Quantum Authentication: Security against Authentication and Verification Queries

Shaoquan Jiang

School of Computer Science, University of Windsor
Email: jiangshq@uwindsor.ca

Abstract. Quantum authentication is a procedure that sends a quantum message to a receiver without being imperceptibly changed in the channel. How to formalize a proper authentication model is a highly non-trivial task. Existing models have various flaws: they either do not capture serious concerns or are over restricted. Most importantly, none of them have addressed the threat from the verification queries. We show that there is a quantum authentication scheme that is secure when no verification query is allowed while it is completely insecure when verification queries are additionally permitted. The threat of verification queries is not artificial. Our attack only needs to know if a forged authentication message is valid or not. It captures the concern that the adversary can watch if the receiver accepts an authentication or not, without even reading the message authenticated. We propose a quantum authentication model that captures the authentication of multiple messages under the same key as well as the verification queries. We allow the attacker to have his own state entangled with the authentication message. Finally, we propose an authentication framework abstracted from the AQA method in Garg et al. (CRYPTO'17) and prove the security in our model. Our result reduces the security of an authentication protocol to certain properties of its component primitives. We also prove that an impersonation attack implies a substitution attack. To our knowledge, this is the first time to confirm this result.

Key Words. Quantum Authentication, Pseudorandom Function, Compressed Quantum Random Oracle, Generalized O2H Lemma.

1 Introduction

As quantum communication technologies mature, ensuring the integrity and authenticity of transmitted quantum information has become a critical concern. In classical communication, message authentication codes (MACs) and digital signatures provide robust tools for detecting tampering and verifying sender identity. Authentication in the classical world has been extensively studied in the literature [18, 19, 6, 13, 37]. However, the quantum authentication is completely different. The unique properties of quantum information, such as the no-cloning theorem and measurement disturbance, render classical techniques inadequate for authenticating quantum messages. The authentication of quantum messages requires fundamentally different approaches that preserve the quantum state while guaranteeing that any tampering behavior can be detected.

We investigate the quantum authentication based on a shared key. The public-key setting is not interesting to us: it was shown in [8] that a public-key digital signature for a quantum state is impossible, although we can still use the standard public-key cryptography to indirectly achieve this (as said in [8]), where we can use the receiver's public-key to encrypt the authentication key and use the sender's digital signature to authenticate this classical ciphertext while the encrypted session key will be used to authenticate the quantum state. This again reduces the problem to the shared key based authentication.

Quantum message authentication schemes aim to protect quantum states transmitted over insecure channels, ensuring that the modification of the state is either harmless or detected. But before we can define any kind of security, we have to investigate what kind of adversarial power should be given. In this paper, we consider the setting where the attacker's state is entangled with the message to be authenticated. We also consider the authentication of multiple quantum messages with the same key. In addition, we also allow an adversary to learn whether his forgery of an authentication message is valid or not. This verification query was first studied by Safavi-Naini and Wild [37] in the classical setting. But the quantum setting has not been studied before.

1.1 Related Works

Barnum et al. [8] first studied the authentication of quantum message in the non-interactive setting. They formalized the security model and proposed an authentication scheme using a novel purity testing code. They also showed that the non-interactive authentication of quantum message implies the encryption. Their security definition essentially states that if the sender sends message state $|\psi\rangle$, then either the receiver rejects or it outputs a state ρ' with high fidelity with $|\psi\rangle\langle\psi|$. Oppenheim and Horodecki [35] adapted the work of Barnum et al to recycle the key. We notice that the message in [8, 35] does not have entanglement with the adversary state. Hayden et al. [29] proved the universal composability [18] of Barnum et al. [8] which essentially implies the total authentication with key leakage in [26]. But it only guarantees the security for *a single execution* of the protocol and the key can be reused only if the receiver accepts the authentication. That is, if the receiver does not receive or if he rejects it, the key can not be used any more. Portmann [36] further studied the full security of these protocols with key recycling in the universal composition model of Mauer and Renner [33]. We stress that Bynum et al. [8] and Garg et al. [26] do not allow to authenticate multiple messages. The key reuse in [35, 36] allows to further authenticate messages but only under the assumption that no authentication has ever failed.

Boneh and Zhandry [10] studied the quantum authentication using the classical authentication code by $|m\rangle \mapsto |m, \mathsf{MAC}_k(m)\rangle$. The security is formalized as after authentication queries of ℓ superposition of messages the attacker can not produce $\ell+1$ authentication codes for $\ell+1$ classical messages. Intuitively, for these ℓ queried message superpositions, attacker might be able to produce ℓ message-tag pairs by measurement. But he should not be able to produce $\ell+1$ such pairs. Their model captures the authentication of multiple messages. But it does not capture the following concerns. First, the attacker might only have queried one superposition with each message starting with 0 but finally it produces an authentication of a classical message starting with 1. Note this does not imply that the attacker can further measure the queried superposition to produce one more message-tag pair because the latter forgery might be transformed from this signed superposition. So the plus-one security is not violated. Second, the attacker might be able to alter the coefficient of the superposition signed by the oracle, in which case the measurement will change the distribution of classical forgery.

Fehr and Salvail [24] studied the authentication of classical messages with key reusability. As first observed by Bennett, Brassard and Breidbart in 1982 [11] (used by celebrated BB84 protocol [12] and also by Damgård et al. [21, 22] and Fehr and Salvail [24]), if the quantum message is not tampered by attacker, then we can reuse the key since the attacker does not learn anything about the key. Fehr and Salvail [24] used the idea in [12] to make the key reusable. But, for a failed execution, it requires an external procedure to share a new key if further authentication is needed.

Quantum message authentication in the sense of Barnum et al. [8] implies the uncloneable encryption of Gottesman [27]. The consequence of this implication is that after the message has been successfully authenticated, the attacker can not learn anything about the message.

Aharonov et al. [4] proposed two authentication schemes provably secure in the model of Barnum et al. [8]. The first scheme directly uses a random Clifford operator to mask the state $|\psi\rangle \otimes |0^\ell\rangle$ and the verification is to remove the masking and check if the last ℓ qubits are zeros. The second one (from Ben-Or et al. [7], with a complete security proof) uses the random signed polynomial code [5] to encode the message $|\psi\rangle \otimes |0^\ell\rangle$ and then mask it by a random Pauli matrix. The verification is to undo this procedure and see if the last qubits are again zeros.

Broadbent et al. [16] proposed a new authentication method called trap code. The idea is to encode the data into a n-qubit system and then introduce another 2 registers of each n-qubit registers with dummy data under different bases. Then it randomly permutes the result. The receiver verifies the authentication by undoing the process. If an error occurs on the dummy data, then the error is caught. Broadbent and Wainewright [17] studied the simulation efficiency in the security proofs of two authentication schemes: the Clifford scheme and the trap code scheme. They showed that if the adversary in the authentication protocol runs in time T, then the ideal adversary simulating it runs in time $O(T)$. This time-preserving is desired for composable security.

Garg et al. [26] formalized two models for authentication security. The first one is called basis-dependent authentication. It reduces the protocol security to an idea adversary that is basis-respecting. This is an adversary that does not tamper the register of the authenticating message. They showed that the security under this model implies the security in the Boneh-Zhandry model [10]. The second model is called total authentication. In this model, it reduces the protocol security to an ideal adversary that either forwards the authentication message faithfully or a completely garbage message to the receiver. This strong authentication also implies the encryption of the input state. They constructed a scheme (called unitary design scheme) with total authentication security. This scheme is to apply a random unitary from unitary t-design [15] to $|\psi\rangle \otimes |\mathbf{0}\rangle$ and the verification is to undo this unitary and check by measurement if the second register contains $\mathbf{0}$. Their model allows the attacker to have side information about the source message as also allowed in [23, 17]. Barnum et al. [8] does not handle this setting and will be strictly weaker than this model.

Dupuis et al. [23] studied the authentication model that is weaker than [26]. Specifically, Dupuis et al. considered the reduction in the average over the secret key. Garg et al. [26] presented a protocol that satisfies the definition of [23] but not the total authentication security in [26]. Haug et al. [28] proposed the quantum authentication from pseudo random unitary PRU [32] and t-design [15]. The PRU is used to encrypt the state and t-design is used to achieve authentication. Alagic and Majenz [1] compared the results in [23] and [26] and relaxed the condition of the authentication scheme in [26] from unitary 8-design to 2-design.

Alagic et al. [3] proposed a blind-unforgeability (BU) for MAC that avoids an attack against plus-one (PO) security of Bohen-Zhandry [10]. They gave an example that is PO-secure but not BU-secure. Their BU security implies the basis-dependent security of [26] but obviously not total-authentication security in [26] as it does not protect the confidentiality of the authenticated message.

We notice that the previous works did not propose a satisfactory model for multiple message authentication. They lie in the plus-one or key reuse model, which, as emphasized, is not sufficient for applications. Especially, the key reuse will not work when one failed or intercepted authentication occurs. Furthermore, we want to emphasize the importance of verification queries. Classically, this was studied by Safavi-Naini and Wild [37], where they obtained some information theoretical bounds. For the classical case, this attack is only useful only if the verifcation is not to re-compute the authentication code; otherwise, the recomputing operation can be done by an authentication query. Most of the classical authentication systems are deterministic and so this attack does not cause further impact beyond the authentication query for such schemes. This is probably why the classical authentication literature does not pay much attention to this issue. However, the quantum setting is different: there is no way to extract the message from the authentication message without disturbing the latter, due to the no-cloning theorem. Thus, the verification by re-computing for a quantum authentication is simply not working. Hence, it is necessary to consider it, no matter the system is determisitic or not. In this paper, we study a security model that admits both authentication and verification queries.

1.2 Contribution

In this work, we propose a new authentication model that captures the authentication of multiple quantum messages under the same secret key. It also admits the verification queries. This verification query is very immportant. We show that there exists a quantum authentication system that is secure when the verification queries are disabled but it is completely insecure when this type of queries are allowed. These capabilities are not captured in the previous model. Our model uses a classic nonce. Introducing a nonce is not strange and it has been widely used in practice (e.g., in blockchain, IPSec and TLS) to prevent a replay attack. The nonce has played a crucial role to avoid the definitional and analytical complication for quantum authentication. We also capture the concern that the attacker might know about the authentication message through entanglement. This conern was addressed by Garg et al. [26]. We prove that a quantum impersonation attack implies a quantum substitution attack. To our knowledge, this is the first formal proof for this result that was questioned in [36]. We finally propose an authentication framework by abstracting the AQA protocol in [26] and show that it is secure under our model if the underlying primitives satisfy certain properties. This reduces the security of the authentication protocol to the properties of the primitives. We give several realizations of these component primitives.

This paper is organized as follows. Section 2 introduces the basic notations and results. Section 3 gives some fundamental lemmas. Section 4 formalizes the authentication model and its security model that captures the authentication queries and verification queries. Section 5 introduces our authentication framework. Section 6 proves the security of the framework. Section 7 constructs a quantum authentication system that is secure when no verification queries are allowed while it is completely insecure when they are additionally permitted. Section 8 proves the reducibility of some functions including t-wise independent functions, hash functions as random oracle, and pseudorandom functions. The last section is a conclusion with open questions.

2 Preliminaries

Notations.

- $A|B$ stands for concatenation of A and B.
- For a finite set S, $a \leftarrow S$ is to sample a uniformly randomly from S.
- In this paper, we use ν to denote the security parameter.
- A non-negative function $\mathbf{negl}(\nu)$ is **negligible** if it vanishes faster than any polynomial fraction. That is, for any positive polynomial $poly(\nu)$, there exists $N > 0$ so that when $\nu > N$, it has $\mathbf{negl}(\nu) < 1/poly(\nu)$. In this paper, $\mathbf{negl}(\nu)$ will represent a negligible function by default.

- A is **indistinguishable** from B, if for any quantum polynomial time distinguisher $\mathcal{D}$, we have $\Pr(\mathcal{D}(A)=1)=\Pr(\mathcal{D}(B)=1)+\mathbf{negl}(\nu)$. We denote it by $A \approx B$.
- We consider vectors indexed by $\mathcal{X}$. We use $(t)_x$ to denote the vector that is t at index x and $\perp$ at every other index. For vector $\mathbf{y}$ with $y_x = \perp$, $\mathbf{y} \cup (t)_x$ is the combined vector $\mathbf{y}'$, where $y'_u = y_u$ for $u \neq x$ and $y'_x = t$. We also use $\mathbb{X}(\mathbf{y})$ to denote the set of index x with $y_x \neq \perp$.

2.1 Introduction to Quantum Computing

We give a brief introduction to quantum computing through a list of notations and some facts, with interpretations if necessary; see [34, 39, 40] for details and [40] for a conceptual taste.

- A quantum system is a finite-dimensional complex vector space (called Hilbert space) $\mathcal{H}$ with an inner product $\langle \cdot | \cdot \rangle$.
- The state of a quantum system is a unit vector $|\psi\rangle$. Its conjugate transpose is denoted by $\langle \psi|$.
- For a finite set $\mathcal{Y}$, $\{|y\rangle\}_{y\in\mathcal{Y}}$ represents an orthonormal basis for $\mathcal{H} = \mathbb{C}^{|\mathcal{Y}|}$. We denote $\mathcal{H}$ by $\mathbb{C}[\mathcal{Y}]$ to emphasize that $\mathcal{H}$ is generated by $\{|y\rangle\}_{y\in\mathcal{Y}}$ and call $|y\rangle_{y\in\mathcal{Y}}$ the **computational basis**.
- For two quantum systems $\mathcal{H}_1$ and $\mathcal{H}_2$, the joint system is a tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$.
- For $|\psi_1\rangle \in \mathcal{H}_1$ and $|\psi_2\rangle \in \mathcal{H}_2$, their product state in $\mathcal{H}_1 \otimes \mathcal{H}_2$ is $|\psi_1\rangle \otimes |\psi_2\rangle$, simplified as $|\psi_1\rangle|\psi_2\rangle$.
- A quantum register is a system holding the quantum state. It is the quantum analogue of a classical processor register. We use $|\psi\rangle_A$ to denote register A with state $|\psi\rangle$.
- For a composite quantum register $Y = MT$ with two sub registers M of domain $\mathcal{M}$ and T of domain $\mathcal{T}$, we use $|m\rangle_Y$ with $m \in \mathcal{M}$ to represent $|m\rangle_M|0\rangle_T$.
- If a quantum system $\mathcal{H}$ has an orthonormal basis $\{|\psi_1\rangle, \cdots, |\psi_n\rangle\}$, then a quantum state $|\psi\rangle \in \mathcal{H}$ can be represented as $|\psi\rangle = \sum_{i=1}^n \lambda_i |\psi_i\rangle$ with $\sum_i |\lambda_i|^2 = 1$.
- For a quantum state $|\psi\rangle$, $|| \, |\psi\rangle \, ||$ is its Euclidean norm.
- A unitary U on $\mathcal{H}$ is an operator from $\mathcal{H}$ to $\mathcal{H}$ with $UU^\dagger = I$, where $U^\dagger$ is the conjugate transpose of U.
- Measurement $M = \{M_i\}_i$ on a quantum state $|\psi\rangle \in \mathcal{H}$ is the operator for extracting the classical information from $|\psi\rangle$, where $\{M_i\}_i$ is required to satisfy the completeness condition $\sum_i M_i^\dagger M_i = I$. When M is applied, it will result in a post-measurement state $M_i|\psi\rangle/||M_i|\psi\rangle||$ with probability $||M_i|\psi\rangle||^2$.
- A quantum algorithm A is represented by a sequence of unitaries/measurements. Due to deferred measurement principle [34, pp. 186], the measurement can be deferred to the end of operations of A. Hence, whenever applicable, we always assume that A before the final measurement is represented by a list of unitaries $U_1, \cdots, U_\ell$.
- If $|1\rangle, \cdots, |n\rangle$ is an orthonormal basis of $\mathcal{H}$, then $P = \sum_{k\in A} |k\rangle\langle k|$ for $A \subset [n]$ is a projector from $\mathcal{H}$ onto the subspace expaned by $\{|k\rangle\}_{k\in A}$.
- The norm of linear operator A on $\mathcal{H}$ is defined as $||A|| = \max_v ||A|v\rangle||$, where $|v\rangle$ goes over all the possible unit vectors in $\mathcal{H}$. By the singular value decomposition theorem, we can write $A = \sum_i \lambda_i |v_i\rangle\langle y_i|$, where $\{|v_i\rangle\}_i$ and $\{|y_i\rangle\}_i$ are respectively a set of orthonormal vectors in $\mathcal{H}$ and $\{\lambda_i\}_i$ is the set of positive singular values of A. Hence, $||A|| = \max_i \lambda_i$.
- For orthonormal states $|\psi_i\rangle$ and $0 < \lambda_i \leq 1, i = 1, \cdots, n$ with $\sum_{i=1}^n \lambda_i = 1$, $\rho = \sum_{i=1}^n \lambda_i |\psi_i\rangle\langle\psi_i|$ is called a **density matrix**. If $n > 1$, ρ is a **mixed state**; if $n = 1$, ρ is a **pure state**. We can interpret ρ as the result of sampling $|\psi_i\rangle$ with probability λ_i.
- If γ is a pure state, we will use $|\gamma\rangle$ to denote its vector format without a mention.
- For Hilbert space $\mathcal{H}$, $D(\mathcal{H})$ denotes the collection of density matrix of states in $\mathcal{H}$.

- The trace distance between two mixed states ρ, σ is defined as $D_t(\rho, \sigma) = \frac{1}{2}\mathrm{tr}(|\rho - \sigma|)$, where $|A| := \sqrt{A^\dagger A}$. If $\rho = \sum_{i=1}^n p_i |\psi_i\rangle\langle\psi_i|$ and $\sigma = \sum_{i=1}^n q_i |\psi_i\rangle\langle\psi_i|$ for orthonormal basis $\{|\psi_i\rangle\}_i$, then $D_t(\rho, \sigma) = \frac{1}{2}\sum_{i=1}^n |p_i - q_i|$, which coincides with the statistical distance of distributions $P = (p_1, \cdots, q_n)$ and $Q = (q_1, \cdots, q_n)$. The trace distance between ρ and σ is also conveniently denoted by $|\rho - \sigma|_1$, where $|\cdot|_1$ is the trace norm [34].
- for an operator A on a register Y, we use A_Y to explicitly indicate this. But when the context is clear, we always remove the subscript Y for brevity.
- If an operator A only operates on register Y in a composite system ZY, we always use A to represent $I_Z \otimes A_Y$ for the composite system.
- Register D is a **control register** in the orthonormal basis $\{|y\rangle\}_y$ for operator B of register WD, if B can be written as $B = \sum_y B_y \otimes |y\rangle\langle y|_D$ where B_y operates on register W. We have two properties for a control register.
  - If A operates on registers XD while B operates on registers YD with D being a control register in the same basis $\{|y\rangle\}_{y\in D}$ for both A and B, then $AB = BA$.
  - If A is a projector on D in basis $\{|y\rangle\}_y$ and B operates on YD with D being a control register in the same basis, then $AB = BA$.
- For $y \in \{0,1\}^n$, ϕ_y is reserved in this paper to denote $\mathsf{QFT}|y\rangle$, where QFT is the quantum Frourier transform and $\{\phi_y\}_y$ is called the **Fourier basis**.
- For a keyed function $f_k : \mathcal{M} \to \{0,1\}^\lambda$ with $k \leftarrow \mathcal{K}$, we say f_k is a **pseudorandom function** if for any quantum polynomial time oracle algorithm $\mathcal{D}$, it holds that

$$\Pr(\mathcal{D}^{f_k}() = 1) = \Pr(\mathcal{D}^{P}() = 1) + \mathbf{negl}(\nu), \tag{1}$$

  where P is a uniformly random function from $\mathcal{M}$ to $\{0,1\}^\lambda$. We also say f_k is q**-pseudorandom** if the equation holds when $\mathcal{D}$ only makes at most q function queries.

2.2 Quantum Random Oracles

In this section, we will introduce the quantum random oracles. We first introduce standard random oracle. That is the classical random oracle extended to the quantum setting. Then, we introduce Zhandry's compressed random oracle [42] ($CStO$). The advantage of this oracle is that it allows a simulator to detect if an input x has been queried to the oracle or not.

Random Oracle and Standard Random Oracle In the random oracle model, a cryptographic hash function $H : \mathcal{X} \to \{0,1\}^n$ is treated as an external oracle so that whenever one needs to compute $H(x)$, he queries x to this oracle and receives $H(x)$. We assume $\mathcal{X}$ has a finite bit-length. The oracle uses a random function from $\mathcal{X}$ to $\mathcal{Y}$ to answer the queries. Let $\mathcal{X} = \{x_1, \cdots, x_N\}$ be an ordered set with $x_1 < x_2 < \cdots < x_N$. Function H can be represented by its truth table $H(x_1), H(x_2), \cdots, H(x_N)$. In the *quantum random oracle* model, H is represented by state $|H\rangle$ (using its truth table). An algorithm $\mathcal{A}$ can query a superposition to random oracle RO. For query $|x\rangle|y\rangle$, RO maps $|x\rangle|y\rangle|H\rangle$ to $|x\rangle|y \oplus H(x)\rangle|H\rangle$.

The *standard random oracle StO* has an initial state in a uniform superposition $\frac{1}{\sqrt{2^{n|\mathcal{X}|}}}\sum_H |H\rangle$. For query $|x\rangle|y\rangle$, StO maps $\frac{1}{\sqrt{2^{n|\mathcal{X}|}}}\sum_H |x\rangle|y\rangle|H\rangle$ to $\frac{1}{\sqrt{2^{n|\mathcal{X}|}}}\sum_H |x\rangle|y \oplus H(x)\rangle|H\rangle$. Notice that RO can be obtained from StO by starting with a projective measurement on oracle register (resulting in $|H\rangle$). Even though RO and StO are different, no adversary can distinguish them [42].

Fact 1 *Let $\mathcal{A}$ be a quantum algorithm with oracle access to the quantum random oracle. Then,* $\Pr(\mathcal{A}^{RO}() = 1) = \Pr(\mathcal{A}^{StO}() = 1)$.

