

# TUOV

## Triangular Unbalanced Oil and Vinegar

Boru Gong

[tuovsig@gmail.com](mailto:tuovsig@gmail.com), [Jintai.Ding@gmail.com](mailto:Jintai.Ding@gmail.com)

September 4, 2023

# Outline

## 1 Preliminaries

- Multivariate Signature Scheme
- MQ Problem

## 2 UOV Scheme

- Description
- Attacks on UOV
- Parameters and Performances

## 3 TUOV Scheme

- Design Rationale
- Description
- Security Proof for TUOV
- Attacks on TUOV
- Parameters and Performances

## 4 Summary

# Multivariate Signature Scheme

- **Public Key:**

$$\mathcal{P}(x_1, \dots, x_n) = (p_1(x_1, \dots, x_n), \dots, p_m(x_1, \dots, x_n)),$$

where each  $p_i$  is a multivariate polynomial over  $\mathbb{F}_q$ .

- **Private Key:** a way to compute  $\mathcal{P}^{-1}$ .
- **Signing a hash of a document:**

$$(x_1, \dots, x_n) \in \mathcal{P}^{-1}(y_1, \dots, y_m).$$

- **Verifying:**

$$(y_1, \dots, y_m) \stackrel{?}{=} \mathcal{P}(x_1, \dots, x_n).$$

# Multivariate Signature Schemes

- The public key  $\mathcal{P}(x_1, \dots, x_n)$  should be *almost surjective*;
  - $n \geq m$  is necessary.
- The signing and verification should be *efficient*;
- Key sizes should be as *small* as possible.

# Theoretical Foundation

Direct attack is to solve the set of equations:

$$\mathcal{P}(x_1, \dots, x_n) = (y'_1, \dots, y'_m).$$

# Quadratic Constructions

- **Efficiency considerations** lead to mainly *quadratic* constructions.

$$p_\ell(x_1, \dots, x_n) = \sum_{i,j} \alpha_{\ell ij} x_i x_j + \sum_i \beta_{\ell i} x_i + \gamma_\ell.$$

- **Mathematical structure consideration:** any set of *higher* degree polynomial equations can be reduced to a set of quadratic equations.
  - For instance,  $x_1 x_2 x_3 = 5$  is equivalent to

$$\begin{aligned} x_1 x_2 - y &= 0 \\ y x_3 &= 5. \end{aligned}$$

# What is MQ Problem

## MQ Problem

**Given:** a multivariate quadratic equation system with  $m$  equations and  $n$  variables over  $\mathbb{F}_q$ .

**Find:** a solution over  $\mathbb{F}_q$ .

## Hardness of MQ Problem

### Theorem

- 1 MQ problem is NP-complete. [Garey-Johnson 1979]
- 2 MQ problem can be solved in polynomial time,  
when:  $m \leq n^2/2$  [Shamir 1999],  
when:  $n \geq m^2/2$  and  $\text{char}(\mathbb{F}_q) = 2$  [Miura-Hashimoto-Takagi 2013].

Moreover, it is *believed*  $\exists 0 < \varepsilon < 1/2$  s.t. MQ problem is hard when  $n = \alpha m^2$  and  $\varepsilon < \alpha < 1/2$ .



## (Unbalanced) Oil and Vinegar Scheme

- The original *balanced* version was introduced by Jacques Patarin in 1997.
  - Inspired by **linearization attack** to Matsumoto-Imai cryptosystem.
  - $n = 2m$ .
- Kipnis and Shamir proposed an attack which breaks this balanced OV scheme in 1998.
- Kipnis, Patarin and Goubin proposed a modified scheme called *Unbalanced Oil and Vinegar (UOV)* signature scheme in 1999.
  - $n > 2m$ .

## (Unbalanced) Oil and Vinegar Scheme

- Let  $\mathbb{F} = \mathbb{F}_q$  be a finite field with  $q$  elements;
- $v$ : the number of *vinegar* variables;  
 $m = o$ : the number of *oil* variables;  
 $n := v + m$ : the number of variables.
- $V = \{1, \dots, v\}$ ,  $O = \{v + 1, \dots, n\}$ .  
 We denote the variables  $x_i$  ( $i \in V$ ) as **Vinegar variables**,  
 $x_{v+1}, \dots, x_n$  as **Oil variables**.

