MQ DOES NOT REDUCE TO TUOV

LAURA MADDISON

Abstract. The submission of the Triangular Unbalanced Oil and Vinegar (TUOV) digital signature scheme to the NIST competition in 2023 claims that if the Multivariate Quadratic (MQ) problem (with suitable parameters) is hard, then the TUOV problem must also be hard. We show why the proof fails and why the claimed theorem cannot be true in general.

1. Introduction

Triangular Unbalanced Oil and Vinegar (TUOV) is one of the multivariate-based algorithms submitted to Round 1 of the NIST call for additional digital signature schemes in 2023 [1]. It is a modification of the well-studied Unbalanced Oil and Vinegar (UOV) digital signature scheme. Much of the cryptanalysis conducted by the authors in [1] consists of demonstrating the resilience of the TUOV scheme against standard attacks for UOV, such as Kipnis-Shamir and MinRank attacks. However, they make the novel claim that the hardness of the MQ problem implies the hardness of the TUOV problem.

In this note, we discuss the error in their proof of this claim and show why it cannot be true in general.

Throughout this paper, q denotes a prime power and $\mathbb{F}_q$ denotes the field with q elements. We also let n and m denote positive integers.

2. Background

We first define the general Multivariate Quadratic (MQ) Problem, which is widely used as the basis for many proposed post-quantum digital signature schemes.

Definition 2.1. An (n, q)-***MQ-polynomial*** $f \in \mathbb{F}_q[x_1, \cdots, x_n]$ is a quadratic polynomial in n variables

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{i,j} x_i x_j + \sum_{i=1}^{n} \beta_i x_i + \gamma.$$

It has unique matrix representation as $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x} + \gamma$ where $\mathbf{A} \in \mathbb{F}_q^{n \times n}$ is upper triangular, $\mathbf{b} \in \mathbb{F}_q^n$, and $\gamma \in \mathbb{F}_q$. An (n, m, q)-***MQ-map*** $\mathcal{M} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ is an m-tuple of (n, q)-MQ-polynomials $(f^{(1)}, \ldots, f^{(m)})$. We call $\mathcal{M}$ ***random*** if the coefficients of all $f^{(i)}$ are chosen uniformly at random from $\mathbb{F}_q$.

Fix a function Setup'$(\cdot)$ that outputs (n, m, q) on input 1^κ.

Date: February 8, 2024.

Definition 2.2. The ***MQ Problem*** in relation to Setup'$(\cdot)$ is (t,ϵ)-***hard*** if there exists no algorithm that, given security parameter κ, params $= (n,m,q) \leftarrow \text{Setup}'(1^\kappa)$ and a random params-MQ-map $\mathcal{M}:\mathbb{F}_q^n \to \mathbb{F}_q^m$ in relation to params on input $\mathbf{y} := \mathcal{M}(\mathbf{w})$ with $\mathbf{w} \xleftarrow{\$} \mathbb{F}_q^n$, outputs $\mathbf{w}'$ such that $\mathcal{M}(\mathbf{w}') = \mathbf{y}$ with probability no less that $\epsilon(\kappa)$ in processing time $t(\kappa)$.

We now define Oil and Vinegar (OV) polynomials, which have become a well-known tool in multivariate cryptography since their introduction by Patarin in 1997 [2]. We follow this with the definition of TOV polynomials, which are a slight generalization from [1]. These two types of polynomials form the basis of the TUOV digital signature scheme.

Definition 2.3. Given $1 \le m < n$, an (n,m)-***OV-polynomial*** f over $\mathbb{F}_q$ is an (n,q)-MQ-polynomial of the restricted form

$$f(x_1,\dots,x_n) = \sum_{i=1}^{n-m}\sum_{j=1}^{n}\alpha_{i,j}x_ix_j + \sum_{i=1}^{n}\beta_i x_i + \gamma.$$

It thus has unique matrix representation of the form

$$f(\mathbf{x}) = \mathbf{x}^\top \begin{bmatrix} \mathbf{A}^{(1)} & \mathbf{A}^{(2)} \\ \mathbf{0}_{m\times(n-m)} & \mathbf{0}_{m\times m}\end{bmatrix}\mathbf{x} + \mathbf{b}^\top\mathbf{x} + \gamma,$$

where $\mathbf{A}^{(1)} \in \mathbb{F}_q^{(n-m)\times(n-m)}$ is upper-triangular, $\mathbf{A}^{(2)} \in \mathbb{F}_q^{(n-m)\times m}$, $\mathbf{b} \in \mathbb{F}_q^n$ and $\gamma \in \mathbb{F}_q$.

A slightly larger set of MQ-polynomials was defined in [1] as follows.

