over an extension of degree $2n$, where π_2 is a rank n projection onto the complement of the subfield of size q^n and Q is an arbitrary quadratic. Then any minus variant in which the image of π_2 is the kernel of T is a C^{*-} public key, but one with multiplicative symmetry. In

particular, any map L representing muliplication by an element in the intermediate extension of degree n would satisfy

$$D(T \circ f \circ U)(U^{-1}La, x) + D(T \circ f \circ U)(a, U^{-1}Lx) = (L^{q^\theta} + L)D(T \circ f \circ U)(a, x).$$

Thus the minus scheme has a multiplicative symmetry even though the original scheme provably does not. In fact, even more strongly, we have computed functions of the form of f above over a degree 6 extension of $GF(2)$ for which no linear differential symmetry of any form exists, but under projection onto the degree 3 subfield, the *multiplicative* symmetry is exhibited.

In the case of PFLASH, we may attempt the strategy of the previous section for proving security. We may always model the removal of r equations as the application of a polynomial $\pi(x) = \sum_{i=0}^{r} a_i x^{q^i}$ to the central map. If only a few equations are removed, then the analysis proceeds just like in [12], because $f \circ \pi$ is a low rank albeit high degree polynomial. Since no parameters suggested for PFLASH are near this range, however, this analysis does not apply. When we perform this analysis with $r \approx \frac{n}{3}$ and $f \circ \pi$, however, the methods of the previous section fail to generate a provably secure class of private keys.

Fortunately, there is an easy heuristic argument revealing a simple relationship between symmetries of the central map and symmetries of a map with the minus modifier that shows that symmetry should be statistically no more likely for any minus modified scheme than for the original. Let T' be the minus projection composed with the inclusion mapping with domain $\mathbb{F}_q^{n-r}$ and codomain $\mathbb{K}$. Suppose that $T' \circ f \circ \pi$ has a differential symmetry. Then

$$D(T' \circ f)(\pi a, Mx) + D(T' \circ f)(Ma, \pi x) = \Lambda_M D(T' \circ f)(\pi a, \pi x)$$

$$T' \left[Df(\pi a, Mx) + Df(Ma, \pi x) \right] = \Lambda_M T' Df(\pi a, \pi x).$$

Since the left is clearly in $T'\mathbb{K}$, the right must be as well. Thus, with high probability, that is, when $Span_{a,x}(Df(\pi a, \pi x)) = \mathbb{K}$, we have that $\Lambda_M T'\mathbb{K} = T'\mathbb{K}$.

We know from linear algebra that in this case there exists at least one invertible transformation Λ'_M such that $\Lambda_M T' = T'\Lambda'_M$. Therefore, we obtain the relation

$$Df(\pi a, Mx) + Df(Ma, \pi x) = \Lambda'_M Df(\pi a, \pi x) \ \ (mod\ ker(T')). \tag{5.1}$$

Clearly, this argument is not reversible for any Λ'_M satisfying (5.1); therefore, we cannot in general conclude that the scheme with the minus modifier inherits any differential symmetry from the central map. On the other hand, satisfying (5.1) imposes $n - r$ constraints on Λ_M, while the "commuting" of Λ_M with T' imposes another r constraints. Thus, the existence of a symmetry in the minus case imposes the same number of constraints on Λ_M as for the central map and so we expect the probability of the existence of a differential symmetry to be no higher than for the central map.

5.2.2 Rank Analysis

One can consider PFLASH to be a high degree version of HFE$^-$ by absorbing the projection of the variables into the central map. Notice that the rank of the composition is still only two, thus PFLASH must achieve its security from the minus modifier.

Recently, in [27], a key recovery attack valid for all parameters of HFE$^-$ is presented. For an HFE$^-$ instance with parameters (q, n, D, r), the complexity is noted as $\mathcal{O}(\binom{n+\lceil log_q(D)\rceil+1}{\lceil log_q(D)\rceil+r+1}^{\omega})$.

In application to PFLASH, there are two things to note about this attack. First, the attack produces an equivalent HFE$^-$ key, not a pC^{*-} key. This fact may not limit the attack, because it will still recover a central map of rank two of the form $f \circ \pi$ which we may then attack as a pC^* scheme in the manner of [28]. Second, the quantity $\lceil log_q(D)\rceil$ in the complexity estimate is derived from the rank structure that the degree bound of HFE implies, not directly from the degree bound itself.

Thus, the rank of the C^* monomial, which is two, plays the role of $\lceil log_q(D) \rceil$ in the application of the techniques of [27] to PFLASH.

In fact, instances of PFLASH with quite inappropriate but still large parameters can be broken with this method. In particular we note that for a PFLASH$(256, 44, 3, 1)$ that the complexity of the attack is roughly estimated $44^{(3+2+1)\omega} \sim 2^{78}$. For large values of r, however, such as in all parameter sets in [5], this attack is infeasible. For example, the smallest parameters suggested in [5] still resist this attack to dozens of orders of magnitude beyond brute force. Thus, for sensible parameters with r sufficiently large, PFLASH is secure.

5.2.3 Security Estimates

Now with a refined security analysis, we can eliminate differential attacks for a larger set of parameters, thus doubling the entropy of the key space for PFLASH. In addition, with the complexity estimate of $\mathscr{O}(n^{(r+3)\omega})$ and practical values of r, PFLASH is quite secure against the new attack on HFE$^-$ schemes. In conjunction with the invariant analysis of [5], we conclude that the security of PFLASH is determined by its resistance to algebraic and brute force attacks.

Viewing PFLASH as an HFE$^-$ scheme, we may use the bound in [22] to estimate the degree of regularity of PFLASH. This upper bound can be computed

$$\frac{(q-1)(R+r)}{2} + 2,$$

where R is the rank of the central map; in the case of PFLASH, this quantity is two. Though this is an upper bound, empirical evidence suggests that it is tight for random systems of rank R. Thus the degree of regularity is far too high for practical schemes to be weakened. Furthermore, direct algebraic attacks for large schemes are impractical even with smaller complexity bounds because the space complexity of the best algorithms are too large to be practical.

Therefore, we corroborate the claims of [5] that brute force collision attacks are the greatest threat to PFLASH schemes. The evidence from our increase of the entropy of the key space and the verification that PFLASH resists recent weaknesses revealed in HFE^- suggest the security levels in Table 5.1 (all of which are in agreement with [5]).

Scheme	Public Key (B)	Security (b)
PFLASH$(16, 62, 22, 1)$	39,040	80
PFLASH$(16, 74, 22, 1)$	72,124	104
PFLASH$(16, 94, 30, 1)$	142,848	128

Table 5.1: Security levels for standard parameters of PFLASH

CHAPTER 6
EFLASH: A NEW MULTIVARIATE ENCRYPTION SCHEME

Many of the previously discussed multivariate schemes are used as signature schemes. Recently we have seen new candidates and strategies emerge for multivariate encryption. Previously, multivariate schemes centered around bijective functions that map from vector spaces of size n back into a vector space of size n. The problem with this strategy is that there are not many bijective quadratic maps. Furthermore, of the maps that do exist, many of these functions were either too hard to invert, or too easy to invert. The common practice to try to overcome this downfall was to try to hide an easily invertible function by composing the bijective function with affine maps.

In 2013, Tao et al. proposed relaxing the *bijective* condition for the central function and replacing it with an *injective* map with a much larger codomain in [29]. In theory, this would make hiding the structure of the map while maintaining efficient inversion easier to accomplish. The recent resurgence of multivariate encryption is due primarily to this change in philosophy. Many schemes have been proposed along these apparently promising lines.

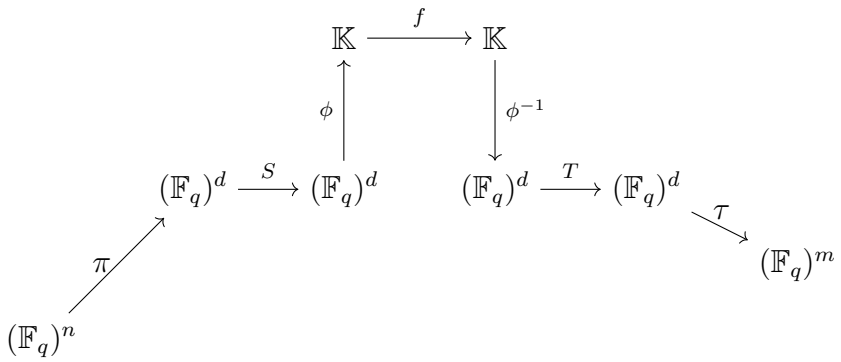
Some notable schemes that increase the codomain size of the central mappings include the ABC Simple Matrix scheme, see [29], which utilizes a large matrix algebra structure; ZHFE, see [30],which is similar to a high degree version of HFE with a single variable over the extension; and SRP, see [31], which combines the Square encryption scheme, Rainbow signature scheme, and Plus method. Although these schemes appear promising, many of these schemes have subsequently been

the victims of surprising (if not disabling) cryptanalysis. The attacks on ABC from [25, 26, 32] work well if the base field is small, and both ZHFE and SRP were broken in [33] and [34], respectively.

In the following chapter, we will introduce a new multivariate encryption scheme, EFLASH. Our scheme will be a new parameterization of a projected C^{*-}. A major difference between our scheme and PFLASH is the size of the projection. The size of our projection π will be much larger.

6.1 Algebraic Structure

We will let n be the number of variables and $d > n$ be the degree of the extension field over $\mathbb{F}_q$. We will let $m \geq n$ be the number of equations ($m < d$) and denote the number of equations removed by $a = d - m$. We will compose our central map $f(x) = x^{q^\theta+1}$ with affine maps S and T from $\mathbb{F}_q^d$ to $\mathbb{F}_q^d$. We let ϕ be a vector space isomorphism from $\mathbb{F}_q^d$ to $\mathbb{K}$, π be a linear embedding from $\mathbb{F}_q^n$ to $\mathbb{F}_q^d$, and τ be a linear projection from $\mathbb{F}_q^d$ to $\mathbb{F}_q^m$.



Our public equations P can be found by computing $P = \tau \circ T \circ \phi^{-1} \circ f \circ \phi \circ S \circ \pi$.

6.2 Encryption and Decryption

To encrypt a message $\overline{x}$, the sender would just compute $P(\overline{x}) = \tau \circ T \circ \phi^{-1} \circ f \circ \phi \circ S \circ \pi(\overline{x}) = \overline{y}$ to get ciphertext $\overline{y}$. To decrypt the message we will take

advantage of some of the weaknesses that an unmodified C^* scheme possesses.

To decrypt, we exploit the more efficient method of inversion Patarin developed in his linearization equations attack from [2].

