Cryptanalysis of the SNOVA Signature Scheme

Peigen Li[1,2(✉)] and Jintai Ding[1,2]

[1] Beijing Institute of Mathematical Sciences and Applications, Beijing, China
lpg22@bimsa.cn

[2] Yau Mathematical Sciences Center, Tsinghua University, Beijing, China

Abstract. SNOVA is a variant of a UOV-type signature scheme over a noncommutative ring. In this article, we demonstrate that certain parameters provided by authors in SNOVA fail to meet the NIST security level, and the complexities are lower than those claimed by SNOVA.

Keywords: multivariate public key cryptography · UOV · SNOVA

1 Introduction

Public key cryptosystems currently used such as RSA and ECC can be broken by a quantum computer executing Shor's algorithm [24] in polynomial time. Therefore, cryptosystems resistant to quantum computers are gaining increasing importance. There are many post-quantum cryptosystems based on different theory such as lattice theory, algebraic geometry, coding theory, and the isogeny theory of elliptic curves.

In 2022, the U.S. National Institute for Standards and Technology (NIST) on post-quantum cryptography (PQC) posted a call for additional digital signature proposals to be considered in the PQC standardization process. In 2023, 50 different signature schemes were submitted, including code-based signatures, isogeny signatures, lattice-based signatures, multivariate signatures, and others.

A multivariate public key cryptosystem (MPKC) has a set of quadratic polynomials over a finite field as its public key. Its security based on the difficulty of solving a system of multivariate quadratic polynomial equations over a finite field (MQ problem). See [9]. Garey and Johnson proved [17] that MQ problem is NP-complete in general.

The oil and vinegar and later derived unbalanced oil and vinegar signature schemes(UOV) [20,22], are well-known signature schemes known for their efficiency and short signature. The UOV scheme has withstood attacks for more than 20 years and is still regarded as a secure signature scheme. It is worth mentioning that the HFE scheme [21] also withstood long-term attacks, but was attacked by Tao et al. in [26]. Notably, the Rainbow signature scheme proposed by Ding and Schmidt [10], a multilayer UOV variant, was selected as a third-round finalist in the NIST PQC project. Although some parameters of Rainbow

M.-J. Saarinen and D. Smith-Tone (Eds.): PQCrypto 2024, LNCS 14772, pp. 79–91, 2024.
https://doi.org/10.1007/978-3-031-62746-0_4

schemes are broken by Beullens, see [4], the structure of UOV is still safe by now. However, both UOV and Rainbow suffer from the disadvantage of having large public key size compared to other PQC candidates, for example, lattice-based signature schemes.

For multivariate signature schemes, the size of public key mainly depends on the number of variables, the number of equations, and the size of the finite field. Depending on different influencing factors, there are different research approaches to develop UOV variants. The first approach does not change the original design of UOV scheme, but only changes the way of key generation. The compression technique [23] developed by Petzoldt et al., which is based on the fact that a part of public key can be arbitrarily chosen before generating the secret key. This implies that a part of public key can be generated using a seed of pseudo-random number generator and the size of public key mainly depends on the dimension of the oil space, the number of equations and the size of the finite field. Note that this technique can be applied to various UOV variants. The second approach is to use polynomials defined over small field as the public key, while the signature and message spaces are defined over the extension field, see LUOV in [5]. But several of its parameters were broken by Ding et al. [12]. The third approach is to reduce the dimension of oil space in the **KeyGen** step. In the **Sign** step, they use different methods to induce a new oil space from the original oil space such that the dimension of the new oil space is greater or equal to the number of equations, for example, QR-UOV [15], MAYO [3], SNOVA [28]. The authors of QR-UOV [15] construct oil space over the extension field then mapping it into the vector space over base field by trace function or tensor product, see also [18]. The signature and message spaces are defined over the base field. BAC-UOV [25] is similar with QR-UOV but it is broken by Furue et al. [16]. For MAYO [3], they increase the dimension of oil space by whipping up the oil and vinegar map $\mathcal{P} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ into a larger map $\mathcal{P}^* : \mathbb{F}_q^{kn} \to \mathbb{F}_q^m$. The authors of SNOVA [28] choose the noncommutative matrix $\mathcal{R}$ of $l \times l$ matrices over $\mathbb{F}_q$ to be the coefficient ring and they construct a UOV-like scheme with coefficients in $\mathcal{R}$. Actually, we can construct oil space in the space $\mathbb{F}_q^{nl}$ and make Kronecker product with $\mathbb{F}_q^l$ to map such oil space into a new oil space of $\mathcal{R}^n$.

Our Contribution. In this paper, our focus is on the multivariate signature SNOVA scheme [28]. We observe that an SNOVA(v, o, q, l) scheme over $\mathcal{R}$ can be viewed as a UOV(lv, lo, q) scheme with l^2o equations over $\mathbb{F}_q$, rather than a UOV(l^2v, l^2o, q) scheme over $\mathbb{F}_q$ as claimed by the authors in [28]. See Sect. 2.2. Consequently, we demonstrate that some parameters provided by the authors in SNOVA can't meet the NIST security level, and the complexities are lower than they claimed, see Table 1. Additionally, the coefficient matrices of these l^2o equations induced by the SNOVA(v, o, q, l) scheme exhibit special forms and are not randomly generated. In most cases, we observe that the l^2o equations induced by SNOVA have more solutions than l^2o random equations from a UOV scheme. Therefore, the actual complexity of SNOVA may be lower than theoretically estimated. Applying the same method, we find that NOVA [27] also has lower complexities claimed by the authors in their article.

2 SNOVA Scheme

2.1 Description of SNOVA Scheme

In [28], the authors introduce a UOV-type signature scheme over a noncommutative ring, which is called SNOVA.

