Quantum-access security of the Winternitz one-time signature scheme

Christian Majenz*[1], Chanelle Matadah Manfouo†[2], and Maris Ozols‡[3]

[1] *Centrum Wiskunde & Informatica and QuSoft, The Netherlands*
[2] *African Institute for Mathematical Science & Quantum Leap Africa, Rwanda*
[3] *Institute for Logic, Language, and Computation, Korteweg-de Vries Institute for Mathematics, and Institute for Theoretical Physics, University of Amsterdam and QuSoft, The Netherlands*

March 23, 2021

Abstract

Quantum-access security, where an attacker is granted superposition access to secret-keyed functionalities, is a fundamental security model and its study has inspired results in post-quantum security. We revisit, and fill a gap in, the quantum-access security analysis of the Lamport one-time signature scheme (OTS) in the quantum random oracle model (QROM) by Alagic et al. (Eurocrypt 2020). We then go on to generalize the technique to the Winternitz OTS. Along the way, we develop a tool for the analysis of hash chains in the QROM based on the superposition oracle technique by Zhandry (Crypto 2019) which might be of independent interest.

Contents

*christian.majenz@cwi.nl
†cmatadah@quantumleapafrica.org
‡marozols@gmail.com

1 Overview

1.1 Introduction

Recently, research and development efforts towards building a universal quantum computer have intensified. As quantum computers will break currently deployed public-key cryptosystems [Sho94], finding adequate replacement schemes (called *post-quantum* secure) has been increasingly a priority, too, as reflected by the ongoing NIST standardization effort for post-quantum secure digital signature schemes and key encapsulation mechanisms [AASA+20].

Quantum-access security. While post-quantum security is the most important attack model involving quantum computers, the stronger *quantum-access* or *quantum world* attack model [BZ13, GHS16], where attackers are granted quantum access to secret-keyed functionalities, has received considerable attention, too. There are a number of reasons why this stronger attack model is important. On the one hand, it is of theoretical importance because it captures the strongest-known achievable security notions for standard classical cryptographic primitives. On the other hand, there are a number of conceivable scenarios where they become relevant, e.g. for composability with obfuscation or when constructing quantum-cryptographic schemes, or to prevent implementation-level vulnerabilities in a future hybrid quantum-classical computing infrastructure. Finally, results in the quantum access model can inform post-quantum cryptographic research, as exemplified by the offline Simon's algorithm attack [BHNP+19].

Blind unforgeability. In this work, we study the security of signature schemes under quantum-access attacks, in the quantum random oracle model (QROM) [BDF+11]. Here, generalizing the standard notion of existential unforgeability under chosen message attacks, the attacker is granted quantum query access to the signing algorithm. In the end, the adversary should output a forgery that they did not obtain from a query. Formalizing such a security notion is complicated due to the so-called *quantum no-cloning principle* according to which quantum states cannot be copied. We use the notion of blind unforgeability introduced in [AMRS20] (see [BZ13, GYZ17] for previous and complementary notions). Informally, blind unforgeability credits an adversary with a successful break of, e.g., a digital signature scheme, if it outputs a valid message-signature pair given a modified signing oracle that is "blinded" on a random subset of all messages in the sense that it outputs a dummy symbol instead of a signature, and if the output message is among these blinded messages (see section Section 2.4 for details).

Hash-based signature schemes. Hash-based signature schemes are prominent candidates for the replacement of digital signature schemes based on quantum-broken number-theoretic hardness assumptions. In particular, the stateful hash-based signature scheme XMSS [BDH11] has been standardized as RFC8391 [HBG+18], and the stateless hash-based signature scheme SPHINCS+ [BHK+19] is an alternate candidate in the ongoing NIST standardization process for post-quantum cryptographic schemes [AASA+20]. The security of hash-based signature schemes can be based on weak computational assumptions, like e.g. the one-wayness of the underlying hash function. Common hash based signature schemes, including the mentioned examples, are constructed using one-time[1] signature (OTS) schemes in combination with a hash-based authentication graph (e.g. a Merkle tree). The most well-known OTSs are the Lamport [Lam79a] and Winternitz [Mer89] OTS. Variations of the latter are used in both XMSS and SPHINCS+.

Previous work. In [AMRS20], the Lamport OTS is studied in the context of blind-unforgeability. More precisely, a proof of one-time blind-unforgeability in the QROM is provided. That proof, however, contains an imprecision in the analysis of the adversarial success. In particular, an auxiliary measurement is used to "collapse" an invariant property that holds *in superposition* into holding *classically*, but the effect of the dependence of this auxiliary measurement on the forgery message is not analyzed.

Related work. Quantum-access security for encryption is an active research area, and generalizing chosen-ciphertext security notions to the quantum access setting has posed, and poses, similar challenges as the ones encountered in the authenticity setting [BZ13, GHS16, GKS20]. On the negative side, key recovery attacks in the quantum access model against a number of symmetric-key primitives that are secure in the respective standard attack models have been discovered [SS17, KLLNP16], and have lead to the discovery of quantum attacks that can be performed without quantum access to secret-keyed functionalities [BHNP$^+$19].

There are a number of works that prove query lower bounds using variants of the superposition oracle technique [LZ19, CFHL20, GHHM20, Unr21]. In the concurrent work [Unr21], a similar reasoning using approximately invariant subspaces is employed to prove query lower bounds for random permutation oracles.

1.2 Summary of results

The Lamport OTS is blind-unforgeable. We revisit the analysis of the Lamport OTS in the QROM presented in [AMRS20] and give a complete proof of blind unforgeability as stated in the following theorem.

Theorem 1 (Blind unforgeability of the Lamport OTS, informal)**.** *The Lamport OTS is blind-unforgeable if the underlying hash function h is modeled as a quantum-accessible random oracle. More precisely, the success probability of any blind unforgeability adversary $\mathcal{A}$ against the Lamport OTS that makes $q > 0$ quantum queries to the random oracle is bounded as*

$$\Pr[\mathcal{A} \text{ succeeds}] \leq C_L q^2 l^3 \cdot 2^{-n},$$

where C_L is a constant, n is the security parameter of the Lamport OTS and l is the message length.

Compared to [AMRS20], our security proof features the following improvements:

- We streamline the usage of the superposition oracle technique of Zhandry [Zha19]. In particular, our analysis only uses (a variant of) the superposition oracle technique to sample the secret key. We reprogram, *in superposition*, the standard random oracle at inputs contained in the secret key. This technique represents a general tool to analyze hash chains in the QROM and might be of independent interest.
- We give a full analysis of the success probability using an auxiliary measurement idea from [AMRS20]. To tackle the problem mentioned above, we introduce a novel technique of tracking an invariant property *in superposition* using projectors and commutators.

The Winternitz OTS is blind-unforgeable. With the full blind unforgeability analysis of the Lamport OTS in hand, we generalize the approach to the Winternitz OTS.

Theorem 2 (Blind unforgeability of the Winternitz OTS, informal)**.** *The Winternitz OTS is blind-unforgeable if the underlying hash function h is modeled as a quantum-accessible random oracle. More precisely, the success probability of any blind unforgeability adversary $\mathcal{A}$ against the Winternitz OTS that makes $q > 0$ quantum queries to the random oracle is bounded as*

$$\Pr[\mathcal{A} \text{ succeeds}] \leq C_W q^2 a^3 \frac{w^4}{\log^3 w} \cdot 2^{-n},$$

where C_W is a constant, n is the security parameter of the Winternitz OTS, a is the message length and $w \geq 2$ is the Winternitz parameter used to trade off signature size versus signing and verification time.

While the simplified analysis of hash chains in the QROM described above was advantageous in proving the blind unforgeability security of the Lamport OTS, it is indispensable in the analysis of the Winternitz scheme. Here, long hash chains are considered and the technique of using the superposition oracle to detect which hash chain elements are known to the adversary relies on the oracle register (or rather here: the hash chain register) being in a product state.

[1] And sometimes few-time signature schemes, e.g. in SPHINCS+.

1.3 Technical overview

In this technical overview, we give a high-level description of our techniques for analyzing the blind unforgeability security of the Lamport and Winternitz OTSs in the QROM.

The superposition oracle technique and hash chains. As in many contexts that concern message authenticity and integrity, the main roadblock we have to overcome in our analysis is the so-called *recording barrier*: quantum oracle queries can, in general, not be recorded for later use. In particular, after a single quantum signing query, it is not possible to reason about the unused parts of the secret key. This is because, in general, all secret key strings have been used in some part of the superposition.

In [AMRS20], Zhandry's superposition oracle technique is used in a novel way to recover the ability to reason about which secret key strings are (un)known to the adversary. There, the secret key of the Lamport scheme, which is a $2 \times l$ array of independent uniformly random n-bit strings, is essentially regarded as a random function from $\{0,1\} \times \{1,\ldots,l\}$. This function, as well as the hash function the Lamport OTS is constructed from, is then modelled using the superposition oracle technique.

We improve this technique as follows. We use the fact that sampling two correlated random variables X and Y can be done by first sampling X, and then Y according to the conditional distribution, or vice versa. In the context of *hash chains* in the (Q)ROM, i.e. sequences of strings $x_0, x_1 = H(x_0), x_2 = H(x_1), \ldots$ for a random oracle H, this means that instead of sampling x_0 and H, and then computing the remaining hash chain elements, we can as well sample $x_0, x_1, \ldots$ from their joint distribution, sample H, and *reprogram H to be consistent with the x_i*. This allows us to i) change the distribution of the x_i to a simpler one that is close in total variational distance, and ii) refrain from using the full superposition oracle technique for H. In particular, we use i) to replace the hash chains that are generated by the key generation algorithms of the Lamport and Winternitz schemes by tuples of independent random strings. This incurs only a small error, as the uniform distribution and the distribution of a hash chain in the (Q)ROM with random starting value x_0 are equal conditioned on all x_i being distinct. But collisions between different hash chain elements are unlikely.

Now that the hash chain elements are independent strings, we can use the full power of the superposition oracle technique. In particular, the one-to-one correspondence between the adversary's ignorance of a hash chain element, and the corresponding superposition oracle register being in uniform superposition, is restored.

Throughout the paper, and in the rest of this technical overview, we perform the analysis in a world where hash chains are formed using a superposition oracle modeling independent uniformly random strings, and the random oracle is reprogrammed accordingly. We call this the Quantum independent world. To conclude our analysis, we make use of the approximate indistinguishability of the Real and Quantum independent world.

Blind unforgeability and classical invariants in superposition. With the tools for analyzing hash chains in the QROM in hand, the next challenge consists of generalizing the classical security arguments for the Blind Unforgeability (BU) of the Lamport and Winternitz OTSs to the quantum access setting. The core of these security arguments is, at a high level, that for each unqueried message, any valid signature contains a string that is unknown to the adversary.[2] As mentioned above, this kind of reasoning does not generalize to the quantum access setting, as here an adversary can query all messages in superposition.

In the security game for the notion of BU, however, the adversary is not provided with an oracle for the full signing algorithm functionality. Instead, the adversary is provided with an oracle for a modified signing algorithm that is "blinded" on a random subset of the messages, in the sense that for these messages it outputs a dummy symbol instead of a signature. These "blinded messages" can now replace the unqueried messages in security arguments, as by definition the adversary is prevented from obtaining a valid signature for them from the blinded signing oracle.

For obtaining a quantum generalization, we need to reformulate this argument. The statement that for each unqueried message any valid signature contains a string unknown to the adversary, is equivalent to saying that, for each fixed message m^* and all $m \neq m^*$, some information related to the secret key is necessary to compute the signature for m^* that is not revealed by the signature for m. For Blind Unforgeability, it suffices to consider blinded m^* and unblinded m. In the superposition oracle framework, the statement "there exists an unblinded message such that the registers corresponding to all parts of the secret key that the signature for that message does not reveal, are in the uniform superposition state" defines a subspace I. By definition, the global state after a BU-adversary makes a single query to the blinded signing oracle, and no queries to the random oracle, is in that subspace.

The crucial step in our analysis is to show that the adversary-oracle state approximately remains in the subspace I, even if the adversary performs a moderate number of quantum queries to the random oracle.

[2] When basing the security on one-wayness, "unknown" is to be taken in a computational sense, but as this paper is about security in the (Q)ROM, it is sufficient to interpret "unknown to" as "independent of the state of".

This means the subspace I can serve as an *invariant*.

Random oracle queries and commutators. To analyze the "leakage" from the invariant subspace I, we use bounds on the norm of matrix commutators: to prove that the final oracle-adversary state is approximately in the invariant subspace I, we can equivalently show that applying the corresponding projector Π_I does not change the state by a lot. We know, however, that the projector does not change the state at all before any random oracle queries have been made. Therefore it suffices to bound the operator norm of the commutator between the projector Π_I and the unitary operator that facilitates random oracle queries in the Quantum independent world. We derive such a norm bound (see e.g. eq. (64) for the Lamport case), and the proof follows the classical intuition about the one-wayness of the random oracle.

2 Preliminaries

Let us introduce some notation that will be used throughout the paper. In this document, quantum systems are associated with finite-dimensional complex Euclidean vector spaces endowed with an inner product. Registers will be denoted by capital letters. We say that $\epsilon = \epsilon(n)$ is negligible if, for all polynomials $p(n)$, $\epsilon(n) < 1/p(n)$ for large enough $n \in \mathbb{N}$. We use the notation $x \overset{\$}{\leftarrow} D$ to say that x is chosen uniformly at random from a set D. We write S^c to denote the complement of set S (in a superset that is clear from the context). We write $s \parallel t$ to denote the *concatenation* of strings s and t, and $[A, B] = AB - BA$ to denote the *commutator* of operators A and B. Throughout this paper, quantum adversaries refer to quantum poly-time algorithms and are denoted by $\mathcal{A}$.

2.1 Quantum computing

In this section, we introduce some basic notions from quantum computing. We refer the reader to [NC02] for more details.

Quantum states. To a given quantum system with d distinguished states we associate a d-dimensional complex Euclidean vector space $\mathcal{H} = \mathbb{C}^d$ endowed with an inner product $\langle \cdot | \cdot \rangle$. The *state* of such system is described by a unit vector, i.e., a vector $|\psi\rangle \in \mathcal{H}$ such that $\langle \psi | \psi \rangle = 1$. For example, a *qubit* state is described by a vector $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \in \mathbb{C}^2$ such that $|\alpha|^2 + |\beta|^2 = 1$, where $|0\rangle$ and $|1\rangle$ are the *computational basis* vectors. The corresponding *dual vector* is given by $\langle\psi| = \bar{\alpha}\langle 0| + \bar{\beta}\langle 1|$, which can also be expressed through entry-wise complex conjugation and transpose: $\langle\psi| = |\psi\rangle^\dagger = \overline{|\psi\rangle}^\mathsf{T}$.

Given two quantum systems A and B with state spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the composite system AB has state space $\mathcal{H}_A \otimes \mathcal{H}_B$ described by the tensor product. In particular, if $|\psi\rangle_A \in \mathcal{H}_A$ and $|\psi\rangle_B \in \mathcal{H}_B$ are states of the two individual systems, then their joint state is given by $|\psi\rangle_A \otimes |\psi\rangle_B$ or simply $|\psi\rangle_A|\psi\rangle_B$. We will often refer to the subsystems A and B as *registers*. For example, an n-qubit system consists of n qubit registers and its computational basis is given by $|x\rangle = |x_1\rangle \cdots |x_n\rangle$ where $x = x_1 \dots x_n \in \{0,1\}^n$. A general n-qubit state is then a linear combination of the computational basis states:

$$\sum_{x \in \{0,1\}^n} \alpha_x |x\rangle \quad \text{with} \quad \sum_{x \in \{0,1\}^n} |\alpha_x|^2 = 1.$$

In particular, when all α_x are equal to $2^{-n/2}$, we call this the *uniform superposition*. Throughout this paper, we will denote this state and the corresponding projector by

$$|\Phi\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle, \qquad \Phi = |\Phi\rangle\langle\Phi|, \tag{1}$$

where the latter corresponds to $2^n \times 2^n$ matrix with all entries equal to $1/2^n$.

Quantum computation proceeds by applying unitary transformations to the state. The information is then read out by applying a measurement.

Unitary transformations. The evolution of a d-dimensional quantum system is described by a *unitary transformation*, i.e., a complex $d \times d$ matrix U such that $UU^\dagger = I$, where $U^\dagger = \bar{U}^\mathsf{T}$ denotes the conjugate transpose of U. If a unitary U is applied only on the A register of a joint system AB that is in state $|\psi\rangle_{AB}$, we write $(U_A \otimes \mathbb{1}_B)|\psi\rangle_{AB}$ where $\mathbb{1}$ denotes the *identity transformation*. We will often abbreviate this as $U_A|\psi\rangle_{AB}$.

Measurement. We can extract information from a quantum state $|\psi\rangle$ by performing a measurement. For our purpose it will be enough to consider only projective measurements. A *projective measurement* is described by a set $\{P_1, \dots, P_k\}$ of orthogonal projectors ($P_i^\dagger = P_i$ and $P_i^2 = P_i$) such that $\sum_{i=1}^k P_i = \mathbb{1}$. When performing a measurement on a quantum state $|\psi\rangle$, the probability of getting outcome i is $p(i) = \langle\psi|P_i|\psi\rangle$. Upon getting outcome i, the state $|\psi\rangle$ collapses to $P_i|\psi\rangle/\sqrt{p(i)}$. Given a composite system AB, a measurement on the subsystem A has operators of the form $(P_i)_A \otimes \mathbb{1}_B$, and the outcome probabilities and post-measurement states are determined analogously [Wat18].

Quantum-accessible oracles. On a quantum computer, a function can be evaluated on a superposition of inputs. The standard way of modelling superposition black-box access to a function $f : \{0,1\}^n \to \{0,1\}^m$ is by providing an oracle for the unitary operation O_f that acts on $n + m$ qubits and is defined by

$$O_f|x\rangle|y\rangle \mapsto |x\rangle|y \oplus f(x)\rangle, \tag{2}$$

for all $x \in \{0,1\}^n$ and $y \in \{0,1\}^m$. Without loss of generality, an algorithm $\mathcal{A}$ that makes q queries to such an oracle has the following form:

$$U_q O_f \cdots U_1 O_f U_0 |\Psi_0\rangle = V_{\mathcal{A}}^{O_f}|\Psi_0\rangle = |\Psi\rangle,$$

possibly followed by a measurement. Here, $|\Psi_0\rangle$ is an initial state U_i are arbitrary unitary operations that do not depend on f.

In this work, we will deal with algorithms that have two oracles, O_1 and O_2, but may only query O_2 at most once (O_1 will be a random oracle and O_2 a signing oracle for a one-time signature scheme). In this case, we can regard an algorithm $\mathcal{A}^{O_1,O_2} = (\mathcal{A}_0^{O_1}, \mathcal{A}_1^{O_1})$ as a two-stage process: $\mathcal{A}_0^{O_1}$ prepares the input for O_2 and an internal register, while $\mathcal{A}_1^{O_1}$ receives the internal state and the output of O_2, and produces the final output of $\mathcal{A}$. In other words, the execution of $\mathcal{A}$ results in the state

$$|\Psi\rangle = V_{\mathcal{A}_1}^{O_1} O_2 V_{\mathcal{A}_0}^{O_1}|\Psi_0\rangle.$$

The most well-known situation in cryptography that features a quantum oracle is the so-called *quantum random oracle model* (QROM) [BDF+11]. In the QROM, just as in the classical random oracle model (ROM) [BR93], a hash function is modeled as a uniformly random function h that all agents have oracle access to, meaning that quantum oracle access to the unitary O_h defined in eq. (2) is provided. This model is used to prove cryptographic security against quantum adversaries when basing security on concrete properties like, e.g., collision resistance, is hard or inefficient.

2.2 Tools from linear algebra

In this section, we state a couple of simple lemmas used in security proofs in Sections 4 and 5. For the first lemma, we use the formulation from [BZ13] (Lemma 2.1), and the proof is also provided in the same reference.

Lemma 1 (Special case of the pinching lemma [Hay02])**.** *Let $\mathcal{A}$ be a quantum algorithm and x any output value of $\mathcal{A}$. Let $\mathcal{A}_0$ be another quantum algorithm obtained from $\mathcal{A}$ by pausing $\mathcal{A}$ in an arbitrary stage of execution, performing a projective measurement that obtains one of k outcomes, and then resuming $\mathcal{A}$. Then,*

$$\Pr[\mathcal{A}_0(1^n) = x] \geq \frac{\Pr[\mathcal{A}(1^n) = x]}{k}. \tag{3}$$

Lemma 2. *Let A and $\{B_i\}_{i=1}^n$ be operators, acting on the same space, such that $\|A\|_\infty, \|B_i\|_\infty \leq 1$. Then*

$$\Big\| \Big[A, \prod_{i=1}^n B_i\Big] \Big\|_\infty \leq \sum_{i=1}^n \|[A, B_i]\|_\infty.$$

Proof. Note that

$$[A, BC] = [A, B]C + B[A, C] \tag{4}$$

for any operators A, B, C acting on the same space. Hence,

$$\begin{aligned}\Big\| \Big[A, \prod_{i=1}^{n} B_i\Big] \Big\|_\infty &= \Big\| [A, B_1] \prod_{i=2}^{n} B_i + B_1 \Big[A, \prod_{i=2}^{n} B_i\Big] \Big\|_\infty \\ &\leq \Big\| [A, B_1] \prod_{i=2}^{n} B_i \Big\|_\infty + \Big\| B_1 \Big[A, \prod_{i=2}^{n} B_i\Big] \Big\|_\infty \\ &\leq \| [A, B_1] \|_\infty \Big\| \prod_{i=2}^{n} B_i \Big\|_\infty + \| B_1 \|_\infty \Big\| \Big[A, \prod_{i=2}^{n} B_i\Big] \Big\|_\infty \\ &\leq \| [A, B_1] \|_\infty + \Big\| \Big[A, \prod_{i=2}^{n} B_i\Big] \Big\|_\infty ,\end{aligned}$$

where the first two inequalities follow from the triangle inequality and the sub-multiplicative property of the norm, respectively, and the last inequality holds because $\|B_1\|_\infty \leq 1$ and

$$\Big\| \prod_{i=2}^{n} B_i \Big\|_\infty \leq \|B_2\|_\infty \|B_3\|_\infty \cdots \|B_n\|_\infty \leq 1.$$

The desired inequality follows by applying the same argument inductively. □

Lemma 3. *Let X and Y be two n-qubit quantum systems and let*

$$P^{=}_{XY} = \sum_{x \in \{0,1\}^n} |x\rangle\langle x|_X \otimes |x\rangle\langle x|_Y$$

be the projector onto the subspace spanned by those computational basis vectors where the two registers are equal. Let $\Phi = |\Phi\rangle\langle\Phi|$ denotes the projector onto the uniform superposition, see eq. (1). Then

$$\|P^{=}_{XY} \Phi_Y\|_\infty = 2^{-n/2}. \tag{5}$$

Proof. Recall that $|\Phi\rangle = 2^{-n/2} \sum_{x \in \{0,1\}^n} |x\rangle$, so $|x\rangle\langle x|\Phi\rangle\langle\Phi| = 2^{-n/2} |x\rangle\langle\Phi|$ for any $x \in \{0,1\}^n$. Hence

$$\begin{aligned}\|P^{=}_{XY} \Phi_Y\|_\infty &= 2^{-n/2} \Big\| \sum_{x \in \{0,1\}^n} |x\rangle\langle x|_X \otimes |x\rangle\langle\Phi|_Y \Big\|_\infty & (6)\\ &= 2^{-n/2} \max_{x \in \{0,1\}^n} \big\| |x\rangle\langle\Phi|_Y \big\|_\infty & (7)\\ &= 2^{-n/2}, & (8)\end{aligned}$$

where the second equality holds since the matrix is block diagonal and the last line follows from the fact that $|x\rangle$ and $|\Phi\rangle$ are unit vectors. □

By applying the triangle inequality, Lemma 3 implies the following bound on the commutator:

$$\|[P^{=}_{XY}, \Phi_Y]\|_\infty \leq 2 \cdot 2^{-n/2}. \tag{9}$$

2.3 Hash-based one-time signature schemes

Hash-based signature schemes [Lam79a, Mer89] are among the digital signature schemes whose security relies on the weakest assumptions. In this paper, we study hash-based one-time signatures (OTSs) which are digital signature schemes that use a pair of keys to sign and verify a single message. Their classical security can be based on the existence of a family of hash functions with certain properties such as one-wayness, collision resistance, pre-image resistance and/or second pre-image resistance. In this section, we present the Lamport OTS and the Winternitz OTS.

2.3.1 The Lamport OTS

The Lamport OTS (also known as Lamport–Diffie OTS), first introduced in [Lam79b], is used as basis for many other hash-based signature schemes. This scheme uses a hash function $h : \{0,1\}^n \to \{0,1\}^n$ for key generation and verification, where $n \in \mathbb{N}$ denotes the security parameter.

1. *Parameters.* Security parameter $n \in \mathbb{N}$ and message length $l \in \mathbb{N}$.

