

# Quantum One-Time Protection of any Randomized Algorithm

Sam Gunn and Ramis Movassagh

Google Quantum AI, Venice CA, 90291

November 6, 2024

## Abstract

The meteoric rise in power and popularity of machine learning models dependent on valuable training data has reignited a basic tension between the power of running a program locally and the risk of exposing details of that program to the user. At the same time, fundamental properties of quantum states offer new solutions to data and program security that can require strikingly few quantum resources to exploit, and offer advantages outside of mere computational run time. In this work, we demonstrate such a solution with *quantum one-time tokens*.

A quantum one-time token is a quantum state that permits a certain program to be evaluated *exactly once*. One-time security guarantees, roughly, that the token cannot be used to evaluate the program more than once. We propose a scheme for building quantum one-time tokens for any randomized classical program, which include generative AI models. We prove that the scheme satisfies an interesting definition of one-time security *as long as outputs of the classical algorithm have high enough min-entropy*, in a black box model.

Importantly, the classical program being protected does not need to be implemented coherently on a quantum computer. In fact, the size and complexity of the quantum one-time token is *independent* of the program being protected, and additional quantum resources serve only to increase the security of the protocol. Due to this flexibility in adjusting the security, we believe that our proposal is parsimonious enough to serve as a promising candidate for a near-term useful demonstration of quantum computing in either the NISQ or early fault tolerant regime.

## 1 Introduction

Commercializing software presents a central dilemma: How does a proprietor distribute software, without forfeiting ownership? On the one hand, software must be made available to users for them to use it; on the other, once the software is distributed, an unlimited number of unauthorized copies can be made. This problem is more acute in the age of generative AI, where the software can be extremely valuable and potentially reveal private information.

There are two widely-adopted solutions to this problem:

1. The proprietor can distribute an obfuscated version of the software. However, this opens up the possibility of piracy, and the user can always run the software an unlimited number of times.
2. The proprietor can allow queries to the software, instead of distributing the software itself. But this requires communication between the user and the proprietor every time the software is used.

Sometimes a combination of these solutions is employed — for instance, an internet browser might be obfuscated and made public while the search engine itself is kept on servers.

There is an apparent trade-off between usability and exclusivity in these solutions. Distributing obfuscated software makes it highly usable, but it is impossible to impose any limits on the number of times it can be used. Keeping the software on servers and responding to user queries has high exclusivity, but lower usability

because the user needs to communicate with the server to perform computations. Furthermore, this solution is only available to proprietors with enough resources to host a reliable server.

This trade-off between usability and exclusivity is inherent in the classical setting, because classical information can always be copied. As soon as software is distributed, the user can query or copy it as many times as they want. Therefore it is natural to look to quantum mechanics for improved solutions. In contrast to classical information, quantum information cannot be cloned. Indeed, unclonable cryptography is able to make use of this principle to improve both of the above solutions.<sup>1</sup>

**Improving Solution (1) with copy protection.** In copy protection, introduced by [Aar09], a quantum state is created that allows a program to be run an unlimited number of times, but not copied. In a classical oracle model, quantum copy protection is possible for any unlearnable program [ALL<sup>+</sup>21]. In the standard model, there exist unlearnable programs that cannot be copy protected [AP21]; nonetheless, it is known how to copy protect a small number of particular functionalities in the standard model [CLLZ21, LLQZ22, CG24, AB24].

Quantum copy protection addresses the piracy concern in Solution (1), but it does not apply in the setting where the proprietor wishes to distribute *individual queries* to the software. This is particularly relevant in the current age of generative AI, where the pay-per-query model is prevalent.

**Improving Solution (2) with one-time programs.** For the pay-per-query model, one-time programs are a more suitable solution. A one-time program token is a quantum state that enables a program to be evaluated *once*. One-time programs therefore remove the need for online communication in Solution (2): The proprietor distributes tokens, and the users consume the tokens offline at their leisure.

**This work.** In this work, we propose the first general-purpose tokenized program scheme using quantum information. Prior work was restricted to protecting specific cryptographic functionalities, and in particular did not apply to non-cryptographic programs like generative AI models.

Our method uses a one-time signature scheme and a program obfuscator to compile any randomized algorithm into a one-time token that allows the algorithm to be executed on any input of the user’s choice. We note that the quantum capabilities required for our scheme are *completely independent of the program being protected*. Therefore, our scheme could plausibly be used to protect a large program using only a small, noisy quantum device — even one that is incapable of demonstrating quantum computational advantage!

**Prior and concurrent work on one-time programs.** One-time programs were initially studied in [GKR08], but without quantum computation. The idea of quantum one-time programs was originally suggested in [BGS13], where it was shown that it is impossible to build quantum one-time tokens for deterministic programs without specialized hardware assumptions. This is because a simple rewinding attack would allow a user to make arbitrarily many queries to the algorithm, given any state that allows it to be run once. Later, [BS23] presented a scheme to one-time protect signature functions in the standard model.

In concurrent and independent work, [GLR<sup>+</sup>24] also present a scheme for the one-time protection of arbitrary randomized algorithms. Their scheme is essentially identical to ours, except that they instantiate the quantum one-time authentication scheme with the particular construction of [CLLZ21]. With respect to security definitions and proofs, their results are generally much more extensive, including stronger (albeit more complex) definitions of one-time security and new impossibility results. However, they take a very different and complementary approach to proving security. It would be interesting to know whether our results (and in particular Lemma 1, which is simple and self-contained) say anything about security in their setting, or whether their results imply our Lemma 1.

---

<sup>1</sup>We note that unclonable cryptography has led to many additional interesting and varied protocols, including quantum money [Wie83], quantum key distribution [BB14], and certified deletion [BI20].

