New Attacks on Lifted Unbalanced Oil Vinegar

Jintai Ding, Zheng Zhang, Joshua Deaton, Kurt Schmidt, FNU Vishakha

Abstract

In 2017, Ward Beullens *et al* submitted Lifted Unbalanced Oil and Vinegar (LUOV)[1], a signature scheme based on the famous multivariate public key cryptosystem (MPKC) called Unbalanced Oil and Vinegar (UOV), to NIST for the competition for post-quantum public key scheme standardization. The defining feature of LUOV is that, though the public key $\mathscr{P}$ works in the extension field of degree r of $\mathbb{F}_2$, the coefficients of $\mathscr{P}$ come from $\mathbb{F}_2$. This is done to significantly reduce the size of $\mathscr{P}$. This is a totally new design which was not therefore under any scrutiny before the submission. The LUOV scheme is now in the second round of the NIST PQC standardization process.

In this paper we introduce a new attack on LUOV. The main idea is to consider some special differentials to develop new approaches to attack the systems.

1 Introduction

1.1 Background and Post-Quantum Cryptography Standardization

A crucial building block for any free, secure, and *digital* society is the ability to authenticate digital messages. In their seminal 1976 paper, Whitfield Diffie and Martin Hellman described the mathematical framework to do such, which is now called a digital signature scheme. They proposed the existence of a function F so that any party can easily check whether $F(X) = D$, *i.e* verify a signature, but only one party, who has a secret key, can find a X so that $F(X) = D$, *i.e* sign a document. Such a function F is called a trapdoor function. Following this idea, Rivest, Shamir, and Adleman proposed the first proof of concept of a signature scheme based off of their now famous RSA public key encryption scheme, which relies on the difficulty of the discrete logarithm problem.

Up to to 2013, the National Institute of Standards and Technology (NIST)'s guidlines allowed for three different types of signature schemes: the Digital Signature Algorithm (DSA), RSA Digital Signature Algorithm, and The Elliptic Curve Digital Signature Algorithm [7]. However, a mayor drawback to these signature schemes is that in 1999 Peter Shor showed that they were weak to a sufficiently powerful quantum computer [14]. As research towards developing a fully fledged quantum computer continues, it has become increasingly clear that there is a significant need to prepare our current communication infrastructure for a post-quantum world. For it is not easy nor quick undergoing to transition our current infrastructure into a post quantum one. Thus a significant effort will be required in order to develop, standardize, and deploy new post-quantum signature schemes.

As such in December 2016 NIST, under the direction of the NSA, put out a call for proposals of new, post-quantum cryptosystems. NIST expects to perform multiple rounds of evaluations over a period of three to five years. The goal of this process is to select a number of acceptable candidate cryptosystems for standardization. These new standards will be used as quantum resistant counterparts to existing standards. The evaluation will be based on the following three criteria: Security, Cost, and Algorithm and Implementation Characteristics. We are currently in the second round of this process, and out of the original twenty-three signature schemes there are only nine left, of which LUOV is one of them.

An additional complication to designing a post-quantum cryptosystem is quantifying securities levels in a post quantum world. For neither the full the capabilities nor limitations of a quantum computer is fully understood. In [11], NIST addresses this issue and quantifies the security strength of a given cryptosystem by comparing it to existing NIST standards in symmetric cryptography, which NIST expects to offer significant resistance to quantum cryptanalysis. Below are the relevant NIST security strength categories which we present logarithm base 2 of the complexity.

NIST Level	Security Description	Complexity
II	At least as hard to break as AES128 (exhaustive key search)	146
IV	At least as hard to break as SHA384 (collision search)	210
V	At least as hard to break as AES265 (exhaustive key search)	272

Table 1: Description of different NIST security strength categories.

1.2 Multivariate Public Key Cryptosystems

Since the work of Diffie and Hellman, mathematicians have found many other groups of cryptosystems that do not rely on Number Theory based problems. Some of these seem to be good candidates for a post-quantum system. One such group is Multivariate Public Key Cryptosystems (MPKC)[3][4]. The security of a MPKC depends on the difficulty of solving a system of multivariate polynomials over a finite field. Usually, these polynomials are of degree two. Solving a set of random multivariate polynomial equations over a finite field, in general, is proven to be an NP-hard problem [8], thus lending a solid foundation for a post-quantum signature scheme. Furthermore, MPKCs in general can be computationally much more efficient than those many other systems.

A breakthrough in MPCK was proposed by Matsumoto and Imai in 1988. Instead of looking for a invertible map between the vector space k^n for a finite field k, they looked at the bigger field K, which is of degree n extension over k, where an inverse map can be constructed [10]. This scheme was broken by Patarin by using the Linearization Equation Attack [12]. However, inspired by the attack, Patarin introduced the Oil and Vinegar scheme [13]. This has been one of the most studied schemes for multivariate cryptography and is the basis for LUOV.

1.3 A Brief Sketch and History of Oil and Vinegar Schemes

One of the most well known multivariate public key signature schemes is the Oil and Vinegar scheme. As the only difference LUOV introduced is what field some coefficients come from, we will only give a brief, verbal description of Oil and Vinegar schemes in general. The key idea of the Oil and Vinegar signature scheme is to reduce signing a document into a linear system. This is done by separating the variables into two collections, the vinegar variables and the oil variables. Then when one constructs the central map, one ensures that oil variables are never multiplied together. That way if one guesses for all the vinegar variables, one is left with a linear system that has high probability of being consistent. To hide the oil and vinegar variables, one simply composes on the right of the central map an affine transformation to change the basis.

