

# Non-malleability for quantum public-key encryption

Christian Majenz<sup>1,2</sup>, Christian Schaffner<sup>1,2</sup>, and Jeroen van Wier<sup>1,2,3</sup>

<sup>1</sup> Institute for Logic, Language and Computation, University of Amsterdam,  
Amsterdam, Netherlands, {c.majenz,c.schaffner}@uva.nl

<sup>2</sup> QuSoft, Amsterdam, Netherlands

<sup>3</sup> Interdisciplinary Centre for Security, Reliability and Trust, University of  
Luxembourg, Esch-sur-Alzette, Luxembourg, jeroen.vanwier@uni.lu

**Abstract.** Non-malleability is an important security property for public-key encryption (PKE). Its significance is due to the fundamental unachievability of integrity and authenticity guarantees in this setting, rendering it the strongest integrity-like property achievable using only PKE, without digital signatures. In this work, we generalize this notion to the setting of *quantum* public-key encryption. Overcoming the notorious “recording barrier” known from generalizing other integrity-like security notions to quantum encryption, we generalize one of the equivalent classical definitions, comparison-based non-malleability, and show how it can be fulfilled. In addition, we explore one-time non-malleability notions for symmetric-key encryption from the literature by defining plaintext and ciphertext variants and by characterizing their relation.

## 1 Introduction

The development of quantum information processing technology has accelerated recently, with many large public and private players investing heavily [Wal18]. A future where communication networks include at least some high-capacity quantum channels and fault-tolerant quantum computers seems therefore more and more likely. How will we secure communication over the resulting “quantum internet” [WEH18]? One approach is to rely on features inherent to quantum theory to get unconditional security, e.g. by using teleportation. Such methods are, however, a far cry from the classical standard internet cryptography in terms of efficiency, as they require interaction. A different and more efficient approach is to generalize modern private- and public-key cryptography to the quantum realm.

In this paper, we focus on the notion of non-malleability, which captures the idea that an encrypted message cannot be altered by a third party in a structured manner. This notion, first introduced by Dolev, Dwork and Naor [DDN03], derives its importance from the fact that it is the strongest integrity-like notion that is achievable using public-key encryption only. The aim of this work is to generalize this notion to public-key encryption of quantum data. A recent attack that exemplifies the relevance of the concept of non-malleability is

the “efail”-attack on the PGP protocol for confidential and authenticated e-mail communication [Pod+18]. This kind of attack, where an attacker is not directly able to learn the message yet still able to modify it, is exactly what non-malleable encryption secures against.

The classical notion of non-malleability is based on the notion of related plaintexts. For a non-malleable encryption scheme, it should, roughly speaking, be hard for an adversary to transform an encryption of a message  $m$  into a *different* ciphertext that decrypts to a *related* message  $m'$ . Here, “related” just means that the adversary has some control over the transformation that is applied to the plaintext underlying the ciphertext he attacks. Generalizing this notion to the quantum case is complicated by the quantum no-cloning theorem: After a message has been encrypted and modified by the adversary and subsequently decrypted, it cannot be compared with the result anymore. In addition, it cannot be checked in a straightforward manner whether the adversary has indeed modified the ciphertext.

In this work, we overcome these obstacles. The key ideas are the following. In the classical security game, an adversary is first asked to submit a distribution from which a plaintext is sampled. In the quantum setting, any message sampling procedure can be implemented by first performing a unitary quantum computation, and then discarding the contents of an auxiliary register. Instead of discarding this register, we view it as an extra record that is created along with the message. This extra record is then used instead of the original plaintext for evaluating the quantum analog of a relation. The test whether the adversary has indeed modified the ciphertext is performed by running the sampling- and encryption computations backwards on the attacked ciphertext. If the ciphertext was *not* modified, this returns the registers into their initial blank state, which can be detected.

We establish confidence in the new security notion by showing that it becomes equivalent to the classical notion when restricted to the post-quantum setting, i.e. to classical PKE schemes and classical plaintexts and ciphertexts. We also show how to satisfy the new security notion using a classical-quantum hybrid construction.

Along the way, we chart the landscape of one-time non-malleability notions for symmetric-key quantum encryption. We propose definitions for plaintext and ciphertext non-malleability and explore their relationship with existing definitions. In particular, we present evidence that these notions are the right ones.