Let v, o, l be positive integers with $v > o$ and $\mathbb{F}_q$ a finite field with q elements. Let $\mathcal{R}$ be the ring of $l \times l$ matrices over the finite field $\mathbb{F}_q$. Set $n = v + o$ and $m = o$, $\mathbf{x} = (x_1, \cdots, x_n)^t, \mathbf{u} = (u_1, \cdots, u_n)^t \in \mathcal{R}^n$, $[P], [F]$ denote some $n \times n$ matrices whose entries are elements of $\mathcal{R}$. For each $Q \in \mathcal{R}$, $[\Lambda_Q]$ denote the $n \times n$ matrix in $\mathbf{M}_{n\times n}(\mathcal{R})$ whose diagonal elements are Q.

The Space $\mathbb{F}_q[s]$. We first randomly choose an $l \times l$ symmetric matrix s such that the characteristic polynomial of s is irreducible. Set

$$\mathbb{F}_q[s] = \left\{a_0 + \cdots + a_{l-1}s^{l-1} : \ a_0, \cdots, a_{l-1} \in \mathbb{F}_q\right\}.$$

Note that $\dim_{\mathbb{F}_q} \mathbb{F}_q[s] = l$ and each nonzero element in $\mathbb{F}_q[s]$ is invertible and symmetric. In particular, $\mathbb{F}_q[s]$ is a subfield of $\mathcal{R}$.

Central Map. The central map of SNOVA scheme is $F = (F_1, \cdots, F_m) : \mathcal{R}^n \to \mathcal{R}^m$. Set $\Omega = \{(j,k) : \ 1 \le j,k \le n\} - \{(j,k) : \ m+1 \le j,k \le n\}$. For each i, F_i is the form of

$$\begin{aligned} F_i(x_1, \cdots, x_n) &= \sum_{\alpha=1}^{l^2} A_\alpha \cdot \Big(\sum_{(j,k)\in\Omega} x_j^t (Q_{\alpha 1} F_{i,jk} Q_{\alpha 2}) x_k \Big) \cdot B_\alpha \\ &= \sum_{\alpha=1}^{l^2} A_\alpha \cdot \mathbf{x}^t ([\Lambda_{Q_{\alpha 1}}][F_i][\Lambda_{Q_{\alpha 2}}])\mathbf{x} \cdot B_\alpha \end{aligned}$$

where $F_{i,jk}$ are randomly chosen from $\mathcal{R}$, A_α and B_α are invertible elements randomly chosen from $\mathcal{R}$, and $Q_{\alpha 1}, Q_{\alpha 2}$ are invertible elements randomly chosen from $\mathbb{F}_q[s]$. Indeed, $[F_i] = (F_{i,jk})$ is the form of

$$[F_i] = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & 0_{o\times o} \end{pmatrix} \in \mathbf{M}_{n\times n}(\mathcal{R}),$$

with $F_{11} \in \mathbf{M}_{v\times v}(\mathcal{R})$, $F_{12} \in \mathbf{M}_{v\times o}(\mathcal{R})$ and $F_{21} \in \mathbf{M}_{o\times v}(\mathcal{R})$.

Public Key and Private Key. Let $T : \mathcal{R}^n \to \mathcal{R}^n$ be the map corresponding to the matrix

$$[T] = \begin{pmatrix} I_{v\times v} & T_{v\times o} \\ 0 & I_{o\times o} \end{pmatrix},$$

where $T_{v\times o}$ is a $v \times o$ matrix whose entries are chosen randomly from $\mathbb{F}_q[s]$. $I_{v\times v}$ and $I_{o\times o}$ are the diagonal matrices with all entries being the identity matrix in $\mathcal{R}$.

Let $P = F \circ T$. Set $\mathbf{x} = [T] \cdot \mathbf{u}$ and $P_i = F_i \circ T$. We get

$$\begin{aligned} P_i(\mathbf{u}) &= \sum_{\alpha=1}^{l^2} \sum_{d_j=1}^{n} \sum_{d_k=1}^{n} A_\alpha \cdot u_{d_j}^t (Q_{\alpha 1} P_{i,d_j d_k} Q_{\alpha 2}) u_{d_k} \cdot B_\alpha \\ &= \sum_{\alpha=1}^{l^2} A_\alpha \cdot \mathbf{u}^t ([\Lambda_{Q_{\alpha 1}}][P_i][\Lambda_{Q_{\alpha 2}}])\mathbf{u} \cdot B_\alpha \end{aligned}$$

where $P_{i,d_j d_k} = \sum_{(j,k)\in\Omega} t_{j,d_j} \cdot F_{i,jk} \cdot t_{k,d_k}$. Note that

$$[P_i] = [P_{i,d_j d_k}] = [T]^t [F_i][T], \ i = 1, \cdots, m.$$

The public key of SNOVA consists of the map $P : \mathcal{R}^n \to \mathcal{R}^m$, i.e., the corresponding matrices $[P_i]$ for $i = 1, \cdots, m$, and matrices $A_\alpha, B_\alpha, Q_{\alpha k}$ for $\alpha = 1, \cdots, l^2$ and $k = 1, 2$. The private key of SNOVA is (F, T), i.e., the matrix $[T]$ and the matrices $[F_i]$ for $i = 1, \cdots, m$.

Signature. Let **Message** be the message to be signed. Set $Hash(\mathbf{Message}) = \mathbf{y} = (y_1, \cdots, y_m)^t \in \mathcal{R}^m$. We first choose random values $a_1, \cdots, a_v \in \mathcal{R}$ as the vinegar variables. Then, the following equation $F(\mathbf{a}, \mathbf{x}_o) = \mathbf{y}$ is a linear system of $\mathbf{x}_o \in \mathcal{R}^o$, and we can obtain a solution $\mathbf{x}_o = (a_{v+1}, \cdots, a_n)$ for the equation

$$F(a_1, \cdots, a_v, x_{v+1}, \cdots, x_n) = \mathbf{y}.$$

If there is no solution to the equation, we choose new random values $a'_1, \cdots, a'_v \in \mathcal{R}$ and repeat the procedure. Set $\mathbf{x} = (a_1, \cdots, a_v, a_{v+1}, \cdots, a_n)^t$. Secondly, the signature is $\mathbf{sign} = T^{-1}(\mathbf{x})$.