As shown in 2.2.2, if $\overline{v} = (\phi^{-1} \circ f \circ \phi)\overline{u}$, then we know there exists a bilinear relationship between $\overline{v}$ and $\overline{u}$. Let $\overline{y}' = T \circ \phi^{-1} \circ f \circ \phi \circ S \circ \pi \circ \overline{x}$, $\hat{v} = T^{-1} \circ \overline{y}'$, and $\hat{u} = S \circ \pi \circ \overline{x}$.Then, $\hat{v} = (\phi^{-1} \circ f \circ \phi)\hat{u}$. Thus, we see we have a bilinear relationship between our plaintext and ciphertext variable. This tells us there is a system of d polynomials of the form

$$\sum_{0 \leq i,j < d} \alpha_{i,j,\ell} u_i v_j + \sum_{0 \leq i < d} \beta_{i,\ell} u_i + \sum_{0 \leq i < d} \gamma_{i,\ell} v_i + \delta_\ell$$

in the coefficients of $\overline{u}$ and $\overline{v}$ which are simultaneously zero.

Given access to the private key the calculation of this bilinear relation is immediate. Adding the linearization equations to the private key can be considered a drawback as it increases the private key size, but is an important aspect for our algorithm.

Inversion, given the ciphertext $\overline{y}$, is then accomplished by concatenating every possible suffix $\overline{y}_a$ to discover $\overline{y}' = \overline{y}||\overline{y}_a$. Success is determined by solving the affine system in $\overline{x}$ induced from the linearization equations upon input $\overline{y}'$. If the affine system has a solution, $\overline{x}$, we can be assured that $P(\overline{x}) = \overline{y}$.

6.3 Decryption Failure Rate

We want to find the probability that there are multiple preimages of y under τ, which would result in a decryption failure. Specifically, we want to compute the probability that $x_1, x_2, y \in \mathbb{F}_q$ exists such that $P(x_1) = P(x_2) = y$, given that $P(x_1) = y$. Given our function $P(x) = \tau \circ T \circ \phi^{-1} \circ f \circ \phi \circ S \circ \pi(x)$, it is clear that the only part of this function that is not injective is τ, and that π is the only additional map that is not bijective. Thus we compute the probability of decryption failure

under the simplifying heuristic that the central map $\hat{P}(x) = T \circ \phi^{-1} \circ f \circ \phi \circ S(x)$ is a random bijection. This assumption is obviously false as f is a quadratic map, but we believe this heuristic to be statistically useful. Let $A = \text{image}(\pi)$, $|A| = q^n$. We can consider B to be the preimage of y under τ, so under our simplifying heuristic B is a random set of q^a elements from $\mathbb{F}_q^d$.

We will use Bernoulli trials to estimate the probability that y is the image of at least two distinct elements of $\mathbb{F}_q^n$, given that it is the image of at least one. If $Pr(\hat{P}(x) \in B : \hat{P}(x) \in A) = p$, then the probability of k elements in A being in B is $\binom{q^n}{k}(1-p)^{(q^n-k)}p^k$.

The probability of $\hat{P}(x) \in B$ is $\frac{q^a}{q^d} = q^{-m}$, and the probability that $\hat{P}(x)$ is not in B is $1 - q^{-m}$. Thus we compute:

$$\begin{aligned} Pr(|A \cap B| \geq 2 \mid |A \cap B| \geq 1) &= \frac{Pr(|A \cap B| \geq 2)}{Pr(|A \cap B| \geq 1)} \\ &= \frac{1 - \Big(Pr(|\mathbf{G} \cap \tau^{-1}(y)| = 0) + Pr(|\mathbf{G} \cap \tau^{-1}(y)| = 1)\Big)}{1 - Pr(|\mathbf{G} \cap \tau^{-1}(y)| = 0)} \end{aligned}$$

Therefore we find $Pr(|A \cap B| \geq 2 \mid |A \cap B| \geq 1)$ to be $p = \frac{1-(1-q^{-m})q^n - q^{n-m}(1-q^{-m})^{q^n-1}}{1-(1-q^{-m})^{q^n}}$. To find an upper bound for the probability p, we find an upper bound for the numerator, and a lower bound for the denominator.

Claim 2. $\binom{a}{i+1}(q^{-(i+1)m}) < \binom{a}{i}(q^{-im})$ *when* $a < q^m$

Proof. Notice that $\binom{a}{i+1}(q^{-(i+1)m}) = \frac{a!}{(i+1)!(a-i-1)!q^{(i+1)m}}$ has the same numerator as $\binom{a}{i}(q^{-im}) = \frac{a!}{(i)!(a-i)!q^{im}}$, so we will prove the claim by showing the denominator of the left hand side is larger than the denominator of the right hand side.

Clearly $(i+1)! > i!$, and $q^{(i+1)m} > q^{im}$ by a factor of q^m. We see that $(a-i-1)! < (a-i)!$ by a factor of $a-i$, but we know that $a - i < a < q^m$. Thus we can conclude $(i+1)!(a-i-1)!q^{(i+1)m} > (i)!(a-i)!q^{im}$ and therefore $\binom{a}{i+1}(q^{-(i+1)m}) < \binom{a}{i}(q^{-im})$ when $a < q^m$. □

Bounding the numerator: $1-(1-q^{-m})^{q^n}-q^{n-m}(1-q^{-m})^{q^n-1}$.

Using binomial coefficients and the above claim, we see that:

$$(1-q^{-m})^{q^n} = (1-\tbinom{q^n}{1}q^{-m}+\tbinom{q^n}{2}q^{-2m}-\cdots) \geq 1-q^n q^{-m}.$$

Thus $1-(1-q^{-m})^{q^n} \leq 1-(1-q^{n-m})$.

By the same argument, we are given:

$$(1-q^{-m})^{q^n-1} = (1-\tbinom{q^n-1}{1}q^{-m}+\tbinom{q^n-1}{2}q^{-2m}-\cdots) \geq 1-(q^n-1)q^{-m}.$$

Therefore, $-q^{n-m}(1-q^{-m})^{q^n-1} \leq -q^{n-m}(1-(q^n-1)q^{-m})$. Thus the numerator is bounded above by $1-(1-q^{n-m})-q^{n-m}(1-(q^n-1)q^{-m})$.

Bounding the denominator: $1-(1-q^{-m})^{q^n}$ Similar to our argument for bounding the numerator, we will use binomial coefficients and claim 1 to find:

$$(1-q^{-m})^{q^n} = (1-\tbinom{q^n}{1}q^{-m}+\tbinom{q^n}{2}q^{-2m}-\cdots) \leq 1-\tbinom{q^n}{1}q^{-m}+\tbinom{q^n}{2}q^{-2m}$$

Hence the denominator is bounded below by $1-(1-q^{n-m}+\frac{q^n q^n-1}{2}q^{-2m})$.

Finding a bound for the probability, p

$$\begin{aligned}
p &= \frac{1-(1-q^{-m})^{q^n}-q^{n-m}(1-q^{-m})^{q^n-1}}{1-(1-q^{-m})^{q^n}} \\
&\leq \frac{1-(1-q^{n-m})-q^{n-m}(1-(q^n-1)q^{-m})}{1-(1-q^{n-m}+\frac{q^n q^n-1}{2}q^{-2m})} \\
&= \frac{1-1+q^{n-m}-q^{n-m}+q^{n-m}(q^n-1)q^{-m}}{1-1+q^{n-m}-\frac{q^n(q^n-1)}{2}q^{-2m}} = \frac{q^{n-m}(q^n-1)q^{-m}}{q^{n-m}-\frac{q^n(q^n-1)}{2}q^{-2m}} \\
&= \frac{q^{n-m}(q^n-1)q^{-m}}{q^{n-m}-q^{n-m}(q^{-(n-m)}\frac{q^n(q^n-1)}{2}q^{-2m})} \\
&= \frac{q^{n-m}-q^{-m}}{1-(\frac{q^{n-m}-q^{-m}}{2})}
\end{aligned}$$

When $q=2$, empirical evidence shows we can approximate this by 2^{n-m-1}. The data to support this claim are shown in Table 7.1.

q	n	d	a	m	$n-m$	decrypt fail rate
2	14	34	8	26	-12	$2^{-13.13}$
2	14	35	8	27	-13	$2^{-13.94}$
2	14	36	8	28	-14	$2^{-14.94}$
2	14	37	8	29	-15	$2^{-15.94}$
2	14	38	8	30	-16	$2^{-17.64}$

Table 6.1: Probability of decryption failure for specific parameters of EFLASH.

6.4 Resistance to Known Attacks

The security analysis of EFLASH is quite related to that of PFLASH because of the similar algebraic structure. There are three attack methods that must be considered. Since the scheme requires more equations than variables to ensure a low probability of decryption failure, we require a careful analysis of the direct algebraic attack to ensure that the degree of regularity of the scheme is not too low. Second, in light of the attack on HFE- schemes, see [9], we require a MinRank analysis. Finally, given the history of the lineage of the C^* family, we require an analysis of symmetric differential methods.

6.4.1 Algebraic Attack

The first relevant attack for EFLASH is the direct algebraic attack. Algebraically, EFLASH is a high degree projected HFE- scheme, in the sense that EFLASH has a low Q-rank like HFE. Applying a projection to the input variables cannot increase the Q-rank, so we analyze the Q-rank of the central map composed with the minus modifier.

The key observation is that, unlike the case of HFE in which removing one equation in general increases the Q-rank by one, since the quadratic form associated

with the central map is so sparse, the removal of one equation in general increases the rank by *two*. To see this, note that the coefficients of the quadratic form associated with HFE are restricted to a square submatrix whose size is typically the Q-rank of the map. A codimension one projection allows these coefficients to bleed into another row and column, which increases the size of the square by one. In contrast, the size of the smallest square containing the nonzero values in the quadratic form of the EFLASH central map is usually much larger than the Q-rank of EFLASH; in fact, the codimension one projection can produce two elements in original rows and columns, see Figure 7.1.

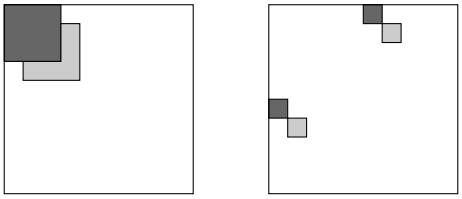


Figure 6.1: The shape of the matrices representing the central maps of HFE- and C^{*-}. The darkly shaded regions represent nonzero values of the central map without the minus modifier, the lightly shaded regions represent new nonzero values introduced by the removal of one equation. Unshaded areas have coefficients of zero.