Compressed Random Oracle The *compressed* random oracle $CStO$ was introduced in [42]. We follow the description in [20]. It is a useful tool for security analysis in the quantum random oracle model. Let $\mathcal{Y} = \{0,1\}^n$ and $\bar{\mathcal{Y}} = \mathcal{Y} \cup \{\perp\}$. Recall that $\phi_y = \mathsf{QFT}|y\rangle$ for $y \in \{0,1\}^n$. Since $\{|y\rangle\}_{y\in\{0,1\}^n}$ is orthonormal and $\mathsf{QFT}^2 = I$, we know that $\{|\phi_y\rangle\}_{y\in\{0,1\}^n}$ is also orthonormal. Then, we define an unitary operator F over $\mathbb{C}[\bar{\mathcal{Y}}]$ such that

$$F|\perp\rangle = |\phi_0\rangle, \quad F|\phi_0\rangle = |\perp\rangle, \quad F|\phi_y\rangle = |\phi_y\rangle, \quad \forall y \in \mathcal{Y} - \{\mathbf{0}\}. \tag{2}$$

It is Hermitian (i.e., $F^\dagger = F$) because $F = |\phi_0\rangle\langle\perp| + |\perp\rangle\langle\phi_0| + \sum_{y\neq 0} |\phi_y\rangle\langle\phi_y|$. Further, notice that $|y\rangle = 2^{-n/2}\sum_{\eta\in\{0,1\}^n}(-1)^{y\cdot\eta}|\phi_\eta\rangle$. This implies that $F|y\rangle = |y\rangle + 2^{-n/2}(|\perp\rangle - |\phi_0\rangle)$.

We consider the multi-register $D = \{D_x\}_{x\in\mathcal{X}}$ for the random oracle, where D_x has a state space $\mathbb{C}[\bar{\mathcal{Y}}]$, spanned by the computational basis $\{|y\rangle\}_{y\in\mathcal{Y}} \cup \{|\perp\rangle\}$. The initial state of D is $\otimes_x|\perp\rangle_{D_x}$. We assume that the adversary has a query register X, response register Y and a work register W. To query the oracle, adversary provides XY registers to oracle who then applies unitary

$$CStO_{XYD} = \sum_{x\in\mathcal{X}} |x\rangle\langle x|_X \otimes CStO_{YD_x} \tag{3}$$

on XYD, where $CStO_{YD_x} = F_{D_x} \cdot \text{CNOT}_{YD_x} \cdot F_{D_x}$ and $\text{CNOT}|y\rangle_Y|u\rangle_{D_x} = |y+u\rangle_Y|u\rangle_{D_x}$. This oracle has property that if $|x\rangle$ has never been queried, then D_x will remain as $|\perp\rangle_{D_x}$. The $\mathcal{X}$-indexed vector $\mathbf{y}$ means $\mathbf{y} = \{y_x\}_{x\in\mathcal{X}}$, where $y_x \in \overline{\mathcal{Y}}$. We remind that $\mathbb{X}(\mathbf{y})$ denotes the set of index in $\mathbf{y}$ with $y_x \neq\perp$ and that $(t)_x$ denotes the vector that is t at index x and $\perp$ for the remaining indexes. We also remind that $F_{D_x}|\mathbf{y}\rangle_D$ represents the operator $(\otimes_{x'\neq x} I_{D_{x'}} \otimes F_{D_x})|\mathbf{y}\rangle_D$. Also, as shown in the following lemma by Zhandry [42], an (unbounded) attacker can not distanguish StO and $CStO$.

Lemma 1. *[42] Let $\mathcal{A}$ be a quantum algorithm with oracle access to the quantum random oracle. Then,* $\Pr(\mathcal{A}^{StO}() = 1) = \Pr(\mathcal{A}^{CStO}() = 1)$.

Zhandry [42] showed how to make the compressed oracle represented efficiently. He applied the standard classical sparse encoding to quantum states to make the representation and maintance efficient. For simplicity, we will only express the oracle state in terms of the inefficient variant of the compressed oracle (i.e., as a possibly entangled superposition of $|\mathbf{y}\rangle$ for $\mathbf{y} \in \overline{\mathcal{Y}}^{\mathcal{X}}$ with each non-queried regiester D_x being $|\perp\rangle_{D_x}$). But we emphasize that using his efficient encoding can make all the unitaries and measurements in this paper efficiently computed.

3 Useful Lemmas

3.1 Basics

The following lemma states that if a quantum oracle algorithm only makes q queries to a $2q$-wise independent function, then the algorithm final output has no difference from that with instead making queries to a quantum random oracle. Consequently, we can always replace a $2q$-wise independent function by a random oracle in the security analysis without any penalty.

Lemma 2. *[43] Let Q be an oracle taken from uniformly random $2d$-wise independent function. Then any algorithm $\mathcal{A}$ making at most d queries in distinguishing Q from a uniformly random function of the same domain and range will have a zero-advantage.*

We also need the relation [34, Eq. (9.100)] between trace distance and the Euclidean distance.

Lemma 3. *Let $|u\rangle, |v\rangle$ be two states for a quantum system. Then, $D_t(|u\rangle\langle u|, |v\rangle\langle v|) \leq || \, |u\rangle - |v\rangle \, ||$.*

3.2 Reduction from $(H(\cdot), H(k, \cdot))$-oracle Game to $H(k, \cdot)$-oracle Game

In this section, we will consider an adversary that has oracle access to random oracle H and tag oracle $H(k, \cdot)$ (where k is a secret). Our result is that $H(\cdot)$ is redundant and thus can be removed without affecting the attacker's output distribution. The main tool to prove this is a generalization of the O2H lemma in [38] to the vector input setting. The proof is similar to the original one and we put it in Appendix A.

Lemma 4. *Assume that $(S_1, \cdots, S_K)$ is an equipartition of $\{0,1\}^n$ with $|S_i| = 2^n/K$ for all i. Let $H : \{0,1\}^n \to \{0,1\}^m$ be a quantum random oracle and $A^H(\mathbf{x}, \mathbf{y})$ be a quantum algorithm with input ($\mathbf{x}$, $\mathbf{y}$), where $\mathbf{x} \subseteq \{0,1\}^n$ of size $2^n/K$ and $y_i \in \{0,1\}^m$ and it makes at most q queries to H. Let B be a quantum algorithm with input $\mathbf{x} = S_t$ for some t: take $i \leftarrow \{1, \cdots, q\}$, $\mathbf{y} \leftarrow (\{0,1\}^m)^{2^n/K}$ and run $A^H(\mathbf{x}, \mathbf{y})$ till the i-th query to H in which case B measures the query in the computational basis and outputs the measurement result. If A makes less than i queries to H, B outputs $\perp$. Then, for $\mathbf{x} = S_k$ with $k \leftarrow [K]$, $\mathbf{y} \leftarrow (\{0,1\}^m)^{2^n/K}$ and a purely random H,*

$$|\Pr(A^H(\mathbf{x}, H(\mathbf{x})) = 1) - \Pr(A^H(\mathbf{x}, \mathbf{y}) = 1)| \leq 2q\sqrt{\delta}. \tag{4}$$

where $\delta = \Pr(x' \in \mathbf{x} : x' = B^H(\mathbf{x}))$.

We keep $H : \{0,1\}^n \to \{0,1\}^m$ as a random oracle. Further, define key domain $\mathcal{K} = \{0,1\}^\nu$ and message domain $\mathcal{M} = \{0,1\}^{n-\nu}$. We consider a random oracle game between an adversary $\mathcal{A}$ and a challenger as follows. $\mathcal{A}$ holds registers $XYMTZ$ with query register X, response register Y, tag query register M, tag register T and the working register Z. $\mathcal{A}$ has access to the random oracle $O_H : |x\rangle_X|u\rangle_Y \mapsto |x\rangle_X|u + H(x)\rangle_Y$ and a tag oracle $O_{H(k,\cdot)} : |m\rangle_M|u\rangle_T \mapsto |m\rangle_M|u + H(k, m)\rangle_T$. Both O_H and $O_{H(k,\cdot)}$ are maintained by the challenger who holds the secret key $k \leftarrow \mathcal{K}$. Initially, the joint state is $|0\rangle_{XYMTZ}$. Then, $\mathcal{A}$ consists of a sequence of oracle queries and unitaries, ending with a final measurement. Between subsequent oracle queries, $\mathcal{A}$ will apply a unitary on registers $XYMTZ$. The oracle query is either a random oracle query or a tag query. For the former, it sends registers XY to challenger; for the latter, it sends registers MT to the challenger. The challenger then responds accordingly. The final measurement is on $XYMTZ$ with outcome either 0 or 1, which is denoted by $\mathcal{A}^{H,H(k,\cdot)}$. We also consider adversary $\mathcal{A}$ that only has access to tag oracle $O_{H(k,\cdot)}$ and denote its output by $\mathcal{A}^{H(k,\cdot)}$.

The game with two oracles is a little annoying, as $\mathcal{A}$ could query $|k, u\rangle_X|0\rangle_Y$ to O_H and we must keep the two oracles consistent. Fortunately, we have the following result that essentially states that O_H does not contribute much to the output of $\mathcal{A}$ and so we can remove it safely without affecting the output distribution.

Proposition 1. *Let $\mathcal{A}$ be an adversary in the above random oracle game. Assume it makes at most q_H random oracle queries and q_T tag queries. Then, there exists an adversary $\mathcal{B}$ that only makes q_T tag queries so that*

$$|\Pr(\mathcal{A}^{H,H(k,\cdot)} = 1) - \Pr(\mathcal{B}^{H(k,\cdot)} = 1)| \leq 2q_H/\sqrt{K}, \tag{5}$$

where the probability is over k, H and the final measurement uncertainty for each fixed k, H.

Proof. We first prove that

$$|\Pr(\mathcal{A}^{H,H(k,\cdot)} = 1) - \Pr(\mathcal{A}^{H,H'(k,\cdot)} = 1)| \leq 2q/\sqrt{K}, \tag{6}$$

where both H and H' are random oracles from $\{0,1\}^n$ to $\{0,1\}^m$ but they are completely independent. Let $S_k = \{k|u\}_{u\in\{0,1\}^{n-\nu}}$ be a set alphabetically sorted according to u, where $\nu = \log K$ and $k \in \{0,1\}^\nu$. Then, $\mathcal{A}^{H,H(k,\cdot)}$ can be simulated by an adversary $\mathcal{A}'$ with oracle access to O_H and input $H(S_k)$ and k; $\mathcal{A}^{H,H'(k,\cdot)}$ can be simulated by adversary $\mathcal{A}'$ with oracle access to O_H and input $\mathbf{y}$ for $\mathbf{y} = H'(S_k)$ and k. The strategy of $\mathcal{A}'$ is as follows. Upon input $\mathbf{y}'$ (either $H(S_k)$ or $H'(S_k)$) and k, $\mathcal{A}'$ forwards the random oracle query to his own oracle O_H and asnwers the tag query u with y'_u. Finally, $\mathcal{A}'$ outputs whatever $\mathcal{A}$ does. We note that k is a representation of S_k and H' is independent of H. Hence, Lemma 4 can be applied to bound the probability gap between the outputs of $\mathcal{A}'$ for the two cases, which is at most $2q\sqrt{\delta}$. It remains to show that δ in our setting is $1/K$. By the description of $\mathcal{A}'$, δ is actually the probability that x' has a prefix k, where x' is generated in the process: $i \leftarrow [q]$ and $\mathbf{y}$ uniformly random and run $\mathcal{A}$ with access to O_H and $O_{H'(k,\cdot)}$ (where $H'(S_k) := \mathbf{y}$) and then measure the ith query of $\mathcal{A}$ with outcome x'. Here since $\mathbf{y}$ is uniformly random, H' is properly distributed (as independent of H). It clear that, prior to the measurement on the ith query, the view of $\mathcal{A}$ is independent of k (as $\mathbf{y}$ is independent of k). Since k is uniformly random, it equals the prefix of x' with probability exactly $1/K$. Therefore, $\delta = \sum_{i=1}^{q} \frac{1}{q} \cdot \frac{1}{K} = 1/K$.

To complete the theorem, it suffices to show that there exists $\mathcal{B}$ so that $\Pr(\mathcal{A}^{H,H'(k,\cdot)} = 1) = \Pr(\mathcal{B}^{H(k,\cdot)} = 1)$. Since H and H' are i.i.d, it is equivalent to show that $\Pr(\mathcal{A}^{H',H(k,\cdot)} = 1) = \Pr(\mathcal{B}^{H(k,\cdot)} = 1)$. Since H and H' are independent, $\mathcal{B}$ can run $\mathcal{A}$ and maintain $O_{H'}$ by himself with a purely random H' and forwards the tag queries to his own tag oracle $O_{H(k,\cdot)}$, deisred! ■

3.3 Security Separation of Authentication with and without Verification Queries: the Classical Setting

In this section, we show the impact of verification queries on the security of a classic message authentication code. The verification query captures a natural attack: an attacker creates an authentication message (from an existing authentication or from scratch) and sends to the verifier and sees if the receiver accepts. It was first investigated by Safavi-Naini and Wild [37], where they found an information theoretical lower bound on the adversarial success probability. The verification query is a very weak but practical threat. For instance, the result can be indicated by the receiver's subsequent behavior. In this section, we construct a MAC scheme that is secure without verification queries but it is completely insecure when these queries are permitted. Although we only consider the classical case here, the idea will later be generalized to the quantum setting.

The authentication system (**Auth**, **Ver**) with key space $\mathcal{K} \subset \{0,1\}^n$, using a pseudorandom function F with key space $\mathcal{K}$ is desribed as follows. Let $k \leftarrow \mathcal{K}$.

$\mathbf{Auth}_k(m)$. Upon m, compute $tag = F_k(m), z = 0, i = 0, u = 0$ and output $(m, z|i|u, tag)$.
$\mathbf{Ver}_k(m, z|i|u, tag)$. Upon $(m, z|i|u, tag)$, first verify if $tag = F_k(u)$. If not, reject; otherwise, it does the following. If $z = 0$ or $(z = 1 \& k_i = u)$, it accepts and outputs m; otherwise, it rejects.

Proposition 2. *The above scheme is existentially unforgeable but it is totally insecure if attacker is additionally allowed to make n verification queries.*

Proof. If no verification queries are allowed, then no verification is executed. So attacker can only create a forgery of a new message m under the help of some tags $F_k(m_i)$ from tag queries $m_i, i = 1, \cdots, q$ (existential unforgeability is well-known; see [31] for formal definition). But the pseudorandomness of F guarantees that this can succeed only negligibly. Now we show that is not secure when n verification queries are allowed. Indeed, an attacker can make a tag query for message m and receive $(m, z|i|u, tag)$ with $z|i|u = 0|0|0$. Then, it modifies to $(m, 1|i|0, tag)$ for $i = 1, \cdots, n$ and sends it for verifications. If it is rejected, then $k_i = 1$; otherwise, $k_i = 0$. After n queries, k is recovered and attacker can forge a tag for any message. ■

4 Authentication Model

In this section, we introduce our authentication model. We allow the adversary to have some side information about the message, as advocated in [23, 26]. Our model also captures the authentication of multiple quantum messages. This explicitly indicates the key reuse capability. Key reuse in the literature (e.g., [36, 29]) demands the key after authentication to preserve the most entropy. But this is possible only if we assume that the authentication message is always delivered faithfully. In other words, if the attacker retains one authentication message or delivers an invalid one, then the key can not be reused. In our model, even if an attacker intercepts or tampers multiple authentication messages, the key still can be reused. We also allow the attacker to request for verification of his authentication forgeries. This captures the concern that the attacker could learn the key by only observing if tampered authentications are accepted or not.

4.1 Syntax

Spaces. Let $\mathcal{K}$ denote the *key* domain, $\mathcal{M}$ denote the *message* domain, $\mathcal{F} = \{acc, rej\}$ denote the set of decision: acceptance (acc) and rejection (rej), $\mathcal{Y}$ denote the *authentication message* domain and $\mathcal{R}$ denote the domain of *nonce*. For a set $\mathcal{G}$, $\mathcal{H}_\mathcal{G}$ denotes Hilbert space $\mathcal{H}[\mathcal{G}]$. We also use $\mathcal{Z}$ to denote the domain for attacker's state. Hence, $\mathcal{H}_\mathcal{M}, \mathcal{H}_\mathcal{Y}$ and $\mathcal{H}_\mathcal{Z}$, $\mathcal{H}_\mathcal{F}$ respectively denotes the message space, authentication message space, attacker's state space and decision space. We remind the difference between space and domain in this paper: space stands for Hilbert space and domain stands for a set.

Nonce. In the classic authentication, nonce is an element to avoid the replay attack. In Euthereum, the nonce of an account is the number of transactions sent from this account. The signature generation of the transaction will take this nonce as part of its input. After that, it increases by one. In TLS, the MAC field is generated with the TLS sequence number as part of input to guarantee the uniqueness of the input in the current session. In both cases, the message is concatenated with a nonce (or sequence number) before being authenticated. Since the nonce (or sequence number) is known to the receiver, the replay attack is avoided. In our model, we assume that a unique *classical* nonce is known to both sender and receiver before authenticating a message. We do not restrict how the nonce will be initialized and how it will evolve. But we require that *within the lifetime of the shared secret key between sender and receiver, the nonce at the sender is unique and the nonce at the receiver is also unique.* Especially, if a receiver receives an authentication message with a previously checked nonce, it will simply reject without verification; if a sender receives an attacker's request to generate an authentication message for a previously used nonce, he will simply ignore it.

With this restriction in mind, sender and receiver in a particular session can do the following to evolve the nonce (but our exposition in this paper does not depend on it).

1. Initially, sender chooses a random number $r_s \in \{0,1\}^{\nu/3}$ and receiver chooses a random number $r_v \in \{0,1\}^{\nu/3}$, sending his number to each other. The nonce is initially defined as $r_s|r_v|0^{\nu/3}$ and later for each message it increases by 1. But keep in mind: if attacker intervenes, sender and receiver could see a different initial nonce. That is, the exchange of r_s and r_v is not authenticated.
2. Our uniqueness restriction at sender and receiver seems to require the sequential communication during the procedure to authenticate multiple messages. But it is not necessary. We can adopt the TCP-like strategy. The receiver maintains a window size w and a current nonce$_c$ (which is the lower bound of currently acceptable nonces). Whenever receiving an authentication message with nonce r, it checks if $nonce_c \leq r \leq nonce_c + w$ holds and r has not been processed so far. If it does not hold, it rejects; otherwise, it verifies the authentication message and marks nonce r has been used. If $r = nonce_c$, update $nonce_c = nonce_c + 1$. If the receiver has a large buffer, we can also set $w = \infty$.
3. If a quantum message is lost (e.g., by observing that authentication message with nonce$_c$ remains unceived longer than a timout), it advances $nonce_c = nonce_c + 1$. Note that we can not expect the sender to retransmit a lost message because the sender generally does not have a copy for it, due to the no-cloning theorem.
4. When sender and receiver are out of this moderate synchronization, they re-initialize the nonce as in item 1.

Remark. It is important that the procedure of nonce evolving is not authenticated. Especially, sender can have an initial nonce *nonce* and the receiver can have an initial nonce *nonce*$'$. But when an attacker asks the sender to generate an authentication message, the sender will use *nonce*; when he sends this (maybe modified) authentication message to the receiver, the receiver will verify it using *nonce*$'$. So unless the authentication system is insecure and can be successfully verified with different nonces, this nonce inconsistency attack does not seem to be useful. But certainly, this behavior is admitted in our security model.

Authentication. An authentication system is a pair of algorithms $(\text{Auth}_k(r,\cdot), \text{Ver}_k(r,\cdot))$ with secret $k \in \mathcal{K}$ and nonce r, where Auth_k is the authentication operator from $\mathcal{H}_\mathcal{M} \times \mathcal{R}$ to $\mathcal{H}_\mathcal{Y}$ and Ver_k is the verification operator from $\mathcal{H}_\mathcal{Y} \times \mathcal{R}$ to $\mathcal{H}_\mathcal{M} \otimes \mathcal{H}_\mathcal{F}$. When the context is clear, we also omit the nonce r. The authentication scheme must be correct: for any $k \in \mathcal{K}, r \in \mathcal{R}$ and any $\rho \in D(\mathcal{H}_\mathcal{M})$, $\text{Ver}_k \circ \text{Auth}_k(r,\rho) = \rho \otimes |acc\rangle\langle acc|$, where Auth_k and Ver_k use the same nonce r. We assume that $\mathcal{H}_\mathcal{M}$ resides in register M and $\mathcal{H}_\mathcal{Y}$ resides in register Y. For simplicity, assume $\mathcal{M} = \{0,1\}^{\ell_1}$ for some ℓ_1 and $\mathcal{Y} = \{0,1\}^{\ell_2}$ with $\ell_2 > \ell_1$. Thus, $\mathcal{H}_\mathcal{M} \otimes |0^{\ell_2-\ell_1}\rangle \subseteq \mathcal{H}_\mathcal{Y}$. By our notational convention, $|\psi\rangle_M \in \mathcal{H}_\mathcal{M}$ can be conveniently written as $|\psi\rangle_Y$ without a confusion.

For given nonce r, let $\mathcal{V}_k$ be the space of valid authentication message under the secret key k and $\mathit{\Pi}_{\mathcal{V}_k}$ be the projector to $\mathcal{V}_k$. Then, the natural way to define Ver_k is as follows:

$$\rho \mapsto (\text{Auth}_k^{-1} \circ \mathit{\Pi}_{\mathcal{V}_k})(\rho) \otimes |acc\rangle\langle acc|_F + \text{tr}((I - \mathit{\Pi}_{\mathcal{V}_k})\rho)|0\rangle\langle 0|_Y \otimes |rej\rangle\langle rej|_F. \tag{7}$$

When the input ρ is an entanglement between register Y and attacker's register Z, this equation should be adjusted as

$$\rho \mapsto (\mathrm{Auth}_k^{-1} \circ \mathit{\Pi}_{\mathcal{V}_k})(\rho) \otimes |acc\rangle\langle acc|_F + \mathrm{tr}_Y((I - \mathit{\Pi}_{\mathcal{V}_k})\rho) \otimes |0\rangle\langle 0|_Y \otimes |rej\rangle\langle rej|_F. \tag{8}$$

where $(I - \mathit{\Pi}_{\mathcal{V}_k})$ represents $(I - \mathit{\Pi}_{\mathcal{V}_k})_Y$ by our notational convention in Section 2.1 and F is the classical register for verification decision. In this paper, we define $\mathrm{Ver}_k(\rho)$ as the mapping in Eq. (8). We also use $\mathrm{Ver}_k^-(\rho)$ that is the mapping that only contains the accepted output:

$$\rho \mapsto (\mathrm{Auth}_k^{-1} \circ \mathit{\Pi}_{\mathcal{V}_k})(\rho). \tag{9}$$

4.2 Security

In this section, we define the security for quantum authentication. We will capture the following adversarial behaviors. The adversary might try to learn the secret key through authentication queries and verification queries. After the learning stage, he can issue a challenge authentication query and receive the authentication message. The attacker then tries to tamper it to a valid authentication of another message. Toward this, the attacker could have side information about the challenge message, through entanglement. The protocol is secure if the attacker can not do better than one who does not even "read" the authentication register (once the challenge authentication has been generated), not to say tampering it. We remind that the attacker can still impact the authentication message through operating on its own register that is entangled with the authenticating message. But this will not cause an *authentication issue* and it is unavoidable (no matter what authentication techniques are used to protect). We also allow the attacker to forge an authentication message without issuing the challenging authentication query and succeed if this forgery is valid. In this case, we require the nonce for the forgery is not used before.