Verification. Let $\mathbf{sign} = (s_1, \cdots, s_n)$ be the signature to be verified. If $Hash(\mathbf{Message}) = P(\mathbf{sign})$, then the signature is accepted, otherwise rejected.

2.2 Structure of SNOVA

The authors assert in [28] that an SNOVA(v, o, q, l) scheme over $\mathcal{R}$ can be considered as a UOV$(l^2 v, l^2 o, q)$ scheme over $\mathbb{F}_q$. However, we argue that it should only be regarded as a UOV(lv, lo, q) scheme with $l^2 o$ equations over $\mathbb{F}_q$. In the second part of this section, we claim that an SNOVA(v, o, q, l) can induce a standard UOV(v, o, q^l) scheme. Here standard means that the dimension of oil space equals to the number of equations.

Claim 1. An SNOVA(v, o, q, l) scheme can be regarded as a UOV(lv, lo, q) scheme with $l^2 o$ equations over $\mathbb{F}_q$.

In fact, all the matrices $[F_i]$, $[T]$, and $[P_i]$ in the SNOVA scheme can be viewed as $ln \times ln$ matrices in $\mathbf{M}_{ln\times ln}(\mathbb{F}_q)$ with

$$[P_i] = [T]^t \cdot [F_i] \cdot [T] \in \mathbf{M}_{ln\times ln}(\mathbb{F}_q), \quad i = 1, \cdots, m.$$

Based on the design of the central map F, the lower-right $lo \times lo$ block is zero block for each $[F_i]$. Therefore there exists a common oil space of $[P_i]$ over $\mathbb{F}_q$ with dimension lo for all i. Set

$$\mathcal{O}_1 = \left\{(0, \cdots, 0, a_{lv+1}, \cdots, a_{ln})^t \in \mathbb{F}_q^{ln} : \ a_i \in \mathbb{F}_q\right\} \text{ and } \mathcal{O} = [T]^{-1}(\mathcal{O}_1) \subset \mathbb{F}_q^{ln}.$$

Note that $\dim_{\mathbb{F}_q} \mathcal{O} = lo$ and for any $\mathbf{u}, \mathbf{v} \in \mathcal{O}$, $0 \leq j, k \leq l-1$, we have

$$\mathbf{u}^t \cdot \left([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\right) \cdot \mathbf{v} = 0 \in \mathbb{F}_q \text{ for } i = 1, \cdots, m. \tag{2.1}$$

That is, each $[\Lambda_{s^j}][P_i][\Lambda_{s^k}]$ sends $\mathcal{O}$ into its own orthogonal complement $\mathcal{O}^{\perp}$. Since $[T]$ and $[\Lambda_s]$ are commutative and $\mathcal{O}_1$ is stable under $[\Lambda_s]$, $\mathcal{O}$ is stable under $[\Lambda_s]$. Therefore an SNOVA(v, o, q, l) scheme induces a UOV(lv, lo, q) scheme whose oil space is $\mathcal{O}$ with l^2o equations given by (2.1).

Next, we will explain how we use the oil space $\mathcal{O}$ of the UOV(lv, lo, q) scheme with l^2o equations given by (2.1) obtained from a SNOVA(v, o, q, l) to recover the oil space of SNOVA scheme. Set

$$\mathcal{O} \otimes \mathbb{F}_q^l := \{ \sum_{1 \leq i \leq lo, 1 \leq j \leq l} a_{ij}\mathbf{u}_i \otimes e_j^t \in \mathcal{R}^n : \ \mathbf{u}_i \in \mathcal{O}, e_j \in \mathbb{F}_q^l, a_{ij} \in \mathbb{F}_q\} \subset \mathcal{R}^n,$$

where $\{\mathbf{u}_i\}_{1 \leq i \leq lo}$(resp. $\{e_j\}_{1 \leq j \leq l}$) is a basis of $\mathcal{O}$(resp. $\mathbb{F}_q^l$) over $\mathbb{F}_q$ and $\otimes$ denotes the Kronecker product of matrices. We have $\dim_{\mathbb{F}_q} \mathcal{O} \otimes \mathbb{F}_q^l = lo^2$ and for any $\mathbf{x} \in \mathcal{O} \otimes \mathbb{F}_q^l$,

$$P(\mathbf{x}) = 0 \in \mathcal{R}^m.$$

Indeed, for any $Q_{\alpha 1}, Q_{\alpha 2} \in \mathbb{F}_q[s]$, $[\Lambda_{Q_{\alpha 1}}][P_i][\Lambda_{Q_{\alpha 2}}]$ can be written as a linear combination of $\{[\Lambda_{s^j}][P_i][\Lambda_{s^k}]\}_{0 \leq j,k \leq l-1}$, and each column of $\mathbf{x}$ in $\mathcal{O} \otimes \mathbb{F}_q^l$ belongs to $\mathcal{O}$. By (2.1), we have $\mathbf{x}^t([\Lambda_{Q_{\alpha 1}}][P_i][\Lambda_{Q_{\alpha 2}}])\mathbf{x} = 0 \in \mathcal{R}$ for each i. Hence $P(\mathbf{x}) = 0 \in \mathcal{R}^m$.

Combining with the fact that $\dim_{\mathbb{F}_q} \mathcal{O} \otimes \mathbb{F}_q^l = lo^2$, a UOV$(lv, lo, q)$ scheme with l^2o equations over $\mathbb{F}_q$ obtained from a SNOVA(v, o, q, l) can easily recover the oil space $\mathcal{O} \otimes \mathbb{F}_q^l$ of SNOVA scheme. Therefore we only need to consider the system of Eq. (2.1), which is a UOV(lv, lo, q) scheme with l^2o equations over $\mathbb{F}_q$.

Claim 2. An SNOVA(v, o, q, l) scheme can induce a standard UOV(v, o, q^l) scheme.