Thus, the central map of EFLASH has Q-rank $2 + 2a$. By the formula provided in [22], we compute an upper bound on the degree of regularity,

$$d_{reg} \leq (q-1)(a+1) + 2. \tag{6.1}$$

When q is small this bound is known to be fairly tight. The complexity of the algebraic attack on EFLASH is therefore estimated to be $\mathscr{O}\left(\binom{n+d_{reg}}{d_{reg}}^{\omega}\right)$, where $2 \leq \omega \leq 3$ is the linear algebra constant.

Experiments were conducted on some small scale instances of EFLASH to study the behavior of the degree of regularity for values of n and $m = d - a$ of a

similar ratio to a full sized scheme with a low decryption failure rate. The results are shown in Table 6.2.

n	d	a	m	d_{reg}	d_{reg} (RANDOM)
16	28	9	19	4	4
24	37	9	28	4	5
32	47	9	38	5	6
40	56	9	47	≥ 6	7

Table 6.2: The degree of regularity of small scale EFLASH parameters in comparison to that of random systems of the same size.

The data show that the degree of regularity grows with the size of the system when a is fixed. Until our resource permissions were limited on the machine, each sufficiently large system exhibited a degree of regularity at most one less than that of a random system. We do not have a solid theoretical argument for why the degree of regularity should be bounded thusly; however, for the sizes of schemes necessary to achieve security, the upper bound provided by Equation (6.1) is already strictly less than the degree of regularity of random systems of the same size.

6.4.2 MinRank Attack

We can denote the calculations used to find our public equations P as matrix multiplications. Let $\mathbf{F}^{*\mathbf{i}}$ be the matrix representation of the i^{th} Frobenius power of the central map f. Then the matrix $\mathbf{F}^{*\mathbf{0}}$ represents our central map f, and is the $d \times d$ matrix with 1's in the $(0, \theta)$ and $(\theta, 0)$ coordinates and zeros elsewhere. Matrices $\mathbf{S}$ and $\mathbf{T}$ are $d \times d$ affine maps. We can also consider π as a linear embedding from $(\mathbb{F}_q)^n$ to $(\mathbb{F}_q)^d$, and τ as a linear projection from $(\mathbb{F}_q)^d$ to $(\mathbb{F}_q)^m$. Let σ be a primitive element of the extension, and thus $\{1, \sigma, \sigma^2, \ldots, \sigma^{d-1}\}$ is a basis vector over $\mathbb{F}_q$. Then mappings of ϕ and ϕ^{-1} can be represented as multiplication of M_d

and M_d^{-1}, respectively, where

$$M_d = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \sigma & \sigma^q & \dots & \sigma^{q^{d-1}} \\ \vdots & \vdots & \dots & \vdots \\ \sigma^{d-1} & \sigma^{(d-1)q} & \dots & \sigma^{(d-1)q^{d-1}} \end{pmatrix}$$

We can express the actions of τ by the following $d \times d$ matrix,

$$\tau^* = \begin{bmatrix} I_m & 0_{m\times a} \\ 0_{a\times m} & 0_{a\times a} \end{bmatrix}$$

.

Notice that $\tau^* : (\mathbb{F}_q)^d \to (\mathbb{F}_q)^d$. We will call $P^* := \tau^* \circ T \circ \phi^{-1} \circ f \circ \phi \circ S \circ \pi$. P and P^* will be comprised of the same m public equations, but P^* will then have a rows of 0 appended to it.

Consider $R = \phi \circ \tau^* \circ T \circ \phi^{-1}$. Then $R : \mathbb{F}_{q^d} \to \mathbb{F}_{q^d}$ is $\mathbb{F}_q$-linear. If we let $\widetilde{\tau}(x) = \Pi_{r\in\ker(R)}(x - r)$, then we know by proposition 2 in [12], there exists a nonsingular linear map $\widetilde{R}$ from $\mathbb{F}_{q^d}$ to $\mathbb{F}_{q^d}$ such that $Rx = \widetilde{R}\widetilde{\tau}x$. Let $\widetilde{T} = \phi^{-1}\circ\widetilde{R}\circ\widetilde{\tau}\circ\phi$. This brings us to the following claim.

Claim 3. $P^*(x) = \tau^* \circ T \circ \phi^{-1} \circ f \circ \phi \circ S \circ \pi x = \widetilde{T} \circ \phi^{-1} \circ f \circ \phi \circ S \circ \pi x$

Proof.

$$\begin{aligned} \widetilde{T} \circ \phi^{-1} \circ f \circ \phi \circ S \circ \pi &= \phi^{-1} \circ \widetilde{R} \circ \widetilde{\tau} \circ \phi \circ \phi^{-1} \circ f \circ \phi \circ S \circ \pi \\ &= \phi^{-1} \circ \widetilde{R} \circ \widetilde{\tau} \circ f \circ \phi \circ S \circ \pi \qquad (*) \\ &= \phi^{-1} \circ R \circ f \circ \phi \circ S \circ \pi \\ &= \phi^{-1} \circ \phi \circ \tau^* \circ T \circ \phi^{-1} \circ f \circ \phi \circ S \circ \pi \\ &= \tau^* \circ T \circ \phi^{-1} \circ f \circ \phi \circ S \circ \pi \\ &= P^* \end{aligned}$$

□

Now, let us reconsider $(*)$. We know that our public key is equivalent to $(*)$, so we see that

$$\begin{aligned} P^* &= \phi^{-1} \circ \widetilde{R} \circ \widetilde{\tau} \circ f \circ \phi \circ S \circ \pi \\ &= \phi^{-1} \circ \widetilde{R} \circ \phi \circ \phi^{-1} \circ \widetilde{\tau} \circ f \circ \phi \circ S \circ \pi \\ &= \widehat{T} \circ \phi^{-1} \circ \hat{f} \circ \phi \circ S \circ \pi \end{aligned}$$

Where $\hat{f}$ is our new central map and $\hat{f} = \widetilde{\tau} \circ f$ and $\widehat{T} = \phi^{-1} \circ \widetilde{R} \circ \phi$. We now consider $\widehat{\mathbf{F}}^{*\mathbf{i}}$ to be the i^{th} Frobenius power of the new central map $\hat{f} = \widetilde{\tau} \circ f$. If we denote $h = \phi^{-1} \circ \hat{f} \circ \phi$, then we can find symmetric matrices $(\mathbf{H_1}, \ldots, \mathbf{H_d}) \in (\mathbb{F}_q)^d$ such that $h_i = \overline{x}\mathbf{H_i}\overline{x}^\top$. As shown in [11] we see,

$$(\mathbf{H_1}, \ldots, \mathbf{H_d}) = (\mathbf{M_d}\widehat{\mathbf{F}}^{*\mathbf{0}}\mathbf{M_d}^\top, \ldots, \mathbf{M_d}\widehat{\mathbf{F}}^{*(\mathbf{d-1})}\mathbf{M_d}^\top)\mathbf{M_d^{-1}}. \tag{6.2}$$

If we denote the public key by $P = (g_1, g_2, \ldots, g_m)^\top$, then we can consider the symmetric matrices $(\mathbf{G_1}, \mathbf{G_2}, \ldots, \mathbf{G_m})$ that correspond to the public polynomials, such that $g_i = \overline{x}\mathbf{G_i}\overline{x}$. By analysis in [11] we find,

$$(\mathbf{G_1}, \ldots, \mathbf{G_m}) = (\pi\mathbf{S}\mathbf{M_d}\widetilde{\mathbf{F}}^{*\mathbf{0}}\mathbf{M_d}^\top\mathbf{S}^\top\pi^\top, \ldots, \pi\mathbf{S}\mathbf{M_d}\widetilde{\mathbf{F}}^{*(\mathbf{d-1})}\mathbf{M_d}^\top\mathbf{S}^\top\pi^\top)\mathbf{M_d^{-1}}\widetilde{\mathbf{T}} \tag{6.3}$$

When we consider our original central map, we saw that $\mathbf{F}^{*\mathbf{0}}$ has rank 2. Looking at our new central map $\hat{f}$, we see that $\widetilde{\tau}$ increases the rank. If we insist that θ is between $a+1$ and $d-a-1$, then $\widehat{\mathbf{F}}^{*0}$ has rank $2(a+1)$.

Notice that the embedding $\pi : (\mathbb{F}_q)^n \to (\mathbb{F}_q)^d$, and the affine map S will not *increase* the rank of the right hand side of (6.3), so it will not affect our MinRank attack. Applying $\widehat{T}$ normally does increase the rank, but it does not increase the min-Q-rank because it just produces new linear combinations of these matrices.

Using these facts and the analysis from [11] we find that we are solving the

MinRank problem:

$$\operatorname{rank}\Big(\sum_{k=0}^{m-1} \lambda_i \mathbf{G_i}\Big) \leq 2(a+1)$$

By the analysis in [35] and [36], the complexity of solving MinRank with the given parameters is $\mathscr{O}\Big(\binom{m+d_{reg}}{d_{reg}}^{\omega}\Big)$, where d_{reg} is the degree of regularity of the minors system and ω is the linear algebra constant. Treating EFLASH as a special case of HFE-, we may derive the degree of regularity of the minors system from [9, Conjecture 2] by using the Q-rank in place of the sum of the logarithm of the degree bound and the number of equations removed . Then we may estimate that the degree of regularity of the minors system is $d_{reg} = 2a + 3$.

6.4.3 Discrete Differential Attack

In Chapter 5 it is shown that almost all parameters of PFLASH are secure against differential adversaries. The proof relies on the fact that the corank of the projection is relatively small. Since EFLASH uses a corank $d - n$ projection, the security proof does not apply and so we must use other arguments.

By the symmetric argument to that in [9], we can express π under the appropriate basis as a polynomial in $\mathbb{K}$ of degree q^{d-n}. Thus, the central quadratic form can be considered a quadratic form in the $d - n$ "variables" $\pi(x)^{q^i}$, for $0 \leq i \leq d - n$. In characteristic two, there are at least as many linearly independent quadratic monomials as in GF(2); thus, there are at least $\binom{d-n+1}{2}$ linearly independent quadratic monomials in $\pi(x)^{q^i}$, for $0 \leq i \leq d - n$ over $\mathbb{K}$.

We expect that the locus of stabilizing pairs of matrices is zero-dimensional over $\mathbb{K}$, though it is necessarily positive dimensional over $\mathbb{F}_q$ since scalar multiples induce symmetry for any map. We performed experiments and found that the solution space was zero-dimensional over $\mathbb{K}$ in all cases. We conclude that the space of linear maps inducing symmetry on EFLASH is too small to be exploited like in

the attack on SFLASH of [37].