Generally, a nonce can not be reused in authentication queries and can not be reused in verification queries, either. But it can appear in one authentication query and one verification query. This requirement is very generic and does not assume how nonce is initialized and how it evolves. In our model, the authentication oracle will be queried with a message and a nonce while the verification oracle will be queried with an authentication message and a nonce. Authentication query captures attacker's behavior of getting a message authenticated by the sender; verification query captures the attacker's behavior of seeing if an authentication message is accepted by the receiver or not.

Formally, the adversary model is formulated as a game between an adversary O and a challenger. Challenger initially takes $k \leftarrow \mathcal{K}$ and samples system parameters (such as random oracle or public-keys if any). Let Z be the register for O and Y be the register for the communication between O and the challenger. Initially, $\omega_{\ell_0} = |0\rangle\langle 0|_{ZY} \otimes |acc\rangle\langle acc|_F$ with $\ell_0 = 0$. The model will iterate the procedure of stage $i+1$ for $i = 0, 1, \cdots$ below, where in each stage only the last query is a verification query and the other queries are authentication queries. In the model, we assume that adversary O will reset the F register to $|acc\rangle_F$ state after the previous stage's verification (e.g., by swapping it with another unused qubit $|0\rangle$ in its working register) and keep this state until the challenger changes it upon the next verification query. Also, ω_{ℓ_i} includes register F while the remaining states in the model such as ω'_{ℓ_i+1} do not include this register (to emphasize that it remains in state $|acc\rangle_F$).

- Adversary O applies a unitary to ω_{ℓ_i}, resulting in $\omega'_{\ell_i+1} \otimes |acc\rangle\langle acc|_F$. Then, he sends Y register with nonce r_{ℓ_i+1} to challenger for an *authentication* query.
- Challenger then computes $\omega_{\ell_i+1} = \mathrm{Auth}_k(\omega'_{\ell_i+1})$ (with nonce r_{ℓ_i+1}) and returns the Y register to O.

$$\vdots$$

- O applies some unitary to $\omega_{\ell_{i+1}-2}$, resulting in $\omega'_{\ell_{i+1}-1}$, and sends Y register with nonce $r_{\ell_{i+1}-1}$ for the *authentication* query.
- Challenger then computes $\omega_{\ell_{i+1}-1} = \mathrm{Auth}_k(\omega'_{\ell_{i+1}-1})$ (with nonce $r_{\ell_{i+1}-1}$) and returns the Y register to O.
- O applies some unitary to $\omega_{\ell_{i+1}-1}$, resulting in $\omega'_{\ell_{i+1}}$, and sends YF registers with nonce $r_{\ell_{i+1}}$ for *verification.*
- Challenger then computes $\omega_{\ell_{i+1}} = \mathrm{Ver}_k(\omega'_{\ell_{i+1}})$ (with nonce $r_{\ell_{i+1}}$) and returns the YF registers to O.

Fig. 1. Stage $i + 1$

The complete adversary model that iterates Stage $i + 1$ in Fig. 1 is described as follows.

- $\omega_{\ell_0} = |0\rangle\langle 0|_{ZY} \otimes |acc\rangle\langle acc|_F$ and $\ell_0 = 0$.
- Stage 1
- Stage 2

$$\vdots$$

- Stage q_v (with exception)
- Let the total number of authentication queries be q_s and total number of authentication/verification queries be q. Then, $q_s + q_v = q$ and $\ell_{q_v} = q$. So the final state is ω_q. Assume for simplicity that if nonce r_q appeared in an authentication query, then that query is the challenge authentication query (so $r_{q-1} = r_q$).
  / When attacker plans to take the challenge using nonce r_q, he could issue a challenge authentication query to help or simply generate the forgery without this query. In the former case, $r_q = r_{q-1}$ (otherwise, it contradicts the fact that r_{q-1} is the nonce for the challenge query); in the latter case, r_q has never occurred in a query. */*
- **Exception in Stage q_v:** Updating ω_{q-1} to ω'_q will consist of a unitary and also a projector, instead of a unitary only in a regular stage. Besides, the verification query at this stage will output $\mathrm{Ver}_k^-(\omega'_q)$, instead of $\mathrm{Ver}_k(\omega'_q)$ in a regular stage.

Fig. 2. Adversary Model

When $q_s = 0$ and $q_v = 1$, the attacker does not make any authentication query (including a challenge authentication query) but does make one verification query. This is an *impersonation attack*. In this case, $\omega'_1 \otimes |acc\rangle\langle acc|$ is sent for verification.

The projector by O applied to ω_{q-1} is a natural treatment to capture the accumulated effect of some (possibly delayed) measurements by the attacker. In the final challenge verification, we use Ver_k^- to verify because we only care about the accepted state in this query and the game no longer continues after this; in the learning stage, we use Ver_k to verify because the game needs to continue even if the verification result is a reject. We remark that if the rejected part in the challenge verification is indeed important, then the attacker should choose to stay longer in the learning stage and then come back to the challenge verification later, where only a success is interested. Our choice of Ver_k^- is in line with [26] and makes the protocol analysis simpler.

We call an adversary in the above model a **full adversary**. The class of full adversaries for an authenticaiton protocol Ξ is denoted by $\mathcal{FULL}(\Xi)$. When Ξ is clear in the context, we denote it by $\mathcal{FULL}$ only. In addition, we denote the class of adversaries that make at most q_s authentication and at most q_v verification queries for fixed q_s, q_v, by $\mathcal{FULL}_{q_s,q_v}$. We also define $\mathcal{FULL}_q = \cup_{q_s+q_v\le q}\mathcal{FULL}_{q_s,q_v}$. We also define an *ideal adversary* with respect to $O \in \mathcal{FULL}$ as follows.

Definition 1. *A quantum algorithm $\mathcal{S}$ is an* **ideal adversary** *for an authentication protocol Σ with respect to $O \in \mathcal{FULL}$, if it first follows O to obtain ω_{q-1} and then updates it to ω'_q as follows. If query $q-1$ is the challenge authentication query, it only operates on Z register of ω_{q-1}; otherwise, it simply traces Y in ω_{q-1} to a dummy invalid authentication state (say, $|\aleph\rangle\langle\aleph|$). Finally, it sends YF for verification and the challenger returns $\mathrm{Ver}_k^-(\omega'_q)\otimes|acc\rangle\langle acc|_F$. Further, $\mathcal{S}$ is initialized with secret key k. The class of ideal adversaries with respect to O is denoted by $\mathcal{IDEAL}_O$. The union of $\mathcal{IDEAL}_O$ for all $O \in \mathcal{FULL}$ is denoted by $\mathcal{IDEAL}$.*

It is clear that an ideal adversary can not even see the authentication register Y if a challenge authentication query is indeed issued and hence can not tamper it. If a challenge authentication query is not issued, then $\mathcal{S}$ should be allowed to work on Y register arbitrarily to create a valid forgery with nonce r_q but he should not be able to succeed (i.e., the forgery should always be rejected, which is well captured by setting the forgery to an invalid dummy state). This ideal adversary perfectly captures the meaning of message authentication. We thus define the authentication security as, for any $O \in \mathcal{FULL}$, there is an ideal adversary $\mathcal{S} \in \mathcal{IDEAL}_O$ achieving the similar performance.

Definition 2. *An authentication protocol $\Sigma = (Auth, Ver)$ is* **secure** *if the following holds.*

- Correctness. *For any $\rho \in D(\mathcal{H}_\mathcal{M})$ and any $r \in \mathcal{R}$, $\mathrm{Ver}_k(r, \mathrm{Auth}_k(r,\rho)) = \rho \otimes |acc\rangle\langle acc|$.*
- Soundness. *Let ω_A be ω_q involving adversary A in the adversary model. Then, for any $O \in \mathcal{FULL}$, there exists $\mathcal{I} \in \mathcal{IDEAL}_O$ so that ω_O and $\omega_\mathcal{I}$ are indistinguishable.*

*Further, given $q \ge 1$, Σ is q-***secure** *if it is correct and sound for any $O \in \mathcal{FULL}_q$.*

Remark. A few subtle issues in the definition deserve a discussion.

First, the ideal adversary is given the secret key k. Recall our protection is only the authentication of the quantum message. We only guarantee that when the receiver accepts, the decoded message should be exactly the same as whatever sent by the sender. Since the ideal attacker can not operate passively or actively on register Y after the challenge query (i.e., the final authentication query in Stage q_v) is issued, the message authenticity is certainly preserved even if the attacker has the key k. The attacker's operation on Z register does not cause an authentication problem as the channel (i.e., Y register) is not under any adversarial operation. If a challenging authentication query is not issued, then the ideal adversary can only be allowed to send an invalid authentication message for verification and so the secret key k is still useless.

Second, a careful reader might notice that the ideal adversary $\mathcal{S}$ is referenced to an adversary $O \in \mathcal{FULL}$. Especially, they reach the same challenge authentication query ω'_{q-1} (if any). This is justified as follows. We are *only* concerned with the authentication property and want to guarantee that one that tampers, can not do better than one that does not, when the challenge authentication query with nonce r_{q-1} is indeed issued. We want to compare an ideal adversary with a real attacker when they use the same challenge authentication query. We can not allow $\mathcal{S}$ to generate a challenge authentication query freely: he can simulate the challenger interacting with O and when O outputs the final tampered authentication message, $\mathcal{S}$ verifies it using k (which will undo the authenticated part of the tampered authentication and trace the invalid part of this tampering to a dummy state). For the former, he sends it as the challenge authentication query (by measuring the decision bit) and sends the response for the challenge verification; for the latter case, he can create an invalid state for verification. It is easy to see that $\mathcal{S}$ will achieve the same result as O, no matter whether the authentication system is flawed or not. Essentially, this is because $\mathcal{S}$ can make the authentication message effectively the same as the tampered one by O in the real attack, by sending the verified output of O's final challenge verification query as his own challenge authentication query. In our definition, this attack is prevented as they will send the same challenge authentication query (if any).

Third, we require ω_O and $\omega_{\mathcal{S}}$ to be indistinguishable. Since they are parameterized by k implicitly, this essentially aligns the definition with Dupuis et al. [23] that averages distinguisher over k. This was commented in [26] as a flawed definition. They gave an example that the real attacker O can copy part of the secret appended to the true authentication message (in register Y) to register Z while the ideal adversary $\mathcal{S}$ can not (as he can not access the authentication register after the challenge authentication query) is issued. But this attack is useful only if O can use the leaked information to create some harm (such as forging a new authentication). However, their example does not leak any secret information that is used to generate the essential part of the authentication message (that is used in the verification). Specifically, this essential part remains unchanged in their attack. That is, the verification is not affected by their attack. So the authentication is in fact preserved. However, their definition regards this as insecure. We believe that this definition is overrestricted. Further, our indistinguishability averaging over k is line with quantum pseudorandom states [30] and quantum pseudorandom functions [44], as well as *all* the classical indistinguishability-based security notations such as IND-CCA and IND-CPA for encryptions [25] (although the classical setting is probably a different story).

Fourth, when a challenge authentication query has never been issued, the ideal adversary will define ω'_q in this setting to be a dummy state and hence the verification result is 0. Then, to be seure, the attack involving O must have $\omega_q \approx 0$. As a special case, it requires an impersonation attacker (i.e., $O \in \mathcal{FULL}_{0,1}$) to succeed negligibly only.

4.3 An Impersonation Attack Implies a Substitution Attack.

In the classical setting, an impersonation attack implies a substitution attack: the attacker can first take a random message $m \in \mathcal{M}$ and submits for an authentication query; then, it activates an impersonation attacker to generate an authentication σ_0 and outputs it as his forgery. It is successful if σ_0 is valid and $\mathrm{Ver}_k(\sigma_0) \neq m$. This occurs with probability at least $\Pr(\mathrm{Ver}_k(\sigma_0) \neq \perp)/2$ (i.e., the half impersonation success probability), as m is random and $|\mathcal{M}| \geq 2$ (note: $|\mathcal{M}| = 1$ does not need an authentication at all). Motivated by the classical idea, we transform a quantum impersonation attack to a quantum substitution attack.

Theorem 1. *Let $|\mathcal{M}| \geq 2$. If authentication protocol* (Auth, Ver) *is insecure against some $\mathcal{O} \in \mathcal{FULL}_{0,1}$ for arbitrary given nonce, then it is insecure against some $\mathcal{A} \in \mathcal{FULL}_{1,1} \backslash \mathcal{FULL}_{0,1}$, w.r.t. an unbounded distinguisher.*

Proof. For $m = 0, 1$, let us specify $\mathcal{A}_m \in \mathcal{FULL}_{1,1}\backslash\mathcal{FULL}_{0,1}$ that uses $\mathcal{O} \in \mathcal{FULL}_{0,1}$ as a subroutine. We fix nonce r. $\mathcal{A}_m$ sends $|m\rangle$ with r for a challenge authentication query and obtains $|c_{mk}\rangle_Y$, where k is the secret. Then, it activates $\mathcal{O}$ to generate a $|\varphi\rangle_{Z'Y'}$ w.r.t. nonce r (assume it is a pure state and if not, it can be purified with a reference system as part of Z'), which will result in $|\varphi'\rangle_{Z'Y'}$ if verified by Ver_k^- on Y' register. Let $|\psi_m\rangle_{Z'Y'Y} = |\varphi\rangle_{Z'Y'} \otimes |c_{mk}\rangle_Y$. $\mathcal{A}_m$ applies a swap gate on $|\psi_m\rangle_{Z'Y'Y}$ to obtain $|\psi'_m\rangle_{Z'Y'Y} = |\varphi\rangle_{Z'Y} \otimes |c_{mk}\rangle_{Y'}$. It outputs $|\psi'_m\rangle_{Z'Y'Y}$ as its forgery. The verification using Ver_k^- will turn it to $|\psi''_m\rangle = |\varphi'\rangle_{Z'Y} \otimes |c_{mk}\rangle_{Y'}$. Let $\varPsi_m = \mathbf{E}_{\varTheta}(|\varphi'\rangle\langle\varphi'|_{Z'Y} \otimes |c_{mk}\rangle\langle c_{mk}|_{Y'})$. Now recall that right after the challenge authentication query (i.e., before activating $\mathcal{O}$ above), the joint state becomes $|c_{mk}\rangle_Y \otimes |0\rangle_Z$, where $Z = Z'Y'$. $\mathcal{I}_m \in \mathcal{IDEAL}_{\mathcal{A}_m}$ will operate on Z register of this state and presents Y register for verification. Then, the final joint mixed state will become $\varPsi'_m = \mathbf{E}_{\varTheta}(|m\rangle\langle m|_Y \otimes \omega_m(k)_Z)$ for some ω. We want to show that $|\varPsi_m - \varPsi'_m|_1$ is non-negligible for at least one $m \in \{0, 1\}$ and **arbitrary** $\mathcal{I}_m$. In fact, this equals to

$$|\mathbf{E}_{\varTheta}(\, |\varphi'\rangle\langle\varphi'|_{Z'Y} \otimes |c_{mk}\rangle\langle c_{mk}|_{Y'} - |m\rangle\langle m|_Y \otimes \omega_m(k)_Z \,)|_1, \tag{10}$$

which, by tracing out $Y'Z'$ and using **Fact** 2, is lower bounded by

$$|\rho_Y - |m\rangle\langle m| t_m\,|_1, \tag{11}$$

where $\rho_Y := tr_{Z'}(\mathbf{E}_{\varTheta}(\varphi'_{Z'Y}))$ and $t_m = tr_Z(\mathbf{E}_{\varTheta}(\omega_m(k)_Z)) \leq 1$. Let $\rho_Y = \begin{pmatrix} A & C \\ C^* & B \end{pmatrix}$ with A a 2×2 matrix and C a $2 \times (n-2)$ matrix and B a $(n-2) \times (n-2)$ matrix, where $n = |\mathcal{M}|$. Let $W_m = \rho_Y - |m\rangle\langle m| t_m$. Then, $W_0 = \begin{pmatrix} A - \mathrm{DIAG}(t_0, 0) & C \\ C^* & B \end{pmatrix}$ and $W_1 = \begin{pmatrix} A - \mathrm{DIAG}(0, t_0) & C \\ C^* & B \end{pmatrix}$. Let $\hat{W}_0 = \begin{pmatrix} A - \mathrm{DIAG}(t_0, 0) & 0 \\ 0 & B \end{pmatrix}$. Then, $\hat{W}_0 = \frac{1}{2}(W_0 + \mathrm{DIAG}(I_2, -I_{n-2}) W_0 \mathrm{DIAG}(I_2, -I_{n-2}))$. Hence, by triangle inequality and also noticing that $\mathrm{DIAG}(I_2, -I_{n-2})$ is unitary, we have $|\hat{W}_0|_1 \leq |W_0|_1$. Also, $|\hat{W}_0|_1 = tr(|A - \mathrm{DIAG}(t_0, 0)|) + tr(|B|) = tr(|A - \mathrm{DIAG}(t_0, 0)|) + tr(B)$, as $B \geq 0$ (by noting that $0 \leq (0, \alpha)\rho_Y(0, \alpha)^T = \alpha B \alpha^T$ for any α of dimension $n-2$). Similarly, we can define $\hat{W}_1$ from W_1 and have $|\hat{W}_1|_1 = tr(|A - \mathrm{DIAG}(0, t_1)|) + tr(B)$.

Let $A = \begin{pmatrix} a & c \\ c^* & b \end{pmatrix}$ (w.r.t., $|0\rangle, |1\rangle$). Then, $A_0 := A - \mathrm{DIAG}(t_0, 0) = \begin{pmatrix} a - t_0 & c \\ c^* & b \end{pmatrix}$ and $A_1 := A - \mathrm{DIAG}(0, t_1) = \begin{pmatrix} a & c \\ c^* & b - t_1 \end{pmatrix}$. Eigenvalues of A_0 are

$$\lambda^0_{\pm} = \frac{a + b - t_0}{2} \pm \sqrt{\Big(\frac{a - t_0 - b}{2}\Big)^2 + |c|^2}. \tag{12}$$

Thus, $|A_0|_1 = |\lambda^0_+| + |\lambda^0_-| \geq \max(|a + b - t_0|, |a - t_0 - b|) \geq b$ (as $(|x| + |y|)/2 \geq |x - y|/2$). Similarly, we have $|A_1|_1 \geq \max(|a + b - t_1|, |b - t_1 - a|) \geq a$. Therefore,

$$\max(|A_0|_1, |A_1|_1) \geq (a + b)/2 = tr(A)/2. \tag{13}$$

Hence,

$$\max(|\hat{W}_0|_1, |\hat{W}_1|_1) \geq tr(A)/2 + tr(B) \geq tr(\rho_Y)/2. \tag{14}$$

That is, Eq. (11) (and further $|\varPsi_m - \varPsi'_m|_1$) for either $m = 0$ or $m = 1$, is lower bounded by $tr(\rho_Y)/2$, which is non-negligible (as $tr(\rho_Y)$ is the impersonation attack advantage over ideal adversary who simply outputs an invalid authentication (resulting in zero after verification)). Since trace distance is the distinguishing advantage for an unbounded distinguisher, we know that there exists an attacker $\mathcal{A} \in \{\mathcal{A}_0, \mathcal{A}_1\}$ that reaches the final state Ψ so that no ideal attacker $\mathcal{I} \in \mathcal{IDEAL}_{\mathcal{A}}$ can reach the final state Ψ' with Ψ and Ψ' indistinguishable. □

Fact 2 *Let A be a Hermitian operator on system XY. Then $|tr_X(A)|_1 \leq |A|_1$.*

Proof. Let $A = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ be the spectral decomposition of A with eigenvalues λ_i's. Then, $|A|_1 = \sum_i |\lambda_i|$. Now $tr_X(A) = \sum_i \lambda_i tr_X(|\psi_i\rangle\langle\psi_i|) = \sum_i \lambda_i \rho_i$, where $\rho_i = tr_X(|\psi_i\rangle\langle\psi_i|)$ is the density operator of system Y in $|\psi_i\rangle_{XY}$. Thus,

$$|tr_X(A)|_1 \leq \sum_i |\lambda_i||\rho_i|_1 = \sum_i |\lambda_i| = |A|_1, \tag{15}$$

where we use the fact $|\rho_i|_1 = 1$ for any density matrix ρ_i. This concludes the claim. ■

5 The AQA Framework

In this section, we present an Auth-QFT-Auth (AQA) authentication framework, which is a simple generalization of the AQA protocol [26] by replacing the original Wegman-Carter hash functions with general functions and introducing a nonce. We will show that as long as these functions satisfy certain conditions, the framework will be secure.

Let $k = (k_1, k_2) \leftarrow \mathcal{K}^2$ be the secret. Let $g(k_1, \cdot) : \mathcal{M} \times \{0,1\}^\nu \rightarrow \mathcal{T}_1$ be a keyed function with secret k_1 and $f(k_2, \cdot) : \mathcal{M} \times \mathcal{T}_1 \times \{0,1\}^\nu \rightarrow \mathcal{T}_2$ be a keyed function with secret k_2. Let $|\rho\rangle_{YZ}$ be the authentication input that is an entanglement on register Z and Y with $Y = MT_1T_2$, where M is the message register and T_1, T_2 are the tag registers for g and f respectively. Let the Schmidt decomposition of $|\rho\rangle_{YZ}$ be

$$|\rho\rangle_{YZ} = \sum_z \sqrt{\lambda_z}(\sum_{m\in\mathcal{M}} \alpha_{zm}|m\rangle_Y) \otimes |\psi_z\rangle_Z, \tag{16}$$

where $\{|\psi_z\rangle\}_z$ is a set of orthonormal states of size at most $|\mathcal{M}|$ and $\sum_z \lambda_z = 1$ and Recall that, as a convention, $|m\rangle_Y$ represents $|m\rangle_M|0\rangle_{T_1T_2}$. Assume that Alice will authenticate $|\rho\rangle_{YZ}$ to Bob using nonce $r \in \{0,1\}^\nu$. The protocol proceeds as follows.