We know that all the eigenvalues of s lie in $\mathbb{F}_{q^l}$ due to the characteristic polynomial of s being irreducible. Let $\lambda \in \mathbb{F}_{q^l}$ be an eigenvalue of s and $\xi \in (\mathbb{F}_{q^l})^l$ an eigenvector corresponding to λ. Let τ be the Frobenius element $z \mapsto z^q$ in the Galois group $\mathrm{Gal}(\mathbb{F}_{q^l}/\mathbb{F}_q)$. For $j = 0, \cdots, l-1$, we have

$$s\tau^j(\xi) = \tau^j(\lambda)\tau^j(\xi).$$

Thus for each j, $\tau^j(\xi)$ is an eigenvector corresponding to the eigenvalue $\tau^j(\lambda)$. In particular, $\{\xi, \tau^1(\xi), \cdots, \tau^{l-1}(\xi)\}$ are linear independent and so

$$\mathrm{Tr}(\xi) := \sum_{j=0}^{l-1} \tau^j(\xi) \in \mathbb{F}_q^l - \{0\}.$$

Lemma 2.1. *With the notations above. Suppose that*

$$\mathcal{O}_1 = \left\{(0, \cdots, 0, a_{lv+1}, \cdots, a_{ln})^t \in \mathbb{F}_q^{ln} : \ a_i \in \mathbb{F}_q\right\} \ \text{and} \ \mathcal{O} = [T]^{-1}(\mathcal{O}_1) \subset \mathbb{F}_q^{ln}.$$

There is a subspace $\mathcal{O}_2$ of $\mathbb{F}_{q^l}^{nl}$ such that

$$\mathcal{O} \otimes_{\mathbb{F}_q} \mathbb{F}_{q^l} = \mathcal{O}_2 \oplus \tau(\mathcal{O}_2) \oplus \cdots \oplus \tau^{l-1}(\mathcal{O}_2) \tag{2.2}$$

and $\dim_{\mathbb{F}_{q^l}} \mathcal{O}_2 = o$, where τ is induced from the Frobenius element. Here we use the same notation.

Proof. Take

$$\mathcal{O}_1' := \left\{(0, \cdots, 0, a_{v+1}\xi^t, \cdots, a_n\xi^t)^t \in \mathbb{F}_{q^l}^{ln} : \ a_i \in \mathbb{F}_{q^l}\right\} \subset \mathcal{O}_1 \otimes_{\mathbb{F}_q} \mathbb{F}_{q^l}$$

and

$$\mathcal{O}_2 := [T]^{-1}(\mathcal{O}_1') \subset \mathcal{O} \otimes_{\mathbb{F}_q} \mathbb{F}_{q^l}.$$

It is easy to know $\dim_{\mathbb{F}_{q^l}} \mathcal{O}_2 = o$. We claim that such $\mathcal{O}_2$ is what we want. Indeed, since each entry of $[T]$ belongs to $\mathbb{F}_q[s]$, we have $\mathcal{O} \otimes_{\mathbb{F}_q} \mathbb{F}_{q^l} = [T]^{-1}(\mathcal{O}_1 \otimes_{\mathbb{F}_q} \mathbb{F}_{q^l})$ and

$$[T]^{-1}(\tau^k(\mathcal{O}_1')) = \tau^k([T]^{-1}(\mathcal{O}_1')) = \tau^k(\mathcal{O}_2)$$

for $k = 1, \cdots, l-1$. Thus, it is necessary to show

$$\mathcal{O}_1 \otimes_{\mathbb{F}_q} \mathbb{F}_{q^l} = \mathcal{O}_1' \oplus \tau(\mathcal{O}_1') \oplus \cdots \oplus \tau^{l-1}(\mathcal{O}_1').$$

Since $\{\xi, \tau^1(\xi), \cdots, \tau^{l-1}(\xi)\}$ are linear independent, we have

$$\mathbb{F}_{q^l}^l = \mathbb{F}_{q^l} \cdot \xi \oplus \mathbb{F}_{q^l} \cdot \tau(\xi) \oplus \cdots \oplus \mathbb{F}_{q^l} \cdot \tau^{l-1}(\xi).$$

Thus

$$\begin{aligned}\mathcal{O}_{1i} :&= \left\{(0, \cdots, 0, a_{l(i-1)+1}, \cdots, a_{li}, 0, \cdots, 0)^t : \ a_{li+j} \in \mathbb{F}_{q^l}\right\} \cong \mathbb{F}_{q^l}^l \\ &= \mathcal{O}_{1i}' \oplus \tau(\mathcal{O}_{1i}') \cdots \oplus \tau^{l-1}(\mathcal{O}_{1i}'),\end{aligned}$$

where $\mathcal{O}_{1i}' := \left\{(0, \cdots, 0, a_i\xi^t, 0, \cdots, 0, \cdots, 0)^t : \ a_i \in \mathbb{F}_{q^l}\right\} \cong \mathbb{F}_{q^l}$. Then we have

$$\begin{aligned}\mathcal{O}_1 \otimes_{\mathbb{F}_q} \mathbb{F}_{q^l} = \bigoplus_{i=v+1}^{n} \mathcal{O}_{1i} &= \bigoplus_{i=v+1}^{n} \left(\mathcal{O}_{1i}' \oplus \tau(\mathcal{O}_{1i}') \oplus \cdots \oplus \tau^{l-1}(\mathcal{O}_{1i}')\right) \\ &= \bigoplus_{i=v+1}^{n} \mathcal{O}_{1i}' \oplus \bigoplus_{i=v+1}^{n} \tau(\mathcal{O}_{1i}') \oplus \cdots \oplus \bigoplus_{i=v+1}^{n} \tau^{l-1}(\mathcal{O}_{1i}') \\ &= \mathcal{O}_1' \oplus \tau(\mathcal{O}_1') \oplus \cdots \oplus \tau^{l-1}(\mathcal{O}_1').\end{aligned}$$

