Post-quantum Security of Plain OAEP Transform

Ehsan Ebrahimi

Department of Computer Science, University of Luxembourg

Abstract. In this paper, we show that OAEP transform is indistinguishable under chosen ciphertext attack in the quantum random oracle model if the underlying trapdoor permutation is quantum partial-domain one-way. The existing post-quantum security of OAEP (TCC 2016-B [14]) requires a modification to the OAEP transform using an extra hash function. We prove the security of the OAEP transform without any modification and this answers an open question in one of the finalists of NIST competition, NTRU submission [6], affirmatively.

Keywords. Post-quantum Security, OAEP, Quantum Random Oracle Model

1 Introduction

The rapid progress on quantum computing and the existence of quantum algorithms like Shor's algorithm [12] has sparked the necessity of replacing old cryptography with post-quantum cryptography. Toward this goal, the National Institute of Standards and Technology (NIST) has initiated a competition for post-quantum cryptography. In this paper we address an open question in one of the finalists of NIST competition, NTRU submission [6]. The security of (unmodified) Optimal Asymmetric Encryption Padding (OAEP) in the quantum random oracle model has been mentioned as an interesting open question in [6][1]. The existing post-quantum security proof of OAEP [14] requires a modification to OAEP transform. (See details below.)

The random oracle model [1] is a powerful model in which the security of a cryptographic scheme is proven assuming the existence of a truly random function that is accessible by all parties including the adversary. But in real world applications, the random oracle will be replaced with a cryptographic hash function and the code of this function is public and known to the adversary. Following [4], we use the quantum random oracle model in which the adversary can make queries to the random oracle in superposition (that is, given a superposition of inputs, he can get a superposition of output values). This is necessary since a quantum adversary attacking a scheme based on a real hash function is necessarily

[1] In the subsection 2.4.5 (titled: An IND-CCA2 PKE using Q-OAEP) of the version dated September 2020.

able to evaluate that function in superposition. Hence the random oracle model must reflect that ability if one requests post-quantum security.

Bellare and Rogaway [2] proposed OAEP transform, for converting a trapdoor permutation into a public-key encryption scheme using two random oracles. It was believed that the OAEP-cryptosystem is provable secure in the random oracle model based on one-wayness of trapdoor permutation, but Shoup [13] showed it is an unjustified belief. Later, Fujisaki et al. [9] proved IND-CCA security of the OAEP-cryptosystem based on a stronger assumption, namely, partial-domain one-wayness of the underlying permutation.

Is OAEP transform secure in the standard model? A recent work to study this question [5] shows that a full instantiation of RSA-OAEP is only possible for two variants of RSA-OAEP (called 't-clear' and 's-clear'). Also, we emphasize that the positive results in [5] hold against a classical adversary and one needs to investigate the possibility of such instantiation in the post-quantum setting. For instance, the partial instantiations are based on algebraic properties of the RSA assumption that trivially does not hold in the post-quantum setting. Or the full instantiation of t-clear RSA-OAEP is based on non-standard assumptions (called 'XOR-type' assumptions) for which an intuitive justifications has been only given in light of the multiplicative structure of RSA, and etc. Even though the post-quantum instantiation of the random oracles in OAEP is a relevant research question, it is not in the scope of this paper and we leave a further investigation as an open question. Here, we investigate the security of OAEP transform in the quantum random oracle model.

Post-quantum security of OAEP transform has been studied in [14]. The authors modified OAEP transform (called it Q-OAEP) using an extra hash function that is length-preserving and show that Q-OAEP is IND-CCA secure in the quantum random oracle model. The extra hash function in Q-OAEP is used to extract the preimage of a random oracle queries in the security proof. In this work, we show that this extra hash function is unnecessary. We use Zhandry's compressed oracle technique [17] to prove IND-qCCA security of OAEP transform (without any modification) in the quantum random oracle model. IND-qCCA notion introduced in [3] is an adaptation of IND-CCA in which the adversary is allowed to make quantum decryption queries, but, the challenge query is restricted to be classical. Since security in the sense of IND-qCCA implies IND-CCA security, our result answers an open question in one of the finalists of NIST competition, NTRU [6], affirmatively.

Note that in the IND-qCCA notion, the adversary's challenge queries are restricted to be classical. Proposing a quantum IND-CCA notion that grants the adversary the possibility of submitting quantum challenge queries is a challenging task with some partial successes [7, 10]. We postpone verifying the security of OAEP transform in the sense of definitions in [7, 10] until a definite definition is given.

Organization. In Section 2, we present some basics of quantum information and computation, security definitions needed in the paper and an introduction for the Compressed Standard Oracle that has been introduced in [17] which we

use in the paper. In Section 3, we present the OAEP scheme and show that it is IND-qCCA secure in the quantum random oracle model.

2 Preliminaries

Notations. Let $\mathtt{MSP}$ shows the message space. The notation $x \xleftarrow{\$} X$ means that x is chosen uniformly at random from the set X. For a natural number n, $[n]$ means the set $\{1, \cdots, n\}$. $\Pr[P : G]$ is the probability that the predicate P holds true where free variables in P are assigned according to the program in G. The function $\mathsf{negl(n)}$ is any non-negative function that is smaller than the inverse of any non-negative polynomial $p(n)$ for sufficiently large n. That is, $\lim_{n\to\infty} \mathsf{negl(n)}p(n) = 0$ for any polynomial $p(n)$. For a function f, f_x denotes the evaluation of f on the input x, that is $f(x)$. For a bit-string x of size more-than-equal k, $[x]_k$ are the k least significant bits of x and $[x]^k$ are the k most significant bits of x. For two bits b and b', $[b = b']$ is 1 if $b = b'$ and it is 0 otherwise.

2.1 Quantum Computing

We present basics of quantum computing in this subsection. The interested reader can refer to [11] for more information. For two vectors $|\Psi\rangle = (\psi_1, \psi_2, \cdots, \psi_n)$ and $|\Phi\rangle = (\phi_1, \phi_2, \cdots, \phi_n)$ in $\mathbb{C}^n$, the inner product is defined as $\langle \Psi, \Phi\rangle = \sum_i \psi_i^* \phi_i$ where ψ_i^* is the complex conjugate of ψ_i. Norm of $|\Phi\rangle$ is defined as $\| |\Phi\rangle \| = \sqrt{\langle \Phi, \Phi\rangle}$. The n-dimensional Hilbert space $\mathcal{H}$ is the complex vector space $\mathbb{C}^n$ with the inner product defined above. A quantum system is a Hilbert space $\mathcal{H}$ and a quantum state $|\psi\rangle$ is a vector $|\psi\rangle$ in $\mathcal{H}$ with norm 1. A unitary operation over $\mathcal{H}$ is a transformation $\mathbb{U}$ such that $\mathbb{U}\mathbb{U}^\dagger = \mathbb{U}^\dagger\mathbb{U} = \mathbb{I}$ where $\mathbb{U}^\dagger$ is the Hermitian transpose of $\mathbb{U}$ and $\mathbb{I}$ is the identity operator over $\mathcal{H}$. Norm of an operator $\mathbb{U}$ is $\|\mathbb{U}\| = \max_{|\psi\rangle} \|\mathbb{U}|\psi\rangle\|$. The computational basis for $\mathcal{H}$ consists of $\log n$ vectors $|b_i\rangle$ of length $\log n$ with 1 in the position i and 0 elsewhere. With this basis, the Hadamard unitary is defined as

$$\mathbb{H} : |b\rangle \to \frac{1}{\sqrt{2}}(|\bar{b}\rangle + (-1)^b |b\rangle),$$

for $b \in \{0, 1\}$ where $\bar{b} = 1 - b$. The controlled-swap unitary is defined as

$$|b\rangle |\psi_0\rangle |\psi_1\rangle \to |b\rangle |\psi_b\rangle |\psi_{\bar{b}}\rangle ,$$

for $b \in \{0, 1\}$. The controlled-unitary $\mathbb{U}$ ($c\mathbb{U}$) is define as:

$$c\mathbb{U} |b\rangle |\Psi\rangle \to \begin{cases} |b\rangle \mathbb{U} |\Psi\rangle & \text{if } b = 1 \\ |b\rangle |\Psi\rangle & \text{if } b = 0 \end{cases}.$$

The bit-flip unitary $\mathbb{X}$ maps $|b\rangle$ to $|\bar{b}\rangle$ for $b \in \{0, 1\}$. An orthogonal projection $\mathbb{P}$ over $\mathcal{H}$ is a linear transformation such that $\mathbb{P}^2 = \mathbb{P} = \mathbb{P}^\dagger$. A measurement on a

Hilbert space is defined with a family of projectors that are pairwise orthogonal. An example of measurement is the computational basis measurement in which any projection is defined by a basis vector. The output of computational measurement on a state $|\Psi\rangle$ is i with probability $\|\langle b_i, \Psi\rangle\|^2$ and the post measurement state is $|b_i\rangle$. For a general measurement $\{\mathbb{P}_i\}_i$, the output of this measurement on a state $|\Psi\rangle$ is i with probability $\|\mathbb{P}_i |\Psi\rangle\|^2$ and the post measurement state is $\frac{\mathbb{P}_i|\Psi\rangle}{\|\mathbb{P}_i|\Psi\rangle\|}$.

