Trapdoor-Free Multivariate One-Time Signatures in the Standard Model

Zijun Zhuang[1][0009−0007−7729−5790], Yingjie Zhang[2][0000−0001−9745−0558], and Jintai Ding[3][0000−0003−1257−7598]

[1] Tsinghua University, 30 Shuangqing Rd, Haidian District, Beijing 100084, China
zhuangzj22@mails.tsinghua.edu.cn
[2] Beijing Institute of Mathematical Sciences and Applications, Yanqi Lake West Rd, Huairou District, Beijing 101408, China
zyj.crypto@outlook.com
[3] Xi'an Jiaotong-Liverpool University, 111 Ren'ai Rd, Suzhou Industrial Park, Suzhou 215123, China
jintai.ding@gmail.com

Abstract. Constructing a multivariate signature scheme that is both trapdoor-free and provably secure in the standard model has been a long-standing challenge. In this paper, we present MQ-based W-OTS$^+$, the first multivariate one-time signature scheme to achieve these properties, and prove its *existential unforgeability under chosen-message attack* via reductions to the *MQ one-wayness (MQOW) assumption*.
Our analysis focuses on the three properties that the MQ function family must satisfy within the W-OTS$^+$ framework: *one-wayness* (OW), *second-preimage resistance* (SPR), and *undetectability* (UD). For SPR, we give a tight reduction from SPR to its OW. For UD, we provide a tighter reduction from UD to its OW, reducing the security loss from the previous bound of $O(\epsilon^{-4})$ to $O(\epsilon^{-2})$ over arbitrary finite fields.
MQ-based W-OTS$^+$ consolidates the three assumptions required by W-OTS$^+$ (OW, SPR, and UD) into one standard multivariate assumption (MQOW). This shows that we can leverage some desirable properties of the MQ function family to construct a multivariate one-time signature in the standard model, instead of relying on trapdoors or zero-knowledge proofs.

Keywords: Post-Quantum Cryptography · Multivariate Cryptography · One-Time Signature Schemes · W-OTS$^+$

1 Introduction

Multivariate cryptosystems are compelling candidates for post-quantum signatures, rooted in the hardness of solving multivariate quadratic (MQ) systems over finite fields [23]. However, existing multivariate signatures typically fall into two categories [4]: they either rely on hidden algebraic trapdoors [31, 34–37], whose underlying structures have repeatedly been targeted by cryptanalysis [8, 11, 12, 15, 26, 39, 41], or are based on zero-knowledge proofs [5, 7, 13, 14, 22, 40],

with known security proofs only in the random oracle model [5,13] or the quantum random oracle model [7,14,20,22]. Constructing a multivariate signature scheme that is both *trapdoor-free* and provably secure in the *standard model* has remained a long-standing challenge.

We address this challenge in the one-time setting, focusing on a core building block of hash-based signatures (HBS): the one-time signature (OTS). Specifically, we leverage the W-OTS$^+$ [30] framework, which only relies on the forward evaluation of the underlying function family and avoids the need for trapdoors. Furthermore, while the security of the original W-OTS scheme relied on *collision resistance* (CR) [19], W-OTS$^+$ eliminates the need for this demanding property. Instead, Hülsing showed that [30] in the standard model, the *existential unforgeability under chosen-message attack* (EU-CMA) of W-OTS$^+$ reduces to three properties of the underlying function family: *one-wayness* (OW), *second-preimage resistance* (SPR), and *undetectability* (UD, which characterizes the pesudorandomness of a function family). This relaxation directly benefits MQ functions, whose inherent difficulty in achieving CR [9,17] has historically limited their use as general-purpose hash functions. By embedding MQ functions into the W-OTS$^+$ framework, we circumvent this limitation and obtain a trapdoor-free multivariate one-time signature with standard-model security (since OW of the MQ function family is widely believed to hold in multivariate cryptography [6,13], the remaining issues are SPR and UD).

This integration also takes a first step toward addressing a persistent gap within HBS. Although HBS frameworks provide reductionist security proofs, their concrete instantiations typically treat hash functions as black boxes. While standard hash functions enjoy strong empirical security with respect to basic properties (OW, SPR, and CR) after withstanding decades of cryptanalysis [27,43,44], they are not specifically designed to satisfy the specialized properties demanded by HBS proofs. This mismatch has already surfaced in concrete HBS analyses and attacks [2,38]. By instantiating W-OTS$^+$ with MQ functions, we ground the security of the resulting scheme in *an explicit algebraic computational assumption.*

Our Contributions

We use $\mathcal{Q}_{n,m,q}$ to denote the keyed function family consisting of all MQ functions $f : \mathbb{F}_q^n \to \mathbb{F}_q^m$, and the key space consists of the coefficients of all MQ functions. Let $\mathrm{InSec}^{\mathsf{P}}(\mathcal{Q}_{n,m,q}; t)$ denote the maximum success probability of any adversary running in time t against property P of $\mathcal{Q}_{n,m,q}$.

In this work, we analyze the OW, SPR, and UD of $\mathcal{Q}_{n,m,q}$, and then use $\mathcal{Q}_{n,m,q}$ as the underlying function family in the W-OTS$^+$ framework. This yields MQ-based W-OTS$^+$, which is the first multivariate one-time signature scheme that is both trapdoor-free and provably secure in the standard model. Our main contributions are as follows:

– We give a tight reduction from SPR to OW: for each t, we have

$$\mathrm{InSec}^{\mathsf{SPR}}(\mathcal{Q}_{n,m,q}; t) \leq 2q \cdot \mathrm{InSec}^{\mathsf{OW}}(\mathcal{Q}_{n,m,q}; t'),$$

where $t' = t + \mathrm{poly}(n, m, \log q)$.
- We give a tighter reduction from UD to OW for arbitrary q, improving the $O(\epsilon^{-4})$ loss in [10] to $O(\epsilon^{-2})$: for each t, we have

$$\mathrm{InSec}^{\mathsf{UD}}\left(\mathcal{Q}_{n,m,q}; t\right) \leq \inf_{0<\epsilon\leq 1} \max\left\{\epsilon,\ 8(q-1)\cdot \mathrm{InSec}^{\mathsf{OW}}\left(\mathcal{Q}_{n,m,q}; t_\epsilon\right)\right\},$$

where $t_\epsilon = \frac{72\, n^2 q^2 (q-1)^3}{\epsilon^2} t + \frac{1}{\epsilon^2}\,\mathrm{poly}\left(n, m, q, \log\frac{1}{\epsilon}\right)$.
- We prove that MQ-based W-OTS$^+$ is EU-CMA secure under the MQ one-wayness (MQOW) assumption. This shifts the security foundation from empirically justified black-box assumptions on symmetric primitives to an explicit algebraic hardness assumption.
- Compared with W-OTS$^+$ instantiated with symmetric primitives, where OW, SPR, and UD are treated as independent assumptions, MQ-based W-OTS$^+$ consolidates these three assumptions into one standard multivariate assumption (MQOW).

Taken together, our results constitute a first step toward a new direction for multivariate signatures: the MQ function family has desirable properties and can serve as a forward-evaluable primitive for standard-model one-time signatures, without relying on hidden trapdoors or zero-knowledge proofs.

Related Work

One-Wayness (OW) In the OW problem, the challenge is generated as $\mathbf{y} = f_k(\mathbf{x})$ for uniformly sampled k and $\mathbf{x}$.

In multivariate cryptography, it is widely believed [6, 13, 16] that $\mathcal{Q}_{n,m,q}$ is one-way (hence we refer to it as the MQOW assumption) when $m = \Theta(n)$. In fact, all known algorithms have exponential complexity when $m = \Theta(n)$ [4, 21, 42].

Second-Preimage Resistance (SPR) In the SPR problem, the challenge is a uniformly sampled pair $(k, \mathbf{x})$, and the adversary needs to find $\mathbf{x}' \neq \mathbf{x}$ such that $f_k(\mathbf{x}') = f_k(\mathbf{x})$.

In 2017, [3] showed that the SPR of the MQ function family $\mathcal{Q}_{n,m,q}$ can be tightly reduced to finding a *non-zero* solution of the following MQ system (we refer to it as the zero-constant MQ (ZCMQ) system):

$$\mathbf{x}^\mathsf{T}\mathbf{A}_i\mathbf{x} + \mathbf{b}_i^\mathsf{T}\mathbf{x} = 0 \qquad \text{for all } i = 1, \ldots, m,$$

where all $\mathbf{A}_i, \mathbf{b}_i$ are uniformly sampled.

Undetectability (UD) UD [19] is, in fact, another term of pseudorandomness: for $\mathcal{Q}_{n,m,q}$, UD refers to the computational indistinguishability of the output distribution from the uniform distribution, i.e.,

$$(k, f_k(\mathbf{x})) \approx_c (k, \mathbf{u}),$$

where k, $\mathbf{x}$, and $\mathbf{u}$ are sampled uniformly from their respective spaces.

In 2006, [6] showed that for $q = 2$, the UD of $\mathcal{Q}_{n,m,q}$ can be reduced to its OW, with an $O(\epsilon^{-2})$ security loss. The reduction is, in fact, a realization of the Goldreich–Levin theorem [24]. In 2008, [33] proved this decision-to-search result for a general q. However, since it encountered some difficulty (we will analyze it later) in generalizing the method in [6], it utilized the algorithm in [25], yielding an $O(\epsilon^{-5})$ loss. Recently, [10] utilized the large-field Goldreich–Levin theorem [18] and obtain an $O(\epsilon^{-4})$ loss.

Overview of Techniques

The main technical results in this paper are the tight reduction from SPR to OW, and the tighter reduction from UD to OW.

Tight Reduction from SPR to OW Our reduction is based on the reduction from SPR to ZCMQ [3]. We show that ZCMQ can, in turn, be reduced to OW. The idea is to randomly mask the linear coefficients of an MQOW instance, transform the constants of the MQOW instance into the linear coefficients of a ZCMQ instance, call the ZCMQ solver on this instance and receive a vector $\mathbf{z}$. If $\mathbf{z}$ is the non-zero solution of the ZCMQ instance, then it is a preimage of the OW instance. The reduction invokes the ZCMQ solver only once and loses a factor of at most $2q$ in success probability, yielding a tight reduction, as q is a small constant in multivariate cryptography.

Tighter Reduction from UD to OW Our reduction is a generalization of the reduction in [6]. Firstly, we recall the fact that a UD distinguisher for $\mathcal{Q}_{n,m,q}$ can be used to construct a linear form predictor [33]. Secondly, we show how to turn a linear form predictor with a fixed function-element pair $(H, H(\mathbf{x}))$ into an inverter for a general function H, with an $O(\epsilon^{-2})$ security loss for arbitrary q. Finally, we use a linear form predictor to construct an inverter for $\mathcal{Q}_{n,m,q}$ in the average case, maintaining the loss.

The second step is our main technical contribution. We show that for any fixed function $H : \mathbb{F}_q^n \to \mathbb{F}_q^m$, if there exists an algorithm $\mathcal{P}$ which, on input a linear form $R : \mathbb{F}_q^n \to \mathbb{F}_q$ and $H(\mathbf{x})$, can predict $R(\mathbf{x})$ with success probability at least $1/q+\epsilon$, then one can efficiently construct a short list of candidate preimages of $H(\mathbf{x})$ that contains $\mathbf{x}$ with constant probability.

Our algorithm is a q-ary analogue of the Goldreich–Levin reconstruction procedure for $\mathbb{F}_2$ used in [6]. The $\mathbb{F}_2$ case uses majority voting technique, and uses the fast Walsh–Hadamard transform to accelerate the vote computation. Over $\mathbb{F}_q$, however, [33] pointed out that a prediction probability larger than $1/q$ does not by itself imply that the correct symbol is the most frequent one, which is the main barrier to reuse the majority voting technique.

To solve this problem, our reconstruction algorithm introduces an extra scalar $s \in \mathbb{F}_q^*$ in the predictor queries and ranges over all nonzero choices of s. In the probability analysis, this is equivalent to averaging over a uniformly random

$s \leftarrow \mathbb{F}_q^*$. This averaging symmetrizes the nonzero prediction errors and makes all wrong symbols have the same expected vote. Then we can apply the fast q-ary Fourier transform to to accelerate the vote computation and obtain a tighter reduction.

Organization

The rest of the paper proceeds as follows. We establish the necessary preliminaries in Section 2, followed by an analysis of the cryptographic properties of the MQ function family in Section 3. Section 4 presents the MQ-based W-OTS$^+$ construction, proves its EU-CMA security, and provides a proof-of-concept implementation. Section 5 concludes the paper.

2 Preliminaries

We begin by recalling standard notation, following [32]. Let $x \xleftarrow{\$} X$ denote that x is sampled uniformly at random from a finite set X. Let $y \leftarrow \mathsf{Alg}(x)$ denote the output of the algorithm Alg on input x. We write log for $\log_2$, and $[n]$ for the set $\{1, 2, \ldots, n\}$. The symbol $\oplus$ denotes bitwise exclusive-or, and $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the floor and ceiling functions, respectively.

2.1 Security Notions

In this subsection, we revisit the formal security notions of a general keyed function family. From now on, we assume that

$$\mathcal{F}_\lambda = \{f_k : \mathcal{X}_\lambda \to \mathcal{Y}_\lambda\}_{k \in \mathcal{K}_\lambda}$$

is a keyed function family, where $\mathcal{X}_\lambda$ and $\mathcal{Y}_\lambda$ denote the input and output spaces, respectively, and $\mathcal{K}_\lambda$ denotes the key space. We assume that these spaces are finite and equipped with efficient uniform samplers.

The security of W-OTS$^+$ relies on the following properties of the underlying function family $\mathcal{F}_\lambda$: one-wayness, second-preimage resistance, and undetectability (see Section 2.2). The following definitions of these security notions are adapted from [32] to general finite input and output spaces.

Definition 1 (One-Wayness). *We say that $\mathcal{F}_\lambda$ is one-way (OW) if for every probabilistic polynomial-time adversary $\mathcal{A}$,*

$$\Pr\big[k \xleftarrow{\$} \mathcal{K}_\lambda,\, x \xleftarrow{\$} \mathcal{X}_\lambda,\, y = f_k(x),\, x' \leftarrow \mathcal{A}(k, y) : \\ x' \in \mathcal{X}_\lambda \ \wedge \ y = f_k(x')\big] \le \mathrm{negl}(\lambda), \tag{1}$$

where the probability is taken over the randomness of k, x, and $\mathcal{A}$. Equivalently,

$$\mathrm{InSec}^{\mathsf{OW}}(\mathcal{F}_\lambda; t) := \sup_{\mathcal{A}} \mathrm{Succ}^{\mathsf{OW}}_{\mathcal{F}_\lambda}(\mathcal{A}) \le \mathrm{negl}(\lambda),$$

where $\mathrm{Succ}^{\mathsf{OW}}_{\mathcal{F}_\lambda}(\mathcal{A})$ *denotes the probability in Equation* (1)*, the supremum is taken over all probabilistic adversaries* $\mathcal{A}$ *running in time at most* t*, and* $t = t(\lambda)$ *is polynomially bounded.*

Definition 2 (Second-Preimage Resistance). *We say that* $\mathcal{F}_\lambda$ *is second-preimage resistant (SPR) if for every probabilistic polynomial-time adversary* $\mathcal{A}$*,*

$$\Pr\big[k \xleftarrow{\$} \mathcal{K}_\lambda,\, x \xleftarrow{\$} \mathcal{X}_\lambda,\, x' \leftarrow \mathcal{A}(k,x) : \\ x' \in \mathcal{X}_\lambda \ \wedge\ f_k(x) = f_k(x') \ \wedge\ x \neq x'\big] \leq \mathrm{negl}(\lambda), \tag{2}$$

where the probability is taken over the randomness of k*,* x*, and* $\mathcal{A}$*. Equivalently,*

$$\mathrm{InSec}^{\mathsf{SPR}}(\mathcal{F}_\lambda; t) := \sup_{\mathcal{A}} \mathrm{Succ}^{\mathsf{SPR}}_{\mathcal{F}_\lambda}(\mathcal{A}) \leq \mathrm{negl}(\lambda),$$

where $\mathrm{Succ}^{\mathsf{SPR}}_{\mathcal{F}_\lambda}(\mathcal{A})$ *denotes the probability in Equation* (2)*, the supremum is taken over all probabilistic adversaries* $\mathcal{A}$ *running in time at most* t*, and* $t = t(\lambda)$ *is polynomially bounded.*

To define undetectability, we first recall the notion of distinguishing advantage.