□

Note that each element of $\mathcal{O}_2$ has the form of

$$\mathbf{u} = (\lambda_1\xi^t, \cdots, \lambda_v\xi^t, a_1\xi^t, \cdots, a_o\xi^t)^t \in \mathbb{F}_{q^l}^{ln} \text{ with } \lambda_i, a_i \in \mathbb{F}_{q^l} \tag{2.3}$$

due to the fact that $[T]^{-1} \in \mathbf{M}_{n\times n}(\mathbb{F}_q[s])$ and ξ is a common eigenvector of the elements of $\mathbb{F}_q[s]$. We claim that many equations in (2.1) are redundant if $\mathbf{u}$ and $\mathbf{v}$ are elements in $\mathcal{O}_2$. Indeed, take $\mathbf{u}, \mathbf{v} \in \mathcal{O}_2$, we have

$$\mathbf{u}^t \cdot \left([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\right) \cdot \mathbf{v} = \lambda^{j+k}\mathbf{u}^t \cdot [P_i] \cdot \mathbf{v},$$

where λ is the eigenvalue corresponding to ξ. Thus $\mathbf{u}^t \cdot [P_i] \cdot \mathbf{v} = 0$ will imply that $\mathbf{u}^t \cdot \left([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\right) \cdot \mathbf{v} = 0$ for $0 \le j, k \le l-1$ and only m equations in (2.1) are effective when $\mathbf{u}, \mathbf{v} \in \mathcal{O}_2$. Therefore, after linear transformation, there is a standard UOV(v, o, q^l) scheme induced by SNOVA(v, o, q, l) scheme. In fact, take

$$\mathcal{O}_2' = \{(a_1, \cdots, a_n) : \ (a_1\xi^t, \cdots, a_n\xi^t)^t \in \mathcal{O}_2\} \subset \mathbb{F}_{q^l}^n.$$

It is easy to know $\dim \mathcal{O}_2' = o$. Let $[P_i] = [P_{i,jk}]$ be the public key of SNOVA scheme with $P_{i,jk} \in \mathcal{R}$. Set

$$\widetilde{P}_{i,jk} = \xi^t \cdot P_{i,jk} \cdot \xi, \quad [\widetilde{P}_i] = (\widetilde{P}_{i,jk}) \in \mathbf{M}_{n\times n}(\mathbb{F}_{q^l}). \tag{2.4}$$

We have

$$\tilde{\mathbf{u}}^t \cdot [\widetilde{P}_i] \cdot \tilde{\mathbf{v}} = (\tilde{\mathbf{u}} \otimes \xi)^t \cdot [P_i] \cdot (\tilde{\mathbf{v}} \otimes \xi) = 0 \tag{2.5}$$

for any $\tilde{\mathbf{u}}, \tilde{\mathbf{v}} \in \mathcal{O}_2'$ and $i = 1, \cdots, m$. The second equality holds because $\tilde{\mathbf{u}} \otimes \xi, \tilde{\mathbf{v}} \otimes \xi \in \mathcal{O}_2 \subset \mathcal{O} \otimes_{\mathbb{F}_q} \mathbb{F}_{q^l}$. Hence, the public keys of the induced UOV(v, o, q^l) scheme are given by $\{[\widetilde{P}_i]\}_{i=1}^m$, as defined in (2.4). The oil space of the induced UOV scheme is $\mathcal{O}_2'$, and the elements of $\mathcal{O}_2'$ satisfy equation (2.5).

Remark 2.1. If there is only one subspace of $\mathbb{F}_{q^l}^n$, whose elements satisfy equation (2.5), and the dimension of such subspace is o, it must equal to the oil space $\mathcal{O}_2'$ of the induced UOV(v, o, q^l) scheme described above. Then the oil space of SNOVA can be recovered by (2.2). Applying equivalent key attack to the induced UOV(v, o, q^l) scheme, we need solve a system of o^3 quadratic equations in vo variables over $\mathbb{F}_{q^l}$. According to the parameters given in [28], o^3 is always much larger than vo. Therefore, it may be effective for us to attack SNOVA by recovering the oil space of the induced UOV scheme. But we are not going to use the induced UOV(v, o, q^l) scheme to give specific complexity of SNOVA(v, o, q, l) scheme.

3 Security Analysis

3.1 Complexity

Given a homogeneous multivariate quadratic map $P : \mathbb{F}_q^N \to \mathbb{F}_q^M$, we use $MQ(N, M, q)$ to denote the complexity of finding a non-trivial solution $\mathbf{u}$ satisfying $P(\mathbf{u}) = 0$ if such solution exists. Several algorithms for algebraically

solving the quadratic system by computing Gröbner basis [6] include F_4 [13], F_5 [14] and XL [8]. In this paper, we estimate the complexity of solving M homogeneous quadratic equations in N variables [7] as

$$3 \cdot \binom{N-1+d_{reg}}{d_{reg}}^2 \cdot \binom{N+1}{2}$$

field multiplications, where d_{reg} is equal to the degree of the first non-positive term in the series generated by

$$\frac{(1-t^2)^M}{(1-t)^N}.$$

The hybrid approach [1], which randomly guesses k ($k = 0, \cdots, N$) variables before computing a Gröbner basis. Hence the complexity are

$$\min_k q^k \cdot MQ(N-k+1, M, q)$$

field multiplications.

An underdetermined system can be reduced to an overdetermined system, then apply hybrid approach. There are many approaches listed in [28].

3.2 K-S Attack

In the UOV(v, o, q) scheme, the K-S attack [20] obtains the oil space. To obtain the oil space, the K-S attack chooses two invertible matrices W_1, W_2 from the set of linear combinations of the public keys $P_1, \cdots, P_m$ of the UOV scheme. Then, it probabilistically recovers a part of the oil space. The complexity of K-S attack is estimated by

$$\text{Comp}_{\text{K-S}}\ \text{UOV} = q^{v-o}$$

field multiplications.