## 1.1 Related Work

Non-malleability has been studied extensively in the classical setting, see [BS99; PV06] and references therein. In quantum cryptography, non-malleability has been, to our knowledge, subject of only two earlier works [ABW09; AM17], which were only concerned with one-time security for symmetric-key encryption.

Quantum public-key encryption has been studied in [BJ15; Ala+16] with respect to confidentiality.

Problems due to quantum no-cloning and the destructive nature of quantum measurement similar to the ones we face in this work have arisen before in the literature. In particular, devising security notions for quantum encryption where the classical security definition requires copying and comparing plaintexts or ciphertexts [AGM18b; AGM18a], as well as in some quantum attack models for classical cryptography [BZ13a; BZ13b; Ala+18] requires tackling similar obstacles. Another important case where the generalization of classical techniques is complicated by the mentioned features of quantum theory is that of rewinding and reprogramming [Unr12; Wat18; Don+19].

## 1.2 Summary of Contributions

The contributions presented in this paper can be divided into two categories, depending on whether they concern symmetric-key encryption (SKE) or public-key encryption (PKE). While we consider our results of the latter kind our main contribution, they build upon the former results. We therefore begin by presenting our results on one-time non-malleability of quantum SKE in Section 3, after which we continue with the results on many-time non-malleability for quantum PKE in Section 4.

**Symmetric-Key Non-Malleability** All security notions that are concerned with malleability attacks come in two flavors, a *plaintext* and a *ciphertext* one. The difference is that in the former, an attack that modifies a ciphertext into a different one that decrypts to the same plaintext is considered harmless, while the latter considers any modification a successful attack. We refine the non-malleability notion NM introduced in [AM17], to obtain a definition for both ciphertext and plaintext non-malleability, while staying in the *effective-map based framework*. The effective map resulting from a one-time malleability attack is the map on the plaintext space, that is implemented by the procedure of encrypting the input, applying the attack and subsequently decrypting the result again.

**Definition 1.1 (PNM and CiNM, informal).** *A scheme is plaintext non-malleable, (PNM), if for any attack  $A$  the effective map  $\tilde{A}$  consists of replacing the plaintext with a random decryption with some probability  $p$ , and leaving it unchanged otherwise. If the probability  $p$  is equal to the probability that  $A$  acts as the identity on a random ciphertext, it is even ciphertext non-malleable (CiNM).*

There is one important subtlety that we would like to highlight here. The notion NM from [AM17] is very similar to CiNM. The only difference is that in the former,  $p$  is derived from the probability that  $A$  acts as identity on a random element of the ciphertext space, including ciphertexts that are not even valid. For CiNM, on the other hand,  $p$  is the probability that  $A$  acts as identity on a ciphertext that is generated by picking a random plaintext and then encrypting it.

We continue by exploring the relationship between NM, CiNM, and PNM. In particular, we present separating examples between NM and CiNM, NM and

PNM, and CiNM and PNM, and show that both notions of ciphertext non-malleability, NM and CiNM, imply plaintext non-malleability,

**Theorem 1.2 (3.5).** *Any  $\varepsilon$ -NM or  $\varepsilon$ -CiNM SKQES is  $\varepsilon$ -PNM.*

Additionally, we give a simplifying characterization of PNM that allows for efficient simulation following [BW16]. We also show that for encryption schemes with unitary encryption map<sup>4</sup>, all three notions are equivalent.

**Theorem 1.3 (3.8, informal).** *For symmetric-key encryption schemes with unitary encryption map, PNM, CiNM, and NM are equivalent.*

Finally, we show that one can construct a quantum authentication scheme according to the security definition from [DNS12] from a PNM scheme (and therefore, by Theorem 1.2, also from a CiNM scheme). This is done by adding a tag to the plaintext during encryption, which is checked during decryption, as proposed for NM schemes in [AM17].

**Theorem 1.4 (3.10).** *From any PNM scheme, one can construct a  $2^{2-r}$ -DNS-authenticating scheme using  $r$  tag qubits.*

**Public-Key Non-Malleability** We propose a definition for public-key quantum non-malleability in a computational setting, by adapting the classical definition for comparison-based non-malleability found in [BS99], a real-vs-ideal definition. In the following, we describe informally what main challenges the generalization of the classical security experiments (the real and the ideal one) to the quantum setting poses, and how we resolved them.