Patarin originally proposed that the number of oil variables would equal the number of vinegar variables. However, the Balanced Oil Vinegar scheme was broken by Kipnis and Shamir using the method of invariant subspaces [9]. This attack, however, is thwarted by making the number of vinegar variables greater then the number of oil variables. Though,

proposed nearly twenty years ago, the Unbalanced Oil and Vinegar (UOV) scheme still remains unbroken. Further, this simple and elegant signature scheme boasts small signatures and fast signing times. Arguably the only drawback to UOV is its rather large public key size. There have been several attempt to reduce the size of the public key but keeping the strength of UOV like HIMQ-3, a round one NIST submission. However, this is in general hard to due as can be seen from the soon to be published singularity attack on HIMQ-3 by Ding *et al.*

Out of the nine signature schemes that were accepted to round two of the NIST standardization program, two (LUOV and Rainbow) are based off of UOV with modifications to reduce key size. Rainbow, originally proposed in 2005, reduces its keysize by forming multiple layers of UOV schemes, where oil variables in a higher layer become vinegar variables in the lower layers [5, 6]. LUOV achieved a reduction in key size by forcing all the coefficients of the public key to either be 0 or 1. This is a totally new design which was not therefore under any scrutiny before the submission. In this paper we will show that such modifications used by LUOV allows for algebraic manipulations that result in a underdetermined quadratic system over a much smaller finite field. We will further show that Rainbow and other UOV schemes is immune to such attacks.

1.4 Our Contributions

We will present a new attack method called the Subfield Differential Attack (SDA). This attack does not rely on the Oil and Vinegar structure of LUOV but merely that the coefficients of the quadratic terms are contained in a small subfield. We will show that the attack will make it impossible for LUOV to fulfill the NIST security levels requirements.

First, we will recall the design of LUOV. Afterwards, we will argue heuristically that there is a high probability that there exists a solution in a smaller field. More precisely, for public key $\mathscr{P} : \mathbb{F}_{2^r}^n \to \mathbb{F}_{2^r}^o$, we assert the for any $\mathbf{x}' \in \mathbb{F}_{2^r}^n$ and $\mathbf{y} \in \mathbb{F}_{2^r}^o$ there exists $\bar{\mathbf{x}} \in \mathbb{F}_{2^d}^n$ such that $\mathscr{P}(\mathbf{x}' + \bar{\mathbf{x}}) = \mathbf{y}$, where $\mathbb{F}_{2^d}$ is a small subfield of $\mathbb{F}_{2^r}$. We call $\mathbf{x}' + \bar{\mathbf{x}}$ the differential of $\mathscr{P}$. We will also provide experimental evidence that overwhelming supports our assertion. With the differential and the fact that the coefficients of $\mathscr{P}$ are either 0 or 1, we will show that by viewing $\mathscr{P}$ in the polynomial ring over an small subfield modded by an irreducible polynomial and by comparing coefficients, we have reduced the problem of solving an underdetermined quadratic over $\mathbb{F}_{2^d}$. The complexity required for such is well under our target. For each proposed set of parameters, we will explicitly apply our attack. We will provided a small toy example. Finally, we will explain how UOV and Rainbow are unaffected by our attack.

2 Lifted Unbalanced Oil Vinegar Scheme

One can tell from the name that LUOV is a modification of the original unbalanced oil vinegar scheme. The main difference between these two schemes is that in the original unbalanced oil vinegar scheme there is no restriction on the coefficients of the public key. However in the LUOV, the scheme uses two finite fields, one is the binary field of two elements, the other is its extension of degree r. In order to shorten the size of the public, the coefficients of the public key are all from the base field, but the document and the signatures may contain elements from the extension field.

Let $\mathbb{F}_2$ be the binary field and $\mathbb{F}_{2^r}$ be its extension of degree r. Let o and v be two positive integers such that $o < v$ and $n = o + v$.

The central map $\mathscr{F} : \mathbb{F}_{2^r}^n \to \mathbb{F}_{2^r}^o$ is a quadratic map whose components $f_1, \cdots, f_o$ are in the form:

$$f_k(\mathbf{x}) = \sum_{i=1}^{v} \sum_{j=i}^{n} \alpha_{i,j,k} x_i x_j + \sum_{i=1}^{n} \beta_{i,k} x_i + \gamma_k.$$

where the coefficients $\alpha_{i,j,k}, \beta_{i,k}$ and γ_k are from the base field $\mathbb{F}_2$. One can easily see that these polynomials are in the oil vinegar form, meaning that oil is never multiplied with oil. To hide the oil vinegar structure of these polynomials, an invertible linear map $\mathscr{T} : \mathbb{F}_{2^r}^n \to \mathbb{F}_{2^r}^n$ is used to mix the variables. In particular, the authors choose $\mathscr{T}$ in the form:

$$\begin{bmatrix} \mathbf{1}_v & \mathbf{T} \\ \mathbf{0} & \mathbf{1}_o \end{bmatrix}$$

where $\mathbf{T}$ is a $v \times o$ matrix whose entries are also from the small field $\mathbb{F}_2$, to speed up the key generation and signing process as well as keep the size of the private key small.
The public and private keys of LUOV are given by:

$$\textbf{Public key}: \mathscr{P} = \mathscr{F} \circ \mathscr{T}. \qquad \textbf{Private key}: \mathscr{T}.$$

The way to invert the central map $\mathscr{F}$ is the same as UOV. Given an element $\mathbf{y} \in \mathbb{F}_{2^r}^o$, one wants to find an $\mathbf{x} \in \mathbb{F}_{2^r}^n$ such that $\mathscr{F}(\mathbf{x}) = \mathbf{y}$. First, one assigns random values from $\mathbb{F}_{2^r}$ to the vinegar variables $x_1, \cdots, x_v$. By substituting them in the equations, the remaining system becomes a linear system, which can be solved easily by Gaussian elimination. If no solution exists, repeat the process by choosing different values for the vinegar variables.