For two operators $\mathbb{U}_1$ and $\mathbb{U}_2$, the commutator is $[\mathbb{U}_1, \mathbb{U}_2] = \mathbb{U}_1\mathbb{U}_2 - \mathbb{U}_2\mathbb{U}_1$. For two quantum systems $\mathcal{H}_1$ and $\mathcal{H}_2$, the composition of them is defined by the tensor product and it is $\mathcal{H}_1 \otimes \mathcal{H}_2$. For two unitary $\mathbb{U}_1$ and $\mathbb{U}_2$ defined over $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, $(\mathbb{U}_1 \otimes \mathbb{U}_2)(\mathcal{H}_1 \otimes \mathcal{H}_2) = \mathbb{U}_1(\mathcal{H}_1) \otimes \mathbb{U}_2(\mathcal{H}_2)$. In this paper, QFT over an n-qubits system is $\mathbb{H}^{\otimes n}$.

If a system is in the state $|\Psi_i\rangle$ with the probability p_i, we interpret this with a quantum ensemble $E = \{(|\Psi_i\rangle, p_i)\}_i$. Different outputs of a quantum algorithm can be represented as a quantum ensemble. The density operator corresponding with the ensemble E is $\rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|$ where $|\Psi_i\rangle\langle\Psi_i|$ is the operator acting as $|\Psi_i\rangle\langle\Psi_i| : |\Phi\rangle \rightarrow \langle\Psi_i, \Phi\rangle|\Psi_i\rangle$. The trace distance of two density operators ρ_1, ρ_2 is defined as $\mathrm{TD}(\rho_1, \rho_2) := \frac{1}{2}\mathrm{tr}|\rho_1 - \rho_2|$ where tr is the trace of a square matrix (the sum of entries on the main diagonal) and $|\rho_1 - \rho_2| := \sqrt{(\rho_1 - \rho_2)^\dagger(\rho_1 - \rho_2)}$. Note that the trace distance of two pure states $|\Psi\rangle, |\Phi\rangle$ is defined as $\mathrm{TD}(|\Psi\rangle\langle\Psi|, |\Phi\rangle\langle\Phi|)$.

Any classical function $f : X \rightarrow Y$ can be implemented as a unitary operator $\mathbb{U}_f$ in a quantum computer where $\mathbb{U}_f : |x, y\rangle \rightarrow |x, y \oplus f(x)\rangle$ and it is clear that $\mathbb{U}_f^\dagger = \mathbb{U}_f$. A quantum adversary has standard oracle access to a classical function f if it can query the unitary $\mathbb{U}_f$.

2.2 Definitions

Here, we define a public-key encryption scheme, the IND-qCCA security notion and the quantum partial-domain one-wayness.

Definition 1. *A public-key encryption scheme $\mathcal{E}$ consists of three polynomial-time (in the security parameter n) algorithms, $\mathcal{E} = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$, such that:*

1. Gen*, the key generation algorithm, is a probabilistic algorithm which on input 1^n outputs a pair of keys, $(pk, sk) \leftarrow \mathrm{Gen}(1^n)$, called the public key and the secret key for the encryption scheme, respectively.*
2. Enc*, the encryption algorithm, is a probabilistic algorithm which takes as input a public key pk and a message $m \in \mathtt{MSP}$ and outputs a ciphertext $c \leftarrow \mathrm{Enc}_{pk}(m)$. The message space, $\mathtt{MSP}$, may depend on pk.*
3. Dec*, the decryption algorithm, is a deterministic algorithm that takes as input a secret key sk and a ciphertext c and returns the message $m := \mathrm{Dec}_{sk}(c)$. It is required that the decryption algorithm returns the original message, i.e., $\mathrm{Dec}_{sk}(\mathrm{Enc}_{pk}(m)) = m$, for every $(pk, sk) \leftarrow \mathrm{Gen}(1^n)$ and every $m \in \mathtt{MSP}$. The algorithm* Dec *returns $\perp$ if ciphertext c is not decryptable.*

In the following, we define the IND-qCCA security notion [3] in the quantum random oracle model. The IND-qCCA security notion for a public-key encryption

scheme allows the adversary to make quantum decryption queries but the challenge query is classical. We define $\mathbb{U}_{\mathrm{Dec}}$ as:

$$\mathbb{U}_{\mathrm{Dec}}\,|c,y\rangle \rightarrow \begin{cases} |c, y\oplus \perp\rangle & \text{if } c^* \text{ is defined} \quad \wedge \;\; c = c^* \\ |c, y \oplus \mathrm{Dec}_{sk}(c)\rangle & \text{otherwise} \end{cases},$$

where c^* is the challenge ciphertext and $\perp$ is a value outside of the output space. We say that a quantum algorithm $\mathcal{A}$ has quantum access to the random oracle H if $\mathcal{A}$ can submit queries in superposition and the oracle H answers to these queries by applying a unitary transformation that maps $|x,\ y\rangle$ to $|x,\ y \oplus H(x)\rangle$.

Definition 2 (IND-qCCA in the quantum random oracle model). *A public-key encryption scheme $\mathcal{E} = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ is IND-qCCA secure if for any **quantum** polynomial-time adversary $\mathcal{A}$*

$$\Pr\Big[b = 1 : b \leftarrow \mathsf{Exp}^{qCCA,qRO}_{\mathcal{A},\mathcal{E}}(n)\Big] \leq 1/2 + \mathsf{negl(n)},$$

where $\mathsf{Exp}^{qCCA,qRO}_{\mathcal{A},\mathcal{E}}(n)$ game is define as:
$\mathsf{Exp}^{qCCA,qRO}_{\mathcal{A},\mathcal{E}}(n)$ *game:*
***Key Gen:** The challenger runs $\mathrm{Gen}(1^n)$ to obtain a pair of keys (pk, sk) and chooses random oracles.*

***Query:** The adversary $\mathcal{A}$ given the public key pk, the **quantum** oracle access to $\mathbb{U}_{\mathrm{Dec}}$ and the **quantum** access to the random oracles, chooses two **classical** messages m_0, m_1 of the same length and sends them to the challenger. The challenger chooses a random bit b and responds with $c^* \leftarrow \mathrm{Enc}_{pk}(m_b)$.*

***Guess:** The adversary $\mathcal{A}$ continues to query the decryption oracle and the random oracles. Finally, the adversary $\mathcal{A}$ produces a bit b'. The output of the game is $[b = b']$.*

Definition 3 (Quantum partial-domain one-way function). *We say a permutation $f : \{0,1\}^{n+k_1} \times \{0,1\}^{k_0} \rightarrow \{0,1\}^m$ is quantum partial-domain one-way if for any polynomial-time quantum adversary A,*

$$\Pr\Big[\tilde{s} = s : s \xleftarrow{\$} \{0,1\}^{n+k_1},\ t \xleftarrow{\$} \{0,1\}^{k_0},\ \tilde{s} \leftarrow A(f(s,t))\Big] \leq \mathsf{negl(n)}.$$

2.3 Compressed Standard Oracle

In this section, we briefly present the Compressed Standard Oracle (CStO) that has been introduced in [17]. The interested reader can refer to [8, 17] for more details.