In the SNOVA scheme, we have claimed that SNOVA(v, o, q, l) scheme over $\mathcal{R}$ can be regarded as a UOV(lv, lo, q) scheme in Claim 1. In such case, we have

$$\text{Comp}_{\text{K-S}}\ \text{SNOVA} = q^{lv-lo}$$

field multiplications.

3.3 Reconciliation Attack

The reconciliation attack [11] for UOV is similar to the K-S attack, trying to find an element of the oil space and hence basis of oil space can be recovered. In Sect. 2.2, we have proved that SNOVA(v, o, q, l) scheme can be regarded as a UOV(lv, lo, q) scheme with l^2o equations and the elements of oil space satisfy equation (2.1). Therefore, the reconciliation attack can be decomposed into a

series of steps. Firstly, we may find an element $\mathbf{u} = (u_1, \cdots, u_{lv}, 0, \cdots, 0, 1)^t \in \mathbb{F}_q^{ln}$ such that

$$\mathbf{u}^t \cdot \Big([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\Big) \cdot \mathbf{u} = 0 \in \mathbb{F}_q \tag{3.1}$$

for $i = 1, \cdots, o$ and $0 \leq j, k \leq l-1$. There are $l^2 o$ homogeneous quadratic equations in $lv+1$ variables in (3.1). Secondly, using Eq. (2.1), we get $2 \cdot o \cdot l^2$ linear equations for the other elements of $\mathcal{O}$. Hence the complexity of reconciliation attack is mainly centered on solving Eq. (3.1). Note that in the case of $vl + 1 > l^2 o$, Eq. (3.1) has a lot of solutions not in the space $\mathcal{O}$. Therefore, the complexity of the reconciliation attack is evaluated by

$$\text{Comp}_{\text{Reconciliation}}\ \text{SNOVA} = \min q^k MQ(lv + 1 - k, l^2 o, q), \tag{3.2}$$

where $\max\{0, lv+1-l^2 o\} \leq k \leq lv$ is the number of fixed variables in the hybrid approach.

Remark 3.1. We observe that finding solutions to Eq. (3.1) is easier in the extension field $\mathbb{F}_{q^l}$. This phenomenon does not exist in the UOV scheme. In the following, we will explain our observation. Indeed, applying the standard discussion of the reconciliation attack to the induced UOV(v, o, q^l) scheme in Claim 2, with the same notations as Sect. 2.2, Eq. (3.1) will reduce to the following equation

$$\tilde{\mathbf{u}}^t \cdot [\widetilde{P}_i] \cdot \tilde{\mathbf{u}} = 0,\ \ \tilde{\mathbf{u}} = (\lambda_1, \cdots, \lambda_v, 0, \cdots, 0, 1)^t \in \mathbb{F}_{q^l}^n \tag{3.3}$$

when we take

$$\mathbf{u} = \tilde{\mathbf{u}} \otimes \xi = (\lambda_1 \xi^t, \cdots, \lambda_v \xi^t, 0, \cdots, 0, \xi^t)^t \in \mathbb{F}_{q^l}^{ln}$$

in (3.1). There are only m quadratic equations and v variables over $\mathbb{F}_{q^l}$ in (3.3).

Unfortunately, we only prove the fact that Eq. (3.1) is easier to find a solution over the extension field $\mathbb{F}_{q^l}$. We don't know how to use the solution over the extension field $\mathbb{F}_{q^l}$ to get the solution over the base field $\mathbb{F}_q$. Thus, we will not use the induced UOV(v, o, q^l) scheme to give the specific complexity of SNOVA scheme. Indeed, we only have the following lemma:

Lemma 3.1. *Let* $\mathbf{u} = (\lambda_1 \xi^t, \cdots, \lambda_v \xi^t, 0, \cdots, 0, \xi^t)^t \in \mathbb{F}_{q^l}^{ln}$ *satisfying*

$$\mathbf{u}^t \cdot [P_i] \cdot \tau^j(\mathbf{u}) = 0 \in \mathbb{F}_{q^l} \tag{3.4}$$

for $i = 1, \cdots, m$ *and* $j = 0, \cdots, l-1$. *Set*

$$\mathbf{v} := \text{Tr}(\mathbf{u}) = \mathbf{u} + \tau(\mathbf{u}) + \cdots + \tau^{l-1}(\mathbf{u}) \in \mathbb{F}_q^{ln}.$$

We have $\mathbf{v} \neq 0$ *and*

$$\mathbf{v}^t \cdot \Big([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\Big) \cdot \mathbf{v} = 0$$

for $i = 1, \cdots, m$ *and* $0 \leq j, k \leq l-1$.

Proof. Note that $\mathbf{v} \neq 0$ due to $\mathrm{Tr}(\xi) \neq 0$. We have

$$\begin{aligned}\mathbf{v}^t \cdot \left([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\right) \cdot \mathbf{v} &= \sum_{0 \le a,b \le l-1} \tau^a(\mathbf{u}^t) \cdot \left([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\right) \cdot \tau^b(\mathbf{u}) \\ &= \sum_{0 \le a,b \le l-1} \tau^a(\lambda^j)\tau^b(\lambda^k)\tau^a\left(\mathbf{u}^t \cdot [P_i] \cdot \tau^{b-a}(\mathbf{u})\right) \\ &= 0.\end{aligned}$$

The last equality holds because Eq. (3.4) implies that $\mathbf{u}^t \cdot [P_i] \cdot \tau^c(\mathbf{u}) = 0$ for any $c \in \mathbb{Z}$. □

According to Lemma 3.1 above, if we want to get a solution over $\mathbb{F}_q$ from a solution $\mathbf{u}$ over $\mathbb{F}_{q^l}$ for Eq. (3.1), maybe we need to solve Eq. (3.4). But the equations in (3.4) are of degree $1 + q^j$. We could not apply the complexity formula in §3.1 directly.