This is a totally new design which was not therefore under any scrutiny before the submission.

The LUOV scheme actually made into the second round of the NIST PQC standardization process.

3 The Subfield Differential Attack on LUOV

3.1 The general idea of the attack

The key to our attack is the structure of finite field extension. The authors ignore the use of the small subfields. The small $\mathbb{F}_{2^d}$ connects the base field $\mathbb{F}_2$ and the extension field $\mathbb{F}_{2^r}$ and what is more important is that we can construct isomorphisms between the extension field $\mathbb{F}_{2^r}$ and the quotient of polynomial ring over $\mathbb{F}_{2^d}$ modded by an irreducible polynomial $f(t)$. So every element in the extension field can be represented by a polynomial in $\mathbb{F}_{2^d}[t]/f(t)$.

The next question is how we use the field $\mathbb{F}_{2^d}[t]/f(t)$. This is where the differential comes in. Suppose $P : \mathbb{F}_{2^r}^n \to \mathbb{F}_{2^r}^o$ is a quadratic multivariate polynomial having coefficients from $\mathbb{F}_2$. Let $\mathbf{y}$ be an arbitrary element in $\mathbb{F}_{2^r}^o$. The differential is defined to be $\mathbf{x}' + \bar{\mathbf{x}}$ where $\mathbf{x}'$ is an arbitrary point we choose from $\mathbb{F}_{2^r}^n$ and $\bar{\mathbf{x}}$ is an indeterminate point in $\mathbb{F}_{2^d}^n$. If we evaluate the quadratic multivariate polynomial $\mathscr{P}$ at the differential and set it equal to $\mathbf{y}$, the quadratic part of the polynomial P will produce terms ${x_i}'{x_j}'$, ${x_i}'\bar{x_j}$ and $\bar{x_i}\bar{x_j}$, where ${x_i}'$ are known and $\bar{x_j}$ are unknown. We are expressing $\mathbb{F}_{2^r}$ as $\mathbb{F}_{2^d}[t]/f(t)$, so every element in the field can be expressed as a polynomial in t of degree at most $r/d - 1$. Applying such isomorphism on both sides and comparing the coefficients of t, $t^2, \cdots, t^{r/d-1}$, we almost obtain a linear equation. This is because in this equation, the $\bar{x_i}\bar{x_j}$ is a product of two elements from the small subfield. Thus if we apply this method on every component of the public key we almost have a system of linear equations. So the idea of solving this equation by differential is that we associate the elements from extension field to the linear part where it is easy to solve, and the elements from the small subfield to the quadratic part where we can search for them.

3.2 The Goal of the Attack

Let $\mathscr{P} = \mathscr{F} \circ \mathscr{T} : \mathbb{F}_{2^r}^n \to \mathbb{F}_{2^r}^o$ be a given LUOV public key as described in section 2.1. Then following the construction of all Oil Vinegar Schemes, $\mathscr{P}$ appears to be a random quadratic system, except in this case the coefficients will be from the field $\mathbb{F}_2$ embedded in $\mathbb{F}_{2^r}$.

$$\mathscr{P}(\mathbf{x}) = \begin{cases} \tilde{f}_1(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=i}^{n} \alpha_{i,j,1} x_i x_j + \sum_{i=1}^{n} \beta_{i,1} x_i + \gamma_1 \\ \tilde{f}_2(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=i}^{n} \alpha_{i,j,2} x_i x_j + \sum_{i=1}^{n} \beta_{i,2} x_i + \gamma_2 \\ \vdots \\ \tilde{f}_o(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=i}^{n} \alpha_{i,j,o} x_i x_j + \sum_{i=1}^{n} \beta_{i,o} x_i + \gamma_o \end{cases}$$

As is found in basic field theory $\mathbb{F}_2$ is not the only subfield embedded in $\mathbb{F}_{2^r}$. For if d is any positive divisor of r where $ds = r$, then for any irreducible degree s polynomial $f(t) \in \mathbb{F}_{r^d}[t]$ we have that $\mathbb{F}_{2^d}[t]/f(t) \cong \mathbb{F}_{2^r}$. Here $\mathbb{F}_{2^d}$ is embedded as the set of constant polynomials in $\mathbb{F}_{2^d}[t]/f(t)$. This allows us to change question like finding preimages (and thus signatures) from being about elements from the larger field $\mathbb{F}_{2^r}$ to being about elements from the smaller field $\mathbb{F}_{2^d}$.

The first such question is that given a document $\mathbf{y} = (y_1, \cdots, y_o) \in \mathbb{F}_{2^r}^o$ and an arbitrary $\mathbf{x}' \in \mathbb{F}_{2^d}^n$ does there exist a reasonable small integer d such that there will also exist a $\bar{\mathbf{x}} \in \mathbb{F}_{2^d}^n \subset \mathbb{F}_{2^r}^n$ where $P(\mathbf{x}' + \bar{\mathbf{x}}) = \mathbf{y}$? It turns out that for all given parameters of LUOV that the answer is yes.