In the standard quantum random oracle model, a function $H : \{0,1\}^m \rightarrow \{0,1\}^n$ is chosen uniformly at random from the set of all functions (lets call it

Ω_H) and superposition queries will be answered by the unitary $\mathbb{U}_H$ that maps $|x,\ y\rangle$ to $|x,\ y \oplus H(x)\rangle$. Another perspective to consider this is that the oracle puts the superposition of all functions on his private register[2] and a query is implemented as

$$\mathsf{StO} : |x,\ y\rangle \sum_H \frac{1}{\sqrt{|\Omega_H|}} |H\rangle \to \sum_H \frac{1}{\sqrt{|\Omega_H|}} |x,\ y \oplus H(x)\rangle |H\rangle .$$

Note that if the oracle measures its internal state in the computational basis, this corresponds to choosing H uniformly at random from Ω_H and answer with $\mathbb{U}_H$. So these two oracles are perfectly indistinguishable. Now if we apply QFT to the output register before and after applying StO, we will get the Phase oracle that operates as follows:

$$\mathsf{PhO} : |x, y\rangle \sum_H \frac{1}{\sqrt{|\Omega_H|}} |H\rangle \to \sum_H \frac{1}{\sqrt{|\Omega_H|}} (-1)^{y \cdot H(x)} |x, y\rangle |H\rangle .$$

Let $\mathfrak{D}$ represent the truth table of the function H and $P_{x,y}$ represent the truth table of the point function that is y on the input x and it is zero elsewhere. With this notation we can write the query above as follows:

$$\mathsf{PhO} : |x,\ y\rangle \sum_{\mathfrak{D}} \frac{1}{\sqrt{|\Omega_H|}} |\mathfrak{D}\rangle \to \sum_{\mathfrak{D}} \frac{1}{\sqrt{|\Omega_H|}} (-1)^{P_{x,y} \cdot \mathfrak{D}} |x, y\rangle |\mathfrak{D}\rangle .$$

Now if the oracle applies QFT to the oracle register after applying PhO, it will get:

$$\mathsf{QFT}_{\mathfrak{D}}\mathsf{PhO} : |x,\ y\rangle \sum_{\mathfrak{D}} \frac{1}{\sqrt{|\Omega_H|}} |\mathfrak{D}\rangle \to |x, y\rangle |P_{x,y}\rangle .$$

Note that $\mathsf{QFT}_{\mathfrak{D}}$ only effects the oracle state and it is undetectable to the adversary. At this stage, the oracle will symmetrically store the inputs/outputs of the adversary's queries in its private register. Informally, if the oracle is able to move the entry that is not zero in the database $P_{x,y}$ to the beginning of its private register and remove all the zero slots (without the adversary's detection), the private register of the oracle can contain a polynomial number of registers.

$$\mathsf{RmoV}_{\mathfrak{D}}\mathsf{MoV}_{\mathfrak{D}}\mathsf{QFT}_{\mathfrak{D}}\mathsf{PhO} : \sum_{x,y} \alpha_{x,y} |x, y\rangle \sum_{\mathfrak{D}} \frac{1}{\sqrt{|\Omega_H|}} |\mathfrak{D}\rangle \to \sum_{x,y} \alpha_{x,y} |x, y\rangle |x, y\rangle .^{3}$$

Following the perspective above, Zhandry [17] developed the CStO that its private register can be implemented efficiently, symmetrically stores the inputs/outputs of the adversary's queries in its private register and it is perfectly indistinguishable from the standard oracle (StO).

Lemma 1 (Lemma 4 in [17]). CStO *and* StO *are perfectly indistinguishable.*

[2]This requires an exponential number of registers that is not efficient.

[3]This informal 'move' and 'remove' operations are detectable to the adversary and they are given only to build the intuition behind CStO.

For the rest, we import the representation of CStO from [8]. Let $\mathfrak{D} = \otimes_{x\in X}\mathfrak{D}_x$ be the oracle register. The state space of $\mathfrak{D}_x$ is generated with vectors $|y\rangle$ for $y \in Y \cup \{\perp\}$. Let $F_{\mathfrak{D}_x}$ be a unitary acting on $\mathfrak{D}_x$ that maps $|\perp\rangle$ to $\mathsf{QFT}\,|0\rangle$ and vice versa. And for any vector orthogonal to $|\perp\rangle$ and $\mathsf{QFT}\,|0\rangle$, F is identity. We define CStO to be the following unitary acting on the input register, the output register and the $\mathfrak{D}$ register.

$$\mathsf{CStO} = \sum_x |x\rangle\langle x| \otimes F_{\mathfrak{D}_x} CNOT_{Y\mathfrak{D}_x} F_{\mathfrak{D}_x},$$

where $CNOT_{Y\mathfrak{D}_x}\,|y, y_x\rangle = |y \oplus y_x, y_x\rangle$ for $y, y_x \in Y$ and it is identity on $|y, \perp\rangle$. The initial state of $\mathfrak{D}$ register is $\otimes_{x\in X}\,|\perp\rangle$.

We call a query to CStO 'dummy' if its output register is set to the uniform superposition. Note that for such a query $CNOT_{Y\mathfrak{D}_x}$ is identity and therefore CStO is identity.

In the following, we present preliminaries for Theorem 3.1 in [8] that will be used in the security proof in Section 3. For a fixed relation $R \subset X \times Y$, Γ_R is the maximum number of y's that fulfill the relation R where the maximum is taken over all $x \in X$:

$$\Gamma_R = \max_{x\in X} |\{y \in Y | (x, y) \in R\}|.$$

We define a projector $\Pi^x_{\mathfrak{D}_x}$ that checks if the register $\mathfrak{D}_x$ contains a value $y \neq \perp$ such that $(x, y) \in R$:

$$\Pi^x_{\mathfrak{D}_x} := \sum_{y\ s.t.\ (x,y)\in R} |y\rangle\langle y|_{\mathfrak{D}_x}.$$

Let $\bar{\Pi}^x_{\mathfrak{D}_x} = \mathbb{I}_{\mathfrak{D}_x} - \Pi^x_{\mathfrak{D}_x}$. We define the measurement $\mathbb{M}$ to be the set of projectors $\{\Sigma^x\}_{x\in X\cup\{\emptyset\}}$ where

$$\Sigma^x := \bigotimes_{x'<x} \bar{\Pi}^{x'}_{\mathfrak{D}_{x'}} \otimes \Pi^x_{\mathfrak{D}_x} \text{ for } x \in X \text{ and } \Sigma^\emptyset := \mathbb{I} - \sum_x \Sigma^x. \tag{1}$$

Informally, the measurement $\mathbb{M}$ checks for the smallest x for which $\mathfrak{D}_x$ contains a value $y \neq \perp$ such that $(x, y) \in R$. If no register $\mathfrak{D}_x$ contains a value $y \neq \perp$ such that $(x, y) \in R$, the outcome of $\mathbb{M}$ is $\emptyset$. We define a purified measurement $\mathbb{M}_{\mathfrak{D}P}$ corresponding to $\mathbb{M}$ that XORs the outcome of the measurement to an ancillary register:

$$\mathbb{M}_{\mathfrak{D}P}\,|\phi, z\rangle_{\mathfrak{D}P} \to \sum_{x\in X\cup\{\emptyset\}} \Sigma^x\,|\phi\rangle_{\mathfrak{D}}\,|z \oplus x\rangle_P.$$

The following lemma states that CStO and $\mathbb{M}_{\mathfrak{D}P}$ almost commute if Γ_R is small proportional to the size of Y.

Lemma 2 (Theorem 3.1 in [8]). *For any relation R and Γ_R defined above, the commutator $[\mathsf{CStO}, \mathbb{M}_{\mathfrak{D}P}]$ is bounded as follows:*

$$\|[\mathsf{CStO}, \mathbb{M}_{\mathfrak{D}P}]\| \leq 8 \cdot 2^{-n/2}\sqrt{2\Gamma_R}.$$

It has been shown that a quantum adversary needs an exponential number of quantum queries to a random oracle to find a collision [16]. As an immediate corollary, a random injective function is indistinguishable from a random oracle for a quantum polynomial-time adversary. We use this corollary in the security proof of OAEP.

Lemma 3 (From [16]). *Any quantum adversary making q queries to a random oracle $H : \{0,1\}^m \to \{0,1\}^n$ outputs a collision for H with probability at most $C(q+1)^3/2^n$ where C is a universal constant.*

In addition to the lemmas above, we use the 'gentle-measurement lemma' [15] in the proof. Informally, it states that if an output of a measurement is almost certain for a quantum state, the measurement does not disturb the state much.

Lemma 4 (gentle-measurement lemma). *Let $\mathbb{M} = \{\mathbb{P}_i\}_i$ is a measurement. For any state $|\Psi\rangle$, if there exists an i such that $\|\mathbb{P}_i |\Psi\rangle\|^2 \geq 1-\epsilon$, then $\mathrm{TD}(|\Psi\rangle, \mathbb{M}|\Psi\rangle) \leq \sqrt{\epsilon} + \epsilon$.*

3 Security of OAEP

In this section, we define OAEP transformation and prove that it is IND-qCCA secure in the quantum random oracle model if the underlying trapdoor permutation is quantum partial-domain one-way. (Since IND-qCCA security trivially implies IND-CCA security, our result shows that OAEP transform is IND-CCA in the quantum random oracle model if the underlying trapdoor permutation is quantum partial-domain one-way.)