# (Unbalanced) Oil and Vinegar Scheme

## OV-polynomial

An  $(n, m)$ -**OV-polynomial**  $f$  over  $\mathbb{F}$  is defined as

$$\sum_{i=1}^{n-m} \sum_{j=1}^n \alpha_{i,j} \cdot x_i x_j + \sum_{i=1}^n \beta_i \cdot x_i + \gamma$$

The homogeneous quadratic part of  $f$  can be uniquely represented in an upper-triangular quadratic form:

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

Note that there is **no "Oil  $\times$  Oil"** part in an OV-polynomial.

# Key Generation

## UOV Central Map

$\mathcal{F} = (f_1, \dots, f_m)$ , where each  $f_k$  is an

$(n, m)$ -OV-polynomial,  $k = 1, \dots, m$ .

- **Private Key** is  $(\mathcal{T}, \mathcal{F})$ , where
  - $\mathcal{F} : \mathbb{F}^n \rightarrow \mathbb{F}^m$  is a UOV central map.
    - $\mathcal{F} = (f_1, \dots, f_m)$ .
  - $\mathcal{T} : \mathbb{F}^n \rightarrow \mathbb{F}^n$  is an affine map;
- **Public Key** is  $\mathcal{P} = \mathcal{F} \circ \mathcal{T}$ .

# Signature Generation

$\mathbf{z} \leftarrow \text{Sign}(\mu)$ :

- 1 Use a hash function  $\mathcal{H} : \{0, 1\}^* \rightarrow \mathbb{F}^m$  to compute a digest  $\mathbf{w} = \mathcal{H}(\mu)$ .
- 2 Find a pre-image  $\mathbf{y} \in \mathbb{F}^n$  of  $\mathbf{w}$  under the central map  $\mathcal{F}$ .
- 3 Compute the signature  $\mathbf{z} \in \mathbb{F}^n$  by  $\mathbf{z} = \mathcal{T}^{-1}(\mathbf{y})$ .

# How to find $\mathcal{F}^{-1}$

Inversion of  $\mathcal{F}^{-1}$  is efficient indeed:

- 1 Fix values for vinegar variables  $x'_1, \dots, x'_v$ .
- 2  $f_k = \sum \alpha_{i,j}^{(k)} x_i x'_j + \sum \alpha_{i,j}^{(k)} x'_i x'_j + \sum \beta_i^{(k)} x_i + \sum \beta_i^{(k)} x'_i + \gamma^{(k)}$
- 3  $\mathcal{F}(x'_1, \dots, x'_v, \dots) : \mathbb{F}_q^m \rightarrow \mathbb{F}_q^m$  is a linear system in oil variables  $x_{v+1}, \dots, x_n$ .

# Signature Verification

Signature verification is fast:

- Check whether  $\mathcal{H}(d) = \mathcal{P}(\mathbf{z})$ . If so, then the signature  $\mathbf{z}$  is accepted, otherwise rejected.

# Attacks on UOV

- Kipnis-Shamir attack (*i.e.*, UOV attack)
- Reconciliation attack
- Collision attack
- Direct attack
- Intersection attack
- MinRank attack
- Quantum attack



# UOV: Parameter Sets

	NIST S.L.	$n$	$m$	$q$	$ \text{epk} $ (bytes)	$ \text{esk} $ (bytes)	$ \text{cpk} $ (bytes)	$ \text{csk} $ (bytes)	signature (bytes)
uov-1p	1	112	44	256	278 432	237 896	43 576	48	128
uov-1s	1	160	64	16	412 160	348 704	66 576	48	96
uov-III	3	184	72	256	1 225 440	1 044 320	189 232	48	200
uov-V	5	244	96	256	2 869 440	2 436 704	446 992	48	260

