Cryptanalysis of the SNOVA Signature Scheme

Peigen Li

Beijing Institute of Mathematical Sciences and Applications

Joint work with Jingtai Ding
lpg22@bimsa.cn

PQcrypto-June 12,2024

Overview

SNOVA is a multivariate signature scheme (based on UOV scheme) submitted to the additional NIST PQC standardization project started in 2022. It has lower public key size compared to UOV.

Description of UOV[Patarin1999]

UOV scheme is based on a trapdoor multivariate quadratic function. The central map

$$\mathcal{F} = (f_1, \cdots, f_m) : \mathbb{F}_q^n \to \mathbb{F}_q^m$$

is designed such that each $f_1, \cdots, f_m$ is a quadratic polynomial of the form

$$f_k(x) = \sum_{i=1}^{n}\sum_{j=1}^{v} \alpha_{ij}^k x_i x_j = (\overrightarrow{x_v}^t \quad \overrightarrow{x_o}^t) \begin{pmatrix} *_{v\times v} & *_{v\times m} \\ *_{m\times v} & 0 \end{pmatrix} \begin{pmatrix} \overrightarrow{x_v} \\ \overrightarrow{x_o} \end{pmatrix}. \tag{1.1}$$

where $\alpha_{ij}^k \in \mathbb{F}_q$, $v = n - m$. $\overrightarrow{x_v}$ is called vinegar variables, $\overrightarrow{x_o}$ is called oil variables. Fix $\overrightarrow{x_v} = a$, the system $f_k(a, \overrightarrow{x_o}) = y_k$ becomes a linear system.

A linear map $\mathcal{S} : \mathbb{F}_q^n \to \mathbb{F}_q^n$ is then randomly chosen. Next, the public key map is

$$\mathcal{P} = \mathcal{F} \circ \mathcal{S} = (p_1, \cdots, p_m) : \mathbb{F}_q^n \to \mathbb{F}_q^m.$$

The secret key is $\mathcal{S}$.

A linear map $\mathcal{S} : \mathbb{F}_q^n \to \mathbb{F}_q^n$ is then randomly chosen. Next, the public key map is

$$\mathcal{P} = \mathcal{F} \circ \mathcal{S} = (p_1, \cdots, p_m) : \mathbb{F}_q^n \to \mathbb{F}_q^m.$$

The secret key is $\mathcal{S}$.
In general, it's hard to find a preimage of $\mathcal{P}$ unless we know a basis of the oil space $\mathcal{O} := \mathcal{S}^{-1}\{(0, \cdots, 0, a_{v+1}, \cdots, a_n)^t : \ a_i \in \mathbb{F}_q\}$.
Note that

$$\mathcal{P}(\mathcal{O}) = 0 \in \mathbb{F}_q^m.$$

- Advantage: short signature and short execution time.
- Disadvantage: large public key size.

	pk	sig
UOV 1	94 Kb	98 Bytes
Dilithium 2	1.28 Kb	2420 Bytes
SNOVA 1	2.25 Kb	148 Bytes

- The idea of SNOVA is replacing the coefficients ring $\mathbb{F}_q$ by matrix ring to reduce the size of public key.

Description of SNOVA scheme

- $\mathcal{R} = \mathrm{Mat}_{l\times l}(\mathbb{F}_q)$:
  matrix ring consisting by $l \times l$ matrices over $\mathbb{F}_q$, $l = 2, 3, 4$.
- $[P]$, $[F]$ denote some $n \times n$ matrices whose entries are elements of $\mathcal{R}$
- For each $Q \in \mathcal{R}$, we use

$$[\Lambda_Q] = \begin{bmatrix} Q & & \\ & \ddots & \\ & & Q \end{bmatrix}$$

  to denote the $n \times n$ block matrix whose diagonal elements are Q

Description of SNOVA scheme

- The subring $\mathbb{F}_q[s]$.
  $\mathbb{F}_q[s] = \{a_0 + a_1 s + \cdots + a_{l-1} s^{l-1} : a_0, a_1, \cdots, a_{l-1} \in \mathbb{F}_q\}$
  where s is an $l \times l$ symmetric matrix randomly chosen from $\mathcal{R}$, and the characteristic polynomial of s is irreducible.
  Note that $\mathbb{F}_q[s]$ is a subfield of $\mathcal{R}$ and $\dim_{\mathbb{F}_q} \mathbb{F}_q[s] = l$.