3.3 The Probability of $\mathbf{x}' + \bar{\mathbf{x}}$ Existing

Our first step is to calculate the probability of it existing. Fix $\mathbf{x}' \in \mathbb{F}_{2^r}^n$ and to consider the function $\mathscr{P}' : \mathbb{F}_{2^d}^n \to \mathbb{F}_{2^r}^o$ given by $\mathscr{P}'(\bar{\mathbf{x}}) = \mathscr{P}(\mathbf{x}' + \bar{\mathbf{x}})$ where $d|r$. Notice that $\mathscr{P}'$ is a quadratic system of n variable with o equations over $\mathbb{F}_{2^d}$. For the sake of argument, we will assume that $\mathscr{P}'$ acts as a random map from $\mathbb{F}_{2^d}^n \to \mathbb{F}_{2^r}^o$. This mean that the outputs of $\mathscr{P}'$ are uniform in $\mathbb{F}_{2^r}^o$. Now we arbitrarily choose $\mathbf{y} \in \mathbb{F}_{2^r}^o$. We are interested in the probability that there does not exist $\bar{\mathbf{x}} \in \mathbb{F}_{2^d}^n$ such that $\mathscr{P}'(\bar{\mathbf{x}}) = \mathbf{y}$. Since $|\mathbb{F}_{2^d}^n| = 2^{d \cdot n}$ and $|\mathbb{F}_{2^r}^o| = 2^{r \cdot o}$, we find that for any $\bar{\mathbf{x}} \in \mathbb{F}_{2^d}^n$ that the probability that $\mathscr{P}'(\bar{\mathbf{x}}) \neq \mathbf{y}$ is $1 - \frac{1}{2^{r \cdot o}}$. Since the outputs of $\mathscr{P}'$ are independent, exhausting every element of $\mathbb{F}_{2^d}^n$ we find we can estimate our desired probability as

$$\left(1 - \frac{1}{2^{r \cdot o}}\right)^{2^{d \cdot n}} = \left(\left(1 - \frac{1}{2^{r \cdot o}}\right)^{2^{r \cdot o}}\right)^{2^{(d \cdot n) - (r \cdot o)}} \approx e^{-2^{(d \cdot n) - (r \cdot o)}},$$

because $\lim_{n \to \infty}(1 - \frac{1}{n})^n = e^{-1}$.

In tables 2 and 3 we calculate the probability of failure for the first set parameters as given in the LUOV description. The table 2 is given on parameters designed to reduce the size of signatures. These parameters are used in situations where many signatures are needed. For these we always chose $d = 2$ based on the relatively small size of the extension field $\mathbb{F}_{2^r}$. The table 3 is given on parameters designed to the cost of both signatures and public keys. These parameters are used when communicating both signatures and public keys is needed. Due to the larger size of $\mathbb{F}_{2^r}$, our choice of d is the smallest such that the probability of failure is small. The smallness of d will mean finding a signature will be easier later.

3.4 On how to find $\bar{\mathbf{x}}$

So we can say with confidence that such a $\bar{\mathbf{x}}$ will exists for the given parameters of LUOV.

NIST Security Level	r	o	v	n	d	Probability of Failure
II	8	58	237	295	2	$\exp(-2^{126})$
IV	8	82	323	405	2	$\exp(-2^{154})$
V	8	107	371	478	2	$\exp(-2^{100})$

Table 2: Estimated Probabilities of Failure for Parameters Designed to Minimize the Size of the Signature

NIST Security Level	r	o	v	n	d	Probability of Failure
II	48	43	222	265	8	$\exp(-2^{56})$
IV	64	61	302	363	16	$\exp(-2^{1904})$
V	80	76	363	439	16	$\exp(-2^{944})$

Table 3: Estimated Probabilities of Failure for Parameters Designed to Minimize the Size of the Signature and Public Key

As the size of $\mathbb{F}_{2^d}^n$ is still far to large for a brute force search, we will narrow down the possibilities by examining the components of $P(\mathbf{x}' + \bar{\mathbf{x}})$'s coefficients and $\mathbf{y}$'s components when viewed as elements of $\mathbb{F}_{2^d}[t]/f(t)$. Let our arbitrary $\mathbf{x}' \in \mathbb{F}_{2^r}^n$ be equal to $(x_1', \cdots, x_n')$. We see that the kth component of $\mathscr{P}(\mathbf{x}' + \bar{\mathbf{x}})$ is

$$\tilde{f}_k(\mathbf{x}' + \bar{\mathbf{x}}) = \sum_{i=1}^{n} \sum_{j=i}^{n} \alpha_{i,j,k}(x_i' + \bar{x}_i)(x_j' + \bar{x}_j) + \sum_{i=1}^{n} \beta_{i,k}(x_i' + \bar{x}_i) + \gamma_k = y_k$$

Expanding the above and separating the quadratic terms leads to

$$\begin{aligned} \tilde{f}_k(\mathbf{x}' + \bar{\mathbf{x}}) = & \sum_{i=1}^{n} \sum_{j=i}^{n} \alpha_{i,j,k}(x_i' x_j' + x_i' \bar{x}_i + x_j' \bar{x}_j) + \sum_{i=1}^{n} \beta_{i,k}(x_i' + \bar{x}_i) + \gamma_k \\ & + \sum_{i=1}^{n} \sum_{j=i}^{n} \alpha_{i,j,k} \bar{x}_i \bar{x}_j \\ = & \ y_k \end{aligned}$$