Definition 4 (OAEP). *Let $G : \{0,1\}^{k_0} \to \{0,1\}^{n+k_1}$, $H : \{0,1\}^{n+k_1} \to \{0,1\}^{k_0}$ be random oracles. The encryption scheme $OAEP = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ is defined as:*

1. Gen*: Specifies an instance of the injective function f and its inverse f^{-1}. Therefore, the public key and secret key are f and f^{-1} respectively.*
2. Enc*: Given a message $m \in \{0,1\}^n$, the encryption algorithm computes*

$$s := m||0^{k_1} \oplus G(r) \quad \textit{and} \quad t := r \oplus H(s),$$

   where $r \xleftarrow{\$} \{0,1\}^{k_0}$, and outputs the ciphertext $c := f(s,t)$[4]*.*
3. Dec*: Given a ciphertext c, the decryption algorithm does the following: Compute $f^{-1}(c) = (s,t)$ and then,*
   (a) *query the random oracle H on input s, query the random oracle G on input $t \oplus H(s)$ and compute $M := s \oplus G(t \oplus H(s))$. In addition it submits two dummy queries to the random oracle G* [5]*.*

[4] Q-OAEP in [14] outputs the ciphertext $c := \big(f(s,t), H'(s,t)\big)$ for a fresh random oracle H'.

[5] Note that these dummy queries are required to make the number of queries submitted to G equal in the Games 1 and 2 in the security proof.

(b) if the k_1 least significant bits of M are zero then return the n most significant bits of M, otherwise return $\perp$.

Note that k_0 and k_1 depend on the security parameter n.

We prove the security of OAEP for the parameters $k_0 - n = O(n)$ (this is needed to show that Games 1 and 2 are indistinguishable) and $n + k_1 \geq k_0$ (because we need to replace the random oracle G with a random injective function in Game 1).

Here we sketch the main ideas to prove the IND-qCCA security of OAEP in the quantum random oracle. We start with the IND-qCCA game in QROM in which the adversary wins if he guesses the challenge bit b correctly. By introducing some (indistinguishable) intermediate games we reach the last game in which the adversary's success probability is 1/2. In the last game, the adversary is not allowed to query the randomness r^* that is used to obtain the challenge ciphertext c^*. (Since queries are quantum, this is prevented by measuring the input register of the queries to G by the projective measurement $\mathbb{M}_{r^*} = \{\mathbb{P}_1 = |r^*\rangle\langle r^*|, \mathbb{P}_0 = \mathbb{I} - \mathbb{P}_1\}$ and aborting if the outcome is 1.) Therefore, $G(r^*)$ is a random value for the adversary and $m_b||0^{k_1} \oplus G(r^*)$ hides the challenge bit b information-theoretically.

Note that at some steps of the proof, the indistinguishability of two games (specifically two last games in the proof) needs to be reduced to the partial-domain one-wayness of the underlying permutation. A reduction adversary to break the partial-domain one-wayness of the underlying permutation needs to answer the decryption queries without knowing f^{-1}. In this step, the reduction adversary uses the databases of the compressed standard oracles corresponding to the random oracles H, G for decryption. (On the input c it searches over the inputs/outputs of the random oracle queries in the databases of H, G that satisfies $c = f(s, r \oplus H_s)$ and $[G_r \oplus s]_{k_1} = 0^{k_1}$ and outputs $[G_r \oplus s]^n$.) However, it is not straightforward to show that this new decryption algorithm is indistinguishable from the decryption algorithm of the OAEP scheme. This is because a decryption algorithm that uses the databases to decrypt may cause detectable effects on the databases. In other words, the extraction of data from the databases may be detectable to the adversary. Here we use Lemma 2 to show that the oracle can extract information from the databases without an adversary's detection. We show this indistinguishability by modifying the decryption algorithm of the OAEP scheme step by step to reach the decryption algorithm that only uses the databases.

Theorem 1. *If the underlying permutation is quantum partial-domain one-way, then the OAEP scheme is IND-qCCA secure in the quantum random oracle model.*

Proof. Let Ω_H and Ω_G be the set of all function $G : \{0,1\}^{k_0} \to \{0,1\}^{n+k_1}$ and $H : \{0,1\}^{n+k_1} \to \{0,1\}^{k_0}$, respectively. Let S_G shows the set of all injective functions from $\{0,1\}^{k_0}$ to $\{0,1\}^{n+k_1}$. Let A be a polynomial-time quantum adversary that attacks the OAEP-cryptosystem in the sense of IND-qCCA in the quantum random oracle model and makes at most q_H and q_G queries to the random oracles

H and G respectively and q_{dec} decryption queries.

Game 0: This is IND-qCCA game in qROM, $\mathsf{Exp}^{qCCA,qRO}_{\mathcal{A},\mathcal{OAEP}}(n)$.

Game 0:

let $(pk, sk) \leftarrow \mathrm{Gen}(1^n)$, $r^* \xleftarrow{\$} \{0,1\}^{k_0}$, $b \xleftarrow{\$} \{0,1\}$, $H \xleftarrow{\$} \Omega_H, G \xleftarrow{\$} \Omega_G$
let $m_0, m_1 \leftarrow A^{H,G,\mathbb{U}_{\mathrm{Dec}}}(pk)$
let $s^* := m_b||0^{k_1} \oplus G(r^*)$, $t^* := r^* \oplus H(s^*)$, $c^* := f(s^*, t^*)$
let $b' \leftarrow A^{H,G,\mathbb{U}_{\mathrm{Dec}}}(c^*)$
return $[b = b']$

Game 1: In this game, we consider H is being implemented as the compressed standard oracles CStO_H and G is replaced with a random injective function.

Game 1:

let $(pk, sk) \leftarrow \mathrm{Gen}(1^n)$, $r^* \xleftarrow{\$} \{0,1\}^{k_0}$, $b \xleftarrow{\$} \{0,1\}$, $G \xleftarrow{\$} S_G$
let $m_0, m_1 \leftarrow A^{\mathsf{CStO}_H,G,\mathbb{U}_{\mathrm{Dec}}}(pk)$
let $s^* := m_b||0^{k_1} \oplus G(r^*)$, $t^* := r^* \oplus H(s^*)$, $c^* := f(s^*, t^*)$
let $b' \leftarrow A^{\mathsf{CStO}_H,G,\mathbb{U}_{\mathrm{Dec}}}(c^*)$
return $[b = b']$

Since CStO_H and the standard oracles StO_H are perfectly indistinguishable by Lemma 1, this change does not effect the adversary's success probability. And changing the random oracle G to a random injective function is distinguishable by a probability at most $C(q_G+3q_{dec}+2)^3/2^{n+k_1}$ by Lemma 3. (Each decryption query makes three random oracle queries to G, so the total number of queries to G is at most $q_G + 3q_{dec}$ plus 1 for the challenge query.)

Game 2: In this game we change $\mathbb{U}_{\mathrm{Dec}}$ oracle to $\mathbb{U}_{\mathrm{Dec}^{(1)}}$ described below. Let $\mathfrak{D}_H$ denotes the database of CStO_H. We define the relation R^H_c to be the set of all (s, H_s) such that $[G(H_s \oplus [f^{-1}(c)]_{k_0}) \oplus s]_{k_1} = 0^{k_1}$. Given the relation R^H_c, the projectors Σ^s_c for $s \in \{0,1\}^{n+k_1}$ and $\Sigma^{\emptyset}_c$ are defined similar to Equation (1). Now the measurement $\mathbb{M}^H = \{\Sigma^s_c\}_{s\in\{0,1\}^{n+k_1}\cup\{\emptyset\}}$ checks if there exists a pair in $\mathfrak{D}_H$ satisfying the relation R^H_c or not. If there is more than one pair satisfying the relation R^H_c, the smallest s will be the output of $\mathbb{M}^H$. If there is no such a pair the output of $\mathbb{M}^H$ is $\emptyset$. Let $\mathbb{M}^c_{\mathfrak{D}_H,P_H}$ be the following purified measurement corresponding to $\mathbb{M}^H$:

$$\mathbb{M}^c_{\mathfrak{D}_H,P_H} |\phi, z\rangle_{\mathfrak{D}_H P_H} \to \sum_{s\in\{0,1\}^{n+k_1}\cup\{\emptyset\}} \Sigma^s_c |\phi\rangle_{\mathfrak{D}_H} |z \oplus s\rangle_{P_H} .$$

We define the unitary $\mathbb{M}_{\mathfrak{D}_H,P_H}$ that operates on the ciphertext, $\mathfrak{D}_H$ and P_H registers as:

$$\mathbb{M}_{\mathfrak{D}_H,P_H} |c\rangle |\phi, z\rangle_{\mathfrak{D}_H P} \to |c\rangle \otimes \mathbb{M}^c_{\mathfrak{D}_H,P_H} |\phi, z\rangle_{\mathfrak{D}_H P_H} .$$

Note that $\mathbb{M}_{\mathfrak{D}_H,P_H}$ is an involution, that is, $\mathbb{M}_{\mathfrak{D}_H,P_H}\mathbb{M}_{\mathfrak{D}_H,P_H} = \mathbb{I}$. For each decryption query, $\mathbb{U}_{\mathrm{Dec}^{(1)}}$ first applies the $\mathbb{M}_{\mathfrak{D}_H,P_H}$ unitary with the P_H register initiated with 0. Then it executes $\mathbb{U}_{\mathrm{Dec}}$ without submitting the two dummy queries to the random oracle G. We denote this slightly modified decryption algorithm by $\mathbb{U}'_{\mathrm{Dec}}$. (We omit these dummy queries since $\mathbb{M}_{\mathfrak{D}_H,P_H}$ makes two queries to G in each decryption query.) Finally it applies the $\mathbb{M}_{\mathfrak{D}_H,P_H}$ again.