3.4 Intersection Attack

Beullens proposed a new attack against UOV called the intersection attack in [2]. The intersection attack attempts to obtain an equivalent key by recovering the subspace $\mathcal{O}$ defined in Sect. 2.2. Let M_1, M_2 be two invertible matrices in the set of linear combinations of $\{[\Lambda_{s^j}][P_i][\Lambda_{s^k}]\}_{1 \le i \le m, 0 \le j,k \le l-1}$. By (2.1), we know that $M_1\mathcal{O}$ and $M_2\mathcal{O}$ are both subspaces of $\mathcal{O}^\perp$. Although $M_1\mathcal{O} \neq M_2\mathcal{O}$, we still have

$$\begin{aligned}\dim(M_1\mathcal{O} \cap M_2\mathcal{O}) &= \dim(M_1\mathcal{O}) + \dim(M_2\mathcal{O}) - \dim(M_1\mathcal{O} + M_2\mathcal{O}) \\ &\ge 2lo - \dim(\mathcal{O}^\perp) \\ &= 2lo - lv.\end{aligned}$$

In the Case of $2o > v$. Let $\mathbf{x}$ be an element in the intersection $M_1\mathcal{O} \cap M_2\mathcal{O}$, then both $M_1^{-1}\mathbf{x}$ and $M_2^{-1}\mathbf{x}$ are in $\mathcal{O}$. Therefore, $\mathbf{x}$ is a solution to the following system of quadratic equations

$$\begin{cases} (M_1^{-1}\mathbf{x})^t \cdot \left([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\right) \cdot (M_1^{-1}\mathbf{x}) = 0 \\ (M_2^{-1}\mathbf{x})^t \cdot \left([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\right) \cdot (M_2^{-1}\mathbf{x}) = 0 \\ (M_1^{-1}\mathbf{x})^t \cdot \left([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\right) \cdot (M_2^{-1}\mathbf{x}) = 0 \\ (M_2^{-1}\mathbf{x})^t \cdot \left([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\right) \cdot (M_1^{-1}\mathbf{x}) = 0 \end{cases} \tag{3.5}$$

Note that the third and the fourth equations in (3.5) are same when $[P_i]$ is symmetric. [19] and [28] both pointed out that there are $2l$ redundant equations in (3.5), see also [2]. Since there is a $2lo - lv$ dimensional subspace of solutions, we can impose $2lo - lv$ affine constraints on $\mathbf{x}$. Then the attack is reduced to find a solution to the above system of $4l^2o - 2l$ quadratic homogeneous equations in $ln - (2lo - lv - 1) = 2lv - lo + 1$ variables. Therefore the complexity is

$$\text{Comp}_{\text{Intersection}} = \min_k q^k MQ(2lv - lo + 1 - k, 4l^2o - 2l, q) \tag{3.6}$$

field multiplications, where k is the number of fixed variables in the hybrid approach.

In the Case of $2o \leq v$. The intersection $M_1\mathcal{O} \cap M_2\mathcal{O}$ may have no nontrivial vector. If $M_1\mathcal{O}$ and $M_2\mathcal{O}$ are uniformly random subspaces of $\mathcal{O}^\perp$, then the probability that they have non-trivial intersection is approximately $q^{-lv+2lo-1}$. Therefore, the attack becomes a probabilistic algorithm for solving the system (3.5) with a probability of approximately $q^{-lv+2lo-1}$. Therefore the complexity is

$$\text{Comp}_{\text{Intersection}} = \min_k q^{lv-2lo+1} q^k MQ(ln - k + 1, 4l^2o - 2l, q) \tag{3.7}$$

field multiplications, where k is the number of fixed variables in the hybrid approach.

3.5 Security

Table 1 presents the classical complexity of respective attacks against the parameters submitted in [28], where the number of gates required for an attack is computed by

$$\#\text{gates} = \#\text{field multiplications} \cdot (2 \cdot (\log_2 q)^2 + \log_2 q).$$

In each pair of complexities, the left one denotes the complexity using the analysis results in this article, the right one denotes the complexity given by [28], where k is the number of fixed variables in the hybrid approach. Complexities that do not meet the security level of the NIST PQC project are highlighted in bold fonts. Furthermore, Table 1 also indicates that the complexity of SNOVA is generally lower than what the authors claimed in [28].

Table 1. Table of classical complexity in $\log_2(\#\text{gates})$

SL	(v, o, q, l)	K-S	Reconciliation	Intersection
I	(28, 17, 16, 2)	**93**/181	**132**/192 ($k = 2$)	**83**/275 ($k = 0$)
	(25, 8, 16, 3)	209/617	201/231 ($k = 15$)	221/819 ($k = 0$)
	(24, 5, 16, 4)	309/1221	270/286 ($k = 30$)	349/1439 ($k = 0$)
III	(43, 25, 16, 2)	**149**/293	**193**/279 ($k = 6$)	**116**/439 ($k = 0$)
	(49, 11, 16, 3)	461/1373	438/530 ($k = 66$)	529/1631 ($k = 0$)
	(37, 8, 16, 4)	469/1861	388/424 ($k = 45$)	507/2192 ($k = 0$)
V	(61, 33, 16, 2)	**229**/453	277/386 ($k = 17$)	**166**/727 ($k = 0$)
	(66, 15, 16, 3)	617/1841	575/707 ($k = 87$)	690/2178 ($k = 0$)
	(60, 10, 16, 4)	805/3205	695/812 ($k = 112$)	922/3602 ($k = 0$)

Acknowledgments. This work is supported by National Key R&D Program of China (No. 2021YFB3100100).