**Figure:** Recommended parameter sets and the corresponding key/signature sizes for UOV.

## UOV: Performances

	Haswell			Skylake		
	KeyGen	Sign	Verify	KeyGen	Sign	Verify
uov-1p-classic	3 311 188	116 624	82 668	2 903 434	105 324	90 336
uov-1p-pkc	3 393 872		311 720	2 858 724		224 006
uov-1p-pkc+skc	3 287 336	2 251 440		2 848 774	1 876 442	
uov-1s-classic	4 945 376	123 376	60 832	4 332 050	109 314	58 274
uov-1s-pkc	5 002 756		398 596	4 376 338		276 520
uov-1s-pkc+skc	5 448 272	3 042 756		4 450 838	2 473 254	
uov-III-classic	22 046 680	346 424	275 216	17 603 360	299 316	241 588
uov-III-pkc	22 389 144		1 280 160	17 534 058		917 402
uov-III-pkc+skc	21 779 704	11 381 092		17 157 802	9 965 110	
uov-V-classic	58 162 124	690 752	514 100	48 480 444	591 812	470 886
uov-V-pkc	57 315 504		2 842 416	46 656 796		2 032 992
uov-V-pkc+skc	57 306 980	26 021 784		45 492 216	22 992 816	
Dilithium 2 <sup>†</sup> [28]	97 621*	281 078*	108 711*	70 548	194 892	72 633
Falcon-512 [44]	19 189 801*	792 360*	103 281*	26 604 000	948 132	81 036
SPHINCS+ <sup>‡</sup> [25]	1 334 220	33 651 546	2 150 290	1 510 712*	50 084 397*	2 254 495*

<sup>†</sup> Security level II. <sup>‡</sup> Sphincs+-SHA2-128f-simple. \* Data from SUPERCOP [20].

Figure: Benchmarking results of AVX2 implementations. Numbers are the median CPU cycles of 1000 executions each.

## Why we do TUOV?

- The hardness of the UOV scheme relies on the **UOV assumption**, *i.e.*, it is hard to find a pre-image of  $\mathcal{P}$ .
- On the one hand, it is known that
  - find a pre-image of  $\mathcal{P} \leq$  solve MQ problem.
- On the other hand, it is not known whether
  - find a pre-image of  $\mathcal{P} \stackrel{?}{=} \text{solve MQ problem.}$
- a large part of coefficients in  $\mathcal{F}$  are zeroes, which makes  $\mathcal{P} = \mathcal{F} \circ \mathcal{T}$  failed to be proved as random as MQ map.

# Why TUOV?

Our solution:

- add some nonzero parts in  $\mathcal{F}$ ; and
- keep  $\mathcal{F}$  efficiently invertible.

Our result: the *Triangular* Unbalanced Oil and Vinegar (TUOV) scheme.

# Triangular Map

## Definition

The **T**riangular in the name TUOV refers to a triangular map (or, *de Jonqui ere* map) as

$$\mathcal{J} : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^n,$$
$$\mathbf{x} \mapsto (x_1, x_2 + g_2(x_1), \dots, x_n + g_n(x_1, \dots, x_{n-1}))$$

where  $g_i$  is a polynomial over  $\mathbb{F}_q$ .

The triangular map  $\mathcal{J}$  is efficiently invertible.

# TUOV notations

- parameters: let  $m, v, o_1$  be integers and the number of variables is given by  $n = v + o_1 + (m - o_1)$ .
- index sets:  
 $V = \{1, \dots, v\},$   
 $O_1 = \{v + 1, \dots, v + o_1\},$   
 $O_2 = \{v + o_1 + 1, \dots, n\}.$
- variables: denote  $x_i$  ( $i \in V$ ) as Vinegar variables, and  $x_{v+1}, \dots, x_n$  ( $O_1 \cup O_2$ ) as **Oil variables**.

# TUOV-polynomial

## Definition

For  $d \geq 1$ , an  $(n, m, d)$ -**TUOV-polynomial**  $f$  over  $\mathbb{F} = \mathbb{F}_q$  is defined as

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

The homogeneous quadratic part of  $f$ :

$$\begin{bmatrix} \mathbf{A}^{(1)} & \mathbf{A}^{(2)} & \mathbf{A}^{(3)} \\ \mathbf{0} & \mathbf{A}^{(5)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

$(n, m, d)$ -TUOV-polynomial adds an " $O_1 \times O_1$  ( $o_1 = d$ )" part.