Definition 3 (Distinguishing Advantage). *Let* $\mathcal{D}_1$ *and* $\mathcal{D}_2$ *be two distributions, and let an adversary* $\mathcal{A}$ *be given a sample drawn from one of them, chosen uniformly at random. The adversary outputs* 1 *if it believes the sample was drawn from* $\mathcal{D}_1$ *and* 0 *otherwise. The distinguishing advantage of* $\mathcal{A}$ *between* $\mathcal{D}_1$ *and* $\mathcal{D}_2$ *is defined as*

$$\mathrm{Adv}_{\mathcal{D}_1,\mathcal{D}_2}(\mathcal{A}) := \Big|\Pr_{z\leftarrow\mathcal{D}_1}[\mathcal{A}(z)=1] - \Pr_{z\leftarrow\mathcal{D}_2}[\mathcal{A}(z)=1]\Big|,$$

where the probabilities are taken over the sampled input and the randomness of $\mathcal{A}$*.*

We consider the following two distributions over $\mathcal{K}_\lambda \times \mathcal{Y}_\lambda$:

- $\mathcal{D}_{\mathsf{UD},\mathcal{U}}$: sample $k \xleftarrow{\$} \mathcal{K}_\lambda$ and $u \xleftarrow{\$} \mathcal{Y}_\lambda$, and output (k,u);
- $\mathcal{D}_{\mathsf{UD},\mathcal{F}_\lambda}$: sample $k \xleftarrow{\$} \mathcal{K}_\lambda$ and $x \xleftarrow{\$} \mathcal{X}_\lambda$, set $u = f_k(x)$, and output (k,u).

The advantage of an adversary $\mathcal{A}$ against the undetectability of $\mathcal{F}_\lambda$ is defined as

$$\mathrm{Adv}^{\mathsf{UD}}_{\mathcal{F}_\lambda}(\mathcal{A}) := \mathrm{Adv}_{\mathcal{D}_{\mathsf{UD},\mathcal{U}},\mathcal{D}_{\mathsf{UD},\mathcal{F}_\lambda}}(\mathcal{A}).$$

Definition 4 (Undetectability). *We say that* $\mathcal{F}_\lambda$ *is undetectable (UD) if*

$$\mathrm{InSec}^{\mathsf{UD}}(\mathcal{F}_\lambda; t) := \sup_{\mathcal{A}} \mathrm{Adv}^{\mathsf{UD}}_{\mathcal{F}_\lambda}(\mathcal{A}) \leq \mathrm{negl}(\lambda),$$

where the supremum is taken over all probabilistic adversaries $\mathcal{A}$ *running in time at most* t*, and* $t = t(\lambda)$ *is polynomially bounded.*

2.2 W-OTS$^+$ Scheme and its Security

We briefly recall the W-OTS$^+$ signature scheme and its security theorem, following the presentation of Kudinov et al. [32], in order to make explicit the exact assumptions required of the underlying function family.

Construction

In the construction of W-OTS$^+$, we take $\mathcal{X}_\lambda = \mathcal{Y}_\lambda = \{0,1\}^\lambda$. We describe the W-OTS$^+$ signature scheme in [32].

Parameters. The parameters of the W-OTS$^+$ scheme are positive integers defined as follows:

- λ: the security parameter, which also determines the input and output length of the functions in $\mathcal{F}_\lambda$;
- l_{msg}: the bit-length of messages to be signed;
- w: the Winternitz parameter, which determines the base-w representation and the compression level;
- l: the number of λ-bit strings contained in a secret key, public key, and signature. It is computed as

$$l_1 = \left\lceil \frac{l_{\mathrm{msg}}}{\log(w)} \right\rceil, \quad l_2 = \left\lfloor \frac{\log(l_1(w-1))}{\log(w)} \right\rfloor + 1, \quad l = l_1 + l_2.$$

Chaining Functions. The function family $\mathcal{F}_\lambda$ is used to define the chaining functions. For $x \in \{0,1\}^\lambda$, a randomness vector $\mathbf{r} = (r_1, \ldots, r_j) \in \{0,1\}^{(\lambda,j)}$, and integers $j \geq i \geq 0$, we define

$$c_k^i(x, \mathbf{r}) = \begin{cases} x, & \text{if } i = 0, \\ f_k\big(c_k^{i-1}(x, \mathbf{r}) \oplus r_i\big), & \text{if } i > 0. \end{cases}$$

Here, $\mathbf{r} \in \{0,1\}^{(\lambda,j)}$ denotes a sequence of j bit strings of length λ.

Key Generation $\mathsf{Kg}(1^\lambda)$. First, sample a function key $k \xleftarrow{\$} \mathcal{K}_\lambda$ and randomness $\mathbf{r} = (r_1, \ldots, r_{w-1}) \xleftarrow{\$} \{0,1\}^{(\lambda,w-1)}$. Next, generate the secret key

$$\mathsf{sk} = (\mathsf{sk}_1, \ldots, \mathsf{sk}_l) \xleftarrow{\$} \{0,1\}^{(\lambda,l)}.$$

The public key is computed as

$$\mathsf{pk} = (\mathsf{pk}_0, \mathsf{pk}_1, \ldots, \mathsf{pk}_l) = \big((\mathbf{r}, k),\, c_k^{w-1}(\mathsf{sk}_1, \mathbf{r}), \ldots, c_k^{w-1}(\mathsf{sk}_l, \mathbf{r})\big).$$

Signature Generation $\mathsf{Sign}(\mathsf{sk}, M)$. The signer first converts the message M into its base-w representation $M = (M_1, \ldots, M_{l_1})$, where $M_i \in \{0, \ldots, w-1\}$. The checksum is computed as

$$C = \sum_{i=1}^{l_1} (w - 1 - M_i),$$

and represented in base w as $C = (C_1, \ldots, C_{l_2})$. Let $B = (b_1, \ldots, b_l) = M \| C$. The signature on M is then computed as

$$\sigma = (\sigma_1, \ldots, \sigma_l) = \left(c_k^{b_1}(\mathsf{sk}_1, \mathbf{r}), \ldots, c_k^{b_l}(\mathsf{sk}_l, \mathbf{r})\right).$$

Signature Verification $\mathsf{Vf}(\mathsf{pk}, \sigma, M)$. The verifier first computes $B = (b_1, \ldots, b_l)$ from M as above, and then checks whether

$$(\mathsf{pk}_1, \ldots, \mathsf{pk}_l) = \left(c_k^{w-1-b_1}(\sigma_1, \mathbf{r}_{b_1+1,w-1}), \ldots, c_k^{w-1-b_l}(\sigma_l, \mathbf{r}_{b_l+1,w-1})\right),$$

where

$$\mathbf{r}_{a,b} = \begin{cases} (r_a, r_{a+1}, \ldots, r_b), & \text{if } b \geq a, \\ \epsilon, & \text{otherwise.} \end{cases}$$

Security Theorem

The existential unforgeability under chosen-message attacks (EU-CMA) security of a digital signature scheme $\mathsf{Dss} = (\mathsf{Kg}, \mathsf{Sign}, \mathsf{Vf})$ is defined via the following experiment [32].

Experiment $\mathsf{Exp}^{\mathsf{EU\text{-}CMA}}_{\mathsf{Dss}(1^\lambda)}(\mathcal{A})$

- $(\mathsf{sk}, \mathsf{pk}) \leftarrow \mathsf{Kg}(1^\lambda)$;
- $(M^*, \sigma^*) \leftarrow \mathcal{A}^{\mathsf{Sign}(\mathsf{sk}, \cdot)}(\mathsf{pk})$;
- Let $\{(M_i, \sigma_i)\}_{i=1}^{q_s}$ denote the query–response pairs corresponding to the adversary's signing oracle queries;
- Return 1 if and only if $\mathsf{Vf}(\mathsf{pk}, \sigma^*, M^*) = 1$ and $M^* \notin \{M_i\}_{i=1}^{q_s}$.

The success probability of $\mathcal{A}$ in the above experiment is defined as

$$\mathrm{Succ}^{\mathsf{EU\text{-}CMA}}_{\mathsf{Dss}(1^\lambda), q_s}(\mathcal{A}) = \Pr\left[\mathsf{Exp}^{\mathsf{EU\text{-}CMA}}_{\mathsf{Dss}(1^\lambda)}(\mathcal{A}) = 1\right],$$

where the probability is taken over the randomness of Kg, Sign, and $\mathcal{A}$, and where q_s denotes the number of signing oracle queries made by $\mathcal{A}$.

Definition 5 (EU-CMA Security). *We say that* Dss *is EU-CMA-secure if its insecurity is negligible in* λ*, namely,*

$$\mathrm{InSec}^{\mathsf{EU\text{-}CMA}}(\mathsf{Dss}(1^\lambda); t, q_s) := \sup_{\mathcal{A}} \mathrm{Succ}^{\mathsf{EU\text{-}CMA}}_{\mathsf{Dss}(1^\lambda), q_s}(\mathcal{A}) = \mathrm{negl}(\lambda),$$

where the supremum is taken over all probabilistic adversaries $\mathcal{A}$ *running in time at most* t *and making at most* q_s *signing oracle queries.*

When analyzing the EU-CMA security of W-OTS$^+$, the adversary is allowed at most one signing query, i.e., $q_s = 1$. This restriction is inherent to the scheme: with two carefully chosen signing queries, an adversary can forge a signature under the same W-OTS$^+$ key with probability 1.

In this paper, we use the following security bound of W-OTS$^+$ [32]:

Theorem 1. *Suppose $w, l_{\mathrm{msg}} = \mathrm{poly}(\lambda)$. Let $\mathcal{F}_\lambda$ be a one-way, second-preimage resistant, and undetectable function family. Then the insecurity of the W-OTS$^+$ scheme against EU-CMA attacks is bounded by*

$$\mathrm{InSec}^{\mathsf{EU\text{-}CMA}}\big(\text{W-OTS}^+(1^\lambda, w, l_{\mathrm{msg}}); t, q_s = 1\big) < lw\Big(\mathrm{InSec}^{\mathsf{OW}}(\mathcal{F}_\lambda; \tilde{t}) + w \cdot \mathrm{InSec}^{\mathsf{SPR}}(\mathcal{F}_\lambda; \tilde{t}) + w \cdot \mathrm{InSec}^{\mathsf{UD}}(\mathcal{F}_\lambda; \tilde{t})\Big),$$

where $\tilde{t} = t + 3lw + w - 2$, and time is measured in the number of evaluations of functions from $\mathcal{F}_\lambda$.

3 Cryptographic Properties of the MQ Function Family

We consider the following keyed function family of multivariate quadratic (MQ) functions:

$$\mathcal{Q}_{n,m,q} := \left\{ f_{\mathbf{A},\mathbf{b},\boldsymbol{c}} : \mathbb{F}_q^n \to \mathbb{F}_q^m \;\middle|\; \begin{array}{c} f_{\mathbf{A},\mathbf{b},\boldsymbol{c}}(\mathbf{x}) = \left(\mathbf{x}^\mathsf{T}\mathbf{A}_i\mathbf{x} + \mathbf{b}_i^\mathsf{T}\mathbf{x} + c_i\right)_{i=1}^m, \\ \text{where } \mathbf{A}_i \in \mathbb{F}_q^{n\times n} \text{ is upper triangular,} \\ \mathbf{b}_i \in \mathbb{F}_q^n,\ c_i \in \mathbb{F}_q, \quad i = 1, \ldots, m. \end{array} \right\},$$

where we use the coefficient tuple $k = (\mathbf{A}, \mathbf{b}, \boldsymbol{c}) = (\mathbf{A}_i, \mathbf{b}_i, c_i)_{i=1}^m \in \mathbb{F}_q^{N_{n,m}} \triangleq \mathcal{K}$ as the key, with $N_{n,m} = m\left(\frac{n(n+1)}{2} + n + 1\right)$. For $q = 2$, this coefficient-tuple representation is redundant since $z^2 = z$ as functions over $\mathbb{F}_2$. This redundancy does not affect our security definitions or reductions, since the key is the coefficient tuple itself and random keys are sampled uniformly from $\mathcal{K}$.

In this section, we study the cryptographic properties of the MQ function family $\mathcal{Q}_{n,m,q}$ that are required by the security framework of W-OTS$^+$ as stated in Theorem 1: one-wayness (OW), second-preimage resistance (SPR), and undetectability (UD).

In multivariate cryptography, one-wayness of $\mathcal{Q}_{n,m,q}$ is a standard assumption [13, 40]:

Definition 6 (MQOW Assumption [13]). *The (n, m, q)-MQOW assumption states that the MQ function family $\mathcal{Q}_{n,m,q}$ is one-way, in the sense of Definition 1. Specifically:*

Given $k = (\mathbf{A}_i, \mathbf{b}_i, c_i)_{i=1}^m \xleftarrow{\$} \mathbb{F}_q^{N_{n,m}}$, and $\mathbf{y} = f_k(\mathbf{x})$, where $\mathbf{x} \xleftarrow{\$} \mathbb{F}_q^n$ is unknown, no probabilistic polynomial-time algorithm outputs, with non-negligible probability, an $\mathbf{x}' \in \mathbb{F}_q^n$ satisfying $f_k(\mathbf{x}') = \mathbf{y}$.

When $m = \Theta(n)$, research on specific algorithms indicates that MQOW assumptions is widely believed to hold [4, 21, 42].

3.1 Second-Preimage Resistance

This subsection provides a tight reduction from the SPR property for $\mathcal{Q}_{n,m,q}$ to its OW.

As an intermediate step, we consider the following family of multivariate quadratic functions without constant terms:

$$\tilde{\mathcal{Q}}_{n,m,q} := \left\{ f_{\mathbf{A},\mathbf{b}} : \mathbb{F}_q^n \to \mathbb{F}_q^m \;\middle|\; f_{\mathbf{A},\mathbf{b}}(\mathbf{x}) = \left(\mathbf{x}^\mathsf{T}\mathbf{A}_i\mathbf{x} + \mathbf{b}_i^\mathsf{T}\mathbf{x}\right)_{i=1}^m \right\},$$

where each $\mathbf{A}_i \in \mathbb{F}_q^{n\times n}$ is upper triangular and $\mathbf{b}_i \in \mathbb{F}_q^n$.