Definition 2.4. Given $1 \le d < m < n$, an (n,m,d)-***TOV-polynomial*** f over $\mathbb{F}_q$ is an (n,q)-MQ-polynomial of the form

$$\sum_{i=n-m+1}^{n-m+d}\sum_{j=n-m+1}^{n-m+d}\alpha_{i,j}x_ix_j + \sum_{i=1}^{n-m}\sum_{j=1}^{n}\alpha_{i,j}x_ix_j + \sum_{i=1}^{n}\beta_i x_i + \gamma.$$

It has a unique matrix representation as

$$f(\mathbf{x}) = \mathbf{x}^\top \begin{bmatrix} \mathbf{A}^{(1)} & \mathbf{A}^{(2d1)} & \mathbf{A}^{(2d2)} \\ \mathbf{0}_{d\times(n-m)} & \mathbf{A}^{(4d1)} & \mathbf{0}_{d\times(m-d)} \\ \mathbf{0}_{(m-d)\times(n-m)} & \mathbf{0}_{(m-d)\times d} & \mathbf{0}_{(m-d)\times(m-d)}\end{bmatrix}\mathbf{x} + \mathbf{b}^\top\mathbf{x} + \gamma$$

where $\mathbf{A}^{(1)} \in \mathbb{F}_q^{(n-m)\times(n-m)}$ and $\mathbf{A}^{(4d1)} \in \mathbb{F}_q^{d\times d}$ are upper-triangular, $\mathbf{A}^{(2d1)} \in \mathbb{F}_q^{(n-m)\times d}$, $\mathbf{A}^{(2d2)} \in \mathbb{F}_q^{(n-m)\times(m-d)}$, $\mathbf{b} \in \mathbb{F}_q^n$ and $\gamma \in \mathbb{F}_q$.

In essence, TOV-polynomials are a slight generalization of OV-polynomials, in which more non-zero entries are permitted in the matrices representing the quadratic part. The MQ-maps used in [1] are instead TUOV-maps, as follows.

Definition 2.5. Given $1 \le m_1 < m_2 < m < n$, a ***TUOV central map*** in relation to params $= (n,m,m_1,m_2,q)$ is a function $\mathcal{F}:\mathbb{F}_q^n \to \mathbb{F}_q^m$ sending $\mathbf{x} \mapsto \mathcal{F}(\mathbf{x}) = (f_1(\mathbf{x}),\dots,f_m(\mathbf{x}))^\top$, where

$$f_k \text{ is an } \begin{cases} (n,m_1)-\text{OV-polynomial}, & \text{if } k \in [m_1] \\ (n,m,k-m_1)-\text{TOV-polynomial}, & \text{if } k \in [m_2]\setminus[m_1] \\ (n,m-m_2+m_1-1)-\text{OV-polynomial}, & \text{if } k \in [m]\setminus[m_2]. \end{cases}$$

A ***TUOV map*** in relation to params $= (n, m, m_1, m_2, q)$ is $\mathcal{P} = \mathcal{S} \circ \mathcal{F} \circ \mathcal{T} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ where $\mathcal{S} : \mathbb{F}_q^m \to \mathbb{F}_q^m$ and $\mathcal{T} : \mathbb{F}_q^n \to \mathbb{F}_q^n$ are invertible affine transformations and $\mathcal{F} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ is a TUOV central map in relation to params. We call $\mathcal{P}$ ***random*** if the coefficients of the polynomials f_i in $\mathcal{F}$ as well as the affine transformations $\mathcal{S}$ and $\mathcal{T}$ are chosen uniformly at random.

Fix Setup$(\cdot)$ that outputs params $= (n, m, m_1, m_2, q)$ on input 1^κ.

Definition 2.6. The ***TUOV problem*** in relation to Setup$(\cdot)$ is (t, ϵ)-***hard*** if there exists no algorithm that, given params and a random TUOV map $\mathcal{P} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ in relation to params, on input $\mathbf{z} = \mathcal{P}(\mathbf{w})$ with $\mathbf{w} \xleftarrow{\$} \mathbb{F}_q^n$, outputs $\mathbf{w}'$ such that $\mathcal{P}(\mathbf{w}') = \mathbf{z}$ with probability no less than $\epsilon(\kappa)$ in processing time $t(\kappa)$.

The security of the TUOV signature scheme is conjecturally based on the hardness of the TUOV problem.

3. Invalidity of the Security Reduction

The authors of [1] claim that, under certain parameters, the MQ problem reduces to the TUOV problem. Their precise statement is as follows.