# TUOV Central Map

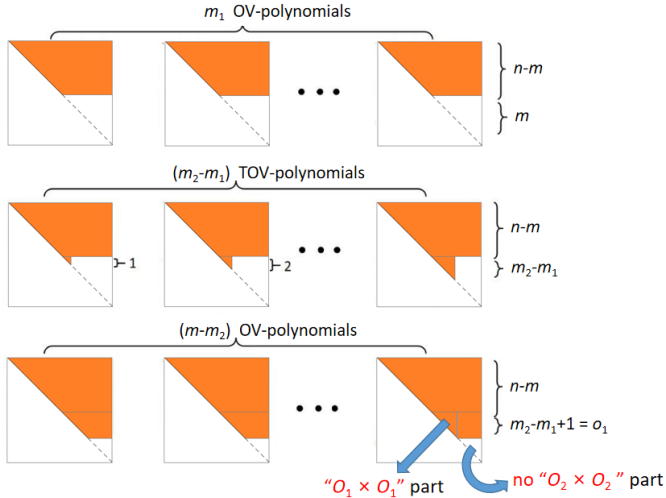
## Definition

A *TUOV central map* with parameters  $(n, m, m_1, m_2, q)$  is  $\mathcal{F} = (f_1, \dots, f_m)$ , where  $f_k$  is

$$\begin{cases} (n, m)\text{-OV-polynomial,} & k = 1, \dots, m_1 \\ (n, m, k - m_1)\text{-TOV-polynomial,} & k = m_1 + 1, \dots, m_2 \\ (n, m - m_2 + m_1 - 1)\text{-OV-polynomial,} & k = m_2 + 1, \dots, m. \end{cases}$$



# TUOV Central Map



# Key Generation

- **Private Key:**  $(S, \mathcal{F}, \mathcal{T})$ , where
  - $S : \mathbb{F}^m \rightarrow \mathbb{F}^m$ ,  $\mathcal{T} : \mathbb{F}^n \rightarrow \mathbb{F}^n$  are affine maps;
  - $\mathcal{F} : \mathbb{F}^n \rightarrow \mathbb{F}^m$  is a TUOV central map.
- **Public Key:**  $\mathcal{P} = S \circ \mathcal{F} \circ \mathcal{T}$ .

# Signature Generation

$\mathbf{z} \leftarrow \text{Sign}(\mu)$ :

- 1 Use a hash function  $\mathcal{H} : \{0, 1\}^* \rightarrow \mathbb{F}^m$  to compute a digest  $\mathbf{w} = \mathcal{H}(\mu)$ .
- 2 Compute  $\mathbf{x} = \mathcal{S}^{-1}(\mathbf{w}) \in \mathbb{F}^m$ .
- 3 Find a pre-image  $\mathbf{y} \in \mathbb{F}^n$  of  $\mathbf{x}$  under the central map  $\mathcal{F}$ .
- 4 Compute the signature  $\mathbf{z} \in \mathbb{F}^n$  by  $\mathbf{z} = \mathcal{T}^{-1}(\mathbf{y})$ .

## How to find the pre-image under $\mathcal{F}$ (1/2)

**Step 1:** Use one  $(n, m)$ -OV-polynomial and all TOV-polynomials in  $\mathcal{F}$ :

$$g(x_{v+1}) + \sum_{j=v+1}^n L_{1,j}(x_1, \dots, x_v)x_j + Q_1(x_1, \dots, x_v)$$

$$g(x_{v+1}) + \sum_{j=v+1}^n L_{2,j}(x_1, \dots, x_v)x_j + Q_2(x_1, \dots, x_v)$$

$$\dots$$

$$g(x_{v+1}, \dots, x_{v+o_1-1}) + \sum_{j=v+1}^n L_{o_1,j}(x_1, \dots, x_v)x_j + Q_{o_1}(x_1, \dots, x_v)$$

Solve linear equations of  $x_i$  ( $i \in V$ ), so that the **red parts** = 0 and **blue parts** = 1. Substitute the solution, then we can solve  $x_i$  ( $i \in O_1$ ) efficiently – by triangular map.