Since shifting the output by a constant vector does not affect the occurrence of second preimages, the SPR of the full function family $\mathcal{Q}_{n,m,q}$ is equivalent to that of $\tilde{\mathcal{Q}}_{n,m,q}$, and we can establish the SPR of $\mathcal{Q}_{n,m,q}$ by analyzing $\tilde{\mathcal{Q}}_{n,m,q}$.

The hardness of finding second preimages for $\tilde{\mathcal{Q}}_{n,m,q}$ has been established under a variant of the MQOW assumption, which we refer to as the zero-constant MQ (ZCMQ) assumption.

Definition 7 (Zero-Constant MQ (ZCMQ) Assumption [3]). *The (n,m,q)-ZCMQ assumption states that it is computationally infeasible to solve the following problem: Given m upper triangular matrices $\mathbf{A}_1,\ldots,\mathbf{A}_m \in \mathbb{F}_q^{n\times n}$ and m vectors $\mathbf{b}_1,\ldots,\mathbf{b}_m \in \mathbb{F}_q^n$, all chosen independently and uniformly at random, find a non-zero $\mathbf{x} \in \mathbb{F}_q^n$ satisfying*

$$\mathbf{x}^\mathsf{T}\mathbf{A}_i\mathbf{x} + \mathbf{b}_i^\mathsf{T}\mathbf{x} = 0 \qquad \text{for all } i = 1,\ldots,m.$$

Lemma 1 (ZCMQ Assumption ⇒ SPR for $\tilde{\mathcal{Q}}_{n,m,q}$ [3]). *Under the (n,m,q)-ZCMQ assumption, the family $\tilde{\mathcal{Q}}_{n,m,q}$ is second-preimage resistant. Specifically, any adversary $\mathcal{A}$ that breaks the SPR property of $\tilde{\mathcal{Q}}_{n,m,q}$ in time t with success probability ϵ can be converted into an algorithm $\mathcal{B}$ that breaks the (n,m,q)-ZCMQ assumption with time $t' = t + \mathrm{poly}(n,m,\log q)$ and success probability $\epsilon' = \epsilon$.*

We refer to [3] for the proof of Lemma 1.

Now we show that the MQOW assumption implies the ZCMQ assumption.

Theorem 2 (MQOW Assumption ⇒ ZCMQ Assumption). *Under the (n,m,q)-MQOW assumption, the (n,m,q)-ZCMQ assumption holds.*

More concretely, suppose that a probabilistic algorithm $\mathcal{B}$ solves a uniformly random (n,m,q)-ZCMQ instance in time t with success probability ϵ. Then there exists a probabilistic algorithm $\mathcal{C}^{\mathcal{B}}$ that breaks the one-wayness of $\mathcal{Q}_{n,m,q}$ in time $t' = t + \mathrm{poly}(n,m,\log q)$, with success probability $\epsilon' \geq \left(1 - \frac{1}{q}\right)\frac{q^{n-1}}{q^n-1}\epsilon$.

Proof. We construct $\mathcal{C}^{\mathcal{B}}$ in Algorithm 1.

Overview of $\mathcal{C}^{\mathcal{B}}$. Algorithm 1 transforms the constants of the MQOW instance into the linear coefficients of an ZCMQ instance $k^\star$: for each $i \in [m]$ it computes $d_i = c_i - y_i$ and replaces the linear coefficient $\mathbf{b}_i$ by $\mathbf{b}_i + d_i\mathbf{u}$, where the nonzero random vector $\mathbf{u}$ is used for all m equations. If $\mathcal{B}$ returns a valid solution $\mathbf{z}$ of the ZCMQ instance $k^\star$, and $\mathbf{u}^\mathsf{T}\mathbf{z} = 1$, then $\mathbf{z}$ is a preimage of the original MQOW instance.

We prove three properties: the transformation is correct, the $k^\star$ is distributed as a random ZCMQ instance, and the extraction succeeds with the claimed probability.

Algorithm 1 The Algorithm $\mathcal{C}^{\mathcal{B}}$

Require: An (n, m, q)-MQOW challenge $(k, \mathbf{y})$, where $\mathbf{y} = f_k(\mathbf{x})$ for a uniformly random and unknown $\mathbf{x} \in \mathbb{F}_q^n$; access to an (n, m, q)-ZCMQ solver $\mathcal{B}$
Ensure: A preimage $\mathbf{x}' \in \mathbb{F}_q^n$ satisfying $f_k(\mathbf{x}') = \mathbf{y}$, or $\perp$

Parse $k = (\mathbf{A}_i, \mathbf{b}_i, c_i)_{i=1}^m$
Sample $\mathbf{u} \xleftarrow{\$} \mathbb{F}_q^n \setminus \{\mathbf{0}\}$
for $i = 1, \ldots, m$ **do**
    Set $d_i \leftarrow c_i - y_i$
    Set $\mathbf{A}_i^\star \leftarrow \mathbf{A}_i$
    Set $\mathbf{b}_i^\star \leftarrow \mathbf{b}_i + d_i\mathbf{u}$
end for
Let $k^\star \leftarrow (\mathbf{A}_i^\star, \mathbf{b}_i^\star)_{i=1}^m$
Query $\mathcal{B}$ on $k^\star$ and let $\mathbf{z} \leftarrow \mathcal{B}(k^\star)$
if $\mathbf{z} \notin \mathbb{F}_q^n \setminus \{\mathbf{0}\}$ **then**
    return $\perp$
end if
if $\mathbf{z}^\mathsf{T}\mathbf{A}_i^\star\mathbf{z} + (\mathbf{b}_i^\star)^\mathsf{T}\mathbf{z} \neq 0$ for some $i \in [m]$ **then**
    return $\perp$
end if
if $\mathbf{u}^\mathsf{T}\mathbf{z} \neq 1$ **then**
    return $\perp$
end if
return $\mathbf{z}$

Correctness of the transformation. Suppose that $\mathcal{B}$ returns a non-zero vector $\mathbf{z}$ satisfying $\mathbf{z}^\mathsf{T}\mathbf{A}_i^\star\mathbf{z} + (\mathbf{b}_i^\star)^\mathsf{T}\mathbf{z} = 0$ for every $i \in [m]$, and suppose additionally that $\mathbf{u}^\mathsf{T}\mathbf{z} = 1$. Then, for every $i \in [m]$ we have

$$0 = \mathbf{z}^\mathsf{T}\mathbf{A}_i^\star\mathbf{z} + (\mathbf{b}_i^\star)^\mathsf{T}\mathbf{z} = \mathbf{z}^\mathsf{T}\mathbf{A}_i\mathbf{z} + \mathbf{b}_i^\mathsf{T}\mathbf{z} + (c_i - y_i)\mathbf{u}^\mathsf{T}\mathbf{z} = \mathbf{z}^\mathsf{T}\mathbf{A}_i\mathbf{z} + \mathbf{b}_i^\mathsf{T}\mathbf{z} + c_i - y_i,$$

which means that $f_{k,i}(\mathbf{z}) = y_i$ for every $i \in [m]$. Therefore, $f_k(\mathbf{z}) = \mathbf{y}$, and the output of $\mathcal{C}^{\mathcal{B}}$ is a valid preimage of the MQOW challenge.

Distribution of the constructed ZCMQ instance. Let the hidden input in the MQOW experiment be $\mathbf{x} \in \mathbb{F}_q^n$. For every $i \in [m]$, the challenge output satisfies $y_i = \mathbf{x}^\mathsf{T}\mathbf{A}_i\mathbf{x} + \mathbf{b}_i^\mathsf{T}\mathbf{x} + c_i$. Consequently, $d_i = c_i - y_i = -(\mathbf{x}^\mathsf{T}\mathbf{A}_i\mathbf{x} + \mathbf{b}_i^\mathsf{T}\mathbf{x})$, and

$$\mathbf{b}_i^\star = \mathbf{b}_i - \left(\mathbf{x}^\mathsf{T}\mathbf{A}_i\mathbf{x} + \mathbf{b}_i^\mathsf{T}\mathbf{x}\right)\mathbf{u} = \left(\mathbf{I}_n - \mathbf{u}\mathbf{x}^\mathsf{T}\right)\mathbf{b}_i - \left(\mathbf{x}^\mathsf{T}\mathbf{A}_i\mathbf{x}\right)\mathbf{u}. \tag{3}$$

Let G be the event that $\mathbf{u}^\mathsf{T}\mathbf{x} \neq 1$. For every fixed non-zero $\mathbf{u}$, since $\mathbf{x}$ is uniform over $\mathbb{F}_q^n$, $\mathbf{u}^\mathsf{T}\mathbf{x}$ is uniformly distributed over $\mathbb{F}_q$ when $\mathbf{x}$ is uniform. Therefore, $\Pr[G] = 1 - \frac{1}{q}$.

Fix arbitrary values of $\mathbf{x}$ and $\mathbf{u}$ satisfying G, and fix an arbitrary upper-triangular matrix $\mathbf{A}_i$. Linear algebra (see [28] Corollary 18.1.3) tells us that $\det\left(\mathbf{I}_n - \mathbf{u}\mathbf{x}^\mathsf{T}\right) = 1 - \mathbf{x}^\mathsf{T}\mathbf{u}$. Thus, on the event G, the matrix $\mathbf{I}_n - \mathbf{u}\mathbf{x}^\mathsf{T}$ is invertible. Denote $\mathcal{K}_{\mathsf{ZCMQ}}$ to be the coefficient space of (n, m, q)-ZCMQ instances, then by Equation (3), we know that the map $\phi_{\mathbf{x},\mathbf{u}}\colon \mathcal{K}_{\mathsf{ZCMQ}} \to \mathcal{K}_{\mathsf{ZCMQ}}, (\mathbf{A}_i, \mathbf{b}_i)_{i=1}^m \mapsto$

$(\mathbf{A}_i^*, b_i^\star)_{i=1}^m$ is a bijection, with inverse

$$(\mathbf{A}_i^*, b_i^\star)_{i=1}^m \mapsto (\mathbf{A}_i, \mathbf{b}_i)_{i=1}^m = (\mathbf{A}_i^*, \left(\mathbf{I}_n - \mathbf{u}\mathbf{x}^\mathsf{T}\right)^{-1} \left(\mathbf{b}_i^\star + \left(\mathbf{x}^\mathsf{T}\mathbf{A}_i^\star\mathbf{x}\right)\mathbf{u}\right))_{i=1}^m.$$

Hence, conditioned on G, since $(\mathbf{A}_i, \mathbf{b}_i)_{i=1}^m$ is uniformly random over $\mathcal{K}_{\mathsf{ZCMQ}}$, it follows that $(\mathbf{A}_i, \mathbf{b}_i)_{i=1}^m$ is also uniformly random over $\mathcal{K}_{\mathsf{ZCMQ}}$, and is a uniformly random (n, m, q)-ZCMQ instance.

Let S denote the event that $\mathcal{B}$ outputs a non-zero solution of $k^\star$. Since the instance sent to $\mathcal{B}$ is uniformly distributed conditioned on G, the assumed success probability of $\mathcal{B}$ gives $\Pr[S \mid G] = \epsilon$.

Extraction probability. We firstly show that the $\mathcal{B}$'s view does not reveal any information on $\mathbf{u}$. Fix any $\mathbf{u}_0 \in \mathbb{F}_q^n \setminus \{\mathbf{0}\}$ and any ZCMQ instance $k_0 \in \mathcal{K}_{\mathsf{ZCMQ}}$. By the law of total probability, we obtain

$$\begin{aligned}
&\Pr[k^\star = k_0 \mid \mathbf{u} = \mathbf{u}_0, G] \\
&= \sum_{\mathbf{x}' \in \mathbb{F}_q^n, \mathbf{u}_0^\mathsf{T}\mathbf{x}' \neq 1} \Pr[\, k^\star = k_0, \mathbf{x} = \mathbf{x}' \mid \mathbf{u} = \mathbf{u}_0, G] \\
&= \sum_{\mathbf{x}' \in \mathbb{F}_q^n, \mathbf{u}_0^\mathsf{T}\mathbf{x}' \neq 1} \Pr[\, k^\star = k_0 \mid \mathbf{x} = \mathbf{x}', \mathbf{u} = \mathbf{u}_0, G] \Pr[\, \mathbf{x} = \mathbf{x}' \mid \mathbf{u} = \mathbf{u}_0, G] \\
&= \frac{1}{|\mathcal{K}_{\mathsf{ZCMQ}}|}.
\end{aligned}$$

Since the above probability does not depend on $\mathbf{u}_0$, it follows that, conditioned on G, the instance $k^\star$ is uniformly distributed and independent of $\mathbf{u}$. Equivalently, for every $k_0 \in \mathcal{K}_{\mathsf{ZCMQ}}$ and every non-zero $\mathbf{u}_0$, Bayes' rule gives

$$\begin{aligned}
\Pr[\mathbf{u} = \mathbf{u}_0 \mid k^\star = k_0, G] &= \frac{\Pr[k^\star = k_0 \mid \mathbf{u} = \mathbf{u}_0, G] \Pr[\mathbf{u} = \mathbf{u}_0 \mid G]}{\Pr[k^\star = k_0 \mid G]} \\
&= \Pr[\mathbf{u} = \mathbf{u}_0 \mid G].
\end{aligned}$$

Thus, observing the constructed ZCMQ instance does not change the posterior distribution of $\mathbf{u}$.

Now we determine the conditional distribution of $\mathbf{u}$ under G. For every fixed $\mathbf{u}_0 \in \mathbb{F}_q^n \setminus \{\mathbf{0}\}$, we have $\Pr[G \mid \mathbf{u} = \mathbf{u}_0] = 1 - 1/q$. This probability is independent of $\mathbf{u}_0$. Since $\mathbf{u}$ is sampled uniformly from $\mathbb{F}_q^n \setminus \{\mathbf{0}\}$, we have $\Pr[\mathbf{u} = \mathbf{u}_0] = 1/(q^n - 1)$. Since $\Pr[G] = 1 - 1/q$, we have (by Bayes' rule)

$$\Pr[\mathbf{u} = \mathbf{u}_0 \mid G] = \frac{\Pr[G \mid \mathbf{u} = \mathbf{u}_0] \Pr[\mathbf{u} = \mathbf{u}_0]}{\Pr[G]} = \frac{\left(1 - \frac{1}{q}\right)\frac{1}{q^n - 1}}{1 - \frac{1}{q}} = \frac{1}{q^n - 1}.$$

Thus, conditioned on G, the vector $\mathbf{u}$ remains uniformly distributed over $\mathbb{F}_q^n \setminus \{\mathbf{0}\}$.

Therefore, $\Pr[\mathbf{u} = \mathbf{u}_0 \mid k^\star = k_0, G] = \Pr[\mathbf{u} = \mathbf{u}_0 \mid G] = 1/(q^n - 1)$, which means that, conditioned on G, after observing the constructed ZCMQ instance $k^\star$, $\mathcal{B}$'s conditional view leaves $\mathbf{u}$ uniformly distributed over $\mathbb{F}_q^n \setminus \{\mathbf{0}\}$, and the $\mathcal{B}$'s view does not reveal any information on $\mathbf{u}$.