**Removing quantum communication.** The protocol described below requires quantum communication between the proprietor and the user, because the proprietor needs to send the user the quantum one-time program token. However, if the one-time signatures are instantiated with those from [CLLZ21], then the results of [Shm22, CHV23] allow us to replace this quantum communication with a remote state preparation protocol that uses only classical communication. That is, there exists a polynomial-time interactive protocol between a quantum user and a *classical* proprietor that results in the user holding a quantum one-time token which they can only use once. Unfortunately, this protocol is highly complex and not likely to be feasible on a near-term quantum device.<sup>2</sup>

## 1.1 Acknowledgements

We thank Fermi Ma and Ryan Babbush for helpful discussions. We thank Jarrod McClean for detailed comments on an earlier draft of this work.

## 2 The construction

In this section, we present our scheme for one-time protecting an arbitrary randomized algorithm. Any randomized algorithm can be described by a function  $f$ , which takes the input together with a random string, and outputs the result of the computation.

Our scheme relies on three cryptographic primitives:

- (AuthKeyGen, AuthTokenGen, Sign, Verify), a quantum one-time authentication scheme. For instance, any quantum one-time signature scheme will suffice — although it is not required to be publicly verifiable. We define a quantum one-time authentication scheme in Definition 1.
- Obf, a classical program obfuscator. This can be heuristically instantiated with any classical obfuscation algorithm (for instance, a candidate indistinguishability obfuscator), but we model it as black-box obfuscation for our security proof.
- $H$ , a hash function modeled as a random oracle. This can be replaced with a pseudorandom function without affecting the security proof, because we model Obf as black-box obfuscation.

The software proprietor will generate a program token which can be used to evaluate  $f$  one time. A program token consists of two parts:

- $|\psi\rangle$ , an unclonable quantum state that does not depend on  $f$ .
- Obf( $P$ ), an obfuscated classical circuit  $P$  that depends on  $f$ . This part is completely classical.

We emphasize that  $|\psi\rangle$  *does not depend on the program being obfuscated*. In particular, this means that the quantum capabilities required of the user and the software proprietor do not depend on the complexity of the computation being one-time protected. Highly-complex computations like evaluation of a generative AI model can be protected, even if the quantum capabilities of both the user and software proprietor are limited. Furthermore, if we use the one-time signature scheme from [CLLZ21] for one-time authentication, then the only quantum capabilities required of the user are to store a constant-sized quantum state and to measure it in the standard and Hadamard bases.

At a high level, our scheme works by requiring the user to one-time authenticate any input they wish to evaluate  $f$  on. The program  $P$  will not evaluate  $f$  unless a valid input-authentication pair is provided. Crucially,  $P$  will determine the randomness to use for  $f$  by computing a hash of the input-authentication pair.

---

<sup>2</sup>It also requires the assumptions that LWE is sub-exponentially hard for quantum computers and that indistinguishability obfuscation for classical circuits exists with sub-exponential security against quantum polynomial-time adversaries.

## 2.1 One-time authentication

Before we state our one-time program construction in Construction 2, let us define quantum one-time authentication. This is a secret-key version of the definition of a quantum one-time signature scheme from [BS23]. In the following definition, the notation  $\mathcal{A}^{\text{Verify}(\text{sk}, \cdot)}$  means that  $\mathcal{A}$  is given quantum query access to the function  $(x, z) \mapsto \text{Verify}(\text{sk}, (x, z))$ .

**Definition 1** (Quantum one-time authentication). We say that a tuple of four polynomial-time algorithms  $(\text{AuthKeyGen}, \text{AuthTokenGen}, \text{Sign}, \text{Verify})$  is a *quantum one-time authentication scheme* if the following hold:<sup>3</sup>

- (Correctness)  $\text{Verify}(\text{sk}, (x, z)) = 1$  if  $z \leftarrow \text{Sign}(x, |\text{tk}\rangle)$ ,  $|\text{tk}\rangle \leftarrow \text{TokenGen}(\text{sk})$ , and  $\text{sk} \leftarrow \text{AuthKeyGen}(1^\lambda)$ .
- (One-time unforgeability) For any polynomial-time adversary  $\mathcal{A}$ ,

$$\Pr_{\substack{(x_1, z_1, x_2, z_2) \leftarrow \mathcal{A}^{\text{Verify}(\text{sk}, \cdot)}(|\text{tk}\rangle) \\ |\text{tk}\rangle \leftarrow \text{AuthTokenGen}(\text{sk}) \\ \text{sk} \leftarrow \text{AuthKeyGen}(1^\lambda)}} [\text{Verify}(\text{sk}, (x_1, z_1)) = \text{Verify}(\text{sk}, (x_2, z_2)) = 1 \text{ and } x_1 \neq x_2] = \text{negl}(\lambda).$$

Since one-time authentication schemes can be easily built from any one-time signature scheme by simply considering the public key to be part of the secret key, the results of [BS23, CLLZ21] imply their existence. In fact, [BS23, CLLZ21] show the existence of *information-theoretically secure* quantum one-time authentication schemes — that is, these constructions are secure against computationally unbounded adversaries, as long as they are only allowed to make  $\text{poly}(\lambda)$  queries to  $\text{Verify}$ . We describe one such scheme for single-bit messages now. The case for multi-bit messages works by simply using several single-bit schemes in parallel.

**Construction 1** (Single-bit quantum one-time authentication from hidden subspace states, [BS23]). The algorithms are defined as follows.

$\text{AuthKeyGen}(1^\lambda)$

1. Sample a uniformly random subspace  $A \subseteq \mathbb{F}_2^\lambda$  of dimension  $\lambda/2$ .
2. Output  $A$ .

$\text{AuthTokenGen}(A)$

1. Output  $|A\rangle = 2^{-\lambda/4} \sum_{a \in A} |a\rangle$ .

$\text{Sign}(x, |A\rangle)$

1. If  $x = 0$ , measure  $A$  in the standard basis and output the result.
2. If  $x = 1$ , measure  $A$  in the Hadamard basis and output the result.