2. *Key generation algorithm* (KeyGen). On input of the security parameter n in unary, KeyGen outputs a secret signing key sk and a public verification key pk: $(\mathsf{sk}, \mathsf{pk}) \leftarrow \mathsf{KeyGen}(1^n)$ as follows,

$$\begin{aligned} \mathsf{sk} &= (s_i^j)_{i=1,\ldots,l}^{j=0,1} \quad \text{with} \quad s_i^j \xleftarrow{\$} \{0,1\}^n, \\ \mathsf{pk} &= (p_i^j)_{i=1,\ldots,l}^{j=0,1} \quad \text{where} \quad p_i^j = h(s_i^j) \in \{0,1\}^n. \end{aligned}$$

3. *Signature algorithm* ($\mathsf{Sign}_{\mathsf{sk}}$). On input of a message $m = m_1 \ldots m_l \in \{0,1\}^l$ of length l, $\mathsf{Sign}_{\mathsf{sk}}$ outputs the following signature:

$$\mathsf{Sign}_{\mathsf{sk}}(m) = \sigma = \sigma_1 \ldots \sigma_l \quad \text{where} \quad \sigma_i = s_i^{m_i} \in \{0,1\}^n.$$

4. *Verification procedure* ($\mathsf{Ver}_{\mathsf{pk}}$). The verification procedure checks the correctness of the signature using the public key pk. Upon receiving a message m and a signature $\sigma = \sigma_1 \ldots \sigma_l$, $\mathsf{Ver}_{\mathsf{pk}}$ outputs the following:

$$\mathsf{Ver}_{\mathsf{pk}}(m, \sigma) = \begin{cases} \texttt{acc} & \text{if } h(\sigma_i) = p_i^{m_i} \text{ for all } i \in \{1, \ldots, l\}, \\ \texttt{rej} & \text{otherwise.} \end{cases}$$

2.3.2 The Winternitz OTS

The Winternitz OTS was first introduced by Merkle [Mer89] and many variants of the Winternitz OTS have been proposed since then. In this work, we study the Winternitz OTS that uses function chains. Before describing the scheme, we define the notion of function chains in line with [DSS05].

Definition 1 (Function chain). *Let $n \in \mathbb{N}$ denote a security parameter and let $\mathcal{D}$ and $\mathcal{K}$, called domain and key space, be sets whose elements have a description length polynomial in n. A function chain $\mathcal{C} = (\mathcal{I}, \mathcal{E})$ consists of a pair of probabilistic polynomial time algorithms $\mathcal{I}$ and $\mathcal{E}$.*

- Initialization algorithm $\mathcal{I}$: *given the security parameter n in unary and a chain length parameter $z \in \mathbb{N}$, it returns a public chain key $\mathsf{ck} = \mathcal{I}(1^n, z) \in \mathcal{K}$ to be used for chains of length z.*
- Evaluation algorithm $\mathcal{E}_{\mathsf{ck}}$: *given an interval $\{i, \ldots, j\} \subseteq \{0, \ldots, z\}$ and assuming $x \in \mathcal{D}$ is the i-th value of the chain, $\mathcal{E}_{\mathsf{ck}}$ uses the chain key $\mathsf{ck} \in \mathcal{K}$ to compute the j-th value $\mathcal{E}_{\mathsf{ck}}^{i,j}(x) \in \mathcal{D}$ of the same chain. We will abbreviate $\mathcal{E}_{\mathsf{ck}}^{0,j}$ as $\mathcal{E}_{\mathsf{ck}}^{j}$.*

These two algorithms fulfill the following correctness condition: for any parameters $n, z \in \mathbb{N}$, public chain key $\mathsf{ck} \leftarrow \mathcal{I}(1^n, z)$, positions $0 \le i \le j \le k \le z$ in the chain, and value $x \in \mathcal{D}$ at position i,

$$\mathcal{E}_{\mathsf{ck}}^{j,k}(\mathcal{E}_{\mathsf{ck}}^{i,j}(x)) = \mathcal{E}_{\mathsf{ck}}^{i,k}(x).$$

Now, we give a description of the Winternitz OTS scheme. It involves several parameters and consists of three probabilistic polynomial-time algorithms.

1. *Parameters*. The scheme is parameterized by a security parameter n, binary message length a, and the Winternitz parameter $w \ge 2$ that determines the time-memory trade-off (these parameters are integers and are publicly known). The parameters a and w are used to compute

$$l_1 = \left\lceil \frac{a}{\log(w)} \right\rceil, \qquad l_2 = \left\lfloor \frac{\log(l_1(w-1))}{\log(w)} \right\rfloor + 1, \qquad l = l_1 + l_2. \tag{10}$$

The Winternitz OTS uses a function chain $\mathcal{C} = (\mathcal{I}, \mathcal{E})$ with chain length $w - 1$ and domain $\mathcal{D} = \{0,1\}^n$ as described in Definition 1.

2. *Key generation algorithm* (KeyGen). On input of security parameter n, the key generation algorithm first chooses uniformly at random l values $\mathsf{sk} = (s_1, \ldots, s_l) \xleftarrow{\$} \mathcal{D}^l$ that form the secret signing key. Next, it initializes the function chain $\mathsf{ck} \leftarrow \mathcal{I}(1^n, w-1)$ by generating a chain public key ck. Finally, it computes the public verification key pk as follows:

$$\mathsf{pk} = (p_0, p_1, \ldots, p_l) = \left(\mathsf{ck}, \mathcal{E}_{\mathsf{ck}}^{w-1}(s_1), \ldots, \mathcal{E}_{\mathsf{ck}}^{w-1}(s_l)\right).$$

3. *Signature algorithm* ($\mathsf{Sign}_{\mathrm{sk}}$). For a given input message $x \in \{0,1\}^a$ and secret key sk, the signature algorithm first computes a base-w representation of x with l_1 digits: $m = (b_1, \ldots, b_{l_1})$ where $b_i \in \{0, \ldots, w-1\}$. Next, it computes the checksum

$$C(m) = \sum_{i=1}^{l_1} (w - 1 - b_i)$$

and represents it as l_2 digits $C(m) = (b_{l_1+1}, \ldots, b_l)$ in base w. Note that the length of the base-w representation of $C(m)$ is at most l_2 since $C(m) \leq l_1(w-1)$. We set

$$b(m) = (b_1, \ldots, b_l) = m \parallel C(m), \tag{11}$$

the concatenation of the base-w representations of m and $C(m)$. The signature is then computed as

$$\sigma = (\sigma_1, \ldots, \sigma_l) = (\mathcal{E}_{\mathrm{ck}}^{b_1}(s_1), \ldots, \mathcal{E}_{\mathrm{ck}}^{b_l}(s_l)).$$

Notice that the checksum can be considered as an intermediate verification step. Given the value of b corresponding to a message m, it guarantees that the b' corresponding to any other message $m' \neq m$ contains at least one $b'_i < b_i$, $1 \leq i \leq l$.

4. *Verification algorithm* ($\mathsf{Ver}_{\mathrm{pk}}$). Given a message m of binary length a, a signature σ and the public verification key pk, the verification algorithm first computes the $(b_1, \ldots, b_l)$ as described above and then checks whether the value of $\mathcal{E}_{\mathrm{ck}}^{b_i, w-1}(\sigma_i)$ agrees with the public key p_i:

$$\mathsf{Ver}_{\mathrm{pk}}(m, \sigma) = \begin{cases} \texttt{acc} & \text{if } p_i = \mathcal{E}_{\mathrm{ck}}^{b_i, w-1}(\sigma_i) \text{ for all } i \in \{1, \ldots, l\}, \\ \texttt{rej} & \text{otherwise.} \end{cases}$$

The simplest function chains which are commonly used to instantiate the Winternitz OTS are defined using a hash function h. In case h is a real-world hash function, $\mathcal{I}$ is trivial and $\mathcal{E}^{i,j} = h^{j-i}$. We will study the Winternitz OTS using such a function chain, in the QROM, i.e. when h is a quantum-accessible oracle for a random function.

2.4 Blind unforgeability

Blind unforgeability (BU) [AMRS20] is a quantum-access replacement for the standard security notion of EU-CMA for message authentication codes and digital signature schemes. Before we recall the definition of this notion, we need to introduce some additional background, including the concept of *blinding* a function and the *blind forgery* experiment.

Definition 2 (Blinding a function). *Let $f : X \to Y$ be a function and $B \subset X$ a subset of X. The* blinded *function Bf with respect to the* blinding set *B is defined as*

$$Bf(x) = \begin{cases} \perp & \textit{if } x \in B, \\ f(x) & \textit{otherwise,} \end{cases} \tag{12}$$

where $\perp$ is a special blinding symbol. One concrete way to instantiate this is by means of an extra bit. In that case, given a function $f : \{0,1\}^n \to \{0,1\}^m$, we define $Bf : \{0,1\}^n \to \{0,1\}^{m+1}$ by

$$Bf(x) = \begin{cases} 0^n \parallel 1 & \textit{if } x \in B, \\ f(x) \parallel 0 & \textit{otherwise.} \end{cases} \tag{13}$$

The second definition is more convenient because it enables us to easily measure and control from this bit without modifying the output of the function f.

Blinding a signing algorithm. Let $\mathsf{Sign}_{\mathrm{sk}}$ be a signing algorithm for a signature scheme with message space M. Now, sample a blinding set $B \subseteq M$ by adding every input with probability ϵ, independently. Then, the blinded signing algorithm is given by

$$B\,\mathsf{Sign}_{\mathrm{sk}}(m) = \begin{cases} \perp & \text{if } m \in B, \\ \mathsf{Sign}_{\mathrm{sk}}(m) & \text{otherwise.} \end{cases} \tag{14}$$

Note that in the rest of the document, we refer to B^c as the subspace of un-blinded messages.

The blind forgery (BF) experiment. Let $S = (\mathsf{KeyGen}, \mathsf{Sign}, \mathsf{Ver})$ be a digital signature scheme with a security parameter n and message space M. Let $\mathcal{A}$ be an adversary and let $\epsilon : \mathbb{N} \to \mathbb{R}_+$ be a negligible function. We define the *blind forgery* experiment $\mathsf{BlindForge}_{S,\mathcal{A}}(n, \epsilon)$ as follows:

- *Key generation*: $(\mathsf{sk}, \mathsf{pk}) \leftarrow \mathsf{KeyGen}(1^n)$;
- *Generation of blinding set*: select the blinding set $B \subseteq M$ by choosing each $m \in M$ independently at random with probability ϵ;
- *Forgery*: $(m, \sigma) \leftarrow \mathcal{A}^{B\,\mathsf{Sign}_{\mathsf{sk}}}(1^n)$;
- *Outcome*: win if $\mathsf{Ver}_{\mathsf{pk}}(m, \sigma) = \mathtt{acc}$ and $m \in B$, and lose otherwise.

Definition 3 (Blind unforgeability (BU))**.** *A digital signature scheme S is* q-BU secure *if for any adversary $\mathcal{A}$ making at most q queries to $B\,\mathsf{Sign}_{\mathsf{sk}}$, the success probability of winning the BF experiment is negligible in the security parameter n, i.e.*

$$\Pr\left[\mathcal{A} \text{ wins } \mathsf{BlindForge}_{S,\mathcal{A}}(n, \epsilon)\right] \leq \mu(n) \tag{15}$$

for some negligible function μ.

This paper is concerned with one-time signature schemes for which the pair of keys is used only once. That is, a BlindForge notion of security for one-time signature schemes in which the adversary is allowed to query the $B\,\mathsf{Sign}_{\mathsf{sk}}$ algorithm only once, i.e. $q = 1$.

3 Hash chains in the QROM

3.1 Quantum hash chain sampling

In this section, we introduce hash chains, several closely related worlds, and show that we can work in the one that is the easiest to handle. More precisely, we describe this technical tool that we will use to prove BU security for the Lamport and Winternitz OTSs. For both OTSs, the KeyGen routine computes so-called *hash chains*, i.e. sequences of strings obtained by iteratively applying a hash function.

In the (Q)ROM, to generate a hash chain based on a hash function h, we first sample an initial string s_0 uniformly at random and then compute $s_i = h(s_{i-1})$ for $i = 1, \ldots, w-1$ to obtain a hash chain of length w. For key generation in the Lamport or Winternitz OTS, the secret key sk is a tuple of l initial strings $s_1, \ldots, s_l \xleftarrow{\$} \{0,1\}^n$ sampled uniformly at random. Then a tuple of hash chains $\gamma = (\gamma_i^j)_{i=1,\ldots,l}^{j=0,\ldots,w-1}$ is obtained by querying the hash function h on each string of the secret key $w-1$ times, i.e.

$$\gamma_i^0 = s_i, \quad \gamma_i^j = h^j(\gamma_i^0), \quad p_i = \gamma_i^{w-1} = h^{w-1}(\gamma_i^0), \quad j = 0, \ldots, w-1, \quad i = 1, \ldots, l,$$

where w is the length of the hash chain ($w = 2$ for Lamport) and l is the number of hash chains. The final entry of each chain is used as a public key.

In the BlindForge game, the secret key is only used by the blinded signing oracle. When analyzing this experiment, we can thus modify the key generation, signing and random oracle algorithms in an arbitrary way, as long as the modified triple is indistinguishable from the real one to an adversary.

In the proofs in Sections 4 and 5 we make use of the following modified triple, which we will refer to as defining the Quantum independent world. We construct the secret key and the intermediate hash chain elements initially in uniform superposition. That is, we prepare each hash chain register $(\Gamma_i^j)_{i=1,\ldots,l}^{j=0,\ldots,w-2}$ in the uniform superposition state $|\Phi\rangle$, with the intention of measuring them to sample the strings γ_i^j in mind. Then, we sample the final hash chain at random. The random oracle is then "reprogrammed in superposition" to be approximately consistent with the hash chains.

We proceed to show that the way of implementing the hash chain and the random oracle in the Real world and in the Quantum independent world are indistinguishable. For that purpose, we first formally define both worlds and some intermediate worlds between them. Each world is specified by two oracles, H and Sign, replacing the random oracle h and the signing oracle in the Real world (the KeyGen algorithm is implicitly replaced by the setup described in each world below, that generates the initial state and the public key). The oracles of the Quantum independent world are described below as well.

Real world. In the Real world, the first element γ_i^0 of each hash chain γ_i is generated at random and the hash function is queried to generate the rest of the hash chain, i.e.,

$$s_i = \gamma_i^0 \xleftarrow{\$} \{0,1\}^n, \gamma_i^j = h^j(\gamma_i^0) \ ; \ p_i = \gamma_i^{w-1} = h^{w-1}(\gamma_i^0); \ j = 0, \ldots, w-1; \ i = 1, \ldots, l.$$

Here, the random oracle is implemented at random, i.e. $H = h$, the Sign oracle uses the secret key sk defined above.

Intermediate world 1. Here, the first hash chain element is generated at random, the following hash chain elements are successively sampled uniformly except for the collision tuples. That is

$$s_i = \gamma_i^0 \xleftarrow{\$} \{0,1\}^n \ ; \ \gamma_i^1 \text{ is uniform except for the case where } \gamma_i^1 = \gamma_{i'}^1 \text{ if } \gamma_i^0 = \gamma_{i'}^0,$$

$$\gamma_i^2 \text{ is uniform except for the cases where } \gamma_i^2 = \gamma_{i'}^1 \text{ if } \gamma_i^1 = \gamma_{i'}^0 \ ; \ \gamma_{i'}^2 = \gamma_i^2 \text{ if } \gamma_i^1 = \gamma_{i'}^1, \cdots$$

This world is very similar to the Real world, the only difference is that here we first sample the secret and public key (hash chain), then we reprogram the random oracle according to the secret and public key that we sampled, i.e.,

$$H(x) = \begin{cases} \gamma_i^{j+1} & \text{if } x = \gamma_i^j \text{ with } j \leq w-2, \\ h(x) & \text{else.} \end{cases}$$

The Sign oracle is the same as in the Real world.

Intermediate world 2. In this world, the hash chain elements γ_i^j are first sampled uniformly at random with possible collision tuples. It means that the γ_i^j are uniformly independent strings. Afterwards, the random oracle is reprogrammed such that it is consistent with the secret and public keys. When the random oracle is queried, it compares the input with the hash chain. If the input is not equal to any of the hash chain elements, the random oracle answers with a random function $\hat{h}$. Otherwise, for each hash chain element the input is equal to, it adds the next hash chain element into the output register. If there are two hash chain elements that are the same, the random oracle adds both next hash chain elements into the output register. More formally,

$$H(x) = \begin{cases} \displaystyle\bigoplus_{\substack{i,j: j \leq w-2 \\ \text{and } \gamma_i^j = x}} \gamma_i^{j+1} & \text{if there exists } (i,j) \text{ such that } \gamma_i^j = x \text{ with } j = 0, \ldots, w-2, \\ h(x) & \text{otherwise.} \end{cases}$$

In this case, the Sign oracle uses the full list of hash chains $(\gamma_i^j)_{i=1,\ldots,l}^{j=0,\ldots,w-1}$ to answer the query with all the hash chain elements consistent with the input.

Quantum independent world. In this world, the hash chain registers $(\Gamma_i^j)_{i=1,\ldots,l}^{j=0,\ldots,w-2}$ are initially prepared in the uniform superposition $|\Phi\rangle$, and the last hash chain elements $(\gamma_i^{w-1})_{i=1,\ldots,l}$ are sampled uniformly at random. The random oracle is constructed in such a way that it is compatible with the hash chain. When queried with register X and Y, the random oracle compares the X and Γ registers, then answers the query in the Y register. Abstractly speaking, H is implemented as in the Intermediate world 2, except that the comparison and XOR operations involving γ_i^j are replaced by controlled unitary operations with Γ as the control register. Here, X is input register of length n, Y is the output register which is of length n, and Γ represents the hash chain register and is of length lw.

To be more specific, let us describe in detail the behavior of the random oracle. The following definitions are for the Winternitz OTS, with the Lamport OTS being a special case. For each $i \in \{1, \ldots, l\}$ and $j \in \{0, \ldots, w-2\}$, let U_i^j be the following unitary that compares the input register X with the hash chain register Γ_i^j and places the contents of the subsequent register Γ_i^{j+1} in Y if they are equal:

$$(U_i^j)_{XY\Gamma_i^j\Gamma_i^{j+1}} = P^{=}_{X\Gamma_i^j} \otimes (\mathrm{CNOT}^{\otimes n})_{\Gamma_i^{j+1}:Y} + P^{\neq}_{X\Gamma_i^j} \otimes \mathbb{1}_{\Gamma_i^{j+1}Y}, \tag{16}$$

where the controlled-NOT gates use Γ_i^{j+1} as control and Y as target, and the projectors $P_=$ and $P_{\neq}$ check whether the input register X is equal to the corresponding hash chain register Γ_i^j:

$$P^{=}_{X\Gamma_i^j} = \sum_{x \in \{0,1\}^n} |x\rangle\langle x|_X \otimes |x\rangle\langle x|_{\Gamma_i^j}, \qquad P^{\neq}_{X\Gamma_i^j} = \mathbb{1} - P^{=}_{X\Gamma_i^j}. \tag{17}$$

Combining these equations, we can equivalently write

$$(U_i^j)_{XY\Gamma_i^j\Gamma_i^{j+1}} = P^{=}_{X\Gamma_i^j} \otimes \left((\mathrm{CNOT}^{\otimes n})_{\Gamma_i^{j+1}:Y} - \mathbb{1} \right) + \mathbb{1}, \tag{18}$$

In case $j+1 = w-1$, the above definition of U_i^j still applies in the sense that we can take Γ_i^{w-1} to be the register that stores the i-th block of the public key.

The overall unitary that is applied upon a hash query is

$$(U_h)_{XY\Gamma} = \left(\prod_{i=1}^{l} \prod_{j=0}^{w-2} (U_i^j)_{XY\Gamma_i^j\Gamma_i^{j+1}} \right) U^{\neq}_{XY\Gamma} \tag{19}$$

where the unitary $U^{\neq}_{XY\Gamma}$ corresponds to the case where the input x is not equal to any part of the hash chain register Γ:

$$\begin{aligned} U^{\neq}_{XY\Gamma} &= P^{\neq}_{X\Gamma} U'_{XY} + \left(\mathbb{1}_{X\Gamma} - P^{\neq}_{X\Gamma} \right) \otimes \mathbb{1}_Y & (20) \\ &= P^{\neq}_{X\Gamma} \cdot \left(U'_{XY} - \mathbb{1} \right) + \mathbb{1} & (21) \end{aligned}$$

where, with a slight abuse of notation,

$$P^{\neq}_{X\Gamma} = \prod_{i=1}^{l} \prod_{j=0}^{w-2} P^{\neq}_{X\Gamma_i^j} \tag{22}$$

denotes the projector onto the subspace of $X\Gamma$ where X is not equal to any of the Γ_i^j registers, and U'_{XY} is the standard random oracle unitary that answers the query by XOR-ing the hash value $h(x)$ in the Y register regardless of the entire hash chain register:

$$U'_{XY}|x\rangle_X|y\rangle_Y = |x\rangle_X|y \oplus h(x)\rangle_Y. \tag{23}$$

The conditions on the control registers in eq. (19) are such that, for any input x, only one of the unitaries in the product is applied.

3.1.1 Additional details for the Lamport and Winternitz OTS

As mentioned, the above definitions specialize to the Lamport OTS. Here, the pair of indices (i, j), specifying a message bit's position and its value, replace the index i in the Winternitz setting, and the hash chains have only length two, with the first ($j = 0$ above), and second ($j = 1$ above) elements given by the secret key strings, and public key strings, respectively.

Lastly, we describe the behavior of the blinded signing oracle. We define the action of the oracle for inputs where the register M is in a computational basis state $|m\rangle$, which is sufficient by the linearity of the quantum oracle for the blinded signing function.

When queried with register M and Σ, the signing oracle controls on the message m not being in the blinding set B and answers the query by XOR-ing the signature into the Σ register. For ease of notation, let Γ_i^{w-1} be registers prepared in state $|p_i\rangle$ for $i = 1, \ldots, l$. Then, for a fixed message m, the signing oracle for the Winternitz OTS operates as follows:

$$B\,\mathsf{Sign}_{\mathsf{sk}}\,|m\rangle_M = \begin{cases} |m\rangle_M \otimes \mathbb{1} & \text{if } m \in B, \\ |m\rangle_M \otimes \left(\bigotimes_{i=1}^{l} \mathrm{CNOT}^{\otimes n}_{\Gamma_i^{b_i}:\Sigma_i} \right) & \text{otherwise.} \end{cases}$$

For the Lamport OTS, in the Real world, when queried with an input m of length l, the signing oracle answers the query with an l n-secret keys stings corresponding to each bit of the message input. In contrast, in the $\mathsf{Quantum\ independent\ world}$, the queried message is a quantum state. Thus the signing oracle answers the query by XOR-ing the corresponding secret key sub-registers in the output register Y. Specifically, for a fixed message m, the signing oracle acts as follows:

$$B\,\mathsf{Sign}_{\mathsf{sk}}\,|m\rangle_M = \begin{cases} |m\rangle_M \otimes \mathbb{1} & \text{if } m \in B, \\ |m\rangle_M \otimes \left(\bigotimes_{i=1}^{l} \mathrm{CNOT}^{\otimes n}_{S_i^{m_i}:\Sigma} \right) & \text{otherwise.} \end{cases}$$

3.2 Indistinguishability

Finally, we prove a number of lemmas which together allow us to conclude the indistinguishability of the Real world and the Quantum independent world.

Lemma 4. *Let p and q be the output distributions of an algorithm $\mathcal{A}$ interacting with the* Real world *and the* Quantum independent world*, respectively. Then*

$$\|p - q\|_1 \leq \frac{3(wl)^2}{2^n}. \tag{24}$$

We prove Lemma 4 via a sequence of lemmas.

Lemma 5. *The* Real world *and the* Intermediate world 1 *are indistinguishable.*

Lemma 6. *The distribution p and q of hash chains in the* Intermediate worlds 1 *and* 2 *are close:*

$$\|p - q\|_1 \leq \frac{3(wl)^2}{2^n}. \tag{25}$$

Proof. Let p and q be hash chain probability distributions corresponding to the Intermediate worlds 1 and 2, respectively:

$$\begin{aligned} &p : \underline{\gamma_i^0} \xleftarrow{\$} \{0,1\}^n, \quad i = 1, \ldots, l, \qquad \underline{\gamma_i^j} = H^j(\underline{\gamma_i^0}), \quad \underline{\gamma} = (\gamma_i^j)_{i=1,\ldots,l}^{j=0,\ldots,w-1}, \text{ and} \\ &q : \underline{\gamma_i^j} \xleftarrow{\$} \{0,1\}^n, \quad i = 1, \ldots, l, \quad j = 0, \ldots, w-1 \end{aligned}$$

We want to show that p and q are close and that the probability that collisions occur in both distributions is small. Let $C \subset (\{0,1\}^n)^{lw}$ denote the subset of tuples containing a collision, i.e. $\gamma \in C$ iff there exist i, j, i', j' such that $\gamma_i^j = \gamma_{i'}^{j'}$. One can easily check that p and q are equal, conditioned on the subset C^c of collision-free tuples. Then the total variation distance between p and q is

$$\|p - q\|_1 \leq p(C) + q(C) + \max\{p(C), q(C)\}.$$

Given that we can easily compute the probability of collision-free tuples in both distributions, we can first compute $p(C^c)$ and $q(C^c)$, and deduce the probabilities $p(C)$ and $q(C)$.

In the distribution q, the γ_i^j are independent random n-bit strings. Thus,

$$q(C^c) = \binom{2^n}{lw}(lw)!2^{-nlw} = \frac{(2^n)!2^{-nlw}}{(2^n - lw)!} = 2^n(2^n - 1)\cdots(2^n - lw + 1)2^{-nlw}. \tag{26}$$

Now, let us find a lower bound of the probability of collision-free tuples in the distribution q so that we can have an insight on the range of possible collisions occurring in q. We know that

$$2^n(2^n - 1)\cdots(2^n - lw + 1) \geq (2^n - wl)^{wl} = 2^{nwl}\left(1 - \frac{wl}{2^n}\right)^{wl}.$$

But, setting $f(x) = (1 - x)^{wl}$ with $x = \frac{wl}{2^n}$, one can easily see that f is convex and differentiable on the interval $[0, 1]$. In particular, f is differentiable at 0 and $f(x) \geq xf'(0) + f(0) = -\frac{(wl)^2}{2^n} + 1$. Thus,

$$(2^n - wl)^{wl} = 2^{nwl}\left(1 - \frac{wl}{2^n}\right)^{wl} = 2^{nwl}f(wl/2^n) \geq 2^{nwl}\left(1 - \frac{(wl)^2}{2^n}\right).$$

Hence

$$q(C^c) \geq 1 - \frac{(wl)^2}{2^n}, \qquad\qquad q(C) \leq \frac{(wl)^2}{2^n}.$$

Next, we compute $p(C)$. Here we first compute $p(C^c)$ as well. To derive the probability of collision-free tuples, we can sample the elements of the distribution one by one, starting with γ_1^0 and choosing the γ_i^0 before moving onto γ_1^1, etc. The crucial observation is that, when conditioning on collision-free tuples, each γ_i^j is choosen uniformly from the set of strings that have not yet occurred. We hence get the same probability for the set C^c as for the uniform distribution q:

$$p(C^c) = q(C^c) \geq 1 - \frac{(wl)^2}{2^n}, \qquad\qquad p(C) = q(C) \leq \frac{(wl)^2}{2^n}.$$

Therefore, the probability of collision occurring in both distributions can be at most $(wl)^2/2^n$ which is negligible because w is constant, l is polynomial in n, and n quite large.