Assertion 3.1. [1, Theorem 1] *Given Setup$(\cdot)$ that ouputs*

$$\text{params} = (n = \frac{1}{2}m^2, m, m_1 = \frac{1}{2}m, m_2 = \frac{3}{4}m, q)$$

on input 1^κ and its restriction Setup' that outputs

$$\text{params'} = (n = \frac{1}{2}m^2, m, q),$$

if the MQ problem in relation to Setup'$(\cdot)$ is (t, ϵ)-hard, then the TUOV problem in relation to Setup$(\cdot)$ is (t, ϵ)-hard.

We disprove Assertion 3.1 here by demonstrating where the proof in [1] fails, and in the next section show why such a claim cannot be true.

To prove their claim, they must show how an arbitrary $(n = \frac{1}{2}m^2, m, q)$-MQ map $\mathcal{M}$ can be efficiently transformed into a $(n = \frac{1}{2}m^2, m, m_1 = \frac{1}{2}m, m_2 = \frac{3}{4}m, q)$-TUOV central map. More precisely, they must find an invertible affine transformation $\mathcal{Q} : \mathbb{F}_q^n \to \mathbb{F}_q^n$ such that $\mathcal{F} := \mathcal{M} \circ \mathcal{Q}$ is a TUOV central map. Since the only structural restrictions on a TUOV central map come from the quadratic terms, it suffices to specify $\mathcal{Q}$ based on its transformation on only the quadratic part of polynomials in $\mathcal{M}$, as it can be applied on the linear terms without affecting the specifications of a TUOV central map.

They consider the matrix representation of the quadratic part of the k-th polynomial f_k in $\mathcal{M}$,

$$\mathbf{M}_k = \begin{bmatrix} \mathbf{M}_k^{(1)} & \mathbf{M}_k^{(2)} \\ \mathbf{0}_{m\times(n-m)} & \mathbf{M}_k^{(4)} \end{bmatrix},$$

where $\mathbf{M}_k^{(1)} \in \mathbb{F}_q^{(n-m)\times(n-m)}$ and $\mathbf{M}_k^{(4)} \in \mathbb{F}_q^{m\times m}$ are upper triangular and $\mathbf{M}_k^{(2)} \in \mathbb{F}_q^{(n-m)\times m}$.

Then, they consider the matrix representation of an arbitrary invertible affine transformation $\mathcal{Q} : \mathbb{F}_q^n \to \mathbb{F}_q^n$ of the form

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}^{(1)} & \mathbf{Q}^{(2)} \\ \mathbf{0}_{m\times(n-m)} & \mathbf{I}_m \end{bmatrix}.$$

Then the quadratic part of the k-th polynomial of $\mathcal{M} \circ \mathcal{Q}$ has matrix representation

$$\mathbf{Q}^\top \mathbf{M}_k \mathbf{Q} = \begin{bmatrix} \mathbf{Q}^{(1)\top}\mathbf{M}_k^{(1)}\mathbf{Q}^{(1)} & \mathbf{Q}^{(1\top}\mathbf{M}_k^{(1)}\mathbf{Q}^{(2)} + \mathbf{Q}^{(1)\top}\mathbf{M}_k^{(2)} \\ \mathbf{Q}^{(2)\top}\mathbf{M}_k^{(1)}\mathbf{Q}^{(1)} & \mathbf{Q}^{(2)\top}\mathbf{M}_k^{(1)}\mathbf{Q}^{(2)} + \mathbf{Q}^{(2)\top}\mathbf{M}_k^{(2)} + \mathbf{M}_k^{(4)} \end{bmatrix}.$$

In [1], the authors had the incorrect expression $\mathbf{Q}^{(2)\top}\mathbf{M}_k^{(1)}\mathbf{Q}^{(1)} + \mathbf{Q}^{(2)\top}\mathbf{M}_k^{(2)} + \mathbf{M}_k^{(4)}$ in the bottom right block instead. We will now evaluate the impact of this error on the remainder of their proof.

Since they want this to represent the k-th polynomial in a TUOV central map, $\mathbf{A}_k = \mathbf{Q}^\top \mathbf{M}_k \mathbf{Q}$ should have the form

$$\begin{cases} \begin{bmatrix} \mathbf{A}_k^{(1)} & \mathbf{A}_k^{(2)} \\ \mathbf{0}_{m_1\times(n-m_1)} & \mathbf{0}_{m_1\times m_1} \end{bmatrix}, & k \in [m_1] \\ \begin{bmatrix} \mathbf{A}_k^{(1)} & \mathbf{A}_k^{(2d1)} & \mathbf{A}_k^{(2d2)} \\ \mathbf{0}_{d\times(n-m)} & \mathbf{A}_k^{(4d1)} & \mathbf{0}_{d\times(m-d)} \\ \mathbf{0}_{(m-d)\times(n-m)} & \mathbf{0}_{(m-d)\times d} & \mathbf{0}_{(m-d)\times(m-d)} \end{bmatrix}, & k \in [m_2] \setminus [m_1] \text{ and } d = k - m_1 \\ \begin{bmatrix} \mathbf{A}_k^{(1)} & \mathbf{A}_k^{(2)} \\ \mathbf{0}_{l\times(n-l)} & \mathbf{0}_{l\times l} \end{bmatrix}, & k \in [m] \setminus [m_2] \text{ and } l = m - m_2 + m_1 - 1. \end{cases}$$