Let R denote the internal randomness of $\mathcal{B}$, sampled independently of all the random variables in the reduction. Although $k^\star$ is constructed using $\mathbf{u}$, the calculation above shows that, conditioned on G, the distribution of $(k^\star, R)$ is independent of $\mathbf{u}$. More precisely, for every $k_0 \in \mathcal{K}_{\mathsf{ZCMQ}}$, every possible value r of R, and every $\mathbf{u}_0 \in \mathbb{F}_q^n \setminus \{\mathbf{0}\}$, we have

$$\begin{aligned}\Pr[k^\star = k_0, R = r \mid \mathbf{u} = \mathbf{u}_0, G] &= \Pr[k^\star = k_0 \mid \mathbf{u} = \mathbf{u}_0, G] \Pr[R = r] \\ &= \frac{1}{|\mathcal{K}_{\mathsf{ZCMQ}}|} \Pr[R = r],\end{aligned}$$

and the right-hand side does not depend on $\mathbf{u}_0$.

Let $\mathbf{z} = \mathcal{B}(k^\star; R)$, and let S denote the event that $\mathbf{z}$ is a non-zero solution of $k^\star$. Since both $\mathbf{z}$ and S are determined by $(k^\star, R)$, the conditional independence established above implies that, for every $\mathbf{u}_0 \in \mathbb{F}_q^n \setminus \{\mathbf{0}\}$ and every non-zero $\mathbf{z}_0$,

$$\Pr[\mathbf{u} = \mathbf{u}_0 \mid G, S, \mathbf{z} = \mathbf{z}_0] = \Pr[\mathbf{u} = \mathbf{u}_0 \mid G] = \frac{1}{q^n - 1}.$$

Thus, after $\mathcal{B}$ observes $k^\star$, uses its internal randomness, and returns a non-zero solution $\mathbf{z}_0$ of $k^\star$, the vector $\mathbf{u}$ remains uniformly distributed over $\mathbb{F}_q^n \setminus \{\mathbf{0}\}$. Since the affine equation $\mathbf{u}^\mathsf{T}\mathbf{z}_0 = 1$ has exactly q^{n-1} non-zero solutions in $\mathbb{F}_q^n$, we have

$$\Pr[\mathbf{u}^\mathsf{T}\mathbf{z}_0 = 1 \mid G, S, \mathbf{z} = \mathbf{z}_0] = \frac{q^{n-1}}{q^n - 1}.$$

Since this probability is the same for every possible non-zero output $\mathbf{z}_0$, the law of total probability gives

$$\begin{aligned}\Pr[\mathbf{u}^\mathsf{T}\mathbf{z} = 1 \mid G, S] &= \sum_{\mathbf{z}_0 \in \mathbb{F}_q^n \setminus \{\mathbf{0}\}} \Pr[\mathbf{u}^\mathsf{T}\mathbf{z}_0 = 1 \mid G, S, \mathbf{z} = \mathbf{z}_0] \Pr[\mathbf{z} = \mathbf{z}_0 \mid G, S] \\ &= \frac{q^{n-1}}{q^n - 1} \sum_{\mathbf{z}_0 \in \mathbb{F}_q^n \setminus \{\mathbf{0}\}} \Pr[\mathbf{z} = \mathbf{z}_0 \mid G, S] = \frac{q^{n-1}}{q^n - 1}.\end{aligned}$$

In all, we obtain

$$\begin{aligned}\epsilon' = \Pr[\mathcal{C}^{\mathcal{B}} \text{ inverts the MQOW challenge}] &\geq \Pr[G] \Pr[S \mid G] \Pr[\mathbf{u}^\mathsf{T}\mathbf{z} = 1 \mid G, S] \\ &= \left(1 - \frac{1}{q}\right) \frac{q^{n-1}}{q^n - 1} \epsilon.\end{aligned}$$

Since $\mathcal{C}$ invokes $\mathcal{B}$ once and performing all remaining operations require only $\mathrm{poly}(n, m, \log q)$ time, we have $t' = t + \mathrm{poly}(n, m, \log q)$, completing the proof. □

Combining Lemma 1 with Theorem 2, we obtain the following concrete security bound for the second-preimage resistance of $\mathcal{Q}_{n,m,q}$.

Theorem 3 (MQOW Assumption ⇒ SPR of $\mathcal{Q}_{n,m,q}$). *Under the (n,m,q)-MQOW assumption, the function family $\mathcal{Q}_{n,m,q}$ is second-preimage resistant.*

More concretely, for each t, we have

$$\mathrm{InSec}^{\mathsf{SPR}}(\mathcal{Q}_{n,m,q};t) \le \frac{q^2-q^{2-n}}{q-1}\cdot \mathrm{InSec}^{\mathsf{OW}}(\mathcal{Q}_{n,m,q};t') \le 2q\cdot \mathrm{InSec}^{\mathsf{OW}}(\mathcal{Q}_{n,m,q};t'),$$

where $t' = t + \mathrm{poly}(n,m,\log q)$.

Proof. Let $\mathcal{A}$ be an adversary that breaks the SPR property of $\mathcal{Q}_{n,m,q}$ in time t with success probability ϵ.

First, observe that the constant terms of an MQ function do not affect its second preimages. In fact, for every $k=(\mathbf{A},\mathbf{b},\boldsymbol{c})$ and every $\mathbf{x},\mathbf{x}'\in\mathbb{F}_q^n$, we have

$$f_{\mathbf{A},\mathbf{b},\boldsymbol{c}}(\mathbf{x}') = f_{\mathbf{A},\mathbf{b},\boldsymbol{c}}(\mathbf{x}) \quad\Longleftrightarrow\quad f_{\mathbf{A},\mathbf{b}}(\mathbf{x}') = f_{\mathbf{A},\mathbf{b}}(\mathbf{x}).$$

Moreover, when the key $(\mathbf{A},\mathbf{b},\boldsymbol{c})$ of $\mathcal{Q}_{n,m,q}$ is sampled uniformly, the coefficient tuple $(\mathbf{A},\mathbf{b})$ is distributed as a uniform key of $\tilde{\mathcal{Q}}_{n,m,q}$. Thus, $\mathcal{A}$ can be viewed as an SPR adversary against $\tilde{\mathcal{Q}}_{n,m,q}$ with the same running time and success probability.

By Lemma 1, the adversary $\mathcal{A}$ can be converted into an algorithm $\mathcal{B}$ that solves a uniformly random (n,m,q)-ZCMQ instance with success probability $\epsilon_{\mathsf{ZCMQ}} = \epsilon$ and running time $t_{\mathsf{ZCMQ}} = t + \mathrm{poly}(n,m,\log q)$.

Applying Theorem 2 to $\mathcal{B}$, we obtain an algorithm $\mathcal{C}$ that inverts a random (n,m,q)-MQOW instance with success probability $\epsilon_{\mathsf{MQOW}} \ge \left(1-\frac{1}{q}\right)\frac{q^{n-1}}{q^n-1}\epsilon_{\mathsf{ZCMQ}}$ $= \left(1-\frac{1}{q}\right)\frac{q^{n-1}}{q^n-1}\epsilon$. The running time of $\mathcal{C}^{\mathcal{B}}$ is $t' = t + \mathrm{poly}(n,m,\log q)$.

Since $\mathcal{A}$ was arbitrary, taking the supremum over all SPR adversaries running in time t yields

$$\mathrm{InSec}^{\mathsf{SPR}}(\mathcal{Q}_{n,m,q};t) \le \frac{q^2-q^{2-n}}{q-1}\cdot \mathrm{InSec}^{\mathsf{OW}}(\mathcal{Q}_{n,m,q};t'),$$

For every $q\ge 2$, we have $q/(q-1)\le 2$, and hence $\frac{q^n-1}{\left(1-\frac{1}{q}\right)q^{n-1}} = \frac{q^2-q^{2-n}}{q-1} \le \frac{q^2}{q-1} \le 2q$. □

Remark 1 (Tightness). The security loss in Theorem 3 is $\frac{q^2-q^{2-n}}{q-1} \le 2q$. In the context of multivariate cryptography, where q is a small constant and n grows with the security parameter, our reduction is essentially tight.

Based on this reduction, the SPR of $\mathcal{Q}_{n,m,q}$ is reduced to the (n,m,q)-MQOW assumption. In the standard parameter regime where $m=\Theta(n)$, $\mathcal{Q}_{n,m,q}$ provides a robust level of SPR security in practice.

3.2 Undetectability

In this subsection, we revisit the UD property of $\mathcal{Q}_{n,m,q}$, and present a tighter reduction from UD to its OW for a general q. In [6], Berbain et al. presented this reduction for $q = 2$ with an $O(\epsilon^{-2})$ security loss. In [33], Liu et al. presented this reduction for a general q with an $O(\epsilon^{-5})$ loss. Recently, [10] utilized the large-field Goldreich–Levin theorem [18] and obtain an $O(\epsilon^{-4})$ loss for a general q. We will show that the reduction for a general q can be tightened, giving an $O(\epsilon^{-2})$ loss.

This subsection is divided into four parts. In Part 1, we recall the fact that a UD distinguisher for $\mathcal{Q}_{n,m,q}$ can be used to construct a linear form predictor. In Part 2, the main technical step, we show how to turn a linear form predictor with a fixed function-element pair $(H, H(\mathbf{x}))$ into an inverter for a general function H, with an $O(\epsilon^{-2})$ security loss for arbitrary q. In Part 3, we use a linear form predictor to construct an inverter for $\mathcal{Q}_{n,m,q}$ in the average case, maintaining the loss. In Part 4, we combine the three preceding parts to give a reduction from UD to OW for $\mathcal{Q}_{n,m,q}$ together with a concrete security bound.

Part 1: From a UD Distinguisher to a Linear Form Predictor

In this paper, a linear form refers to a $\mathbb{F}_q$-linear map $R : \mathbb{F}_q^n \to \mathbb{F}_q$, i.e., $R(\mathbf{x}) = \sum_{i=1}^{n} r_i x_i$, where $r_i \in \mathbb{F}_q, i = 1, 2, ..., n$.

Lemma 2 (Distinguisher-to-Predictor over $\mathbb{F}_q$). *Let $\mathcal{A}$ be an algorithm distinguishing the distributions $\mathcal{D}_{\mathsf{UD},\mathcal{Q}_{n,m,q}}$ and $\mathcal{D}_{\mathsf{UD},\mathcal{U}}$ with advantage $\epsilon > 0$ and running time t. Then there exists a predictor $\mathcal{P}^{\mathcal{A}}$ such that, on input a randomly chosen $k \in \mathcal{K}$, $f_k(\mathbf{x})$ (with an unknown $\mathbf{x}$ uniformly chosen from $\mathbb{F}_q^n$), and a uniformly chosen linear form $R : \mathbb{F}_q^n \to \mathbb{F}_q$, predicts $R(\mathbf{x})$ with success probability $\frac{1}{q}(1 + \frac{\epsilon}{2})$ and running time $t' = t + 2T_{f_k}$. Here, the probability is over the randomness of $k, \mathbf{x}, R, \mathcal{A}$ and $\mathcal{P}$, and T_{f_k} is the time taken to evaluate f_k.*

The proof of Lemma 2 can be found in Proposition 2 of [33].

Part 2: From a Fixed-Pair Linear Form Predictor to a Reconstructor

This is the core technical part in the reduction from UD to OW. We show that for any fixed function $H : \mathbb{F}_q^n \to \mathbb{F}_q^m$, if there exists an algorithm $\mathcal{P}$ which, on input a linear form $R : \mathbb{F}_q^n \to \mathbb{F}_q$ and $H(\mathbf{x})$, can predict $R(\mathbf{x})$ with success probability at least $1/q + \epsilon$, then one can efficiently construct a short list of candidate preimages of $H(\mathbf{x})$ that contains $\mathbf{x}$ with constant probability.

Lemma 3 (Fixed-Pair q-ary Reconstruction). *Let $H : \mathbb{F}_q^n \to \mathbb{F}_q^m$ be a fixed function, $\mathbf{x} \in \mathbb{F}_q^n$ be fixed and unknown, and $\mathbf{y} = H(\mathbf{x})$. Suppose that there exists a randomized predictor $\mathcal{P}$ which, on input $H, \mathbf{y}$, and a uniformly chosen linear form $R : \mathbb{F}_q^n \to \mathbb{F}_q$, outputs $R(\mathbf{x})$ with success probability at least $1/q + \epsilon$, where the probability is over the randomness of R and the internal randomness of $\mathcal{P}$.*

Denote $\Delta_\epsilon = q\epsilon/(q-1), \ell_\epsilon = \left\lceil \log_q \left(9nq/2\Delta_\epsilon^2\right)\right\rceil$, *and* $M_\epsilon = q^{\ell_\epsilon}$. *Then there exists a reconstruction algorithm* $\mathcal{B}^{\mathcal{P}}$ *that on input* $H, \mathbf{y}$, *outputs a list* $\mathcal{L} \subseteq \mathbb{F}_q^n$ *such that* $\Pr[\mathbf{x} \in \mathcal{L}] \geq \frac{1}{2}$, *where the probability is over the randomness of* $\mathcal{B}$ *and* $\mathcal{P}$. *Moreover, we have* $|\mathcal{L}| \leq qM_\epsilon$, *and the number of calls to* $\mathcal{P}$ *is* $n(q-1)M_\epsilon$. *If one call to* $\mathcal{P}$ *takes time* $T_{\mathcal{P}}$, *and one evaluation of* H *takes time* T_H, *then the running time of* $\mathcal{B}^{\mathcal{P}}$ *is bounded by* $T_{\mathcal{B}} \leq n(q-1)M_\epsilon T_{\mathcal{P}} + qM_\epsilon T_H + M_\epsilon \cdot \mathrm{poly}(n, q, \ell_\epsilon)$.

Proof. We present $\mathcal{B}^{\mathcal{P}}$ in Algorithm 2.

Overview of $\mathcal{B}^{\mathcal{P}}$. The construction of $\mathcal{B}^{\mathcal{P}}$ has four phases.

Phase 1: Sampling a random affine family of linear forms. $\mathcal{B}^{\mathcal{P}}$ first samples $R_0, R_1, \ldots, R_\ell$ as uniformly random linear forms over $\mathbb{F}_q^n$. For every $\alpha = (\alpha_1, \ldots, \alpha_\ell) \in \mathbb{F}_q^\ell$, it defines $U_\alpha := R_0 + \sum_{j=1}^{\ell} \alpha_j R_j$. The extra form R_0 makes each U_α a uniformly random linear form.

Phase 2: Querying the predictor. For each coordinate $i = 1, \ldots, n$, let $L_i(\mathbf{z}) = z_i$ be the i-th coordinate linear form. For every $s \in \mathbb{F}_q^*$ and $\alpha \in \mathbb{F}_q^\ell$, the algorithm queries the predictor on $U_\alpha + sL_i$ and stores $Y_{i,s}(\alpha) := \mathcal{P}(H, \mathbf{y}, U_\alpha + sL_i)$.