1. (*inner authentication*) Upon input $|\rho\rangle_{YZ}$, Alice uses k_1 to update the state as
$$\sum_z \sqrt{\lambda_z}(\sum_{m\in\mathcal{M}} \alpha_{zm}|m, g(k_1, m|r)\rangle_Y) \otimes |\psi_z\rangle_Z, \tag{17}$$
2. (*Fourier transform*) Alice then makes a Fourier transform on Y register and gives
$$\frac{1}{\sqrt{MT_1}} \sum_z \sqrt{\lambda_z}(\sum_{m\in\mathcal{M}, x\in\mathcal{M}\times\mathcal{T}_1} \alpha_{zm}(-1)^{(m,g(k_1,m|r))\cdot x}|x\rangle_Y) \otimes |\psi_z\rangle_Z, \tag{18}$$
where we overload registers MT_1T_2 as $M = |\mathcal{M}|$ and $T_b = |\mathcal{T}_b|$ for $b = 1, 2$ with no confusion.
3. (*outer authentication*) Alice finally uses k_2 to update the state as
$$|\sigma\rangle = \frac{1}{\sqrt{MT_1}} \sum_z \sqrt{\lambda_z}(\sum_{m,x} \alpha_{zm}(-1)^{(m,g(k_1,m|r))\cdot x}|x, f(k_2, x|r)\rangle_Y) \otimes |\psi_z\rangle_Z \tag{19}$$
and sends it to Bob. Bob then verifies $|\sigma\rangle$ and outputs $\mathrm{Ver}_k|\sigma\rangle$.

6 Security of the AQA Framework

In this section, we show that when g and f satisfy some conditions, the AQA framework will be secure. This reduces the authentication problem to individual properties of g and f.

6.1 Definitions

A keyed function $f(k, \cdot)$ naturally induces an authentication protocol: given nonce r, $\mathrm{Auth}_k|m\rangle_Y = |m, f(k, m|r)\rangle$ and $\mathrm{Ver}_k|m, t\rangle_Y = |m\rangle_Y|acc\rangle_F$ or $|0\rangle_Y|rej\rangle_F$, dependent on whether t is valid. In the following, we introduce the *basis-respecting adversary.*

Definition 3. *A computational basis-respecting adversary is an authentication adversary similar to a full adversary, except that the adversarial operators applied to ω_{q-1}, consists of a unitary of form $V = \sum_i |i\rangle\langle i|_Y \otimes V_i$, followed by a projector on Z register, where $\{|i\rangle\}_i$ is the computational basis of $\mathcal{H}_\mathcal{Y}$ and each V_i is an arbitrary unitary on register Z. The class of computational basis respecting adversaries is denoted by $\mathcal{COMP}$.*

Similar to $\mathcal{IDEAL}_O$, we can also define $\mathcal{COMP}_O$ for any $O \in \mathcal{FULL}$: the adversary's operators, till it reaches the state ω_{q-1}, collide with that of O and differ only in the unitary and the projector that are applied to ω_{q-1}. We emphasize that $\mathcal{COMP}_O$ and $\mathcal{COMP}$ adversaries are not given the secret key k, unlike $\mathcal{IDEAL}_O$ or $\mathcal{IDEAL}$. When we mention $\mathcal{COMP}$ or $\mathcal{FULL}$, it is always concerned with the underlying authentication protocol Ξ. Formally, we write $\mathcal{COMP}(\Xi)$ for $\mathcal{COMP}$ and $\mathcal{FULL}(\Xi)$ for $\mathcal{COMP}$. But when the context is clear, we usually omit Ξ. Further, for the authentication protocol Ξ induced by keyed function $g(k, \cdot)$, $\mathcal{COMP}(\Xi)$ is denoted $\mathcal{COMP}(g)$.

We then introduce the oracle adversary class $\mathcal{FULL}^G$. This is similar to $\mathcal{FULL}$ adversaries, except that it can also query the (quantum or classical) oracle G **before** reaching ω_{q-1}. Similarly, we can define $\mathcal{COMP}^G$. In addition, we can define $\mathcal{COMP}^G_O$ for $O \in \mathcal{FULL}^G$, similar to $\mathcal{COMP}_O$ for $O \in \mathcal{FULL}$. Note here we do not allow $\mathcal{FULL}^G$, $\mathcal{COMP}^G$ and $\mathcal{COMP}^G_O$ to query oracle G when updating ω_{q-1} to ω'_q, because we need this relatively restricted adversary in proving our security theorem.

We next introduce the notion of *reduction* to relate $\mathcal{FULL}^G$ and $\mathcal{COMP}^G$. This is to formalize the intuition: whatever can be achieved by $O \in \mathcal{FULL}^G$, can also be achieved by some adversary in $\mathcal{COMP}^G_O$.

Definition 4. *Let* $\Xi = (\mathrm{Auth}, \mathrm{Ver})$ *be an authentication scheme and G is an oracle.* $(\mathrm{Auth}, \mathrm{Ver})$ *is said* $(\mathcal{FULL}^G,\mathcal{COMP}^G)$**-reducible** *if for any adversary $O \in \mathcal{FULL}^G$, there exists $A \in \mathcal{COMP}^G_O$ so that ω_O and ω_A are indistinguishable, where ω_E is ω_q in the security game with adversary E and distinguisher can make one query to G.*

The authentication protocol from a keyed function f naturally induces the reducibility of f.

Definition 5. *For a keyed function $f(k, \cdot)$ and an oracle G, f is said $(\mathcal{FULL}^G,\mathcal{COMP}^G)$-reducible if the underlying authentication protocol is $(\mathcal{FULL}^G,\mathcal{COMP}^G)$-reducible.*

6.2 Security Theorem

In the following, we prove the security theorem for the AQA authentication framework. The idea is as follow. For any adversary $O \in \mathcal{FULL}$, we need to present an ideal adversary $\mathcal{I}_{ideal} \in \mathcal{IDEAL}_O$ so that the challenge verification query will generate the similar output. The proof uses the sequence of game technique. First of all, $f(k_2,\cdot)$ is $(\mathcal{FULL}^{g(k_1,\cdot)}, \mathcal{COMP}^{g(k_1,\cdot)})$-reducible and so we only need to consider a computational-basis-preserving adversary $\mathcal{I}$. Second, since $g(k_1,\cdot)$ is pseudorandom, we only need to consider the game with $g(k_1,\cdot)$ replaced by a complete random function (i.e., random oracle) $g(\cdot)$. Further, since a random oracle is perfectly inditinguishable from its compressed random oracle (see [42]), we can replace $g(\cdot)$ by its compressed random oracle counterpart. To be more intuitive, suppose $\mathcal{I}$ will issue a challenge authentication query ω'_{q-1} and this state is a pure state. Then, by Schmidt decomposition, we can always write it into a format like

$$|\omega'_{q-1}\rangle = \sum_{z\mathbf{y}} \sqrt{\lambda_{z\mathbf{y}}} \sum_{m} \alpha_{zm\mathbf{y}} |m,0\rangle_Y |\psi_{z\mathbf{y}}\rangle |\mathbf{y}\rangle_D, \tag{20}$$

where $\sum_m |\alpha_{zm\mathbf{y}}|^2 = 1$ and $\sum_{z\mathbf{y}} \lambda_{z\mathbf{y}} = 1$ and $|| \, |\psi_{z\mathbf{y}}\rangle \, || = 1$. After the query, it comes

$$|\omega_{q-1}\rangle = \frac{1}{\sqrt{MT_1}} \sum_{z\mathbf{y}} \sqrt{\lambda_{z\mathbf{y}}} \sum_{mtu} \alpha_{zm\mathbf{y}} (-1)^{(m,t)\cdot u} |u, f(k_2, u|r_q)\rangle_Y |\psi_{z\mathbf{y}}\rangle F_{D_{m|r_q}} |\mathbf{y} \cup (t)_{m|r_q}\rangle_D. \tag{21}$$

Since $\mathcal{I}$ is computational basis-respecting, after the operations by $\mathcal{I}$, it will have

$$|\omega'_q\rangle = \frac{1}{\sqrt{MT_1}} \sum_{z\mathbf{y}} \sqrt{\lambda_{z\mathbf{y}}} \sum_{mtu} \alpha_{zm\mathbf{y}} (-1)^{(m,t)\cdot u} |u, f(k_2, u|r_q)\rangle_Y |\varphi_{uz\mathbf{y}}\rangle F_{D_{m|r_q}} |\mathbf{y} \cup (t)_{m|r_q}\rangle_D. \tag{22}$$

Clearly, the outer authentication tag is valid and hence after the outer layer verification and inverse Fourier transform, it becomes

$$\frac{1}{MT_1} \sum_{z\mathbf{y}} \sqrt{\lambda_{z\mathbf{y}}} \sum_{mtuab} \alpha_{zm\mathbf{y}} (-1)^{(m+a,t+b)\cdot u} |a,b\rangle_Y |\varphi_{uz\mathbf{y}}\rangle F_{D_{m|r_q}} |\mathbf{y} \cup (t)_{m|r_q}\rangle_D. \tag{23}$$

Now we need to perform the inner verification. In this case, if $a = m$, then b must equal t in order to be valid. The verification of this part becomes

$$\sum_{z\mathbf{y}} \sqrt{\lambda_{z\mathbf{y}}} \sum_{m} \alpha_{zm\mathbf{y}} |m\rangle_Y \otimes \frac{1}{\sqrt{MT_1}} \sum_{u} |\varphi_{uz\mathbf{y}}\rangle \otimes |\mathbf{y}\rangle_D, \tag{24}$$

where we have used the fact that $\frac{1}{\sqrt{MT_1}} \sum_t F|t\rangle = |\bot\rangle$. Notice that Eq. (24) can be obtained from Eq. (20) by only operating on Z register, which can be done by an ideal adversary.

For the part with $a \neq m$, it has that $a|r_q$ was not queried and so $a|r_q \notin \mathbb{X}(\mathbf{y})$. Hence, the verification result will be

$$\frac{1}{(MT_1)^{3/2}} \sum_{z\mathbf{y}} \sqrt{\lambda_{z\mathbf{y}}} \sum_{mtuab:a\neq m} \alpha_{zm\mathbf{y}} (-1)^{(m+a,t+b)\cdot u} |a\rangle_Y |\varphi_{uz\mathbf{y}}\rangle F_{D_{m|r_q,a|r_q}} |\mathbf{y} \cup (t)_{m|r_q} \cup (b)_{a|r_q}\rangle_D.$$

By careful calculation, we will show that this part has a very small norm. Thus, the total verification result is almost equal to whatever can be done by an ideal adversary. This completes the proof idea.

Theorem 2. *For keyed functions $g(k_1,\cdot) : \mathcal{M} \times \{0,1\}^\nu \to \mathcal{T}_1$ and $f(k_2,\cdot) : \mathcal{M} \times \mathcal{T}_1 \times \{0,1\}^\nu \to \mathcal{T}_2$ with $k_1, k_2 \leftarrow \mathcal{K}$, $g(k_1,\cdot)$ is a pseudorandom function with $M = O(T_1^{1/2} 2^{-\log^2 \nu})$ and $f(k_2,\cdot)$ is $(\mathcal{FULL}^{g(k_1,\cdot)}, \mathcal{COMP}^{g(k_1,\cdot)})$-reducible. Then, the AQA framework with f and g is secure.*

Proof. The correctness holds obviously. We now focus on the soundness. Let ω_A be the finally verified state ω_q involving adversary A with implicit parameter set Θ (including $k = (k_1, k_2)$ and sampled parameters specifying g and f). Note here we do not choose to analyze the security directly on the averaged ω_A over Θ and instead analayze it over ω_A parametered by Θ. For simplicity, we assume $\Theta = \{k\}$ only, as the proof does not explicitly use other parameters and will be almost identical to this simplified case. In the original security game, attacker O makes q_s authentication queries and q_v verification queries for some polynomially bounded q_s, q_v, including a challenge verification query ω'_q. Denote ω'_q involving adversary A by ω'_A. We will use $\mathrm{Ver}_2(\cdot)$ to represent the verification of the outer authentication (i.e., the authentication using $f(k_2, \cdot)$) and use $\mathrm{Ver}_1(\cdot)$ to represent the verification of the inner authentication (i.e., the authentication using $g(k_1, \cdot)$). In both cases, the verification is defined via Eq. (8) for the first $q_v - 1$ verification queries and via Eq. (9) for the challenge verification query.

Given $O \in \mathcal{FULL}$ against the AQA framework, we can formulate it to $O_f \in \mathcal{FULL}^{g(k_1,\cdot)}(f)$. Specifically, when given access to $g(k_1, \cdot)$ oracle, O_f invokes O. Each authentication query from O can be answered by an oracle query to $g(k_1, \cdot)$ oracle, followed by Fourier transform on register Y by himself and then an authentication query to $f(k_2, \cdot)$-oracle and the final authentication query to $f(k_2, \cdot)$, if any (due to the challenge authentication query from O), is the challenge authentication query of O_f. Further, each verification query from O can be decomposed in a similar but reverse manner. The final verification query to $f(k_2, \cdot)$ by oracle O_f (due to the challenge verification query from O) will be his challenge verification query. Besides, O_f follows O for the remaining (intermediate) operators. From our formulation of O_f, we can see that $\omega'_O = \omega'_{O_f}$. By the $(\mathcal{FULL}^{g(k_1,\cdot)}, \mathcal{COMP}^{g(k_1,\cdot)})$-reducibility of f, there exists $\mathcal{I}_f \in \mathcal{COMP}^{g(k_1,\cdot)}_{O_f}(f)$ so that

$$\mathrm{Ver}_2(\omega'_{O_f}) \text{ and } \mathrm{Ver}_2(\omega'_{\mathcal{I}_f}). \tag{25}$$

are indistinguishable, where, by this reducibility definition, the distinguisher is also allowed to make one query to $g(k_1, \cdot)$ oracle. This $\mathcal{I}_f$ gives an adversary $\mathcal{I} \in \mathcal{COMP}_O$ against AQA protocol: $\mathcal{I}_f$ follows O_f till reaching ω_{q-1} and so $\mathcal{I}$ follows O will reach the same ω_{q-1}, then $\mathcal{I}$ applies basis-preserving $V_{\mathcal{I}_f}$ on YZ and projector Π on Z, on ω_{q-1} to obtain $\omega'_{\mathcal{I}}$ (which equals $\omega'_{\mathcal{I}_f}$), where we especially notice that no query to $g(k_1, \cdot)$ by $\mathcal{I}_f$ during the transition from ω_{q-1} to $\omega'_{\mathcal{I}_f}$ (by definition of $\mathcal{COMP}^G$ and $\mathcal{COMP}^G_O$). By the definition of $\mathcal{I}_f$, $\omega'_{\mathcal{I}}$ is exactly identical to $\omega'_{\mathcal{I}_f}$. Thus, summarizing the results so far, we have

$$\mathrm{Ver}_2(\omega'_O) = \mathrm{Ver}_2(\omega'_{O_f}) \approx \mathrm{Ver}_2(\omega'_{\mathcal{I}_f}) = \mathrm{Ver}_2(\omega'_{\mathcal{I}}). \tag{26}$$

Especially, $\mathrm{Ver}_2(\omega'_O) \approx \mathrm{Ver}_2(\omega'_{\mathcal{I}})$, which is still indistinguishable if further applying Inverse Fourier transform and making one oacle query to $g(k_1, \cdot)$, as the distinguisher of Eq. (25) is allowed to make one query to $g(k_1, \cdot)$. That is, $\omega_O \approx \omega_{\mathcal{I}}$. Denote the game with $\mathcal{I}$ and oracle $g(k_1, \cdot)$ by $\mathbf{G}_0$. We use $A(\mathbf{G})$ to indicate variable A in game $\mathbf{G}$. So we have

$$\omega_O(\mathbf{G}_0) \approx \omega_{\mathcal{I}}(\mathbf{G}_0). \tag{27}$$

Then, we modify the protocol so that $g(k_1, \cdot)$ is replaced by a completely random function g. Denote this game by $\mathbf{G}_1$. By the pseudorandomness of $g(k_1, \cdot)$,

$$\omega_{\mathcal{I}}(\mathbf{G}_1) \text{ and } \omega_{\mathcal{I}}(\mathbf{G}_0) \tag{28}$$

are indistinguishable, given the oracle access to $g(\cdot)/g(k_1, \cdot)$ (where $g(\cdot)$ represents the purely random function from $\mathcal{M}$ to $\mathcal{T}_1$).

Now let us analyze $\omega'_{\mathcal{I}}(\mathbf{G}_1)$. We start with the expression of $\omega'_{\mathcal{I}}(\mathbf{G}_1)$, a state in $\mathcal{H}_{\mathcal{M}\times\mathcal{T}_1\times\mathcal{T}_2}$. We first assume it is a pure state in $\mathcal{H}_{\mathcal{M}\times\mathcal{T}_1\times\mathcal{T}_2}$ (parameterized by g, k_1) and later will generalize to the mixed state case. Recall that if a state γ is a pure state, our notational convention has a representation $|\gamma\rangle$ for it. So $|\omega'_{\mathcal{I}}\rangle$ is a superposition over states with symbols in $\mathcal{M} \times \mathcal{T}_1 \times \mathcal{T}_2$, parameterized by g and k_2. According to our adversary model, the nonce r_q is either not queried before or it is equal to the nonce r_{q-1} in the challenge authentication query. We first consider the latter case and leave the former case for later. Then, by Schmidt decomposition, we can write the challenge authentication query and its response as

$$|\omega'_{q-1}(\mathbf{G}_1)\rangle = \sum_z \sqrt{\lambda_z} \sum_m \alpha_{zm} |m\rangle_Y |\psi_z\rangle_Z, \tag{29}$$

$$|\omega_{q-1}(\mathbf{G}_1)\rangle = \frac{1}{\sqrt{MT_1}} \sum_{zu} \sqrt{\lambda_z} \sum_m \alpha_{zm} (-1)^{(m,t)\cdot u} |u, f(k_2, u|r_q)\rangle_Y |\psi_z\rangle_Z, \tag{30}$$

where $\sum_m |\alpha_{zm}|^2 = 1, \sum_z \lambda_z = 1$, $\{\psi_z\}_z$ orthonormal with dimension of z at most M and $t = g(m, r_q)$ with g a purely random function.

Observing that $\mathcal{I}$ is computational-basis-respecting, after applying $\mathit{\Pi} V_{\mathcal{I}_f}$, we have

$$|\omega'_{\mathcal{I}}(\mathbf{G}_1)\rangle = \frac{1}{\sqrt{MT_1}} \sum_{zu} \sqrt{\lambda_z} \sum_m \alpha_{zm} (-1)^{(m,t)\cdot u} |u, f(k_2, u|r_q)\rangle_Y |\psi_{uz}\rangle_Z, \tag{31}$$

where $|| \, |\psi_{uz}\rangle_Z \, || \le 1$ (due to the projector $\mathit{\Pi}$). After Ver$_2$, it becomes

$$\frac{1}{\sqrt{MT_1}} \sum_{zu} \sqrt{\lambda_z} \sum_m \alpha_{zm} (-1)^{(m,t)\cdot u} |u\rangle_Y |\psi_{uz}\rangle_Z, \tag{32}$$

After inverse Fourier Transform, it becomes

$$\frac{1}{MT_1} \sum_{zu} \sqrt{\lambda_z} \sum_{mab} \alpha_{zm} (-1)^{(m+a,t+b)\cdot u} |a, b\rangle_Y |\psi_{uz}\rangle_Z, \tag{33}$$

with $t = g(m, r_q)$. After Ver$_1$, it becomes

$$|\omega_{\mathcal{I}}(\mathbf{G}_1)\rangle = \frac{1}{MT_1} \sum_{zu} \sqrt{\lambda_z} \sum_{ma} \alpha_{zm} (-1)^{(m+a,t+b)\cdot u} |a\rangle_Y |\psi_{uz}\rangle_Z, \tag{34}$$

where $t = g(m, r_q)$ and $b = g(a, r_q)$. Let us write $|\omega_{\mathcal{I}}(\mathbf{G}_1)\rangle = |\omega_{ideal}\rangle + |\omega_{error}\rangle$, where

$$|\omega_{ideal}\rangle = \sum_z \sqrt{\lambda_z} \sum_m \alpha_{zm} |m\rangle_Y \otimes \frac{1}{MT_1} \sum_u |\psi_{uz}\rangle_Z, \tag{35}$$

$$|\omega_{error}\rangle = \frac{1}{MT_1} \sum_{zu} \sqrt{\lambda_z} \sum_{ma:\ m\neq a} \alpha_{zm} (-1)^{(m+a,t+b)\cdot u} |a\rangle_Y |\psi_{uz}\rangle_Z. \tag{36}$$

The following lemma claims that $\mathbf{E}_{g,k_2}(|| \, |\omega_{error}\rangle \, ||^2)$ is small (see the proof in Appendix B).

Lemma 5. *$\mathbf{E}_{g,k_2}(||\,|\omega_{error}\rangle\,||^2)$ is negligible.*

So far we have considered the case where r_q has been queried. When it was not queried before, we let the verified result $|\omega_{\mathcal{I}}(\mathbf{G}_1)\rangle = |\omega_{ideal}\rangle + |\omega_{error}\rangle$, where $|\omega_{ideal}\rangle := 0$. We will show that $\mathbf{E}_{g,k_2}(|||\omega_{error}\rangle||^2)$ is still small (which is still done in Lemma 5).