Plugging the probabilities $p(C)$and $q(C)$ of collisions occurring in both distributions into our above expression of total variation distance give

$$\|p - q\|_1 \leq p(C) + q(C) + \max\{p(C), q(C)\} \leq \frac{3(wl)^2}{2^n} \tag{27}$$

as desired. □

Lemma 7. *The way of implementing the random oracle in the* Intermediate world 2 *and in the* Quantum independent world *are indistinguishable.*

Proof. In the Intermediate world 2, the random oracle contains the hash chain, so that when queried, it compares the input with the hash chain and answers the query in the output register. Specifically, the hash chain is only used as a comparison tool, so it is not modified. Similarly, in quantum setting the random oracle is implemented in such a way that it contains the hash chain register. When it receives a query, it controls the hash chain register to see whether there is similitude between the hash chain and the queried input, and answers the query by acting on the output register. Those controlled operations commute with computational basis measurements. From this fact it is easy to see that Intermediate world 2 and Quantum independent world are exactly indistinguishable, see [Zha15] for details. □

Proof of Lemma 4. Lemma 4 follows directly from Lemmas 5 to 7. □

Next, we prove the security of the Lamport and Winternitz OTS from Sections 2.3.1 and 2.3.2, respectively, in the case where adversaries are granted both quantum access to the signing oracle and random oracle.

4 One-time BU security of the Lamport OTS

In the BlindForge experiment, the adversary is granted both quantum access to a blinding signing oracle for the digital signature scheme and a the random oracle. For one-time signature schemes, the adversary is allowed only to query the signing oracle at most once. So, to produce a forged message-signature pair, the adversary can make a desired number of quantum queries to the random oracle, then query the signing oracle once, and then query again the random oracle as many times as desired. Our goal is to prove that the success probability of any efficient adversary in producing a valid fresh forged message-signature pair is small. Equivalently, we want to show that the probability that an adversary outputs a correct forged signature on a valid forged message is negligible.

In Lamport OTS, the signature algorithm uses only half of the secret key to produce the signature, and the unused part constitutes the invariant of the secret key. Classically, the property that enables security is that the adversary does not have any information about the invariant of the secret key. Quantumly, our intuition is that since in the BlindForge experiment the forged message must be outside the queried region, for any queried message, there exists at least one bit in which the forged and queried messages are different. Thus, the secret key corresponding to that specific bit should still be in its initial state, hence in the invariant of the secret key. Therefore, we want to show that regardless of the number of queries to the random oracle and to the blinded signing oracle, no adversary can win the BlindForge experiment except with negligible probability. Towards that end, we separately analyze three cases: *hash queries before* Sign *query*, Sign query and *hash queries after* Sign *query*.

We start by describing the overall strategy that we will use to achieve our goal. For *hash queries before* Sign *query*, we know that before any query the entire secret key is in uniform superposition state, thus we define a projector of the secret key register being in uniform superposition, and we establish that this projector approximately commutes with the random oracle unitary. This means that after a moderate number of queries, the secret key registers will still be in uniform superposition. The interpretation of this fact is that the adversary learns almost no information about the secret key.

In the Sign query case, the first step is to track the unused part of the secret key. This part can be easily determined in the classical setting since the adversary queries only one message in each query. In contrast, in the quantum setting, since we are looking at quantum queries we have to track the invariant in superposition over the different queried messages. This is difficult because the invariant is different within each term of the superposition, so we cannot simply describe the invariant for the whole state. To address this problem, we perform a partial measurement that tracks the unused part of the secret key register. Then, we show that for any forgery pair, the outcome where *none of the secret key registers relevant to the forged signature is in the invariant* can never occur. Next, we use this result to show that if there are no hash queries, no adversary can produce a valid forgery pair except with small probability. If there are hash queries before the Sign query, then we define the invariant projector that tracks the invariant of the secret keys after queries and show that this projector is orthogonal to the projector corresponding to the outcome where *none of the secret key registers relevant to the forged signature belong to the invariant*. Then, we show that if there is only Sign query, this new projector does not change the adversary state immediately after the signature. We also establish that if the adversary state after producing forgery is in the range of this new projector, then the adversary has negligible probability to win the BlindForge game. Besides, we prove that the new projector approximately commutes with the random oracle unitary.

Finally, for the case of *hash queries after* Sign *query*, we use the latter argument of the commutator to prove that after hash queries the final adversary state remains roughly in the image of the invariant projector of the secret key. This implies that *hash queries after* Sign *query* do not help the adversary to get relevant information about the secret key. Those results show that even with hash queries before and after the Sign query, any query-limited adversary has only small probability to win the BlindForge experiment.

In the remainder of this section, we prove the following theorem.

Theorem 3. *The Lamport OTS in Section 2.3.1 is 1-*BU *secure if the hash function h is modeled as a quantum-accessible random oracle. More precisely, let $\mathcal{A}$ be an adversary that plays the* BlindForge *game for the Lamport OTS, making a total of q queries to the random oracle. Then $\mathcal{A}$ succeeds with a probability bounded as*

$$\begin{aligned}\Pr[\mathcal{A} \text{ wins } \mathsf{BlindForge}] &\leq l^2 \cdot 2^{-n} \left(3137q^2(l+1)+12\right)\\ &\leq 6286q^2l^3 \cdot 2^{-n}\end{aligned} \tag{28}$$

where n is the security parameter of the Lamport OTS, l is the message length, and the simplified bound in the last line holds for $q > 0$.

The proof of Theorem 3 is presented in steps in the following subsections. In particular, we prove Theorem 3 in the Quantum independent world first, and conclude the statement in the Real world via an application of Lemma 4. In the remainder of the article, we use a subscript QI to indicate that a probability statement holds in the Quantum independent world.

We begin by presenting some concepts and tools which will be used in the proof. Subsequently, we prove the steps outlined above as separate lemmas. Finally, we combine them to prove Theorem 3.

4.1 Q measurement for Lamport OTS

Recall the superposition hash chain formalism from Section 3, in particular the special case of the Lamport OTS key generation discussed in Section 3.1.1. Our proof will make use of a projective measurement to track an invariant on the secret key register for the verification of the forged message in the case of no hash queries. Let (m^*, σ^*) be a forged message-signature pair with $\sigma^* = s_1^{m_1^*} \cdots s_l^{m_l^*}$, where $(s_i^j)_{i=1,\dots,l}^{j=0,1}$ is the secret key and l is the message length.

For any message $m^* \in \{0,1\}^l$ let us define an $(l+1)$-outcome projective measurement $\{Q_1^{m^*}, \dots, Q_{l+1}^{m^*}\}$ acting on the secret key registers $(S_i^j)_{i=1,\dots,l}^{j=0,1}$. It finds the smallest index $i^* \in \{1,\dots,l\}$ for which the register $S_{i^*}^{m_{i^*}^*}$ *is in uniform superposition*, or determines that *none of the relevant secret key registers are in uniform superposition* (this corresponds to the outcome $l+1$). The projectors $Q_i^{m^*}$ are defined in terms of projectors

$$\Phi = |\Phi\rangle\langle\Phi|, \qquad\qquad \Phi^\perp = \mathbb{1} - |\Phi\rangle\langle\Phi| \tag{29}$$

that correspond to the uniform superposition $|\Phi\rangle$ and its orthogonal complement. We place them onto different registers depending on the message m^*:

$$Q_{i^*}^{m^*} = \Phi^\perp_{S_1^{m_1^*}} \otimes \cdots \otimes \Phi^\perp_{S_{i^*-1}^{m_{i^*-1}^*}} \otimes \Phi_{S_{i^*}^{m_{i^*}^*}}, \qquad\qquad Q_{l+1}^{m^*} = \bigotimes_{i=1}^{l} \Phi^\perp_{S_i^{m_i^*}} \tag{30}$$

where $i^* \in \{1,\dots,l\}$. These operators act as $\mathbb{1}$ on all other registers S_i^j that are not specified.

4.2 Invariant projector

In the hash queries part of our proof, we will need a separate projector P_S to track an invariant of the secret key register. In this section we define this projector and state its several properties as lemmas.

Let $\alpha = (\alpha_i^j)_{i=1,\dots,l}^{j=0,1}$ be a $2l$-bit string whose each bit $\alpha_i^j \in \{0,1\}$ indicates that the projector $\Phi(\alpha_i^j)$ is applied on the corresponding secret key register S_i^j where

$$\Phi(0) = \Phi, \qquad\qquad \Phi(1) = \Phi^{\perp}. \tag{31}$$

For each string α, we define the associated projector $\Phi(\alpha)$ on the whole secret key register S as

$$\Phi(\alpha)_S = \bigotimes_{i=1}^{l}\bigotimes_{j=0}^{1}\Phi(\alpha_i^j)_{S_i^j}. \tag{32}$$

Note that this is a complete set of projectors, i.e., $\sum_{\alpha\in\{0,1\}^{2l}}\Phi(\alpha)_S = \mathbb{1}_S$.

Since we are interested in the unused part of the secret key register S, we need to filter those α's for which S_i^j is in state $|\Phi\rangle$. Recall from our discussion of blind unforgeability in Section 2.4 that B denotes the set of blinded messages. Since the blinded signing oracle has signed (at most) a single, un-blinded message, the state after the signing oracle call can be written as a superposition of states where, for some un-blinded message $m \in B^c$, the secret key register of the complementary value $\bar{m}_i$ is still in the uniform superposition $|\Phi\rangle$, for all i. We collect all strings α that are consistent with no blinded messages having been signed in

$$\widehat{B^c} = \bigcup_{m\in B^c}\left\{\alpha \in \{0,1\}^{2l} \;\middle|\; \alpha_i^{\bar{m}_i} = 0 \text{ for all } i = 1,\dots,l\right\}. \tag{33}$$

These strings indicate which secret key registers were not used during hash queries and Sign query. Finally, we define

$$P_S = \sum_{\alpha\in\widehat{B^c}}\Phi(\alpha)_S \tag{34}$$

as the projector on the subspace compatible with $\widehat{B^c}$. Note that P_S is indeed a projector since it is a sum of mutually orthogonal projectors.

Now, we state some ingredients that we will need to prove our main results both for Lamport OTS and Winternitz OTS. Proofs of these lemmas are provided in Appendix A.

Lemma 8. *Let U_h be the random oracle unitary for any given function h (see Section 3) and let $\Phi = |\Phi\rangle\langle\Phi|$ denote the projector onto the uniform superposition. Then, for any $i \in \{1,\dots,l\}$ and $j \in \{0,1\}$,*

$$\left\|\left[U_h, \Phi_{S_i^j}\right]\right\|_\infty \le \frac{6}{2^{n/2}} = \epsilon_L(n) \tag{35}$$

is negligible in n.

For any message in the blinding set B, there exists at least one secret key necessary for its corresponding signature in the invariant of the secret key register. In other words, for any valid forged message, at least one of the secret keys needed for its corresponding forged signature is in the uniform superposition state $|\Phi\rangle$. This implies the following lemma

Lemma 9. *For all $m^* \in B$, the projectors $Q_{l+1}^{m^*}$ and P_S defined in eqs. (30) and (34) are orthogonal:*

$$Q_{l+1}^{m^*}P_S = 0. \tag{36}$$

Lemma 10. *Let $B\,\mathsf{Sign}_{\mathsf{sk}}$ be the blinded signing oracle for the Lamport OTS and let $|\psi_0\rangle$ be the adversary's state before the* Sign *query. If there are no hash queries, then after making at most one* Sign *query the adversary's state $|\psi_1\rangle = B\,\mathsf{Sign}_{\mathsf{sk}}|\psi_0\rangle$ is completely in the image of the invariant projector P_S defined in eq. (34). That is,*

$$P_S B\,\mathsf{Sign}_{\mathsf{sk}}|\psi_0\rangle = B\,\mathsf{Sign}_{\mathsf{sk}}|\psi_0\rangle. \tag{37}$$

Lemma 11. *The invariant projector P_S defined in eq. (34) and the random oracle unitary U_h in the* Quantum independent world, *see eq. (19), approximately commute, i.e.,*

$$\left\|[U_h, P_S]\right\|_\infty \le \delta_L(n), \tag{38}$$

where

$$\delta_L(n) = \frac{32l}{2^{n/2}}$$

is negligible in n.

Next, we use the above lemmas to prove our main results. In the following sections, we analyze the situation where the adversary makes q_0 hash queries before the Sign query and q_1 hash queries after, maximizing the resulting bound under the condition that $q_0 + q_1 = q$.

4.3 Hash queries before Sign query

In this section, we study the impact of hash queries before Sign query on the secret key register S. Our main goal is to show that, for a moderate number of queries to the random oracle, no adversary can learn a significant amount of information about the secret key. Therefore, she cannot produce a valid forgery except with small probability.

Register	Meaning
X	adversary's input
Y	adversary's output
M	Sign query input
Σ	Sign query output
E	adversary's internal workspace
S	secret key

Table 1: Registers used in the analysis.

Let $|\psi\rangle_{XYM\Sigma E}$ be adversary's initial state before any queries (see table 1 for a summary of registers and their roles). Before any query is performed, the whole secret key register S is in the uniform superposition state $|\Phi\rangle^{\otimes 2l}$. Assume the adversary $\mathcal{A}_0$ queries the random oracle q_0 times before querying the signing oracle. If V^i_{XYE} denotes the unitary she performs after the i-th query, the final adversary state after q_0 hash queries is

$$|\psi_0\rangle_{XYM\Sigma ES} = V^{q_0}_{XYE}(U_h)_{XYS}V^{q_0-1}_{XYE}\cdots V^2_{XYE}(U_h)_{XYS}V^1_{XYE}(U_h)_{XYS}|\psi\rangle_{XYM\Sigma E}|\Phi\rangle_S^{\otimes 2l} \tag{39}$$

where U_h is the random oracle unitary used to answer hash queries. The following lemma shows that the secret key registers of this state are still close to being in uniform superposition.

Lemma 12. *In the* Quantum independent world, *without querying the B* Sign *oracle, hash queries leave the state of the secret key registers approximately unchanged:*

$$\left\| \Phi_S^{\otimes 2l}|\psi_0\rangle_{XYM\Sigma ES} - |\psi_0\rangle_{XYM\Sigma ES} \right\|_2 \le 2lq_0\epsilon_L(n).$$

Proof. We want to show that after q_0 hash queries, the state of the secret key register S is still approximately in the uniform superposition state $|\Phi\rangle^{\otimes 2l}$. Let us abbreviate the overall unitary in eq. (39) by W_{XYES}. Since the only operations in W_{XYES} that act on the S register are the hash queries U_h, and they are in fact controlled by the S register, we have $W_{XYES}\Phi_S^{\otimes 2l} = W_{XYES}$. Using this, we get

$$\begin{aligned}
&\left\| \Phi_S^{\otimes 2l}|\psi_0\rangle_{XYM\Sigma ES} - |\psi_0\rangle_{XYM\Sigma ES} \right\|_2 \\
&= \left\| \Phi_S^{\otimes 2l} W_{XYES}|\psi\rangle_{XYM\Sigma E}|\Phi\rangle_S^{\otimes 2l} - W_{XYES}\Phi_S^{\otimes 2l}|\psi\rangle_{XYM\Sigma E}|\Phi\rangle_S^{\otimes 2l} \right\|_2 && (40)\\
&= \left\| \left[\Phi_S^{\otimes 2l}, W_{XYES}\right]|\psi\rangle_{XYM\Sigma E}|\Phi\rangle_S^{\otimes 2l} \right\|_2 && (41)\\
&\le \left\| \left[\Phi_S^{\otimes 2l}, W_{XYES}\right] \right\|_\infty \underbrace{\left\| |\psi\rangle_{XYM\Sigma E}|\Phi\rangle_S^{\otimes 2l} \right\|_2}_{=1} && (42)\\
&= \left\| \left[\Phi_S^{\otimes 2l}, V^{q_0}_{XYE}(U_h)_{XYS}V^{q_0-1}_{XYE}\cdots V^2_{XYE}(U_h)_{XYS}V^1_{XYE}(U_h)_{XYS}\right] \right\|_\infty && (43)\\
&\le q_0\left\| \left[\Phi_S^{\otimes 2l}, (U_h)_{XYS}\right] \right\|_\infty + \sum_{i=1}^{q_0}\left\| \left[\Phi_S^{\otimes 2l}, V^i_{XYE}\right] \right\|_\infty, && (44)
\end{aligned}$$

where eq. (42) follows from the definition of the operator norm and the last inequality follows from Lemma 2.

The first term in eq. (44) can be bounded as follows:

$$\left\| \left[\Phi_S^{\otimes 2l}, (U_h)_{XYS}\right] \right\|_\infty \le \sum_{\substack{i\in\{1,\ldots,l\}\\ j\in\{0,1\}}} \left\| \left[\Phi_{S_i^j}, (U_h)_{XYS}\right] \right\|_\infty \le 2l\epsilon_L(n),$$

which follows by first applying Lemma 2 and then Lemma 8. Since $\Phi_S^{\otimes 2l}$ and V^i_{XYE} act on different registers, they commute and the second term in eq. (44) vanishes. Hence

$$\left\| \Phi_S^{\otimes 2l}|\psi_0\rangle_{XYM\Sigma ES} - |\psi_0\rangle_{XYM\Sigma ES} \right\|_2 \le 2lq_0\epsilon_L(n) \tag{45}$$

as desired. □

Recall from eq. (35) that $\epsilon_L(n) = 4/2^{n/2}$ is negligible in n. Since l is constant, the magnitude of $2lq_0\epsilon_L(n)$ is determined only by the number of queries q_0. The bound in eq. (45) is negligible for any adversary making 2^{cn} queries to the random oracle when $c < 1/2$. Therefore $\Phi_S^{\otimes 2l}|\psi_0\rangle_{XYM\Sigma ES}$ and $|\psi_0\rangle_{XYM\Sigma ES}$ are close, which means that hash queries before Sign query do not significantly change the secret key register. Equivalently, it means that the adversary learns almost no information about the secret key.

4.4 Query to the signing oracle

Now that we have control over the advantage an adversary can gain from making hash queries before the sign query, we need to analyze the possible advantage from hash queries after the sign query and bound the overall success probability using Lemma 12.

A crucial property of the Lamport OTS when analyzing classical security is that for all messages m that have not been queried, there exists an index j such that $s_j^{m_j}$ is hidden from the adversary by the one-wayness of the used hash function. In blind-unforgeability (for classical adversaries), this property holds for all *blinded messages*. In the setting of quantum queries, we have to track this property in superposition while the adversary is making hash queries after the sign query. As this is complicated by the "for all"-quantifier, we begin by analyzing the case where the adversary makes no hash queries after the sign query to ease the reader into our proof technique.

The discussion in this section does not concern the random oracle, so we absorb the random oracle query registers XY into E for the purpose of this section. In the 1-BlindForge game, an adversary $\mathcal{A}$ is allowed to query the Sign-oracle at most once to produce a valid forged message-signature pair (m^*, σ^*). To analyze the interaction between $\mathcal{A}$ and the signing oracle, we will break it into the following steps:

$$|\psi_0\rangle_{M\Sigma BES} \xmapsto{B\,\mathsf{Sign}_{\mathrm{sk}}} |\psi_1\rangle_{M\Sigma BES} \xmapsto{U_{M\Sigma E}} |\psi_2\rangle_{M\Sigma BES} \xmapsto{\langle m^*|_M} |\psi_3(m^*)\rangle_{\Sigma BES} \xmapsto{\langle \sigma^*|_\Sigma} |\psi_4(m^*, \sigma^*)\rangle_{BES}. \tag{46}$$

They correspond to applying the Sign-oracle and an arbitrary unitary $U_{M\Sigma E}$, followed by measuring the message and signature registers M and Σ. Let us now analyze these steps in more detail and write down the corresponding quantum states.

First, $\mathcal{A}$ prepares her input state as an arbitrary superposition of messages:

$$|\psi_0\rangle_{M\Sigma BES} = \left(\sum_{m \in \{0,1\}^l} \sum_{\sigma \in (\{0,1\}^n)^l} \sum_{b \in \{0,1\}} \kappa_{m\sigma b} |m\rangle_M |\sigma\rangle_\Sigma |b\rangle_B |\alpha_{m\sigma b}\rangle_E \right) \otimes \left(|\Phi\rangle^{\otimes 2l} \right)_S \tag{47}$$

where the B will indicate whether the message is blinded or not ($|1\rangle_B$ for blinded and $|0\rangle_B$ for un-blinded).

Next, the adversary supplies this to the Sign oracle which produces the following signed state:

$$|\psi_1\rangle_{M\Sigma BES} = B\,\mathsf{Sign}_{\mathrm{sk}}\,|\psi_0\rangle_{M\Sigma BES} = |\psi_1^1\rangle_{M\Sigma BES} + |\psi_1^0\rangle_{M\Sigma BES} \tag{48}$$

where the superscripts 1 and 0 refer to blinded (B) and un-blinded (B^c) messages, respectively:

$$\begin{aligned} |\psi_1^1\rangle_{M\Sigma BES} &= \sum_{m \in B} \sum_{\sigma \in (\{0,1\}^n)^l} \kappa_{m\sigma 1} |m\rangle_M |\sigma\rangle_\Sigma |1\rangle_B |\alpha_{m\sigma 1}\rangle_E |\Phi\rangle_S^{\otimes 2l}, \\ |\psi_1^0\rangle_{M\Sigma BES} &= \sum_{m \in B^c} \sum_{\sigma \in (\{0,1\}^n)^l} \frac{1}{2^{nl/2}} \sum_{s \in (\{0,1\}^n)^l} \kappa_{m\sigma 0} |m\rangle_M |\sigma \oplus s\rangle_\Sigma |0\rangle_B |\alpha_{m\sigma 0}\rangle_E |\Omega(s,m)\rangle_S, \end{aligned}$$

where $m = m_1 \dots m_l$, $\sigma = \sigma_1 \dots \sigma_l$, and

$$|\Omega(s,m)\rangle_S = |s_1^{m_1}\rangle_{S_1^{m_1}} \cdots |s_l^{m_l}\rangle_{S_l^{m_l}} |\Phi\rangle_{S_1^{\bar{m}_1}} \cdots |\Phi\rangle_{S_l^{\bar{m}_l}}. \tag{49}$$

Once the adversary $\mathcal{A}$ gets the signed state $|\psi_1\rangle_{M\Sigma BES}$, she performs some operations with the intention of producing a forgery message m^*. Intuitively, those operations can be considered as applying some arbitrary unitary $U_{M\Sigma E}$ to $|\psi_1\rangle_{M\Sigma BES}$. Let us denote the resulting state by

$$|\psi_2\rangle_{M\Sigma BES} = U_{M\Sigma E} |\psi_1\rangle_{M\Sigma BES}.$$

Then $\mathcal{A}$ measures the message register M, which yields outcome $m^* \in \{0,1\}^l$. After the measurement, the state $|\psi_2\rangle_{M\Sigma BES}$ collapses to the (unnormalized) state

$$\begin{aligned} |\psi_3(m^*)\rangle_{\Sigma BES} &= \langle m^*|_M |\psi_2^1\rangle_{M\Sigma BES} + \langle m^*|_M |\psi_2^0\rangle_{M\Sigma BES} \\ &= |\psi_3^1(m^*)\rangle_{\Sigma BES} + |\psi_3^0(m^*)\rangle_{\Sigma BES} \end{aligned}$$

where

$$|\psi_3^1(m^*)\rangle_{\Sigma BES} = \sum_{m\in B}\sum_{\sigma\in(\{0,1\}^n)^l} \kappa_{m\sigma 1}\langle m^*|_M U_{M\Sigma E}|m\rangle_M|\sigma\rangle_\Sigma|1\rangle_B|\alpha_{m\sigma 1}\rangle_E|\Phi\rangle_S^{\otimes 2l},$$

$$|\psi_3^0(m^*)\rangle_{\Sigma BES} = \sum_{m\in B^c}\sum_{\sigma\in(\{0,1\}^n)^l} \frac{1}{2^{nl/2}}\sum_{s\in(\{0,1\}^n)^l} \kappa_{m\sigma 0}\langle m^*|_M U_{M\Sigma E}|m\rangle_M|\sigma\oplus s\rangle_\Sigma|0\rangle_B|\alpha_{m\sigma 0}\rangle_E|\Omega(s,m)\rangle_S.$$

Having obtained m^*, the purpose of the adversary $\mathcal{A}$ is to produce a forged signature σ^* that corresponds to m^*. To that end, she measures the signature register Σ of $|\psi_3(m^*)\rangle_{\Sigma BES}$, getting outcome $\sigma^* \in (\{0,1\}^n)^l$. The (unnormalized) post-measurement state is

$$\begin{aligned}|\psi_4(m^*,\sigma^*)\rangle_{BES} &= \langle\sigma^*|_\Sigma|\psi_3^1(m^*)\rangle_{\Sigma BES} + \langle\sigma^*|_\Sigma|\psi_3^0(m^*)\rangle_{\Sigma BES} \\ &= |\psi_4^1(m^*,\sigma^*)\rangle_{BES} + |\psi_4^0(m^*,\sigma^*)\rangle_{BES}\end{aligned} \tag{50}$$

where

$$|\psi_4^1(m^*,\sigma^*)\rangle_{BES} = \sum_{m\in B}\sum_{\sigma\in(\{0,1\}^n)^l} \kappa_{m\sigma 1}\langle m^*|_M\langle\sigma^*|_\Sigma U_{M\Sigma E}|m\rangle_M|\sigma\rangle_\Sigma|1\rangle_B|\alpha_{m\sigma 1}\rangle_E|\Phi\rangle_S^{\otimes 2l},$$

$$|\psi_4^0(m^*,\sigma^*)\rangle_{BES} = \sum_{m\in B^c}\sum_{\sigma\in(\{0,1\}^n)^l} \frac{1}{2^{nl/2}}\sum_{s\in(\{0,1\}^n)^l} \kappa_{m\sigma 0}\langle m^*|_M\langle\sigma^*|_\Sigma U_{M\Sigma E}|m\rangle_M|\sigma\oplus s\rangle_\Sigma|0\rangle_B|\alpha_{m\sigma 0}\rangle_E|\Omega(s,m)\rangle_S.$$

For the sake of simplicity, let us rewrite $|\psi_4^0\rangle_{BES}$ as follows:

$$|\psi_4^0(m^*,\sigma^*)\rangle_{BES} = \sum_{m\in B^c}\frac{1}{2^{nl/2}}\sum_{s\in(\{0,1\}^n)^l}|\eta(m,s)\rangle_{BE}|\Omega(s,m)\rangle_S \tag{51}$$

where only $|\eta(m,s)\rangle_{BE}$ depends on m^* and σ^*:

$$|\eta(m,s)\rangle_{BE} = \sum_{\sigma\in(\{0,1\}^n)^l}\kappa_{m\sigma 0}\langle m^*|_M\langle\sigma^*|_\Sigma U_{M\Sigma E}|m\rangle_M|\sigma\oplus s\rangle_\Sigma|0\rangle_B|\alpha_{m\sigma 0}\rangle_E.$$

Finally, the adversary $\mathcal{A}$ outputs the forged message-signature pair (m^*,σ^*). The probability of producing this pair is $\||\psi_4^0(m^*,\sigma^*)\rangle_{BES}\|^2$.