$\text{Verify}(A, (x, z))$

1. If  $z = 0^\lambda$ , reject (output  $\perp$ ).
2. If  $x = 0$  and  $z \in A$ , accept (output 1).
3. If  $x = 1$  and  $z \in A^\perp$ , accept (output 1).
4. Otherwise, reject (output  $\perp$ ).

---

<sup>3</sup>The notation  $1^\lambda$  just means  $\lambda$  repeated 1's; we sometimes use  $1^\lambda$  instead of  $\lambda$  as the input to an algorithm because we want the algorithm to run in time that is polynomial in the size of the input.

Note that Construction 1 is correct because

$$H^\lambda |A\rangle = 2^{-\lambda/4} \sum_{\substack{b \in \mathbb{F}_2^n \\ b \cdot a = 0 \ \forall a \in A}} |b\rangle = |A^\perp\rangle.$$

Security was proven in [BS23].

**Theorem 1** (Adapted from Theorem 16 of [BS23]). *Construction 1 is an information-theoretically secure quantum one-time authentication scheme.*

## 2.2 Our one-time program construction

Let  $f$  be the program we wish to one-time protect. Our scheme works by publishing a program  $\hat{P}$  which contains within it the description of  $f$ . Crucially, this program will have the following properties:

- $\hat{P}$  uses obfuscation so as not to reveal the code of  $f$ ;
- $\hat{P}$  will compute  $f$  on any input, but only if the input is authenticated with a one-time authentication scheme; and
- $\hat{P}$  suitably “mixes” the input, the authentication tag, and the output together in such a way that, if the output is measured, the input and authentication tag effectively collapse.

While the first property follows from our use of an oracle model, and the second property is by construction, the third will prove to be quite difficult to establish. The statement of this third property is formalized in Lemma 1, which allows us to prove the one-time security of our scheme.

**Construction 2** (General-purpose one-time programs). Let  $f : \mathcal{X} \times \mathcal{R} \rightarrow \mathcal{Y}$  be the function to be one-time protected, described as a bit-string representing a classical circuit that computes  $f$ . Let  $\text{Obf}$  be an obfuscation algorithm for classical circuits, and let  $(\text{AuthKeyGen}, \text{AuthTokenGen}, \text{Sign}, \text{Verify})$  be a quantum one-time authentication scheme as defined in Definition 1. Suppose that the one-time authentication scheme produces authentications of length  $m$  for all messages  $x \in \mathcal{X}$ , and let  $H : \mathcal{X} \times \{0, 1\}^m \rightarrow \mathcal{R}$  be a random oracle.

Our one-time protection scheme is specified by the algorithms  $\text{KeyGen}$ ,  $\text{TokenGen}$ , and  $\text{TokenEval}$ , defined as follows.

$\text{KeyGen}(1^\lambda, f)$

1. Sample  $\text{sk} \leftarrow \text{AuthKeyGen}(1^\lambda)$ .
2. Let  $P : \mathcal{X} \times \{0, 1\}^m \rightarrow \mathcal{Y} \cup \{\perp\}$  be a classical circuit that does the following on input  $(x, z)$ :
  - (a) Compute  $\text{Verify}(\text{sk}, (x, z))$ . If it rejects, output the special “reject” symbol  $\perp$ .
  - (b) Otherwise, output  $f(x; H(x, z))$ .
3. Compute the obfuscation  $\hat{P} = \text{Obf}(P)$  of  $P$ .
4. Output  $(\text{sk}, \hat{P})$ .

$\text{TokenGen}(\text{sk})$

1. Compute the one-time authentication token  $|\text{tk}\rangle \leftarrow \text{AuthTokenGen}(\text{sk})$ .
2. Output  $|\text{tk}\rangle$  as the one-time program token.

$\text{TokenEval}(x, |\text{tk}\rangle, \hat{P})$

1. Compute  $z \leftarrow \text{Sign}(x, |\text{tk}\rangle)$ .
2. Compute  $\hat{P}(x, z)$  and output the result.

After **KeyGen** is run, the obfuscated program  $\hat{P} = \text{Obf}(P)$  is published as a public key. Given a one-time program token  $|\text{tk}\rangle$ , together with this public key, a user can evaluate  $f$  on any input  $x$  of their choice.

### 3 Security

We formalize one-time security of our one-time program scheme with the following *black-box one-time program* game. Essentially, security says that once an adversary produces a measured output of  $\hat{P}$ , the adversary cannot find new accepting inputs to  $\hat{P}$ .

$\mathcal{G}_{\text{BB-OTP}}(|\psi\rangle, \hat{P})$ :

1. The adversary is given  $|\psi\rangle$ . The adversary is also given (quantum) oracle access to  $\hat{P}$ . The adversary is allowed to access this oracle throughout the game.
2. The adversary submits a quantum query to the challenger on register  $Q$ . The challenger does the following:
  - (a) Compute  $\hat{P}$  on  $Q$ , placing the result of the computation onto a register  $R$ .
  - (b) Measure  $R$ , obtaining outcome  $y$ .
  - (c) If  $y = \perp$ , the game is aborted and the adversary loses.
  - (d) Otherwise, return  $Q$  to the adversary.
3. The adversary submits a quantum query to the challenger on register  $Q$ . The challenger does the following:
  - (a) Compute  $\hat{P}$  on  $Q$ , placing the result of the computation onto a register  $R$ .
  - (b) Measure  $R$ , obtaining outcome  $y'$ .
4. The adversary wins if  $y' \notin \{y, \perp\}$ .

We say that a scheme is black-box one-time secure if no polynomial-time adversary can win  $\mathcal{G}_{\text{BB-OTP}}$  with inverse polynomial probability.