$$\mathbb{U}_{\mathrm{Dec}^{(1)}} = \mathbb{M}_{\mathfrak{D}_H,P_H}\ \mathbb{U}'_{\mathrm{Dec}}\ \ \mathbb{M}_{\mathfrak{D}_H,P_H}.$$

Game 2:

let $(pk, sk) \leftarrow \mathrm{Gen}(1^n)$, $r^* \xleftarrow{\$} \{0,1\}^{k_0}$, $b \xleftarrow{\$} \{0,1\}$, $G \xleftarrow{\$} S_G$
let $m_0, m_1 \leftarrow A^{\mathsf{CStO}_H, G, \mathbb{U}_{\mathrm{Dec}^{(1)}}}(pk)$
let $s^* := m_b||0^{k_1} \oplus G(r^*)$, $t^* := r^* \oplus H(s^*)$, $c^* := f(s^*, t^*)$
let $b' \leftarrow A^{\mathsf{CStO}_H, G, \mathbb{U}_{\mathrm{Dec}^{(1)}}}(c^*)$
return $[b = b']$

We prove $\mathbb{M}_{\mathfrak{D}_H,P_H}$ and $\mathbb{U}'_{\mathrm{Dec}}$ almost commute to show the indistinguishability of these two games. Note that $\mathbb{M}_{\mathfrak{D}_H,P_H}$ only interfaces with $\mathbb{U}'_{\mathrm{Dec}}$ when $\mathbb{U}'_{\mathrm{Dec}}$ makes a query to the random oracle H. In other words, the reason that $\mathbb{M}_{\mathfrak{D}_H,P_H}$ does not commute with $\mathbb{U}'_{\mathrm{Dec}}$ is that $\mathbb{U}'_{\mathrm{Dec}}$ makes a random oracle query to H in each decryption query. By Lemma 2, if we commute $\mathbb{M}^c_{\mathfrak{D}_H,P_H}$ and $\mathbb{U}_{\mathrm{Dec}}$, this will be distinguishable to the adversary with a probability at most $8 \cdot 2^{-\frac{k_0}{2}} \sqrt{2\Gamma_{R^c_H}}$. Since G is an injective function $\Gamma_{R^c_H} = 2^n$. Therefore the distinguishing advantage of the adversary is at most $2^{\frac{n-k_0}{2}+\frac{7}{2}}$ that is negligible because $k_0 - n = O(n)$. The overall advantage of the adversary in distinguishing these two games is at most $q_{dec}2^{\frac{n-k_0}{2}+\frac{7}{2}}$.

Game 3: In this game we replace the random injective function with a compressed standard oracle CStO_G. (First we replace the random injective function with a random oracle and then we change it to a compressed standard oracle. We do these two changes in one game in favor of reducing the total number of games in the proof.)

Game 3:

let $(pk, sk) \leftarrow \mathrm{Gen}(1^n)$, $r^* \xleftarrow{\$} \{0,1\}^{k_0}$, $b \xleftarrow{\$} \{0,1\}$,
let $m_0, m_1 \leftarrow A^{\mathsf{CStO}_H, \mathsf{CStO}_G, \mathbb{U}_{\mathrm{Dec}^{(1)}}}(pk)$
let $s^* := m_b||0^{k_1} \oplus G(r^*)$, $t^* := r^* \oplus H(s^*)$, $c^* := f(s^*, t^*)$
let $b' \leftarrow A^{\mathsf{CStO}_H, \mathsf{CStO}_G, \mathbb{U}_{\mathrm{Dec}^{(1)}}}(c^*)$
return $[b = b']$

Replacing the random injective function with a random oracle G is distinguishable with a probability at most $C(q_G + 3q_{dec} + 1)^3/2^{n+k_1}$ by Lemma 3.

Then, a random oracle G is perfectly indistinguishable from a CStO_G by Lemma 1.

Game 4: In this game we change $\mathbb{U}_{\mathrm{Dec}^{(1)}}$ oracle to $\mathbb{U}_{\mathrm{Dec}^{(2)}}$ described below. Let $\mathfrak{D}_G$ denotes the database of CStO_G. We define the relation R_c^G to be the set of all (r, G_r) such that $[[f^{-1}(c)]^{n+k_1} \oplus G_r]_{k_1} = 0^{k_1}$. Given the relation R_c^G, the projectors Σ_c^r for $r \in \{0,1\}^{k_0}$ and $\Sigma_c^{\emptyset}$ are defined similar to Equation (1). Now the measurement $\mathbb{M}^G = \{\Sigma_c^r\}_{r\in\{0,1\}^{k_0}\cup\{\emptyset\}}$ checks if there exists a pair in $\mathfrak{D}_G$ satisfying the relation R_c^G or not. If there are more than one pair satisfying the relation R_c^G, the smallest r will be the output of $\mathbb{M}^G$. If there is no such a pair the output of $\mathbb{M}^G$ is $\emptyset$. Let $\mathbb{M}^c_{\mathfrak{D}_G,P_G}$ be the following purified measurement corresponding to $\mathbb{M}^G$:

$$\mathbb{M}^c_{\mathfrak{D}_G,P_G} |\phi, z\rangle_{\mathfrak{D}_G P_G} \to \sum_{r\in\{0,1\}^{k_0}\cup\{\emptyset\}} \Sigma_c^r |\phi\rangle_{\mathfrak{D}_G} |z \oplus r\rangle_{P_G} .$$

We define the unitary $\mathbb{M}_{\mathfrak{D}_G,P_G}$ that operates on the ciphertext, $\mathfrak{D}_G$ and P_G registers as:

$$\mathbb{M}_{\mathfrak{D}_G,P_G} |c\rangle |\phi, z\rangle_{\mathfrak{D}_G,P_G} \to |c\rangle \otimes \mathbb{M}^c_{\mathfrak{D}_G,P_G} |\phi, z\rangle_{\mathfrak{D}_G P_G} .$$

Note that $\mathbb{M}_{\mathfrak{D}_G,P_G}$ is an involution. For each decryption query, $\mathbb{U}_{\mathrm{Dec}^{(2)}}$ first applies the $\mathbb{M}_{\mathfrak{D}_H,P_H}$ unitary with the P_H register initiated with 0. Then it applies the $\mathbb{M}_{\mathfrak{D}_G,P_G}$ unitary with the P_G register initiated with 0. Then it executes $\mathbb{U}'_{\mathrm{Dec}}$. And finally it applies $\mathbb{M}_{\mathfrak{D}_G,P_G}$ and $\mathbb{M}_{\mathfrak{D}_H,P_H}$ again.

$$\mathbb{U}_{\mathrm{Dec}^{(2)}} = \mathbb{M}_{\mathfrak{D}_H,P_H}\ \mathbb{M}_{\mathfrak{D}_G,P_G}\ \mathbb{U}'_{\mathrm{Dec}}\ \mathbb{M}_{\mathfrak{D}_G,P_G}\ \mathbb{M}_{\mathfrak{D}_H,P_H}.$$

Game 4:

let $(pk, sk) \leftarrow \mathrm{Gen}(1^n)$, $r^* \xleftarrow{\$} \{0,1\}^{k_0}$, $b \xleftarrow{\$} \{0,1\}$
let $m_0, m_1 \leftarrow A^{\mathsf{CStO}_H,\mathsf{CStO}_G,\mathbb{U}_{\mathrm{Dec}^{(2)}}}(pk)$
let $s^* := m_b||0^{k_1} \oplus G(r^*)$, $t^* := r^* \oplus H(s^*)$, $c^* := f(s^*, t^*)$
let $b' \leftarrow A^{\mathsf{CStO}_H,\mathsf{CStO}_G,\mathbb{U}_{\mathrm{Dec}^{(2)}}}(c^*)$
return $[b = b']$

In order to show the indistinguishability of two games, we show that $\mathbb{U}'_{\mathrm{Dec}}$ and $\mathbb{M}_{\mathfrak{D}_G,P_G}$ almost commutes (then $\mathbb{M}_{\mathfrak{D}_G,P_G}$ will cancel out with its second application and we will get $\mathbb{U}_{\mathrm{Dec}^{(1)}}$). Note that $\mathbb{U}'_{\mathrm{Dec}}$ does not commute with $\mathbb{M}_{\mathfrak{D}_G,P_G}$ because it makes a random oracle query to G in each decryption query. In other words, $\mathbb{U}'_{\mathrm{Dec}}$ would commute with $\mathbb{M}_{\mathfrak{D}_G,P_G}$ if $\mathbb{U}'_{\mathrm{Dec}}$ had not made a random oracle query to G. By Lemma 2, if we commute $\mathbb{M}^c_{\mathfrak{D}_G,P_G}$ and $\mathbb{U}_{\mathrm{Dec}}$, this will be distinguishable to the adversary with a probability at most $8 \cdot 2^{-\frac{n+k_1}{2}} \sqrt{2\Gamma_{R_G^c}}$. Since $\Gamma_{R_G^c} = 2^n$, the overall distinguishing advantage of the adversary is at most $q_{dec} 2^{-\frac{k_1}{2}+\frac{7}{2}}$.