The next step is to analyse the probability that $\mathcal{A}$'s forgery candidate (m^*,σ^*) is correct. For that purpose, we consider two cases. The first case, namely when $m^* \notin B$, is trivial since then $\mathcal{A}$ has lost the BlindForge experiment because m^* must be blinded by definition. The rest of this section is devoted to analyzing the second case.

If $m^* \in B$, the forged message m^* has not been signed since the blinded signing oracle signs only unblinded messages. Hence, for any message $m \notin B$, there exists at least one index $i \in \{1,\ldots,l\}$ such that $m_i \neq m_i^*$. This implies that for some index $i^* \in \{1,\ldots,l\}$ the register $S_{i^*}^{m_{i^*}^*}$ has not been used for the signature of the adversary's queried message and is therefore still in the uniform superposition state $|\Phi\rangle$. Note that this holds only in superposition over m. Indeed, i^* depends on m and is in general different for each term of the superposition.

We know that the secret key register S consists of $2l$ n-qubit registers out of which only l are used for the signature procedure while the other l are still in the uniform superposition $|\Phi\rangle$. Despite the secret key being in superposition, we want to track the invariant part of the secret key and show that some of the secret key sub-registers relevant for the forged signature satisfy this invariant and are thus unknown to the adversary, so it is unlikely that the adversary would have used the correct secret key sub-register to produce the forged signature.

For that purpose, we analyze a modified BlindForge experiment, where an additional measurement is performed on the secret key register after the adversary has output their forgery, but before the secret key register is measured to actually sample the secret key as required in the Quantum independent world. This additional measurement was defined in eq. (30) and we will refer to it as the *Q-measurement*. Since it has few outcomes, its effect on the adversary's winning probability is limited and can be bounded by the pinching lemma.

If the Q-measurement yields outcome $i^* \in \{1,\ldots,l\}$, then the secret key sub-register $S_{i^*}^{m_{i^*}}$ is in uniform superposition, and the adversary is bound to fail as σ^* is independent of the secret key string $s_{i^*}^{m_{i^*}}$ (the result of measuring $S_{i^*}^{m_{i^*}}$).

It remains to analyze the outcome $l+1$ that corresponds to the projector $Q_{l+1}^m = (\Phi^\perp)^{\otimes l}$, see eq. (30), where $\Phi^\perp = \mathbb{1} - |\Phi\rangle\langle\Phi|$ is the projector onto the orthogonal complement of $|\Phi\rangle$. The final adversary state after the measurement, see eq. (50), contains both blinded and un-blinded terms. If we apply $\Phi^{\perp}$ to any secret key register of the blinded term $|\psi_4^1(m^*,\sigma^*)\rangle_{BES}$, we get 0 since all secret key sub-registers are in state $|\Phi\rangle$.

For the rest of our analysis, we fix the message m^* and focus on the un-blinded term $|\psi_4^0(m^*,\sigma^*)\rangle_{BES}$. Given that for each $m \notin B$ there is at least one index $i \in \{1,\dots,l\}$ such that $m_i \neq m_i^*$, we define

$$i(m) = \min\{j \in \{1,\dots,l\} \mid m_j \neq m_j^*\}$$

as the smallest index for which $m \neq m^*$. Intuitively, it is the first sub-register of S that still remains in uniform superposition. In the following, let $S(m) := S_1^{m_1} \cdots S_l^{m_l}$ and recall from eqs. (49) and (51) that the un-blinded term is given by

$$|\psi_4^0(m^*,\sigma^*)\rangle_{BES} = \sum_{m \in B^c} \frac{1}{2^{nl/2}} \sum_{\gamma \in (\{0,1\}^n)^l} |\eta(m,\gamma)\rangle_{BE} |s^m\rangle_{S(m)} |\Phi\rangle_{S(\bar{m})}^{\otimes l}. \tag{52}$$

We want to split the first sum into l parts, one for each value of $i(m)$, so that we can easily evaluate $(\Phi^{\perp})_{S(\bar{m})}^{\otimes l} |\psi_4^0(m^*,\sigma^*)\rangle_{BES}$. For that purpose, we define $B_j^c = \{m \in B^c \mid i(m) = j\}$ and note that $\bigcup_{j=1}^{l} B_j^c = B^c$.

We can now rewrite $|\psi_4^0(m^*,\sigma^*)\rangle_{BES}$ as

$$\begin{aligned} |\psi_4^0(m^*,\sigma^*)\rangle_{BES} &= \sum_{j=1}^{l} \sum_{m \in B_j^c} \frac{1}{2^{nl/2}} \sum_{s \in (\{0,1\}^n)^l} |\eta(m,s)\rangle_{BE} |s^m\rangle_{S(m)} |\Phi\rangle_{S(\bar{m})}^{\otimes l} \\ &= \sum_{j=1}^{l} |\hat{\eta}(m^*,\sigma^*,j)\rangle_{BES_{\{(j,m_j^*)\}^c}} |\Phi\rangle_{S_j^{m_j^*}}, \end{aligned} \tag{53}$$

where we absorbed all registers except for $S_j^{m_j^*}$ into the first system. The remaining register $S_j^{m_j^*}$ is still in the uniform superposition $|\Phi\rangle$ since $j = i(m)$ is the smallest index such that $m_j \neq m_j^*$, meaning that $\bar{m}_j = m_j^*$ and thus $S_j^{m_j^*} = S_j^{\bar{m}_j}$. Applying Q_{l+1} onto the l sub-registers $S(m^*)$ of the register S in eq. (53) gives

$$(Q_{l+1}^{m^*})_{S_1^{m_1^*} \cdots S_l^{m_l^*}} |\psi_4^0(m^*,\sigma^*)\rangle_{BES} = (\Phi^{\perp})_{S_1^{m_1^*} \cdots S_l^{m_l^*}}^{\otimes l} \left(\sum_{j=1}^{l} |\hat{\eta}(m^*,\sigma^*,j)\rangle_{BES_{\{(j,m_j^*)\}^c}} |\Phi\rangle_{S_j^{m_j^*}} \right) = 0, \tag{54}$$

which vanishes because, for each j, the register $S_j^{m_j^*}$ is in state $|\Phi\rangle$ and $\Phi^{\perp}|\Phi\rangle = 0$. Hence, the situation where none of the secret key sub-registers relevant for the verification of the forged signature σ^* is in state $|\Phi\rangle$ can ever occur.

Now, we execute the last part of the BlindForge experiment which consists of checking the correctness of the forged signature σ^*. For this purpose, we perform a computational basis measurement on the entire secret key register S to sample the strings s_i^j.

We recall that in the BlindForge experiment, there is no partial measurement. Therefore, we first evaluate the success probability of $\mathcal{A}$ in case of the modified BlindForge experiment (MBF) in which we performed a partial measurement. Afterwards, we use the Pinching lemma (Lemma 1) to deduct the success probability of the adversary in the real BlindForge experiment from the modified experiment.

Knowing that after applying the partial measurement, at least one of the secret key sub-registers relevant to σ^* is still in the state $|\Phi\rangle$, the probability that the adversary $\mathcal{A}$ used the right $s_i^{m_i^*}$ to produce σ^* is $1/2^n$. In addition, given that the state $|\psi_4(m^*,\sigma^*)\rangle_{BES}$ is unnormalized, the success probability of $\mathcal{A}$ in producing a fixed valid forged message-signature pair (m^*,σ^*) after applying the partial measurement is

$$\begin{aligned} &\Pr_{QI,MBF}\big[\text{Success} \wedge \mathcal{A} \text{ outputs } (m^*,\sigma^*)\big] \\ &= \sum_{j=1}^{l} \Pr\big[\text{Success} \wedge \mathcal{A} \text{ outputs } (m^*,\sigma^*) \wedge Q\text{-measurement returns } j\big] \\ &= \frac{1}{2^n} \big\| |\psi_4(m^*,\sigma^*)\rangle_{EBS} \big\|^2. \end{aligned} \tag{55}$$

Here, the sum over the outcomes of the Q-measurement is restricted to $1,\dots,l$ as the outcome $l+1$ never occurs by eq. (54). Since we made an $(l+1)$-outcome partial measurement on the secret key register previously, by the Pinching (see Lemma 1), this partial measurement can only increase the success probability of

$\mathcal{A}$ in the real BlindForge experiment by at most $l+1$. Thus, the probability that the adversary outputs a valid forged message-signature pair (m^*, σ^*) with respect to the real BlindForge experiment is upper bounded by

$$\Pr_{QI,\mathsf{BlindForge}}\left[\text{Success} \wedge \mathcal{A} \text{ outputs } (m^*, \sigma^*)\right] \leq \frac{l+1}{2^n} \left\| |\psi_4(m^*, \sigma^*)\rangle_{BES} \right\|^2. \tag{56}$$

Therefore, the success probability of $\mathcal{A}$ in producing a valid forged message-signature pair (m^*, σ^*) is

$$\Pr_{QI}\left[\mathcal{A} \text{ wins } \mathsf{BlindForge}\right] = \sum_{m^*, \sigma^*} \Pr_{\mathsf{BlindForge}}\left[\text{Success} \wedge \mathcal{A} \text{ outputs } (m^*, \sigma^*)\right] \leq \frac{l+1}{2^n}. \tag{57}$$

We conclude that the same holds in the Real world, up to a difference as permitted by Lemma 4 with $w = 2$,

$$\Pr\left[\mathcal{A} \text{ wins } \mathsf{BlindForge}\right] \leq \frac{l+1}{2^n} + 12l^2 \cdot 2^{-n}. \tag{58}$$

Hence, the success probability of the adversary $\mathcal{A}$ in winning the BlindForge experiment game is at most $(l+1)/2^n$ which is negligible since l is polynomial in n, and n is large enough. We conclude that Sign query does not help the adversary to get significant information about the secret key.

4.5 Hash queries after Sign query

In this section, we analyse the adversary's *hash queries after* Sign *query* to bound the success probability that an adversary with a given number of queries can achieve in the BlindForge game and thus prove Theorem 3. In this case it is not obvious how to track the invariant of the secret key, i.e. the fact that there is at least one unused part of the secret key that is relevant for the forged signature. Therefore we use a special projector P_S defined in eq. (34) that projects onto the subspace of the secret key register that is consistent with a single blinded sign query and no hash queries. If the final adversary state after producing the forgery candidate is in the image of P_S, then the outcome $l+1$ corresponding to the situation when *none of the secret key sub-registers useful for the forged signature is in state* $|\Phi\rangle$ can never occur, according to Lemma 9. We thus want to show that adversary's final state is approximately in the range of P_S.

If there are no hash queries before the Sign query, then from Lemma 10 the adversary state after the Sign query remains completely in the range of P_S, which means that the outcome $l+1$ cannot occur. That is,

$$P_S|\psi_1\rangle = P_S B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle = |\psi_1\rangle$$

where $|\psi_0\rangle$ and $|\psi_1\rangle$ are respectively the adversary state immediately before and after the Sign query.

Now, assuming there are hash queries before the Sign query, since the projector P_S and the random oracle unitary U_h approximately commute by Lemma 11, it follows that hash queries before Sign query give no significant information to the adversary about the invariant of the secret key register.

Suppose there are hash queries after the Sign query and examine in detail what happen in this case. From the previous case, we know that the adversary's state directly after the Sign query is $|\psi_1\rangle_{M\Sigma XYES}$. Just like for hash queries before the Sign query, suppose that the adversary makes q_1 hash queries after querying the signing oracle. Let $(W^i_{XYE})_{i=1,\ldots,q_1}$ be the unitaries applied between hash queries. Then, let

$$|\psi_1'\rangle_{M\Sigma XYES} = (U_h)_{XYS} W^{q_1}_{XYE} (U_h)_{XYS} W^{q_1 - 1}_{XYE} \cdots W^2_{XYE} (U_h)_{XYS} W^1_{XYE} |\psi_1\rangle_{M\Sigma XYES}$$

be the adversary's state after q_1 hash queries and before performing some unitary operations $U_{M\Sigma E}$ on the post hash queried state or any measurement leading to the forgery candidate. For simplicity, we set

$$T = (U_h)_{XYS} W^{q_1}_{XYE} (U_h)_{XYS} W^{q_1 - 1}_{XYE} \cdots W^2_{XYE} (U_h)_{XYS} W^1_{XYE}.$$

We now prove the following lemma.

Lemma 13. *In the* Quantum independent world, *the state right before the adversary's measurement determining the forgery is applied is approximately in the range of P_S, i.e.*

$$\left\| P_S |\psi_1'\rangle_{M\Sigma XYES} - |\psi_1'\rangle_{M\Sigma XYES} \right\|_2 \leq q_1 \delta_L(n) + 4 l q_1 \epsilon_L(n) = q_1(\delta_L(n) + 4l\epsilon_L(n)) \tag{59}$$

Proof. To see how much those hash queries affect the entire secret register, we compute the difference norm between the states $P_S|\psi_1'\rangle_{M\Sigma XYES}$ and $|\psi_1'\rangle_{M\Sigma XYES}$ to see how closed they are. We have:

$$
\begin{aligned}
&\big\| P_S|\psi_1'\rangle_{M\Sigma XYES} - |\psi_1'\rangle_{M\Sigma XYES} \big\|_2 \\
&= \Big\| P_S T|\psi_1\rangle_{M\Sigma XYES} + TP_S|\psi_1\rangle_{M\Sigma XYES} - TP_S|\psi_1\rangle_{M\Sigma XYES} - T|\psi_1\rangle_{M\Sigma XYES} \Big\|_2 \tag{60}\\
&\leq \big\| [P_S, T] \big\|_\infty \underbrace{\big\| |\psi_1\rangle_{M\Sigma XYES} \big\|_2}_{=1} + \underbrace{\big\| T \big\|_\infty}_{=1} \big\| P_S|\psi_1\rangle_{M\Sigma XYES} - |\psi_1\rangle_{M\Sigma XYES} \big\|_2 \tag{61}\\
&= \big\| [P_S, T] \big\|_\infty + \big\| P_S|\psi_1\rangle_{M\Sigma XYES} - |\psi_1\rangle_{M\Sigma XYES} \big\|_2 \tag{62}\\
&\leq q_1 \big\| [P_S, (U_h)_{XYS}] \big\|_\infty + \sum_{i=1}^{q_1} \big\| \underbrace{[P_S, W^i_{XYE}]}_{=0} \big\|_\infty + \big\| P_S|\psi_1\rangle_{M\Sigma XYES} - |\psi_1\rangle_{M\Sigma XYES} \big\|_2 \tag{63}\\
&\leq q_1 \big\| [(U_h)_{XYS}, P_S] \big\|_\infty + \big\| P_S|\psi_1\rangle_{M\Sigma XYES} - |\psi_1\rangle_{M\Sigma XYES} \big\|_2 \tag{64}\\
&\leq q_1 \delta_L(n) + \big\| P_S|\psi_1\rangle_{M\Sigma XYES} - |\psi_1\rangle_{M\Sigma XYES} \big\|_2 \tag{65}
\end{aligned}
$$

whereby, eqs. (61) and (62) come respectively from the definition of commutator and the definition of the operator norm. On top of that, $\big\| |\psi_1\rangle_{M\Sigma XYES} \big\|_2 = 1$ because $|\psi_1\rangle_{M\Sigma XYES}$ is normalized. Equation (63) follows from Lemma 2 and from the fact that P_S and W^i_{XYE} commute. Finally, the first term of the right-hand side of the last equation follows from Lemma 11.

To evaluate the second term of the right-hand side of the latter equation, we will make use of the following bound on the operator norm from Lemma 12:

$$
\Big\| |\psi_0\rangle_{M\Sigma XYEB} - \Phi_S^{\otimes 2l} |\psi_0\rangle_{M\Sigma XYEB} \Big\|_\infty \leq 2lq_1\epsilon_L(n).
$$

We have:

$$
\begin{aligned}
\Big\| P_S|\psi_1\rangle_{M\Sigma XYES} - |\psi_1\rangle_{M\Sigma XYEBS} \Big\|_2 = \Big\| & P_S B\, \mathsf{Sign}_{\mathsf{sk}}\, |\psi_0\rangle_{M\Sigma XYEBS} - P_S B\, \mathsf{Sign}_{\mathsf{sk}}\, \Phi_S^{\otimes 2l} |\psi_0\rangle_{M\Sigma XYEBS} \\
& + P_S B\, \mathsf{Sign}_{\mathsf{sk}}\, \Phi_S^{\otimes 2l} |\psi_0\rangle_{M\Sigma XYEBS} - B\, \mathsf{Sign}_{\mathsf{sk}}\, \Phi_S^{\otimes 2l} |\psi_0\rangle_{M\Sigma XYEBS} \\
& + B\, \mathsf{Sign}_{\mathsf{sk}}\, \Phi_S^{\otimes 2l} |\psi_0\rangle_{M\Sigma XYEBS} - B\, \mathsf{Sign}_{\mathsf{sk}}\, |\psi_0\rangle_{M\Sigma XYESS} \Big\|_2 .
\end{aligned}
$$

We will use the triangle inequality to split this into three terms and then bound each of them separately.

We can bound the first term as follows:

$$
\begin{aligned}
& \Big\| P_S B\, \mathsf{Sign}_{\mathsf{sk}}\, |\psi_0\rangle_{M\Sigma XYEBS} - P_S B\, \mathsf{Sign}_{\mathsf{sk}}\, \Phi_S^{\otimes 2l} |\psi_0\rangle_{M\Sigma XYEBS} \Big\|_2 \\
& \leq \underbrace{\Big\| P_S B\, \mathsf{Sign}_{\mathsf{sk}} \Big\|_\infty}_{=1} \underbrace{\Big\| |\psi_0\rangle_{M\Sigma XYEBS} - \Phi_S^{\otimes 2l} |\psi_0\rangle_{M\Sigma XYEBS} \Big\|_2}_{\leq 2lq_1\epsilon_L(n)} \\
& \leq 2lq_1\epsilon_L(n)
\end{aligned}
$$

where the first inequality follows by the definition of the operator norm and the final upper bound results from Lemma 12.

Next, we bound the second term. It is exactly the same as the expression in eq. (125), thus,

$$
\Big\| P_S B\, \mathsf{Sign}_{\mathsf{sk}}\, \Phi_S^{\otimes 2l} |\psi_0\rangle_{M\Sigma XYEBS} - B\, \mathsf{Sign}_{\mathsf{sk}}\, \Phi_S^{\otimes 2l} |\psi_0\rangle_{M\Sigma XYEBS} \Big\|_2 = 0.
$$

Finally, looking at the third term, we observe that it is very similar to the first term. Thus, they have the same bound. Therefore,

$$
\Big\| P_S|\psi_1\rangle_{M\Sigma XYES} - |\psi_1\rangle_{M\Sigma XYEBS} \Big\|_2 \leq 2lq_1\epsilon_L(n) + 0 + 2lq_1\epsilon_L(n) = 4lq_1\epsilon_L(n)
$$

and

$$
\big\| P_S|\psi_1'\rangle_{M\Sigma XYES} - |\psi_1'\rangle_{M\Sigma XYES} \big\|_2 \leq q_1\delta_L(n) + 4lq_1\epsilon_L(n) = q_1(\delta_L(n) + 4l\epsilon_L(n)). \tag{66}
$$

□

As long as $q_1 = o(2^{n/2})$, the bound in Lemma 13 is small.

Recall that, just like in Section 4.4, we want to analyze the modified BlindForge experiment where the Q-measurement is applied after the adversary has output a forgery, but before the secret key register is measured to sample the secret key and verify the forgery. It thus remains to show that due to the fact that $|\psi_1'\rangle$ is approximately in the range of P_S, the outcome $l+1$ only occurs with small probability.

To that end, we define a new measurement given by projectors $\tilde{Q}_i$ that performs the Q-measurement controlled on the content of the M-register, i.e.

$$\tilde{Q}_i = \sum_m |m\rangle\langle m|_M \otimes Q_i^m.$$

Now, observe that applying the Q-measurement after the adversary has output a forgery is equivalent to applying the $\tilde{Q}$-measurement right before the adversary's measurement that produces the forgery is applied. To prove that, if $m^* \in B$, the outcome $l+1$ occurs only with small probability in the modified BlindForge experiment, it thus suffices to prove the following lemma.

Lemma 14. *In the* Quantum independent world*, for blinded messages, the outcome $l+1$ only occurs with small probability,*

$$\left\| \tilde{Q}_{l+1} \Pi_M^B |\psi_1'\rangle_{M\Sigma XYES} \right\|_2 \le q_1(\delta_L(n) + 4l\epsilon_L(n)),$$

where

$$\Pi^B = \sum_{m\in B} |m\rangle\langle m|.$$

Proof. By Lemma 9, we have

$$\begin{aligned} Q_{l+1}^m \Pi_M^B P_S &= \sum_m \left(|m\rangle\langle m|_M \Pi_M^B \right) \otimes Q_i^m P_S && (67)\\ &= \sum_{m\in B} |m\rangle\langle m|_M \otimes Q_i^m P_S = 0. && (68)\end{aligned}$$

Therefore we can bound

$$\begin{aligned} \left\| \tilde{Q}_{l+1} \Pi_M^B |\psi_1'\rangle_{M\Sigma XYES} \right\|_2 &= \left\| \tilde{Q}_{l+1} \Pi_M^B \left(|\psi_1'\rangle_{M\Sigma XYES} - P_S |\psi_1'\rangle_{M\Sigma XYES} \right) \right\|_2 && (69)\\ &\le \left\| |\psi_1'\rangle_{M\Sigma XYES} - P_S |\psi_1'\rangle_{M\Sigma XYES} \right\|_2 && (70)\\ &\le q_1(\delta_L(n) + 4l\epsilon_L(n)), && (71)\end{aligned}$$

where we have used the fact that $\|\tilde{Q}_{l+1}\Pi_M^B\|_\infty \le 1$ in the first and Lemma 13 in the second inequality. □

We are now ready to prove Theorem 3.

Proof of Theorem 3. We begin by bounding the success probability of the adversary in the modified BlindForge experiment, in the Quantum independent world. Analogously to eq. (55), we bound, abbreviating the modified BlindForge experiment as MBF,

$$\begin{aligned} \Pr_{QI,MBF}[\mathcal{A} \text{ succeeds}] &= \sum_{i=1}^{l+1} \Pr_{QI,MBF}[\mathcal{A} \text{ succeeds} \wedge \text{outcome } i] \\ &= \sum_{i=1}^{l} \Pr_{QI,MBF}[\mathcal{A} \text{ succeeds} \wedge \text{outcome } i] + \Pr_{QI,MBF}[\mathcal{A} \text{ succeeds} \wedge \text{outcome } l+1] \\ &\le \sum_{i=1}^{l} \Pr_{QI,MBF}[\text{outcome } i] \times 2^{-n} + \Pr_{QI,MBF}[\text{outcome } l+1] && (72)\\ &\le 2^{-n} + q^2(\delta_L(n) + 4l\epsilon_L(n))^2, && (73)\end{aligned}$$

where "outcome i" is the event that the Q-measurement yields outcome i, the first inequality uses the fact that σ^* and $s_i^{m_i^*}$ are independent conditioned on outcome i, and the last inequality uses the square of the inequality from Lemma 14.