Then, we will construct an ideal adversary $\mathcal{I}_{ideal}$ so that $|\omega_{\mathcal{I}_{ideal}}\rangle = |\omega_{ideal}\rangle$. Once we have this, we know that

$$\begin{aligned}
&\mathbf{E}_{g,k_2}(|\,|\omega_{\mathcal{I}}\rangle\langle\omega_{\mathcal{I}}| - |\omega_{ideal}\rangle\langle\omega_{ideal}|\,|_1) && (37)\\
\le&\mathbf{E}_{g,k_2}(2|\,|\omega_{ideal}\rangle\langle\omega_{error}|\,|_1 + |\,|\omega_{error}\rangle\langle\omega_{error}|\,|_1) && (38)\\
\le&\mathbf{E}_{g,k_2}(2\sqrt{\langle\omega_{ideal}|\omega_{ideal}\rangle\cdot\langle\omega_{error}|\omega_{error}\rangle} + \langle\omega_{error}|\omega_{error}\rangle) && (39)\\
\le&3\mathbf{E}_{g,k_2}(\sqrt{\langle\omega_{error}|\omega_{error}\rangle}) \le 3\sqrt{\mathbf{E}_{g,k_2}(\langle\omega_{error}|\omega_{error}\rangle)} && (40)\\
=&3\sqrt{\mathbf{E}_{g,k_2}(||\,|\omega_{error}\rangle\,||^2)} = \mathbf{negl}(\nu), && (41)
\end{aligned}$$

where Eqs. (39)(40) holds from Cauchy-Schwarz inequality. As $||\,|\omega_{error}\rangle\,||^2 = |\,|\omega_{error}\rangle\langle\omega_{error}|\,|_1$ and Lemma 5 holds for pure state case, by joint convexity of trace distance from [34, Theorem 9.3], we have $\mathbf{E}_{g,k_2}(|\omega_{error}|_1)$ is negligible for mixed state. Thus, when $\omega'_{\mathcal{I}}(\mathbf{G}_1)$ has evolved from a mixed state, we still have $\mathbf{E}_{g,k_2}(|\omega_{\mathcal{I}} - \omega_{ideal}|_1)$ negligible, where $\omega_{\mathcal{I}}$ (resp. ω_{ideal}) is the mixed state of $|\omega_{\mathcal{I}}\rangle$ (resp. $|\omega_{ideal}\rangle$).

We will present the ideal adversary $\mathcal{I}_{ideal}$ so that $\omega_{\mathcal{I}_{ideal}} = \omega_{ideal}$. If $\mathcal{I}_{ideal}$ can do this with access to $g(\cdot)/g(k_1,\cdot)$ oracle, then ω_{ideal} with $g(k_1,\cdot)$ and ω_{ideal} with purely random g are indistinguishable by the pseudorandomness of $g(k_1,\cdot)$. That is, $\omega_{ideal}(\mathbf{G}_1) \approx \omega_{ideal}(\mathbf{G}_0)$. We have already know previously that $\omega_{\mathcal{I}}(\mathbf{G}_1) \approx \omega_{\mathcal{I}}(\mathbf{G}_0)$ (from pseudorandomness of $g(k,\cdot)$). Since $\mathbf{E}_{g,k_2}(|\omega_{\mathcal{I}}(\mathbf{G}_1) - \omega_{ideal}(\mathbf{G}_1)|_1)$ is negligibly, it follows $\omega_{ideal}(\mathbf{G}_1)$ and $\omega_{\mathcal{I}}(\mathbf{G}_1)$ are indistinguishable and hence $\omega_{ideal}(\mathbf{G}_0)$ and $\omega_{\mathcal{I}}(\mathbf{G}_0)$ are indistinguishable. This concludes our theorem.

It remains to construct $\mathcal{I}_{ideal}$ with access to oracle $g(k_1,\cdot)/g(\cdot)$ (in $\mathbf{G}_0$ / $\mathbf{G}_1$). The strategy is modified from [26, Theorem 16]. Essentially, $\mathcal{I}_{ideal}$ can be constructed from $\mathcal{I}$ as follows (details will be given soon). Till query $q-1$, $\mathcal{I}_{ideal}$ follows $\mathcal{I}$ (or O) as it is not restricted (in comparison with O or $\mathcal{I}$). If $\mathcal{I}$ does not make a challenge authentication query, then $\omega_{ideal} = 0$. Otherwise, after receiving the reply $\omega_{(q-1)\mathcal{I}}$ from the challenge authentication query, assume that $\mathcal{I}$ will apply unitary $V_{\mathcal{I}_f}$ on YZ, followed by a projector Π on register Z. Then, $\mathcal{I}_{ideal}$ performs $V_{\mathcal{I}_f}$ on AZ and Π on Z to try to force the global state to achieve the desired result, where register A will be a simulated copy of Y register by $\mathcal{I}_{ideal}$. The details of $\mathcal{I}_{ideal}$ are as follows.

1. $\mathcal{I}_{ideal}$ has the secret keys and can simulate the oracles $g(\cdot)/g(k_1,\cdot)$ and $f(k_2,\cdot)$. He then follows $\mathcal{I}$ until it receives the reply of the challenge authentication query. Now the global state is $\omega_{(q-1)\mathcal{I}}$ (i.e., ω_{q-1} in the security game with adversary $\mathcal{I}$).
2. $\mathcal{I}_{ideal}$ then creates an entangled state $|\Phi_{k_2}\rangle_{AA'} = \frac{1}{\sqrt{MT_1}}\sum_x |x, f(k_2, x|r_q)\rangle_A |x, f(k_2, x|r_q)\rangle_{A'}$ (with access to oracle $f(k_2,\cdot)$, maintained by himself), where $A \otimes A'$ is isomorphic to $Y \otimes Y$.
3. $\mathcal{I}_{ideal}$ then applies $V_{\mathcal{I}_f}$ to register AZ and Π on Z. Then, $\mathcal{I}_{ideal}$ makes the projector $|\Phi_{k_2}\rangle\langle\Phi_{k_2}|$ (i.e., it applies projective measurement $(|\Phi_{k_2}\rangle\langle\Phi_{k_2}|, I - |\Phi_{k_2}\rangle\langle\Phi_{k_2}|)$ and discards the outcome 1). From Lemma below, the result is $\omega_{ideal} \otimes \Phi_{k_2}$. Tracing out Φ_{k_2} gives $|\omega_{ideal}\rangle$.

The mixed state case follows by linearity of $\mathcal{I}_{ideal}$. This concludes our theorem. □

Lemma 6. $\omega_{\mathcal{I}_{ideal}}(\mathbf{G}_i) = \omega_{ideal}(\mathbf{G}_i), i = 0, 1.$

Proof. We consider $\mathbf{G}_1$ only as $\mathbf{G}_0$ is similar. For simplicity, we assume that $\omega_{(q-1)\mathcal{I}}$ is a pure state. From Eq. (30), we have

$$|\omega_{(q-1)\mathcal{I}}(\mathbf{G}_1)\rangle = \frac{1}{\sqrt{MT_1}} \sum_{zu} \sqrt{\lambda_z} \sum_{m} \alpha_{zm}(-1)^{(m,t)\cdot u}|u, f(k_2, u|r_q)\rangle_Y |\psi_z\rangle_Z. \tag{42}$$

Consider $|\omega_{(q-1)\mathcal{I}}(\mathbf{G}_1)\rangle \otimes |\Phi_k\rangle_{AA'}$. After $V_{\mathcal{I}_f}$ on AZ and Π on Z, by Eq. (31), it becomes

$$\frac{1}{MT_1} \sum_{zu} \sqrt{\lambda_z} \sum_{mx} \alpha_{zm}(-1)^{(m,t)\cdot u}|u, f(k_2, u|r_q)\rangle_Y |\psi_{xz}\rangle_Z \otimes |x, s\rangle_A |x, s\rangle_{A'}, \tag{43}$$

where $s = f(k_2, x|r_q)$ (as $\mathcal{I}$ is computational-basis respecting). Projecting register AA' on Φ_{k_2} and using the fact $\langle x, s, x, s|\Phi_{k_2}\rangle = \frac{1}{\sqrt{MT_1}}$, it becomes

$$\frac{1}{MT_1} \sum_{zu} \sqrt{\lambda_z} \sum_{mx} \alpha_{zm}(-1)^{(m,t)\cdot u}|u, f(k_2, u|r_q)\rangle_Y |\psi_{xz}\rangle_Z \otimes \Phi_{k_2}^{AA'}, \tag{44}$$

which is

$$\sum_{zu} \sqrt{\lambda_z} \sum_{m} \alpha_{zm}(-1)^{(m,t)\cdot u}|u, f(k_2, u|r_q)\rangle_Y \otimes \frac{1}{MT_1} \sum_{x} |\psi_{xz}\rangle_Z \otimes \Phi_{k_2}^{AA'}, \tag{45}$$

Tracing out AA' and verifying Y (using Ver_2, $\mathrm{QFT}^\dagger$ and Ver_1), it becomes

$$\sum_{zu} \sqrt{\lambda_z} \sum_{m} \alpha_{zm}|m\rangle_Y \otimes \frac{1}{MT_1} \sum_{x} |\psi_{xz}\rangle_Z. \tag{46}$$

This is exactly $|\omega_{ideal}\rangle$. The mixed state case follows by linearity. ■

When the pseudorandomness of $g(k_1, \cdot)$ holds only for q queries, we immediately have the following corollary.

Corollary 1. *For keyed functions $g(k_1, \cdot) : \mathcal{M} \times \{0,1\}^\nu \to \mathcal{T}_1$ and $f(k_2, \cdot) : \mathcal{M} \times \mathcal{T}_1 \times \{0,1\}^\nu \to \mathcal{T}_2$ with $k_1, k_2 \leftarrow \mathcal{K}$, let $g(k_1, \cdot)$ be q-pseudorandom and that $f(k_2, \cdot)$ be $(\mathcal{FULL}_q^{g(k_1,\cdot)}, \mathcal{COMP}_q^{g(k_1,\cdot)})$-reducible. Then, the AQA authentication framework is secure against $\mathcal{FULL}_q^{g(k_1,\cdot)}$.*

7 Security Separation of Authentication with and without Verification Queries: the Quantum Setting

In Section 3.3, we showed that there exists a classical MAC scheme that is existentially unforgeable while it is completely insecure when verification queries are additionally allowed. In that scheme, the attacker can use each verification query to extract one bit of the secret key and eventually crack the whole key. In this section, we extend this idea to the quantum setting.

7.1 Notations

Define $\mathcal{FULL}^{\text{-}}$ (resp. $\mathcal{COMP}^{\text{-}}$) to be a subset of $\mathcal{FULL}$ (resp. $\mathcal{COMP}$) that does not make a verification query other than the final challenge verification query. Similarly, $\mathcal{FULL}^{\text{-}g(k_1,\cdot)}$ (resp. $\mathcal{COMP}^{\text{-}g(k_1,\cdot)}$) can be adapted from $\mathcal{FULL}^{g(k_1,\cdot)}$ (resp. $\mathcal{COMP}^{g(k_1,\cdot)}$). Further, we can extend the $(\mathcal{FULL}^{g(k_1,\cdot)}, \mathcal{COMP}^{g(k_1,\cdot)})$-reducibility of f daptively: $f(k, \cdot)$ is said $(\mathcal{FULL}^{\text{-}g(k',\cdot)}, \mathcal{COMP}^{\text{-}g(k',\cdot)})$-**reducible** for $k, k \leftarrow \mathcal{K}$ if the induced authentication protocol of f is $(\mathcal{FULL}^{\text{-}g(k_1,\cdot)}, \mathcal{COMP}^{\text{-}g(k_1,\cdot)})$-reducible.

7.2 Variant of AQA and Authentication protocol of f

Variant of AQA. We now mmodify AQA to AQA$'$ so that it has the limited security claimed above. The construction is to basically follow AQA except with the following changes.

- First, assume AQA has a key space $\mathcal{K}^2 = \{0,1\}^{2\nu}$ and then the modified protocol has a key space $\{0,1\}^{2\nu+1}$. The key is sampled as $s = (s_1, \cdots, s_{2\nu+1}) \leftarrow \{0,1\}^{2\nu+1}$. It defines $(k_1, k_2) = (s_1 \oplus s_{2\nu+1}, \ldots, s_{2\nu} \oplus s_{2\nu+1})$ and use (k_1, k_2) as the secret key in AQA protocol.
- Second, the protocol execution follows AQA with (k_1, k_2), except we attach $|z, i, u\rangle_A = |0, 0, 0\rangle_A$ to the authentication message. That is, the authentication message is now $|\sigma\rangle|0, 0, 0\rangle_A$, where $|\sigma\rangle$ is defined at Eq. (19) and we remind that now the register of authentication is now $Y' = YA$.. The verification is the same as AQA, except that it first measures A register in the computational basis, resulting in the classical value $|z, i, u\rangle_A$, and then checks if (z, i, u) is consistent: either $z = 0$ or ($z = 1$ & $s_i = u$) (similar to the verification in Section 3.3), followed by the normal verification in AQA.

Variant of Induced Authentication of f. The authentication protocol induced by function f can be modified similarly: take $s \leftarrow \{0,1\}^{2\nu+1}$ and define $k = (k_1, k_2)$ as above; use k_2 as the secret key of f; the authentication message and verification are adapted similarly as in AQA$'$. Note that k_1 is not explicitly used in the protocol but s is used in the verification. We call it the *modified authentication protocol of f.*

7.3 Security Analysis

In this section, we analyze the security of AQA$'$. We show that it is secure when no verification queries (other than the challenge one) is issued while it is completely insecure when they are permitted. Our proof needs the following lemma. It basically states that if f is $(\mathcal{FULL}^{\text{-}g(k_1,\cdot)}, \mathcal{COMP}^{\text{-}g(k_1,\cdot)})$-reducible, then the modified authentiction protocol of f (especifially, with $k = (k_1, k_2)$ as specified there) also has the same property. The idea is that an adversary in $\mathcal{FULL}^{\text{-}}$ (resp. $\mathcal{COMP}^{\text{-}}$) can not make a 'learning' verification query and register A returned from the challenger is always dummy $|0\rangle_A$, the protocol before the final verificaiton query is essentially the protocol induced by f. In addition, since the actual key $k = (k_1, k_2)$ for the protocol reducibility is a one-time pad encryption, any bit of s remains uniformly random, given k. So attacker can not succeed better than a random guess of s_i which can be done by $\mathcal{COMP}^{\text{-}g(k_1,\cdot)}$ attacker. So the reducibility of f guarantees that a $\mathcal{COMP}^{\text{-}g(k_1,\cdot)}$ adversary can achieve the same performance as a $\mathcal{FULL}^{\text{-}g(k_1,\cdot)}$ adversary. The detailed proof is to turn this idea into a formal reduction.

Lemma 7. *If f is $(\mathcal{FULL}^{\text{-}g(k_1,\cdot)}, \mathcal{COMP}^{\text{-}g(k_1,\cdot)})$-reducible, then the modified authentication protocol of f is also $(\mathcal{FULL}^{\text{-}g(k_1,\cdot)}, \mathcal{COMP}^{\text{-}g(k_1,\cdot)})$-reducible with an $\mathcal{I} \in \mathcal{COMP}_O^{\text{-}g(k_1,\cdot)}$ for any $O \in \mathcal{FULL}^{\text{-}g(k_1,\cdot)}$ so that $\mathcal{I}$ always keeps A register as $|0, 0, 0\rangle_A$.*

Proof. For any adversary $O \in \mathcal{FULL}^{\text{-}g(k_1,\cdot)}$ against the modified protocol, we can construct an adversary $O' \in \mathcal{FULL}^{\text{-}g(k_1,\cdot)}$ against the original protocol (without register A). O' does this as follows.

- For any authentication query from O, O' simply forwards the register Y to his own challenger and keeps the register A (reset to $|0\rangle_A$, for instance, by swapping with an unused qubit (with state $|0\rangle$) in his own register Z'). When receiving the reply, it forwards YA back to O.

- For the final **challenge verification** query from O, O' first makes a measurement in the computational basis on register A. If $|z, i, u\rangle$ has $z = 1$, then with probability $1/2$, it traces Y to an invalid symbol $|\aleph\rangle\langle\aleph|$; with probability $1/2$, it keeps Y unchanged. Next, it sends Y for verification and also resets A register to $|0, 0, 0\rangle$. When receiving the verification result, it forwards YA registers back to O. Given the authentication queries, any s_i for any i is independent of the adversary view. Hence, the view of O is according to the real game distribution and $\omega_O = \omega_{O'} \otimes |0, 0, 0\rangle\langle 0, 0, 0|_A$

By the reducibility of f in the original protocol, there exists adversary $\mathcal{I}' \in \mathcal{COMP}_{O'}^{\text{-}g(k_1,\cdot)}$ for the original protocol so that the final state $\omega_{O'}$ and $\omega_{\mathcal{I}'}$ are indistinguishable, where the distinguisher is allowed to make one query to $g(k_1, \cdot)$ oracle. Next, we reformulate $\mathcal{I}'$ against the original protocol as an adversary $\mathcal{I} \in \mathcal{COMP}_{O}^{\text{-}g(k_1,\cdot)}$ against the modified protocol. The operation of $\mathcal{I}$ is very simple: it simply follows $\mathcal{I}'$ except that it has an extra register A that always has $|0, 0, 0\rangle_A$ and for each authentication query and the (challenge) verification query, it also forwards register A to the challenger. Thus, we have that $\omega_{\mathcal{I}} = \omega_{\mathcal{I}'} \otimes |0, 0, 0\rangle\langle 0, 0, 0|_A$. Now since $\omega_{\mathcal{I}'} \approx \omega_{O'}$ with distinguisher allowed to query $g(k_1, \cdot)$ once, we know that $\omega_{\mathcal{I}} \approx \omega_{O'} \otimes |0, 0, 0\rangle\langle 0, 0, 0|_A = \omega_O$ still with distinguisher allowed to query $g(k_1, \cdot)$ once. This concludes the proof. ■

Theorem 3. *AQA$'$ scheme is secure without verification queries but insecure when $2\nu + 1$ verification queries are additionally allowed.*

Proof. We first show that it is insecure when $2\nu+1$ verification queries are additionally allowed, besides the authentication queries. The strategy is similar to the classical case, except that each quantum authentication message is modified only once for the verification query. That is, it issues query for message m_i with nonce r_i and upon receiving the authentication message $\omega_i \otimes |0, 0, 0\rangle\langle 0, 0, 0|_A$, it modified to $\omega_i \otimes |1, i, 0\rangle\langle 1, i, 0|_A$ as the verification query. It is accepted if and only if $s_i = 0$. After $2\nu + 1$ authentication queries and verification queries, s is recovered and the attacker can forge any authentication message for any quantum state.

Now we prove that AQA$'$ is secure against attacker without a verification query. Fix the nonce vector $\mathbf{r}$. Recall that ω_B is the final accepted state involving adversary B with implicit parameter set Θ (including $s = (s_1, s_2)$). For simplicity, we assume $\Theta = \{s\}$ only as otherwise the proof does not explicitly use other parameters and will be almost identical to this simple case. In the original security game, attacker O makes q queries including the challenge verification query. Recall that the final state before the challenge verification adversary B is ω'_B.

Given $O \in \mathcal{FULL}^{\text{-}}$ against the AQA$'$ framework, we can formulate it as $O_f \in \mathcal{FULL}^{\text{-}g(k_1,\cdot)}$ (i.e., $\mathcal{FULL}^{\text{-}}(f)$ with access to oracle $g(k_1, \cdot)$), against the reducibility of f, where each tag query from O can be answered by a query to $g(k_1, \cdot)$ oracle, followed by Fourier transform on register Y by himself and then a query to $f(k_2, \cdot)$-oracle and the final query to $f(k_2, \cdot)$ is the challenge verification query of O_f. Beside these, O_f follows O for the remaining (intermediate) operators. From our formulation of O_f, we can see that $\omega'_O = \omega'_{O_f}$. By the $(\mathcal{FULL}^{\text{-}g(k_1,\cdot)}, \mathcal{COMP}^{\text{-}g(k_1,\cdot)})$-reducibility of f and Lemma 7, there exists $\mathcal{I}_f \in \mathcal{COMP}_{O_f}^{\text{-}g(k_1,\cdot)}$ for the modified protocol of f, so that

$$\mathrm{Ver}_2(\omega'_{O_f}) \text{ and } \mathrm{Ver}_2(\omega'_{\mathcal{I}_f}). \tag{47}$$

are indistinguishable and also $\mathcal{I}_f$ always keeps register A as $|0, 0, 0\rangle_A$, where the indistinguishability still holds after QFT and Ver_1 by accessing to $g(k_1, \cdot)$ (as the distinguisher is allowed to make one query to $g(k_1,\cdot)$ oracle). This $\mathcal{I}_f$ gives an adversary $\mathcal{I} \in \mathcal{COMP}^{\text{-}}$ against AQA$'$ protocol, where it follows O till it reaches $\omega_{(q-1)O}$ and then it follows $\mathcal{I}_f$ to apply some unitary $V_{\mathcal{I}_f}$ and projector $\mathit{\Pi}$ to obtain $\omega'_{\mathcal{I}}$. By the definitions of $\mathcal{I}_f$ and $\mathcal{COM}_O^{g(k_1,\cdot)}$, $\omega'_{\mathcal{I}} = \omega'_{\mathcal{I}_f}$ (hence summarizing the results so far gives $\omega_O \approx \omega_{\mathcal{I}}$) and A register is always kept as $|0,0,0\rangle_A$. Especially, the authentication queries and the (challenge) verification query all have $|0,0,0\rangle_A$. Thus, the remaining proof exactly following Theorem 2 can give an ideal adversary $\mathcal{I}_{ideal}$ with $\omega_{\mathcal{I}_{ideal}} \approx \omega_{\mathcal{I}}$ and hence $\omega_{\mathcal{I}_{ideal}} \approx \omega_O$. This completes our proof. □

8 Some Reducible Functions

In this section, we show that if $f(k,\cdot) : \mathcal{M} \to \mathcal{T}$ for $k_2 \leftarrow \mathcal{K}$ is $2q$-wise independent function, then its induced authentication protocol can be reduced to a computational-basis-preserving adversary that makes at most q oracle queries. The idea is as follows. Since f is $2q$-wise independent, by Lemma 2, we can replace it with a random oracle without changing its reducibility. Furthermore, the quantum random oracle and the compressed random oracle $CStO$ are indistinguishable [42]. We can change to work with $CStO$ without changing the reducibility. So we only need to prove the reducibility when the random oracle is $CStO$. In this case, for simplicity, assume that the challenge authentication query is $\sum_{\mathbf{y}} \alpha_{\mathbf{y}} |m,0\rangle_Y |\psi_z\rangle_Z |\mathbf{y}\rangle_D$ (where $\sum_{\mathbf{y}} |\alpha_{\mathbf{y}}|^2 = 1$) with nonce r_{q-1}, then by our model assumption nonce r_{q-1} was not queried before and hence the reply will be

$$\omega_{q-1} = \frac{1}{|\mathcal{T}|^{1/2}} \sum_{t\mathbf{y}} \alpha_{\mathbf{y}} |m,t\rangle_Y |\psi_z\rangle_Z F_{D_{m|r_{q-1}}} |\mathbf{y} \cup (t)_{m|r_{q-1}}\rangle_D. \tag{48}$$