$$\begin{array}{cccccc}
 L_{1,v+1} & L_{1,v+2} & \cdots & L_{1,v+o_1} & \cdots & L_{1,n} \\
 L_{2,v+1} & L_{2,v+2} & \cdots & L_{2,v+o_1} & \cdots & L_{2,n} \\
 \vdots & \vdots & \ddots & \vdots & & \vdots \\
 L_{o_1,v+1} & L_{o_1,v+2} & \cdots & L_{o_1,v+o_1} & \cdots & L_{o_1,n}
 \end{array}$$

## How to find the pre-image under $\mathcal{F}$ (2/2)

**Step 2:** Use remaining OV-polynomials in  $\mathcal{F}$ :

Substitute  $x_i$  ( $i \in V \cup O_1$ ), we can get  $x_i$  ( $i \in O_2$ ) by solving linear system.

### Note

Note that a  $(n, m, d)$ -TOV-polynomial adds a  $(n, m)$ -polynomial is still a  $(n, m, d)$ -TOV-polynomial, thus we can randomly add  $(n, m)$ -polynomials to  $(n, m, d)$ -TOV-polynomial to get random  $x_i$  ( $i \in V$ ) at the beginning.

# Signature Verification

- Check whether  $\mathcal{H}(d) = \mathcal{P}(\mathbf{z})$ . If so, then the signature  $\mathbf{z}$  is accepted, otherwise rejected.

# Security Proof

## Definition: Hardness of MQ problem

The *MQ problem* parameterized by  $(n, m, q)$  is called  $(t, \varepsilon)$ -hard, if there exists no algorithm that, given a random MQ-map

$\mathcal{M} : \mathbb{F}^n \rightarrow \mathbb{F}^m$ , on input  $\mathbf{y} := \mathcal{M}(\mathbf{w})$  with  $\mathbf{w} \xleftarrow{\$} \mathbb{F}^n$ , outputs  $\mathbf{w}'$  such that  $\mathcal{M}(\mathbf{w}') = \mathbf{y}$  with probability no less than  $\varepsilon$  in processing time  $t$ .

# Security Proof

## TUOV map

*TUOV map* is  $\mathcal{P} = \mathcal{S} \circ \mathcal{F} \circ \mathcal{T} : \mathbb{F}^n \rightarrow \mathbb{F}^m$  where  $\mathcal{S} : \mathbb{F}^m \rightarrow \mathbb{F}^m$  and  $\mathcal{T} : \mathbb{F}^n \rightarrow \mathbb{F}^n$  are invertible affine transformations and  $\mathcal{F} : \mathbb{F}^n \rightarrow \mathbb{F}^m$  is a TUOV central map.

## Definition: Hardness of TUOV problem

TUOV problem with params  $(n, m, m_1, m_2, q)$  is  $(t, \varepsilon)$ -hard if there exists no algorithm that, given a random TUOV map

$\mathcal{P} : \mathbb{F}^n \rightarrow \mathbb{F}^m$ , on input  $\mathbf{z} = \mathcal{P}(\mathbf{w})$  with  $\mathbf{w} \xleftarrow{\$} \mathbb{F}^n$ , outputs  $\mathbf{w}'$  such that  $\mathcal{P}(\mathbf{w}') = \mathbf{z}$  with probability no less than  $\varepsilon$  in processing time  $t$ .



# Security Proof

## Theorem

*If MQ problem with params  $(n = \frac{11}{24} \cdot m^2, m, q)$  is  $(t, \varepsilon)$ -hard, then TUOV problem with params  $(n = \frac{1}{2} \cdot m^2, m, m_1 = \frac{1}{2} \cdot m, m_2 = \frac{3}{4} \cdot m, q)$  is  $(t, \varepsilon)$ -hard.*

# Proof Sketch

- Our goal is to prove: it is highly probable that for a random MQ-map  $\mathcal{M}$ , there exists an invertible affine map  $\mathcal{Q}$  and a TUOV central map  $\mathcal{F}$  such that  $\mathcal{M} = \mathcal{F} \circ \mathcal{Q}$ .
- Hence, if there exists an algorithm  $\mathcal{A}$  that solves TUOV problem, we can construct an algorithm  $\mathcal{B}^{\mathcal{A}}$  that solves MQ-problem.
- Specifically, when taking MQ input  $(\mathcal{M}, \mathbf{y})$ ,  $\mathcal{B}^{\mathcal{A}}$  can randomly sample an  $\mathcal{S}$ , make query  $(\mathcal{P} = \mathcal{S} \circ \mathcal{M}, \mathbf{z} = \mathcal{S}(\mathbf{y}))$  on  $\mathcal{A}$  and output the  $\mathbf{w}'$  returned by  $\mathcal{A}$ .