Game 5: In this game we change $\mathbb{U}_{\mathrm{Dec}(2)}$ oracle to $\mathbb{U}_{\mathrm{Dec}^{(3)}}$ described below. For each decryption query, $\mathbb{U}_{\mathrm{Dec}^{(3)}}$ first applies $\mathbb{M}_{\mathfrak{D}_H,P_H}$ and then $\mathbb{M}_{\mathfrak{D}_G,P_G}$ with the P_H and P_G registers initiated with 0. Then, if c^* is defined and $c = c^*$ it XORs $\perp$ to the output register. Otherwise, if the P_H register contains $\emptyset$ or the P_G register contains $\emptyset$ it XORs $\perp$ to the output register and make a dummy query to the random oracles G, H. If the P_H and P_G registers do not contain $\emptyset$, it executes $\mathbb{U}'_{\mathrm{Dec}}$:

$$|c,y\rangle\,|z_1\rangle_{P_H}\,|z_2\rangle_{P_G} \rightarrow \begin{cases} |c, y\oplus \perp\rangle\,|z_1\rangle\,|z_2\rangle & \text{if } c^* \text{ is defined } \ \wedge\ c = c^* \\ |c, y\oplus \perp\rangle\,|z_1\rangle\,|z_2\rangle & \text{if } z_1 = \emptyset \vee z_2 = \emptyset \\ \big|c, y \oplus \mathrm{Dec}_{f^{-1}}(c)\big\rangle\,|z_1\rangle\,|z_2\rangle & \text{if } z_1 \neq \emptyset \ \wedge\ z_2 \neq \emptyset \end{cases}.$$

Finally, it applies the unitary $\mathbb{M}_{\mathfrak{D}_G,P_G}$ and $\mathbb{M}_{\mathfrak{D}_H,P_H}$.

Game 5:

let $(pk, sk) \leftarrow \mathrm{Gen}(1^n)$, $r^* \xleftarrow{\$} \{0,1\}^{k_0}$, $b \xleftarrow{\$} \{0,1\}$
let $m_0, m_1 \leftarrow A^{\mathsf{CStO}_H,\mathsf{CStO}_G,\mathbb{U}_{\mathrm{Dec}^{(3)}}}(pk)$
let $s^* := m_b||0^{k_1} \oplus G(r^*)$, $t^* := r^* \oplus H(s^*)$, $c^* := f(s^*, t^*)$
let $b' \leftarrow A^{\mathsf{CStO}_H,\mathsf{CStO}_G,\mathbb{U}_{\mathrm{Dec}^{(3)}}}(c^*)$
return $[b = b']$

We show that $\mathbb{U}_{\mathrm{Dec}^{(2)}}$ and $\mathbb{U}_{\mathrm{Dec}^{(3)}}$ algorithms are indistinguishable. Below, we recall a bit modified version of the decryption algorithm $\mathbb{U}_{\mathrm{Dec}^{(2)}}$:

$$|c,y\rangle\,|z_1\rangle_{P_H}\,|z_2\rangle_{P_G} \rightarrow \begin{cases} |c, y\oplus \perp\rangle\,|z_1\rangle\,|z_2\rangle & \text{if } c^* \text{ is defined } \ \wedge\ c = c^* \\ \big|c, y \oplus \mathrm{Dec}_{f^{-1}}(c)\big\rangle\,|z_1\rangle\,|z_2\rangle & \text{if } z_1 = \emptyset \vee z_2 = \emptyset \\ \big|c, y \oplus \mathrm{Dec}_{f^{-1}}(c)\big\rangle\,|z_1\rangle\,|z_2\rangle & \text{if } z_1 \neq \emptyset \ \wedge\ z_2 \neq \emptyset \end{cases}.$$

Note that if for any ciphertext c for which $z_1 = \emptyset$ or $z_2 = \emptyset$, $\mathrm{Dec}_{f^{-1}}$ (on input c) returns $\perp$ then the algorithms $\mathbb{U}_{\mathrm{Dec}^{(2)}}$ and $\mathbb{U}_{\mathrm{Dec}^{(3)}}$ return the same output in all three cases. In the claim below, we show that if $z_1 = \emptyset$ or $z_2 = \emptyset$, $\mathrm{Dec}_{f^{-1}}(c)$ returns $\perp$ with an overwhelming probability. The high-level argument to prove this claim is that the adversary is not able to output a valid ciphertex (we call a ciphertext c valid if $\mathrm{Dec}_{f^{-1}}(c) \neq \perp$) with an overwhelming probability unless it executes the encryption oracle, that is, unless it executes the random oracle queries. (Note that the ciphertex space is $\{0,1\}^{n+k_0+k_1}$ and the total number of the valid ciphertexts is 2^{n+k_0}. So a random ciphertext is a valid ciphertext with a probability at most $1/2^{k_1}$.) Then we show that if an adversary can distinguish these two games with a non-negligible probability, then a reduction adversary can output a valid ciphertext c for which $z_1 = \emptyset$ or $z_2 = \emptyset$ with a non-negligible probability and this is a contradiction to the claim shown below.

Claim. A ciphertext c for which $z_1 = \emptyset$ or $z_2 = \emptyset$ is a valid ciphertext with a probability at most $1/2^{k_1}$.

Proof. Let c is a ciphertext for which $z_1 = \emptyset$ and let $f^{-1}(c) = (s', t')$. Note that since $z_1 = \emptyset$, there is no pair (s, H_s) in $\mathfrak{D}_H$ that satisfies $[s \oplus G(H_s \oplus t')]_{k_1} = 0^{k_1}$. This means that either $[s' \oplus G(H_{s'} \oplus t')]_{k_1} \neq 0^{k_1}$ or the adversary has not queried the input s' to H. Clearly if $[s' \oplus G(H_{s'} \oplus t')]_{k_1} \neq 0^{k_1}$, $\mathrm{Dec}_{f^{-1}}(c) = \perp$ and c is an invalid ciphertext. And if s' has not been queried to H, $H_{s'}$ is a random value from the adversary's point of view. Therefore, $[s' \oplus G(t' \oplus H_{s'})]_{k_1} = 0^{k_1}$ holds with a probability at most $1/2^{k_1}$. That is, c is a valid ciphertext with a probability at most $1/2^{k_1}$.

Let c is a ciphertext for which $z_2 = \emptyset$ and let $f^{-1}(c) = (s', t')$. Note that since $z_2 = \emptyset$, there is no pair (r, G_r) in $\mathfrak{D}_G$ that satisfies $[s' \oplus G_r]_{k_1} = 0^{k_1}$. This means that either $[s' \oplus G(t' \oplus H_{s'})]_{k_1} \neq 0^{k_1}$ or the adversary has not queried the input $t' \oplus H_{s'}$ to G. If $[s' \oplus G_r]_{k_1} \neq 0^{k_1}$, $\mathrm{Dec}_{f^{-1}}(c) = \perp$ and c is an invalid ciphertext. If $t' \oplus H_{s'}$ has not been queried to G, since G is a random oracle, the probability that $[s' \oplus G(t' \oplus H_{s'})]_{k_1} = 0^{k_1}$ is at most $1/2^{k_1}$. That is, c is a valid ciphertext with a probability at most $1/2^{k_1}$. □

Now let $\mathcal{A}$ is an adversary that distinguishes these two games with a non-negligible advantage. That is, at least one of the $\mathcal{A}$'s decryption queries is of the form

$$\sum_{c_i \text{ for which } z_1=\emptyset \text{ or } z_2=\emptyset,\ j} \alpha_{i,j} \left|c_i\right\rangle \left|y_j\right\rangle + \left|\Psi\right\rangle,$$

where for any i $\mathrm{Dec}_{f^{-1}}(c_i) \neq \perp$ and $\sum_{i,j} \left\|\alpha_{i,j}\right\|^2$ is non-negligible. (Note that if there is no such a query, one can exclude the ciphertexts c_i for which $\mathrm{Dec}_{f^{-1}}(c_i) \neq \perp$ and $z_1 = \emptyset$ or $z_2 = \emptyset$ from the query using an appropriate projective measurement without the adversary's detection (by Lemma 4) in each decryption query and therefore two games will be indistinguishable.)