Central map $F : \mathcal{R}^n \longrightarrow \mathcal{R}^m$

For $i = 1, \cdots, m$,

$$F_i(x_1, \ldots, x_n) = \sum_{\alpha=1}^{l^2} A_\alpha \cdot \mathrm{x}^t \left([\Lambda_{Q_{\alpha 1}}] [F_i] [\Lambda_{Q_{\alpha 2}}] \right) \mathrm{x} \cdot B_\alpha$$

$[F_i]$ is the form of

$$[F_i] = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & 0 \end{bmatrix} \in \mathrm{Mat}_{l \times l}(\mathcal{R}),$$

with $F_{11} \in \mathrm{M}_{v \times v}(\mathcal{R})$, $F_{12} \in \mathrm{M}_{v \times m}(\mathcal{R})$ and $F_{21} \in \mathrm{M}_{m \times v}(\mathcal{R})$. A_α, B_α are chosen randomly from $\mathcal{R}$ and $Q_{\alpha 1}$, $Q_{\alpha 2}$ are chosen randomly from $\mathbb{F}_q[s] \setminus \{0\}$.

Central map and private key T

- Invertible linear map $T : \mathcal{R}^n \longrightarrow \mathcal{R}^n$, T is the map that corresponding to the matrix

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

with $T_{v\times o} \in M_{v\times m}(\mathbb{F}_q[s])$, $I_{v\times v}$ and $I_{m\times m}$ are the diagonal matrices with all entries being the identity matrix in $\mathcal{R}$.

Public key $P = F \circ T$

Let $P = F \circ T$. Set $\mathrm{x} = [T] \cdot \mathrm{u}$. For $i = 1, 2, \ldots, m$,

$$P_i(\mathrm{u}) = F_i(T(\mathrm{u})) = \sum_{\alpha=1}^{l^2} A_\alpha \cdot \mathrm{u}^t \left([\Lambda_{Q_{\alpha 1}}] [P_i] [\Lambda_{Q_{\alpha 2}}] \right) \mathrm{u} \cdot B_\alpha$$

Note that

$$[P_i] = [T]^t [F_i] [T].$$

Public key and private key of SNOVA

- Public key: The map $P : \mathcal{R}^n \longrightarrow \mathcal{R}^m$,
  i.e., the corresponding matrices $[P_i]$, $i = 1, \cdots, m$ and

$$A_\alpha, B_\alpha, Q_{\alpha 1}, Q_{\alpha 2}, \ \alpha = 1, 2, \ldots, l^2$$

- Private key: (F, T), i.e., the matrix $[T]$, the matrices $[F_i]$ for $i = 1, \ldots, m$.

Sign and Verify

Sign Choose random values $a_1, \ldots, a_v \in \mathcal{R}$ as the vinegar variables $x_1, \cdots, x_v$, the system

$$F(a_1, \ldots, a_v, x_{v+1}, \ldots, x_n) = Hash(message) \in \mathcal{R}^m$$

is a linear system with $l^2 m$ equations and $l^2 m$ variables since each x_i will provide l^2 variables for $i = v+1, \cdots, n$.
sign $= T^{-1}(\mathrm{x}) = T^{-1}(a_1, \ldots, a_v, a_{v+1}, \ldots, a_n)^t$ is the signature.

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

Advantage of SNOVA

- Small public key size and signature size.

	pk	sig
UOV 1	94 Kb	98 Bytes
Dilithium 2	1.28 Kb	2420 Bytes
SNOVA 1	2.25 Kb	148 Bytes

Structure of SNOVA

Claim 1. An SNOVA(v, m, q, l) scheme can be regarded as a UOV(lv, lm, q) scheme with $l^2 m$ equations over $\mathbb{F}_q$, rather than a UOV$(l^2 v, l^2 m, q)$ scheme over $\mathbb{F}_q$ as claimed by the authors.

- all the matrices $[F_i]$, $[T]$, and $[P_i]$ can be viewed as $ln \times ln$ matrices in $\mathrm{M}_{ln\times ln}(\mathbb{F}_q)$
- the lower-right $lm \times lm$ block is zero block for each $[F_i]$. Therefore there exists a common oil space of $[F_i]$ over $\mathbb{F}_q$ with dimension lm. Set

$$\mathcal{O}_1 = \left\{(0, \cdots, 0, a_{lv+1}, \cdots, a_{ln})^t \in \mathbb{F}_q^{ln} : \ a_i \in \mathbb{F}_q \right\}$$

  and $\mathcal{O} = [T]^{-1}(\mathcal{O}_1) \subset \mathbb{F}_q^{ln}$.
- for any $\mathrm{u}, \mathrm{v} \in \mathcal{O}$, $0 \leq j, k \leq l-1$, we have

$$\mathrm{u}^t \cdot \left([\Lambda_{s^j}][P_i][\Lambda_{s^k}]\right) \cdot \mathrm{v} = 0 \in \mathbb{F}_q \text{ for } i = 1, \cdots, m. \qquad (1.2)$$

Structure of SNOVA

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

Structure of SNOVA

Claim 2. An SNOVA(v, m, q, l) scheme can induce a standard UOV(v, m, q^l) scheme.
We can recover the oil space of SNOVA by recovering the oil space of the induced UOV(v, m, q^l) scheme.
But we are not going to use the induced UOV(v, m, q^l) scheme to give specific complexity of SNOVA(v, m, q, l) scheme since the complexity is still large.