We notice that the coefficients of the quadratic terms are still from the field $\mathbb{F}_2$, meaning they will be represented by the constant terms 0 or 1 as we view these over $\mathbb{F}_{2^d}[t]/f(t)$. On the other hand the x_i' and the y_k are arbitrary elements of $\mathbb{F}_{2^r}$ are thus represented by degree at most $s-1$ polynomials in $\mathbb{F}_{2^d}[t]/f(t)$. We can thus regroup the above equation in terms of the powers of t, where the quadratic part is confined in the constant term. Meaning for some $w_{i,k} \in \mathbb{F}_{2^d}$, some linear polynomials $g_{i,k}(\bar{x}_1, \cdots, \bar{x}_n) \in \mathbb{F}_{2^d}[\bar{x}_1, \cdots, \bar{x}_n]$, and some quadratic polynomial $Q_k(\bar{x}_1, \cdots, \bar{x}_n) \in \mathbb{F}_{2^d}[\bar{x}_1, \cdots, \bar{x}_n]$ we have that

$$\tilde{f}_k(\mathbf{x}' + \bar{\mathbf{x}}) = \sum_{i=1}^{s-1} g_{i,k}(\bar{x}_1, \cdots, \bar{x}_n) t^i + Q_k(\bar{x}_1, \cdots, \bar{x}_n) = y_k = \sum_{i=0}^{s-1} w_{i,k} t^i.$$

We thus find $s-1$ linear equations $g_{i,k}(\bar{x}_1, \cdots, \bar{x}_n) = w_{i,k}$ for each $\tilde{f}_k$ in the public key $\mathscr{P}$ leading to a system of $(s-1)o$ linear equations with n variables. Let us denote this by system by

$$A\mathbf{x} = \mathbf{y}$$

where $A\mathbf{x}$ is a matrix times a variable vector representing the linear equations $g_{i,k}(\bar{x}_1, \cdots, \bar{x}_n)$. Our desired $\bar{\mathbf{x}}$ will be in the set S of solutions of this system, though certainly most of the solution space will not be equivalent to $\bar{\mathbf{x}}$ as they also must satisfy each quadratic equation $Q_k(\bar{x}_1, \cdots, \bar{x}_n) = w_{0,k}$ to be a valid signature for $\mathbf{y}$. As each $g_{i,k}(\bar{x}_1, \cdots, \bar{x}_n)$ is essentially a

random linear polynomial, there is a good probability for the rank of A to be $(s-1)o$ (or full rank if $(s-1)o \geq n$, but this makes the existence of $\bar{\mathbf{x}}$ unlikely and need not be considered).

In general we see that

$$|S| = n - \text{rank}(A) = n - (s-1)o$$

by the Rank Nullity Theorem and the fact that there are n variables. If the size of S is sufficiently small (which will depend on the parameters of LUOV and the choice of d), a brute force search now is applicable. If the size of S is too large, then more sophisticated searches are needed. We see that our problem thus reduces to solving o equations with $n-(s-1)o$ variables in a search for a solution from S.

Below we will estimate the complexity for finding a signature for the various parameters given in tables 1 and 2. We will use the method of Thomae and Wolf [15]. This reduces our system of o equations and $n-(s-1)o$ variables to one of m equations *and* variables where

$$m = o - \left\lfloor \frac{n-(s-1)o}{o} \right\rfloor.$$

We will then guess for a certain number of the variables forming a overdetermined system. This system will have a certain degree of regularity. We can then use an algorithm such as XL to solve this system. The most complex part of this process is solving a sparse linear equation over a finite field [2]. We can use the block Wiedermann algorithm to solve this, thus finding our signature. Below we give a table describing this process for the various parameters describing each system as (number of equations) × (number of variables).

Table and Security	**Finite Field**	**Original System**	**New System**	**# of Guesses**	**Degree of Regularity**	**Log_2 Complexity**
(2, II)	$\mathbb{F}_{2^2}$	58×121	56×56	24	7	107
(2, IV)	$\mathbb{F}_{2^2}$	82×259	79×79	33	9	143
(2, V)	$\mathbb{F}_{2^2}$	107×157	106×106	51	9	184
(3, II)	$\mathbb{F}_{2^8}$	43×50	42×42	3	19	135
(3, IV)	$\mathbb{F}_{2^{16}}$	61×180	59×59	2	31	202
(3, V)	$\mathbb{F}_{2^{16}}$	76×131	75×75	2	38	244

Table 4: Results

The degree of regularity estimates and the complexity estiamtes are based on the works [16, 17] and the complexity estimates are based on Proposition 4 in [16].
Recalling that NIST requires complexity $(2^{146}, 2^{210}, 2^{272})$ for security levels (II, IV, V) respectively, we see that LUOV fails to meet the security level requirements in all parameter sets given for their targeted security.
The two schemes which claim to be of Level II security do not even satisfies the Level I security, which is supposed to be 2^{143}.

3.5 Toy Example

Let $o = 2$, $v = 8$, and $n = 10$. The size the large extension field chosen by the public key generator will be $2^8 = 256$. In the attack we will use as our small field $\mathbb{F}_{2^2}$ denoting its elements by $\{0, 1, w_1, w_2\}$. We then will represent the field $\mathbb{F}_{2^8}$ by $\mathbb{F}_{2^4}[t]/f(t)$ where $f(t) = t^4 + t^2 + w_1 t + 1$.