Phase 3: Preprocessing. Since $\mathcal{B}^{\mathcal{P}}$ does not know the values $R_0(\mathbf{x}), R_1(\mathbf{x}), \ldots, R_\ell(\mathbf{x})$, it will later enumerate all possible hypotheses for them. A direct computation of all vote counts for all hypotheses would loosen the reduction. To avoid this, $\mathcal{B}^{\mathcal{P}}$ performs a q-ary Fourier preprocessing:

Fix a nontrivial additive character $\chi : \mathbb{F}_q \to \mathbb{C}^*$ (for example, if $q = p^r$ for some prime number p, we take $\chi(a) = \exp(\frac{2\pi i}{p} \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(a))$). For every $i \in [n]$, $s \in \mathbb{F}_q^*$, $c \in \mathbb{F}_q^*$ and $\alpha \in \mathbb{F}_q^\ell$, define $A_{i,s,c}(\alpha) := \chi\big(cY_{i,s}(\alpha)\big)$ and store them. Then, $\mathcal{B}^{\mathcal{P}}$ applies the fast q-ary Fourier transform to the array $A_{i,s,c}$ and stores the results:

$$\widehat{A}_{i,s,c}(u) := \sum_{\alpha \in \mathbb{F}_q^\ell} A_{i,s,c}(\alpha)\chi(-u \cdot \alpha), \qquad \text{for each } u \in \mathbb{F}_q^\ell.$$

Phase 4: Decoding each hypothesis and constructing the list. $\mathcal{B}^{\mathcal{P}}$ enumerates every hypothesis $\rho = (\rho_0, \rho_1, \ldots, \rho_\ell) \in \mathbb{F}_q^{\ell+1}$.

For a fixed hypothesis ρ, denote $\rho' = (\rho_1, \ldots, \rho_\ell)$. If ρ is correct, then $U_\alpha(\mathbf{x}) = \rho_0 + \alpha \cdot \rho'$. For a candidate value $b \in \mathbb{F}_q$, the equality $Y_{i,s}(\alpha) = \rho_0 + \alpha \cdot \rho' + sb$ is interpreted as a vote for $x_i = b$.

Using the Fourier data from Phase 3, the vote count is computed as

$$V_i(\rho, b) = \frac{(q-1)M}{q} + \frac{1}{q} \sum_{s \in \mathbb{F}_q^*} \sum_{c \in \mathbb{F}_q^*} \chi\big(-c(\rho_0 + sb)\big)\widehat{A}_{i,s,c}(c\rho').$$

Then $\mathcal{B}^{\mathcal{P}}$ sets $z_i(\rho)$ to be the b such that $V_i(\rho, b)$ achieves the maximum over all $b \in \mathbb{F}_q$. Doing this for all $i = 1, \ldots, n$ gives a candidate vector $\mathbf{z}(\rho) = (z_1(\rho), \ldots, z_n(\rho))$. Finally, $\mathcal{B}^{\mathcal{P}}$ verifies the candidate by checking whether $H(\mathbf{z}(\rho)) = \mathbf{y}$, and only the verified candidates are kept in the output list $\mathcal{L}$.

Algorithm 2 The Reconstructor Algorithm $\mathcal{B}^{\mathcal{P}}$

Require: A function $H : \mathbb{F}_q^n \to \mathbb{F}_q^m$, a value $\mathbf{y} = H(\mathbf{x})$, an advantage parameter $\epsilon > 0$, and oracle access to $\mathcal{P}(H, \mathbf{y}, \cdot)$

Ensure: A list $\mathcal{L} \subseteq \mathbb{F}_q^n$ of candidate preimages of $\mathbf{y}$

Compute $\Delta \leftarrow q\epsilon/(q-1)$, $\ell \leftarrow \lceil \log_q \left(9nq/2\Delta^2\right) \rceil$, $M \leftarrow q^\ell$

▷ *Phase 1: Sampling a random affine family of linear forms*

for $j = 0, 1, \ldots, \ell$ **do**

  Sample $\mathbf{r}_j \xleftarrow{\$} \mathbb{F}_q^n$

  Define the linear form $R_j(\mathbf{z}) := \mathbf{r}_j \cdot \mathbf{z}$

end for

for $\alpha = (\alpha_1, \ldots, \alpha_\ell) \in \mathbb{F}_q^\ell$ **do**

  Define the linear form $U_\alpha := R_0 + \sum_{j=1}^{\ell} \alpha_j R_j$

end for

▷ *Phase 2: Querying the predictor*

for $(i, s, \alpha) \in [n] \times \mathbb{F}_q^* \times \mathbb{F}_q^\ell$ **do**

  Query the predictor $Y_{i,s}(\alpha) \leftarrow \mathcal{P}\big(H, \mathbf{y}, U_\alpha + sL_i\big)$, where L_i is the i-th coordinate linear form $L_i(\mathbf{z}) := z_i$

end for

▷ *Phase 3: Preprocessing for simultaneous vote computation*

Fix a nontrivial additive character $\chi : \mathbb{F}_q \to \mathbb{C}^*$.

for $(i, s, c) \in [n] \times \mathbb{F}_q^* \times \mathbb{F}_q^*$ **do**

  for $\alpha \in \mathbb{F}_q^\ell$ **do**

    Compute $A_{i,s,c}(\alpha) \leftarrow \chi\big(cY_{i,s}(\alpha)\big)$

  end for

  Apply one fast q-ary Fourier transform to the array $A_{i,s,c}$, and store all values $\widehat{A}_{i,s,c}(u) := \sum_{\alpha \in \mathbb{F}_q^\ell} A_{i,s,c}(\alpha)\chi(-u \cdot \alpha), u \in \mathbb{F}_q^\ell$.

end for

▷ *Phase 4: Decoding all hypotheses and constructing the list*

$\mathcal{L} \leftarrow \varnothing$

for $\rho = (\rho_0, \rho_1, \ldots, \rho_\ell) \in \mathbb{F}_q^{\ell+1}$ **do**

  Define $\rho' := (\rho_1, \ldots, \rho_\ell)$

  for $i = 1, 2, \ldots, n$ **do**

    for $b \in \mathbb{F}_q$ **do**

      Compute $V_i(\rho, b) \leftarrow \frac{(q-1)M}{q} + \frac{1}{q} \sum_{s \in \mathbb{F}_q^*} \sum_{c \in \mathbb{F}_q^*} \chi\big(-c(\rho_0 + sb)\big) \widehat{A}_{i,s,c}(c\rho')$.

    end for

    Let $z_i(\rho)$ be the b such that $V_i(\rho, b)$ achieves the maximum over all $b \in \mathbb{F}_q$

  end for

  Let $\mathbf{z}(\rho) = (z_1(\rho), \ldots, z_n(\rho))$

  if $H(\mathbf{z}(\rho)) = \mathbf{y}$ **then**

    $\mathcal{L} \leftarrow \mathcal{L} \cup \{\mathbf{z}(\rho)\}$

  end if

end for

return $\mathcal{L}$

Success probability. For the rest of the proof, denote $\Delta = \Delta_\epsilon, \ell = \ell_\epsilon$, and $M = M_\epsilon$. We first verify that the Fourier-computed vote counts are equal to the direct

vote counts

$$\widetilde{V}_i(\rho, b) := \sum_{s\in\mathbb{F}_q^*} \sum_{\alpha\in\mathbb{F}_q^\ell} \mathbf{1}\left[Y_{i,s}(\alpha) = \rho_0 + \alpha\cdot\rho' + sb\right].$$

Algorithm 2 does not compute these vote counts by summing over α separately for every ρ. Instead, it computes the same quantities by a q-ary Fourier transform. Indeed, by the orthogonality relation $\mathbf{1}[z=0] = \frac{1}{q}\sum_{c\in\mathbb{F}_q}\chi(cz)$, we have

$$\widetilde{V}_i(\rho, b) = \frac{(q-1)M}{q} + \frac{1}{q}\sum_{s\in\mathbb{F}_q^*}\sum_{c\in\mathbb{F}_q^*} \chi\big(-c(\rho_0+sb)\big)\widehat{A}_{i,s,c}(c\rho'),$$

and the value $V_i(\rho, b)$ computed by Algorithm 2 is exactly $\widetilde{V}_i(\rho, b)$. The probability analysis below therefore works with the direct vote count $\widetilde{V}_i$, while the q-ary Fourier transform only serves to compute it efficiently.

Let $\rho^* := (R_0(\mathbf{x}), R_1(\mathbf{x}), \ldots, R_\ell(\mathbf{x}))$ be the correct hypothesis. We show that the candidate $\mathbf{z}(\rho^*)$ equals $\mathbf{x}$ with probability at least $1/2$.

For each $(i, b) \in [n]\times\mathbb{F}_q$, define $X_{i,\alpha}(b) := \frac{1}{q-1}\sum_{s\in\mathbb{F}_q^*}\mathbf{1}\left[Y_{i,s}(\alpha) = U_\alpha(\mathbf{x}) + sb\right]$ and $\overline{X}_i(b) := \frac{1}{M}\sum_{\alpha\in\mathbb{F}_q^\ell} X_{i,\alpha}(b)$. Under the correct hypothesis ρ^*, the direct vote count satisfies $\widetilde{V}_i(\rho^*, b) = (q-1)M\overline{X}_i(b)$ by the definition of $X_{i,\alpha}(b)$. Since $V_i = \widetilde{V}_i$, maximizing $V_i(\rho^*, b)$ is equivalent to maximizing $\overline{X}_i(b)$.

For brevity, all probabilities $\Pr[\cdot]$ and expectations $\mathbb{E}[\cdot]$ are taken over the uniform choice of the random linear forms (e.g., R and U_α) as well as the internal randomness of the predictor $\mathcal{P}$, unless explicitly stated otherwise. Let $p_0 := \Pr\big[\mathcal{P}(H, \mathbf{y}, R) = R(\mathbf{x})\big]$. By assumption, $p_0 \geq \frac{1}{q} + \epsilon$.

First consider the correct value $b = x_i$. For every fixed $\alpha \in \mathbb{F}_q^\ell$ and $s \in \mathbb{F}_q^*$, since R_0 is a uniformly random, the form $U_\alpha + sL_i = R_0 + \sum_{j=1}^{\ell}\alpha_j R_j + sL_i$ is uniformly distributed over all linear forms. This gives $\mathbb{E}[X_{i,\alpha}(x_i)] = p_0$.

Now we consider a wrong value $b \neq x_i$. Define the prediction error $E(R) := \mathcal{P}(H, \mathbf{y}, R) - R(\mathbf{x})$. The condition $Y_{i,s}(\alpha) = U_\alpha(\mathbf{x}) + sb$ is equivalent to $E(U_\alpha + sL_i) = s(b - x_i)$. Since $b - x_i \neq 0$, as s ranges over $\mathbb{F}_q^*$, the value $s(b-x_i)$ also ranges over $\mathbb{F}_q^*$. Replacing $U_\alpha + sL_i$ with the uniform form R as before, we have

$$\mathbb{E}[X_{i,\alpha}(b)] = \frac{1}{q-1}\sum_{s\in\mathbb{F}_q^*}\Pr[E(R) = s(b-x_i)] = \frac{1}{q-1}\sum_{e\in\mathbb{F}_q^*}\Pr[E(R) = e] = \frac{1-p_0}{q-1}.$$

It follows that, for every $b \neq x_i$,

$$\mathbb{E}[X_{i,\alpha}(x_i)] - \mathbb{E}[X_{i,\alpha}(b)] = p_0 - \frac{1-p_0}{q-1} = \frac{qp_0 - 1}{q-1} \geq \frac{q\epsilon}{q-1} = \Delta.$$

We now calculate the probability of the event that $\mathbf{x} \in \mathcal{L}$. We first claim that, for fixed i and b, the random variables $\{X_{i,\alpha}(b)\}_{\alpha\in\mathbb{F}_q^\ell}$ are pairwise independent. Let $\mathcal{V} := \mathrm{Hom}(\mathbb{F}_q^n, \mathbb{F}_q)$ be the vector space of linear forms. For any distinct $\alpha, \beta \in \mathbb{F}_q^\ell$, define a linear map $T_{\alpha,\beta} : \mathcal{V}^{\ell+1} \to \mathcal{V}^2$ by: $T_{\alpha,\beta}(v_0, v_1, \ldots, v_\ell) :=$

$\left(v_0 + \sum_{j=1}^{\ell} \alpha_j v_j,\ v_0 + \sum_{j=1}^{\ell} \beta_j v_j\right)$. The coefficient matrix of $T_{\alpha,\beta}$ is given by $\begin{pmatrix} 1 & \alpha_1 & \cdots & \alpha_\ell \\ 1 & \beta_1 & \cdots & \beta_\ell \end{pmatrix}$. Since $\alpha \neq \beta$, the two rows are linearly independent over $\mathbb{F}_q$. This map is therefore surjective onto $\mathcal{V}^2$, and every pair in $\mathcal{V}^2$ has the same number of preimages. Thus the pair (U_α, U_β) is uniformly distributed over $\mathcal{V}^2$, i.e., U_α and U_β are independent uniform linear forms.

For fixed i, the two query families $\{U_\alpha + sL_i : s \in \mathbb{F}_q^*\}$ and $\{U_\beta + sL_i : s \in \mathbb{F}_q^*\}$ are independent. Fresh randomness in each call $\mathcal{P}$ makes the corresponding response families $\{Y_{i,s}(\alpha)\}_{s\in\mathbb{F}_q^*}$ and $\{Y_{i,s}(\beta)\}_{s\in\mathbb{F}_q^*}$ independent. Since $X_{i,\alpha}(b)$ and $X_{i,\beta}(b)$ are functions of these two independent response families, they are independent.

For each $\alpha \in \mathbb{F}_q^\ell$, since the random variable $X_{i,\alpha}(b)$ lies in the interval $[0,1]$, we get $\mathrm{Var}(X_{i,\alpha}(b)) \leq \frac{1}{4}$. Pairwise independence gives $\mathrm{Var}(\overline{X}_i(b)) = \frac{1}{M^2}\sum_{\alpha\in\mathbb{F}_q^\ell} \mathrm{Var}(X_{i,\alpha}(b)) \leq \frac{1}{4M}$. By Chebyshev's inequality, we obtain that

$$\Pr\left[\left|\overline{X}_i(b) - \mathbb{E}[\overline{X}_i(b)]\right| \geq \frac{\Delta}{3}\right] \leq \frac{9}{4M\Delta^2}.$$

Taking a union bound over all $b \in \mathbb{F}_q$, the probability that there exists some b for which $\left|\overline{X}_i(b) - \mathbb{E}[\overline{X}_i(b)]\right| \geq \frac{\Delta}{3}$ is at most $\frac{9q}{4M\Delta^2}$.

On the complementary event, for every $b \in \mathbb{F}_q$, $\left|\overline{X}_i(b) - \mathbb{E}[\overline{X}_i(b)]\right| < \frac{\Delta}{3}$. In particular, for every wrong value $b \neq x_i$, we have

$$\overline{X}_i(x_i) - \overline{X}_i(b) > \mathbb{E}[\overline{X}_i(x_i)] - \mathbb{E}[\overline{X}_i(b)] - \frac{2\Delta}{3} \geq \Delta - \frac{2\Delta}{3} = \frac{\Delta}{3} > 0.$$

Hence $V_i(\rho^*, x_i) = (q-1)M\overline{X}_i(x_i) > (q-1)M\overline{X}_i(b) = V_i(\rho^*, b)$, and $z_i(\rho^*) = x_i$.