**Definition 2** (One-time program). A one-time program scheme consists of three polynomial-time algorithms **KeyGen**, **TokenGen**, and **TokenEval**. We say that it is *black-box one-time secure* for a function  $f$  if, for all polynomial-time adversaries  $\mathcal{A}$  making at most  $\text{poly}(\lambda)$  quantum oracle queries,

$$\Pr_{\substack{(\text{sk}, \hat{P}) \leftarrow \text{KeyGen}(1^\lambda, f) \\ |\text{tk}\rangle \leftarrow \text{TokenGen}(\text{sk})}} [\mathcal{A}^{\hat{P}} \text{ wins } \mathcal{G}_{\text{BB-OTP}}(|\text{tk}\rangle, \hat{P})] = \text{negl}(\lambda).$$

We do not believe that this is the final say in definitions of one-time security, but we believe it is a useful starting point. The rest of this paper is devoted to proving the following theorem.

**Theorem 2** (Construction 2 is one-time secure). *Suppose that, for every  $x \in \mathcal{X}$ , the min-entropy of  $f(x; r)$  for random  $r \leftarrow \mathcal{R}$  is at least  $\text{poly}(\lambda)$ . Suppose further that the quantum one-time authentication scheme in Construction 2 is secure. Then Construction 2 is black-box one-time secure for  $f$ .*

Our proof of Theorem 2 will center on the notion of a *collapsing hash function*. The definition of a collapsing hash function is due to [Unr16]; informally, we say that  $g$  is collapsing if measuring the output of  $g$  collapses the input register from the perspective of any polynomial-time adversary.

Formally, let  $g^H$  be a function that is computed by making calls to a random oracle  $H$ . In this case we define collapsing with the following game.

$\mathcal{G}_{\text{Collapsing}}(b, g^H)$ :

1. The adversary is given a full description of  $g$  and oracle access to  $H$ .
2. The adversary sends the challenger a query on register  $Q$ .
3. The challenger computes  $g^H$  on register  $Q$  and measures the outcome. If  $b = 1$ , the challenger measures  $Q$  as well.
4. The challenger returns  $Q$  to the adversary.
5. The adversary outputs a bit  $b'$ .

**Definition 3** (Collapsing hash function [Unr16]). Let  $g^H$  be a function that is computed by making calls to a random oracle  $H$ . We say that  $g^H$  is *collapsing* if, for all adversaries  $\mathcal{A}$  making at most  $\text{poly}(\lambda)$  quantum oracle queries to  $H$ ,

$$\Pr_{\substack{H \\ b \leftarrow \{0,1\}}} [\mathcal{A}^H \text{ outputs } b' = b \text{ in } \mathcal{G}_{\text{Collapsing}}(b, g^H)] \leq \frac{1}{2} + \text{negl}(\lambda).$$

The key idea in our proof of Theorem 2 is to show that the function  $g : x \mapsto f(x; H(x))$  is collapsing if  $H$  is a random function and  $f$  has significant min-entropy for every  $x$ . This is stated as Lemma 1.

**Lemma 1.** *Let  $f : \{0,1\}^m \times \{0,1\}^n \rightarrow \mathcal{Y}$ . Suppose that for all  $x \in \{0,1\}^m$ , the min-entropy of  $f(x; r)$  for random  $r \leftarrow \{0,1\}^n$  is at least  $\tau$ . Then if  $H : \{0,1\}^m \rightarrow \{0,1\}^n$  is a random oracle, the function  $g^H : x \mapsto f(x; H(x))$  is collapsing with advantage  $O(q^3 \cdot 2^{-\tau})$ .*

We will use the compressed oracles technique of [Zha19] to prove Lemma 1 in Section 3.2. This proof is somewhat involved, so before we present it we will show in Section 3.1 that random oracles are collapsing as a warm-up. The proof of Lemma 1 is essentially identical, except that the components need to be updated.

Once we have Lemma 1, Theorem 2 will follow immediately, as we will now see.

*Proof of Theorem 2.* Suppose that an adversary  $\mathcal{A}$  making  $\text{poly}(\lambda)$  many oracle queries satisfies

$$\Pr_{\substack{(\text{sk}, \hat{P}) \leftarrow \text{KeyGen}(1^\lambda, f) \\ |\text{tk}\rangle \leftarrow \text{TokenGen}(\text{sk})}} [\mathcal{A}^{\hat{P}} \text{ wins } \mathcal{G}_{\text{BB-OTP}}(|\text{tk}\rangle, \hat{P})] = \varepsilon$$

for some  $\varepsilon > 0$ . We will use Lemma 1 to give a reduction  $\mathcal{R}$  making  $\text{poly}(\lambda)$  many queries to  $\text{Verify}$  such that

$$\Pr_{\substack{\text{sk} \leftarrow \text{AuthKeyGen}(1^\lambda) \\ |\text{tk}\rangle \leftarrow \text{AuthTokenGen}(\text{sk})}} [\mathcal{R}^{\text{Verify}} \text{ wins } \mathcal{G}_{\text{Auth}}(|\text{tk}\rangle)] \geq \varepsilon - \text{negl}(\lambda),$$

which will complete the proof.

The reduction behaves as follows.

$\mathcal{R}$ :

1. Receive  $|\text{tk}\rangle$  from the challenger and forward it to  $\mathcal{A}$ . Every time the adversary makes a query to  $P$ , use the oracle access to  $\text{Verify}(\text{sk}, \cdot)$  to compute the response.
2. Receive  $Q$  from the adversary. Measure  $Q$ , obtaining outcome  $(x_1, z_1)$ ; return the collapsed register  $Q$  to the adversary.
3. Receive  $Q$  from the adversary again. Measure  $Q$ , obtaining outcome  $(x_2, z_2)$ .

#### 4. Output $(x_1, z_1, x_2, z_2)$ .

The only thing that's different about the adversary's view in  $\mathcal{G}_{\text{BB-OTP}}$  and in the game with  $\mathcal{R}$  is that  $Q$  is measured in the game with  $\mathcal{R}$ . But by Lemma 1, this is indistinguishable to the adversary!