6.5 Parameter Selection

In choosing parameters for EFLASH, we need to consider security against the direct algebraic attack, the MinRank attack, and fault attacks exploiting decryption failure. We address the constraints each of these attacks places on parameters, as well as efficiency concerns.

The complexity of both the direct attack and the MinRank attack is directly related to the Q-rank of the public key. In the case of very small fields, such as GF(2), the degree of regularity is little larger than the Q-rank, $2a + 2$; thus, several equations must be removed to achieve security. Over GF(2), each increase in a doubles decryption time while making the direct attack approximately n times harder and the MinRank attack approximately $2m$ times harder.

To address decryption failures, we note that the probability estimate of Section 6.3 is approximately q^{n-m-1}. We set a reasonable bound 2^{-B} on the probability of decryption failure and may set $m = n + \frac{B}{lg(q)}$ to achieve this bound.

For larger q, the MinRank attack seems to be the most concerning. For efficiency reasons, it is impractical to have a large a; therefore, an instance with large q is vulnerable to MinRank. For this reason, we recommend the choice $q = 2$ with a and n sufficiently large to resist the algebraic attack. Our specific parameter selections for classical security levels are summarized in Table 6.3.It is important to note that our implementation is a proof of concept, and not at all optimized. This is a magma implementation, and we are only using one core.

In principle, Grover search should affect the security of these schemes, but at this time we are not aware of a result that indicates a Grover search would be feasible for such large parameters. It is possible that Grover search could halve the

Scheme	Security	Public Key B	Enc. (ms)	Dec. (ms)	Dec. Failure
EFLASH$(2, 80, 101, 5)$	80-bit	38892	0.7	194	2^{-17}
EFLASH$(2, 134, 159, 9)$	128-bit	169613	1.3	12758	2^{-17}

Table 6.3: Parameters and unoptimized performance of EFLASH(q, n, d, a) at the 80-bit and 128-bit classical security levels.

dimension of the preimage search space. Thus, we may have to roughly double the size of the plaintext. To protect against the possible threat of Grover search we consider the parameter selections shown in Table 6.4.

On the other hand, we may consider the possibility of the cryptosystem being implemented on a quantum device so that the search step in decryption may be Groverized. Therefore Grover's algorithm may, in fact, improve efficiency.

Scheme	Security	Public Key B	Enc. (ms)	Dec. (ms)	Dec. Failure
EFLASH$(2, 160, 181, 5)$	80-bit	141691	1.9	1140	2^{-17}
EFLASH$(2, 256, 279, 7)$	128-bit	559249	5.3	16177	2^{-17}

Table 6.4: Parameters and unoptomized performance of EFLASH(q, n, d, a) at the 80-bit and 128-bit quantum security levels.

CHAPTER 7
ALL IN THE C^* FAMILY

Section 2.2.3 introduces some of the modifiers that have been proposed to try to secure C^* schemes. In this chapter we will look at each of these modifiers in more detail, and consider new combinations of them. We will start by reviewing the modifiers, and introducing the notation we will use throughout the chapter.

One such modifier is the minus modifier $(-)$, which eliminates a equations from the public key. The plus modifier $(+)$ adds p random equations to the public key. Another modifier introduced is the projection modifier (p). The idea of projection is to fix the value of $d - n$ input variables. We call the codimension of this projection $t := d - n$.

The vinegar modifier (v) adds extra variables into the public key that can be assigned random values upon inversion. The effect of adding vinegar variables is that new quadratic terms, formed from both products of vinegar variables and C^* variables and products among vinegar variables, increase the Q-rank of the public key, that is, its rank as a quadratic form over $\mathbb{K}$. Vinegar variables are typically applied to Hidden Field Equation (HFE) schemes, as discussed in Chapter 4.

The final modifier we will discuss is internal perturbation (ip). For this modifier, we will consider a public key $P : \mathbb{F}_q^n \to \mathbb{F}_q^m$. We will denote the i^{th} equation of the public key as p_i, and plain-text vector $\overline{x} := (x_1, x_2, \ldots, x_n) \in \mathbb{F}_q^n$. Consider an affine map $S : \mathbb{F}_q^n \to \mathbb{F}_q^s$, and denote $\overline{z} := S(\overline{x})$. Additionally, consider $Q : \mathbb{F}_q^s \to \mathbb{F}_q^m$ to be a set of m quadratic equations, where q_i denotes the i^{th}. equation. We can create the internally perturbed public key $\tilde{P}$ by defining $\tilde{p}_i(\overline{x}) = p_i(\overline{x}) + q_i(\overline{z})$ for

each $0 < i \leq m$. The support dimension of (ip) will be denoted as s.

7.1 Known Combinations of Modifiers

The unmodified C^* scheme is easy to attack with the help of hindsight and twenty years of serious development in multivariate cryptography. Not only is C^* vulnerable to Patarin's linearization equations attack of [2], the scheme is also weak against differential methods such as [38], can be broken by finding the extension field structure via the techniques of [37] and is easily defeated by MinRank methods such as [11].

Several attempts at encryption and signatures derived from C^* have been proposed over the years using the modifiers of the previous section. Each of the modifiers has a critical weakness. The minus modifier is vulnerable to differential cryptanalysis as the practical attack on SFLASH of [39] shows. The plus modifier does not increase the MinRank of the scheme and so MinRank attacks are effective and reveal the output basis, breaking the scheme. The projected C^* scheme is still vulnerable to a differential attack as shown in [28], as is the internally perturbed C^* scheme, see [38]. The vinegar variant transforms C^* into a particularly bad HFEv scheme unless some hack making it similar to (ip) is applied. Interestingly, none of the modifiers alone are sufficient to secure C^*.

Combinations of these modifiers have seen some greater success. After the attack in [38] of the original PMI scheme of [40], an ipC^* scheme, PMI+, proposing the combination of the (ip) and (+) modifiers was presented in [41]. Similarly, PFLASH and EFLASH use both the projection and minus modifieres.

We present in Appendix II a summary of the security properties of C^* schemes with various combinations of modifiers. For brevity we are not exhaustive, and as a general rule it seems that combinations of modifiers each of which are

weak against attack A tend to remain weak against attack A.

Two things are important to note about this summary. First is that there is a notable exception to the rule of the previous paragraph. While it has been shown that pC^* schemes and C^{*-} schemes are both vulnerable to differential attacks, it was proven in [17] and generalized in [42] that the combination of projection with the minus modifier renders C^* invincible from differential attacks. Thus the combination (p-) is resistant to differential adversaries though (p) and (-) are weak. Second, resistance to these attack models is typically not binary but parameterized by the modifier. In particular, though the (ip) modifier provides resistance to MinRank attacks, the original parameters of PMI+ are easily broken by a simple modification of [11] and still larger parameters can be defeated by the new MinRank techniques developed in [43]. (Embarrassingly, we can find no reference to either such attack on PMI+.) We will revisit this analysis in Section 7.3 where we will generalize the analysis of the family of C^* schemes to the full array of possible schemes with all of their parameters.

7.2 The C^* Schema

In this section, we describe a C^* construction that is as general as possible using the modifiers of the previous section. We parameterize this generalized C^* scheme with the values:

- n, number of variables
- d, degree of extension
- s, support dimension of the (ip) modifier
- a, number of public equations removed
- p, number of plus polynomials

- t, corank of projection $(d-n)$

We call the resulting scheme a (n, d, s, a, p, t) scheme.

The only modifier whose use we do not consider in the C^* framework is the vinegar modifier, (v). Directly applying the vinegar modifier produces a degenerate HFEv scheme, since inversion must be accomplished via Berlekamp's Algorithm, see [8], or something similar. We do not consider this modifier as a C$^*v$ scheme would be no better performing than an HFEv scheme, while the C$^*v$ would have a smaller key space and worse security properties. Moreover, since the direct application of the vinegar modifiers results in an inversion process that does not use the structure of the C^* map, we do not consider this a C^* scheme, but rather a bad HFE scheme.

Let $\mathbb{K}$ be a degree d extension of $\mathbb{F}_q$ and let $\phi : \mathbb{F}_q^d \to \mathbb{K}$ be an $\mathbb{F}_q$-vector space isomorphism. Let $U : \mathbb{F}_q^n \to \mathbb{F}_q^n$ be an invertible $\mathbb{F}_q$-affine map, let $L_t : \mathbb{F}_q^n \to \mathbb{F}_q^d$ be a corank t embedding, let $S : \mathbb{F}_q^n \to \mathbb{F}_q^s$ be the projection onto the first s coordinates and let $T : \mathbb{F}_q^{d+p} \to \mathbb{F}_q^{d+p-a}$ be a full rank $\mathbb{F}_q$-affine map.

Define $f_c : \mathbb{F}_q^n \to \mathbb{F}_q^d$ by

$$f_c = \phi^{-1} \circ f \circ \phi \circ L_t,$$

where $f : \mathbb{K} \to \mathbb{K}$ is given by $f(X) = X^{q^\theta+1}$. Further, let $f_{ip} : \mathbb{F}_q^n \to \mathbb{F}_q^d$ be given by

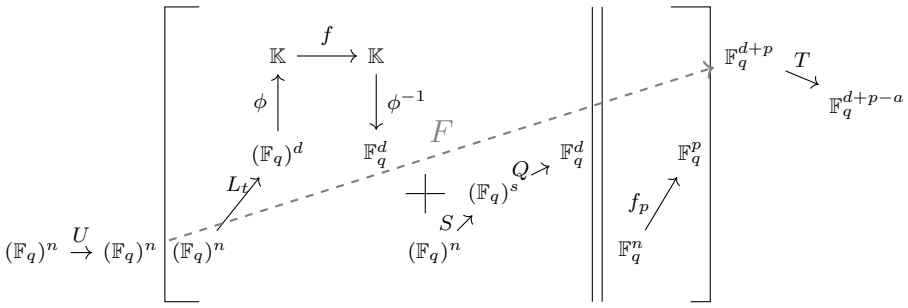
$$f_{ip} = Q \circ S,$$

where $Q : \mathbb{F}_q^s \to \mathbb{F}_q^d$ is a random quadratic. Finally, let $f_p : \mathbb{F}_q^n \to \mathbb{F}_q^p$ be a random quadratic.

The central map is constructed as

$$F = (f_c + f_{ip}) \| f_p.$$

The public key is then given by $\mathscr{P} = T \circ F \circ U$, illustrated below.



Remark 1. *We note that since ϕ is a parameter of the system, that the choice of ϕ amounts to a choice of basis for $\mathbb{K}$ over $\mathbb{F}_q$ and that L_t is random, the choice of defining f_{ip} using only the first s variables still captures the full generality of the (ip) modifier.*

The degree of the extension field d is a parameter that can vary widely depending on the application. For most studied signature schemes of this type, $d = n$. For encryption, it is viable to allow d to be much larger than n, as is the case for EFLASH, see [44].