Taking a further union bound over all n coordinates, the probability that under the correct hypothesis ρ^*, there exists at least one coordinate i such that $z_i(\rho^*) \neq x_i$ is at most $\frac{9nq}{4M\Delta^2}$. By the choice of $M = M_\epsilon$, we have $M \geq \frac{9nq}{2\Delta^2}$ and $\frac{9nq}{4M\Delta^2} \leq \frac{1}{2}$, so $\Pr[\mathbf{z}(\rho^*) = \mathbf{x}] \geq \frac{1}{2}$.

The algorithm $\mathcal{B}^{\mathcal{P}}$ enumerates all $\rho \in \mathbb{F}_q^{\ell+1}$, including ρ^*. If $\mathbf{z}(\rho^*) = \mathbf{x}$, then $H(\mathbf{z}(\rho^*)) = H(\mathbf{x}) = \mathbf{y}$, so $\mathbf{x}$ is inserted into the output list $\mathcal{L}$. Hence $\Pr[\mathbf{x} \in \mathcal{L}] \geq \frac{1}{2}$. The list contains at most one candidate for each hypothesis ρ, so $|\mathcal{L}| \leq q^{\ell+1} = qM = qM_\epsilon$.

Running time and oracle calls. All calls to the predictor $\mathcal{P}$ are made in Algorithm 2. The indices of these calls are $(i, s, \alpha) \in [n] \times \mathbb{F}_q^* \times \mathbb{F}_q^\ell$, so the number of calls to $\mathcal{P}$ is $N_{\mathcal{P}} = n(q-1)q^\ell = n(q-1)M$.

The remaining operations are post-processing operations. First, for each $(i, s, c) \in [n] \times \mathbb{F}_q^* \times \mathbb{F}_q^*$, the algorithm forms the array $A_{i,s,c}(\alpha) = \chi(cY_{i,s}(\alpha)), \alpha \in \mathbb{F}_q^\ell$, which requires $n(q-1)^2M$ character evaluations.

Next, for each fixed (i, s, c), the algorithm applies one fast q-ary Fourier transform of length $M = q^\ell$. Using the fast q-ary Fourier transform, one such transform can be computed using at most ℓqM scalar multiplications and $\ell(q-1)M$ scalar additions. Since there are $n(q-1)^2$ such transforms, the total Fourier

cost is at most $n(q-1)^2\ell qM$ scalar multiplications and $n(q-1)^3\ell M$ scalar additions.

The vote computation in Phase 4 considers $\rho \in \mathbb{F}_q^{\ell+1}, i \in [n], b \in \mathbb{F}_q$. For each triple (ρ, i, b), it sums over $(s, c) \in \mathbb{F}_q^* \times \mathbb{F}_q^*$. Thus the number of summands used in the vote computation is $q^{\ell+1} \cdot n \cdot q \cdot (q-1)^2 = nq^2(q-1)^2M$. The algorithm verifies at most one candidate for each hypothesis ρ, so it performs at most $q^{\ell+1} = qM$ evaluations of H.

Finally, if linear forms are represented by their coefficient vectors, then sampling $R_0, \ldots, R_\ell$, forming all U_α, and representing the queries $U_\alpha + sL_i$ costs at most $T_{\mathsf{form}} = M \cdot \mathrm{poly}(n, q, \ell)$ field operations.

Let $T_{\mathcal{P}}$ be the time for one call to $\mathcal{P}$, T_H be the time for one evaluation of H, T_χ be the cost of one character evaluation, $T_\times$ be the cost of one scalar multiplication, T_+ be the cost of one scalar addition, and T_{sum} be the cost of processing one scalar summand in the final vote computation. Then

$$\begin{aligned} T_{\mathcal{B}} &\le n(q-1)M\,T_{\mathcal{P}} + qM\,T_H + T_{\mathsf{form}} + n(q-1)^2M\,T_\chi \\ &\quad + n(q-1)^2\ell qM\,T_\times + n(q-1)^3\ell M\,T_+ + nq^2(q-1)^2M\,T_{\mathrm{sum}}. \\ &= n(q-1)M\,T_{\mathcal{P}} + qM\,T_H + M \cdot \mathrm{poly}(n, q, \ell). \end{aligned}$$

Since $M = M_\epsilon$ and $\ell = \ell_\epsilon$, the claimed bound follows. □

Part 3: From an Average Linear Form Predictor to an Inverter

This step is the q-ary analogue of the averaging step in Theorem 3 of [6]. The averaging argument is the same, but we utilize Lemma 3 instead of the binary Goldreich–Levin reconstruction lemma. To get the security bound, we include the proof to track the constants and the concrete running time.

Proposition 1 (Average-Case Predictor-to-Inverter Reduction). *Suppose that a probabilistic predictor $\mathcal{P}$ which on input $(k, f_k(\mathbf{x}), R)$, where $k, \mathbf{x}$ are uniformly random, and $R : \mathbb{F}_q^n \to \mathbb{F}_q$ is a uniformly random linear form, predicts $R(\mathbf{x})$ with probability at least $1/q + \epsilon$, and the probability is over the randomness of $k, \mathbf{x}, R$, and the internal randomness of $\mathcal{P}$.*

Denote $\epsilon_{\mathsf{fp}} = \epsilon/2, \Delta_{\mathsf{fp}} = q\epsilon_{\mathsf{fp}}/(q-1) = q\epsilon/(2(q-1))$, $\ell_{\mathsf{fp}} = \lceil \log_q(9nq/(2\Delta_{\mathsf{fp}}^2)) \rceil = \left\lceil \log_q\left(18n(q-1)^2/(q\epsilon^2)\right) \right\rceil, M_{\mathsf{fp}} := q^{\ell_{\mathsf{fp}}}$. Then there exists an inverter $\mathcal{C}^{\mathcal{P}}$ such that, on input a one-wayness challenge $(k, \mathbf{y}), \mathbf{y} = f_k(\mathbf{x})$, it outputs a vector $\mathbf{x}' \in \mathbb{F}_q^n$ satisfying $f_k(\mathbf{x}') = \mathbf{y}$ with probability at least $q\epsilon/(4(q-1))$. Moreover, $\mathcal{C}^{\mathcal{P}}$ makes $n(q-1)M_{\mathsf{fp}}$ calls to $\mathcal{P}$. If one call to $\mathcal{P}$ takes time $T_{\mathcal{P}}$, and one evaluation of f_k takes time T_{f_k}, then the running time of $\mathcal{C}^{\mathcal{P}}$ is bounded by $T_{\mathcal{C}} \le n(q-1)M_{\mathsf{fp}}T_{\mathcal{P}} + qM_{\mathsf{fp}}T_{f_k} + M_{\mathsf{fp}} \cdot \mathrm{poly}(n, q, \ell_{\mathsf{fp}})$.

Proof. We define $\mathcal{C}^{\mathcal{P}}$ in Algorithm 3. It queries $\mathcal{B}^{\mathcal{P}}$ constructed in Lemma 3 for the given challenge, chooses any element in the returned list, and outputs it.

For every pair $(k, \mathbf{x})$, define the success probability of $\mathcal{P}$ for $(k, \mathbf{x})$ as $p_{k,\mathbf{x}} = \Pr_R\left[\mathcal{P}(k, f_k(\mathbf{x}), R) = R(\mathbf{x})\right]$, where the probability is over the random linear

Algorithm 3 The Average-Case Inverter $\mathcal{C}^{\mathcal{P}}$

Require: A one-wayness challenge $(k, \mathbf{y})$, where $\mathbf{y} = f_k(\mathbf{x})$ for a uniformly random $\mathbf{x} \in \mathbb{F}_q^n$; oracle access to $\mathcal{P}$; an advantage parameter $\epsilon > 0$
Ensure: A preimage $\mathbf{x}' \in \mathbb{F}_q^n$ satisfying $f_k(\mathbf{x}') = \mathbf{y}$, or $\perp$
1: Set $\epsilon_{\mathsf{fp}} \leftarrow \epsilon/2$
2: Query $\mathcal{B}^{\mathcal{P}}$ constructed in Lemma 3 on $(f_k, \mathbf{y}, \epsilon_{\mathsf{fp}})$
3: Let $\mathcal{L} \leftarrow \mathcal{B}^{\mathcal{P}}(f_k, \mathbf{y}, \epsilon_{\mathsf{fp}})$
4: **if** $\mathcal{L} = \varnothing$ **then**
5: **return** $\perp$
6: **else**
7: Choose any $\mathbf{z} \in \mathcal{L}$
8: **return** $\mathbf{z}$
9: **end if**

form R and the internal randomness of $\mathcal{P}$. By assumption, we have $\mathbb{E}_{k,\mathbf{x}}[p_{k,\mathbf{x}}] \geq 1/q + \epsilon$.

Let $\mathsf{Good} = \{(k, \mathbf{x}) : p_{k,\mathbf{x}} \geq 1/q + \epsilon/2\}$, and denote $\gamma := \Pr_{k,\mathbf{x}}[(k, \mathbf{x}) \in \mathsf{Good}]$. Since $p_{k,\mathbf{x}} \leq 1$, we have $1/q + \epsilon \leq (1-\gamma)(1/q + \epsilon/2) + \gamma$. This gives

$$\gamma \geq \frac{\epsilon/2}{1 - 1/q - \epsilon/2} \geq \frac{\epsilon/2}{1 - 1/q} = \frac{q\epsilon}{2(q-1)}.$$

For every good pair $(k, \mathbf{x})$, Lemma 3 applies to the fixed function $(H, \mathbf{y}) = (f_k, f_k(\mathbf{x}))$ with advantage parameter $\epsilon_{\mathsf{fp}} := \epsilon/2$, so the reconstruction list contains $\mathbf{x}$ with probability at least $1/2$. Since each $\mathbf{z} \in \mathcal{L}$ satisfies $f_k(\mathbf{z}) = \mathbf{y}$, $\mathcal{C}^{\mathcal{P}}$ outputs a valid preimage on every good pair $(k, \mathbf{x})$ with probability at least $1/2$. Hence

$$\Pr\left[f_k(\mathcal{C}^{\mathcal{P}}(k, f_k(\mathbf{x}))) = f_k(\mathbf{x})\right] \geq \frac{\gamma}{2} \geq \frac{q\epsilon}{4(q-1)}.$$

The algorithm $\mathcal{C}^{\mathcal{P}}$ calls $\mathcal{B}^{\mathcal{P}}$ in Lemma 3 once with advantage parameter $\epsilon_{\mathsf{fp}} = \epsilon/2$, giving $\Delta_{\mathsf{fp}} = q\epsilon_{\mathsf{fp}}/(q-1) = q\epsilon/(2(q-1))$, and $M_{\mathsf{fp}} = q^{\lceil \log_q(9nq/(2\Delta_{\mathsf{fp}}^2)) \rceil} = q^{\lceil \log_q(18n(q-1)^2/(q\epsilon^2)) \rceil}$. By Lemma 3, the number of calls to $\mathcal{P}$ is exactly $n(q-1)M_{\mathsf{fp}}$, and the running time is at most $n(q-1)M_{\mathsf{fp}}T_{\mathcal{P}} + qM_{\mathsf{fp}}T_{f_k} + M_{\mathsf{fp}} \cdot \mathrm{poly}(n, q, \ell_{\mathsf{fp}})$. This proves the proposition. □

Part 4: Reduction from Undetectability to One-Wayness for $\mathcal{Q}_{n,m,q}$

We combine the three reduction steps. To state the running time explicitly, let T_{mask} denote an upper bound for one execution of the masking operation used in Lemma 2, excluding the call to the distinguisher $\mathcal{A}$. Let $T_{\mathsf{rec}}(n, m, q, M)$ denote an upper bound for the running time of Algorithms 2, excluding the oracle calls to the predictor $\mathcal{P}$. This term includes the sampling of the random linear forms, the q-ary Fourier transforms, the vote computations, and the verification of at most qM candidates.

For a distinguisher running in time t, define

$$T_{\mathsf{red}}(t, M) := n(q-1)M\big(t + T_{\mathsf{mask}}\big) + T_{\mathsf{rec}}(n, m, q, M).$$

Theorem 4 (Tighter Reduction from UD to OW for $\mathcal{Q}_{n,m,q}$). *Let $0 < \epsilon \le 1$, and $M_\epsilon = q^{\lceil \log_q(72\,nq(q-1)^2/(\epsilon^2))\rceil}$. If there exists a UD distinguisher $\mathcal{A}$ against $\mathcal{Q}_{n,m,q}$ running in time at most t and satisfying $\mathrm{Adv}^{\mathsf{UD}}_{\mathcal{Q}_{n,m,q}}(\mathcal{A}) \triangleq \delta \ge \epsilon$, then there exists an OW inverter $\mathcal{C}$ against $\mathcal{Q}_{n,m,q}$ running in time at most $t' = \frac{72\,n^2q^2(q-1)^3}{\epsilon^2}t + \frac{1}{\epsilon^2}\,\mathrm{poly}(n, m, q, \log\frac{1}{\epsilon})$ with success probability $\delta' \ge \delta/(8(q-1)) \ge \epsilon/(8(q-1))$.*

Proof. By Lemma 2, the distinguisher $\mathcal{A}$ yields a predictor $\mathcal{P}^{\mathcal{A}}$ for random linear forms with success probability at least $1/q + \alpha$, where $\alpha = \delta/(2q)$.

For prediction advantage α, Proposition 1 requires $M \ge \frac{18n(q-1)^2}{q\alpha^2} = \frac{72\,nq(q-1)^2}{\delta^2}$. The choice of M_ϵ meets this required lower bound, so we can construct an inverter $\mathcal{C}$ with success probability at least $\frac{q\alpha}{4(q-1)} = \frac{q}{4(q-1)} \cdot \frac{\delta}{2q} = \frac{\delta}{8(q-1)} \ge \frac{\epsilon}{8(q-1)}$, and its running time is bounded by $T_{\mathcal{C}} \le T_{\mathsf{red}}(t, M_\epsilon)$.

Since $M_\epsilon \le \frac{72\,nq^2(q-1)^2}{\epsilon^2}$, the dominant term in $T_{\mathsf{red}}(t, M_\epsilon)$ is $n(q-1)M_\epsilon t \le \frac{72\,n^2q^2(q-1)^3}{\epsilon^2}t$. From Lemma 2 and Lemma 3 we know that the remaining term is $\frac{1}{\epsilon^2}\,\mathrm{poly}(n, m, q, \ell_\epsilon)$, and it is bounded by $\frac{1}{\epsilon^2}\,\mathrm{poly}(n, m, q, \log\frac{1}{\epsilon})$ since $\ell_\epsilon = \left\lceil \log_q\left(72\,nq(q-1)^2/(\epsilon^2)\right)\right\rceil$. □

Remark 2 (Security Loss). Following the standard formalization of security loss (e.g., as explicitly defined by Abe et al. [1]), the security loss factor L is calculated as follows:

$$L = \frac{t'/\delta'}{t/\delta} \le \frac{8(q-1)t'}{t} = O\left(\epsilon^{-2}\right).$$

This gives an $O(\epsilon^{-2})$ loss up to polynomial factors in n, m, q, improving the $O(\epsilon^{-4})$ dependence in previous reductions for general q [10].