Now observe that if  $\mathcal{A}$  wins  $\mathcal{G}_{\text{BB-OTP}}$ , then  $\mathcal{R}$  breaks the one-time unforgeability of the one-time authentication scheme  $(\text{AuthKeyGen}, \text{AuthTokenGen}, \text{Sign}, \text{Verify})$ . This is because  $y \neq y'$  both must not be  $\perp$ , which means that  $\text{Verify}(\text{sk}, (x_1, z_1)) = \text{Verify}(\text{sk}, (x_2, z_2)) = 1$  and  $x_1 \neq x_2$ . This completes the proof.  $\square$

It only remains to prove Lemma 1.

### 3.1 Warm-up: Random oracles are collapsing

As a warm-up to proving Lemma 1, we use the compressed oracles technique to present a direct proof that random oracles are collapsing. It is already known that random oracles are collapsing by a proof of [Unr16], but that proof makes use of multiple lemmas that are proven using the compressed oracle technique — here, we prove the result directly using this technique.

**Proposition 1.** *A random oracle  $R : \{0, 1\}^m \rightarrow \{0, 1\}^n$  is collapsing with advantage  $O(q^3/2^n)$ , where  $q$  is the number of queries made by the adversary.*

*Proof.* We use the compressed oracle technique due to [Zha19], and we define  $\text{CPhsO}$ ,  $\text{CPhsO}'$  and the notation for compressed oracle databases as in that paper. Suppose that the adversary makes  $q_0$  queries before the challenge query and  $q_1$  queries after. Depending on whether or not the challenger records the challenge query input, the state just after the challenge query is

$$\sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x\rangle \otimes |D, D(x)\rangle$$

or

$$\sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x\rangle \otimes |D, D(x), x\rangle.$$

Here, the first register contains the adversary's state; the second is the query register; the third is the database register; and the fourth and fifth are the registers into which the challenger records the challenge query output and (in the second case) input.

Next, the adversary makes  $q_1$  further queries to  $\text{CPhsO}$ . Let  $U$  be the unitary describing the evolution of the game state under these queries.

Let  $\text{Invert}$  be the unitary that reads  $|D, y\rangle$  and writes the lexicographically first  $x$  such that  $(x, y) \in D$  onto a new register, if there is any such  $x$ . This function is almost identical to  $\text{FindInput}$  from [Zha19]; one can compute either of  $\text{Invert}$  or  $\text{FindInput}$  with one call to the other.

Our proof strategy is to show that applying  $\text{Invert}$  to  $|D, D(x)\rangle$  at the end of the game in the  $b = 0$  case yields the state in the  $b = 1$  case. Since  $\text{Invert}$  is applied only to the challenger's side, this will complete the proof.

Since  $\text{Invert}$  commutes with  $\text{CPhsO}'$ , [Zha19, Lemma 7] implies that  $\text{Invert}$   $O(1/2^n)$ -almost commutes with  $\text{CPhsO}$ . Therefore

$$\left\| \text{Invert} \circ U \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, D(x)\rangle - U \circ \text{Invert} \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, D(x)\rangle \right\| = O\left(q/\sqrt{2^n}\right).$$



By [Zha19, Theorem 2], the mass on branches  $D$  with collisions is at most  $O(\sqrt{q^3/2^n})$  in the state on the right, so:

$$\left\| U \circ \text{Invert} \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, D(x)\rangle - U \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, D(x), x\rangle \right\| = O\left(\sqrt{q^3/2^n}\right).$$

Putting it together, we have

$$\left\| \text{Invert} \circ U \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, D(x)\rangle - U \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, D(x), x\rangle \right\| = O\left(\sqrt{q^3/2^n}\right),$$

completing the proof.  $\square$

### 3.2 Proof of Lemma 1

In this section we will prove Lemma 1, which we reproduce below.

**Lemma 1.** *Let  $f : \{0,1\}^m \times \{0,1\}^n \rightarrow \mathcal{Y}$ . Suppose that for all  $x \in \{0,1\}^m$ , the min-entropy of  $f(x;r)$  for random  $r \leftarrow \{0,1\}^n$  is at least  $\tau$ . Then if  $H : \{0,1\}^m \rightarrow \{0,1\}^n$  is a random oracle, the function  $g : x \mapsto f(x; H(x))$  is collapsing with advantage  $O(q^3 \cdot 2^{-\tau})$ .*

*Proof.* We will follow the same proof strategy as Proposition 1, except that we will need to generalize the components to handle the case where the high-entropy function  $f$  is applied to the random oracle output. We give our generalizations of [Zha19, Lemma 7] and [Zha19, Theorem 2] in Claim 2 and Claim 1, respectively.

Suppose that the adversary makes  $q_0$  queries before the challenge query and  $q_1$  queries after. Depending on whether or not the challenger records the challenge query input, the state just after the challenge query is

$$\sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x\rangle \otimes |D, f(x; D(x))\rangle$$

or

$$\sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x\rangle \otimes |D, f(x; D(x)), x\rangle.$$

The first register contains the adversary's state; the second is the query register; the third is the database register; and the fourth and fifth are the registers into which the challenger records the challenge query output and (in the second case) input.

Next, the adversary makes  $q_1$  further queries to CPhsO. Let  $U$  be the unitary describing the evolution of the game state under these queries.

Let  $\text{Invert}_f$  be the unitary that, given  $D, y$ , writes the lexicographically first  $x$  such that  $f(x; D(x)) = y$  onto a new register, if such an  $x$  exists. Our proof strategy is to show that applying  $\text{Invert}_f$  to  $|D, f(x; D(x))\rangle$  at the end of the game in the  $b = 0$  case yields the state in the  $b = 1$  case. Since  $\text{Invert}_f$  is applied only to the challenger's side, this will complete the proof.