Now a reduction adversary $\mathcal{B}$ runs $\mathcal{A}$ and measures one of its decryption queries at random. It is clear that $\mathcal{B}$ is able to output a valid ciphertex c for which $z_1 = \emptyset$ or $z_2 = \emptyset$ with a non-negligible probability. And this is a contradiction to the claim above.

Game 6: The decryption algorithm $\mathbb{U}_{\mathrm{Dec}^{(3)}}$ in Game 5 searches over databases $\mathfrak{D}_H, \mathfrak{D}_G$ to find pairs $(s, H_s), (r, G_r)$ such that $[G(H_s \oplus [f^{-1}(c)]_{k_0}) \oplus s]_{k_1} = 0^{k_1}$ and $[[f^{-1}(c)]^{n+k_1} \oplus G_r]_{k_1} = 0^{k_1}$ respectively. Instead of using f^{-1}, we can simply search for pairs $(s, H_s), (r, G_r)$ that satisfy $c = f(s, r \oplus H_s)$ and $[G_r \oplus s]_{k_1} = 0^{k_1}$. In this game, we change $\mathbb{U}_{\mathrm{Dec}^{(3)}}$ a new decryption oracle $\mathbb{U}_{\mathrm{Dec}^{(4)}}$ that searches the databases $\mathfrak{D}_H$ and $\mathfrak{D}_G$ to decrypt. Let Search be a function that on input $(c, \mathfrak{D}_H, \mathfrak{D}_G)$ searches for the pairs (s, H_s) in $\mathfrak{D}_H$ and (r, G_r) in $\mathfrak{D}_G$ such that $c = f(s, r \oplus H_s)$ and $[G_r \oplus s]_{k_1} = 0^{k_1}$. If it finds such pairs, it returns $(1, [G_r \oplus s]^n)$, otherwise it returns $(0, \perp)$.

Let $Q_{b'} Q_m$ be quantum registers of size $(n+1)$ that are initiated with zero. The unitary $\mathbb{U}_{\mathrm{Dec}^{(4)}}$ first applies the unitary $\mathbb{U}_{\mathsf{Search}}$ where its output is stored in

$Q_{b'}Q_m$ registers. Then it does as the following:

$$|c,y\rangle\,|b',m\rangle_{Q_{b'}Q_m} \to \begin{cases} |c,y\oplus \perp\rangle\,|b',m\rangle & \text{if } c^* \text{ is defined} \quad \wedge \quad c=c^* \\ |c,y\oplus \perp\rangle\,|b',m\rangle & \text{if } b'=0 \\ |c,y\oplus m\rangle\,|b',m\rangle & \text{if } b'=1 \end{cases},$$

it submits two dummy queries to the random oracle G in all cases, and it submits a dummy query to the random oracles G, H when $b'=0$ and when $b'=1$. Finally, it applies $\mathbb{U}_{\mathsf{Search}}$ to undo $Q_{b'}Q_m$ registers to zero.

Game 6:

let $(pk, sk) \leftarrow \mathrm{Gen}(1^n)$, $r^* \xleftarrow{\$} \{0,1\}^{k_0}$, $b \xleftarrow{\$} \{0,1\}$
let $m_0, m_1 \leftarrow A^{\mathsf{CStO}_H, \mathsf{CStO}_G, \mathbb{U}_{\mathrm{Dec}^{(4)}}}(pk)$
let $s^* := m_b || 0^{k_1} \oplus G(r^*)$, $t^* := r^* \oplus H(s^*)$, $c^* := f(s^*, t^*)$
let $b' \leftarrow A^{\mathsf{CStO}_H, \mathsf{CStO}_G, \mathbb{U}_{\mathrm{Dec}^{(4)}}}(c^*)$
return $[b = b']$

We show that $\mathbb{U}_{\mathrm{Dec}^{(3)}}$ and $\mathbb{U}_{\mathrm{Dec}^{(4)}}$ are indistinguishable.

1. When c^* is defined and $c = c^*$, both algorithms XOR $\perp$ to the output register and make two random oracle queries to G.
2. When $b' = 0$, it is clear that either z_1 is $\emptyset$ or z_2 is $\emptyset$. Both algorithms XOR $\perp$ to the output register and make three random oracle queries to G and one random oracle query to H.
3. When $b' = 1$, it is clear that $z_1 \neq \emptyset$ and $z_2 \neq \emptyset$. So both algorithms XOR $[G_r \oplus s]^n$ to the output register and make three random oracle queries to G and one random oracle query to H.

Game 7: This is identical to Game 6, except it measures all the queries to CStO_G with the projective measurements $\mathbb{M}_{r^*}$. If there is an 1-output measurement, it aborts and returns a random bit.

Game 7:

let $(pk, sk) \leftarrow \mathrm{Gen}(1^n)$, $r^* \xleftarrow{\$} \{0,1\}^{k_0}$, $b \xleftarrow{\$} \{0,1\}$
$\mathbb{M}_{r^*} = \{\mathbb{P}_1 = |r^*\rangle\langle r^*|, \mathbb{P}_0 = \mathbb{I} - \mathbb{P}_1\}$,
run until *there is an* 1*-output measurement with* $\mathbb{M}_{r^*}$
  let $m_0, m_1 \leftarrow A^{\mathsf{CStO}_H, \mathsf{CStO}_G, \mathbb{U}_{\mathrm{Dec}^{(4)}}}(pk)$
  let $s^* := m_b || 0^{k_1} \oplus G(r^*)$, $t^* := r^* \oplus H(s^*)$, $c^* := f(s^*, t^*)$
  let $b' \leftarrow A^{\mathsf{CStO}_H, \mathsf{CStO}_G, \mathbb{U}_{\mathrm{Dec}^{(4)}}}(c^*)$
return $[b = b']$

Let q_{G1} be the total number of queries submitted to G before the challenge query. Let q_{G2} be the total number of queries submitted to G after the challenge query. ($q_{G1} + q_{G2} = q_G + 3q_{dec}$.) If there is no query to CStO_G with a non-negligible weight on the state $|r^*\rangle$, we can use Lemma 4 (gentle-measurement lemma) to show that these two games are indistinguishable. In more details, let ρ_i is the

state of the i-th query (for $i \in [q_G + 3q_{dec}]$) and let $\mathbb{M}_{r^*}(\rho_i)$ returns 1 with the probability ϵ_i. By the gentle-measurement lemma, the trace distance between $\mathbb{M}_{r^*}(\rho_i)$ and ρ_i is at most $\sqrt{\epsilon_i}+\epsilon_i$. So overall, these two games are distinguishable with the advantage of at most $2(q_G + 3q_{dec})\sqrt{\max_i\{\epsilon_i\}}$. Therefore, if $\max_i\{\epsilon_i\}$ is negligible, two games are indistinguishable.

Since r^* is a random value that has not been used before the challenge query $\mathbb{M}_{r^*}(\rho_i)$ returns 1 with a probability at most $1/2^{k_0}$ for any $i \in [q_{G1}]$. So the measurements before the challenge query are distinguishable with a probability at most $2q_{G1}\sqrt{2^{-k_0}}$ that is negligible.

It is left to show that the measurements after the challenge query are indistinguishable. Proof by contrary, let assume $\mathcal{A}$ makes a query to CStO_G after the challenge query with a non-negligible weight on $|r^*\rangle$. From $\mathcal{A}$, we can construct an adversary $\mathcal{B}$ that breaks the quantum partial-domain one-wayness of f. In more details, $\mathcal{B}$ on input $c^*\big(:= f(s^*, t^*)$ for uniformly random $s^*, t^*\big)$, chooses a random element i from $[q_{G2}]$ and a random bit b, runs the adversary $\mathcal{A}$, answers the random oracle queries and decryption queries using two compressed oracles CStO_H, CStO_G and finally it measures the input register of the i-th query to CStO_G and the database $\mathfrak{D}_H$ with the computational basis measurement, returns an output and aborts. In the following we describe $\mathcal{B}$ in more details.