Inversion of the public map given the private key is straightforward. First, at key generation, compute the image of f_{ip} so that its elements can be efficiently enumerated. Given $\mathbf{y} \in \mathbb{F}_q^{d+p-a}$, compute the preimage of $\mathbf{y}$ under T and parse each vector as $\mathbf{w} = \mathbf{w}_1 \| \mathbf{w}_2 = (w_1, \ldots, w_d) \| (w_{d+1}, \ldots, w_{d+p})$. For all elements $\mathbf{w}_s \in \mathrm{Im}(f_{ip})$ and for all such $\mathbf{w}_1$, enumerate all preimages $\mathbf{u}$ of $\mathbf{w}_1 + \mathbf{w}_s$ under f_c. Check that $f_{ip}(\mathbf{u}) = \mathbf{w}_s$ and that $f_p(\mathbf{u}) = \mathbf{w}_2$: if either fails retry with another pair; if the check succeeds then output $U^{-1}(\mathbf{u})$. All of the details are explicitly provided in Algorithm 1, which calls on a subroutine to invert the map f_c, since this process can differ depending on whether the scheme is parameterized as a signature or an encryption scheme.

7.3 Security Analyses of the Schema

Here we consider the known attacks on multivariate schemes and discuss

Algorithm 1 (n, d, s, a, p, t) scheme Public Key Inversion

Input: $T, \phi, f_c, f_{ip}, f_p, U^{-1}, \mathrm{Im}(f_{ip})$ and $\mathbf{y} \in \mathbb{F}_q^{d+p-a}$.
Output: $\mathbf{x}$ such that $\mathscr{P}(\mathbf{x}) = \mathbf{y}$.

$\mathbf{x} = \perp$
$W_T \leftarrow \{\mathbf{w} \in \mathbb{F}_q^{d+p} : T(\mathbf{w}) = \mathbf{y}\}$
for all $\mathbf{w} \in W_T$ **do**
  $\mathbf{w}_1 \leftarrow \mathbf{w}[1:d]$, $\mathbf{w}_2 \leftarrow \mathbf{w}[d+1:d+p]$
  for all $\mathbf{w}_s \in \mathrm{Im}(f_{ip})$ **do**
    $\mathbf{w}_c \leftarrow \mathbf{w}_1 + \mathbf{w}_s$
    $U_c \leftarrow \mathrm{Inv} f_c(\mathbf{w}_c)$
    for all $\mathbf{u} \in U_c$ **do**
      if $f_{ip}(\mathbf{u}) = \mathbf{w}_s$ **and** $f_p(\mathbf{u}) = \mathbf{w}_2$ **then**
        $\mathbf{x} \leftarrow U^{-1}(\mathbf{u})$
        break
      end if
    end for
  end for
end for
return $\mathbf{x}$

their application on the general C^* framework presented in the previous section. In particular, we discuss what combinations of modifiers prevent attacks and highlight what modifiers are vulnerable to attack without a companion modifier.

7.3.1 Differential Analysis

Differential symmetry attacks also broke SFLASH, SQUARE and PMI, see [28, 37, 38]. The vulnerability of C^* to the differential symmetric attack is provably removed with the combination of the projection and minus modifiers as shown in [5, 17] and generalized in [42]. Thus, for any generalized C^* scheme with both the projection and minus modifiers to be vulnerable to such attacks, the additional modifiers must somehow reintroduce this weakness. Since the remaining modifiers introduce random coefficients or equations to the central map, the likelihood of such an occurrence is very remote. In particular, both of the random (ip) and (+) modifiers move the distribution of $(n, d, 0, a, 0, t)$ schemes towards the uniform

distribution on quadratic maps in statistical distance. Under the loose heuristic that the distribution of such (n, d, s, a, p, t) schemes is very close to uniform, we can derive an approximation of the probability that the result has a differential symmetry.

To calculate the probability that a random function has a differential symmetry, we note that quantified over all quadratic functions, we obtain all possible differential symmetry relations of the form

$$DP(Ma, x) + DP(a, Mx) = \Lambda_M DP(a, x). \tag{7.1}$$

in the unknown coefficients of M and Λ_M. We can see that M is being applied to the plaintext variables a and x, which are vectors of length n. Therefore, M will be an $n \times n$ matrix. Lambda on the other hand, will be applied to the output variables, and will therefore have dimension $d - a + p \times d - a + p$. So the two matrices together will have $(d - a + p)^2 + n^2$ variables. To determine the number of equations, it is easier to think about this relationship coordinate wise. If we denote DP_i as the differential of the i^{th} public key equation, and consider solving $DP_i(My, x) + DP_i(y, Mx) = \Lambda_{M_i} DP(y, x)$, we can see that we will have a set of $d - a + p$ equations. For each coordinate i, we get one equation. But, we can interpret each equation as a matrix equation, $M^\top \mathbf{DP_i} + \mathbf{DP_i} M = \Lambda_{Mi} \mathbf{DP}$. The left hand side of the equation will give me a symmetric matrix as $\mathbf{DP_i}$ is symmetric, so we only need to test the upper triangular section of the matrix, including the diagonal, which gives us $\binom{n+1}{2}$ coordinates. Thus, there are $(d - a + p)$ times $\binom{n+1}{2}$ equations. Therefore the symmetric discrete differential equation has a solution with probability $q^{(d-a+p)^2+n^2-(d-a+p)\binom{n}{2}} = q^{\mathcal{O}(-n^3)}$.

For the sake of caution, we will consider schemes with at most one of the projection and minus modifiers and at most one of the (ip) and (+) modifiers to be insecure. The reason is that schemes with at most one of (p) and (-) exhibit a

form of differential symmetry and it is plausible that statistical techniques similar to [38] or projection techniques similar to [43] may be effective against (ip) and (+) modifications of these schemes, respectively.

Thus the space of (n, d, s, a, p, t) schemes resistant to differential attacks are those with $at > 0$ or $sp > 0$.

7.3.2 MinRank

The complexity of MinRank is dependent upon Q-rank, so it is clear that if we increase the rank, we decrease the effectiveness of the MinRank attack. The minus modifier increases the rank of the central map. Since the quadratic form associated with the central map is so sparse, the removal of one equation in general increases the rank by two. The minus modifier can be viewed as a codimension one projection, which can produce two elements in original rows and columns, see Figure 7.1.

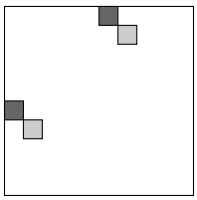


Figure 7.1: The shape of the matrix representing the central map of C^{*-}. The darkly shaded regions represent nonzero values of the central map without the minus modifier, the lightly shaded regions represent new nonzero values introduced by the removal of one equation. Unshaded areas have coefficients of zero.

The projection modifier cannot increase the Q-rank of the central map. Since we expect the Q-rank of the projected scheme to be the same as the original, we expect the security of a (n, d, s, a, p, t) scheme against MinRank attacks to be the same as that of a $(n, d, s, a, p, 0)$ scheme.

The effect of (ip) depends on the support dimension s. Since the support dimension of the (ip) summand is s, there is an s dimensional space in which the (ip) modifier adds extra randomness to the quadratic form $\mathbf{F}$. Under an appropriate change of basis, all of this contribution can be contained in the upper left $s \times s$ block of $\mathbf{F}$; hence, (ip) increases the Q-rank by at most s. We can apply projections on the input to try to kill the (ip) support, which once again makes the Q-rank low. Therefore, techniques similar to [43] may limit the effectiveness of the (ip) modifier in preventing a MinRank attack, a fact that once again highlights the close relationship between (ip) and (v).

Finally, the (+) modifier has no effect on min-Q-rank. While any nonzero linear combination of the (+) polynomials is likely to have a high rank, any linear combination of the public polynomials that eliminates the contribution of the (+) polynomials has a rank independent of the (+) modifier. Thus, if the Q-rank of a $(n, d, s, a, 0, t)$ scheme is sufficiently low, the (+) modifier adds no significant security; it merely increases the number of equations.

Thus, the Q-rank of a generic (n, d, s, a, p, t) scheme is $2 + 2a + s$ at most, and this is a tight bound for most realistic parameters. Thus, we can conclude that the complexity of a MinRank attack on (n, d, s, a, p, t) schemes is

$$\mathcal{O}\left(\binom{d+a+p+s+3}{3+2a+s}^{\omega}\right) \approx \mathcal{O}\left((d-a+p)^{(3+2a+s)\omega}\right).$$

7.3.3 Algebraic

The complexity of algebraic attacks rely on the degree of regularity, defined as the smallest degree at which a nontrivial degree fall is generated in the Gröbner basis algorithm. We can find an estimate of the degree of regularity depending on the Q-rank of the system in [22].

As discussed in the previous section, the Q-rank of a C^* scheme with a

equations removed is $2 + 2a$. The projection and plus modifier cannot increase the rank of the central map. Thus we compute an upper bound on the degree of regularity on a (n, d, s, a, p, t) scheme to be

$$d_{reg} \leq (q-1)(a+s+1) + 2.$$

We note that d_{reg} monotonically decreases in p, thus we cannot add too many plus polynomials.

If we are more conservative, we may consider projection attacks such as those in [43] attempting to eliminate some of the support for the (ip) modifier. We note that with appropriately chosen parameters, this attack can be rendered no more effective than the standard algebraic attack.

Therefore, similar to the analysis done in [44], the complexity of the algebraic attack on a (n, d, s, a, p, t) scheme is estimated to be $\mathcal{O}\left(\binom{n+d_{reg}}{d_{reg}}^{\omega}\right)$, where $2 \leq \omega \leq 3$ is the linear algebra constant.

7.4 Performance Analyses

The performance characteristics of C^* schemes under various modifiers are straightforward to derive. There are some significant differences, however, in the effects of the modifiers when employed for signatures versus encryption. We, thus, treat each case independently.

7.4.1 Key Size

For a generic (n, d, s, a, p, t) scheme, the derivation of the public key size is straightforward. There are $d - a + p$ public equations in n variables. Each of those equations will be a quadratic function with $\binom{n}{2}$ possible quadratic terms (recall, $x^2 = x$ in $\mathbb{F}_2$), n linear terms, and 1 monomial. So each equation will have $\binom{n}{2} + n + 1$ elements. The public key therefore consists of $(d - a + p)\left(\binom{n+1}{2} + 1\right)$

elements from $\mathbb{F}_q$. This observation is not all there is to say about public key size, however.