Exactly as in in the simplified case in Section 4.4, we can bound the success probability in the actual BlindForge experiment using the pinching lemma, Lemma 1,

$$\Pr_{QI,\mathsf{BlindForge}}[\mathcal{A} \text{ succeeds}] \le (l+1)\left(2^{-n} + q^2(\delta_L(n) + 4l\epsilon_L(n))^2\right).$$

Finally, plugging in the functions $\epsilon_L(n)$ and $\delta_L(n)$ from Lemmas 8 and 11, and applying Lemma 4 for $w = 2$, we obtain

$$\begin{aligned}\Pr_{\mathsf{BlindForge}}[\mathcal{A}\text{ succeeds}] &\leq (l+1)\left(2^{-n} + q^2\left(\frac{32l}{2^{n/2}} + 4l\frac{6}{2^{n/2}}\right)^2\right) + 12l^2 2^{-n} \\ &\leq l^2 \cdot 2^{-n}\left(3137q^2(l+1) + 12\right).\end{aligned}$$

□

5 One-time BU security of the Winternitz OTS

The Lamport OTS that we analyzed in the last section is, in some sense, a special case of the Winternitz OTS. Indeed, the Winternitz scheme for $w = 2$ is fairly similar to the Lamport OTS, except that the public key is used to sign the bits that are equal to 1, which is compensated for by the checksum encoding. As a result, the analysis of the Winternitz OTS in the QROM is, in a similar sense, a generalization of the one of the Lamport OTS.

Before getting started, we give and overview of our strategy. In this section, we use the same register labels as in the table 1 of Lamport OTS section except that the secret key register S is now replaced by the hash chain register Γ. We remark that the general overview of the proof for Lamport OTS given at the beginning of Section 4 is similar for the Winternitz OTS except that the security argument is different. More precisely, in the Winternitz OTS, the signature algorithm uses the hash chain registers above the queried position to produce the signature. Classically, the property that enables security is that the adversary does not have any information about the part of the hash chain below the queried position, and this represents the invariant of the hash chain. Quantumly, our intuition is that since in the BlindForge experiment the forged message must be outside the queried region, and since by construction of the checksum, for any queried message there exists at least one position at which the block corresponding to the forged message is smaller than the one of the queried message. Thus, the hash chain corresponding to that specific block should still be in its initial state, and hence in the invariant of the hash chain. Therefore, we want to show that for a moderate number of queries to the random oracle, no adversary can win the BlindForge experiment with a significant probability. Towards that goal, we follow the same steps as in the Lamport OTS. Specifically, we prove the following theorem.

Theorem 4. *The Winternitz OTS in Section 2.3.2 is 1-*BU *secure if the function chain $\mathcal{C}$ is modeled as a quantum-accessible random oracle. More precisely, let $\mathcal{A}$ be an adversary that plays the* BlindForge *game for the Winternitz OTS, making a total of q queries to the random oracle. Then $\mathcal{A}$ succeeds with a probability bounded as*

$$\begin{aligned}\Pr[\mathcal{A}\text{ wins }\mathsf{BlindForge}] &\leq 2^{-n}\left[\left(1 + q^2l^2(w-1)^2(20w-4)^2\right)(l+1) + 3w^2l^2\right] & (74)\\ &\leq 800w^4q^2l^3 \cdot 2^{-n}. & (75)\end{aligned}$$

Here, l is the length of the encoded message in w-ary, see eq. (10), $w \geq 2$ is the Winternitz parameter, and the simplified bound in the last line holds for $q > 0$.

The main difference between the analyses of the Lamport and Winternitz OTS is as follows. For the Lamport OTS, the public key is obtained from the private key by applying a hash function once. For the Winternitz OTS, on the other hand, the secret an public key consist of the start and end points of length w hash chains, respectively. Thus, while following the same proof strategy, the Q projectors as well as the invariant projector P needs to be defined differently. Thus, we start our analysis by describing the Q projectors and the invariant projector for the Winternitz OTS.

5.1 Q projectors for Winternitz OTS

The Winternitz signature of any message is composed of l hash chain elements. In complete analogy to eq. (30) in Section 4.1, we define a measurement whose projectors correspond respectively to the events that *the i-th hash chain element relevant for the forged signature is in state* $|\Phi\rangle$ and *none of them is in state* $|\Phi\rangle$:

$$Q_{i^*}^{b^*} = \Phi^{\perp}_{\Gamma_1^{b_1^*}} \otimes \cdots \otimes \Phi^{\perp}_{\Gamma_{i^*-1}^{b_{i^*-1}^*}} \otimes \Phi_{\Gamma_{i^*}^{b_{i^*}^*}}, \qquad Q_{l+1}^{b^*} = \bigotimes_{i=1}^{l} \Phi^{\perp}_{\Gamma_i^{b_i^*}} \tag{76}$$

where $i^* \in \{1,\ldots,l\}$, $b_i^* = b_i(m^*)$ and l is the number of blocks of the message and the checksum, see eq. (10). These operators act as $\mathbb{1}$ on all other registers Γ_i^j that are not specified.

5.2 Invariant projector for Winternitz OTS

In this section, we define the invariant projector, denoted by P_Γ, that will be used to track the invariant of the hash chain register. We also state several of its properties.

Recall from our discussion of blind unforgeability in Section 2.4 that B denotes the set of blinded messages and B^c its complement,i.e., the set of un-blinded messages. We also recall from the description of the Winternitz OTS in Section 2.3.2 that a block b of a message is the concatenation of the blocks obtained from the encoding of the message and its corresponding checksum in w-ary.

Define $\alpha = (\alpha_i^j)_{i=1,\dots,l}^{j=0,\dots,w-2}$ as a $l(w-1)$-bit string whose bits $\alpha_i^j \in \{0,1\}$ indicate that the projector $\Phi(\alpha_i^j)$ is applied on the corresponding hash chain register Γ_i^j where

$$\Phi(0) = \Phi, \qquad\qquad \Phi(1) = \Phi^\perp. \tag{77}$$

For each string α, we define the associated projector $\Phi(\alpha)$ on the whole hash chain (except for the last) register Γ as

$$\Phi(\alpha)_\Gamma = \bigotimes_{i=1}^{l} \bigotimes_{j=0}^{w-2} \Phi(\alpha_i^j)_{\Gamma_i^j}. \tag{78}$$

Note that this is a complete set of projectors, i.e., $\sum_{\alpha\in\{0,1\}^{l(w-1)}} \Phi(\alpha)_\Gamma = \mathbb{1}_\Gamma$.

Since we are interested in the unused part of the hash chain register, we need to filter those α's for which Γ_i^j is in state $|\Phi\rangle$. By construction of the checksum, if a block b of a message m is computed, then in the block b' of any other message m', there exists at least one position i at which $b_i' < b_i$, $1 \le i \le l$. Therefore, since the blinded signing oracle signs at most a single un-blinded message, $m \in B^c$, the state after the signing oracle call can be written as a superposition of states where, for some un-blinded message $m' \in B^c$, $b_i' < b_i$ for all i. The latter implies that the hash chain registers corresponding to those b_i' are still in the uniform superposition $|\Phi\rangle$, for all i. Thus, we collect all strings α that are consistent with no blinded messages having been signed in

$$\widehat{B^c} = \bigcup_{m\in B^c} \left\{ \alpha \in \{0,1\}^{l(w-1)} \;\middle|\; \alpha_i^j = 0 \text{ for all } i = 1,\dots,l \text{ and } j < b_i(m) \right\} \tag{79}$$

as the set of strings α that indicate which hash chain registers were not used during hash queries and Sign query, that is those that fulfill the condition $\alpha_i^{\bar{m}_i} = 0$ for all i. Specifically, $\widehat{B^c}$ contains all the strings that are consistent with no blinded messages having been signed. Finally, we define

$$P_\Gamma = \sum_{\alpha\in\widehat{B^c}} \Phi(\alpha)_\Gamma \tag{80}$$

as the projector acting on the invariant hash chain register, specifically on the subspace consistent with $\widehat{B^c}$. Note that P_Γ is indeed a projector since it is a sum of mutually orthogonal projectors.

Using these new definitions of the Q projectors and the invariant projector P_Γ, a set of lemmas similar to Lemmas 8 to 11 forms the basis of the BU security proof for the Winternitz OTS. In fact, Lemma 8 is a special case of Lemma 15 where the register Γ is replaced by S and we set $w = 2$ (see Appendix A.1 for proof). Lemma 9 holds for the new projectors Q_{l+1} and P_Γ by construction. Finally, Lemmas 10 and 11 need to be changed slightly for the Winternitz OTS and are stated below. Lemmas 16 to 18 are proved in Appendix B.

Lemma 15. *Let U_h be the random oracle unitary for any given function h (see Section 3) and let $\Phi = |\Phi\rangle\langle\Phi|$ denote the projector onto the uniform superposition $|\Phi\rangle$. Furthermore, let $\Gamma_i^{\le j} = \Gamma_i^0 \dots \Gamma_i^j$ and*

$$\Phi_{\Gamma_i^{\le j}} = \left(\Phi^{\otimes j}\right)_{\Gamma_i^{\le j}}. \tag{81}$$

Then, for any $i' \in \{1,\dots,l\}$ and $j' \in \{0,\dots,w-2\}$,

$$\left\| \left[(U_h)_{XY\Gamma}, \Phi_{\Gamma_{i'}^{\le j'}} \right] \right\|_\infty \le \frac{6(w-1)}{2^{n/2}} = \epsilon_W(n) \tag{82}$$

is negligible in n.

Lemma 16. *Let $B\,\mathsf{Sign}_{\mathsf{sk}}$ be the blinded signing oracle for the Winternitz OTS, and let $|\psi_0\rangle$ be the adversary's state before the* Sign *query. If there are no hash queries, then after making a single* Sign *query the adversary's state $|\psi_1\rangle = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle$ is completely in the range of the invariant projector P_Γ defined in eq. (80). That is,*

$$P_\Gamma B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle. \tag{83}$$

For every message in the blinding set B, there exists at least one hash chain element necessary for its corresponding signature such that the corresponding hash chain register is in the uniform superposition state $|\Phi\rangle$. This implies the following lemma.

Lemma 17. *Let $m^* \in B$ and $b^* = b(m^*)$, see eq. (11). Then the projectors $Q_{l+1}^{b^*}$ and P_Γ defined in eqs. (76) and (80) are orthogonal:*

$$Q_{l+1}^{b^*} P_\Gamma = 0. \tag{84}$$

Lemma 18. *Let P_Γ and U_h be respectively the invariant projector for the Winternitz OTS and the random oracle unitary defined with respect to the* Quantum independent world. *If there are hash queries after the* Sign *query, then*

$$\left\| [U_h, P_\Gamma] \right\|_\infty \leq \delta_W(n) \tag{85}$$

where

$$\delta_W(n) = \frac{8l(w+1)(w-1)}{2^{n/2}}.$$

Just like in the Lamport OTS, we use the above lemmas to prove our main results. In the following sections, we analyze the situation where the adversary makes q_0 hash queries before the Sign query and q_1 hash queries after, maximizing the resulting bound under the condition that $q_0 + q_1 = q$.

The proof of Theorem 4 is presented in steps in the following subsections. We begin by presenting some concepts and tools which will be used in the proof. Subsequently, we prove the lemmas stated above. Finally, we combine them to prove Theorem 4.

5.3 Hash queries before Sign query

In this section, we study the impact of hash queries before Sign query on the hash chain register Γ. Our goal is to show that, for a moderate number of queries to the random oracle, no adversary can learn a significant amount of information about the hash chain. Therefore, she cannot produce a valid forgery except with small probability.

Let $|\psi\rangle_{XYM\Sigma E}$ be adversary's initial state before any queries. Before any query is performed, the whole hash chain register Γ, except the last, is in the uniform superposition state, i.e,

$$|\nu\rangle_\Gamma = \bigotimes_{i=1}^{l} \bigotimes_{j=0}^{w-2} |\Phi\rangle_{\Gamma_i^j}. \tag{86}$$

Assume the adversary $\mathcal{A}_0$ queries the random oracle q_0 times before querying the signing oracle. If V_{XYE}^i denotes the unitary she performs after the i-th query, the final adversary state after q_0 hash queries is

$$|\psi_0\rangle_{XYM\Sigma E\Gamma} = V_{XYE}^{q_0}(U_h)_{XY\Gamma} V_{XYE}^{q_0-1} \cdots V_{XYE}^{2}(U_h)_{XY\Gamma} V_{XYE}^{1}(U_h)_{XY\Gamma} |\psi\rangle_{XYM\Sigma E} |\nu\rangle_\Gamma \tag{87}$$

where U_h is the random oracle unitary used to answer hash queries. The following lemma shows that the hash chain registers of this state are still close to being in uniform superposition.

Lemma 19. *In the* Quantum independent world, *without querying the B* Sign *oracle, hash queries leave the state of the secret key registers approximately unchanged:*

$$\left\| \Phi_\Gamma^{\otimes l(w-1)} |\psi_0\rangle_{XYM\Sigma E\Gamma} - |\psi_0\rangle_{XYM\Sigma E\Gamma} \right\|_2 \leq l q_0 \epsilon_W(n). \tag{88}$$

Proof. The proof of Lemma 19 is very similar to the proof of Lemma 12 in the Lamport OTS security analysis, except that it uses Lemma 15 for each of the l hash chains, with $j' = w-2$, where the proof of Lemma 12 applies Lemma 8 for each of the $2l$ secret key registers. □

As in for the Lamport OTS, the above Lemma means that the adversary learns almost no information about the hash chain, unless $q_0 = \Omega(2^{n/2})$.

5.4 Query to the signing oracle

Now that we have control over the advantage an adversary can gain from making hash queries before the sign query, we need to analyze the possible advantage from hash queries after the sign query and bound the overall success probability using Lemma 19. The discussion in this section does not concern the random oracle, so we absorb the random oracle query registers XY into E for the purpose of this section.

A key property of the Winternitz OTS when analyzing classical security is that for all messages m that have not been queried, there exists an index j such that $s_j^{m_j}$ is hidden from the adversary by the preimage resistance of the used hash function. In blind-unforgeability (for classical adversaries), this property holds for all *blinded messages*. In the setting of quantum queries, we have to track this property in superposition while the adversary is making hash queries after the sign query. As this is complicated by the "for all"-quantifier, we begin by analyzing the case where the adversary makes no hash queries after the sign query to ease the reader into our proof technique.

In the 1-BlindForge game, an adversary $\mathcal{A}$ is allowed to query the Sign-oracle at most once to produce a valid forged message-signature pair (m^*, σ^*). To analyze the interaction between $\mathcal{A}$ and the signing oracle, we will follow the steps stated in eq. (46). Those steps correspond to applying the Sign-oracle and an arbitrary unitary $U_{M\Sigma E}$, followed by measuring the message and signature registers M and Σ. Let us now analyze these steps in more detail and write down the corresponding quantum states.

First, the adversary $\mathcal{A}$ prepares the state

$$|\psi_0\rangle_{M\Sigma EB\Gamma} = \sum_{m\in\{0,1\}^a} \sum_{\sigma\in(\{0,1\}^n)^l} \kappa_{m\sigma b} |m\rangle_M |\sigma\rangle_\Sigma |\alpha_{m\sigma b}\rangle_E |b\rangle_B |\nu\rangle_\Gamma \tag{89}$$

where $|m\rangle_M = |m_1 \cdots m_a\rangle_M$, $|\sigma\rangle_\Sigma = |\sigma_1\rangle_{\Sigma_1} \cdots |\sigma_l\rangle_{\Sigma_l}$, $|\nu\rangle_\Gamma$ is defined in eq. (86), and b is the amplitude of the blinding register B. Here Γ is a composite register of the form $\Gamma = \{\Gamma_i^j : i \in \{1, \ldots, l\}, j \in \{0, \ldots, w-1\}\}$, B indicates whether the message is blinded or not ($|1\rangle_B$ for blinded and $|0\rangle_B$ for un-blinded), and the remaining registers are defined as in table 1.

Next, $\mathcal{A}$ queries this state to the Sign oracle which answers the query with the following signed state:

$$\begin{aligned} |\psi_1\rangle_{M\Sigma EB\Gamma} &= B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle_{M\Sigma EB\Gamma} \\ &= |\psi_1^1\rangle_{M\Sigma EB\Gamma} + |\psi_1^0\rangle_{M\Sigma EB\Gamma} \end{aligned} \tag{90}$$

where the superscripts 1 and 0 correspond to blinded (B) and un-blinded (B^c) messages. The expression of the first term is given by

$$|\psi_1^1\rangle_{M\Sigma EB\Gamma} = \sum_{m\in B} \sum_{\sigma\in(\{0,1\}^n)^l} \kappa_{m\sigma 1} |m\rangle_M |\sigma\rangle_\Sigma |\alpha_{m\sigma 1}\rangle_E |1\rangle_B |\nu\rangle_\Gamma \tag{91}$$

where the latter follows because for blinded messages ($m \in B$) there is no signature.

From now on, we will describe how the second term $|\psi_1^0\rangle_{M\Sigma EB\Gamma}$ in eq. (90) is obtained. While generally the signature is computed in superposition, we will describe it on a fixed message for the sake of simplicity. The general operation corresponds to extending this description by linearity.

Given a fixed message $|m\rangle_M = |m_1 \cdots m_a\rangle_M \in B^c$ represented in computational basis, the signing oracle first encodes the message in l blocks, each in base-w representation:

$$U_b |m\rangle_M |0\rangle_W |\sigma\rangle_\Sigma |\nu\rangle_\Gamma \mapsto |m\rangle_M |0 \oplus b(m)\rangle_W |\sigma\rangle_\Sigma |\nu\rangle_\Gamma \tag{92}$$

where l is defined in eq. (10), W is an ancilla register used to store $b(m)$, and $b(m)$ is defined in eq. (11) as

$$b(m) = (b_1, \ldots, b_l) = m \parallel C(m)$$

where $C(m)$ is the checksum corresponding to m. Note that the process by which the unitary U_b computes $|b(m)\rangle$ is similar to the classical way described in Section 2.3.2.

Using the blocks of m, the signing oracle computes the signature as follows:

$$\begin{aligned} U_{\mathsf{Sign}_{\mathsf{sk}}\,W\Gamma:\Sigma} |m\rangle_M |b_1 \cdots b_l\rangle_W |\sigma \oplus \gamma\rangle_\Sigma |\Omega(m,\gamma)\rangle_\Gamma &= \mathrm{CNOT}^{\otimes n}_{\Gamma_1^{b_1}:\Sigma_1} \cdots \mathrm{CNOT}^{\otimes n}_{\Gamma_l^{b_l}:\Sigma_l} |m\rangle_M |b_1 \cdots b_l\rangle_W |\sigma\rangle_\Sigma |\nu\rangle_\Gamma \\ &= |m\rangle_M |b_1 \cdots b_l\rangle_W |\sigma \oplus \gamma\rangle_\Sigma |\Omega(m,\gamma)\rangle_\Gamma. \end{aligned}$$

Once the signature is obtained, the ancilla register W is not useful for further analysis so we can remove it from the signed state by applying $U_b^\dagger$. Thus, the final signed state of a fixed message $|m\rangle_M$ is given by

$$|m\rangle_M |\sigma \oplus \gamma\rangle_\Sigma |\Omega(m,\gamma)\rangle_\Gamma \tag{93}$$

with

$$|\sigma \oplus \gamma\rangle_\Sigma := |\sigma_1 \oplus \gamma_1\rangle_{\Sigma_1} \ldots |\sigma_l \oplus \gamma_l\rangle_{\Sigma_l}, \tag{94}$$

$$|\Omega(m,\gamma)\rangle_\Gamma := \left(\bigotimes_{i=1,\ldots,l; b_i \neq w-1} |\gamma_i\rangle_{\Gamma_i^{b_i}} \bigotimes_{j=0,\ldots,w-2; j\neq b_i} |\Phi\rangle_{\Gamma_1^j} \cdots \bigotimes_{j=0,\ldots,w-2; j\neq b_i} |\Phi\rangle_{\Gamma_l^j} \right) |p_1\rangle_{\Gamma_1^{w-1}} \ldots |p_l\rangle_{\Gamma_l^{w-1}}. \tag{95}$$

By linearity, the signature of the un-blinded term which is composed of superposition of messages is given by

$$|\psi_1^0\rangle_{M\Sigma EB\Gamma} = \sum_{m\in B^c} \xi_m \sum_{\sigma\in(\{0,1\}^n)^l} \sum_{\gamma\in(\{0,1\}^n)^l} \kappa_{m\sigma 0}|m\rangle_M|\sigma\oplus\gamma\rangle_\Sigma|\alpha_{m\sigma 0}\rangle_E|0\rangle_B|\Omega(m,\gamma)\rangle_\Gamma \tag{96}$$

where ξ_m is a normalization factor of all hash chain elements used to produce the signature.

Once the adversary receives the signed state $|\psi_1\rangle_{M\Sigma EB\Gamma}$, she carries out some operations with the intention of producing a forged message m^*. Intuitively, those operations can be viewed as applying some arbitrary unitary $U_{M\Sigma E}$ to $|\psi_1\rangle_{M\Sigma EB\Gamma}$:

$$|\psi_2\rangle_{M\Sigma EB\Gamma} = U_{M\Sigma E}|\psi_1\rangle_{M\Sigma EB\Gamma}.$$

Afterwards, the adversary $\mathcal{A}$ measures the M register of the latter state, which gives outcome $m^*\in\{0,1\}^a$. After the measurement, the state $|\psi_2\rangle_{M\Sigma EB\Gamma}$ collapses to the (unnormalized) state

$$\begin{aligned}|\psi_3(m^*)\rangle_{\Sigma EB\Gamma} &= \langle m^*|_M|\psi_2\rangle_{M\Sigma EB\Gamma}\\ &= |\psi_3^1(m^*)\rangle_{\Sigma EB\Gamma} + |\psi_3^0(m^*)\rangle_{\Sigma EB\Gamma}\end{aligned}$$

whereby,

$$|\psi_3^1(m^*)\rangle_{\Sigma EB\Gamma} = \langle m^*|_M U_{M\Sigma E}|\psi_1^1\rangle_{M\Sigma EB\Gamma}$$

and

$$\begin{aligned}|\psi_3^0(m^*)\rangle_{\Sigma EB\Gamma} &= \langle m^*|_M U_{M\Sigma E}|\psi_1^0\rangle_{M\Sigma EB\Gamma}\\ &= \sum_{m\in B^c}\xi_m \sum_{\sigma\in(\{0,1\}^n)^l}\sum_{\gamma\in(\{0,1\}^n)^l} \kappa_{m\sigma 0}\langle m^*|_M U_{M\Sigma E}|m\rangle_M|\sigma\oplus\gamma\rangle_\Sigma|\alpha_{m\sigma 0}\rangle_E|0\rangle_B|\Omega(m,\gamma)\rangle_\Gamma.\end{aligned}$$

The goal of the adversary $\mathcal{A}$ is to produce a forged message σ^* that matches m^*. Towards this end, she measures the Σ register of $|\psi_3(m^*)\rangle_{\Sigma EB\Gamma}$, obtaining outcome $\sigma^*\in(\{0,1\}^n)^l$. Then, the (unnormalized) post-measurement state is

$$\begin{aligned}|\psi_4(m^*,\sigma^*)\rangle_{EB\Gamma} &= \langle\sigma^*|_\Sigma|\psi_3(m^*)\rangle_{\Sigma EB\Gamma}\\ &= \langle\sigma^*|_\Sigma|\psi_3^1(m^*)\rangle_{\Sigma EB\Gamma} + \langle\sigma^*|_\Sigma|\psi_3^0(m^*)\rangle_{\Sigma EB\Gamma}\\ &= |\psi_4^1(m^*)\rangle_{EB\Gamma} + |\psi_4^0(m^*)\rangle_{EB\Gamma}\end{aligned}$$

where

$$|\psi_4^1(m^*,\sigma^*)\rangle_{EB\Gamma} = \langle\sigma^*|_\Sigma\langle m^*|_M U_{M\Sigma E}|\psi_1^1\rangle_{M\Sigma EB\Gamma}$$

and

$$|\psi_4^0(m^*,\sigma^*)\rangle_{EB\Gamma} = \sum_{m\in B^c}\xi_m \sum_{\sigma\in(\{0,1\}^n)^l}\sum_{\gamma\in(\{0,1\}^n)^l} \alpha_{m\sigma 0}\langle\sigma^*|_\Sigma\langle m^*|_M U_{M\Sigma E}|m\rangle_M|\sigma\oplus\gamma\rangle_\Sigma|\alpha_{m\sigma 0}\rangle_E|0\rangle_B|\Omega(m,\gamma)\rangle_\Gamma. \tag{97}$$

For simplicity, we can rewrite $|\psi_4^0(m^*,\sigma^*)\rangle_{EB\Gamma}$ as

$$|\psi_4^0(m^*,\sigma^*)\rangle_{EB\Gamma} = \sum_{m\in B^c}\sum_{\gamma\in(\{0,1\}^n)^l}|\eta(m,\gamma)\rangle_{EB}|\Omega(m,\gamma)\rangle_\Gamma \tag{98}$$

where only $|\eta(m,s)\rangle_{BE}$ depends on m^* and σ^*:

$$|\eta(m,\gamma)\rangle_{EB} = \xi_m\sum_{\sigma\in(\{0,1\}^n)^l}\kappa_{m\sigma 0}\langle\sigma^*|_\Sigma\langle m^*|_M U_{M\Sigma E}|m\rangle_M|\sigma\oplus\gamma\rangle_\Sigma|\alpha_{m\sigma 0}\rangle_E|0\rangle_B.$$

Hence, the adversary $\mathcal{A}$ produces a forged message-signature pair (m^*,σ^*). The probability of producing this pair is $\||\psi_4^0(m^*,\sigma^*)\rangle_{BE\Gamma}\|^2$.

The next step is to analyze the probability that $\mathcal{A}$'s forgery candidate is correct. Recall from eq. (11) that for some un-blinded message $m\in B^c$ or blinded message $m^*\in B$, we denote by $b=b(m)$ and $b^*=b(m^*)$ the base-w representation of m and m^* in term l blocks, each in w-ary.