References

1. Bettale, L., Faugere, J.C., Perret, L.: Hybrid approach for solving multivariate systems over finite fields. J. Math. Cryptol. **3**(3), 177–197 (2009)
2. Beullens, W.: Improved cryptanalysis of UOV and rainbow. In: Canteaut, A., Standaert, F.-X. (eds.) EUROCRYPT 2021. LNCS, vol. 12696, pp. 348–373. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77870-5_13
3. Beullens, W.: MAYO: practical post-quantum signatures from oil-and-vinegar maps. In: AlTawy, R., Hülsing, A. (eds.) SAC 2021. LNCS, vol. 13203, pp. 355–376. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-99277-4_17
4. Beullens, W.: Breaking rainbow takes a weekend on a laptop. In: Dodis, Y., Shrimpton, T. (eds.) CRYPTO 2022. LNCS, vol. 13508, pp. 464–479. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-15979-4_16
5. Beullens, W., Preneel, B.: Field lifting for smaller UOV public keys. In: Patra, A., Smart, N.P. (eds.) INDOCRYPT 2017. LNCS, vol. 10698, pp. 227–246. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-71667-1_12
6. Buchberger, B.: Ein algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal. Ph.D. thesis, Math. Inst., University of Innsbruck (1965)
7. Cheng, C.-M., Chou, T., Niederhagen, R., Yang, B.-Y.: Solving quadratic equations with XL on parallel architectures. In: Prouff, E., Schaumont, P. (eds.) CHES 2012. LNCS, vol. 7428, pp. 356–373. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33027-8_21
8. Courtois, N., Klimov, A., Patarin, J., Shamir, A.: Efficient algorithms for solving overdefined systems of multivariate polynomial equations. In: Preneel, B. (ed.) EUROCRYPT 2000. LNCS, vol. 1807, pp. 392–407. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-45539-6_27
9. Ding, J., Gower, J.E., Schmidt, D.S.: Multivariate Public Key Cryptosystems, vol. 25. Springer, Heidelberg (2006)
10. Ding, J., Schmidt, D.: Rainbow, a new multivariable polynomial signature scheme. In: Ioannidis, J., Keromytis, A., Yung, M. (eds.) ACNS 2005. LNCS, vol. 3531, pp. 164–175. Springer, Heidelberg (2005). https://doi.org/10.1007/11496137_12
11. Ding, J., Yang, B.-Y., Chen, C.-H.O., Chen, M.-S., Cheng, C.-M.: New differential-algebraic attacks and reparametrization of rainbow. In: Bellovin, S.M., Gennaro, R., Keromytis, A., Yung, M. (eds.) ACNS 2008. LNCS, vol. 5037, pp. 242–257. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-68914-0_15
12. Ding, J., Zhang, Z., Deaton, J., Schmidt, K., Vishakha, F.: New attacks on lifted unbalanced oil vinegar. In: the 2nd NIST PQC Standardization Conference (2019)
13. Faugere, J.C.: A new efficient algorithm for computing Gröbner bases (F4). J. Pure Appl. Algebra **139**(1–3), 61–88 (1999)
14. Faugere, J.C.: A new efficient algorithm for computing Gröbner bases without reduction to zero (F5). In: Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation, pp. 75–83 (2002)
15. Furue, H., Ikematsu, Y., Kiyomura, Y., Takagi, T.: A new variant of unbalanced oil and vinegar using quotient ring: QR-UOV. In: Tibouchi, M., Wang, H. (eds.) ASIACRYPT 2021, Part IV. LNCS, vol. 13093, pp. 187–217. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-92068-5_7

16. Furue, H., Kinjo, K., Ikematsu, Y., Wang, Y., Takagi, T.: A structural attack on block-anti-circulant UOV at SAC 2019. In: Ding, J., Tillich, J.-P. (eds.) PQCrypto 2020. LNCS, vol. 12100, pp. 323–339. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-44223-1_18
17. Garey, M.R., Johnson, D.S.: Computers and Intractability, vol. 174. Freeman, San Francisco (1979)
18. Hashimoto, Y.: An elementary construction of QR-UOV. Cryptology ePrint Archive (2022). https://eprint.iacr.org/2022/145
19. Ikematsu, Y., Akiyama, R.: Revisiting the security analysis of SNOVA. Cryptology ePrint Archive (2024). https://eprint.iacr.org/2024/096
20. Kipnis, A., Shamir, A.: Cryptanalysis of the oil and vinegar signature scheme. In: Krawczyk, H. (ed.) CRYPTO 1998. LNCS, vol. 1462, pp. 257–266. Springer, Cham (1998)
21. Patarin, J.: Hidden fields equations (HFE) and isomorphisms of polynomials (IP): two new families of asymmetric algorithms. In: Maurer, U. (ed.) EUROCRYPT 1996. LNCS, vol. 1070, pp. 33–48. Springer, Heidelberg (1996). https://doi.org/10.1007/3-540-68339-9_4
22. Patarin, J.: The oil and vinegar algorithm for signatures. In: Dagstuhl Workshop on Cryptography (1997)
23. Petzoldt, A., Bulygin, S., Buchmann, J.: CyclicRainbow – a multivariate signature scheme with a partially cyclic public key. In: Gong, G., Gupta, K.C. (eds.) INDOCRYPT 2010. LNCS, vol. 6498, pp. 33–48. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-17401-8_4
24. Shor, P.W.: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM Rev. **41**(2), 303–332 (1999)
25. Szepieniec, A., Preneel, B.: Block-anti-circulant unbalanced oil and vinegar. In: Paterson, K.G., Stebila, D. (eds.) SAC 2019. LNCS, vol. 11959, pp. 574–588. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-38471-5_23
26. Tao, C., Petzoldt, A., Ding, J.: Efficient key recovery for all HFE signature variants. In: Malkin, T., Peikert, C. (eds.) CRYPTO 2021, Part I. LNCS, vol. 12825, pp. 70–93. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-84242-0_4
27. Wang, L.C., Tseng, P.E., Kuan, Y.L., Chou, C.Y.: NOVA, a noncommutative-ring based unbalanced oil and vinegar signature scheme with key-randomness alignment. Cryptology ePrint Archive (2022). https://eprint.iacr.org/2022/665
28. Wang, L.C., Tseng, P.E., Kuan, Y.L., Chou, C.Y.: A simple noncommutative UOV scheme. Cryptology ePrint Archive (2022). https://eprint.iacr.org/2022/1742.pdf