Simulation of random oracle queries. For H-queries, the adversary $\mathcal{B}$ uses CStO_H. For G-queries, $\mathcal{B}$ does as follows. Let G' be a random oracle with the same domain and co-domain as G. Let Find be an operator that on inputs $r, c^*, \mathfrak{D}_H$, checks if there exists a pair (s, H_s) in $\mathfrak{D}_H$ such that $c^* = f(s, r \oplus H_s)$. If there exists such a pair it returns $(1, s)$. Otherwise, it returns $(0, 0^{n+k_1})$. Note that since f is a permutation, the Find unitary either returns $(0, 0^{n+k_1})$ or returns $(1, s^*)$. For each query, $\mathcal{B}$ first applies Find operator with an ancillary register $Q_{b'}Q_s$ of $(1 + n + k_1)$ qubits initiated with zero. Then, if the query is conducted before the challenge query or the $Q_{b'}$ is set to 0, it forwards the query to $\mathsf{CStO}_{G'}$, otherwise, it XORs $m_b||0^{k_1} \oplus s^*$ to the output register:

$$G : |r, y\rangle\,|\mathfrak{D}_H\rangle \to \begin{cases} |r, y \oplus G'(r)\rangle & \text{if } m_b \text{ is not defined} \\ |r, y \oplus G'(r)\rangle & \text{if } \mathsf{Find}(r, c^*, \mathfrak{D}_H) = (0, 0^{n+k_1}) \\ |r, y \oplus (m_b||0^{k_1} \oplus s^*)\rangle & \text{if } \mathsf{Find}(r, c^*, \mathfrak{D}_H) = (1, s^*) \end{cases}.$$

And finally it applies the Find operator again. Since f is a permutation, there exists only one r such that $c^* = f(s^*, r \oplus H_{s^*})$ and that is r^*. For any $r \neq r^*$ the oracle G and the random oracle G' are the same, therefore, the simulation of G-queries will be indistinguishable from the random oracle G' unless the adversary submits a post-challenge query with a non-negligible weight on the state $|r^*\rangle$ and $\mathsf{Find}(r^*, c^*, \mathfrak{D}_H) = (1, s^*)$. (And if this happens, it breaks the quantum partial-domain one-wayness of f explained below.)

The challenge query. Upon receiving m_0 and m_1 from $\mathcal{A}$, the adversary $\mathcal{B}$ returns c^* as the challenge ciphertext. (Note that the way we simulate G-queries $G(r^*) := m_b||0^{k_1} \oplus s^*$ and $c^* = f(s^*, r^* \oplus H_{S^*})$ that is a perfect simulation of

the challenge query.)

Simulation of decryption queries. $\mathcal{B}$ uses the oracle $\mathbb{U}_{\mathrm{Dec}^{(4)}}$ on inputs $\mathfrak{D}_H$ and $\mathfrak{D}_{G'}$ for the decryption queries. Note that G and G' only differ on the input r^* for which $c^* = f(s^*, r^* \oplus H_{s^*})$. Since $\mathbb{U}_{\mathrm{Dec}^{(4)}}$ on input c^* does not use its database and returns $\perp$, the simulation of the decryption queries is perfect.

Output of $\mathcal{B}$. The adversary $\mathcal{B}$ measures the $(q_{G1}+i)$-th random oracle query to CStO_G with $\mathbb{M}_{r^*}$ and the database $\mathfrak{D}_H$ with the computational basis measurement. Since there exists a query with a non-negligible weight on the state $|r^*\rangle$, the adversary $\mathcal{B}$ can obtain r^* with a non-negligible probability. Then, the adversary searches over the database $\mathfrak{D}_H$ to find a pair (s^*, H_{s^*}) such that $c^* = f(s^*, r^* \oplus H_{s^*})$. If it finds such a pair, it returns s^* as the partial inverse of f on c^* and aborts. Otherwise, it returns $s^* = G'(r^*) \oplus m_b||0^{k_1}$ as the partial inverse of f on the input c^*. (Note that when there is no pair (s^*, H_{s^*}) in $\mathfrak{D}_H$ such that $c^* = f(s^*, r^* \oplus H_{s^*})$, that is $\mathsf{Find}(r^*, c^*, \mathfrak{D}_H) = (0, 0^{n+k_1})$, the G-queries are answered with the random oracle G'. Therefore, the equation $c^* = f(x, r^* \oplus H(x))$ holds for $x = G'(r^*) \oplus m_b||0^{k_1}$.) Since f is quantum partial-domain one-way, Games 6 and 7 are indistinguishable.

Now, it is clear that Game 7 returns 1 with the probability 1/2 because if one of the measurements returns 1, the output of the game is a random bit. If none of the measurements return 1, $G(r^*)$ remains an uniformly random value for $\mathcal{A}$ and consequently $m_b||0^{k_1} \oplus G(r^*)$ is an uniformly random value for $\mathcal{A}$. So the probability that $\mathcal{A}$ guesses b is 1/2. Finally, since each two consecutive games are indistinguishable, the probability that $\mathcal{A}$ guesses b in Game 0 is $1/2 + \mathsf{negl}(\mathsf{n})$ and this finishes the proof of the theorem. □

Acknowledgment. We would like to thank anonymous reviewers for their useful comments and suggestions.

References

1. M. Bellare and P. Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In D. E. Denning, R. Pyle, R. Ganesan, R. S. Sandhu, and V. Ashby, editors, *CCS '93, Proceedings of the 1st ACM Conference on Computer and Communications Security, Fairfax, Virginia, USA, November 3-5, 1993.*, pages 62–73. ACM, 1993.
2. M. Bellare and P. Rogaway. Optimal asymmetric encryption. In A. D. Santis, editor, *Advances in Cryptology - EUROCRYPT '94, Workshop on the Theory and Application of Cryptographic Techniques, Perugia, Italy, May 9-12, 1994, Proceedings*, volume 950 of *Lecture Notes in Computer Science*, pages 92–111. Springer, 1994.
3. Boneh and M. Zhandry. Secure signatures and chosen ciphertext security in a quantum computing world. In R. Canetti and J. A. Garay, editors, *Advances in Cryptology - CRYPTO 2013 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part II*, volume 8043 of *Lecture Notes in Computer Science*, pages 361–379. Springer, 2013.

4. D. Boneh, Ö. Dagdelen, M. Fischlin, A. Lehmann, C. Schaffner, and M. Zhandry. Random oracles in a quantum world. In D. H. Lee and X. Wang, editors, *Advances in Cryptology - ASIACRYPT 2011 - 17th International Conference on the Theory and Application of Cryptology and Information Security, Seoul, South Korea, December 4-8, 2011. Proceedings*, volume 7073 of *Lecture Notes in Computer Science*, pages 41–69. Springer, 2011.
5. N. Cao, A. O'Neill, and M. Zaheri. Toward RSA-OAEP without random oracles. In A. Kiayias, M. Kohlweiss, P. Wallden, and V. Zikas, editors, *Public-Key Cryptography - PKC 2020 - 23rd IACR International Conference on Practice and Theory of Public-Key Cryptography, Edinburgh, UK, May 4-7, 2020, Proceedings, Part I*, volume 12110 of *Lecture Notes in Computer Science*, pages 279–308. Springer, 2020.
6. C. Chen, O. Danba, J. Hoffstein, A. Hulsing, J. Rijneveld, J. M. Schanck, P. Schwabe, W. Whyte, Z. Zhang, T. Saito, T. Yamakawa, and K. Xagawa. Ntru, 2020. `https://ntru.org`.
7. C. Chevalier, E. Ebrahimi, and Q. H. Vu. On the security notions for encryption in a quantum world. *IACR Cryptol. ePrint Arch.*, 2020:237, 2020.
8. J. Don, S. Fehr, C. Majenz, and C. Schaffner. Online-extractability in the quantum random-oracle model. Cryptology ePrint Archive, Report 2021/280, 2021. `https://eprint.iacr.org/2021/280`.
9. E. Fujisaki, T. Okamoto, D. Pointcheval, and J. Stern. RSA-OAEP is secure under the RSA assumption. *J. Cryptology*, 17(2):81–104, 2004.
10. T. Gagliardoni, J. Krämer, and P. Struck. Quantum indistinguishability for public key encryption. *IACR Cryptol. ePrint Arch.*, 2020:266, 2020.
11. M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information (10th Anniversary edition)*. Cambridge University Press, 2016.
12. P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, 1997.
13. V. Shoup. OAEP reconsidered. In J. Kilian, editor, *Advances in Cryptology - CRYPTO 2001, 21st Annual International Cryptology Conference, Santa Barbara, California, USA, August 19-23, 2001, Proceedings*, volume 2139 of *Lecture Notes in Computer Science*, pages 239–259. Springer, 2001.
14. E. E. Targhi and D. Unruh. Post-quantum security of the fujisaki-okamoto and OAEP transforms. In M. Hirt and A. D. Smith, editors, *Theory of Cryptography - 14th International Conference, TCC 2016-B, Beijing, China, October 31 - November 3, 2016, Proceedings, Part II*, volume 9986 of *Lecture Notes in Computer Science*, pages 192–216, 2016.
15. A. J. Winter. Coding theorem and strong converse for quantum channels. *IEEE Trans. Inf. Theory*, 45(7):2481–2485, 1999.
16. M. Zhandry. A note on the quantum collision and set equality problems. *Quantum Inf. Comput.*, 15(7&8):557–567, 2015.
17. M. Zhandry. How to record quantum queries, and applications to quantum indifferentiability. In A. Boldyreva and D. 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.