Here, we consider two cases. The first case, namely when $m^*\notin B$, is trivial because $\mathcal{A}$ has lost the BlindForge experiment as m^* must be blinded by definition. The rest of this section is devoted to analyzing the second case.

If $m^* \in B$, then the forged message m^* has not been signed since the blinded Sign oracle signs only un-blinded messages. Hence, by construction of the checksum, for any message $m \notin B$ there exists at least one index $i^* \in \{1, \ldots, l\}$ such that the corresponding block b is larger than the block b^* of m^*. This implies that this specific hash chain element $\Gamma_{i^*}^{b_{i^*}}$ has not been used for the signature of the adversary's queried message and is therefore still in its initial state. Note that this holds only in superposition over m. Indeed, i^* depends on m and is in general different for each term of the superposition.

In the Winternitz OTS, we know that the hash sub-chain below the queried position is not used for the signature procedure and is therefore in uniform superposition $|\Phi\rangle$. But given that during the signature process the hash chain is in superposition, our main goal here is to track the invariant of the hash chain and show that some of the hash chain elements relevant for verifying the forged signature satisfy this invariant and are thus unknown to the adversary, so it is unlikely that the adversary would have used the correct hash chain to produce the forged signature.

To that end, we analyze a modified BlindForge experiment, where an additional measurement is performed on the hash chain register after the adversary has output their forgery, but before the hash chain register is measured to actually sample the hash chain as required in the Quantum independent world. This additional Q-measurement was defined in eq. (30), with Winternitz Q projectors defined in Section 5.1. Since it has few outcomes, its effect on the adversary's winning probability is limited and can be bounded by the pinching lemma (Lemma 1).

If the Q-measurement yields outcome $i^* \in \{1, \ldots, l\}$, then the hash chain element $\Gamma_{i^*}^{b_{i^*}}$ is in uniform superposition and the adversary is bound to fail as σ^* is independent of the hash chain n-bit string $\gamma_{i^*}^{b_{i^*}}$ (the result of measuring $\Gamma_{i^*}^{b_{i^*}}$).

It remains to analyze the outcome $l+1$ that corresponds to the projector $Q_{l+1}^m = (\Phi^\perp)^{\otimes l}$ where $\Phi^\perp = \mathbb{1} - |\Phi\rangle\langle\Phi|$. The final adversary state after the measurement, see eq. (97), contains both blinded and un-blinded terms. If we apply $\Phi^\perp$ to any hash chain register of the blinded term $|\psi_4^1(m^*, \sigma^*)\rangle_{BE\Gamma}$, we get 0 since all hash chain registers (except for the last) are in uniform superposition $|\Phi\rangle$ and the checksum guarantees that there exists at least one position i^* such that $b_{i^*} < w - 1$.

For the rest of this section, we fix a message m^* and focus on the un-blinded term $|\psi_4^0(m^*, \sigma^*)\rangle_{BE\Gamma}$. Because $m^* \in B$, it has not been signed. So for any un-blinded message $m \in B^c$ that has been queried to the Sign oracle, the checksum guarantees that there exists at least one index i^* for which $b_i(m^*) < b_i(m)$. Therefore, the hash chain element $\Gamma_{i^*}^{b_{i^*}}$ corresponding to that position is in state $|\Phi\rangle$. To find that position, we define

$$i(b) = \min\{k = 1, \ldots, l \mid b_k^* < b_k\}$$

as the smallest index k for which $b_k(m^*) < b_k(m)$. Intuitively, it is the first hash chain element of Γ that still remains in uniform superposition. In the following, let $\Gamma(b(m)) := \Gamma_1^{b_1(m)} \cdots \Gamma_l^{b_l(m)}$ and recall from eqs. (95) and (98) that the un-blinded term is given by

$$\begin{aligned} |\psi_4^0(m^*, \sigma^*)\rangle_{EB\Gamma} &= \sum_{m \in B^c} \sum_{\gamma \in (\{0,1\}^n)^l} |\eta(m, \gamma)\rangle_{EB} |\Omega(m, \gamma)\rangle_\Gamma |\Phi\rangle_{\Gamma(b(\bar{m}))} \\ &= \sum_{k=1}^{l} |\hat{\eta}(m^*, \sigma^*, k)\rangle_{EB\Gamma_{\{(k, b_k^*)\}^c}} |\Phi\rangle_{\Gamma_k^{b_k^*}}, \end{aligned} \tag{99}$$

where all the registers except for $\Gamma_k^{b_k^*}$ are included into the first system. The state $|\hat{\eta}(m^*, \sigma^*, k)\rangle$ is defined as the part of the superposition over m in $|\psi_4^0(m^*, \sigma^*)\rangle$ with constant $k = i(b)$ (excluding the register $\Gamma_k^{b_k^*}$). Thus the register $\Gamma_k^{b_k^*}$ is still in the uniform superposition $|\Phi\rangle$. Applying Q_{l+1}^m onto the l hash chain elements $\Gamma(b^*)$ of the register Γ in eq. (99) gives

$$(Q_{l+1}^m)_{\Gamma(b^*)} |\psi_4^0(m^*, \sigma^*)\rangle_{EB\Gamma} = ((\Phi^\perp)^{\otimes l})_{\Gamma(b^*)} \left(\sum_{k=1}^{l} |\hat{\eta}(m^*, \sigma^*, k)\rangle_{BE\Gamma_{\{(k, b_k^*)\}^c}} |\Phi\rangle_{\Gamma_k^{b_k^*}} \right) = 0,$$

which vanishes because, for each k, the register $\Gamma_k^{b_k^*}$ is in state $|\Phi\rangle$ and $\Phi^\perp |\Phi\rangle = 0$. Hence, the situation where none of the hash chain elements relevant for the verification of the forged signature σ^* is in state $|\Phi\rangle$ can never occur.

Finally, we bound the success probability of the adversary in winning the BlindForge game in the Real world. Given that this analysis is roughly the same as the one done in Section 4.4 of the Lamport OTS. Following the same steps, we get the following adversary's success probability in producing a valid forged message-signature pair (m^*,σ^*):

$$\Pr_{QI}[\mathcal{A} \text{ wins } \mathsf{BlindForge}] = \sum_{m^*,\sigma^*} \Pr_{\mathsf{BlindForge}}[\text{Success} \wedge \mathcal{A} \text{ outputs } (m^*,\sigma^*)] \leq \frac{l+1}{2^n}.$$

But, given that there is possible collision tuples in Real world, adding the upper bound in Lemma 6 to the latter equation gives

$$\begin{aligned}\Pr[\mathcal{A} \text{ wins } \mathsf{BlindForge}] &= \sum_{m^*,\sigma^*} \Pr_{\mathsf{BlindForge}}[\text{Success} \wedge \mathcal{A} \text{ outputs } (m^*,\sigma^*)] \\ &\leq \frac{l+1}{2^n} + \frac{3(wl)^2}{2^n} = \frac{1+l+3(wl)^2}{2^n}.\end{aligned}$$

5.5 Hash queries after Sign query

In this section, we analyse the adversary's *hash queries after* Sign *query* to bound the success probability that an adversary with a given number of queries can achieve in the BlindForge game and thus prove Theorem 4. In this case it is not obvious how to track the invariant of the hash chain, i.e. the fact that there is at least one unused part of the hash chain element that is relevant for the forged signature. Therefore we use a special projector P_Γ defined in Section 5.2 that projects onto the subspace of the hash chain register that is consistent with a single blinded sign query and no hash queries. If the final adversary state after producing the forgery candidate is in the image of P_Γ, then the outcome $l+1$ corresponding to the situation where *none of the hash chain elements useful for the forged signature is in state* $|\Phi\rangle$ can never occur, according to Lemma 17. We thus want to show that the adversary's final state is approximately in the range of P_Γ.

If there are no hash queries before the Sign query, then from Lemma 16 the adversary state after the Sign query remains completely in the range of P_Γ, which means that the outcome $l+1$ cannot occur. That is,

$$P_\Gamma|\psi_1\rangle = P_\Gamma B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle = |\psi_1\rangle$$

where $|\psi_0\rangle$ and $|\psi_1\rangle$ are respectively the adversary state immediately before and after the Sign query.

Now, assuming there are hash queries before the Sign query, since the projector P_Γ and the random oracle unitary U_h approximately commute by Lemma 18, it follows that hash queries before Sign query give no significant information to the adversary about the invariant of the hash chain register.

Suppose there are hash queries after the Sign query and examine in detail what happen in this case. From the previous case, we know that the adversary's state directly after the Sign query is $|\psi_1\rangle_{M\Sigma XYE\Gamma}$. Just like for hash queries before the Sign query, suppose that the adversary makes q_1 hash queries after querying the signing oracle. Let $(W^i_{XY\Gamma E})_{i=1,\dots,q_1}$ be unitaries applied to the adversary state between hash queries. Then, the adversary's state after q_1 hash queries and before performing some unitary operations $U_{M\Sigma E}$ on the post hash queried state or any measurement leading to the forgery candidate is

$$|\psi_1'\rangle_{XYM\Sigma EB\Gamma} = (U_h)_{XY\Gamma} W^{q_1}_{XYM\Sigma E} (U_h)_{XY\Gamma} W^{q_1-1}_{XYM\Sigma E} \cdots W^2_{XYM\Sigma E} (U_h)_{XY\Gamma} W^1_{XYM\Sigma E} |\psi_1\rangle_{XYM\Sigma EB\Gamma}$$

where $|\psi_1\rangle_{XYM\Sigma EB\Gamma}$ is the adversary's state immediately after Sign query. For sake of simplicity, we set

$$T = (U_h)_{XY\Gamma} W^{q_1}_{XYM\Sigma E} (U_h)_{XY\Gamma} W^{q_1-1}_{XYM\Sigma E} \cdots W^2_{XYM\Sigma E} (U_h)_{XY\Gamma} W^1_{XYM\Sigma E}.$$

Next, we prove the following lemma.

Lemma 20. *In the* Quantum independent world, *the state right before the adversary's measurement determining the forgery is applied is approximately in the range of* P_Γ, *i.e.*

$$\left\| P_\Gamma |\psi_1'\rangle_{M\Sigma XYE\Gamma} - |\psi_1'\rangle_{M\Sigma XYE\Gamma} \right\|_2 \leq q_1 \delta_W(n) + 4lq_1\epsilon_W(n) = q_1\Big(\delta_W(n) + 2l(w-1)\epsilon_W(n)\Big). \qquad (100)$$

Proof. The proof is the same as that of Lemma 13, except that Lemma 18 is used in place of Lemma 11. □

Notice the bound in eq. (100) is small if an adversary makes at most $q_1 = o(2^{n/2})$ queries.

Recall that, just like in Section 5.4, we want to analyze the modified BlindForge experiment, where the Q-measurement is applied after the adversary has output a forgery, but before the secret key register is measured to sample the hash chain and verify the forgery. It thus remains to show that due to the fact that $|\psi_1'\rangle$ is approximately in the range of P_Γ, the outcome $l+1$ only occurs with small probability. Given that this part is similar to the analysis of the last part of the BlindForge experiment made in Section 4.5 after the proof of Lemma 13, we use the same analysis for the Winternitz OTS.

To that end, we define a new measurement given by projectors $\tilde{Q}_i$ that performs the Q-measurement controlled on the content of the M-register, i.e.

$$\tilde{Q}_i = \sum_m |m\rangle\langle m|_M \otimes Q_i^{(m)}.$$

Now, observe that applying the Q-measurement after the adversary has output a forgery is equivalent to applying the $\tilde{Q}$-measurement right before the adversary's measurement that produces the forgery is applied. To prove that, if $m^* \in B$, the outcome $l+1$ occurs only with small probability in the modified BlindForge experiment, it thus suffices to prove the following lemma.

Lemma 21. *In the* Quantum independent world, *for blinded messages, the outcome $l+1$ only occurs with small probability,*

$$\left\| \tilde{Q}_{l+1} \Pi_M^B |\psi_1'\rangle_{M\Sigma XYE\Gamma} \right\|_2 \leq q_1(\delta_W(n) + 2l(w-1)\epsilon_W(n))$$

where

$$\Pi^B = \sum_{m \in B} |m\rangle\langle m|.$$

Proof. The proof of this lemma is exactly the same as the one of Lemma 14 of the Lamport OTS, ecxept that it uses Lemma 20 instead of Lemma 13.

□

Now, all is set to prove Theorem 4.

Proof of Theorem 4. Just like in Section 5.4 we want to bound the final adversary's success probability in winning the BlindForge game in the general case. Towards that end, we remark that the computations are the same as those done in Section 4.5. Thus we proceed as in the Lamport OTS by deriving the bound respectively with respect to the modified BlindForge experiment and the the real experiment. Then, putting all our arguments together, we get the following:

$$\Pr_{QI,\mathsf{BlindForge}}[\mathcal{A} \text{ succeeds}] \leq (l+1)\left(2^{-n} + q^2(\delta_W(n) + 4l\epsilon_W(n))^2\right).$$

Finally, plugging in the functions $\epsilon_W(n)$ and $\delta_W(n)$ from Lemmas 8 and 18, and applying Lemma 4, we obtain

$$\begin{aligned}\Pr_{\mathsf{BlindForge}}[\mathcal{A} \text{ succeeds}] &\leq (l+1)\left(2^{-n} + q^2\left(\frac{8l(w+1)(w-1)}{2^{n/2}} + 2l(w-1)\frac{6(w-1)}{2^{n/2}}\right)^2\right) + \frac{3(wl)^2}{2^n} \\ &\leq 2^{-n}\left[\left(1 + q^2l^2(w-1)^2(20w-4)^2\right)(l+1) + 3w^2l^2\right] \\ &\leq 800w^4q^2l^3 \cdot 2^{-n}.\end{aligned}$$

□

6 Tightness

The notion of blind-unforgeability does not have as close of a relation to the intuitive security property it strives to model as EU-CMA.[3] The concrete security bounds, however, arguably nevertheless provide an indication of concrete security levels. It is hence an interesting question whether the bounds proven in Sections 4 and 5 are tight. In the following, we present an attack against the BU security of the Lamport scheme in the QROM, and analyze its success probability, to show that the bound in Theorem 3 is tight up to a factor l in the number of queries. The attack generalizes to the Winternitz scheme in a straight-forward manner.

We begin by describing the straightforward classical attack based on search. This attack proceeds as follows: To attack the BU security of the Lamport scheme, choose a blinding probability of $\frac{1}{2}$. Now make q distinct queries to the random oracle to search for a preimage of one of the $2l$ public key strings. This succeeds with probability

$$p_{\text{search}}(q) = 1 - \left(1 - \frac{2l}{2^n}\right)^q \geq \frac{2ql}{2^n}. \tag{101}$$

[3] Indeed, it is a nice exercise to show that an adversary against (say, q-time) EU-CMA with success probability ϵ can be used to construct a BU-adversary with success probability $\theta(\frac{\epsilon}{q})$, and this reduction is tight for efficient adversaries if one-way functions exists.

Suppose this search succeeded, finding a preimage y^* of $p_{i^*}^{j^*}$. Then chose $m \in \{0,1\}^l$ such that $m_{i^*} = \bar{j}^*$ and query the oracle to obtain a signature for m. This succeeds with probability $1/2$. Now output m' obtained from m by flipping the i^*th bit, and σ' obtained from σ by replacing σ_{i^*} with y^*. m' is blinded with probability $1/2$, and y^* is equal to the correct secret key string $s_{i^*}^{j^*}$ with constant probability. In summary, the entire attack succeeds with constant probability if

$$q = \Omega\left(\frac{2^n}{l}\right). \tag{102}$$

It is now easy to see that the search step can be replaced by a Grover search in the QROM. Using the analysis of Grover's algorithm for multiple targets from [BHT97], together with a basic analysis of the number of targets (which follows a binomial distribution), it is easy to see that one can achieve a constant success probability if

$$q = \Omega\left(\sqrt{\frac{2^n}{l}}\right). \tag{103}$$

To compare this result with Theorem 3, note that the inequality in Theorem 3, eq. (28), implies that to achieve a constant success probability, at least

$$q \geq C\sqrt{\frac{2^n}{l^3}} \tag{104}$$

are necessary for some constant C, i.e. the upper and lower bounds on the number of queries the optimal attack requires indeed differ by a factor of l up to constant factors. For the Winternitz scheme, the bounds differ by a factor of $w^2 l$.

Acknowledgements. The authors thank Stacey Jeffery for helpful discussions. CM was funded by a NWO VENI grant (Project No. VI.Veni.192.159). CMM deeply thanks the African Institute for Mathematical Science, Quantum Leap Africa Rwanda and QuSoft Amsterdam for their funding and support. MO was supported by NWO Vidi grant VI.Vidi.192.109.

References

[AASA+20] Gorjan Alagic, Jacob Alperin-Sheriff, Daniel Apon, David Cooper, Quynh Dang, John Kelsey, Yi-Kai Liu, Carl Miller, Dustin Moody, Rene Peralta, et al. Status report on the second round of the NIST post-quantum cryptography standardization process. *US Department of Commerce, NIST*, 2020. doi:10.6028/NIST.IR.8309.

[AMRS20] Gorjan Alagic, Christian Majenz, Alexander Russell, and Fang Song. Quantum-access-secure message authentication via blind-unforgeability. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 788–817. Springer, 2020. URL: https://ia.cr/2018/1150, arXiv:1803.03761, doi:10.1007/978-3-030-45727-3_27.

[BDF+11] Dan Boneh, Özgür Dagdelen, Marc Fischlin, Anja Lehmann, Christian Schaffner, and Mark Zhandry. Random oracles in a quantum world. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 41–69. Springer, 2011. URL: https://ia.cr/2010/428, arXiv:1008.0931, doi:10.1007/978-3-642-25385-0_3.

[BDH11] Johannes Buchmann, Erik Dahmen, and Andreas Hülsing. XMSS - a practical forward secure signature scheme based on minimal security assumptions. In Bo-Yin Yang, editor, *Post-Quantum Cryptography*, pages 117–129, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg. URL: https://ia.cr/2011/484, doi:10.1007/978-3-642-25405-5_8.

[BHK+19] Daniel J. Bernstein, Andreas Hülsing, Stefan Kölbl, Ruben Niederhagen, Joost Rijneveld, and Peter Schwabe. The SPHINCS+ signature framework. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, CCS'19, pages 2129–2146, New York, NY, USA, 2019. Association for Computing Machinery. URL: https://ia.cr/2019/1086, doi:10.1145/3319535.3363229.

[BHNP+19] Xavier Bonnetain, Akinori Hosoyamada, María Naya-Plasencia, Yu Sasaki, and André Schrottenloher. Quantum attacks without superposition queries: The offline Simon's algorithm. In Steven D. Galbraith and Shiho Moriai, editors, *Advances in Cryptology – ASIACRYPT 2019*, pages 552–583, Cham, 2019. Springer International Publishing. arXiv:2002.12439, doi:10.1007/978-3-030-34578-5_20.

[BHT97] Gilles Brassard, Peter Hoyer, and Alain Tapp. Quantum algorithm for the collision problem. 1997. URL: https://arxiv.org/pdf/quant-ph/9705002v1.pdf, arXiv:quant-ph/9705002, doi:10.1007/BFb0054319.

[BR93] Mihir Bellare and Phillip Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In *Proceedings of the 1st ACM conference on Computer and communications security*, pages 62–73. ACM, 1993. URL: https://caislab.kaist.ac.kr/lecture/2010/spring/cs548/basic/B11.pdf, doi:10.1145/168588.168596.

[BZ13] Dan Boneh and Mark Zhandry. Secure signatures and chosen ciphertext security in a quantum computing world. In Ran Canetti and Juan A. Garay, editors, *Advances in Cryptology – CRYPTO 2013*, pages 361–379, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg. URL: https://ia.cr/2013/088, doi:10.1007/978-3-642-40084-1_21.

[CFHL20] Kai-Min Chung, Serge Fehr, Yu-Hsuan Huang, and Tai-Ning Liao. On the compressed-oracle technique, and post-quantum security of proofs of sequential work, 2020. URL: https://ia.cr/2020/1305, arXiv:2010.11658.

[DSS05] Chris Dods, Nigel P Smart, and Martijn Stam. Hash based digital signature schemes. In *IMA International Conference on Cryptography and Coding*, pages 96–115. Springer, 2005. doi:10.1007/11586821_8.

[GHHM20] Alex B. Grilo, Kathrin Hövelmanns, Andreas Hülsing, and Christian Majenz. Tight adaptive reprogramming in the QROM, 2020. URL: https://ia.cr/2020/1361, arXiv:2010.15103.

[GHS16] Tommaso Gagliardoni, Andreas Hülsing, and Christian Schaffner. Semantic security and indistinguishability in the quantum world. In Matthew Robshaw and Jonathan Katz, editors, *Advances in Cryptology – CRYPTO 2016*, pages 60–89, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg. URL: https://ia.cr/2015/355, arXiv:1504.05255, doi:10.1007/978-3-662-53015-3_3.

[GKS20] Tommaso Gagliardoni, Juliane Krämer, and Patrick Struck. Quantum indistinguishability for public key encryption, 2020. URL: https://ia.cr/2020/266, arXiv:2003.00578.

[GYZ17] Sumegha Garg, Henry Yuen, and Mark Zhandry. New security notions and feasibility results for authentication of quantum data. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology – CRYPTO 2017*, pages 342–371, Cham, 2017. Springer International Publishing. URL: https://eprint.iacr.org/2017/538.pdf, doi:10.1007/978-3-319-63715-0_12.

[Hay02] Masahito Hayashi. Optimal sequence of quantum measurements in the sense of stein's lemma in quantum hypothesis testing. *Journal of Physics A: Mathematical and General*, 35(50):10759, 2002. arXiv:quant-ph/0208020.

[HBG$^+$18] Andreas Hülsing, Denise Butin, Stefan-Lukas Gazdag, Joost Rijneveld, and Aziz Mohaisen. Xmss: Extended hash-based signatures. rfc 8391, 2018. URL: https://tools.ietf.org/html/rfc8391.

[KLLNP16] Marc Kaplan, Gaëtan Leurent, Anthony Leverrier, and María Naya-Plasencia. Breaking symmetric cryptosystems using quantum period finding. In Matthew Robshaw and Jonathan Katz, editors, *Advances in Cryptology – CRYPTO 2016*, pages 207–237, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg. arXiv:1602.05973, doi:10.1007/978-3-662-53008-5_8.

[Lam79a] Leslie Lamport. Constructing digital signatures from a one way function. Technical Report SRI-CSL-98, SRI International Computer Science Laboratory, 1979. URL: http://lamport.azurewebsites.net/pubs/dig-sig.pdf.

[Lam79b] Leslie Lamport. Constructing digital signatures from a one-way function. Technical report, Technical Report CSL-98, SRI International, 1979.

[LZ19] Qipeng Liu and Mark Zhandry. On finding quantum multi-collisions. In Yuval Ishai and Vincent Rijmen, editors, *Advances in Cryptology – EUROCRYPT 2019*, pages 189–218, Cham, 2019. Springer International Publishing. arXiv:1811.05385, doi:10.1007/978-3-030-17659-4_7.

[Mer89] Ralph C Merkle. A certified digital signature. In *Conference on the Theory and Application of Cryptology*, pages 218–238. Springer, 1989. doi:10.1007/0-387-34805-0_21.

[NC02] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002. URL: http://csis.pace.edu/~ctappert/cs837-19spring/QC-textbook.pdf, doi:10.1023/A:1012603118140.

[Sho94] Peter W Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. IEEE, 1994. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.123.5183&rep=rep1&type=pdf, doi:10.1109/SFCS.1994.365700.

[SS17] Thomas Santoli and Christian Schaffner. Using Simon's algorithm to attack symmetric-key cryptographic primitives. *Quantum Info. Comput.*, 17(1–2):65–78, February 2017. arXiv:1603.07856, doi:10.26421/QIC17.1-2-4.

[Unr21] Dominique Unruh. Compressed permutation oracles (and the collision-resistance of sponge/SHA3). Cryptology ePrint Archive, Report 2021/062, 2021. URL: https://ia.cr/2021/062.

[Wat18] John Watrous. *The theory of quantum information*. Cambridge University Press, 2018. URL: https://cs.uwaterloo.ca/~watrous/TQI/TQI.pdf.

[Zha15] Mark Zhandry. Secure identity-based encryption in the quantum random oracle model. *International Journal of Quantum Information*, 13(04):1550014, 2015. URL: https://ia.cr/2012/076, doi:10.1142/S0219749915500148.

[Zha19] Mark Zhandry. How to record quantum queries, and applications to quantum indifferentiability. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019*, pages 239–268, Cham, 2019. Springer International Publishing. URL: https://www.cs.princeton.edu/~mzhandry/docs/papers/QIndiff.pdf, doi:10.1007/978-3-030-26951-7_9.

A Lemmas used for proving the security of the Lamport OTS

A.1 Proof of Lemmas 8 and 15

Here, we prove that the unitary U_h does not significantly modify the secret key register. Towards that end, we show that the random oracle unitary U_h approximately commute with the projector onto the relevant parts of the secret key register being in state $|\Phi\rangle$.

Lemma 15. *Let U_h be the random oracle unitary for any given function h (see Section 3) and let $\Phi = |\Phi\rangle\langle\Phi|$ denote the projector onto the uniform superposition $|\Phi\rangle$. Furthermore, let $\Gamma_i^{\leq j} = \Gamma_i^0 \dots \Gamma_i^j$ and*

$$\Phi_{\Gamma_i^{\leq j}} = \left(\Phi^{\otimes j}\right)_{\Gamma_i^{\leq j}}. \tag{81}$$

Then, for any $i' \in \{1, \dots, l\}$ and $j' \in \{0, \dots, w-2\}$,

$$\left\| \left[(U_h)_{XY\Gamma}, \Phi_{\Gamma_{i'}^{\leq j'}} \right] \right\|_\infty \leq \frac{6(w-1)}{2^{n/2}} = \epsilon_W(n) \tag{82}$$

is negligible in n.