Consider the LUOV public key $\mathscr{P} : \mathbb{F}_{2^8}^n \to \mathbb{F}_{2^8}^o$ where for simplicity sake will be homogeneous degree two:

$$
\begin{aligned}
\tilde{f}_1(\mathbf{x}) =& x_1x_4 + x_1x_5 + x_1x_6 + x_1x_7 + x_1x_8 + x_1x_9 + x_2x_4 + x_2x_6 + x_2x_9 + x_3^2 \\
&+ x_3x_6 + x_3x_7 + x_3x_{10} + x_4^2 + x_4x_7 + x_4x_8 + x_4x_9 + x_4x_{10} + x_5x_6 + x_6x_{10} \\
&+ x_7^2 + x_7x_8 + x_7x_9 + x_8x_9 + x_8x_{10} + x_9^2 + x_9x_{10} \\
\tilde{f}_2(\mathbf{x}) =& x_1x_3 + x_1x_4 + x_1x_5 + x_1x_9 + x_2x_3 + x_2x_6 + x_2x_7 + x_2x_9 + x_3^2 + x_3x_4 \\
&+ x_3x_5 + x_3x_6 + x_3x_7 + x_3x_9 + x_4^2 + x_4x_5 + x_4x_6 + x_4x_7 + x_4x_{10} + x_5^2 \\
&+ x_5x_6 + x_5x_7 + x_5x_8 + x_5x_{10} + x_6x_7 + x_7x_9 + x_9x_{10} + x_{10}^2
\end{aligned}
$$

We will attempt to find a signature for the message:

$$
\mathbf{y} = \begin{bmatrix} w_1t^3 + w_2t^2 + w_2t \\ w_2t^3 + w_2t^2 + t \end{bmatrix}
$$

First we randomly select our $\mathbf{x}'$ as

$$
\mathbf{x}' = \begin{bmatrix}
t^3 + w_2t \\
w_1t^3 + w_2t^2 + w_2t \\
t^3 + t + 1 \\
w_2t^2 + w_1 \\
t^3 + t^2 + 1 \\
w_2t^3 + t^2 + w_2t + w_2 \\
w_1t^3 + w_2t + w \\
w_1t^2 + w_2t + 1 \\
t^3 + w_2t + w_1 \\
w_2t + w_2
\end{bmatrix}
$$

We then calculate $\mathscr{P}(\mathbf{x} + \bar{\mathbf{x}})$ and represent it as a polynomial of t:

$$
\begin{aligned}
\tilde{f}_1(\mathbf{x}' + \bar{\mathbf{x}}) =& (\bar{x}_1 + w_1\bar{x}_2 + \bar{x}_3 + w_1\bar{x}_5 + w_2\bar{x}_6 + \bar{x}_7 + w_1\bar{x}_8 + \bar{x}_9 + w_2\bar{x}_{10})\,t^3 \\
&+ (\bar{x}_1 + w_1\bar{x}_2 + \bar{x}_3 + \bar{x}_4 + \bar{x}_5 + w_1\bar{x}_6 + \bar{x}_7 + w_2\bar{x}_8 + w_1\bar{x}_9)\,t^2 \\
&+ (w_2\bar{x}_3 + w_1\bar{x}_6 + w_1\bar{x}_7 + w_2\bar{x}_9 + w_1\bar{x}_{10})\,t \\
&+ Q_1(\bar{x}_1, \cdots, \bar{x}_n) \\
\tilde{f}_2(\mathbf{x}' + \bar{\mathbf{x}}) =& (\bar{x}_1 + \bar{x}_2 + w_1\bar{x}_3 + \bar{x}_5 + \bar{x}_8)\,t^3 \\
&+ (w_1\bar{x}_1 + \bar{x}_2 + \bar{x}_6 + \bar{x}_8 + w_2\bar{x}_9 + w_1\bar{x}_{10})\,t^2 \\
&+ (w_1\bar{x}_1 + w_1\bar{x}_2 + w_2\bar{x}_3 + \bar{x}_4 + w_1\bar{x}_5 + \bar{x}_6 + w_1\bar{x}_7 + \bar{x}_9 + w_2\bar{x}_{10})\,t \\
&+ Q_2(\bar{x}_1, \cdots, \bar{x}_n)
\end{aligned}
$$

Where $Q_1(\bar{x}_1, \cdots, \bar{x}_n)$ and $Q_2(\bar{x}_1, \cdots, \bar{x}_n)$ are quadratic polynomials from $\mathbb{F}_{2^2}[\bar{x}_1, \cdots, \bar{x}_n]$. By comparing the coefficients of t^3, t^2, t^1 assuming $\mathscr{P}(\mathbf{x}' + \bar{\mathbf{x}}) = \mathbf{y}$ we arrive at a system of linear equations over $\mathbb{F}_{2^2}$. This can be represented by a matrix equation $\mathbf{Ax} = \mathbf{y}$. In our case this is the following:

$$
\begin{bmatrix}
1 & w_1 & 1 & 0 & w_1 & w_2 & 1 & w_1 & 1 & w_2 \\
1 & w_1 & 1 & 1 & 1 & w_1 & 1 & w_2 & w_1 & 0 \\
0 & 0 & w_2 & 0 & 0 & w_1 & w_1 & 0 & w_2 & w_1 \\
1 & 1 & w_1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
w_1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & w_2 & w_1 \\
w_1 & w_1 & w_2 & 1 & w_1 & 1 & w_1 & 0 & 1 & w_2
\end{bmatrix}
\begin{bmatrix}
\bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_4 \\ \bar{x}_5 \\ \bar{x}_6 \\ \bar{x}_7 \\ \bar{x}_8 \\ \bar{x}_9 \\ \bar{x}_{10}
\end{bmatrix}
=
\begin{bmatrix}
w_1 \\ w_2 \\ w_2 \\ w_2 \\ w_2 \\ 1
\end{bmatrix}
$$