By Claim 2,

$$\left\| \text{Invert}_f \circ U \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, f(x; D(x))\rangle - U \circ \text{Invert}_f \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, f(x; D(x))\rangle \right\| = O\left(q/\sqrt{2^\tau}\right).$$

By Claim 1, the mass on branches  $D$  with collisions  $f(x; D(x)) = f(x'; D(x'))$  is at most  $O(\sqrt{q^3/2^\tau})$  in the state on the right, so

$$\left\| U \circ \text{Invert}_f \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, f(x; D(x))\rangle \right. \\ \left. - U \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, f(x; D(x)), x\rangle \right\| = O\left(\sqrt{q^3/2^\tau}\right).$$

Putting it together, we have

$$\left\| \text{Invert}_f \circ U \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, f(x; D(x))\rangle \right. \\ \left. - U \sum_{D,x} \alpha_{D,x} |\psi_{D,x}\rangle \otimes |x, D, f(x; D(x)), x\rangle \right\| = O\left(\sqrt{q^3/2^\tau}\right),$$

completing the proof.  $\square$

We need a generalization of Theorem 2 from [Zha19].

**Claim 1.** *For any adversary making  $q$  queries to CPhsO, if the database  $D$  is measured after the  $q$  queries, the resulting database will contain distinct inputs  $x, x'$  such that  $f(x; D(x)) = f(x'; D(x'))$  with probability at most  $O(q^3/2^\tau)$ .*

*Proof.* We closely follow the proof of Theorem 2 from [Zha19].

We say that  $D$  contains an  $f$ -collision if it contains distinct inputs  $x, x'$  such that  $f(x; D(x)) = f(x'; D(x'))$ . The compressed oracle's database starts out empty, so the probability of containing an  $f$ -collision starts out as 0. We will show that the probability cannot rise too much with each query. Consider the game state just before the  $q$ th query:

$$|\psi\rangle = \sum_{x,w,z,D} \alpha_{x,w,z,D} |x, w, z\rangle \otimes |D\rangle,$$

where  $D$  is the compressed phase oracle,  $x, w$  are the query registers, and  $z$  is the adversary's internal register.

Let  $P$  be the projection onto the span of basis states  $|x, w, z\rangle \otimes |D\rangle$  where  $D$  contains an  $f$ -collision. We will relate  $\|P|\psi\rangle\|$  to  $\|P \circ \text{CPhsO}|\psi\rangle\|$ .

Let  $Q$  be the projection onto the span of basis states  $|x, w, z\rangle \otimes |D\rangle$  such that  $D$  does not contain an  $f$ -collision,  $w \neq 0$ , and  $D(x) = \perp$ . Let  $R$  be the projection onto the span of states such that  $D$  does not contain an  $f$ -collision,  $w \neq 0$ , and  $D(x) \neq \perp$ . Let  $S$  be the projection onto the span of states such that  $D$  does not contain an  $f$ -collision and  $w = 0$ . Observe that  $P + Q + R + S = I$ .

Applying  $P \circ \text{CPhsO}$  to any state  $\sum_{x,w,z,D} \alpha_{x,w,z,D} |x, w, z\rangle \otimes |D\rangle$  in the support of  $Q$  yields

$$\sum_{x,w,z,D} \alpha_{x,w,z,D} |x, w, z\rangle \otimes 2^{-n/2} \sum_{r: \exists(x', r') \in D \text{ s.t. } f(x; r) = f(x'; r')} (-1)^{w \cdot r} |D \cup (x, r)\rangle.$$

Let  $|D\rangle = |(x_1, r_1), \dots, (x_{q'}, r_{q'})\rangle$ , where  $q' \leq q$ . We can then write the above state as  $2^{-n/2} \sum_{i=1}^q |\phi_i\rangle$ , where

$$|\phi_i\rangle = \sum_{x,w,z,D} \alpha_{x,w,z,D} |x, w, z\rangle \otimes \sum_{r: f(x; r) = f(x_i; r_i)} (-1)^{w \cdot r} |D \cup (x, r)\rangle.$$

The  $|\phi_i\rangle$  are orthogonal and satisfy  $\| |\phi_i\rangle \| \leq \sqrt{2^{n-\tau}} \|Q|\psi\rangle\|$ , because there are at most  $2^{n-\tau}$  choices of  $r$  satisfying  $f(x; r) = f(x_i; r_i)$  due to the min-entropy condition on  $f$ . Therefore,  $\|P \circ \text{CPhsO} \circ Q|\psi\rangle\| \leq \sqrt{q/2^\tau} \|Q|\psi\rangle\|$ .

For states  $R|\psi\rangle = \sum_{x,w,z,D} \alpha_{x,w,z,D} |x,w,z\rangle \otimes |D\rangle$  in the support of  $R$ , let  $D'$  be  $D$  with  $x$  removed and write  $D = D' \cup (x, r)$  for  $r = D(x)$ . Then  $\text{CPhsO} |x,w,z\rangle \otimes |D\rangle$  is

$$|x,w,z\rangle \otimes \left( (-1)^{w \cdot r} \left( |D' \cup (x, r)\rangle + \frac{1}{\sqrt{2^n}} |D'\rangle \right) + \frac{1}{2^n} \sum_{r'} (1 - (-1)^{w \cdot r} - (-1)^{w \cdot r'}) |D' \cup (x, r')\rangle \right),$$

so

$$P \circ \text{CPhsO} \circ R|\psi\rangle = \sum_{x,w,z,D',r} \alpha_{x,w,z,D',r} |x,w,z\rangle \otimes \frac{1}{2^n} \sum_{r': \exists (x'', r'') \in D \text{ s.t. } f(x; r') = f(x''; r'')} (1 - (-1)^{w \cdot r} - (-1)^{w \cdot r'}) |D' \cup (x, r')\rangle.$$

Let  $|D'\rangle = |(x_1, r_1), \dots, (x_{q'}, r_{q'})\rangle$  and

$$|\phi_i\rangle = \frac{1}{2^n} \sum_{x,w,z,D',r} \alpha_{x,w,z,D',r} |x,w,z\rangle \otimes \sum_{r': f(x; r') = f(x_i; r_i)} (1 - (-1)^{w \cdot r} - (-1)^{w \cdot r'}) |D' \cup (x, r')\rangle.$$