Then, attacker will apply unitary V on YZ and $\mathit{\Pi}$ on Z and we can assume the result is

$$|\omega'_q\rangle = \frac{1}{|\mathcal{T}|^{1/2}} \sum_{tab\mathbf{y}} \alpha_{\mathbf{y}} \beta_{ab}^{mtz\mathbf{y}} |a,b\rangle_Y |\gamma_{ab}^{mtz\mathbf{y}}\rangle_Z F_{D_{m|r_{q-1}}} |\mathbf{y} \cup (t)_{m|r_{q-1}}\rangle_D \tag{49}$$

with $\sum_{ab} |\beta_{ab}^{mtz\mathbf{y}}|^2 = 1$ and $|\gamma_{ab}^{mtz\mathbf{y}}\rangle_Z$ is a unit vector. Now the challenger will verify $|\omega'_q\rangle$. Since $a|r_{q-1}$ for $a \neq m$ is not recorded in $\mathbf{y} \cup (t)_{m|r_{q-1}}$ while $m|r_{q-1}$ is, the verification result will be

$$\begin{aligned} |\omega_q\rangle =& \frac{1}{|\mathcal{T}|^{1/2}} \sum_{t\mathbf{y}} \alpha_{\mathbf{y}} \beta_{mt}^{mtz\mathbf{y}} |m\rangle_Y |\gamma_{mt}^{mtz\mathbf{y}}\rangle_Z F_{D_{m|r_{q-1}}} |\mathbf{y} \cup (t)_{m|r_{q-1}}\rangle_D + & (50)\\ & \frac{1}{|\mathcal{T}|} \sum_{t\mathbf{y}ab:a\neq m} \alpha_{\mathbf{y}} \beta_{ab}^{mtz\mathbf{y}} |a\rangle_Y |\gamma_{ab}^{mtz\mathbf{y}}\rangle_Z F_{D_{m|r_{q-1},a|r_{q-1}}} |\mathbf{y} \cup (t)_{m|r_{q-1}} \cup (b)_{a|r_{q-1}}\rangle_D. & (51) \end{aligned}$$

The crucial point is that the part at Eq. (51) is very small. Indeed, $\{|\mathbf{y} \cup (t)_{m|r_{q-1}} \cup (b)_{a|r_{q-1}}\rangle\}_{tb}$ are almost orthonormal. To see the idea, we can assumt they are truly orthonomal. Then, the norm of part of Eq. (51) is at most $1/|\mathcal{T}|^{1/2}$, as $\sum_{ab} |\beta_{ab}|^2 \leq 1$ and $\sum_{\mathbf{y}} |\alpha_{\mathbf{y}}|^2 = 1$. Therefore, $|\omega_q\rangle$ is almost equal to the part at Eq. (50). If we can construct a computational-basis-preserving adversary $\mathcal{I}$ to generate this part, then we are done. Notice that this part is the verified result of partial sum in Eq. (49) with $a = m, b = t$. That is, this part is computational-basis-preserving, compared with ω_{q-1} at Eq. (48) with only changes on Z registers. This can be done by a computional-basis-repserving adversary. In the following, we will implement this proof idea rigorously.

Theorem 4. *If $f(k,\cdot) : \mathcal{M} \to \mathcal{T}$ is a $2q$-wise independent function with $k \leftarrow \mathcal{K}$ and $|\mathcal{M}| = O(|\mathcal{T}|^{1/2} 2^{-\log^2 \lambda})$, then $f(k,\cdot)$ is $(\mathcal{FULL}_q^R, \mathcal{COMP}_q^R)$-reducible with R independent of k.*

Proof. For an adversary $O \in \mathcal{FULL}_q^R$, we need to find $\mathcal{I} \in \mathcal{COMP}_{qO}^R$ to satisfy the reducibility condition: $\omega_O \approx \omega_{\mathcal{I}}$, where the distinguisher is allowed to make one query to oracle R. Toward this, we fix nonces $\mathbf{r}$ for simplicity. Then, O will issue q authentication and verification queries and the final query is a challenge verification query. We will not explicitly mention the query to R oracle as that can be incorporated into operators between authentication and verification queries. Then, O finally will receive the verified state ω_O. Denote this game by $\mathbf{G}_0$.

Game $\mathbf{G}_1$. We modify $\mathbf{G}_0$ so that $f(k, \cdot)$ is replaced by random oracle $f(\cdot)$ (i.e., a purely random function from $\mathcal{M}$ to $\mathcal{T}$). By Lemma 2, the final states in $\mathbf{G}_1$ and $\mathbf{G}_0$ are perfectly indistinguishable (note: since k is independent of R, the oracle access to R by distinguisher does not affect the indistinguishability as he can simulate the oracle by himself; we will ignore R for the similar reason from now). Further, we can replace $|f\rangle$ with its uniform superposition $\frac{1}{\sqrt{N}} \sum_f |f\rangle$ (i.e., the standard random oracle StO): since register $|f\rangle$ is a control register for all the operators in the game (see Section 2), measuring $|f\rangle$ at the beginning or the end of the game does not change the final joint state ω_O. Here N is the total number of f's.

Now before proceeding, we characterize $\mathbf{G}_1$. We consider two cases: query $q-1$ is a challenge authentication query (so $r_q = r_{q-1}$), or, it is not challenge authentication query (so r_q has never been queried).

We first consider the case where query $q-1$ is a challenge authentication query (so $r_q = r_{q-1}$). We first only consider the pure state case (i.e., ω'_{q-1} is a pure state $|\omega'_{q-1}\rangle$). The extension to the mixed state will be considered later. Then, by Schmidt decomposition, we can assume that $|\omega'_{q-1}\rangle$ (the challenge authentication query) can be written as

$$\omega'_{q-1} = \sum_{zf} \sqrt{\lambda_{zf}} \sum_m \alpha_{zmf} |m, 0\rangle_Y |\psi_{zf}\rangle_Z |f\rangle_D, \tag{52}$$

where the distribution of f is absorbed into λ_{zf} (i.e., $\sum_{zg} \lambda_{zf} = 1$) and $\{|\psi_{zf}\rangle\}_z$ for each f is a set of orthonormal states of size at most M and $\sum_m |\alpha_{zmf}|^2 = 1$ for each z, f. After the query, it becomes

$$|\omega_{q-1}\rangle = \sum_{zf} \sqrt{\lambda_{zf}} \sum_m \alpha_{zmf} |m, t\rangle_Y |\psi_{zf}\rangle_Z |f\rangle_D, \tag{53}$$

where $t = f(m, r_q)$. Then, the adversary O applies a unitary V on registers Y and Z, followed by a projector $\mathit{\Pi}$ on Z, so that $|m, t\rangle \otimes |\psi_{zf}\rangle \mapsto |\mathit{\Gamma}_{mtzf}\rangle$. We rewrite $|\mathit{\Gamma}_{mtzf}\rangle$ as

$$|\mathit{\Gamma}_{mtzf}\rangle = \sum_{ab} \beta_{ab}^{mtzf} |a, b\rangle_Y \otimes |\gamma_{ab}^{mtzf}\rangle_Z, \tag{54}$$

where $\{|\gamma_{ab}^{mtzf}\rangle\}_{a,b}$ is a collection of vectors for register Z with norm at most 1 (due to projector $\mathit{\Pi}$) and $\sum_{ab} |\beta_{ab}^{mtzf}|^2 = 1$. So the joint state after applying $\mathit{\Pi} V$ is

$$|\omega'_q\rangle = \sum_{zf} \sqrt{\lambda_{zf}} \sum_{mab} \alpha_{zmf} \beta_{ab}^{mtzf} |a, b\rangle \otimes |\gamma_{ab}^{mtzf}\rangle \otimes |f\rangle_D, \tag{55}$$

Game $\mathbf{G}_2$. We modify $\mathbf{G}_1$ to $\mathbf{G}_2$ so that random oracle StO is replaced by compressed random oracle $CStO$. By [42, Lemma 4], $\mathbf{G}_1$ and $\mathbf{G}_2$ are perfectly indistinguishable, where the distinguisher has access to the final state $|\omega_q\rangle$ on all registers except for the random oracle register. Indeed, the distinguisher is the (attacker, challenger) pair, except the random oracle is separated apart but remains accessible to the distinguisher. When the random oracle is $CStO$, the final state is from $\mathbf{G}_2$; otherwise, it is from $\mathbf{G}_1$. Then, [42, Lemma 4] implies the indistinguishability.

Analysis of $\mathbf{G}_2$. We now analyze $\mathbf{G}_2$ and construct adversar $\mathcal{I} \in \mathcal{COMP}_O^R$ to satisfy the reducibility property. Toward this, we first notice that Eq. (53) in $\mathbf{G}_1$ will be changed in $\mathbf{G}_2$ as

$$|\omega_{q-1}(\mathbf{G}_2)\rangle = \frac{1}{|\mathcal{T}|^{1/2}} \sum_{z\mathbf{y}t} \sqrt{\lambda_{z\mathbf{y}}} \sum_{m} \alpha_{zm\mathbf{y}} |m,t\rangle_Y |\psi_{z\mathbf{y}}\rangle_Z F_{D_{m|r_q}} |\mathbf{y} \cup (t)_{m|r_q}\rangle_D, \tag{56}$$

as $m|r_{q-1}$ was not queried prior to query $q-1$ due to the existence of nonce r_{q-1} in the challenge query and so $r_q = r_{q-1}$ according to our adversary model. Thus, after the operation by O, we have

$$|\omega'_q\rangle = \frac{1}{|\mathcal{T}|^{1/2}} \sum_{z\mathbf{y}t} \sqrt{\lambda_{z\mathbf{y}}} \sum_{mab} \alpha_{zm\mathbf{y}} \beta_{ab}^{mtz\mathbf{y}} |a,b\rangle \otimes |\gamma_{ab}^{mtz\mathbf{y}}\rangle \otimes F_{D_{m|r_q}} |\mathbf{y} \cup (t)_{m|r_q}\rangle_D. \tag{57}$$

Also, since attacker only made q authentication/verification queries, $\mathbb{X}(\mathbf{y})$ has at most q non-$\perp$ elements.

We can divide $|\omega'_q\rangle$ into two parts: $|\omega'_q\rangle = |\omega'_{ideal}\rangle + |\omega'_{error}\rangle$, where

$$|\omega'_{ideal}\rangle = \frac{1}{|\mathcal{T}|^{1/2}} \sum_{z\mathbf{y}t} \sqrt{\lambda_{z\mathbf{y}}} \sum_{mb} \alpha_{zm\mathbf{y}} \beta_{mb}^{mtz\mathbf{y}} |m,b\rangle \otimes |\gamma_{mb}^{mtz\mathbf{y}}\rangle \otimes F_{D_{m|r_q}} |\mathbf{y} \cup (t)_{m|r_q}\rangle_D, \tag{58}$$

$$|\omega'_{error}\rangle = \frac{1}{|\mathcal{T}|^{1/2}} \sum_{z\mathbf{y}t} \sqrt{\lambda_{z\mathbf{y}}} \sum_{mab:a\neq m} \alpha_{zm\mathbf{y}} \beta_{ab}^{mtz\mathbf{y}} |a,b\rangle \otimes |\gamma_{ab}^{mtz\mathbf{y}}\rangle \otimes F_{D_{m|r_q}} |\mathbf{y} \cup (t)_{m|r_q}\rangle_D. \tag{59}$$

After verification, the accepted part in $|\omega'_{ideal}\rangle$ will be

$$|\omega_{ideal}\rangle = \frac{1}{|\mathcal{T}|^{1/2}} \sum_{z\mathbf{y}t} \sqrt{\lambda_{z\mathbf{y}}} \sum_{m} \alpha_{zm\mathbf{y}} \beta_{mt}^{mtz\mathbf{y}} |m\rangle \otimes |\gamma_{mt}^{mtz\mathbf{y}}\rangle \otimes F_{D_{m|r_q}} |\mathbf{y} \cup (t)_{m|r_q}\rangle_D, \tag{60}$$

On the other hand, for each $a \neq m$ in $|\omega'_{error}\rangle$, $a|r_q \notin \mathbb{X}(\mathbf{y} \cup (t)_{m|r_q})$, as it can not be recorded in the previous query due to the uniqueness of r_q. Since $f(a, r_\nu)$ is uniformly random, it follows that after the verification, $|\omega'_{error}\rangle$ becomes

$$|\omega_{error}\rangle = \frac{1}{|\mathcal{T}|} \sum_{z\mathbf{y}t} \sqrt{\lambda_{z\mathbf{y}}} \sum_{mab:a\neq m} \alpha_{zm\mathbf{y}} \beta_{ab}^{mtz\mathbf{y}} |a\rangle \otimes |\gamma_{ab}^{mtz\mathbf{y}}\rangle \otimes F_{D_{\{m|r_q, a|r_q\}}} |\mathbf{y} \cup (b)_{a|r_q} \cup (t)_{m|r_q}\rangle_D.$$

In the following, we upper bound $||\, |\omega_{error}\rangle\, ||$ which turns out to be small. Toward this, let us first look at the inner product between

$$|a\rangle F_{D_{\{m|r_q, a|r_q\}}} |\mathbf{y} \cup (t)_{m|r_q} \cup (b)_{a|r_q}\rangle_D \text{ and } |a'\rangle F_{D_{\{m'|r_q, a'|r_q\}}} |\mathbf{y}' \cup (b')_{a'|r_q} \cup (t')_{m'|r_q}\rangle_D, \tag{61}$$

with $a \notin \mathbb{X}(\mathbf{y})$ and $a' \notin \mathbb{X}(\mathbf{y}')$. We have the following cases:

- If $(a,b) \neq (a',b')$ or $\mathbf{y} \neq \mathbf{y}'$, the inner product is 0.
- If $(a,b) = (a',b')$ and $\mathbf{y} = \mathbf{y}'$ but $m' \neq m$, the inner product is $1/|\mathcal{T}|$.
- If $(a,b) = (a',b')$, $\mathbf{y} = \mathbf{y}'$ and $m' = m$ but $t \neq t'$, the inner product is 0.

– If $(a,b) = (a',b')$ and $\mathbf{y} = \mathbf{y}'$ and $(m,t) = (m',t')$, the inner product is 1.

Therefore, we can write $|| \, |\omega_{error}\rangle \, ||^2 = (A_0 + A_1)/|\mathcal{T}|^2$, where

$$A_0 = \sum_{z'z\mathbf{y}tmab:a\neq m} \sqrt{\lambda_{z\mathbf{y}}\lambda_{z'\mathbf{y}}}\alpha_{zm\mathbf{y}}\bar{\alpha}_{z'm\mathbf{y}}\beta_{ab}^{mtz\mathbf{y}}\bar{\beta}_{ab}^{mtz'\mathbf{y}}\langle\gamma_{ab}^{mtz'\mathbf{y}}|\gamma_{ab}^{mtz\mathbf{y}}\rangle, \tag{62}$$

$$A_1 = \frac{1}{|\mathcal{T}|}\sum_{z'z\mathbf{y}m'mtt'ab:|\{a,m',m\}|=3} \sqrt{\lambda_{z\mathbf{y}}\lambda_{z'\mathbf{y}}}\alpha_{zm\mathbf{y}}\bar{\alpha}_{z'm'\mathbf{y}}\beta_{ab}^{mtz\mathbf{y}}\bar{\beta}_{ab}^{m't'z'\mathbf{y}}\langle\gamma_{ab}^{m't'z'\mathbf{y}}|\gamma_{ab}^{mtz\mathbf{y}}\rangle, \tag{63}$$

Consider A_1 first. We have

$$\begin{aligned}
|A_1| \leq & \frac{1}{|\mathcal{T}|}\sum_{z'z\mathbf{y}tt'm'm:\ m'\neq m} \sqrt{\lambda_{z\mathbf{y}}\lambda_{z'\mathbf{y}}}|\alpha_{zm\mathbf{y}}\bar{\alpha}_{z'm'\mathbf{y}}| \cdot \sum_{ab}|\beta_{ab}^{mtz\mathbf{x}\mathbf{y}}\bar{\beta}_{ab}^{m't'z'\mathbf{y}}| & (64)\\
\leq & |\mathcal{T}|\sum_{z'z\mathbf{y}m'm:m\neq m'} \sqrt{\lambda_{z\mathbf{y}}\lambda_{z'\mathbf{y}}}|\alpha_{zm\mathbf{y}}\bar{\alpha}_{z'm'\mathbf{y}}| \text{ (by Cauchy-Schwarz with } \textstyle\sum_{ab}|\beta_{ab}^{u}|^2 = 1) & (65)\\
\leq & |\mathcal{T}|\sum_{\mathbf{y}zz'}\frac{1}{2}\sum_{m'm}(\lambda_{z\mathbf{y}}|\alpha_{zm\mathbf{y}}|^2 + \lambda_{z'\mathbf{y}}|\bar{\alpha}_{z'm'\mathbf{y}}|^2) & (66)\\
= & |\mathcal{T}| \cdot |\mathcal{M}|\sum_{\mathbf{y}zz'}\lambda_{z\mathbf{y}} = |\mathcal{M}|^2|\mathcal{T}|, & (67)
\end{aligned}$$

where we note that the dimension of Schmidt decomposition from z or z' is at most $|\mathcal{M}|$.

Then, we focus on A_0. Notice that

$$\begin{aligned}
|A_0| \leq & \sum_{z'z\mathbf{y}tmab} \sqrt{\lambda_{z\mathbf{y}}\lambda_{z'\mathbf{y}}}|\alpha_{zm\mathbf{y}}\bar{\alpha}_{z'm\mathbf{y}}\beta_{ab}^{mtz\mathbf{y}}\bar{\beta}_{ab}^{mtz'\mathbf{y}}\langle\gamma_{ab}^{mtz'\mathbf{y}}|\gamma_{ab}^{mtz\mathbf{y}}\rangle|, & (68)\\
\leq & \sum_{z'z\mathbf{y}tmab} \sqrt{\lambda_{z\mathbf{y}}\lambda_{z'\mathbf{y}}}|\alpha_{zm\mathbf{y}}\bar{\alpha}_{z'm\mathbf{y}}\beta_{ab}^{mtz\mathbf{y}}\bar{\beta}_{ab}^{mtz'\mathbf{y}}| & (69)\\
\leq & |\mathcal{T}|\sum_{z'z\mathbf{y}m} \sqrt{\lambda_{z\mathbf{y}}\lambda_{z'\mathbf{y}}}|\alpha_{zm\mathbf{y}}\bar{\alpha}_{z'm\mathbf{y}}|, \quad \text{/* as } \textstyle\sum_{ab}|\beta_{ab}^{u}|^2 = 1 \text{ for any } u \text{ */} & (70)\\
\leq & |\mathcal{T}|\sum_{z'z\mathbf{y}} \sqrt{\lambda_{z\mathbf{y}}\lambda_{z'\mathbf{y}}}, \quad \text{/* as } \textstyle\sum_{m}|\alpha_{zm\mathbf{y}}|^2 = 1 \text{ for any } z\mathbf{y} \text{ */} & (71)\\
= & |\mathcal{T}|\sum_{\mathbf{y}}(\sum_{z}\sqrt{\lambda_{z\mathbf{y}}})^2 \leq |\mathcal{T}||\mathcal{M}|\sum_{\mathbf{y}}\sum_{z}\lambda_{z\mathbf{y}} = |\mathcal{M}| \cdot |\mathcal{T}|. & (72)
\end{aligned}$$

where Eqs. (70)-(72) all have used Cauchy-Schwarz inequality. Therefore,

$$|| \, |\omega_{error}\rangle \, ||^2 \leq \frac{1}{|\mathcal{T}|^2} \cdot (|\mathcal{M}|^2|\mathcal{T}| + |\mathcal{M}||\mathcal{T}|) \leq \frac{2|\mathcal{M}|^2}{|\mathcal{T}|}. \tag{73}$$

Next, we consider the case where no challenge authentication query is made (at query $q-1$). We assume the query q for challenge verification is $|\omega'_q\rangle = \sum_{mt\mathbf{y}}\alpha_{mt\mathbf{y}}|m,t\rangle|\psi_{mt\mathbf{y}}\rangle|\mathbf{y}\rangle$, where $|\psi_{mt\mathbf{y}}\rangle$ is a unit vector and $\sum_{mt\mathbf{y}}|\alpha_{mt\mathbf{y}}|^2 \leq 1$. Then, define $|\omega'_{ideal}\rangle = 0$ and $|\omega'_{error}\rangle = |\omega'_q\rangle$. It suffices to show that $|||\omega_q\rangle||^2$ is small. Note that since $m|r_q$ was not queried before, $m|r_q \notin \mathbb{X}(\mathbf{y})$ and hence after the verification, it becomes

$$|\omega_q\rangle = \frac{1}{|\mathcal{T}|^{1/2}} \sum_{mt\mathbf{y}} \alpha_{mt\mathbf{y}} |m\rangle |\psi_{mt\mathbf{y}}\rangle F_{D_{m|r_q}} |\mathbf{y} \cup (t)_{m|r_q}\rangle. \tag{74}$$

Consider the inner product of

$$|m\rangle F_{D_{m|r_q}} |\mathbf{y} \cup (t)_{m|r_q}\rangle \text{ and } |m'\rangle F_{D_{m'|r_q}} |\mathbf{y}' \cup (t')_{m'|r_q}\rangle,$$

where $*|r_q$ was not queried to the random oracle. We have the following observations:

- If $m \neq m'$ or $\mathbf{y} \neq \mathbf{y}'$, then the inner product will be 0.
- If $m = m'$ and $\mathbf{y} = \mathbf{y}'$ but $t' \neq t$, then the inner product is still 0.
- If $(m, t) = (m', t')$ and $\mathbf{y} = \mathbf{y}'$, then the inner product is 1.