Theorem 5 (Security Bound of Undetectability). *Under the (n, m, q)-MQOW assumption, the function family $\mathcal{Q}_{n,m,q}$ is undetectable.*

More concretely, for every $0 < \epsilon \le 1$ and every time bound t, we have

$$\mathrm{InSec}^{\mathsf{UD}}\big(\mathcal{Q}_{n,m,q}; t\big) \le \max\left\{\epsilon,\ 8(q-1)\ \mathrm{InSec}^{\mathsf{OW}}\big(\mathcal{Q}_{n,m,q}; t_\epsilon\big)\right\},$$

where $t_\epsilon = \frac{72\,n^2q^2(q-1)^3}{\epsilon^2}t + \frac{1}{\epsilon^2}\,\mathrm{poly}\left(n, m, q, \log\frac{1}{\epsilon}\right)$. Consequently,

$$\mathrm{InSec}^{\mathsf{UD}}\big(\mathcal{Q}_{n,m,q}; t\big) \le \inf_{0<\epsilon\le 1} \max\left\{\epsilon,\ 8(q-1)\ \mathrm{InSec}^{\mathsf{OW}}\big(\mathcal{Q}_{n,m,q}; t_\epsilon\big)\right\}.$$

Proof. Fix $0 < \epsilon \le 1$ and denote $D = \mathrm{InSec}^{\mathsf{UD}}\big(\mathcal{Q}_{n,m,q}; t\big)$. If $D \le \epsilon$, the claimed bound is immediate. Otherwise, for every $\rho \in (\epsilon, D)$, there exists a UD distinguisher $\mathcal{A}_\rho$ running in time at most t with $\mathrm{Adv}^{\mathsf{UD}}_{\mathcal{Q}_{n,m,q}}(\mathcal{A}_\rho) \ge \rho$. By Theorem 4, this distinguisher yields an OW inverter running in time at most t_ϵ and succeeding with probability at least $\rho/(8(q-1))$. Hence $\mathrm{InSec}^{\mathsf{OW}}\big(\mathcal{Q}_{n,m,q}; t_\epsilon\big) \ge$

$\rho/(8(q-1))$. Since this holds for every $\rho \in (\epsilon, D)$, letting ρ approach D gives $D \leq 8(q-1)\,\mathrm{InSec}^{\mathsf{OW}}\big(\mathcal{Q}_{n,m,q}; t_\epsilon\big)$. Combining this case with the case $D \leq \epsilon$ proves the first inequality. Taking the infimum over all $0 < \epsilon \leq 1$ gives the second inequality. □

4 The MQ-Based W-OTS$^+$ Scheme

This section presents the MQ-based W-OTS$^+$ scheme, along with its provable security and implementation aspects.

4.1 Construction

In this work, we use the MQ function family $\mathcal{Q}_{n,m,q}$ as the underlying function family $\mathcal{F}_\lambda$ in the W-OTS$^+$ framework. We restrict to binary extension fields $q = 2^s$ and use the square setting $m = n$, so that each $f_k \in \mathcal{Q}_{n,n,q}$ defines a length-preserving map $f_k : \mathbb{F}_q^n \to \mathbb{F}_q^n$. We set $\lambda = ns = n \log q$ and fix a bijection between $\mathbb{F}_q^n$ and $\{0,1\}^\lambda$, so each $f_k : \mathbb{F}_q^n \to \mathbb{F}_q^n$ is regarded as a length-preserving bit-string function $f_k : \{0,1\}^\lambda \longrightarrow \{0,1\}^\lambda$.

Definition 8 (MQ-W-OTS$^+$). *The MQ-based W-OTS$^+$ scheme (MQ-W-OTS$^+$) is the W-OTS$^+$ signature scheme in which the underlying length-preserving function family $\mathcal{F}_\lambda$ is instantiated by the MQ function family $\mathcal{Q}_{n,n,q}$ under the fixed bit-string representation described above.*

4.2 Provable Security

Now we establish the EU-CMA security theorem of the MQ-W-OTS$^+$ scheme.

Theorem 6 (Security of MQ-W-OTS$^+$ under the MQOW Assumption). *Let $w, l_{\mathrm{msg}} = \mathrm{poly}(\lambda)$, and. Let T_{eval} denote the time required to evaluate one function from $\mathcal{Q}_{n,n,q}$.*

Under the (n,n,q)-MQOW assumption, the MQ-W-OTS$^+$ scheme is existentially unforgeable under a one-time chosen-message attack. Specifically,

$$\begin{aligned}\mathrm{InSec}^{\mathsf{EU\text{-}CMA}}\big(\text{MQ-W-OTS}^+(1^\lambda, w, l_{\mathrm{msg}}); t, q_s = 1\big) \\ < lw\Bigg(\mathrm{InSec}^{\mathsf{OW}}\big(\mathcal{Q}_{n,n,q}; t_{\mathsf{base}}\big) + 2wq\,\mathrm{InSec}^{\mathsf{OW}}\big(\mathcal{Q}_{n,n,q}; t_{\mathsf{spr}}\big) \\ + w \inf_{0<\epsilon\leq 1} \max\Big\{\epsilon, 8(q-1)\,\mathrm{InSec}^{\mathsf{OW}}\big(\mathcal{Q}_{n,n,q}; t_\epsilon\big)\Big\}\Bigg).\end{aligned}$$

Here, $t_{\mathsf{base}} = t + (3lw + w - 2)T_{\mathsf{eval}}$, $t_{\mathsf{spr}} = t_{\mathsf{base}} + \mathrm{poly}(n, \log q)$. and for every $0 < \epsilon \leq 1$, $t_\epsilon = \frac{72n^2q^2(q-1)^3}{\epsilon^2} t_{\mathsf{base}} + \frac{1}{\epsilon^2}\,\mathrm{poly}\big(n, q, \log(1/\epsilon)\big)$.

Proof. Let $\mathcal{A}$ be an EU-CMA adversary against MQ-W-OTS$^+$ that runs in time t and makes at most one signing query. We apply the generic security theorem for W-OTS$^+$ with the underlying function family $\mathcal{F}_\lambda = \mathcal{Q}_{n,n,q}$.

Since one such evaluation takes time T_{eval}, the total running time of the derived adversary in Theorem 1 is at most $t_{\mathsf{base}} = t+(3lw+w-2)T_{\mathsf{eval}}$. Therefore, Theorem 1 implies

$$\begin{aligned}\mathrm{InSec}^{\mathsf{EU\text{-}CMA}}\big(\text{MQ-W-OTS}^+(1^\lambda, w, l_{\mathrm{msg}}); t, q_s = 1\big) \\ < lw\Big(\mathrm{InSec}^{\mathsf{OW}}(\mathcal{Q}_{n,n,q}; t_{\mathsf{base}}) + w\,\mathrm{InSec}^{\mathsf{SPR}}(\mathcal{Q}_{n,n,q}; t_{\mathsf{base}}) \\ + w\,\mathrm{InSec}^{\mathsf{UD}}(\mathcal{Q}_{n,n,q}; t_{\mathsf{base}})\Big).\end{aligned}$$

By Theorem 3, $\mathrm{InSec}^{\mathsf{SPR}}(\mathcal{Q}_{n,n,q}; t_{\mathsf{base}}) \le 2q\,\mathrm{InSec}^{\mathsf{OW}}(\mathcal{Q}_{n,n,q}; t_{\mathsf{spr}})$, where $t_{\mathsf{spr}} = t_{\mathsf{base}} + \mathrm{poly}(n, \log q)$. By Theorem 5, $\mathrm{InSec}^{\mathsf{UD}}(\mathcal{Q}_{n,n,q}; t_{\mathsf{base}}) < \inf_{0<\epsilon\le 1} \max\{\epsilon, 8(q-1)\,\mathrm{InSec}^{\mathsf{OW}}\big(\mathcal{Q}_{n,n,q}; t_\epsilon\big)\}$, where $t_\epsilon = \frac{72n^2q^2(q-1)^3}{\epsilon^2}t_{\mathsf{base}} + \frac{1}{\epsilon^2}\,\mathrm{poly}(n, q, \log(1/\epsilon))$. Substituting them into the generic W-OTS$^+$ inequality gives the claimed security bound. □

Advantages of MQ-W-OTS$^+$. MQ-W-OTS$^+$ has two main theoretical advantages:

- *Standard-Model Security without Trapdoors:* MQ-W-OTS$^+$ is trapdoor-free and provably secure in the standard model. The absence of a hidden trapdoor removes the trapdoor-specific attack surface that has affected many multivariate cryptosystems; in our construction, successful forgery is directly reduced to breaking the underlying multivariate assumptions.
- *Explicit and Consolidated Security Foundation:* Conventional W-OTS$^+$ instantiations based on symmetric primitives require three independent black-box assumptions on the underlying function family (OW, SPR, and UD). In contrast, MQ-W-OTS$^+$ consolidates these three assumptions into one standard multivariate hardness assumptions (MQOW).

4.3 Concrete Complexity Analysis

To instantiate the MQ-W-OTS$^+$ scheme with concrete parameters, we must evaluate the explicit security bound derived in Theorem 6. Let $T_{\mathsf{MQ}}(n, m, q)$ denote the computational complexity (e.g., measured in bit operations or gate evaluations, which can be calculated using `MQ-estimator` [4]) required to solve a random (n, m, q)-MQOW instance.

For an adversary operating under a practical time budget $t_{\mathsf{base}} \ll T_{\mathsf{MQ}}(n, n, q)$, we approximate the success probabilities as follows:

$$\mathrm{InSec}^{\mathsf{OW}}(\mathcal{Q}_{n,n,q}; t_{\mathsf{base}}) \approx \frac{t_{\mathsf{base}}}{T_{\mathsf{MQ}}(n, n, q)}. \tag{4}$$

For the last term of the EU-CMA bound in Theorem 6, we choose the parameter ϵ by balancing the two terms inside the maximum. Ignoring polynomial terms

in t_ϵ, we obtain

$$\epsilon_\star = 8(q-1) \cdot \frac{\frac{72\, n^2 q^2 (q-1)^3}{\epsilon_\star^2} t_{\mathsf{base}}}{T_{\mathsf{MQ}}(n,n,q)} \quad \Longrightarrow \quad \epsilon_\star = \sqrt[3]{\frac{576\, n^2 q^2 (q-1)^4 t_{\mathsf{base}}}{T_{\mathsf{MQ}}(n,n,q)}}.$$

Thus

$$\inf_{0<\epsilon\leq 1} \max\left\{\epsilon,\ 8(q-1)\ \mathrm{InSec}^{\mathsf{OW}}\big(\mathcal{Q}_{n,m,q}; t_\epsilon\big)\right\} \approx \sqrt[3]{\frac{576\, n^2 q^2 (q-1)^4 t_{\mathsf{base}}}{T_{\mathsf{MQ}}(n,n,q)}}.$$

Let ϵ_{sig} denote the total forgery advantage. Substituting these terms back into the EU-CMA bound gives:

$$\epsilon_{\mathsf{sig}} \approx lw\left(\frac{t_{\mathsf{base}}}{T_{\mathsf{MQ}}(n,n,q)} + 2wq\frac{t_{\mathsf{base}}}{T_{\mathsf{MQ}}(n,n,q)} + w\sqrt[3]{\frac{576\, n^2 q^2 (q-1)^4 t_{\mathsf{base}}}{T_{\mathsf{MQ}}(n,n,q)}}\right). \quad (5)$$

Asymptotic Dominance. For practical security levels, the cubic root of this ratio is orders of magnitude larger than the linear terms, and the total forgery advantage is dominated by the cubic-root term:

$$\epsilon_{\mathsf{sig}} \approx lw^2 \sqrt[3]{\frac{576\, n^2 q^2 (q-1)^4 t_{\mathsf{base}}}{T_{\mathsf{MQ}}(n,n,q)}}. \quad (6)$$

Security Level. Following standard practices for evaluating the concrete security of signature schemes (e.g., RFC 8391 [29]), we define the effective bit-security level b such that an adversary expending time $t = 2^{b-1}$ achieves a maximum forgery advantage of $\epsilon_{\mathsf{sig}} = 1/2$. This yields $b = \log t + 1$. We assume that the additive time overhead introduced by the reduction is negligible compared to the running time t of any practical attack. Therefore, we have $t \approx t_{\mathsf{base}}$.

Equating our asymptotically dominant bound to $1/2$ and cubing both sides, we isolate the time boundary t:

$$\frac{1}{8} \approx l^3 w^6 \cdot \frac{576 n^2 q^2 (q-1)^4 t}{T_{\mathsf{MQ}}(n,n,q)} \quad \Longrightarrow \quad t \approx \frac{T_{\mathsf{MQ}}(n,n,q)}{4608\, l^3 w^6 n^2 q^2 (q-1)^4}. \quad (7)$$

Applying the security level definition, we explicitly derive the concrete bit-security for the MQ-W-OTS$^+$:

$$\begin{aligned} b &\approx \log t + 1 \\ &\approx \log\big(T_{\mathsf{MQ}}(n,n,q)\big) - \log\big(l^3 w^6 n^2 q^2 (q-1)^4\big) - 8 - 2\log 3. \end{aligned} \quad (8)$$

Equation (8) shows that the effective bit-security of the scheme equals the hardness of the underlying MQ instance, minus a logarithmic term representing the reduction loss.

4.4 Proof-of-Concept Implementation

To demonstrate the feasibility of MQ-W-OTS$^+$, we provide a proof-of-concept (PoC) C implementation.

As described in Section 2.2, the original construction stores the uniformly sampled MQ coefficient tuple k explicitly in the public key, resulting in a large public-key size. We implement both this original version and a seed-compressed version, in which the public key stores a 128-bit public seed in place of k. The coefficient tuple k is generated from this seed during KeyGen and regenerated during Verify. The original version matches the formulation used in our security proof, whereas the seed-compressed version substantially reduces the public-key size at the cost of regenerating the MQ coefficients during verification.

We fix the message length to be $l_{\text{msg}} = 256$ bits. When evaluated on an x86-64 server platform equipped with an Intel Xeon Silver 4214R processor (Cascade Lake microarchitecture) running at 2.40 GHz, using AVX2 intrinsics, the two versions have comparable computational performance across all evaluated parameter sets. For example, for the parameter set $(q = 2, n = 248, w = 16)$, which has an estimated classical bit security of $b = 143.7$ corresponding to NIST Security Level 1, the original version has a public key of 967 417 bytes and KeyGen/Sign/Verify costs of 206×10^6, 99×10^6, and 101×10^6 CPU cycles, respectively. The seed-compressed version reduces the public key by a factor of about 378×, down to 2 558 bytes, with KeyGen/Sign/Verify costs of 210×10^6, 98×10^6, and 111×10^6 CPU cycles, respectively. For both versions, the secret-key size and signature size are both 2 077 bytes.

In the following, we report results for the seed-compressed version. Among the parameter sets evaluated in our PoC, $(q = 2, n = 256, w = 16)$ and $(q = 16, n = 96, w = 16)$ represent two attractive time-space-security trade-offs relative to the baseline $(q = 2, n = 248, w = 16)$. The former aligns naturally with 256-bit SIMD registers, increases the estimated classical bit security from $b = 143.7$ to $b = 150.4$, and achieves an approximate 2× speedup, while increasing both the public-key and signature sizes by only about 1.03×. The latter achieves the best performance among the evaluated parameter sets, with an approximate 5.6× speedup and an estimated classical bit security of $b = 145.4$, still above the baseline $b = 143.7$, at the cost of about 1.55× larger public keys and signatures.