Then the  $|\phi_i\rangle$  are orthogonal and

$$\begin{aligned} \| |\phi_i\rangle \|^2 &= \frac{1}{4^n} \sum_{x,w,z,D',r': f(x; r') = f(x_i; r_i)} \left| \sum_r \alpha_{x,w,z,D',r} (1 - (-1)^{w \cdot r} - (-1)^{w \cdot r'}) \right|^2 \\ &\leq \frac{3}{2^n} \sum_{x,w,z,D',r,r': f(x; r') = f(x_i; r_i)} |\alpha_{x,w,z,D',r}|^2 \\ &\leq \frac{3}{2^\tau} \|R|\psi\rangle\|^2, \end{aligned}$$

so  $\|P \circ \text{CPhsO} \circ R|\psi\rangle\| \leq \sqrt{3q/2^\tau} \|R|\psi\rangle\|$ .

Finally,  $\|P \circ \text{CPhsO} \circ P|\psi\rangle\| \leq \|P|\psi\rangle\|$  and  $\text{CPhsO} \circ S|\psi\rangle = S|\psi\rangle$ . Putting everything together,  $\|P|\psi\rangle\|$  increases by at most  $O(\sqrt{q/2^n})$  with each query. Therefore, after  $q$  queries, the total norm of  $P|\psi\rangle$  is at most  $O(\sqrt{q^3/2^n})$ , completing the proof.  $\square$

The following is similar to Lemma 7 from [Zha19].

**Claim 2.** *Consider a quantum system over  $x, w, D, y$ . Let  $f$  be a function such that for every  $x$ , for uniformly random  $r = D(x)$ ,  $f(x; r)$  has min-entropy at least  $\tau$ . Then the following unitaries  $O(1/\sqrt{2^\tau})$ -almost commute:*

- $\text{CPhsO}$ , acting on the  $x, w, D$  registers.
- $\text{Invert}_f$ , taking as input the  $D, y$  registers and XORing the output onto a new register.

*Proof.* Recall the definitions of  $\text{CPhsO}'$ ,  $\text{StdDecomp}$ , and  $\text{Increase}$  from [Zha19]. In particular, recall that  $\text{CPhsO} = \text{StdDecomp} \circ \text{CPhsO}' \circ \text{StdDecomp} \circ \text{Increase}$ . Since  $\text{Invert}_f$  commutes with  $\text{CPhsO}'$  and  $\text{Increase}$  by construction, it suffices to show that  $\text{Invert}_f$  and  $\text{StdDecomp}$  are  $O(1/\sqrt{2^\tau})$ -almost commuting. To show this, the following intuition is taken from [Zha19], adapted to our setting.

For  $\text{StdDecomp}$  to have any effect, either (1)  $D(x) = \perp$  or (2)  $D(x)$  is in uniform superposition;  $\text{StdDecomp}$  will simply toggle between the two cases. In Case (2), the probability that  $\text{Invert}_f$  will find a match at input  $x$  is at most  $2^{-\tau}$ . And in Case (1),  $\text{Invert}_f$  will never find a match at input  $x$ . Hence, there is an exponentially small error between the action of  $\text{Invert}_f$  on these two cases.

More formally, let  $|\psi\rangle = \sum_{x,D,y} \alpha_{x,D,y} |x, D, y\rangle$ . We omit the  $w$  register because neither  $\text{StdDecomp}$  nor  $\text{Invert}_f$  touch  $w$ . Let  $P$  be the projection onto  $D(x) = \perp$ , and let  $Q$  be the projection onto  $D(x) = |+\rangle$ . Writing  $D = D' \cup (x, \perp)$ ,

$$\text{StdDecomp} \circ \text{Invert}_f \circ P |\psi\rangle = 2^{-n/2} \sum_{x,D',y,r} \alpha_{x,D',y} |x, D' \cup (x, r), y, \text{Invert}_f(D', y)\rangle$$

and

$$\text{Invert}_f \circ \text{StdDecomp} \circ P |\psi\rangle = 2^{-n/2} \sum_{x,D',y,r} \alpha_{x,D',y} |x, D' \cup (x, r), y, \text{Invert}_f(D' \cup (x, r), y)\rangle,$$

where  $\text{Invert}_f(D, y)$  outputs the lexicographically first  $x$  such that  $f(x, D(x)) = y$ , if such an  $x$  exists. Letting  $\Delta := \text{StdDecomp} \circ \text{Invert}_f - \text{Invert}_f \circ \text{StdDecomp}$ , we have

$$\Delta \circ P |\psi\rangle = 2^{-n/2} \sum_{x,D',y,r} \alpha_{x,D',y} |x, D' \cup (x, r), y\rangle \otimes (|\text{Invert}_f(D', y)\rangle - |\text{Invert}_f(D' \cup (x, r), y)\rangle),$$

and

$$\begin{aligned} \|\Delta \circ P |\psi\rangle\|^2 &= 2^{-n} \sum_{x,D',y,r} |\alpha_{x,D',y}|^2 (2 - 2 \cdot \mathbf{1}\{\text{Invert}_f(D', y) = \text{Invert}_f(D' \cup (x, r), y)\}) \\ &= 2\|P |\psi\rangle\|^2 - 2 \sum_{x,D',y} |\alpha_{x,D',y}|^2 \Pr_r[\text{Invert}_f(D', y) = \text{Invert}_f(D' \cup (x, r), y)] \\ &\leq \frac{2}{2^\tau} \|P |\psi\rangle\|^2, \end{aligned}$$

since  $\Pr_r[\text{Invert}_f(D', y) = \text{Invert}_f(D' \cup (x, r), y)] \geq 1 - 2^{-\tau}$ .