In the first step in the classical security experiments, the adversary submits a probability distribution  $p$  over messages. In both experiments, a plaintext from this distribution is sampled, encrypted and sent to the adversary. The adversary now has the opportunity to manipulate (or “malleate”) the ciphertext with the goal that the output decrypts to a *related* plaintext.<sup>5</sup> The relation according to which the plaintexts are related, is supplied by the adversary. Of course there are examples of relations that allow for easy creation of a ciphertext that decrypts to a related plaintext, like e.g. the trivial relation where any plaintext is related to any other plaintext. To not credit an adversary with a break for fulfilling such a relation, her success in two experiments is compared: in the real world, the relation is evaluated on the initial and final plaintexts, but in the ideal world, it is evaluated on the final plaintext and a plaintext that is independently sampled from  $p$ .

Attempting a naive quantum generalization, we face two main challenges: How does the challenger ensure that the ciphertext he received from the adversary is actually modified? And how does he evaluate a relation on the input

<sup>4</sup> More precisely, the encryption with a fixed key is unitary.

<sup>5</sup> In the actual experiments, the adversary is allowed to transform the ciphertext into many attempted manipulated ciphertexts. In this informal description we simplify as no significant additional technical challenges arise from the generalization.

plaintext and the decrypted one? Both questions are complicated by the fact that quantum information cannot be copied. The first question has a rather elegant solution. Instead of asking the adversary to provide a distribution of messages, we ask her to provide a state preparation circuit, a strict generalization of the former. Such a state preparation circuit starts from a blank register and prepares a quantum state on the plaintext register and an auxiliary register. But quantum operations are reversible, which means that to test whether the plaintext has changed after encryption, attack and decryption,<sup>6</sup> we can run preparation backwards and measure whether we got back a blank register. If so, the ciphertext was not changed, and the candidate manipulated plaintext is discarded. If not, we run preparation forward again, recovering (the actually changed part of) the adversary’s candidate malleation.

The second question is solved by exploiting the fact that in the quantum setting, any message-sampling procedure can be implemented by first performing a unitary quantum computation, and then discarding the contents of an auxiliary register. We can therefore ask the adversary to provide a state-sampling unitary and store the auxiliary register as a record indicating which plaintext has been created. After proceeding with the experiment as in the classical case and using the modification test as described above, instead of checking whether the original and the attacked plaintext are related, we can now, in the real world case, check whether the attacked plaintext is related to the record. In the ideal world, the record is replaced with an independently created one.

**Definition 1.5 (QCNM, informal).** *A scheme is quantum comparison-based non-malleable (QCNM) if no adversary can achieve a better than negligible advantage in distinguishing the real and ideal versions of the quantum comparison-based non-malleability experiment described above.*

We go on to show that QCNM is a consistent generalization of CNM.

**Theorem 1.6 (4.4, informal).** *When restricted to the post-quantum setting, QCNM and CNM are equivalent.*

Finally, we show that a QCNM scheme can be constructed from a CiNM scheme. The quantum-classical hybrid construction, which was extensively studied in [AGM18a] with respect to confidentiality and integrity, is obtained by encrypting every plaintext with a symmetric-key, one-time secure quantum encryption scheme and a fresh key, and then encrypting that key with a non-malleable classical public-key encryption scheme and appending it to the ciphertext.

**Theorem 1.7 (4.6, informal).** *Using a CNM classical scheme and a CiNM quantum scheme it is possible to construct a QCNM scheme via quantum-classical hybrid encryption.*

---

<sup>6</sup> Here we have to undo encryption in a different way as decryption is an irreversible process in general, see Section 4 for details.

## 2 Preliminaries

In this section, we introduce the notation and conventions used and provide a very brief overview of background material. For a more general overview of quantum computing see, for example, [Wat18].