In the TUOV specification, the error in the bottom right block of $\mathbf{Q}^\top \mathbf{M}_k \mathbf{Q}$ led to the belief that if the entries of $\mathbf{Q}^{(1)}$ were fixed, then the system to solve would become linear in the entries of $\mathbf{Q}^{(2)}$ [1]. With our correction, we see that even after fixing the entries of $\mathbf{Q}^{(1)}$, the resulting system that needs to be solved is quadratic in the entries of $\mathbf{Q}^{(2)}$.

After fixing the entries of $\mathbf{Q}^{(1)}$ as in [1], the resulting quadratic system in the entries of the $(n-m)\times m$ matrix $\mathbf{Q}^{(2)}$ has $(n-m)m \approx \frac{1}{2}m^3$ variables and approximately $\frac{11}{24}m^3$ equations. The number of equations is determined by counting the number of equations arising from the matrix representation of the quadratic part of each polynomial f_k depending on its desired form after the transformation, and summing for $k \in [m]$. The full details of this can be found in [1]. Thus, we have taken a $(\frac{1}{2}m^2, m, q)$-MQ problem, and transformed it into a presumably larger $(\frac{1}{2}m^3, \frac{11}{24}m^3, q)$-MQ problem, which we cannot solve efficiently for large m.

4. Conclusions

We conclude by demonstrating why such a transformation from MQ to TUOV cannot exist in general, and discussing the security of the TUOV digital signature scheme beyond the invalid reduction.

Lemma 4.1. *An invertible affine transformation $\mathcal{Q} : \mathbb{F}_q^n \to \mathbb{F}_q^n$ such that $\mathcal{M} \circ \mathcal{Q}$ is a TUOV central map does not exist for a general (n, m, q)-MQ map $\mathcal{M}$.*

Proof. Observe that if $\mathcal{Q}$ is invertible and $\mathbf{Q}$ is its matrix representation of its linear part, then $\det(\mathbf{Q}) \neq 0$. Let $\mathbf{M}_k$ be the matrix representation of the quadratic part of of the k-th polynomial f_k

of the MQ map $\mathcal{M}$ and $\mathbf{A}_k$ the matrix representation of the quadratic part of the k-th polynomial of an arbitrary TUOV central map. Suppose $\mathcal{M} \circ \mathcal{Q}$ is a TUOV central map. Then for all $k \in [m]$,

$$\det(\mathbf{A}_k) = \det(\mathbf{Q}^\top \mathbf{M}_k \mathbf{Q}) = (\det(\mathbf{Q}))^2 \det(\mathbf{M}_k),$$

However, $\det(\mathbf{A}_k)$=0 for all k since all matrices in a TUOV central map have a row of zeroes. This implies that $\det(\mathbf{M}_k)$ is equal to 0 for all $k \in [m]$, which is not true in general. □

Thus, the security reduction of the TUOV submission to the NIST competition is not valid. However, this does not necessarily imply that the scheme is insecure. The security analysis of UOV-based digital signature schemes is usually comprised of analysis of several known attacks, such as the Kipnis-Shamir and MinRank attacks, applied to the proposed scheme. The authors of [1] provide a robust security analysis based on these known attacks and choose their parameters accordingly, which seems to imply that TUOV is as secure as standard UOV.

References

[1] Jintai Ding, Boru Gong, Hao Guo, Xiaoou He, Yi Jin, Yuansheng Pan, Dieter Schmidt, Chengdong Tao, Danli Xie, Bo-Yin Yang, and Ziyu Zhao. TUOV: Triangular Unbalanced Oil and Vinegar, 2023. https://csrc.nist.gov/projects/pqc-dig-sig/round-1-additional-signatures.

[2] Jacques Patarin. The Oil and Vinegar signature scheme, 1997. Presented at the Dagstuhl Workshop on Cryptography.

Email address: lmadd036@uottawa.ca

Department of Mathematics and Statistics, University of Ottawa, STEM Complex, 150 Louis-Pasteur Pvt, Ottawa, Ontario, Canada K1N 6N5