# Proof Sketch

Without loss of generality, we consider quadratic polynomials only with their quadratic part.

- Set the matrix representation of  $Q^{-1}$  to be

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

- There are  $(n - m) \times m$  unknown variables.
- Recall our goal is to find  $\mathcal{F}$  and  $Q$  such that  $\mathcal{M} = \mathcal{F} \circ Q$ .
- Hence we want  $\mathcal{M} \circ Q^{-1}$  to be some TUOV central map, i.e. it satisfies some equations.

## Proof Sketch

The total number of equations to solve is

$$\begin{aligned} & \frac{m_1}{2}(m+1)m \\ & + \frac{m_2 - m_1}{6} \left( 3m(m+1) - (m_2 - m_1)^2 - 3(m_2 - m_1) - 2 \right) \\ & + \frac{m - m_2}{2} (m - m_2 + m_1 - 1)(m - m_2 + m_1). \end{aligned}$$

If we pick  $m_1 = \frac{1}{2}m$  and  $m_2 = \frac{3}{4}m$ , *i.e.*,  $m_2 - m_1 = \frac{1}{4}m$ , then we have roughly  $\frac{11}{24}m^3$  equations. As long as the number of variables is no less than equations, *i.e.*,  $n \geq \frac{11}{24} \cdot m^2$ , there exists invertible linear transformation  $Q$  such that  $\mathcal{M} \circ Q^{-1}$  is TUOV central map with high probability.

# Attacks on TUOV

To our knowledge, attacks against UOV is applicable to TUOV, and vice versa.

# TUOV: Parameter Sets

For efficient implementation, we choose  $m_1 = m_2$ .

	NIST Security Level ( $n, m, m_1, q$ )	$ upk $ (bytes)	$ usk $ (bytes)	$ cpk $ (bytes)	$ csk $ (bytes)	$ \sigma $ (bytes)
tuov-1p	1 (112, 44, 22, 256)	278 432	239 391	42 608	48	112
tuov-1s	1 (160, 64, 32, 16)	412 160	350 272	65 552	48	80
tuov-III	3 (184, 72, 36, 256)	1 225 440	1 048 279	186 640	48	184
tuov-V	5 (244, 96, 48, 256)	2 869 440	2 443 711	442 384	48	244

**Figure:** Recommended parameter sets and the corresponding key/signature sizes for TUOV.

## TUOV: Performances

Schemes	Optimized Implementations (AVX2)		
	KeyGen	Sign	Verify
tuov-Ip	10,682,834	220,792	127,722
tuov-Ip-pkc			491,120
tuov-Ip-pkc+skc	6,617,102	6,698,588	
tuov-Is	32,007,930	272,394	103,746
tuov-Is-pkc			570,194
tuov-Is-pkc+skc	15,635,380	21,534,990	
Dilithium-II	113,316	272,332	123,916
tuov-III	57,322,074	608,604	442,770
tuov-III-pkc			1,914,056
tuov-III-pkc+skc	33,336,974	33,409,538	
Dilithium-III	197,026	448,172	199,656
tuov-V	139,948,218	1,133,958	786,450
tuov-V-pkc			4,520,748
tuov-V-pkc+skc	85,778,546	74,923,822	
Dilithium-V	303,434	551,760	313,096

**Figure:** Benchmarking results of AVX2 implementations. Numbers are the median CPU cycles of 10000 executions each.

## Summary

- UOV is competitive with the new NIST standards by most measures, except for public key size.
- with the triangular structure, TUOV has slower signing speed than UOV, while it gains a reduction to hard problem (MQ-problem when  $n = \alpha m^2$  where  $\varepsilon < \alpha < 1/2$ ).



# Thanks and Any Questions?

*[tuovsig@gmail.com](mailto:tuovsig@gmail.com), [Jintai.Ding@gmail.com](mailto:Jintai.Ding@gmail.com)*