Therefore, $||\, |\omega_q\rangle \,||^2 = \frac{1}{|\mathcal{T}|} \sum_{mt\mathbf{y}} |\alpha_{mt\mathbf{y}}|^2 \cdot ||\, |\psi_{mt\mathbf{y}}\rangle \,||^2 \leq 1/|\mathcal{T}|$, as $\sum_{mt\mathbf{y}} |\alpha_{mt\mathbf{y}}|^2 \leq 1$.

Therefore, similar to Eq. (37) (without the expectation), we have

$$|\, |\omega_q\rangle\langle\omega_q| - |\omega_{ideal}\rangle\langle\omega_{ideal}| \,|_1 \leq 3 \cdot \sqrt{\max(2|\mathcal{M}|^2/|\mathcal{T}|, 1/|\mathcal{T}|)} \leq \frac{5|\mathcal{M}|}{|\mathcal{T}|^{1/2}}. \tag{75}$$

By convexity of trace distance, we know that generally $|\omega_q - \omega_{ideal}|_1$ has the same upper bound. Since we are in $\mathbf{G}_2$, this precisely has $|\omega_q(\mathbf{G}_2) - \omega_{ideal}(\mathbf{G}_2)|_1 \leq \frac{5|\mathcal{M}|}{|\mathcal{T}|^{1/2}}$. Now if we can construct $\mathcal{I} \in \mathcal{COMP}_O^R$, with access to the compressed random oracle to achieve state $|\omega_{ideal}\rangle$ (that is, $\omega_{\mathcal{I}}(\mathbf{G}_2) = \omega_{ideal}(\mathbf{G}_2)$), then we have $|\omega_O(\mathbf{G}_2) - \omega_{\mathcal{I}}(\mathbf{G}_2)|_1 \leq \frac{5|\mathcal{M}|}{|\mathcal{T}|^{1/2}}$, negligible! Especially, $\omega_O(\mathbf{G}_2) \approx \omega_{\mathcal{I}}(\mathbf{G}_2)$. When the compressed random oracle of $\mathcal{I}$ is replaced by StO or $f(k, \cdot)$, we have $\omega_{\mathcal{I}}(\mathbf{G}_1)$ or $\omega_{\mathcal{I}}(\mathbf{G}_0)$. We hence have $\omega_{\mathcal{I}}(\mathbf{G}_2) \approx \omega_{\mathcal{I}}(\mathbf{G}_1) \approx \omega_{\mathcal{I}}(\mathbf{G}_0)$, by the indistinguishability between $CStO$ and StO and the $2q$-independence of $f(k, \cdot)$. Since we also have $\omega_O(\mathbf{G}_2) \approx \omega_O(\mathbf{G}_1) \approx \omega_O(\mathbf{G}_0)$, it follows that $\omega_O(\mathbf{G}_0) \approx \omega_{\mathcal{I}}(\mathbf{G}_0)$. This will conclude our theorem.

It remains to construct $\mathcal{I}$ that has the access to $CStO$ to achieve ω_{ideal}. This is as follows.

1. $\mathcal{I}$ first creates auxiliary registers $M'T'Z_2$ where $M'T'$ is isomorphic to MT and Z_2 is a qubit register.
2. $\mathcal{I}$ follows O, until it reaches the state $|\omega_{q-1}\rangle$ (here we assume $|\omega_{q-1}\rangle$ is a pure state and mixed state case is obained by linearity). If query $q-1$ is not a challenge authentication query, it creates a dummy invalid authentication message (such as $|\aleph\rangle$) as $|\omega'_q\rangle$ and results in $|\omega_q\rangle = 0$. Otherwise, $|\omega_{q-1}\rangle$ will be the reply of the challenge authentication query and then $\mathcal{I}$ creates $|\omega_{q-1}\rangle_{MTZD}|0\rangle_{M'T'Z_2}$.
3. Then, $\mathcal{I}$ coherently copies MT registers into $M'T'$. Then apply the final unitary V of the real attacker on $M'T'Z$ registers, followed by projector Π on Z. Then, it compares MT and $M'T'$ in the computational basis. If they are equal, then wrote 0 on Z_2; otherwise, it changes to 1. Then, it applies projector $|0\rangle\langle 0|_{Z_2}$ on Z_2 (i.e., it applies the projector measurement $(|0\rangle\langle 0|, |1\rangle\langle 1|)$ and we are concerned with the outcome is 0).
4. It then applies unitary $|m\rangle_M |t\rangle_T |m'\rangle_{M'} |t'\rangle_{T'} \mapsto |m\rangle_M |t\rangle_T |0\rangle_{M'} |0\rangle_{T'}$, as $m' + m' = t' + t' = 0$. Since the measurement on Z_2 has outcome 0, it always has
$$|m\rangle_M |t\rangle_T |m\rangle_{M'} |t\rangle_{T'} \mapsto |m\rangle_M |t\rangle_T |0\rangle_{M'} |0\rangle_{T'}. \tag{76}$$
That is, after the measurement, the state on $M'T'Z_2$ is $|0\rangle_{M'T'Z_2}$, which is in tensor product with YZ registers and can be ignored. Thus, the state on YZ is perfectly $|\omega_{ideal}\rangle$.

The verification that following the action of $\mathcal{I}$ above (starting from ω_{q-1} in the expression at Eq. (56)) will indeed result in ω_{ideal}, is routine and straightforward. We omit it here. □

Corollary 2. *If $f(k,\cdot):\mathcal{M}\rightarrow\mathcal{T}$ is pseudorandom with $k\leftarrow\mathcal{K}$ and $|\mathcal{M}|=O(|\mathcal{T}|^{1/2}2^{-\log^2\nu})$, then it is ($\mathcal{FULL}^R$,$\mathcal{COMP}^R$)-reducible for any oracle R that is independent of k.*

Proof. The proof almost follows from Theorem 4. In Theorem 4, we replace $2q$-wise independent function with random oracle f without changing the final verified state as the adversary only makes q queries while the pseudorandomness of $f(k,\cdot)$ in our setting allows us to do this for any polynomially bounded q with only changing the final verified state by a negligible distinguishing gap. Once $f(k,\cdot)$ is replaced by purely random f, the remaining proof follows Theorem 4. ■

Corollary 3. *If $H:\mathcal{K}\otimes\mathcal{M}\rightarrow\mathcal{T}$ is a random oracle and $q=O(|\mathcal{K}|^{1/2}2^{-\log^2\nu})$ and $|\mathcal{M}|=O(|\mathcal{T}|^{1/2}2^{-\log^2\nu})$, then $H(k,\cdot)$ is ($\mathcal{FULL}^R$,$\mathcal{COMP}^R$)-reducible for for $k\leftarrow\mathcal{K}$ and any oracle R independent of k, where q is the number of random oracle queries from adversary.*

Proof. This again (as in Corollary 2) follows the proof of Theorem 4. Since H is already a random oracle, the proof can start from $\mathbf{G}_1$. Further, by Proposition 1, we can assume that adversary O does not make any random oracle queries (other than authentication and verifiation queries). The replacement only induces a negligible distinguishing gap due to $q=O(|\mathcal{K}|^{1/2}2^{-\log^2\nu})$. When the random oracle is not queried, the proof exactly follows Theorem 4 and the final distinguishing gape between O and $\mathcal{I}$ constructed there is negligible, due to $|\mathcal{M}|=O(|\mathcal{T}|^{1/2}2^{-\log^2\nu})$. ■

9 Conclusion

This paper studied the quantum message authentication model. We proposed a new model that captures several concerns. First, the attacker's state might be entangled with the authentication state. Second, the same seret key is desired to authenticate multiple messages. Third, it should address the threat from the verification queries. This is the concern that the attacker could create a forgery and see if the forgery is accepted or not. We showed that there exists a quantum authentication system that is secure when no verification queries are allowed while it is completely insecure when some verification queries are additionally admitted. We then proposed a framework by abstracting the protocol of Garg et al. [26] and showed that the framework is secure under our model as long as the component primitives satisfy certain properties. This reduces the design of the authentication protocol to the design of the primitives. There are some questions worth future study. Fehr and Salvail [24] studied the authentication with key reusability using the observation by Bennett, Brassard and Breidbart in 1982 [11] and notice that if a quantum message has not been changed in the transmission, then the attacker did not learn anything from it and so the key can be reused. It is interesting how we can capture this observation in an authentication model. In our model, we can not include it because the attacker's verification query can be the superposition of valid authentication message and invalid authentication message and also the verification result is a mixture (instead of simply a valid message or an invalid symbol). It is also interesting to study the composition security such as quantum universal composition from authenticated channel to unauthenticated channel, similar to its classical counterpart [14].

References

1. Gorjan Alagic and Christian Majenz, Quantum non-malleability and authentication, *CRYPTO'17*, 2017.
2. Andris Ambainis, Michele Mosca, Alain Tapp, Ronald de Wolf, Private Quantum Channels, *FOCS'00*, pp. 547-553, 2000.
3. Gorjan Alagic, Christian Majenz, Alexander Russell, Fang Song, Quantum-Access-Secure Message Authentication via Blind-Unforgeability, *EUROCRYPT'20*, pages 788-817, 2020.
4. Dorit Aharonov, Michael Ben-Or, and Elad Eban. Interactive proofs for quantum computations. *ICS'10*, Tsinghua University Press, 2010.
5. D. Aharonov and M. Ben-Or. Fault-tolerant quantum computation with constant error. *STOC'97*, pp. 176-188, 1997.
6. Mihir Bellare, Ran Canetti, Hugo Krawczyk: A Modular Approach to the Design and Analysis of Authentication and Key Exchange Protocols (Extended Abstract). STOC 1998: 419-428
7. Michael Ben-Or, Claude Crépeau, Daniel Gottesman, Avinatan Hassidim, Adam D. Smith, Secure Multiparty Quantum Computation with (Only) a Strict Honest Majority, *FOCS'06*, pp. 249-260, 2006.
8. Howard Barnum, Claude Crépeau, Daniel Gottesman, Adam Smith, and Alain Tapp. Authentication of quantum messages. *FOCS'02*, pp449-458, IEEE, 2002.
9. Dan Boneh and Mark Zhandry, Quantum-secure message authentication codes, *EUROCRYPT'13*, pp. 592-608, Springer, 2013.
10. Dan Boneh, Mark Zhandry, Secure Signatures and Chosen Ciphertext Security in a Quantum Computing World, *CRYPTO'13*, pages 361-379, 2013.
11. Charles H. Bennett, Gilles Brassard, Seth Breidbart, Quantum Cryptography II: How to re-use a one-time pad safely even if P=NP, *Nat. Comput.* 13(4), pp. 453-458, 2014.
12. C.H. Bennett, and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. *IEEE International Conference on Computers, Systems & Signal Processing*, pp. 175-179, 1984.
13. Ran Canetti, Hugo Krawczyk: Analysis of Key-Exchange Protocols and Their Use for Building Secure Channels. EUROCRYPT 2001: 453-474.
14. Ran Canetti, Hugo Krawczyk: Universally Composable Notions of Key Exchange and Secure Channels. EUROCRYPT 2002: 337-351
15. Fernando GSL Brandao, Aram W Harrow, and Michal Horodecki. Local random quantum circuits are approximate polynomial-designs, ArXiv:1208.0692, 2012.
16. Anne Broadbent, Gus Gutoski, and Douglas Stebila. Quantum one-time programs, *CRYPTO'13*, pp. 344-360, Springer, 2013.
17. Anne Broadbent, Evelyn Wainewright: Efficient Simulation for Quantum Message Authentication. *ICITS 2016*, pp. 72-91, 2016.
18. Ran Canetti. Universally composable security: A new paradigm for cryptographic protocols. *FOCS'01*, pages 136-145. IEEE, 2001.
19. Ran Canetti, Universally Composable Security, *J. ACM* 67(5): 28:1-28:94 (2020)
20. Jelle Don, Serge Fehr, Christian Majenz, Christian Schaffner, Online-Extractability in the Quantum Random-Oracle Model, *EUROCRYPT'22.*
21. I. Damgård, T. Brochmann Pedersen, L. Salvail, A quantum cipher with near optimal key-recycling. *CRYPTO'05*, LNCS 3621, pp. 494-510, 2005.
22. I. Damgård, T. Brochmann Pedersen, L. Salvail. How to re-use a one-time pad safely and almost optimally even if P =NP. *Natural Computing*,13(4), pages 469-486, 2014.
23. Frédéric Dupuis, Jesper Buus Nielsen, and Louis Salvail. Actively secure two-party evaluation of any quantum operation, *CRYPTO'12*, pp. 794-811, Springer, 2012.
24. Serge Fehr, Louis Salvail, Quantum Authentication and Encryption with Key Recycling - Or: How to Re-use a One-Time Pad Even if P=NP - Safely & Feasibly, *EUROCRYPT'17*, pages 311-338, 2017.
25. Oded Goldreich, *Foundations of Cryptography: Basic Applications*, Cambridge University Press, 2004.
26. Sumegha Garg, Henry Yuen, Mark Zhandry, New Security Notions and Feasibility Results for Authentication of Quantum Data, *CRYPTO'17*, pages 342-371, 2017.
27. Daniel Gottesman, Uncloneable encryption. *Quantum Inf. Comput.* 3(6), pages 581-602, 2003.
28. T. Haug, N. Bansal, W. K. Mok, D. E. Koh and K. Bharti, Pseudorandom quantum authentication, arXiv:2501.00951, 2025.
29. Patrick Hayden, Debbie Leung, and Dominic Mayers, The universal composable security of quantum message authentication with key recycling, *QCrypt 2011.* Available at arXiv: 1610.09434.

30. Zhengfeng Ji, Yi-Kai Liu and Fang Song, Pseudorandom Quantum States. *CRYPTO* (3) 2018, pp. 126-152, 2018.
31. Jonathan Katz and Yehuda Lindell, *Introduction to Modern Cryptography*, CRC Press, 2007.
32. Fermi Ma, Hsin-Yuan Huang, How to Construct Random Unitaries, IACR ePrint 2024/1652, 2024.
33. Ueli Maurer and Renato Renner. Abstract cryptography, *ICS 2011*, pp. 1-21. Tsinghua University Press, 2011.
34. Michael A. Nielsen and Isaac L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, New York, 2010.
35. Jonathan Oppenheim and Micha l Horodecki. How to reuse a one-time pad and other notes on authentication, encryption, and protection of quantum information. Physical Review A, 72:042309, October 2005.
36. Christopher Portmann, Quantum Authentication with Key Recycling, *EUROCRYPT'17*, pp. 339-368, 2017.
37. Reihaneh Safavi-Naini, Peter R. Wild: Information Theoretic Bounds on Authentication Systems in Query Model. IEEE Trans. Inf. Theory 54(6): 2426-2436 (2008)
38. Dominique Unruh, Revocable Quantum Timed-Release Encryption. *EUROCRYPT 2014*, pp. 129-146.
39. John Watrous, Quantum Computing, *Lecture notes*, 2006. Available at https://cs.uwaterloo.ca/ watrous/QC-notes/
40. John Watrous. An introduction to quantum information and quantum circuits. *ACM SIGACT News* 42(2): 52-67, 2011.
41. M. N. Wegman and J. L. Carter, "New hash functions and their use in authentication and set equality", Journal of computer and system sciences, vol.22, no. 3, pp. 265-279, June 1981.
42. Mark Zhandry, How to Record Quantum Queries, and Applications to Quantum Indifferentiability, *CRYPTO 2019*, part II, pages 239-268.
43. Mark Zhandry, Secure identity-based encryption in the quantum random oracle model, CRYPTO'12.
44. Mark Zhandry, how to construct quantum random functions, *FOCS'12*.
45. Mark Zhandry, A note on the quantum collision and set equality problems. *Quantum Inf. Comput.* 15(7&8): 557-567 (2015).

A Proof of Lemma 4

Lemma 4. *Assume that $(S_1, \cdots, S_K)$ is an equipartition of $\{0,1\}^n$ with $|S_i| = 2^n/K$ for all i. Let $H : \{0,1\}^n \to \{0,1\}^m$ be a quantum random oracle and $A^H(\mathbf{x}, \mathbf{y})$ be a quantum algorithm with input ($\mathbf{x}$, $\mathbf{y}$), where $\mathbf{x} \subseteq \{0,1\}^n$ of size $2^n/K$ and $y_i \in \{0,1\}^m$ and it makes at most q queries to H. Let B be a quantum algorithm with input $\mathbf{x} = S_t$ for some t: take $i \leftarrow \{1, \cdots, q\}$, $\mathbf{y} \leftarrow (\{0,1\}^m)^{2^n/K}$ and run $A^H(\mathbf{x}, \mathbf{y})$ till the i-th query to H in which case B measures the query in the computational basis and outputs the measurement result. If A makes less than i queries to H, B outputs $\perp$. Then, for $\mathbf{x} = S_k$ with $k \leftarrow [K]$, $\mathbf{y} \leftarrow (\{0,1\}^m)^{2^n/K}$ and a purely random H,*

$$|\Pr(A^H(\mathbf{x}, H(\mathbf{x})) = 1) - \Pr(A^H(\mathbf{x}, \mathbf{y}) = 1)| \leq 2q\sqrt{\delta}. \tag{77}$$

where $\delta = \Pr(x' \in \mathbf{x} : x' = B^H(\mathbf{x}))$ and probability is over the randomness of $\mathbf{xy}H$ and the uncertainty of B's final measurement.

Proof. Let the state of A consist of three quantum systems: X, Y, W for query, response and working registers. Let $|\Psi\rangle_{\mathbf{xy}}$ be the initial state prepared by A when it is given input $\mathbf{x}, \mathbf{y}$. Let $O_H : |x, y, w\rangle \mapsto |x, y+H(x), w\rangle$ be the quantum random oracle. For simplicity, let U be the unitary operator between two random oracle queries. Let $|\Psi^i_{\mathbf{xy},H}\rangle := (UO_H)^i|\Psi_{\mathbf{xy}}\rangle$. Then, the state of A before the final measurement is $|\Psi^q_{\mathbf{xy},H}\rangle$. Let $p_1 = \Pr(A^H(\mathbf{x}, H(\mathbf{x})) = 1)$ and $p_2 = \Pr(A^H(\mathbf{x}, \mathbf{y}) = 1)$. Then, p_2 is the probability of measurement on $|\Psi^q_{\mathbf{xy},H}\rangle$ with outcome 1 for fixed $\mathbf{xy}H$, followed by averaging over $\mathbf{xy}H$. Also, p_1 is the probability of the measurement on $|\Psi^q_{\mathbf{x}H(\mathbf{x}),H}\rangle$ with outcome 1 for fixed $\mathbf{x}H$, followed by averaging over $\mathbf{x}H$. Define $H_{\mathbf{xy}}$ to be a variant of H so that $H(\mathbf{x}) = \mathbf{y}$ while $H_{\mathbf{xy}}(x') = H(x')$ for $x' \notin \mathbf{x}$. Then, p_1 can be also described as the probability of the measurement on $|\Psi^q_{\mathbf{xy},H_{\mathbf{xy}}}\rangle$ with outcome 1 for fixed $\mathbf{xy}H$, followed by averaging over $\mathbf{xy}H$, as $H_{\mathbf{xy}}(\mathbf{x}) = \mathbf{y}$ and $H(\mathbf{x})$

is not used in this modification (so averaging with or without $H(\mathbf{x})$ does not change p_1). Therefore, $|p_1 - p_2|$ is the distinguishing gap between $\mathbf{E}_{\mathbf{xy}H}(|\Psi^q_{\mathbf{xy},H_{\mathbf{xy}}}\rangle\langle\Psi^q_{\mathbf{xy},H_{\mathbf{xy}}}|)$ and $\mathbf{E}_{\mathbf{xy}H}(|\Psi^q_{\mathbf{xy},H}\rangle\langle\Psi^q_{\mathbf{xy},H}|)$, which is bounded by their trace distance, due to [34, Theorem 9.1]. By triangle inequality,

$$|p_1 - p_2| \leq \mathbf{E}_{\mathbf{xy}H}(D_t(|\Psi^q_{\mathbf{xy},H}\rangle\langle\Psi^q_{\mathbf{xy},H}|, |\Psi^q_{\mathbf{xy},H_{\mathbf{xy}}}\rangle\langle\Psi^q_{\mathbf{xy},H_{\mathbf{xy}}}|)). \tag{78}$$

Let $|\Psi^{iq}_{\mathbf{xy},H}\rangle := (UO_{H_{\mathbf{xy}}})^{q-i}(UO_H)^i|\Psi_{\mathbf{xy}}\rangle$. Notice that $|\Psi^{0q}_{\mathbf{xy},H}\rangle = |\Psi^q_{\mathbf{xy},H_{\mathbf{xy}}}\rangle$ and $|\Psi^{qq}_{\mathbf{xy},H}\rangle = |\Psi^q_{\mathbf{xy},H}\rangle$.

By triangle inequality again, we have

$$\begin{aligned}
|p_1 - p_2| \leq & \mathbf{E}_{\mathbf{xy}H}\Big(\sum_{i=0}^{q-1} D_t(|\Psi^{iq}_{\mathbf{xy},H}\rangle\langle\Psi^{iq}_{\mathbf{xy},H}|, |\Psi^{i+1,q}_{\mathbf{xy},H}\rangle\langle\Psi^{i+1,q}_{\mathbf{xy},H}|)\Big) && (79)\\
= & \mathbf{E}_{\mathbf{xy}H}\Big(\sum_{i=0}^{q-1} D_t(UO_{H_{\mathbf{xy}}}|\Psi^{i}_{\mathbf{xy},H}\rangle\langle\Psi^{i}_{\mathbf{xy},H}|O_{H_{\mathbf{xy}}}U^\dagger, UO_H|\Psi^{i}_{\mathbf{xy},H}\rangle\langle\Psi^{i}_{\mathbf{xy},H}|O_H U^\dagger)\Big) && (80)\\
& \text{(applying unitarty operator } (UO_{H_{\mathbf{xy}}})^{q-i-1} \text{ does not change the trace distance)} && (81)\\
= & \mathbf{E}_{\mathbf{xy}H}\Big(\sum_{i=0}^{q-1} D_t(O_{H_{\mathbf{xy}}}|\Psi^{i}_{\mathbf{xy},H}\rangle\langle\Psi^{i}_{\mathbf{xy},H}|O_{H_{\mathbf{xy}}}, O_H|\Psi^{i}_{\mathbf{xy},H}\rangle\langle\Psi^{i}_{\mathbf{xy},H}|O_H)\Big) && (82)\\
& \text{(applying } U \text{ dos not change the trace distance)} && (83)
\end{aligned}$$

Let $O_{\mathbf{xy}}(x)$ be the function such that if $x = x_i$ then the output is y_i; otherwise, it outputs 0. Also let $\neg\mathbf{x} = \{0,1\}^n - \mathbf{x}$. Notice that $H = O_{\mathbf{x}H(\mathbf{x})} + O_{(\neg\mathbf{x})H(\neg\mathbf{x})}$ and $H_{\mathbf{xy}} = O_{\mathbf{xy}} + O_{(\neg\mathbf{x})H(\neg\mathbf{x})}$. To proceed, we need to the following claim.