The solution space for the above equation has dimension 4 over $\mathbb{F}_{2^2}$, as we would expect as $n-(s-1)o=4$. Thus there are only $(2^2)^4=2^8$ possible choices for $\bar{\mathbf{x}}$. A quick search through these, plugging them + $\mathbf{x}'$ into the public key and seeing if the result is $\mathbf{y}$, finds the signature

$$\sigma = \begin{bmatrix} t^3+w_2t+1 \\ w_1t^3+w_2t^2+w_2t+w_1 \\ t^3+t+w_2 \\ w_2t^2 \\ t^3+t^2+1 \\ w_2t^3+t^2+w_2t+1 \\ w_1t^3+w_2t+w_1 \\ w_1t^2+w_2t+1 \\ t^3+w_2t+1 \\ w_2t \end{bmatrix}$$

In order to show that this was not a fluke and that our above heuristic argument on $\mathscr{P}'$ (namely that it acts as a random map) reflects reality, we ran an experiment on a fixed public key with parameters $r=8, o=5, v=20, n=25, d=2$. Generating 10,000 random documents, we were able to find using the method from the toy example a signature for every document.

4 The Inapplicability of the Subfield Differential Attack on Unbalanced Oil Vinegar

Now let us discuss why subfield differential attack does not work on unbalanced oil vinegar. Let us assume that $\mathbb{F}_{q^r}$ contains nontrivial subfield $\mathbb{F}_{q^d}$. If we construct our differential $\mathbf{x}'+\bar{\mathbf{x}}$ with $\mathbf{x}'\in\mathbb{F}_{q^r}$ and $\bar{\mathbf{x}}\in\mathbb{F}_{q^d}$, evaluating the public key at the the differential. In its kth component, we have that

$$\tilde{f}_k(\mathbf{x}'+\bar{\mathbf{x}}) = \sum_{i=1}^{n}\sum_{j=i}^{n}\alpha_{i,j,k}(x_i'+\bar{x}_i)(x_j'+\bar{x}_j)+\sum_{i=1}^{n}\beta_{i,k}(x_i'++\bar{x}_i)+\gamma_k = y_k$$

Note that there is no restrictions on coefficients, $\alpha_{i,j,k},\beta_{i,k}$ and γ_k are randomly chosen from $\mathbb{F}_{q^r}$. If we multiply the polynomial out

$$\begin{aligned}\tilde{f}_k(\mathbf{x}'+\mathbf{x}) = &\sum_{i=1}^{n}\sum_{j=i}^{n}\alpha_{i,j,k}(x_i'x_j'+x_i'\bar{x}_i+x_j'\bar{x}_j)+\sum_{i=1}^{n}\beta_{i,k}(x_i'+\bar{x}_i)+\gamma_k \\ &+\sum_{i=1}^{n}\sum_{j=i}^{n}\alpha_{i,j,k}\bar{x}_i\bar{x}_j = y_k\end{aligned} \tag{1}$$

and apply the isomorphism to express every element in a polynomial in $\mathbb{F}_{q^d}[t]/f(t)$ except the quadratic terms $\bar{x}_i\bar{x}_j$, we immediately lose track of the polynomial expression for $\bar{x}_i$ and $\bar{x}_j$. Because the coefficients $\alpha_{i,j,k},\beta_{i,k}$ and γ_k can also be represented by a polynomial in $\mathbb{F}_{q^d}[t]/f(t)$, multiplication from these coefficients $\alpha_{i,j,k},\beta_{i,k}$ and γ_k in $\tilde{f}_k$ will mix the degrees of the polynomial expression of $\bar{x}_i$ and $\bar{x}_j$ in $\mathbb{F}_{q^d}[t]/f(t)$. Thus comparing the coefficients of all degrees of t is useless. Therefore this attack is not at all applicable to UOV or Rainbow.

5 Prime UOV and subprime UOV

One may immediately realize that if the scheme is constructed with a prime extension, then there will be no intermediate field.This will prevent the subfield differential attack as it is

now.
Using this idea, We will propose a new type UOV, which we call prime UOV (PUOV). It is based on the UOV and LUOV, but with additional conditional requirements.

- Let the base field be $\mathbb{F}_p$, where p is the prime number. Then we require that the extension field which is $\mathbb{F}_{p^r}$ to be a primes extension field of $\mathbb{F}_p$, namely r is a prime number. Therefore here the only subfield inside the extension field is $\mathbb{F}_p$.
- It is a normal UOV, but the coefficients of central OV map, the linear transformation to hide the OV polynomials and the pubic key should be over $\mathbb{F}(p)$ not $\mathbb{F}_{p^r}$, but we will use it to sign a document over $\mathbb{F}_{p^r}$.
- We require that $o + v < ro$.
- We require that $p^{ro-(o+v)} > 55$.
- For higher security, we should increase $p^{ro-(o+v)}$ to the level required.
- The new scheme should have security level required by the usual UOV under the normal UOV attacks.

Another variant of this construction, which we call subprime UOV.
It is based on the UOV and LUOV idea, but with additional conditional requirements.