Proof. Recall from Section 3.1 that U_h compares the input in register X to the secret key in Γ, and stores the output in register Y. Recall from eq. (19) that U_h is defined as follows:

$$(U_h)_{XY\Gamma} = \left(\prod_{i=1}^{l} \prod_{j=0}^{w-2} (U_i^j)_{XY\Gamma_i^j} \right) U^{\neq}_{XY\Gamma}.$$

Using Lemma 2, we get

$$\left\| \left[U_h, \Phi_{\Gamma_{i'}^{\leq j'}} \right] \right\|_\infty \leq \sum_{i=1}^{l} \sum_{j=0}^{w-2} \left\| \left[U_i^j, \Phi_{\Gamma_{i'}^{\leq j'}} \right] \right\|_\infty + \left\| \left[U^{\neq}_{XY\Gamma}, \Phi_{\Gamma_{i'}^{\leq j'}} \right] \right\|_\infty. \tag{105}$$

We will bound the two terms separately.

To deal with the first term, recall from eq. (18) that

$$(U_i^j)_{XY\Gamma_i^j\Gamma_i^{j+1}} = P^{=}_{X\Gamma_i^j} \otimes \left((\mathrm{CNOT}^{\otimes n})_{\Gamma_i^{j+1}:Y} - \mathbb{1}\right) + \mathbb{1}, \tag{106}$$

so the expression in first term of eq. (105) is given by

$$\left\|\left[U_i^j, \Phi_{\Gamma_{i'}^{\leq j'}}\right]\right\|_\infty = \left\|\left[P^{=}_{X\Gamma_i^j} \otimes \left((\mathrm{CNOT}^{\otimes n})_{\Gamma_i^{j+1}:Y} - \mathbb{1}\right), \Phi_{\Gamma_{i'}^{\leq j'}}\right]\right\|_\infty. \tag{107}$$

If $i \neq i'$ or $j' < j$, the commutator vanishes because on every register at least one of the two operators acts as the identity matrix. Let us assume now that $i = i'$ and $j' \geq j$. Then we can upper bound the norm in eq. (107) by noting that $\Phi_{\Gamma_i^{\leq j'}} = \Phi_{\Gamma_i^j} \otimes \Phi'$ where Φ' is a projector on the remaining registers, and then separating $X\Gamma_i^j$ out from the remaining registers. Starting with the triangle inequality and then using this observation,

$$\begin{aligned}
\left\|\left[U_i^j, \Phi_{\Gamma_i^{\leq j'}}\right]\right\|_\infty &\leq 2\left\|\left(P^{=}_{X\Gamma_i^j} \otimes \left((\mathrm{CNOT}^{\otimes n})_{\Gamma_i^{j+1}:Y} - \mathbb{1}\right)\right) \cdot \Phi_{\Gamma_i^{\leq j'}}\right\|_\infty \\
&\leq 2\left\|P^{=}_{X\Gamma_i^j}\Phi_{\Gamma_i^j}\right\|_\infty \left\|(\mathrm{CNOT}^{\otimes n})_{\Gamma_i^{j+1}:Y} - \mathbb{1}\right\|_\infty \|\Phi'\|_\infty \\
&\leq 4\left\|P^{=}_{X\Gamma_i^j}\Phi_{\Gamma_i^j}\right\|_\infty \\
&\leq 4 \cdot 2^{-n/2},
\end{aligned}$$

where we used $\|(\mathrm{CNOT}^{\otimes n})_{\Gamma_i^{j+1}:Y} - \mathbb{1}\|_\infty \leq 2$ and $\|\Phi'\|_\infty \leq 1$ to obtain the third inequality, and then Lemma 3 for the last inequality. In summary,

$$\left\|\left[U_i^j, \Phi_{\Gamma_{i'}^{\leq j'}}\right]\right\|_\infty \leq \delta_{i,i'}\delta_{j' \geq j} \cdot 4 \cdot 2^{-n/2}, \tag{108}$$

so the first term in eq. (105) can be upper bounded as

$$\sum_{i=1}^{l}\sum_{j=0}^{w-2}\left\|\left[U_i^j, \Phi_{\Gamma_{i'}^{\leq j'}}\right]\right\|_\infty \leq 4(j'+1) \cdot 2^{-n/2}. \tag{109}$$

To upper bound the second term in eq. (105), recall from eq. (21) that $U^{\neq}_{XY\Gamma} = P^{\neq}_{X\Gamma} \cdot (U'_{XY} - \mathbb{1}) + \mathbb{1}$. Using Lemma 2, we can inductively expand the projectors $P^{\neq}_{X\Gamma}$ and $\Phi_{\Gamma_{i'}^{\leq j'}}$ defined in eqs. (22) and (81) to get

$$\begin{aligned}
\left\|\left[U^{\neq}_{XY\Gamma}, \Phi_{\Gamma_{i'}^{\leq j'}}\right]\right\|_\infty &= \left\|\left[P^{\neq}_{X\Gamma} \cdot (U'_{XY} - \mathbb{1}), \Phi_{\Gamma_{i'}^{\leq j'}}\right]\right\|_\infty \\
&\leq \sum_{i=1}^{l}\sum_{j=0}^{w-2}\sum_{k=0}^{j'}\left\|\left[P^{\neq}_{X\Gamma_i^j} \cdot (U'_{XY} - \mathbb{1}), \Phi_{\Gamma_{i'}^k}\right]\right\|_\infty.
\end{aligned} \tag{110}$$

We can simplify this further by using Lemma 2 once again:

$$\left\|\left[P^{\neq}_{X\Gamma_i^j} \cdot (U'_{XY} - \mathbb{1}), \Phi_{\Gamma_{i'}^k}\right]\right\|_\infty \leq \left\|\left[P^{\neq}_{X\Gamma_i^j}, \Phi_{\Gamma_{i'}^k}\right]\right\|_\infty + \left\|\left[U'_{XY} - \mathbb{1}, \Phi_{\Gamma_{i'}^k}\right]\right\|_\infty.$$

The second term vanishes since the query operation U'_{XY} from eq. (23) acts trivially on $\Gamma_{i'}^k$. Furthermore, we can replace $P^{\neq}$ by $P^{=}$ in the first term since the two operators differ by $\mathbb{1}$, which commutes with everything. With these observations, eq. (110) simplifies to

$$\left\|\left[U^{\neq}_{XY\Gamma}, \Phi_{\Gamma_{i'}^{\leq j'}}\right]\right\|_\infty \leq \sum_{i=1}^{l}\sum_{j=0}^{w-2}\sum_{k=0}^{j'}\left\|\left[P^{=}_{X\Gamma_i^j}, \Phi_{\Gamma_{i'}^k}\right]\right\|_\infty.$$

Using Lemma 3, we obtain the following bound on each term:

$$\left\|\left[P^{=}_{X\Gamma_i^j}, \Phi_{\Gamma_{i'}^k}\right]\right\|_\infty \leq 2\delta_{i,i'}\delta_{j,k} \cdot 2^{-n/2}.$$

Putting these together, the second term in eq. (105) can be upper bounded as

$$\left\| \left[U^{\neq}_{XY\Gamma}, \Phi_{\Gamma^{\leq j'}_{i'}} \right] \right\|_\infty \leq 2(j'+1) \cdot 2^{-n/2}.$$

Combining this with eq. (109), both terms in eq. (105) can be upper bounded as

$$\left\| \left[U_h, \Phi_{\Gamma^{\leq j'}_{i'}} \right] \right\|_\infty \leq 6(j'+1)2^{-n/2} \leq 6(w-1) \cdot 2^{-n/2} \tag{111}$$

where we used $j' \leq w-2$. □

As a corollary, we obtain the corresponding lemma for the Lamport OTS by setting $w = 2$.

Lemma 8. *Let U_h be the random oracle unitary for any given function h (see Section 3) and let $\Phi = |\Phi\rangle\langle\Phi|$ denote the projector onto the uniform superposition. Then, for any $i \in \{1, \ldots, l\}$ and $j \in \{0, 1\}$,*

$$\left\| \left[U_h, \Phi_{S^j_i} \right] \right\|_\infty \leq \frac{6}{2^{n/2}} = \epsilon_L(n) \tag{35}$$

is negligible in n.

A.2 Proof of Lemma 9

Lemma 9. *For all $m^* \in B$, the projectors $Q^{m^*}_{l+1}$ and P_S defined in eqs. (30) and (34) are orthogonal:*

$$Q^{m^*}_{l+1} P_S = 0. \tag{36}$$

Proof. Recall from eqs. (30), (32) and (34) that

$$Q^{m^*}_{l+1} = \bigotimes_{i=1}^{l} \Phi^\perp_{S_i^{m^*_i}}, \qquad\qquad P_S = \sum_{\alpha \in \widehat{B^c}} \Phi(\alpha)_S = \sum_{\alpha \in \widehat{B^c}} \bigotimes_{i=1}^{l} \bigotimes_{j=0}^{1} \Phi(\alpha^j_i)_{S^j_i}, \tag{112}$$

where $\Phi(0) = |\Phi\rangle\langle\Phi| = \Phi$ and $\Phi(1) = \mathbb{1} - |\Phi\rangle\langle\Phi| = \Phi^\perp$, and the set $\widehat{B^c}$ was defined in eq. (33) as

$$\widehat{B^c} = \bigcup_{m \in B^c} \left\{ \alpha \in \{0,1\}^{2l} \,\middle|\, \alpha_i^{\bar{m}_i} = 0 \text{ for all } i = 1, \ldots, l \right\}. \tag{113}$$

Substituting the expression for $Q^{m^*}_{l+1}$,

$$Q^{m^*}_{l+1} P_S = \sum_{\alpha \in \widehat{B^c}} \left(\left((\Phi^\perp)^{\otimes l} \right)_{S_1^{m^*_1} \ldots S_l^{m^*_l}} \otimes \mathbb{1}_{S_1^{\bar{m}^*_1} \ldots S_l^{\bar{m}^*_l}} \right) \Phi(\alpha). \tag{114}$$

Consider a single term $\alpha \in \widehat{B^c}$ in the sum. By definition of $\widehat{B^c}$, there exists a message $m \in B^c$ such that $\alpha_i^{\bar{m}_i} = 0$ for all i. But as $B \ni m^* \neq m \in B^c$, there exists an index i such that $m^*_i \neq m_i$, or equivalently $m^*_i = \bar{m}_i$. Therefore we can simplify the corresponding term to

$$\begin{aligned}
&\left(\left((\Phi^\perp)^{\otimes l} \right)_{S_1^{m^*_1} \ldots S_l^{m^*_l}} \otimes \mathbb{1}_{S_1^{\bar{m}^*_1} \ldots S_l^{\bar{m}^*_l}} \right) \Phi(\alpha) \\
&= \left(\Phi^\perp_{S_1^{m^*_1}} \otimes \cdots \otimes \Phi^\perp_{S_i^{m^*_i}} \otimes \cdots \otimes \Phi^\perp_{S_l^{m^*_l}} \otimes \mathbb{1}_{S_1^{\bar{m}^*_1} \ldots S_l^{\bar{m}^*_l}} \right) \\
&\quad \left(\Phi_{S_1^{\bar{m}_1}} \otimes \cdots \otimes \Phi_{S_i^{\bar{m}_i}} \otimes \cdots \otimes \Phi_{S_l^{\bar{m}_l}} \otimes \Phi(\alpha_1^{m_1})_{S_1^{m_1}} \otimes \cdots \otimes \Phi(\alpha_l^{m_l})_{S_l^{m_l}} \right) \\
&= \cdots \otimes \underbrace{\left(\Phi^\perp \Phi \right)_{S_i^{m^*_i}}}_{=0} \otimes \cdots \\
&= 0.
\end{aligned}$$

Following the same analysis for all the terms in eq. (114) leads to the required result. □

A.3 Proof of Lemma 10

Here, we show that if there are no hash queries before the Sign query, projecting P_S onto the adversary's state after the Sign query leaves it completely in the range of P_S.

Lemma 10. *Let $B\,\mathsf{Sign}_{\mathsf{sk}}$ be the blinded signing oracle for the Lamport OTS and let $|\psi_0\rangle$ be the adversary's state before the* Sign *query. If there are no hash queries, then after making at most one* Sign *query the adversary's state $|\psi_1\rangle = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle$ is completely in the image of the invariant projector P_S defined in eq. (34). That is,*

$$P_S B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle. \tag{37}$$

Proof. We know that all secret key sub-registers are in state $|\Phi\rangle$ if there are no hash queries before Sign query, i.e.,

$$|\psi_0\rangle_{M\Sigma EBS} - \Phi_S^{\otimes 2l}|\psi_0\rangle_{M\Sigma EBS} = 0$$

where $|\psi_0\rangle_{M\Sigma EBS}$ is the adversary's state immediately before the Sign query. Let us denote the state after a single Sign query by

$$|\psi_1\rangle_{M\Sigma EBS} = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle_{M\Sigma EBS}. \tag{115}$$

We want to show that applying P_S onto the state $|\psi_1\rangle_{M\Sigma EBS}$ leaves it invariant. That is,

$$\left\| |\psi_1\rangle_{M\Sigma EBS} - P_S|\psi_1\rangle_{M\Sigma EBS} \right\|_2 = 0. \tag{116}$$

Recall that

$$\begin{aligned} P_S|\psi_1\rangle_{M\Sigma EBS} &= P_S B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle_{M\Sigma EBS} && (117)\\ &= P_S B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0^1\rangle_{M\Sigma EBS} + P_S B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0^0\rangle_{M\Sigma EBS}, && (118)\end{aligned}$$

where the superscripts 1 and 0 refer to blinded (B) and un-blinded (B^c) messages, respectively. Recall from eq. (47) that:

$$|\psi_0^1\rangle_{M\Sigma BES} = \left(\sum_{m\in B} \sum_{\sigma\in(\{0,1\}^n)^l} \kappa_{m\sigma 1}|m\rangle_M|\sigma\rangle_\Sigma|1\rangle_B|\alpha_{m\sigma 1}\rangle_E \right) \otimes \left(|\Phi\rangle^{\otimes 2l} \right)_S \tag{119}$$

$$|\psi_0^0\rangle_{M\Sigma BES} = \left(\sum_{m\in B^c} \sum_{\sigma\in(\{0,1\}^n)^l} \kappa_{m\sigma 0}|m\rangle_M|\sigma\rangle_\Sigma|0\rangle_B|\alpha_{m\sigma 0}\rangle_E \right) \otimes \left(|\Phi\rangle^{\otimes 2l} \right)_S \tag{120}$$

We start by deriving the first term of the right-hand side of eq. (118). If there are no hash queries before the Sign query, the secret key register is fully in uniform superposition state $|\Phi\rangle$ before the single Sign query. Since we are dealing with the first term that corresponds to blinded messages, there is no signature. Thus, the joint state of all secret key registers is in the range of the projector Φ and

$$P_S B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0^1\rangle_{M\Sigma EBS} = |\psi_0^1\rangle_{M\Sigma EBS}. \tag{121}$$

Next, we evaluate the second term of eq. (118) that corresponds to un-blinded messages. We start by setting

$$\begin{aligned} |\tau\rangle_{M\Sigma BES} &= P_S B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0^0\rangle_{M\Sigma EBS} \\ &= \sum_{m\in B^c} P_S B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\Phi\rangle_S^{\otimes 2l} \end{aligned}$$

where

$$|\gamma(m)\rangle_{\Sigma EB} = \sum_{\sigma\in(\{0,1\}^n)^l} \kappa_{m\sigma 0}|m\rangle_M|\sigma\rangle_\Sigma|0\rangle_B|\alpha_{m\sigma}\rangle_E.$$

From the definition of $\widehat{B^c}$ in eq. (113) we know that for any message $m \in B^c$ there exists an $i \in \{1,\dots,l\}$ for which the secret key register $S_i^{\bar{m}_i}$ corresponding to $\bar{m}_i$ is in the range of the projector Φ. Since the projector $P_S = \sum_{\alpha\in\widehat{B^c}} \Phi(\alpha)_S$ acts on the entire secret key register, we can split it into two sums where the first is over all the projectors for which $\alpha_i^{\bar{m}_i} = 0$ (i.e., $S_i^{\bar{m}_i}$ is in the range of $\Phi(0) = \Phi$) and the second is over the remaining projectors ($\alpha_i^{\bar{m}_i} = 1$ for at least one i). Accordingly, for any message $m \in B^c$, we can split P_S as

$$P_S = \sum_{\alpha\in\widehat{B^c}} \Phi(\alpha) = \sum_{\substack{\alpha\in\widehat{B^c}\\ \forall i:\alpha_i^{\bar{m}_i}=0}} \Phi(\alpha) + \sum_{\substack{\alpha\in\widehat{B^c}\\ \exists i:\alpha_i^{\bar{m}_i}=1}} \Phi(\alpha) = P_S^1(m) + P_S^2(m). \tag{122}$$

Note that this split can differ from one message $m \in B^c$ to another. Therefore, $|\tau\rangle_{M\Sigma BES}$ becomes

$$|\tau\rangle_{M\Sigma BES} = \sum_{m\in B^c} \big(P_S^1(m) + P_S^2(m)\big) B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\Phi\rangle_S^{\otimes 2l}. \tag{123}$$

For a fixed message m, $P_S^1(m)$ can be written as

$$P_S^1(m) = \sum_{\substack{\alpha\in\widehat{B^c}\\ \forall i:\alpha_i^{\bar{m}_i}=0}} \Phi(\alpha) = \mathbb{1}_{S_1^{m_1}\ldots S_l^{m_l}} \otimes (\Phi^{\otimes l})_{S_1^{\bar{m}_1}\ldots S_l^{\bar{m}_l}}.$$

Thus,

$$\begin{aligned} &P_S^1(m) B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\Phi\rangle_S^{\otimes 2l} \\ &= B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}\Big(\mathbb{1}_{S_1^{m_1}\ldots S_l^{m_l}} \otimes (\Phi^{\otimes l})_{S_1^{\bar{m}_1}\ldots S_l^{\bar{m}_l}}\Big)|\Phi\rangle_S^{\otimes 2l} \\ &= B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\Phi\rangle_S^{\otimes 2l} \end{aligned} \tag{124}$$

where the second equality follows because $P_S^1(m)$ and $B\,\mathsf{Sign}_{\mathsf{sk}}$ act on different sub-registers. More precisely, $P_S^1(m)$ acts on the secret key registers corresponding to the complementary message $\bar{m}$ while $B\,\mathsf{Sign}_{\mathsf{sk}}$ acts on the secret key registers corresponding to the message m, i.e.

$$[P_S^1(m), B\,\mathsf{Sign}_{\mathsf{sk}}]|m\rangle_M = 0.$$

This is only true for the fixed message m.

Note that $P_S^1(m)$ and $P_S^2(m)$ are orthogonal, i.e.

$$P_S^2(m)P_S^1(m) = 0.$$

So eq. (124) implies

$$P_S^2(m)B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\Phi\rangle_S^{\otimes 2l} = P_S^2(m)P_S^1(m)B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\Phi\rangle_S^{\otimes 2l} = 0.$$

Adding to this the result obtained in eq. (121), we get

$$P_S B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\Phi\rangle_S^{\otimes 2l} = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\Phi\rangle_S^{\otimes 2l}.$$

Hence,

$$\big\| |\psi_1\rangle_{M\Sigma EBS} - P_S|\psi_1\rangle_{M\Sigma EBS}\big\|_2 = 0 \tag{125}$$

by eq. (123). □

A.4 Proof of Lemma 11

Lemma 11. *The invariant projector P_S defined in eq. (34) and the random oracle unitary U_h in the* Quantum independent world, *see eq. (19), approximately commute, i.e.,*

$$\big\|[U_h, P_S]\big\|_\infty \leq \delta_L(n), \tag{38}$$

where

$$\delta_L(n) = \frac{32l}{2^{n/2}}$$

is negligible in n.

Proof. Recall from eq. (19) that U_h is defined as follows:

$$(U_h)_{XYS} = \left(\prod_{i=1}^{l}\prod_{j=0}^{1}(U_i^j)_{XYS_i^j}\right)U_{XYS}^{\neq}. \tag{126}$$

Similar to eq. (18), the unitary U_i^j is defined as

$$(U_i^j)_{XY\Gamma_i^j\Gamma_i^{j+1}} = P_{X\Gamma_i^j}^{=} \otimes \Big((\mathrm{CNOT}^{\otimes n})_{\Gamma_i^{j+1}:Y} - \mathbb{1}\Big) + \mathbb{1}, \tag{127}$$

where for the Lamport OTS $j = 0$ and Γ_i^{j+1} represents the register that stores the i-th n-bit string of the public key which is described in Section 2.3.1. More precisely, in the Lamport OTS, the pair of indices (i, j) of the secret and public keys takes the role of the index i in the Winternitz scheme, and whether a string is part of the secret key ($j = 0$) or the public key ($j = 1$) in the Lamport scheme determines the value of the index j from the Winternitz scheme. If we introduce a register P_i^j that stores the (i, j)-th block of the public key, eq. (127) for the Lamport OTS becomes

$$(U_{i,j})_{XYS_i^jP_i^j} = P^{=}_{XS_i^j} \otimes \left((\mathrm{CNOT}^{\otimes n})_{P_i^j:Y} - \mathbb{1}\right) + \mathbb{1}, \tag{128}$$

Also, similar to eq. (21),

$$U^{\neq}_{XYS} = P^{\neq}_{XS}U'_{XY} + (\mathbb{1}_{XS} - P^{\neq}_{XS}) \otimes \mathbb{1}_Y. \tag{129}$$

By substituting the formula for U_h from eq. (126), we get

$$\begin{aligned}
\|[U_h, P_S]\|_\infty &= \left\| \left[\Big(\prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j} \Big) U^{\neq}_{XYS}, P_S \right] \right\|_\infty & (130)\\
&= \left\| \left[\prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j}, P_S \right] U^{\neq}_{XYS} + \prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j} \left[U^{\neq}_{XYS}, P_S \right] \right\|_\infty & (131)\\
&\leq \left\| \left[\prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j}, P_S \right] U^{\neq}_{XYS} \right\|_\infty + \left\| \prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j} \left[U^{\neq}_{XYS}, P_S \right] \right\|_\infty & (132)\\
&\leq \left\| \left[\prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j}, P_S \right] \right\|_\infty \left\| U^{\neq}_{XYS} \right\|_\infty + \left\| \prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j} \right\|_\infty \left\| \left[U^{\neq}_{XYS}, P_S \right] \right\|_\infty, & (133)
\end{aligned}$$

where the first equality follows from the definition of U_h, the second follows from eq. (4), and the two inequalities follow respectively from the triangle inequality and the sub-multiplicative property of the operator norm. Since $U^{\neq}_{XYS}$ is an unitary, $\|U^{\neq}_{XYS}\|_\infty = 1$. Similarly,

$$\left\| \prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j} \right\|_\infty \leq \prod_{\substack{i=1,\dots,l\\ j=0,1}} \left\| (U_{i,j})_{XYS_i^jP_i^j} \right\|_\infty \leq 1$$

and hence

$$\|[U_h, P_S]\|_\infty \leq \left\| \left[\prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j}, P_S \right] \right\|_\infty + \left\| \left[U^{\neq}_{XYS}, P_S \right] \right\|_\infty.$$

Substituting $U^{\neq}_{XYS}$ from eq. (129) and using the triangle inequality lead to

$$\begin{aligned}
\|[U_h, P_S]\|_\infty &\leq \left\| \left[\prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j}, P_S \right] \right\|_\infty + \left\| \left[P^{\neq}_{XS}U'_{XY} + (\mathbb{1}_{XS} - P^{\neq}_{XS}) \otimes \mathbb{1}_Y, P_S \right] \right\|_\infty & (134)\\
&\leq \left\| \left[\prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j}, P_S \right] \right\|_\infty + \left\| \left[P^{\neq}_{XS}U'_{XY}, P_S \right] \right\|_\infty + \left\| \left[(\mathbb{1}_{XS} - P^{\neq}_{XS}) \otimes \mathbb{1}_Y, P_S \right] \right\|_\infty & (135)
\end{aligned}$$

where U'_{XY} is the standard random oracle. We will now bound separately each term of the latter equation.