### 2.1 Conventions and Notation

The adjoint of a complex matrix  $M$  is denoted by  $M^\dagger$  and its trace as  $\text{Tr}[M]$ . All Hilbert spaces  $\mathcal{H}_A$  in this work have dimension  $|A| := \dim(\mathcal{H}_A) = 2^m$  for some  $m \in \mathbb{N}$ . For Hilbert spaces  $\mathcal{H}_A$ , and  $\mathcal{H}_B$ , we write  $\mathbb{1}^A$  for the identity matrix on  $\mathcal{H}_A$ , or  $\mathbb{1}$  if the space is clear from context, and  $\mathbf{0}^{A \rightarrow B}$  or  $\mathbf{0}^A$  for the all-zero matrix of dimension  $|A| \times |B|$  or  $|A| \times |A|$  respectively. We denote the set of square matrices that act on  $\mathcal{H}_A$  as  $\mathcal{B}(\mathcal{H}_A)$ . We call a function  $\varepsilon(n)$  negligible (denoted  $\varepsilon \leq \text{negl}(n)$ ) if for every polynomial  $p$  there exists  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$  it holds that  $\varepsilon(n) < \frac{1}{p(n)}$ . Furthermore we use  $\log(x)$  to denote the base-2 logarithm of  $x$ .

### 2.2 Quantum States and Operations

We use bra-ket notation to denote a norm-1 vector  $|\phi\rangle \in \mathcal{H}_A$ , sometimes denoted  $|\phi\rangle^A$  for clarity. The set  $\{|x\rangle^A \mid x \in \{0, 1\}^n\}$  forms a basis of  $\mathcal{H}_A$  with  $|A| = 2^n$ , which is called the *computational basis*. Quantum states are described by *density matrices*, which are positive semi-definite Hermitian matrices with trace 1. The set of density matrices on  $\mathcal{H}_A$  is denoted by  $\mathcal{D}(\mathcal{H}_A)$ . The *maximally mixed state* is defined as  $\tau^A = \frac{\mathbb{1}^A}{|A|}$ . Furthermore we use  $\phi^{+AA'} = |\phi^+\rangle\langle\phi^+|^{AA'}$  to denote the (standard) *maximally entangled state*, where  $|\phi^+\rangle^{AA'} = \frac{1}{\sqrt{|A|}} \sum_{x \in \{0, 1\}^{\log(|A|)}} |xx\rangle$ .

A quantum state can be stored in a *quantum register*, which can be thought of as the quantum equivalent of a variable. A register  $A$  can store a density matrix  $\rho \in \mathcal{D}(\mathcal{H}_A)$ . In a cryptographic setting a “register”  $A$  is often an infinite family of registers, one for each value of the security parameter. The action of a quantum algorithm can be described as a completely positive trace-preserving (CPTP) map (a quantum channel). Sometimes the trace preserving property is relaxed to trace non-increasing, in which case we call it a CPTNI-map. If a quantum algorithm has a classical argument then it is understood that this argument is converted to the computational basis and classical outputs are obtained by measuring in the computational basis. We write  $\Lambda^{A \rightarrow B}$  to mean a CPTP map from register  $A$  to register  $B$ . When a quantum channel  $\Lambda^{A \rightarrow B}$  is evaluated on a state  $\rho^{AC}$ , then it implicitly acts as identity on register  $C$ , meaning  $\Lambda^{A \rightarrow B}(\rho^{AC}) = (\Lambda^{A \rightarrow B} \otimes \text{id}^C)(\rho^{AC})$ . To quantify the difference between quantum channels we will use the diamond norm, or completely bounded trace norm, defined as

$$\|\Lambda^{A \rightarrow B}\|_\diamond = \max_{\rho^{AA'}} \left\| (\Lambda \otimes \text{id}^{A'}) (\rho) \right\|_1,$$

where  $A'$  is a copy of the  $A$  register and  $\|M\|_1 = \text{Tr}[\sqrt{M^\dagger M}]$ . For a quantum state  $\sigma$ , we define the CPTP map  $\langle\sigma\rangle(\cdot) = \sigma \text{Tr}(\cdot)$ , i.e.  $\langle\sigma\rangle$  is the constant quantum channel that maps every input state to  $\sigma$ .