- Let the base field be $\mathbb{F}_p$, where p is the prime number. Then we require that the extension field which is $\mathbb{F}_{p^r}$ to be an extension of $\mathbb{F}_p$, namely r is not necessarily a prime number.
- It is a normal UOV, but the coefficients of central OV map, the linear transformation to hide the OV polynomials and the pubic key should be over $\mathbb{F}_p$ not $\mathbb{F}_{p^r}$, but we will use it to sign a document over $\mathbb{F}_{p^r}$.
- Let d be any factor of r including 1, we require that $d(o + v) < ro$.
- We also require that $p^{ro-d(o+v)} > 55$.
- For higher security, we should increase $p^{ro-(o+v)}$ to the level required.
- The new scheme should have security level required by the usual UOV under the normal UOV attacks.

However our further work indicates that we can do new attacks on such designs without using any subfield but some special **subset** in the large field, which we call subset differential attack. Therefore we believe much more work is needed to understand the security of these new schemes. from our experiences by now, we believe there are possible even more lethal attacks and we hope the research community can use held to extend our work.

6 Conclusion

We proposed a new attack to a NIST round 2 candidate LUOV. This attack only uses basic structure of field extension and a differential $\mathbf{x} + \bar{\mathbf{x}}$ to solve system of equations. The idea of our attack is simple, however it has great potential. First, one can see that the attack does not depend on the design of central map, it can be applied to other scheme with a lifted structure. Furthermore, one may ask if it is possible to choose a better $\mathbf{x}$ rather than randomly selecting one to reduce the complexity of solving the equations. Further work indicates that much more work needs to be done on this type pf new differential attacks.

7 Acknowledgment

We would like to thank Bo-yin Yang for useful discussions, in particular, on various complexity analysis. We would like to thank partial support of NSF and NIST.

References

[1] Ward Beullens and Bart Preneel. Field lifting for smaller uov public keys. In *Progress in Cryptology – INDOCRYPT 2017*, pages 227–246. Springer, 2017.

[2] Chen-Mou Cheng, Tung Chou, Ruben Niederhagen, and Bo-Yin Yang. Solving quadratic equations with xl on parallel architectures - extended version. Cryptology ePrint Archive, Report 2016/412, 2016. https://eprint.iacr.org/2016/412.

[3] Jintai Ding, Jason E. Gower, and Dieter Schmidt. *Multivariate Public Key Cryptosystems*, volume 25 of *Advances in Information Security*. Springer, 2006.

[4] Jintai Ding and Albrecht Petzoldt. Current state of multivariate cryptography. *IEEE Security & Privacy*, 15(4):28–36, 2017.

[5] Jintai Ding and Dieter Schmidt. Rainbow, a new multivariable polynomial signature scheme. In *International Conference on Applied Cryptography and Network Security*, pages 164–175. Springer, 2005.

[6] Jintai Ding, Bo-Yin Yang, Chia-Hsin Owen Chen, Ming-Shing Chen, and Chen-Mou Cheng. New differential-algebraic attacks and reparametrization of rainbow. In *Applied Cryptography and Network Security, 6th International Conference, ACNS 2008, New York, NY, USA, June 3-6, 2008. Proceedings*, pages 242–257, 2008.

[7] Patrick Gallagher. Digital signature standard (dss). *Federal Information Processing Standards Publications, volume FIPS*, pages 186–3, 2013.

[8] David S Johnson and Michael R Garey. *Computers and intractability: A guide to the theory of NP-completeness*. WH Freeman, 1979.

[9] Aviad Kipnis and Adi Shamir. Cryptanalysis of the oil and vinegar signature scheme. In *Annual International Cryptology Conference*, pages 257–266. Springer, 1998.

[10] Tsutomu Matsumoto and Hideki Imai. Public quadratic polynomial-tuples for efficient signature-verification and message-encryption. In *Workshop on the Theory and Application of of Cryptographic Techniques*, pages 419–453. Springer, 1988.

[11] National Institute of Standards and Technology. Submission requirements and evaluation criteria for the post-quantum cryptography standardization process. Technical report, National Institute of Standards and Technology, 2017.

[12] Jacques Patarin. Cryptanalysis of the matsumoto and imai public key scheme of eurocrypt'88. In *Annual International Cryptology Conference*, pages 248–261. Springer, 1995.

[13] Jacques Patarin. The oil and vinegar algorithm for signatures. In *Dagstuhl Workshop on Cryptography, 1997*, 1997.

[14] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41(2):303–332, 1999.

[15] Enrico Thomae and Christopher Wolf. Solving underdetermined systems of multivariate quadratic equations revisited. In *Public Key Cryptography - PKC 2012 - 15th International Conference on Practice and Theory in Public Key Cryptography, Darmstadt, Germany, May 21-23, 2012. Proceedings*, pages 156–171. Springer, 2012.

[16] Bo-Yin Yang, Chia-Hsin Owen Chen, Daniel J. Bernstein, and Jiun-Ming Chen. Analysis of QUAD. In *Fast Software Encryption, 14th International Workshop, FSE 2007, Luxembourg, Luxembourg, March 26-28, 2007, Revised Selected Papers*, pages 290–308. Sringer, 2007.

[17] Bo-Yin Yang and Jiun-Ming Chen. Theoretical analysis of XL over small fields. In *Information Security and Privacy: 9th Australasian Conference, ACISP 2004, Sydney, Australia, July 13-15, 2004. Proceedings*, pages 277–288. Sringer, 2004.