Let us pick some register S_i^j and split up the invariant projector P_S into a sum of two terms according to whether they contain Φ or $\Phi^\perp = \mathbb{1} - \Phi$ on this register:

$$P_S = \Phi_{S_i^j} \otimes \tilde{\Phi}^0_{(S_i^j)^c} + \Phi^\perp_{S_i^j} \otimes \tilde{\Phi}^1_{(S_i^j)^c} \tag{136}$$

where, for $b \in \{0, 1\}$,

$$\tilde{\Phi}^b_{(S_i^j)^c} = \sum_{\substack{\alpha\in\widehat{B^c}\\ \alpha_i^j=b}} \bigotimes_{\substack{i'=1,\dots,l\\ j'=0,1\\ (i',j')\neq(i,j)}} \Phi(\alpha_{i'}^{j'})_{S_{i'}^{j'}}$$

is a sum of mutually orthogonal projectors and hence a projector itself. Consequently,

$$\|\tilde{\Phi}^b_{(S_i^j)^c}\|_\infty = 1. \tag{137}$$

Using Lemma 2 and substituting $U_{i,j}$ from eq. (128) gives,

$$\begin{aligned}
\left\|\left[\prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j}, P_S\right]\right\|_\infty &\leq \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[(U_{i,j})_{XYS_i^jP_i^j}, P_S\right]\right\|_\infty \\
&= \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{=}_{XS_i^j} \otimes \left((\mathrm{CNOT}^{\otimes n})_{P_i^j:Y} - \mathbb{1}\right) + \mathbb{1}, P_S\right]\right\|_\infty \\
&\leq \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{=}_{XS_i^j}, P_S\right]\right\|_\infty \underbrace{\left\|(\mathrm{CNOT}^{\otimes n})_{P_i^j:Y} - \mathbb{1}\right\|_\infty}_{=2} \\
&= 2 \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{=}_{XS_i^j}, P_S\right]\right\|_\infty
\end{aligned}$$

where the last inequality follows from the sub-multiplicative property of norm and the fact that $[\mathbb{1}, P_S] = 0$. Recall that here $j = 0$ and pk_i is the i-th public key prepared in the computational basis. Substituting P_S from eq. (136) gives

$$\begin{aligned}
\sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{=}_{XS_i^j}, P_S\right]\right\|_\infty &= \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{=}_{XS_i^j}, \Phi_{S_i^j} \otimes \tilde{\Phi}^0_{(S_i^j)^c} + \Phi^{\perp}_{S_i^j} \otimes \tilde{\Phi}^1_{(S_i^j)^c}\right]\right\|_\infty \\
&\leq \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{=}_{XS_i^j}, \Phi_{S_i^j}\right]\right\|_\infty \left\|\tilde{\Phi}^0_{(S_i^j)^c}\right\|_\infty + \left\|\left[P^{=}_{XS_i^j}, \mathbb{1} - \Phi_{S_i^j}\right]\right\|_\infty \left\|\tilde{\Phi}^1_{(S_i^j)^c}\right\|_\infty \\
&= 2 \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{=}_{XS_i^j}, \Phi_{S_i^j}\right]\right\|_\infty,
\end{aligned}$$

where we used the triangle inequality, sub-multiplicativity of the norm, and eq. (137). Recall from Lemma 3 that $\|[P^{=}_{XS_i^j}, \Phi_{S_i^j}]\|_\infty = 2 \cdot 2^{-n/2}$. Hence,

$$\sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{=}_{XS_i^j}, P_S\right]\right\|_\infty \leq 2 \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{=}_{XS_i^j}, \Phi_{S_i^j}\right]\right\|_\infty = 2 \cdot 2l \cdot 2 \cdot 2^{-n/2} = 8l \cdot 2^{-n/2}, \tag{138}$$

and the first term in eq. (135) can be upper bounded as

$$\left\|\left[\prod_{\substack{i=1,\dots,l\\ j=0,1}} (U_{i,j})_{XYS_i^jP_i^j}, P_S\right]\right\|_\infty \leq 2 \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{=}_{XS_i^j}, P_S\right]\right\|_\infty \leq 16l \cdot 2^{-n/2}.$$

Next, we bound the second term of eq. (135). Substituting

$$P^{\neq}_{XS} = \prod_{i=1}^{l} \prod_{j=0}^{1} P^{\neq}_{XS_i^j} \tag{139}$$

from eq. (22), we get

$$\begin{aligned}
\left\|\left[P^{\neq}_{XS}U'_{XY}, P_S\right]\right\|_\infty &= \left\|\left[\prod_{\substack{i=1,\dots,l\\ j=0,1}} P^{\neq}_{XS^j_i} U'_{XY}, P_S\right]\right\|_\infty \\
&\leq \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{\neq}_{XS^j_i}, P_S\right]\right\|_\infty + \Big\|\underbrace{\left[U'_{XY}, P_S\right]}_{=0}\Big\|_\infty \\
&= \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[\mathbb{1}_{XS^j_i} - P^{=}_{XS^j_i}, P_S\right]\right\|_\infty \\
&= \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[P^{=}_{XS^j_i}, P_S\right]\right\|_\infty \\
&\leq 8l \cdot 2^{-n/2}
\end{aligned}$$

where we used the triangle inequality, the fact that $\left[U'_{XY}, P_S\right] = 0$ because U'_{XY} and P_S act on different registers, and finally eq. (138) to obtain the last inequality.

Lastly, we bound the third term of eq. (135) as follows:

$$\begin{aligned}
&\left\|\left[\left(\mathbb{1}_{XS} - \prod_{\substack{i=1,\dots,l\\ j=0,1}} P^{\neq}_{XS^j_i}\right) \otimes \mathbb{1}_Y, P_S\right]\right\|_\infty \\
&= \Bigg\|\underbrace{\left[\mathbb{1}_{XS}, P_S\right]}_{=0} \otimes \mathbb{1}_Y - \left[\prod_{\substack{i=1,\dots,l\\ j=0,1}} P^{\neq}_{XS^j_i}, P_S\right] \otimes \mathbb{1}_Y\Bigg\|_\infty \\
&\leq \left\|\left[\prod_{\substack{i=1,\dots,l\\ j=0,1}} P^{\neq}_{XS^j_i}, P_S\right]\right\|_\infty \\
&\leq \sum_{\substack{i=1,\dots,l\\ j=0,1}} \left\|\left[\mathbb{1}_{XS^j_i} - P^{=}_{XS^j_i}, P_S\right]\right\|_\infty \\
&\leq 8l \cdot 2^{-n/2},
\end{aligned}$$

where the reasoning is similar to what we used for bounding the second term.

Replacing the three terms of eq. (135) by their respective bounds gives

$$\left\|\left[U_h, P_S\right]\right\|_\infty \leq 16l \cdot 2^{-n/2} + 8l \cdot 2^{-n/2} + 8l \cdot 2^{-n/2} = 32l \cdot 2^{-n/2},$$

as claimed. □

B Lemmas used for proving the security of the Winternitz OTS

In this appendix, we prove Lemmas 16 to 18 which are the main technical ingredients of Theorem 4 on security of the Winternitz OTS (another ingredient, Lemma 15, was already proved in Appendix A.1). The proofs of these lemmas are similar to those for the Lamport OTS in Appendix A.

B.1 Proof of Lemma 16

The proof of this lemma is similar to the proof of Lemma 10 in Appendix A.3.

Lemma 16. *Let $B\,\mathsf{Sign}_{\mathsf{sk}}$ be the blinded signing oracle for the Winternitz OTS, and let $|\psi_0\rangle$ be the adversary's state before the* Sign *query. If there are no hash queries, then after making a single* Sign *query the adversary's state $|\psi_1\rangle = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle$ is completely in the range of the invariant projector P_Γ defined in eq. (80). That is,*

$$P_\Gamma B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle. \tag{83}$$

Proof. If there are no hash queries before the Sign query, all hash chain registers (except for the last) are in state $|\Phi\rangle$, i.e.

$$|\psi_0\rangle_{M\Sigma EB\Gamma} - \Phi_\Gamma^{\otimes l(w-1)}|\psi_0\rangle_{M\Sigma EB\Gamma} = 0$$

where $|\psi_0\rangle_{M\Sigma EB\Gamma}$ is adversary's state immediately before the Sign query. Our goal is to show that applying the invariant projector P_Γ onto the post-signature state $|\psi_1\rangle_{M\Sigma EB\Gamma} = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle_{M\Sigma EB\Gamma}$ leaves it unchanged, i.e.

$$\left\| |\psi_1\rangle_{M\Sigma EB\Gamma} - P_\Gamma |\psi_1\rangle_{M\Sigma EB\Gamma} \right\|_2 = 0.$$

We have:

$$\begin{aligned} P_\Gamma|\psi_1\rangle_{M\Sigma EB\Gamma} &= P_\Gamma B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0\rangle_{M\Sigma EB\Gamma} \\ &= P_\Gamma B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0^1\rangle_{M\Sigma EB\Gamma} + P_\Gamma B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0^0\rangle_{M\Sigma EB\Gamma}. \end{aligned} \tag{140}$$

Recall from eq. (90) that $|\psi_0^0\rangle$ and $|\psi_0^1\rangle$ refer respectively to the blinded message ($m \in B$) term and the unblinded message ($m \in B^c$) term of the queried message. Note that using eq. (89), their respective expressions is given by:

$$|\psi_0^1\rangle_{M\Sigma EB\Gamma} = \sum_{m\in B} \sum_{\sigma\in(\{0,1\}^n)^l} \kappa_{m\sigma 1}|m\rangle_M|\sigma\rangle_\Sigma|\alpha_{m\sigma 1}\rangle_E|1\rangle_B|\nu\rangle_\Gamma \tag{141}$$

$$|\psi_0^0\rangle_{M\Sigma EB\Gamma} = \sum_{m\in B^c} \sum_{\sigma\in(\{0,1\}^n)^l} \kappa_{m\sigma 0}|m\rangle_M|\sigma\rangle_\Sigma|\alpha_{m\sigma 0}\rangle_E|0\rangle_B|\nu\rangle_\Gamma \tag{142}$$

Let us analyze each term of the right hand-side of eq. (140) separately. Given that the first term corresponds to blinded messages, there is no signature, so all the hash chain registers (except for the last) are in range of the projector Φ. Thus, the state $|\psi_0^0\rangle_{M\Sigma EB\Gamma}$ remains unchanged, that is

$$P_\Gamma B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0^1\rangle_{M\Sigma EB\Gamma} = |\psi_0^1\rangle_{M\Sigma EB\Gamma}. \tag{143}$$

For the second term, we know by the checksum that if a message $m \in B^c$ with corresponding block $b(m)$ has been queried for any other message $m' \neq m$, the corresponding block $b(m')$ contains at least one index i such that $b_i' < b_i$ with $1 \leq i \leq l$ where b_i' is the i-th block of $b(m')$.

Recall from eq. (11) that $b(m)$ is defined by $b = b(m) = (b_1, \ldots, b_l) = m \parallel C(m)$, where $C(m)$ is the checksum corresponding to the message m.

Since P_Γ acts on the whole hash chain register but we are interested only in those sub-registers that are consistent with $\widehat{B^c}$, we can split P_Γ into a sum of two orthogonal projectors, just like we did in eq. (136). Letting

$$A(m) = \{\alpha \in \widehat{B^c} : \alpha_i^j = 0 \text{ for all } (i,j) \text{ with } j < b_i(m)\} \subseteq \widehat{B^c}, \tag{144}$$

we can split P_Γ as follows:

$$\begin{aligned} P_\Gamma = \sum_{\alpha\in\widehat{B^c}} \Phi(\alpha)_\Gamma &= \sum_{\alpha\in A(m)} \Phi(\alpha)_\Gamma + \sum_{\alpha\in\widehat{B^c}\setminus A(m)} \Phi(\alpha)_\Gamma \\ &= P_\Gamma^1(m) + P_\Gamma^2(m). \end{aligned}$$

The first projector $P_\Gamma^1(m)$ simply collects all terms that correspond to those hash chain registers for which $\alpha_i^j = 0$. More precisely, $P_\Gamma^1(m)$ corresponds to the hash chain registers Γ_i^j with $j < b_i$ in range of Φ while $P_\Gamma^2(m)$ corresponds to the remaining hash chain registers. Note that the way P_Γ is split differs from one message to another.

Using this decomposition of P_Γ, the second term of the right-hand side of eq. (140) can be expressed as

$$P_\Gamma B\,\mathsf{Sign}_{\mathsf{sk}}\,|\psi_0^0\rangle_{M\Sigma EB\Gamma} = \sum_{m\in B^c} \left(P_\Gamma^1(m) + P_\Gamma^2(m)\right) B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\nu\rangle_\Gamma \tag{145}$$

where

$$|\gamma(m)\rangle_{\Sigma EB} = \sum_{\sigma\in(\{0,1\}^n)^l} \kappa_{m\sigma 0}|m\rangle_M|\sigma\rangle_\Sigma|\alpha_{m\sigma 0}\rangle_E|0\rangle_B$$

comes from the definition of $|\psi_0^0\rangle_{M\Sigma EB\Gamma}$ in eq. (142) and $|\nu\rangle_\Gamma$ is the initial hash chain block (before any query) as defined in eq. (86). For a fixed message m, $P_\Gamma^1(m)$ can also be written as

$$P_\Gamma^1(m) = \sum_{\alpha\in A(m)} \Phi(\alpha)_\Gamma = \Phi_{\Gamma_1^{<b_1(m)}\cdots\Gamma_l^{<b_l(m)}} \otimes \mathbb{1}_{\Gamma_1^{\geq b_1(m)}\cdots\Gamma_l^{\geq b_l(m)}},$$

where $\Gamma_j^{<j} = \Gamma_i^0\Gamma_i^1\cdots\Gamma_i^{j-1}$ and $\Gamma_i^{\geq j}$ is defined analogously.

Thus,

$$\begin{aligned}&P^1_\Gamma(m)B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\nu\rangle_\Gamma\\&= B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}\left(\Phi_{\Gamma_1^{<b_1(m)}\ldots\Gamma_l^{<b_l(m)}}\otimes\mathbb{1}_{\Gamma_1^{\geq b_1(m)}\ldots\Gamma_l^{\geq b_l(m)}}\right)|\nu\rangle_\Gamma\\&= B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\nu\rangle_\Gamma\end{aligned}\tag{146}$$

where the second equation follows because $P^1_\Gamma(m)$ and $B\,\mathsf{Sign}_{\mathsf{sk}}$ act on different registers and therefore commute. More precisely, $P^1_\Gamma(m)$ acts on the hash chain sub-registers of messages m' whose blocks fulfil the condition $b_i(m') < b_i(m)$ while $B\,\mathsf{Sign}_{\mathsf{sk}}$ acts on hash chain registers corresponding to the message m that was signed, i.e.

$$[P^1_\Gamma(m), B\,\mathsf{Sign}_{\mathsf{sk}}]|m\rangle_M = 0.$$

Note that this is only true for a fixed message m. The last equation follows from the fact that Γ (except for the last sub-chain) is in state $|\Phi\rangle$.

Since P^1_Γ and P^2_Γ are sums of mutually orthogonal projectors,

$$P^2_\Gamma(m)P^1_\Gamma(m) = 0.$$

Therefore, by eq. (146),

$$P^2_\Gamma(m)B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\nu\rangle_\Gamma = 0.$$

Hence, from the latter equation and eq. (143) it follows that

$$P_\Gamma B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\nu\rangle_\Gamma = B\,\mathsf{Sign}_{\mathsf{sk}}\,|\gamma(m)\rangle_{\Sigma EB}|\nu\rangle_\Gamma.$$

Hence, by eq. (145),

$$\left\||\psi_1\rangle_{M\Sigma EB\Gamma} - P_\Gamma|\psi_1\rangle_{M\Sigma EB\Gamma}\right\|_2 = 0.\tag{147}$$

We conclude that the projection of P_Γ onto the post-signature state leaves it unchanged considering that hash queries before Sign query leave the hash chain registers (except for the last) in state $|\Phi\rangle$. □

B.2 Proof of Lemma 17

The proof of this lemma is similar to the proof of Lemma 9 in Appendix A.2.

Lemma 17. *Let $m^* \in B$ and $b^* = b(m^*)$, see eq. (11). Then the projectors $Q^{b^*}_{l+1}$ and P_Γ defined in eqs. (76) and (80) are orthogonal:*

$$Q^{b^*}_{l+1}P_\Gamma = 0.\tag{84}$$

Proof. Recall from eq. (76) that

$$Q^{b^*}_{l+1} = \bigotimes_{i=1}^{l}\Phi^\perp_{\Gamma_i^{b_i^*}}\tag{148}$$

where $\Phi^\perp = \mathbb{1} - \Phi$ and $\Phi = |\Phi\rangle\langle\Phi|$. Also, recall from eq. (79) that

$$\widehat{B^c} = \bigcup_{m\in B^c}\left\{\alpha\in\{0,1\}^{l(w-1)}\;\middle|\;\alpha_i^j = 0 \text{ for all } i = 1,\ldots,l \text{ and } j < b_i(m)\right\}\tag{149}$$

where $b_i(m)$ is the i-th block of m concatenated with its checksum, see Section 2.3.2. Moreover, recall from eqs. (78) and (80) that

$$P_\Gamma = \sum_{\alpha\in\widehat{B^c}}\Phi(\alpha)_\Gamma = \sum_{\alpha\in\widehat{B^c}}\bigotimes_{i=1}^{l}\bigotimes_{j=0}^{w-2}\Phi(\alpha_i^j)_{\Gamma_i^j}.\tag{150}$$

where $\Phi(0) = \Phi$ and $\Phi(1) = \Phi^\perp$.

Combining the expressions of $Q^{b^*}_{l+1}$ and P_Γ, we get

$$Q^{b^*}_{l+1}P_\Gamma = \sum_{\alpha\in\widehat{B^c}}\left(\left((\Phi^\perp)^{\otimes l}\right)_{\Gamma_1^{b_1^*}\ldots\Gamma_l^{b_l^*}}\otimes\mathbb{1}^{\otimes l(w-2)}_{\Gamma_{(b_i^*,i)^c}}\right)\Phi(\alpha).\tag{151}$$

Recall from Section 2.3.2 that by construction of the checksum, if the block b of a message m is computed, the block b' of any other message m' contains at least an index i such that $b'_i < b_i$, $i = 1, \dots, l$. Consider a single term $\alpha \in \widehat{B^c}$ in the sum of eq. (151) and an un-blinded message $m \in B^c$ with associated block $b = b(m)$. By definition of $\widehat{B^c}$, there exists a message $m' \in B^c$ such that $\alpha_i^{b'_i} = 0$ for all i. This implies by construction of the checksum that $b'_i < b_i$ for all i. But as $B \ni m^* \neq m \in B^c$, there exists an index i such that $b^*_i < b_i$, or equivalently $b^*_i = b'_i$. Therefore we can rewrite the corresponding term as follows:

$$
\begin{aligned}
&\left(\left(\Phi^{\perp}_{\Gamma_1^{b_1^*}} \otimes \cdots \otimes \Phi^{\perp}_{\Gamma_l^{b_l^*}}\right) \otimes \mathbb{1}^{\otimes l(w-2)}_{\Gamma_{(b_i^*,i)^c}}\right) \Phi(\alpha) \\
&= \left(\Phi^{\perp}_{\Gamma_1^{b_1^*}} \otimes \cdots \Phi^{\perp}_{\Gamma_i^{b_i^*}} \otimes \cdots \otimes \Phi^{\perp}_{\Gamma_l^{b_l^*}} \otimes \mathbb{1}^{\otimes l(w-2)}_{\Gamma_{(b_i^*,i)^c}}\right) \\
&\quad \left(\Phi_{\Gamma_1^{b'_1}} \otimes \cdots \otimes \Phi_{\Gamma_i^{b'_i}} \otimes \cdots \otimes \Phi_{\Gamma_l^{b'_l}} \otimes \Phi(\alpha_1^{b_1})_{\Gamma_1^{b_1}} \otimes \cdots \otimes \Phi(\alpha_l^{b_l})_{\Gamma_1^{b_l}}\right) \\
&= \cdots \otimes \underbrace{\left(\Phi^{\perp}\Phi\right)_{\Gamma_i^{b_i^*}}}_{=0} \otimes \cdots \\
&= 0.
\end{aligned}
$$

The result follows by applying this argument to each term in eq. (151). □

B.3 Proof of Lemma 18

The proof of this lemma is similar to the proof of Lemma 11 in Appendix A.4.

Lemma 18. *Let P_Γ and U_h be respectively the invariant projector for the Winternitz OTS and the random oracle unitary defined with respect to the* Quantum independent world. *If there are hash queries after the* Sign *query, then*

$$
\left\| [U_h, P_\Gamma] \right\|_\infty \leq \delta_W(n) \tag{85}
$$

where

$$
\delta_W(n) = \frac{8l(w+1)(w-1)}{2^{n/2}}.
$$

Proof. Here, we want to prove that the commutator of the invariant projector P_Γ and the random oracle unitary U_h is small. We remind the reader that we use the convention that operators are tensored with an identity on any missing registers, which should be clear from context. We begin by deriving a decomposition for P_Γ. By definition of $\widehat{B^c}$, for some i and j, if $\alpha \in \widehat{B^c}$ such that $\alpha_i^{j'} = 0$ for $j' < j$ and $\alpha_i^j = 1$, then $\tilde{\alpha} \in \widehat{B^c}$ for all $\tilde{\alpha}$ such that $\alpha_{i'}^{j'} = \tilde{\alpha}_{i'}^{j'}$ if $i' \neq i$ or $j' \leq j$. It follows that for a fixed $i \in \{1, \dots, l\}$, we can write

$$
\begin{aligned}
P_\Gamma &= \sum_{j=0}^{w-1} \left(\Phi^{\otimes j}\right)_{\Gamma_i^{\leq j-1}} \otimes \Phi^{\perp}_{\Gamma_i^j} \otimes P^{(i,j)}_{\Gamma_{i^c}} + \left(\Phi^{\otimes(w-1)}\right)_{\Gamma_i} \otimes P^{(i,w)}_{\Gamma_{i^c}} && (152) \\
&= \sum_{j=0}^{w-1} \left(\left(\Phi^{\otimes j}\right)_{\Gamma_i^{\leq j-1}} - \left(\Phi^{\otimes(j+1)}\right)_{\Gamma_i^{\leq j}}\right) \otimes P^{(i,j)}_{\Gamma_{i^c}} + \left(\Phi^{\otimes(w-1)}\right)_{\Gamma_i} \otimes P^{(i,w)}_{\Gamma_{i^c}} && (153)
\end{aligned}
$$

with $\Phi = |\Phi\rangle\langle\Phi|$ and $\Phi^\perp = \mathbb{1} - \Phi$ with $|\Phi\rangle$ defined in eq. (1), and

$$
P^{(i,j)} = \sum_{\substack{\alpha \in A(m) \\ \alpha_i^j = 1}} \bigotimes_{\substack{i'=1,\dots,l \\ i' \neq i}} \bigotimes_{j'=0,\dots,w-2} \Phi(\alpha_{i'}^{j'}) \tag{154}
$$

where we use the convention that $\alpha_i^w = 1$ and $A(m)$ was defined in eq. (144) in the proof of Lemma 16. We can further rearrange to bring P_Γ into the form

$$
P_\Gamma = \sum_{j=0}^{w-1} \left(\Phi^{\otimes j}\right)_{\Gamma_i^{\leq j-1}} \otimes (A_i^j)_{\Gamma_{i^c}}, \tag{155}
$$

where A_i^j is a difference of two projectors[4]. We begin by bounding

$$
\left\| [U_h, P_\Gamma] \right\|_\infty \leq \sum_{\substack{i=1,\dots,l \\ j=0,\dots,w-2}} \left\| \left[(U_i^j)_{\Gamma_i^j}, P_\Gamma\right] \right\|_\infty + 2 \sum_{\substack{i=1,\dots,l \\ j=0,\dots,w-2}} \left\| \left[P^{\neq}_{X\Gamma_i^j}, P_\Gamma\right] \right\|_\infty \tag{156}
$$

[4] For some j, one of these projectors is the zero projector.

using the same steps as in eqs. (130) to (135). We can bound the first term as

$$\sum_{\substack{i=1,\dots,l\\ j=0,\dots,w-2}} \left\| \left[(U_i^j)_{\Gamma_i XY}, P_\Gamma \right] \right\|_\infty \leq 2 \sum_{j=0}^{w-1} \sum_{\substack{i=1,\dots,l\\ j'=0,\dots,w-2}} \left\| \left[(U_i^j)_{\Gamma_i XY}, (\Phi^{\otimes j'})_{\Gamma_i^{\leq j'-1}} \right] \right\|_\infty,$$

where the inequality follows from the decomposition of P_Γ above, the triangle inequality, and the fact that $\|A_i^j\|_\infty \leq 2$. By eq. (109) in the proof of Lemma 15,

$$\left\| \left[(U_i^j)_{\Gamma_i XY}, (\Phi^{\otimes j'})_{\Gamma_i^{\leq j'-1}} \right] \right\|_\infty \leq 4(w-1)2^{-n/2}. \tag{157}$$

For the second term in eq. (156), we simplify it to

$$\begin{aligned} 2 \sum_{\substack{i=1,\dots,l\\ j=0,\dots,w-2}} \left\| \left[P^{\neq}_{X\Gamma_i^j}, P_\Gamma \right] \right\|_\infty &= 2 \sum_{\substack{i=1,\dots,l\\ j=0,\dots,w-2}} \left\| \left[\mathbb{1}_{X\Gamma} - P^{=}_{X\Gamma_i^j}, P_\Gamma \right] \right\|_\infty \\ &= 2 \sum_{\substack{i=1,\dots,l\\ j=0,\dots,w-2}} \left\| \left[P^{=}_{X\Gamma_i^j}, P_\Gamma \right] \right\|_\infty . \end{aligned} \tag{158}$$

We can now alternatively decompose P_Γ similar to eq. (122), i.e.

$$P_\Gamma = \Phi_{\Gamma_i^j} \otimes \tilde{\Phi}^0_{\Gamma_{(i,j)^c}} + \Phi^\perp_{\Gamma_i^j} \otimes \tilde{\Phi}^1_{\Gamma_{(i,j)^c}} \tag{159}$$

for some projectors $\tilde{\Phi}^b$, $b = 0, 1$. Using this decomposition, we bound

$$2 \sum_{\substack{i=1,\dots,l\\ j=0,\dots,w-2}} \left\| \left[P^{=}_{X\Gamma_i^j}, P_\Gamma \right] \right\|_\infty \leq \frac{8l(w-1)}{2^{n/2}}$$

using the same calculations we performed in the proof of Lemma 11. By replacing the two terms of eq. (156) by their respective bounds, we get

$$\begin{aligned} \|[U_h, P_\Gamma]\|_\infty &= \sum_{\substack{i=1,\dots,l\\ j=0,\dots,w-2}} \left\| \left[(U_i^j)_{\Gamma_i^j}, P_\Gamma \right] \right\|_\infty + 2 \sum_{\substack{i=1,\dots,l\\ j=0,\dots,w-2}} \left\| \left[P^{\neq}_{X\Gamma_i^j}, P_\Gamma \right] \right\|_\infty \\ &\leq \frac{8lw(w-1)}{2^{n/2}} + \frac{8l(w-1)}{2^{n/2}} \\ &= \frac{8l(w+1)(w-1)}{2^{n/2}}, \end{aligned} \tag{160}$$

as desired. □