The (+) modifier significantly reduces the probability of the existence of a preimage for an arbitrary element in the codomain of the public key. We may therefore safely ignore this modifier for digital signature applications.

Similarly, there is a possibility of inversion failure for the public key if the balance of the projection, minus and plus modifiers is not correctly handled. In particular, if d is larger than n the chance of failure is high. Therefore, we will restrict our consideration in the case of signature schemes to $(n, n, s, a, 0, t)$ with t small and in particular $t < a$.

For encryption schemes we need the public key to be statistically injective. To accomplish this goal with modifiers, we either require some redundancy in the plaintext space or to have a larger codomain than the domain. To allow random plaintexts, we use a value of d much larger than n. Then inherently a projection of corank at least $t = d - n$ is required. Since a larger value of t precludes unique inversion, we are forced to set the parameter $t = d - n$.

7.4.2 Complexity of Inversion

Inversion of the public key is accomplished with Algorithm 1. From the algorithm we can see that we have as many as q^{a+s} calls to $\mathrm{Inv} f_c$, plus as many as $q^{a+s}|\ker(L_t)|$ evaluations of f_{ip} and f_p, plus a couple of linear algebra operations. These numbers can vary depending on the parameters and the application.

Consider a $(n, d, s, a, 0, t)$ scheme designed for signatures. With $n = d$, any preimage of $\mathbf{y}$ has a high probability of producing a valid signature and so it is likely that very few of the q^a such preimages will need to be utilized in the inversion. In contrast, for an encryption scheme with n far smaller than d, one may have to

search nearly the entire q^a preimages to find the valid plaintext. Similarly, for a signature scheme, the kernel of L_t may be large, as large as q^t, possibly, whereas for an encryption scheme it is necessary that L_t is an embedding. Thus, we split into two cases to consider inversion complexity.

First, we consider inversion for signatures. Under the above assumptions for valid performance of a (n, d, s, a, p, t) scheme for signatures, we have that $n = d$, $p = 0$, t is small and $t < a$. Since t is small and evaluation of f_{ip} and f_p are extremely efficient, the complexity is dominated by that of $\mathrm{Inv} f_c$. Since we expect to only require a few values of $\mathbf{w}$ to find a valid inverse, the complexity of this step is $\mathscr{O}(q^s)$ times the complexity of $\mathrm{Inv} f_c$. Using an efficient inversion process based on linearization equations, the complexity of the latter is $\mathscr{O}(n^\omega)$. Thus the complexity for signature schemes is

$$\mathscr{O}\left(n^\omega q^s\right).$$

In the case of encryption, we assume that d is much larger than n and that $t = d - n$. In this case, all of $\mathrm{Inv} f_c$ and f_{ip} and f_p are evaluated q^{a+s} times, thus, once again, the complexity of inversion is dominated by that of $\mathrm{Inv} f_c$. In this case, however, it is likely that on the order of q^a preimages of $\mathbf{y}$ under T need to be searched. Thus the complexity of inversion for an encryption scheme is

$$\mathscr{O}\left(n^\omega q^{a+s}\right).$$

7.4.3 Decryption Failure Rate

The decryption failure rate for $(n, d, 0, a, 0, t)$ schemes created with the intention of encryption is discussed in [44]. Using conditional probabilities and Bernoulli trials, it was found that the probability, p, that y is the image of at least two distinct elements of $\mathbb{F}_q^n$, given that it is the image of at least one can be approximated by q^{n-m}, where $m = d - a$.

This probability was found under the simplifying heuristic that considered the $\mathbb{F}_q-$quadratic central maps f to be random injective maps. Considering (ip), we no longer have the injective property, though the (ip) modifier has a large codomain, so it is unlikely to have collisions between it and anything else. Therefore, (ip) should not have an effect on the decryption failure rate because we would just be adding another random summand. Adding a plus modifier will decrease the decryption failure rate because it will add extra equations that need to be satisfied.

We can model the central map without the minus modifier as a random function $G : q^n \to q^{\widehat{m}}$, where $\widehat{m} := d + p - a$. We will again use Bernoulli trials, and compute: $P(|G^{-1}(y)| \geq 2 \mid |G^{-1}(y)| \geq 1) \leq \frac{1-(1-q^{n-\widehat{m}})-q^{n-\widehat{m}}+q^{2n-2\widehat{m}}-q^{n-2\widehat{m}}}{1-(1-q^{n-d}+\frac{1}{2}(q^n(q^{n-1})q^{-2d}))}$. Following an analysis equivalent to that in [44], we see that the numerator is bounded by $q^{2n-2\widehat{m}}$ while the denominator is very close to $q^{n-\widehat{m}}$; thus, a good approximation is about $q^{n-\widehat{m}} = q^{n+a-d-p}$. We can see that the plus modifier reduces the probability of decryption failure approximately by q^p, while (ip) has no significant effect on the failure rate.

We performed a series of experiments on failure rate for (n, d, s, a, p, t) schemes designed for encryption to investigate the rate of decryption failure relative to varying parameters d and p, the degree of extension and number of plus polynomials. The results are reported in Table 7.1. All experiments were performed by encrypting all possible plaintexts and counting the number of plaintexts producing non-unique ciphertexts. In every experiment the failure rates follow the above analysis closely without significant variance.

q	n	d	a	s	p	$\widehat{m}$	$n - \widehat{m}$	decryption failure rate
2	14	26	4	2	2	24	-10	$2^{-9.62}$
2	14	26	4	2	3	25	-11	$2^{-10.71}$
2	14	27	4	2	2	25	-11	$2^{-10.86}$
2	14	27	4	2	3	26	-12	$2^{-12.23}$
2	14	28	4	2	2	26	-12	$2^{-11.68}$
2	14	28	4	2	3	27	-13	$2^{-13.23}$

Table 7.1: Probability of decryption failure for specific parameters of a (n, d, s, a, p, t) scheme.

7.4.4 Parameter Spaces for Encryption and Signatures

From Section 7.3, we obtain the following constraints for a 128-bit secure (n, d, s, a, p, t) scheme.

$$at + sp > 0$$
$$(d - a + p)^{(2+2a+s)\omega} \geq 2^{128}$$
$$n^{((q-1)(a+s+1)+2)\omega} \geq 2^{128}.$$

These constraints assure security against differential, MinRank and algebraic attacks, respectively.

For signature schemes with suggested parameters of the form $(n, d, s, a, 0, t)$ we obtain a public key size of at least $(n - a)\left(\binom{n+1}{2} + 1\right) lg(q)$ bits and a signing time on the order of $n^\omega q^s$ field operations. Both of these quantities seem to be optimized by making $s = 0$, having a fairly large, and having $t = 1$. This choice of parameters, unsurprisingly, produces the PFLASH scheme. It is interesting to note that these data also suggest an optimal choice of q for such schemes of 2 or 4.

For encryption schemes of the form (n, d, s, a, p, t), the public key size has the form $(d - a + p)\left(\binom{n+1}{2} + 1\right) lg(q)$ bits while decryption time is around $n^\omega q^{a+s}$ and

the decryption failure rate is $q^{n+a-d-p}$. Here there is a much more interesting trade-off between different strategies. The quantity $a + s$ needs to remain sufficiently large to provide security but directly impacts the decryption speed. Public key size is reduced if a is increased while s and p are reduced , however this directly and negatively impacts decryption failure rate. Thus there are an array of options offering various optimizations.

REFERENCES

[1] T. Matsumoto and H. Imai. Public quadratic polynomial-tuples for efficient signature verification and message-encryption. *Eurocrypt '88, Springer*, 330:419–545, 1988.

[2] Jacques Patarin. Cryptoanalysis of the Matsumoto and Imai Public Key Scheme of Eurocrypt'88. In Don Coppersmith, editor, *CRYPTO*, volume 963 of *Lecture Notes in Computer Science*, pages 248–261. Springer, 1995.

[3] Jacques Patarin, Louis Goubin, and Nicolas Courtois. C^*_{-+} and HM: variations around two schemes of t. matsumoto and h. imai. In Kazuo Ohta and Dingyi Pei, editors, *Advances in Cryptology - ASIACRYPT '98, International Conference on the Theory and Applications of Cryptology and Information Security, Beijing, China, October 18-22, 1998, Proceedings*, volume 1514 of *Lecture Notes in Computer Science*, pages 35–49. Springer, 1998.

[4] V. Dubois, P.-A. Fouque, A. Shamir, and J. Stern. Practical cryptanalysis of SFLASH. *Advances in Cryptology - CRYPTO 2007, Springer*, 4622:1–12, 2007.

[5] M.-S. Chen, B.-Y. Yang, and D. Smith-Tone. Pflash - secure asymmetric signatures on smart cards. Lightweight Cryptography Workshop 2015, 2015. http://csrc.nist.gov/groups/ST/lwc-workshop2015/papers/session3-smith-tone-paper.pdf.

[6] Jacques Patarin, Louis Goubin, and Nicolas Courtois. Improved algorithms for isomorphisms of polynomials. In Kaisa Nyberg, editor, *Advances in Cryptology*

- EUROCRYPT '98, International Conference on the Theory and Application of Cryptographic Techniques, Espoo, Finland, May 31 - June 4, 1998, Proceeding, volume 1403 of *Lecture Notes in Computer Science*, pages 184–200. Springer, 1998.

[7] Jacques Patarin. Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms. In *EUROCRYPT*, pages 33–48, 1996.

[8] E. R. Berlekamp. Factoring polynomials over large finite fields. *Mathematics of Computation*, 24(111):pp. 713–735, 1970.

[9] Jeremy Vates and Daniel Smith-Tone. Key recovery attack for all parameters of HFE-. In Lange and Takagi [45], pages 272–288.

[10] A. Kipnis and A. Shamir. Cryptanalysis of the hfe public key cryptosystem by relinearization. *Advances in Cryptology - CRYPTO 1999, Springer*, 1666:788, 1999.

[11] Luk Bettale, Jean-Charles Faugère, and Ludovic Perret. Cryptanalysis of hfe, multi-hfe and variants for odd and even characteristic. *Des. Codes Cryptography*, 69(1):1–52, 2013.

[12] Taylor Daniels and Daniel Smith-Tone. Differential properties of the HFE cryptosystem. In Mosca [46], pages 59–75.

[13] Koblitz. *Algebraic Aspects of Cryptography.* Springer-Verlag, 2004.

[14] Bruno Buchberger. Bruno Buchberger's Phd Thesis 1965: An Algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal. *Journal of Symbolic Computation*, 2006.

[15] J. C. Faugere. A new efficient algorithm for computing grobner bases (f4). *Journal of Pure and Applied Algebra*, 139:61–88, 1999.