Comparison with Hash-Based W-OTS$^+$. MQ-W-OTS$^+$ incurs an expected performance overhead compared to traditional hash-based W-OTS$^+$. First, standard hash functions (e.g., SHA-2) benefit from mature acceleration, whereas our PoC is a research-grade AVX2 implementation lacking production-level optimizations. More fundamentally, we trade the high efficiency achieved by relying on empirical, black-box security for explicit algebraic hardness and a consolidated security foundation.

5 Conclusion

In this paper, we presented MQ-based W-OTS$^+$, a trapdoor-free multivariate one-time signature scheme with provable security in the standard model. The

construction relies neither on hidden trapdoors nor on zero-knowledge proofs; instead, it utilizes the desirable cryptographic properties of the MQ function family and treat the MQ function family as a forward-evaluable primitive.

Future work includes exploring other properties of the MQ function family and extending this approach to practical many-time signature schemes.

References

1. Abe, M., Hofheinz, D., Nishimaki, R., Ohkubo, M., Pan, J.: Compact structure-preserving signatures with almost tight security. Journal of Cryptology **36**(4), 37 (2023). https://doi.org/10.1007/s00145-023-09477-z
2. Antonov, S.: Round 3 official comment: Sphincs+. NIST PQC Forum (Google Groups) comment (2022), available at https://groups.google.com/a/list.nist.gov/g/pqc-forum/c/FVItvyRea28/m/mGaRi5iZBwAJ, accessed April 13, 2025
3. Applebaum, B., Haramaty-Krasne, N., Ishai, Y., Kushilevitz, E., Vaikuntanathan, V.: Low-Complexity Cryptographic Hash Functions. In: Papadimitriou, C.H. (ed.) 8th Innovations in Theoretical Computer Science Conference (ITCS 2017). Leibniz International Proceedings in Informatics (LIPIcs), vol. 67, pp. 7:1–7:31. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany (2017). https://doi.org/10.4230/LIPIcs.ITCS.2017.7
4. Bellini, E., Makarim, R.H., Sanna, C., Verbel, J.: An estimator for the hardness of the MQ problem. In: Batina, L., Daemen, J. (eds.) Progress in Cryptology – AFRICACRYPT 2022. Lecture Notes in Computer Science, vol. 13503, pp. 323–347. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-17433-9_14
5. Benadjila, R., Feneuil, T., Rivain, M.: MQ on my mind: Post-quantum signatures from the non-structured multivariate quadratic problem. In: 9th IEEE European Symposium on Security and Privacy, EuroS&P 2024, Vienna, Austria, July 8–12, 2024. pp. 468–485. IEEE (2024). https://doi.org/10.1109/EuroSP60621.2024.00032
6. Berbain, C., Gilbert, H., Patarin, J.: QUAD: A practical stream cipher with provable security. In: Vaudenay, S. (ed.) Advances in Cryptology – EUROCRYPT 2006. Lecture Notes in Computer Science, vol. 4004, pp. 109–128. Springer, Berlin, Heidelberg (2006). https://doi.org/10.1007/11761679_8
7. Beullens, W.: Sigma protocols for MQ, PKP and SIS, and fishy signature schemes. In: Canteaut, A., Ishai, Y. (eds.) Advances in Cryptology – EUROCRYPT 2020. Lecture Notes in Computer Science, vol. 12107, pp. 183–211. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-45727-3_7
8. Beullens, W.: Breaking rainbow takes a weekend on a laptop. In: Dodis, Y., Shrimpton, T. (eds.) Advances in Cryptology – CRYPTO 2022. Lecture Notes in Computer Science, vol. 13508, pp. 464–479. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-15979-4_16
9. Billet, O., Robshaw, M.J.B., Peyrin, T.: On building hash functions from multivariate quadratic equations. In: Pieprzyk, J., Ghodosi, H., Dawson, E. (eds.) Information Security and Privacy – ACISP 2007. Lecture Notes in Computer Science, vol. 4586, pp. 82–95. Springer, Berlin, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73458-1_7
10. Bouillaguet, C., Feneuil, T., Maire, J., Rivain, M., Sauvage, J., Vergnaud, D.: Multivariate commitments and signatures with efficient protocols. Cryptology ePrint Archive (2025)
11. Briaud, P., Bros, M., Perlner, R., Smith-Tone, D.: Practical attack on all parameters of the HPPC signature scheme. In: Boura, C., Mashatan, A., Miri, A. (eds.) Selected Areas in Cryptography – SAC 2025. Lecture Notes in Computer Science, vol. 16207, pp. 519–538. Springer, Cham (2026). https://doi.org/10.1007/978-3-032-10536-3_19
12. Cabarcas, D., Li, P., Verbel, J.A., Villanueva-Polanco, R.: Improved attacks for SNOVA by exploiting stability under a group action. In: Tauman Kalai, Y.,

Kamara, S.F. (eds.) Advances in Cryptology – CRYPTO 2025. Lecture Notes in Computer Science, vol. 16005, pp. 358–389. Springer, Cham (2025). https://doi.org/10.1007/978-3-032-01887-8_12
13. Chen, M.S., Hülsing, A., Rijneveld, J., Samardjiska, S., Schwabe, P.: From 5-pass $\mathcal{MQ}$-based identification to $\mathcal{MQ}$-based signatures. In: Cheon, J.H., Takagi, T. (eds.) Advances in Cryptology – ASIACRYPT 2016. Lecture Notes in Computer Science, vol. 10032, pp. 135–165. Springer, Berlin, Heidelberg (2016). https://doi.org/10.1007/978-3-662-53890-6_5
14. Chen, M.S., Hülsing, A., Rijneveld, J., Samardjiska, S., Schwabe, P.: SOFIA: $\mathcal{MQ}$-based signatures in the QROM. In: Abdalla, M., Dahab, R. (eds.) Public-Key Cryptography – PKC 2018. Lecture Notes in Computer Science, vol. 10770, pp. 3–33. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-76581-5_1
15. Ding, J., Perlner, R., Petzoldt, A., Smith-Tone, D.: Improved cryptanalysis of HFEv- via projection. In: Lange, T., Steinwandt, R. (eds.) Post-Quantum Cryptography – PQCrypto 2018. Lecture Notes in Computer Science, vol. 10786, pp. 375–395. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-79063-3_18
16. Ding, J., Petzoldt, A., Schmidt, D.S.: Multivariate public key cryptosystems. Springer Nature (2020)
17. Ding, J., Yang, B.Y.: Multivariates polynomials for hashing. In: Pei, D., Yung, M., Lin, D., Wu, C. (eds.) Information Security and Cryptology – Inscrypt 2007. Lecture Notes in Computer Science, vol. 4990, pp. 358–371. Springer, Berlin, Heidelberg (2008). https://doi.org/10.1007/978-3-540-79499-8_28
18. Dodis, Y., Goldwasser, S., Tauman Kalai, Y., Peikert, C., Vaikuntanathan, V.: Public-key encryption schemes with auxiliary inputs. In: Micciancio, D. (ed.) Theory of Cryptography – TCC 2010. Lecture Notes in Computer Science, vol. 5978, pp. 361–381. Springer, Berlin, Heidelberg (2010). https://doi.org/10.1007/978-3-642-11799-2_22
19. Dods, C., Smart, N.P., Stam, M.: Hash based digital signature schemes. In: Smart, N.P. (ed.) Cryptography and Coding – IMACC 2005. Lecture Notes in Computer Science, vol. 3796, pp. 96–115. Springer, Berlin, Heidelberg (2005). https://doi.org/10.1007/11586821_8
20. Don, J., Fehr, S., Majenz, C.: The measure-and-reprogram technique 2.0: Multi-round Fiat-Shamir and more. In: Micciancio, D., Ristenpart, T. (eds.) Advances in Cryptology – CRYPTO 2020. Lecture Notes in Computer Science, vol. 12172, pp. 602–631. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-56877-1_21
21. Faugère, J.C., Safey El Din, M., Verron, T.: On the complexity of computing gröbner bases for weighted homogeneous systems. Journal of Symbolic Computation **76**, 107–141 (2016). https://doi.org/10.1016/j.jsc.2015.12.001
22. Furue, H., Duong, D.H., Takagi, T.: An efficient MQ-based signature with tight security proof. International Journal of Networking and Computing **10**(2), 308–324 (2020). https://doi.org/10.15803/ijnc.10.2_308
23. Garey, M.R., Johnson, D.S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman, San Francisco (1979)
24. Goldreich, O., Levin, L.A.: A hard-core predicate for all one-way functions. In: Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing. pp. 25–32. ACM, New York, NY, USA (1989)
25. Goldreich, O., Rubinfeld, R., Sudan, M.: Learning polynomials with queries: The highly noisy case. SIAM Journal on Discrete Mathematics **13**(4), 535–570 (2000). https://doi.org/10.1137/S0895480198344540, https://doi.org/10.1137/S0895480198344540

26. Guo, H., Jin, Y., Pan, Y., He, X., Gong, B., Ding, J.: Practical and theoretical cryptanalysis of VOX. In: Saarinen, M.J., Smith-Tone, D. (eds.) Post-Quantum Cryptography – PQCrypto 2024. Lecture Notes in Computer Science, vol. 14772, pp. 186–208. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62746-0_9
27. Guo, J., Li, H., Liu, M., Wang, S., Zhang, T.: Dual-syncopation meet-in-the-middle attacks: New results on SHA-2 and MD5. In: Daemen, J., Thomé, E. (eds.) Advances in Cryptology – EUROCRYPT 2026. Lecture Notes in Computer Science, vol. 16546, pp. 121–151. Springer, Cham (2026). https://doi.org/10.1007/978-3-032-25333-0_5
28. Harville, D.A.: Matrix Algebra From a Statistician's Perspective. Springer-Verlag, New York, NY, 1 edn. (1997). https://doi.org/10.1007/b98818, https://doi.org/10.1007/b98818
29. Hülsing, A., Butin, D., Gazdag, S., Rijneveld, J., Mohaisen, A.: RFC 8391: XMSS: eXtended Merkle Signature Scheme. RFC Editor (2018), https://www.rfc-editor.org/info/rfc8391
30. Hülsing, A.: W-OTS+ – shorter signatures for hash-based signature schemes. In: Youssef, A., Nitaj, A., Hassanien, A.E. (eds.) Progress in Cryptology – AFRICACRYPT 2013. Lecture Notes in Computer Science, vol. 7918, pp. 173–188. Springer, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38553-7_10
31. Kipnis, A., Patarin, J., Goubin, L.: Unbalanced oil and vinegar signature schemes. In: Stern, J. (ed.) Advances in Cryptology – EUROCRYPT '99. Lecture Notes in Computer Science, vol. 1592, pp. 206–222. Springer, Berlin, Heidelberg (1999). https://doi.org/10.1007/3-540-48910-X_15
32. Kudinov, M.A., Kiktenko, E.O., Fedorov, A.K.: Security analysis of the W-OTS$^+$ signature scheme: Updating security bounds. Matematicheskie Voprosy Kriptografii **12**(2), 129–145 (2021). https://doi.org/10.4213/mvk362, https://www.mathnet.ru/eng/mvk370
33. Liu, F.H., Lu, C.J., Yang, B.Y.: Secure PRNGs from specialized polynomial maps over any $\mathbb{F}_q$. In: Buchmann, J., Ding, J. (eds.) Post-Quantum Cryptography – PQCrypto 2008. Lecture Notes in Computer Science, vol. 5299, pp. 181–202. Springer, Berlin, Heidelberg (2008). https://doi.org/10.1007/978-3-540-88403-3_13
34. Matsumoto, T., Imai, H.: Public quadratic polynomial-tuples for efficient signature-verification and message-encryption. In: Günther, C.G. (ed.) Advances in Cryptology – EUROCRYPT '88. Lecture Notes in Computer Science, vol. 330, pp. 419–453. Springer, Berlin, Heidelberg (1988). https://doi.org/10.1007/3-540-45961-8_39
35. Moh, T.T.: A public key system with signature and master key functions. Communications in Algebra **27**(5), 2207–2222 (1999)
36. Patarin, J.: Hidden fields equations (HFE) and isomorphisms of polynomials (IP): Two new families of asymmetric algorithms. In: Maurer, U. (ed.) Advances in Cryptology – EUROCRYPT '96. Lecture Notes in Computer Science, vol. 1070, pp. 33–48. Springer, Berlin, Heidelberg (1996). https://doi.org/10.1007/3-540-68339-9_4
37. Patarin, J.: The oil and vinegar signature scheme (1997), presented at the Dagstuhl Workshop on Cryptography, September 1997
38. Perlner, R., Kelsey, J., Cooper, D.: Breaking category five SPHINCS+ with SHA-256. In: Cheon, J.H., Johansson, T. (eds.) Post-Quantum Cryptography –

PQCrypto 2022. Lecture Notes in Computer Science, vol. 13512, pp. 501–522. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-17234-2_23
39. Ran, L.: Wedges, oil, and vinegar: An analysis of UOV in the exterior algebra. In: Daemen, J., Thomé, E. (eds.) Advances in Cryptology – EUROCRYPT 2026. Lecture Notes in Computer Science, vol. 16544, pp. 363–388. Springer, Cham (2026). https://doi.org/10.1007/978-3-032-25327-9_13
40. Sakumoto, K., Shirai, T., Hiwatari, H.: Public-key identification schemes based on multivariate quadratic polynomials. In: Rogaway, P. (ed.) Advances in Cryptology – CRYPTO 2011. Lecture Notes in Computer Science, vol. 6841, pp. 706–723. Springer, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22792-9_40
41. Tao, C., Petzoldt, A., Ding, J.: Efficient key recovery for all HFE signature variants. In: Malkin, T., Peikert, C. (eds.) Advances in Cryptology – CRYPTO 2021. Lecture Notes in Computer Science, vol. 12825, pp. 70–93. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-84242-0_4
42. Thomae, E., Wolf, C.: Solving underdetermined systems of multivariate quadratic equations revisited. In: Fischlin, M., Buchmann, J., Manulis, M. (eds.) Public Key Cryptography – PKC 2012. Lecture Notes in Computer Science, vol. 7293, pp. 156–171. Springer, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-30057-8_10
43. Zhang, Z., Hou, C., Liu, M.: Probabilistic linearization: Internal differential collisions in up to 6 rounds of SHA-3. In: Reyzin, L., Stebila, D. (eds.) Advances in Cryptology – CRYPTO 2024. Lecture Notes in Computer Science, vol. 14923, pp. 241–272. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-68385-5_8
44. Zhang, Z., Li, M., Gao, L., Wang, M.: Collision attacks on SHA-256 up to 37 steps with improved trail search. In: Daemen, J., Thomé, E. (eds.) Advances in Cryptology – EUROCRYPT 2026. Lecture Notes in Computer Science, vol. 16546, pp. 91–120. Springer, Cham (2026). https://doi.org/10.1007/978-3-032-25333-0_4