Similarly,  $\|\Delta \circ Q |\psi\rangle\|^2 \leq 2 \cdot 2^{-\tau} \|Q |\psi\rangle\|^2$ . Since  $\Delta \circ (I - P - Q) = 0$ , it follows that  $\|\Delta |\psi\rangle\| = O(1/\sqrt{2^\tau})$ .  $\square$

## References

- [Aar09] Scott Aaronson. Quantum copy-protection and quantum money. In *2009 24th Annual IEEE Conference on Computational Complexity*, pages 229–242. IEEE, 2009.
- [AB24] Prabhanjan Ananth and Amit Behera. A modular approach to unclonable cryptography. In Leonid Reyzin and Douglas Stebila, editors, *Advances in Cryptology - CRYPTO 2024 - 44th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2024, Proceedings, Part VII*, volume 14926 of *Lecture Notes in Computer Science*, pages 3–37. Springer, 2024.
- [ALL<sup>+</sup>21] Scott Aaronson, Jiahui Liu, Qipeng Liu, Mark Zhandry, and Ruizhe Zhang. New approaches for quantum copy-protection. In Tal Malkin and Chris Peikert, editors, *Advances in Cryptology - CRYPTO 2021 - 41st Annual International Cryptology Conference, CRYPTO 2021, Virtual Event, August 16-20, 2021, Proceedings, Part I*, volume 12825 of *Lecture Notes in Computer Science*, pages 526–555. Springer, 2021.
- [AP21] Prabhanjan Ananth and Rolando L. La Placa. Secure software leasing. In Anne Canteaut and François-Xavier Standaert, editors, *Advances in Cryptology - EUROCRYPT 2021 - 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, October 17-21, 2021, Proceedings, Part II*, volume 12697 of *Lecture Notes in Computer Science*, pages 501–530. Springer, 2021.
- [BB14] Charles H Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical computer science*, 560:7–11, 2014.

- [BGS13] Anne Broadbent, Gus Gutoski, and Douglas Stebila. Quantum one-time programs. *IACR Cryptol. ePrint Arch.*, page 343, 2013.
- [BI20] Anne Broadbent and Rabib Islam. Quantum encryption with certified deletion. In Rafael Pass and Krzysztof Pietrzak, editors, *Theory of Cryptography - 18th International Conference, TCC 2020, Durham, NC, USA, November 16-19, 2020, Proceedings, Part III*, volume 12552 of *Lecture Notes in Computer Science*, pages 92–122. Springer, 2020.
- [BS23] Shalev Ben-David and Or Sattath. Quantum tokens for digital signatures. *Quantum*, 7:901, 2023.
- [CG24] Andrea Coladangelo and Sam Gunn. How to use quantum indistinguishability obfuscation. In Bojan Mohar, Igor Shinkar, and Ryan O’Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 1003–1008. ACM, 2024.
- [CHV23] Céline Chevalier, Paul Hermouet, and Quoc-Huy Vu. Semi-quantum copy-protection and more. In Guy N. Rothblum and Hoeteck Wee, editors, *Theory of Cryptography - 21st International Conference, TCC 2023, Taipei, Taiwan, November 29 - December 2, 2023, Proceedings, Part IV*, volume 14372 of *Lecture Notes in Computer Science*, pages 155–182. Springer, 2023.
- [CLLZ21] Andrea Coladangelo, Jiahui Liu, Qipeng Liu, and Mark Zhandry. Hidden cosets and applications to unclonable cryptography. In Tal Malkin and Chris Peikert, editors, *Advances in Cryptology - CRYPTO 2021 - 41st Annual International Cryptology Conference, CRYPTO 2021, Virtual Event, August 16-20, 2021, Proceedings, Part I*, volume 12825 of *Lecture Notes in Computer Science*, pages 556–584. Springer, 2021.
- [GKR08] Shafi Goldwasser, Yael Tauman Kalai, and Guy N. Rothblum. One-time programs. In David A. Wagner, editor, *Advances in Cryptology - CRYPTO 2008, 28th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2008. Proceedings*, volume 5157 of *Lecture Notes in Computer Science*, pages 39–56. Springer, 2008.
- [GLR<sup>+</sup>24] Aparna Gupte, Jiahui Liu, Justin Raizes, Bhaskar Roberts, and Vinod Vaikuntanathan. Quantum one-time programs, revisited. *CoRR*, 2024.
- [LLQZ22] Jiahui Liu, Qipeng Liu, Luowen Qian, and Mark Zhandry. Collusion resistant copy-protection for watermarkable functionalities. In Eike Kiltz and Vinod Vaikuntanathan, editors, *Theory of Cryptography - 20th International Conference, TCC 2022, Chicago, IL, USA, November 7-10, 2022, Proceedings, Part I*, volume 13747 of *Lecture Notes in Computer Science*, pages 294–323. Springer, 2022.
- [Shm22] Omri Shmueli. Semi-quantum tokenized signatures. In Yevgeniy Dodis and Thomas Shrimpton, editors, *Advances in Cryptology - CRYPTO 2022 - 42nd Annual International Cryptology Conference, CRYPTO 2022, Santa Barbara, CA, USA, August 15-18, 2022, Proceedings, Part I*, volume 13507 of *Lecture Notes in Computer Science*, pages 296–319. Springer, 2022.
- [Unr16] Dominique Unruh. Computationally binding quantum commitments. In Marc Fischlin and Jean-Sébastien Coron, editors, *Advances in Cryptology - EUROCRYPT 2016 - 35th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Vienna, Austria, May 8-12, 2016, Proceedings, Part II*, volume 9666 of *Lecture Notes in Computer Science*, pages 497–527. Springer, 2016.
- [Wie83] Stephen Wiesner. Conjugate coding. *ACM Sigact News*, 15(1):78–88, 1983.
- [Zha19] Mark Zhandry. How to record quantum queries, and applications to quantum indistinguishability. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology - CRYPTO 2019 - 39th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18-22,*

2019, *Proceedings, Part II*, volume 11693 of *Lecture Notes in Computer Science*, pages 239–268. Springer, 2019.