[16] V. Dubois, P.-A. Fouque, and J. Stern. Cryptanalysis of SFLASH with slightly modified parameters. *Eurocrypt ?07, Springer*, 4515:264–275, 2007.

[17] Daniel Smith-Tone. On the differential security of multivariate public key cryptosystems. In Bo-Yin Yang, editor, *PQCrypto*, volume 7071 of *Lecture Notes in Computer Science*, pages 130–142. Springer, 2011.

[18] Daniel Smith-Tone. Properties of the discrete differential with cryptographic applications. In Nicolas Sendrier, editor, *PQCrypto*, volume 6061 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2010.

[19] Yang Tao Ding Petzoldt, Chen. Design principles for HFEv- based multivariate signature schemes. *ASIACRYPT 2015. LNCS*, 9452:311–334, 2015.

[20] Ryann Cartor, Ryan Gipson, Daniel Smith-Tone, and Jeremy Vates. On the differential security of the hfev- signature primitive. In Tsuyoshi Takagi, editor, *Post-Quantum Cryptography - 7th International Workshop, PQCrypto 2016, Fukuoka, Japan, February 24-26, 2016, Proceedings*, volume 9606 of *Lecture Notes in Computer Science*, pages 162–181. Springer, 2016.

[21] J. C. Faugere. Algebraic cryptanalysis of hidden field equations (hfe) using grobner bases. *CRYPTO 2003, LNCS*, 2729:44–60, 2003.

[22] Jintai Ding and Thorsten Kleinjung. Degree of regularity for HFE-. *IACR Cryptology ePrint Archive*, 2011:570, 2011.

[23] Christopher Wolf and Bart Preneel. Equivalent keys in multivariate quadratic public key systems. *J. Mathematical Cryptology*, 4(4):375–415, 2011.

[24] Jintai Ding, Vivien Dubois, Bo-Yin Yang, Chia-Hsin Owen Chen, and Chen-Mou Cheng. Could SFLASH be Repaired? In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, *ICALP (2)*, volume 5126 of *Lecture Notes in Computer Science*, pages 691–701. Springer, 2008.

[25] Dustin Moody, Ray A. Perlner, and Daniel Smith-Tone. An asymptotically optimal structural attack on the ABC multivariate encryption scheme. In Mosca [46], pages 180–196.

[26] Dustin Moody, Ray A. Perlner, and Daniel Smith-Tone. *Key Recovery Attack on the Cubic ABC Simple Matrix Multivariate Encryption Scheme*. Springer, 2017.

[27] Jeremy Vates and Daniel Smith-Tone. Key recovery attack for all parameters of hfe-. In Current Submission, 2017.

[28] O. Billet and G. Macario-Rat. Cryptanalysis of the square cryptosystems. *ASIACRYPT 2009, LNCS*, 5912:451–486, 2009.

[29] Chengdong Tao, Adama Diene, Shaohua Tang, and Jintai Ding. Simple matrix scheme for encryption. In Philippe Gaborit, editor, *PQCrypto*, volume 7932 of *Lecture Notes in Computer Science*, pages 231–242. Springer, 2013.

[30] Jaiberth Porras, John Baena, and Jintai Ding. Zhfe, a new multivariate public key encryption scheme. In Mosca [46], pages 229–245.

[31] Takanori Yasuda and Kouichi Sakurai. *A Multivariate Encryption Scheme with Rainbow*, pages 236–251. Springer International Publishing, Cham, 2016.

[32] Dustin Moody, Ray A. Perlner, and Daniel Smith-Tone. Improved attacks for characteristic-2 parameters of the cubic ABC simple matrix encryption scheme.

In *Post-Quantum Cryptography - 8th International Workshop, PQCrypto 2017, Utrecht, The Netherlands, June 26-28, 2017, Proceedings*, pages 255–271, 2017.

[33] Daniel Cabarcas, Daniel Smith-Tone, and Javier A. Verbel. Key recovery attack for ZHFE. In Lange and Takagi [45], pages 289–308.

[34] Ray A. Perlner, Albrecht Petzoldt, and Daniel Smith-Tone. *Total Break of the SRP Encryption Scheme*. Springer, In press., 2017.

[35] Magali Bardet, Jean-Charles Faugere, and Bruno Salvy. On the complexity of gröbner basis computation of semi-regular overdetermined algebraic equations. In *Proceedings of the International Conference on Polynomial System Solving*, 2004.

[36] M Bardet, JC Faugére, B Salvy, and BY Yang. Asymptotic behaviour of the degree of regularity of semi-regular polynomial systems. *Proc. of MEGA 2005, Eighth International Symposium on Effective Methods in Algebraic Geometry*, 2005.

[37] Vivien Dubois, Pierre-Alain Fouque, Adi Shamir, and Jacques Stern. Practical Cryptanalysis of SFLASH. In Alfred Menezes, editor, *CRYPTO*, volume 4622 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2007.

[38] Pierre-Alain Fouque, Louis Granboulan, and Jacques Stern. Differential cryptanalysis for multivariate schemes. In *Advances in Cryptology - EUROCRYPT 2005, 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 22-26, 2005, Proceedings*, pages 341–353, 2005.

[39] Vivien Dubois, Pierre-Alain Fouque, and Jacques Stern. Cryptanalysis of SFLASH with Slightly Modified Parameters. In Moni Naor, editor, *EURO-*

CRYPT, volume 4515 of *Lecture Notes in Computer Science*, pages 264–275. Springer, 2007.

[40] J. Ding. A new variant of the matsumoto-imai cryptosystem through perturbation. *PKC 2004, LNCS*, 2947:305–318, 2004.

[41] J. Ding and J. Gower. Inoculating multivariate schemes against differential attacks. *PKC 2006, LNCS*, 3958:290–301, 2006.

[42] Ryann Cartor and Daniel Smith-Tone. An updated security analysis of PFLASH. In *Post-Quantum Cryptography - 8th International Workshop, PQCrypto 2017, Utrecht, The Netherlands, June 26-28, 2017, Proceedings*, pages 241–254, 2017.

[43] Jintai Ding, Ray A. Perlner, Albrecht Petzoldt, and Daniel Smith-Tone. Improved cryptanalysis of hfev- via projection. In *Post-Quantum Cryptography - 9th International Conference, PQCrypto 2018, Fort Lauderdale, FL, USA, April 9-11, 2018, Proceedings*, pages 375–395, 2018.

[44] Ryann Cartor and Daniel Smith-Tone. EFLASH: A new multivariate encryption scheme. In *Selected Areas in Cryptography - SAC 2018 - 25th International Conference, Calgary, AB, Canada, August 15-17, 2018, Revised Selected Papers*, pages 281–299, 2018.

[45] Tanja Lange and Tsuyoshi Takagi, editors. *Post-Quantum Cryptography - 8th International Workshop, PQCrypto 2017, Utrecht, The Netherlands, June 26-28, 2017, Proceedings*, volume 10346 of *Lecture Notes in Computer Science*. Springer, 2017.

[46] Michele Mosca, editor. *Post-Quantum Cryptography - 6th International Workshop, PQCrypto 2014, Waterloo, ON, Canada, October 1-3, 2014. Proceedings*, volume 8772 of *Lecture Notes in Computer Science*. Springer, 2014.

APPENDIX I

HFEv and HFEv^- Key Check Algorithms

Algorithm 2 HFEvKeyCheck

Input: An HFEv central map f, a flag flg.

Output: Set of indices of coefficients m_i of submatrix m_{00} which are possibly nonzero in a linear map inducing differential symmetry for f.

for monomial $\alpha_{i,j}x^{q^i+q^j}$ in f **do**
  $S_i = \{\}$;
  $S_j = \{\}$;
  for monomial with powers r and s in f **do**
    $S_i = S_i \cup \{r-j, s-j, i-j+r-s, i-j+s-r\}$;
    $S_j = S_j \cup \{r-i, s-i, j-i+r-s, j-i+s-r\}$;
  end for
end for
if flg **then**
  return all S_i
else
  return $\cap S_i$
end if

Algorithm 3 HFEv^-KeyCheck

Input: An HFEv^- central map $\pi(f)$, the corank of π, r.

Output: Set of indices of coefficients m_i of submatrix m_{00} which are possibly nonzero in a linear map inducing differential symmetry for $\pi(f)$.

Call: HFEvKeyCheck(f,1)
for all S_i **do**
  $T_i = \{\}$;
  for j from 0 to $r-1$ **do**
    $T_i = T_i \cup (j + S_i)$;
  end for
end for
return $\cap T_i$

APPENDIX II

C^* Security

Modifier	Algebraic	Differential	Linearization Equation	MinRank
(ip)	+	0	+	+
(-)	+	0	+	+
(+)	-	+	+	0
(p)	-	0	0	0
(v)*	+	+	+	+
(ip+)	-	+	+	+
(+-)	+	+	+	+
(p-)	+-	1	+	+

Table II.1: Resistance of C^* against attacks under certain modifiers. The table can be read as probabilities of resistance to the given attack. Thus 0 means that the modifier(s) provide(s) no security in the attack model, 1 means the modifiers(s) provide(s) provable security, and + or - mean increases, respectively decreases in security. * These schemes use HFEv inversion and are not C^* schemes *per se*.

APPENDIX III

Glossary

- **Algebra.** An *algebra* over a field is a vector space along with a bilinear product.
- **Algebraic Variety.** An *algebraic variety* is the set of solutions of a set of polynomial equations.
- **Basis.** The *basis* of a vector space is a set of linearly independent elements that span the vector space.
- **Codimension** The codimension of $W \subseteq V$ is a V is a vector space is $codim(W) = dim(V) - dim(W)$.
- **Corank.** If an $m \times n$ matrix has rank r, then the *corank* of the matrix is $m - r$.
- **Coset.** For any $N \leq G$ and any $g \in G$ where G is a group, let $gN = \{gn | n \in N\}$ and $Ng = \{ng | n \in N\}$ called respectively a left *coset* and a right *coset* of N in G.
- **Degree of Regularity.** Although there are many ways to define the degree of regularity for a set of polynomials, this work will be using the following definition: the *degree of regularity* for a system of equations is the degree of the first degree fall while using the Buchberger algorithm to the compute Gröbner basis.

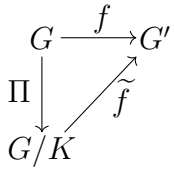
- **Dimension.** The *dimension* of a vector space is the number of elements in the basis of the vector space.

- **Extension field.** If $\mathbb{F}$ is a field, then $\mathbb{K}$ is an *extension field* of $\mathbb{F}$ if $\mathbb{K}$ is a field and $\mathbb{F} \subset \mathbb{K}$.