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. Take

$$\mathcal{O}_2 := [T]^{-1}\Big(\Big\{ (0, \cdots, 0, a_{v+1}\xi^t, \cdots, a_n\xi^t)^t \in \mathbb{F}_{q^l}^{ln} : \ a_i \in \mathbb{F}_{q^l} \Big\} \Big)$$

For any $\mathrm{u}, \mathrm{v} \in \mathcal{O}_2$, we have

$$\mathrm{u}^t \cdot \Big([\Lambda_{s^j}][P_i][\Lambda_{s^k}] \Big) \cdot \mathrm{v} = \lambda^{j+k} \mathrm{u}^t \cdot [P_i] \cdot \mathrm{v}.$$

Take

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

Let $[P_i] = [P_{i,jk}]$ be the public key of SNOVA scheme with $P_{i,jk} \in \mathcal{R}$. Set

$$\widetilde{P}_{i,jk} = \xi^t \cdot P_{i,jk} \cdot \xi \in \mathbb{F}_{q^l}, \quad [\widetilde{P}_i] = (\widetilde{P}_{i,jk}) \in \mathrm{M}_{n\times n}(\mathbb{F}_{q^l}). \tag{1.3}$$

We have

$$\tilde{\mathrm{u}}^t \cdot [\widetilde{P}_i] \cdot \tilde{\mathrm{v}} = (\tilde{\mathrm{u}} \otimes \xi)^t \cdot [P_i] \cdot (\tilde{\mathrm{v}} \otimes \xi) = 0 \tag{1.4}$$

for any $\tilde{\mathrm{u}}, \tilde{\mathrm{v}} \in \mathcal{O}_2'$ and $i = 1, \cdots, m$.

In the following, we only use the claim 1 to give the new complexity of SNOVA.

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

to give a new complexity of SNOVA.

KS attack

In the SNOVA scheme, we have claimed that SNOVA (v, m, q, l) scheme over $\mathcal{R}$ can be regarded as a UOV (lv, lm, q) scheme. The complexity is

$$\text{Comp}_{\text{KS; classical}}^{\text{SNOVA}} = q^{lv-lm-1} \cdot (lm)^4$$

field multiplications.

Reconciliation Attack

The reconciliation attack can be decomposed into a series of steps:

- find an element $\mathrm{u} = (u_1, \ldots, u_{lv}, 0, \ldots, 0, 1)^t \in \mathbb{F}_q^{ln}$ s.t.

$$\mathrm{u}^t \cdot ([\Lambda_{s^j}]\,[P_i]\,[\Lambda_{s^k}]) \cdot \mathrm{u} = 0 \tag{3.1}$$

  $\rightarrow l^2 o$ equations in lv variables
- using the equations $\mathrm{u}^t \cdot ([\Lambda_{s^j}]\,[P_i]\,[\Lambda_{s^k}]) \cdot \mathrm{v} = 0$, we get $2l^2 m$ linear equations for the other elements of $\mathcal{O}$

$$\mathrm{Comp}_{\mathrm{Rec}}\mathrm{SNOVA} = \min_k q^k MQ(lv + 1 - k, l^2 m, q)$$

where $\max\{0, lv + 1 - l^2 m\} \leq k \leq lv$ is the number of fixed variables in the hybrid approach since equation (3.1) has a lot of solutions not in the space $\mathcal{O}$.

Intersection attack

Let M_1, M_2 be two invertible matrices in the set of linear combinations of $\{[\Lambda_{s^j}]\,[P_i]\,[\Lambda_{s^k}]\}_{1\leq i\leq m, 0\leq j,k\leq l-1}$

- Goal: Find $\mathrm{x} \in M_1\mathcal{O} \cap M_2\mathcal{O}$
- **In the case of $2lm > lv$.** x is a solution of

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

The attack is reduced to find a solution to the above system of $4l^2m$ quadratic equations in $2lv - lm$ variables

Intersection attack

- **In the case of $2lm \leq lv$.**
  The intersection $M_1\mathcal{O} \cap M_2\mathcal{O}$ may have no nontrivial vector.
  The probability that the intersection is non-trivial $\approx q^{-lv+2lm-1}$.
  Therefore, the attack becomes a probabilistic algorithm The complexity is

$$\text{Comp}_{\text{Int.}}\text{SNOVA} = q^{lv-2ml+1} MQ(ln, 4l^2 m, q)$$

Table 1 presents the complexity of respective attacks against the parameters submitted by SNOVA. In each pair of complexities, the left one denotes the complexity in classical gates using the analysis results in this article, and the right one denotes the complexity in classical gate given by SNOVA team, where k is the number of fixed variables in the hybrid approach.

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

SL	(v, m, 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($k=0$)

Later, the team of SNOVA submit new parameters for $l = 2$ and modify their complexity.

Thanks for listening!