We write  $y \leftarrow A(x_1, \dots, x_n)$  to mean that  $y$  is the result of running an algorithm  $A$  on inputs  $x_1, \dots, x_n$ , and similarly  $Y \leftarrow A(X_1, \dots, X_n)$  to mean that register  $Y$  holds the state resulting from running the quantum algorithm  $A$  on input registers  $X_1, \dots, X_n$ . We write PPT to denote a uniform polynomial-time family of classical circuits and QPT to denote a uniform polynomial-time family of quantum circuits.

### 2.3 (Quantum) Encryption Schemes

We follow the conventions used in [AGM18b], in particular we use  $\text{Enc}_k = \text{Enc}(k, \cdot)$  and  $\text{Dec}_k = \text{Dec}(k, \cdot)$ . We begin by defining symmetric-key and public-key quantum encryption schemes.

**Definition 2.1.** A symmetric-key quantum encryption scheme (SKQES) is a triple  $(\text{KeyGen}, \text{Enc}, \text{Dec})$ , where

- $\text{KeyGen}$  is a PPT algorithm that given a security parameter  $n \in \mathbb{N}$  outputs a key  $k$ ,
- $\text{Enc}$  is a QPT algorithm which takes as input a classical key  $k$  and a quantum state in register  $M$  and outputs a quantum state in register  $C$ ,
- $\text{Dec}$  is a QPT algorithm which takes as input a classical key  $k$  and a quantum state in register  $C$  and outputs a quantum state in register  $M$  or  $|\perp\rangle\langle\perp|^\perp$ ,

such that  $\left\| \text{Dec}_k \circ \text{Enc}_k - \text{id}^{M \rightarrow M \oplus \perp} \right\|_\diamond \leq \text{negl}(n)$  for all  $k \leftarrow \text{KeyGen}(1^n)$ .

**Definition 2.2.** A public-key quantum encryption scheme (PKQES) is a triple  $(\text{KeyGen}, \text{Enc}, \text{Dec})$ , where

- $\text{KeyGen}$  is a PPT algorithm that given a security parameter  $n \in \mathbb{N}$  outputs a pair of keys  $(pk, sk)$ ,
- $\text{Enc}$  is a QPT algorithm which takes as input a classical public key  $pk$  and a quantum state in register  $M$  and outputs a quantum state in register  $C$ ,
- $\text{Dec}$  is a QPT algorithm which takes as input a classical secret key  $sk$  and a quantum state in register  $C$  and outputs a quantum state in register  $M$  or  $|\perp\rangle\langle\perp|^\perp$ ,

s.t.  $\left\| \text{Dec}_{sk} \circ \text{Enc}_{pk} - \text{id}^{M \rightarrow M \oplus \perp} \right\|_\diamond \leq \text{negl}(n)$  for all  $(pk, sk) \leftarrow \text{KeyGen}(1^n)$ .

It is implicit that  $|M| \leq |C| \leq 2^{q(n)}$  for some polynomial  $q$ . Furthermore we only consider *fixed-length* schemes, which means  $|M|$  is a fixed function of  $n$ . Lastly we adopt the convention that every honest party applies the measurement  $\{|\perp\rangle\langle\perp|, \mathbb{1} - |\perp\rangle\langle\perp|\}$  after running  $\text{Dec}$ , and denote with  $\text{Dec}_k(C) \neq \perp$  the event that this measurement did not measure  $|\perp\rangle\langle\perp|$  and thus produced a valid plaintext. Because of this convention we often state that the output space of  $\text{Dec}$  is  $\mathcal{D}(\mathcal{H}_M)$  although it is technically  $\mathcal{D}(\mathcal{H}_M \oplus \mathcal{H}_\perp)$ , where  $\mathcal{H}_\perp = \mathbb{C}|\perp\rangle$ .

**Theorem 2.3 (Lemma 1 in [AGM18a]).**

Let  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  be a SKQES, then  $\text{Enc}$  and  $\text{Dec}$  have the following form, for all  $k \leftarrow \text{KeyGen}$ :

$$\begin{aligned} \left\| \text{Enc}_k - V_k((\cdot)^M \otimes \sigma_k^T) V_k^\dagger \right\|_\diamond &\leq \varepsilon \\ \left\| \text{Dec}_k(V_k P_{\sigma_k}^T (V_k^\dagger(\cdot)^C V_k) P_{\sigma_k}^T V_k^\dagger) - \text{Tr}_T \left[ P_{\sigma_k}^T (V_k^\dagger(\cdot)^C V_k) P_{\sigma_k}^T \right] \right\|_\diamond &\leq \varepsilon. \end{aligned}$$