- **Field.** A *field* is a set $\mathbb{F}$ along with operations multiplication and addition that satisfy the following properties: associativity, commutativity, the distributive law, existence of additive identity 0, existence of multiplicative identity 1, additive inverses, and multiplicative inverses for everything except 0.

- **First Isomorphism Theorem.** Let $f : G \to G'$ be a surjective homomorphism with kernel K.

  1. The map $\widetilde{f} : G/K \to G'$, defined by $\widetilde{f}(xk) = f(x)$ for every $x \in G$ is well defined
  2. The map $\widetilde{f}$ is an isomorphism
  3. Let $\Pi : G \to G/K$ be the natural map. Then $\widetilde{f} \circ \Pi = f$, i.e., the following diagrams commute



- **Graph Isomorphism Problem (GI).** The *Graph Isomorphism Problem* is the problem of determining whether two finite graphs are isomorphic.

- **Group.** A *group* consists of set and an operation such that the operation satisfies closure, associativity, identity and invertibility.

- **Hilbert's Nullstellensatz:** Let $\mathbb{F}$ be a field and $\mathbb{K}$ and algebraically closed extension field. Let I be an ideal of the polynomial ring $\mathbb{F}[X_1, \ldots, X_n]$, and let

$V(I)$ be the algebraic set of the ideal, defined such that $\forall \mathbf{x} \in V(I)$, $f(\mathbf{x}) = 0$, for every $f \in I$. If some polynomial $p \in \mathbb{F}[X_1, \ldots, X_n]$ vanishes on $V(I)$ (meaning $p(\mathbf{x}) = 0$ for all $\mathbf{x} \in V(I)$), then there exists a natural number r such that $p^r \in I$.

- **Ideal.** Let R be a ring. A subset I of R is called a left (respectively right or 2 sided) *ideal* of R if the following conditions hold:

  1. $0 \in I$
  2. $x, y \in I \Rightarrow x + y \in I$
  3. $r \in R, x \in I \Rightarrow rx \in I$

- **Irreducible Polynomial.** A nonconstant polynomial f is said to be *irreducible* over a field $\mathbb{F}$ if it cannot be factored into a polynomial of lower degree.

- **Minimal polynomial.** The *minimal polynomial* of an element α is the polynomial f of lowest degree such that $f(\alpha) = 0$.

- **Monic.** A polynomial is said to be *monic* if the leading coefficient is 1.

- **Nonsingular.** A square matrix is *nonsingular* if the matrix has a multiplicative inverse (i.e., is invertible). A square matrix is nonsingular when the determinate is nonzero

- **Normal Subgroup.** A subgroup $H \subseteq G$ is normal if the left and right cosets of H in G are equal, i.e., $gH = Hg$ for all $g \in G$

- NP**,** NP**-complete.** A decisional problem belongs to the *class* NP if, given a witness (an example where the problem's answer is yes), we can check that the answer is correct in polynomial time. Stands for "non-deterministic polynomial time". If you have a problem where a witness can be checked in

polynomial time, and a known NP problem can be solved using the given problem with modified input, then your problem is *NP-complete*.

- **P-space.** A decisional problem belongs to the class P if there is a polynomial-time algorithm that solves any solves any instance of the problem in polynomial time. Note, $P \subseteq NP$.

- **Polynomial ring.** A *polynomial ring* over a field $\mathbb{F}$ in the variables $[X_1, X_2, \ldots, X_n]$, denoted $\mathbb{F}[X]$, consists of all finite sums of products of powers of $X_1, \ldots, X_m$ with coefficients in $\mathbb{F}$.

- **Primitive Element.** If the field $\mathbb{K}$ is generated by a single element α over $\mathbb{F}$, $\mathbb{K} = \mathbb{F}(\alpha)$, then α is the *primitive element* of that extension.

- **Projection.** A *projection* on a vector space V is a linear operator $P : V \mapsto V$ such that $P^2 = P$ (i.e., P is idempotent).

- **Quotient Group.** Let N be a normal subgroup of G. We define the set G/N to be the set of all left cosets of N in G, i.e., $G/N = \{aN : a \in G\}$. We define an operation on G/N as $(aN)(bN) = (ab)N$. The set G/N together with the defined operation forms the *quotient group* of G by N.

- **Ring.** A *ring* is a set R along with two binary operators $+$ and $\times$ that satisfy additive associativity, additive commutativity, additive identity, additive inverse, distributivity, and multiplicative associativity.

- **Singular.** A square matrix is *singular* if the matrix has no multiplicative inverse (i.e., is not invertible). Thus, a square matrix is singular when the determinate is equal to zero.

- **Standard Basis.** The *standard basis* of the space containing $K-$dimensional vectors is a basis composed of vectors that have one entry equal to 1 and the

remaining $K - 1$ entries equal to 0.

- **Support.** The *support* of a function $f : A \to B$ is the set of elements $x \in A$ such that $f(x) \neq 0$.

- **Vector Space.** A *vector space* is a set that is closed under finite vector addition and scalar multiplication.

CURRICULUM VITAE

Ryann Cartor
University of Louisville
ryann.cartor@louisville.edu

EDUCATION

University of Louisville, Louisville, KY
Ph. D, Applied Mathematics August 2019
M.A., Applied Mathematics December 2016

Bellarmine University, Louisville, KY.
B.S., Mathematics (Summa Cum Laude) May 2014

PROFESSIONAL EXPERIENCE

Teaching Assistant, University of Louisville 2014 - 2019

Courses independently taught include:

- Elementary Statistics (Math 109)
- College Algebra (Math 111)
- Math for Elementary Education I (Math 151)
- Math for Elementary Education II (Math 152)
- Elements of Calculus (Math 180)
- Precalculus (Math 190)
- Calculus I (Math 205)
- Calculus II (Math 206)

Courses recitation instructor for include:

- Contemporary Mathematics (Math 105)
- Finite Mathematics (Math 107)
- College Algebra (Math 111)
- Elements of Calculus (Math 180)

Adjunct Professor, Bellarmine University Spring 2018

Courses independently taught:

- Math for Liberal Arts (Math 107)

RESEARCH

Publications

Cartor, R., Smith-Tone, D. EFLASH: A New Multivariate Encryption Scheme. 25th Conference on Selected Areas in Cryptography, Calgary, Canada, August 15-17, 2018, Proceedings, Springer (2018)

Cartor, R., Smith-Tone, D. An Updated Security Analysis of PFLASH. 8th International Workshop, PQCrypto 2017, Utrecht, the Netherlands, June 26-28, 2017, Proceedings, Springer (2017)

Cartor, R., Gipson, R., Smith-Tone, D., Vates, J. On the Differential Security of the HFEv- Signature Primitive. 7th International Workshop, PQCrypto 2016, Fukuoka, Japan, February 24-26, 2016, Proceedings, Springer (2016)

Presentations

Cartor, R., "Studying C^*: An Introduction to Multivariate Post Quantum Cryptograhpy," Clemson University (April 2019)

Cartor, R., "An Introduction to Multivariate Cryptography," Butler University (November 2018)

Cartor, R., "EFLASH: A New Multivariate Encryption Scheme," Conference on Selected Areas in Cryptography, University of Calgary (August 2018)

Cartor, R., "EFLASH: Introduction and Analysis," Graduate Student Seminar, University of Louisville (November 2017)

Cartor, R., "A Review of Collective Approval," Graduate Student Seminar, University of Louisville (October 2016)

Cartor, R., "An HFE Example," Graduate Student Seminar, University of Louisville (March 2016)

Cartor, R., Gipson, R., Vates, J., "On the Differential Security of the HFEv- Signature Primitive," Algebra-Combinatorics Seminar, University of Louisville (October 2015)

Cartor, R. "Estimating the Size of a Polynomial," Bluegrass Undergraduate Mathematics Symposium, Centre College (September 2013)

Posters

Cartor, R., "Computational Algebraic Geometry in Finite Fields," Association for Women in Mathematics Research Symposium (April 2017)

Cartor, R., "Estimating the Size of a Polynomial," Undergraduate Scholarship Poster Session, Bellarmine University (Winning Poster–April 2014)

Cartor, R., "Minimum Number of Necessary Sudoku Clues," Undergraduate Scholarship Poster Session, Bellarmine University (April 2014)

Cartor, R., "Estimating the Size of a Polynomial," Nebraska Conference of Undergraduate Women in Mathematics (Jan 31-Feb 2 2014)

PROFESSIONAL DEVELOPMENT

Indiana MAA Sectional Meeting *Hanover College*	October 2018
Selected Areas in Cryptography *University of Calgary-Awarded Conference Funding*	August 2018
AWM Research Symposium *University of California, Los Angeles*	April 8-9 2017
PQCrypto *Fukuoka Japan-Awarded Conference Funding*	Feb 2016
AWM Research Symposium *University of Maryland*	April 2015
Ohio River Analysis Meeting *University of Cincinnati*	February 2015
Nebraska Conference for Undergraduate Women in Math *University of Nebraska*	Jan 31-Feb 2 2014
Bluegrass Undergraduate Mathematics Symposium *Centre College*	September 2013
Research Experience for Undergraduates (REU) *Kent State University*	Summer 2013
Bluegrass Undergraduate Mathematics Symposium *Centre College*	September 2012

DEPARTMENTAL SERVICE

Vice President of the University of Louisville's American Mathematical Society Student Chapter	2017-present
Co-Chair of the Graduate Student Seminars	2015-present

AWARDS/HONORS

Selected Areas in Cryptography Stipend	2018
Partial Support from Department of Mathematics Travel Fund	2017 & 2018
Graduate Student Counsel Travel Fund, University of Louisville	2016 & 2018
PQCrypto 2016 Grant	2016
Faculty Merit Award	2014
Faculty Merit Award–Mathematics	2014
Sister Mary Casilda Science Award	2014
In Veritatis Amore Award Nominee	2014
Chi Alpha Sigma	Inducted 2014
Kappa Gamma Pi	Inducted 2014
Omicron Delta Kappa	Inducted 2014
Dean's List	Fall and Spring 2010-2014
Lenihan Memorial Award for Campus and Community Service	2012
Who's Who Recognition	2012
Monsignor Horrigan Scholarship	2010

PROFESSIONAL MEMBERSHIPS

Association for Women in Mathematics
Mathematical Association of America
American Mathematical Society

COMMUNITY ENGAGEMENT

Head Coach of Presentation Academy Dance Team	2015-2019
Assistant Coach of Presentation Academy Dance Team	2014-2015
Head Coach of Atherton High School Dance Team	2013-2014