Claim. [38, Lem. 11] *Let two states $|\psi_i\rangle = |\omega\rangle + |\psi'_i\rangle$ $(i = 0, 1)$ with $|\psi'_i\rangle$ orthogonal to $|\omega\rangle$. Then,*

$$D_t(|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|) \leq 2||\,|\psi'_0\rangle\,||.$$

From this claim, we know that

$$D_t(O_{H_{\mathbf{xy}}}|\Psi^{i}_{\mathbf{xy},H}\rangle\langle\Psi^{i}_{\mathbf{xy},H}|O_{H_{\mathbf{xy}}}, O_H|\Psi^{i}_{\mathbf{xy},H}\rangle\langle\Psi^{i}_{\mathbf{xy},H}|O_H) \leq 2||\mathbb{O}_{\mathbf{xy}}|\Psi^{i}_{\mathbf{xy},H}\rangle||, \tag{84}$$

where $\mathbb{O}_{\mathbf{xy}}$ is the quantum operator $|x,u\rangle \mapsto |x, u + O_{\mathbf{xy}}(x)\rangle$, followed by a projector $\sum_{u\in\mathbf{x}} |u\rangle\langle u|_X$. Hence, Eq. (82) is upper bounded by

$$\begin{aligned}
& \mathbf{E}_{\mathbf{xy}H}\Big(\sum_{i=0}^{q-1} 2||\mathbb{O}_{\mathbf{xy}}|\Psi^{i}_{\mathbf{xy},H}\rangle||\Big) && (85)\\
\leq & 2\sqrt{q\mathbf{E}_{\mathbf{xy}H}\Big(\sum_{i=0}^{q-1} ||\mathbb{O}_{\mathbf{xy}}|\Psi^{i}_{\mathbf{xy},H}\rangle||^2\Big)} = 2q\sqrt{\delta}, && (86)
\end{aligned}$$

where the last inequality is due to Cauchy-Schwarz inequality (using the fact: if $\mathbf{xy}H$ in $\mathbf{E}$ has the probability weight $\alpha_{\mathbf{xy}H}$, then $\sum_{i\mathbf{xy}H} \frac{\alpha_{\mathbf{xy}H}}{q} = 1$) and the final equality has used the fact that $\delta = \mathbf{E}_{\mathbf{xy}H}\Big(\frac{1}{q}\sum_{i=0}^{q-1} ||\mathbb{O}_{\mathbf{xy}}|\Psi^{i}_{\mathbf{xy},H}\rangle||^2\Big)$, which comes from the fact that the projector $\sum_{u\in\mathbf{x}} |u\rangle\langle u|_X$ commutes with unitary $|x,u\rangle \mapsto |x, u + O_{\mathbf{xy}}(x)\rangle$ and the norm is not changed if unitary $(|x,u\rangle \mapsto |x, u + O_{\mathbf{xy}}(x)\rangle)$ is applied finally. This completes our proof. □

B Proof of Lemma 5

Lemma 5. $\mathbf{E}_{g,k_2}(|| \, |\omega_{error}\rangle \, ||^2)$ *is negligible.*

Proof. Recall that we are considering the pure state case for ω_q' (i.e., the state before the final verification) and thus the state after the final verification in $\mathbf{G}_1$ is $|\omega_{\mathcal{I}}\rangle = |\omega_{ideal}\rangle + |\omega_{error}\rangle$. We need to show that $|||\omega_{error}\rangle||$ is small. We first assume attacker always makes the challenge signing query and will deal with the other case later. Then, in $\mathbf{G}_1$, we know from Eq. (34) that

$$|\omega_{\mathcal{I}}\rangle = \frac{1}{MT_1}\sum_{zu}\sqrt{\lambda_z}\sum_{ma}\alpha_{zm}(-1)^{(m+a,t+b)\cdot u}|a\rangle_Y|\psi_{uz}\rangle_Z, \tag{87}$$

where $t = g(m|r_q)$ and $b = g(a|r_q)$ with g being a purely random function. Further, $|\omega_{ideal}\rangle$ and $|\omega_{error}\rangle$ have the same expression, except with constraint $m = a$ and $m \neq a$ respectively.

Game $\mathbf{G}_2$. We modify $\mathbf{G}_1$ to $\mathbf{G}_2$ so that g appears in a superposition (instead of first sampling g). Then, $|\omega_{\mathcal{I}}(\mathbf{G}_1)\rangle, |\omega_{ideal}(\mathbf{G}_1)\rangle$ and $|\omega_{error}(\mathbf{G}_1)\rangle$ will become $|\omega_{\mathcal{I}}(\mathbf{G}_2)\rangle, |\omega_{ideal}(\mathbf{G}_2)\rangle$ and $|\omega_{error}(\mathbf{G}_2)\rangle$ respectively, where

$$|\omega_{error}(\mathbf{G}_2\rangle = \frac{1}{MT_1}\sum_{zg}\sqrt{\lambda_{zg}}\sum_{uma,m\neq a}\alpha_{zmg}(-1)^{(m+a,t+b)\cdot u}|a\rangle_Y|\varphi_{uzg}\rangle_Z|g\rangle_D, \tag{88}$$

where $t = g(m|r_q)$ and $b = g(a|r_q)$ and the distribution of g is absorbed into λ_{zg}. Further, $|\omega_{ideal}(\mathbf{G}_2\rangle$ and $|\omega_{\mathcal{I}}(\mathbf{G}_2\rangle$ are modified accordingly. Clearly, $\mathbf{E}_{g,k_2}(|||\omega_{error}\rangle(\mathbf{G}_1)||^2) = \mathbf{E}_{k_2}(|| \, |\omega_{error}(\mathbf{G}_2)\rangle \, ||^2)$ as $|g\rangle_D$ is an orthonormal basis for the random oracle.

Game $\mathbf{G}_3$. We modify $\mathbf{G}_2$ to $\mathbf{G}_3$ so that the random oracle is replaced by compressed random oracle *CStO*. We observe that the inefficient representation of *CStO* and *StO* differs only by an unitary F_D (by recalling that, given input $|x\rangle_X|y\rangle_Y$, in *CStO*, before applying CNOT_{YD_x}, it first applies F_{D_x} to D_x register, while *StO* directly applies CNOT_{YD_x}; after the operation in *CStO*, F_{D_x} is applied again to get back to their *CStO* representation). It is clear that applying a unitary to D in Eq. (88) will not change its squared norm. Hence, $|| \, |\omega_{error}(\mathbf{G}_3)\rangle \, ||^2 = || \, |\omega_{error}(\mathbf{G}_2)\rangle \, ||^2$.

Let the global joint state right before verifying tag on register T_1 be denoted by $|\omega''_{\mathcal{I}}(\mathbf{G}_3)\rangle$. Let $|\omega''_{\mathcal{I}}(\mathbf{G}_3)\rangle = |\omega''_{ideal}(\mathbf{G}_3)\rangle + |\omega''_{error}(\mathbf{G}_3)\rangle$, where

$$|\omega''_{ideal}(\mathbf{G}_3)\rangle = \frac{1}{MT_1}\sum_{z\mathbf{y}}\sqrt{\lambda_{z\mathbf{y}}}\sum_{umab:m=a}\alpha_{zm\mathbf{y}}(-1)^{(m+a,t+b)\cdot u}|a,b\rangle_M|\varphi_{uz\mathbf{y}}\rangle_Z F_{D_{m|r_q}}|\mathbf{y}\rangle_D,$$

$$|\omega''_{error}(\mathbf{G}_3)\rangle = \frac{1}{MT_1}\sum_{z\mathbf{y}}\sqrt{\lambda_{z\mathbf{y}}}\sum_{umab:m\neq a}\alpha_{zm\mathbf{y}}(-1)^{(m+a,t+b)\cdot u}|a,b\rangle_M|\varphi_{uz\mathbf{y}}\rangle_Z F_{D_{m|r_q}}|\mathbf{y}\rangle_D,$$

where $y_{m|r_q} = t$ and $D_{m|r_q}$ is represented under a different basis for convenience. The verification of $|\omega''_{error}(\mathbf{G}_3)\rangle$ will give to $|\omega_{error}(\mathbf{G}_3)\rangle$. Since attacker only made q authentication /verification queries, $|\mathbb{X}(\mathbf{y})|$ has at most q elements and it holds similarly for the superposition entry $|\mathbf{y}'\rangle$ in $F_{D_{m|r_q}}|\mathbf{y}\rangle$. For $m = a$ case, $m|r_q$ was queried to the random oracle. So if $b \neq t$, then it will be verified to be invalid and hence removed. For $a \neq m$ case, $a|r_q$ was not recorded in the database that contains $m|r_q$. So $a|r_q \notin \mathbb{X}(\mathbf{y})$. Thus, the verification will make query $a|r_q$ to the random oracle and insert it into the database. This results in

$$|\omega_{error}(\mathbf{G}_3)\rangle = \frac{1}{MT_1^{3/2}}\sum_{z\mathbf{y}utmab:m\neq a}\sqrt{\lambda_{z\mathbf{y}t}}\alpha_{zm\mathbf{y}t}(-1)^{(m+a,t+b)\cdot u}|a\rangle|\varphi_{uz\mathbf{y}t}\rangle F_{D_{m|r_q,a|r_q}}|\mathbf{y}\cup(t)_{m|r_q}\cup(b)_{a|r_q}\rangle,$$

where we separate $(t)_{m|r_q}$ from $\mathbf{y}$ for the convenience of the calculation below.

In the following, we upper bound $|| \, |\omega_{error}(\mathbf{G}_3)\rangle \, ||^2$ which will turn out to be small. Consider the inner product of

$$|a\rangle F_{D_{m|r_q,a|r_q}} |\mathbf{y} \cup (t)_{m|r_q} \cup (b)_{a|r_q}\rangle \text{ and } |a'\rangle F_{D_{m'|r_q,a'|r_q}} |\mathbf{y}' \cup (t')_{m'|r_q} \cup (b')_{a'|r_q}\rangle.$$

The result has several cases:

- If $(a,b) \neq (a',b')$ or $\mathbf{y} \neq \mathbf{y}'$, the inner product is 0.
- If $(a,b) = (a',b')$ and $\mathbf{y} = \mathbf{y}'$ but $m' \neq m$, the inner product is $1/T_1$.
- If $(a,b) = (a',b')$, $\mathbf{y} = \mathbf{y}'$ and $m' = m$ but $t \neq t'$, the inner product is 0.
- If $(a,b) = (a',b')$ and $\mathbf{y} = \mathbf{y}'$ and $(m,t) = (m',t')$, the inner product is 1.

Therefore, we can write $|| \, |\omega_{error}(\mathbf{G}_3)\rangle \, ||^2 = \frac{1}{M^2 T_1^3}(A_0 + A_1)$, where

$$A_0 = \sum_{uu'zz'\mathbf{y}mm'att':|\{a,m,m'\}|=3} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t'}}\alpha_{zm\mathbf{y}t}\bar{\alpha}_{z'm'\mathbf{y}t'} \sum_b (-1)^{(m+a,t+b)\cdot u+(m'+a,t'+b)\cdot u'} \langle\varphi_{uz\mathbf{y}t}|\varphi_{u'z'\mathbf{y}t'}\rangle/T_1, \tag{89}$$

$$A_1 = \sum_{uu'zz'\mathbf{y}mat:a\neq m} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t}}\alpha_{zm\mathbf{y}t}\bar{\alpha}_{z'm\mathbf{y}t} \sum_b (-1)^{(m+a,t+b)\cdot(u+u')} \langle\varphi_{uz\mathbf{y}t}|\varphi_{u'z'\mathbf{y}t}\rangle. \tag{90}$$

Consider A_0 first. Write $u = (u_1, u_2)$ and $u' = (u'_1, u'_2)$. Consider

$$f_{amm'uu'} \overset{def}{=} \sum_b (-1)^{(a+m,b+t)\cdot(u_1,u_2)+(a+m',b+t')\cdot(u'_1,u'_2)} \tag{91}$$

$$= (-1)^{(a+m,t)\cdot(u_1,u_2)+(a+m',t')\cdot(u'_1,u'_2)} \sum_b (-1)^{b\cdot(u_2+u'_2)}. \tag{92}$$

So if $u_2 \neq u'_2$, then $f_{amm'uu'} = 0$. Thus, Eq. (89) can be simplified as

$$A_0 = \sum_{uu'_1zz'\mathbf{y}mm'att':|\{a,m,m'\}|=3} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t'}}\alpha_{zm\mathbf{y}t}\bar{\alpha}_{z'm'\mathbf{y}t'} (-1)^{(m+a,t)\cdot(u_1,u_2)+(m'+a,t')\cdot(u'_1,u_2)} \langle\varphi_{uz\mathbf{y}t}|\varphi_{u'z'\mathbf{y}t'}\rangle. \tag{93}$$

Further, given $m \neq m'$, let

$$f'_{mm'uu'_1} := \sum_{a\notin\{m,m'\}} (-1)^{(a+m,t)\cdot(u_1,u_2)+(a+m',t')\cdot(u'_1,u_2)} \tag{94}$$

$$= (-1)^{(m,t)\cdot(u_1,u_2)+(m',t')\cdot(u'_1,u_2)} \sum_{a\notin\{m,m'\}} (-1)^{a\cdot(u_1+u'_1)}. \tag{95}$$

If $u_1 = u'_1$, then $f'_{mm'uu'_1} = (M-2)(-1)^{(m+m',t+t')\cdot(u_1,u_2)}$.

If $u_1 \neq u'_1$, then

$$f'_{mm'uu'_1} = -(-1)^{(m,t)\cdot(u_1,u_2)+(m',t')\cdot(u'_1,u_2)} \sum_{a\in\{m,m'\}} (-1)^{a\cdot(u_1+u'_1)}. \tag{96}$$

Therefore, we can write $A_0 = \varDelta_0 + \varDelta_1$, where $\varDelta_0$ (resp. $\varDelta_1$) corresponds to the summation in A_0 with $u_1 = u_1'$ (resp. $u_1 \neq u_1'$). Notice that

$$\begin{aligned}
|\varDelta_0| &\leq (M-2) \sum_{uz z' m m' \mathbf{y} t t'} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t'}} |\alpha_{zm\mathbf{y}t}\bar{\alpha}_{z'm'\mathbf{y}t'} \langle \varphi_{uz\mathbf{y}t} | \varphi_{uz'\mathbf{y}t'} \rangle| && (97)\\
&\leq MT_1(M-2) \sum_{z z' m m' \mathbf{y} t t'} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t'}} |\alpha_{zm\mathbf{y}t}\bar{\alpha}_{z'm'\mathbf{y}t'}| && (98)\\
&\leq M^2 T_1(M-2) \sum_{z z' \mathbf{y} t t'} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t'}} = M^2 T_1 (M-2) \sum_{\mathbf{y}} (\sum_{zt} \sqrt{\lambda_{z\mathbf{y}t}})^2 && (99)\\
&\leq M^2 T_1 (M-2) \sum_{\mathbf{y}} MT_1 \sum_{zt} \lambda_{z\mathbf{y}t} \leq M^3 T_1^2 (M-2). && (100)
\end{aligned}$$

where inequality at Eq. (99) follows from Chauchy-Schwarz inequality on α for label m', noticing that $\sum_{m'} |\alpha_{zm'\mathbf{y}t'}|^2 = 1$, and the final inequality uses $\sum_{z\mathbf{y}t} \lambda_{z\mathbf{y}t} = 1$. The second last inequality uses the fact that the dimension from label z in Eq. (31) is at most M.

Now let us focus on $\varDelta_1$. Applying Eq. (96) to Eq. (93) with restriction to $u_1' \neq u_1$, we have

$$\begin{aligned}
|\varDelta_1| \leq 2 & \sum_{uu_1' zz' \mathbf{y} tt' mm' :\ u_1 \neq u_1', m \neq m'} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t'}} |\alpha_{zm\mathbf{y}t}\bar{\alpha}_{z'm'\mathbf{y}t'} \langle \varphi_{uz\mathbf{y}t} | \varphi_{u'z'\mathbf{y}t'} \rangle| && (101)\\
\leq & 2M^2 T_1 \sum_{zz'\mathbf{y}tt'mm' :\ m \neq m'} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t'}} |\alpha_{zm\mathbf{y}t}\bar{\alpha}_{z'm'\mathbf{y}t'}| && (102)\\
\leq & 2M^3 T_1 \sum_{zz'\mathbf{y}tt'} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t'}} \qquad \text{/* by Cauchy-Schwarz ineq. and } \textstyle\sum_m |\alpha_{zm\mathbf{y}t}|^2 = 1 \text{ */} && (103)\\
\leq & 2M^3 T_1 \sum_{\mathbf{y}} MT_1 \sum_{z} \lambda_{z\mathbf{y}t} \leq 2M^4 T_1^2, && (104)
\end{aligned}$$

where the last inequality follows similarly as $\varDelta_0$. Therefore,

$$|A_0| \leq |\varDelta_0| + |\varDelta_1| \leq 3M^4 T_1^2. \tag{105}$$

Now let us consider A_1 in Eq. (90). Again, if $u_2 \neq u_2'$, then the contribution of summation over b will give 0 (similar to the case in A_0). Thus, A_1 can be simplified as

$$A_1 = T_1 \sum_{uu_1' zz' \mathbf{y} mat :\ m \neq a} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t}} \alpha_{zm\mathbf{y}t}\bar{\alpha}_{z'm\mathbf{y}t} (-1)^{(m+a)\cdot(u_1+u_1')} \langle \varphi_{uz\mathbf{y}t} | \varphi_{u'z'\mathbf{y}t} \rangle. \tag{106}$$

Again, this has two cases $u_1 = u_1'$ and $u_1 \neq u_1'$. Write $A_1 = \eta_0 + \eta_1$, where η_0 corresponds to $u_1 = u_1'$ case and η_1 corresponds to $u_1 \neq u_1'$ case. Thus,

$$\begin{aligned}
\eta_0 \leq & T_1^2 M(M-1) \sum_{zz'\mathbf{y}mt} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t}} |\alpha_{zm\mathbf{y}t}\bar{\alpha}_{z'm\mathbf{y}t}| && (107)\\
\leq & T_1^2 M(M-1) \sum_{zz'\mathbf{y}t} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t}} \qquad \text{/* Cauchy-Schwarz ineq. \& } \textstyle\sum_m |\alpha_{zm\mathbf{y}t}|^2 = 1 \text{ */} && (108)\\
= & T_1^2 M(M-1) \sum_{\mathbf{y}t} (\sum_{z} \sqrt{\lambda_{z\mathbf{y}t}})^2 \leq T_1^2 M^2 (M-1) \sum_{\mathbf{y}t} \sum_{z} \lambda_{z\mathbf{y}t} && (109)\\
= & T_1^2 M^2 (M-1). && (110)
\end{aligned}$$

Then, we consider η_1. Notice if no restriction of $m \neq a$, then η_1 should be zero. Thus,

$$\begin{aligned} |\eta_1| \leq & T_1 \sum_{uu_1' zz'\mathbf{y}tm} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t}} |\alpha_{zm\mathbf{y}t}\bar{\alpha}_{z'm\mathbf{y}t}\langle \varphi_{uz\mathbf{y}t}|\varphi_{u'z'\mathbf{y}t}\rangle| & (111) \\ \leq & T_1^2 M^2 \sum_{zz'\mathbf{y}tm} \sqrt{\lambda_{z\mathbf{y}t}\lambda_{z'\mathbf{y}t}} |\alpha_{zm\mathbf{y}t}\bar{\alpha}_{z'm\mathbf{y}t}| & (112) \\ \leq & T_1^2 M^2 \cdot M = T_1^2 M^3, & (113) \end{aligned}$$

Combining the bounds for A_0 and A_1, we have

$$|| \, |\omega_{error}(\mathbf{G}_3)\rangle \, ||^2 \leq \frac{3T_1^2 M^4 + 2M^3 T_1^2}{M^2 T_1^3} \leq \frac{5M^2}{T_1}. \tag{114}$$

This bound is for the case where the $q-1$ query is a challenge authentication query. We now consider the case where no challenge authentication query was queried. In this case, after g replaced by random function and further replaced by the standard quantum random oracle and finall the compressed random oracle. It suffices to show that $\mathbf{E}_{k_2}(|||\omega_{\mathcal{I}}(\mathbf{G}_3\rangle||^2)$ is negligible (recall that $|\omega_{ideal}(\mathbf{G}_3)\rangle = 0$). Further, we assume that after Ver_2 and Fourier inverse transform, the state (prior to Ver_1) becomes

$$\sum_{mt,\mathbf{y}} \alpha_{mt\mathbf{y}} |m,t\rangle |\mathbf{y}\rangle_D, \tag{115}$$

where $\sum_{mt\mathbf{y}} ||\alpha_{mt\mathbf{y}}||^2 \leq 1$. Now since $m|r_q$ was not queried to the random oracle, it follows that $m|r_q \notin \mathbb{X}(\mathbf{y})$. Thus, after the verification of Ver_1, it will become

$$|\omega_{\mathcal{I}}(\mathbf{G}_3)\rangle = \frac{1}{T_1^{1/2}} \sum_{mt\mathbf{y}} \alpha_{mt\mathbf{y}} |m,t\rangle F_{D_{m|r_q}} |\mathbf{y} \cup (t)_{m|r_q}\rangle_D, . \tag{116}$$

This gives $\mathbf{E}_{k_2}(|| \, |\omega_{\mathcal{I}}(\mathbf{G}_3)\rangle \, ||^2) \leq 1/T_1$. This completes the proof of the lemma. ■