Here  $\sigma_k$  is a state on register  $T$ ,  $V_k$  is a unitary  $\varepsilon \leq \text{negl}(n)$ . Furthermore,  $P_{\sigma_k}$  is an orthogonal projectors such that  $\|P_{\sigma_k} \sigma_k P_{\sigma_k} - \sigma_k\|_\diamond \leq \varepsilon$  and  $\bar{P}_{\sigma_k} = \mathbf{1} - P_{\sigma_k}$ .

Furthermore, for every  $k$  there exists a probability distribution  $p_k$  and a family of quantum states  $|\psi_{k,r}\rangle^T$  such that  $\text{Enc}_k$  is  $\varepsilon$ -close to the following algorithm:

1. sample  $r \xleftarrow{p_k} \{0, 1\}^{\log |T|}$ ;
2. apply the map  $\text{Enc}_{k;r}(X^M) = V_k(X^M \otimes \psi_{k,r}^T) V_k^\dagger$ .

In this paper we will only consider schemes where all the actions described in Theorem 2.3 can be implemented by a PPT or QPT algorithm.

## 2.4 Security Definitions

In this paper we will build upon the classical definitions of non-malleability [BS99] and the existing quantum definitions of non-malleability [ABW09; AM17].

Experiment 1: CNM-Real	Experiment 2: CNM-Ideal
<b>Input</b> : $\Pi, \mathcal{A}, n$	<b>Input</b> : $\Pi, \mathcal{A}, n$
<b>Output</b> : $b \in \{0, 1\}$	<b>Output</b> : $b \in \{0, 1\}$
1 $(pk, sk) \leftarrow \text{KeyGen}(1^n)$	1 $(pk, sk) \leftarrow \text{KeyGen}(1^n)$
2 $(M, s) \leftarrow \mathcal{A}_1(pk)$	2 $(M, s) \leftarrow \mathcal{A}_1(pk)$
3 $x \leftarrow M$	3 $x, \tilde{x} \leftarrow M$
4 $y \leftarrow \text{Enc}_{pk}(x)$	4 $\tilde{y} \leftarrow \text{Enc}_{pk}(\tilde{x})$
5 $(R, \mathbf{y}) \leftarrow \mathcal{A}_2(s, y)$	5 $(R, \tilde{\mathbf{y}}) \leftarrow \mathcal{A}_2(s, \tilde{y})$
6 $\mathbf{x} \leftarrow \text{Dec}_{sk}(\mathbf{y})$	6 $\tilde{\mathbf{x}} \leftarrow \text{Dec}_{sk}(\tilde{\mathbf{y}})$
7 Output 1 iff $(y \notin \mathbf{y}) \wedge R(x, \mathbf{x})$	7 Output 1 iff $(\tilde{y} \notin \tilde{\mathbf{y}}) \wedge R(x, \tilde{\mathbf{x}})$

**Definition 2.4 (Definition 2 in [BS99] (CNM-CPA)).** A PKES  $\Pi$  is comparison-based non-malleable for chosen-plaintext attacks (CNM) if for any adversary  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$  it holds that

$$\Pr[\text{CNM-Real}(\Pi, \mathcal{A}, n) = 1] - \Pr[\text{CNM-Ideal}(\Pi, \mathcal{A}, n) = 1] \leq \text{negl}(n),$$

if  $\mathcal{A}$  is such that:

- $\mathcal{A}_1$  and  $\mathcal{A}_2$  are PPT
- $\mathcal{A}_1$  outputs a valid message space  $M$  which can be sampled by a PPT algorithm
- $\mathcal{A}_2$  outputs a relation  $R$  computable by a PPT algorithm



- $\mathcal{A}_2$  outputs a vector  $\mathbf{y}$  such that  $\perp \notin \text{Dec}_{sk}(\mathbf{y})$

For comparison-based non-malleability, we consider adversaries that are split into two stages, where each stage is a probabilistic algorithm. The first stage takes as input the public key and produces a message distribution, which is (a description of) a probabilistic algorithm that produces a plaintext. The second stage takes as input one ciphertext of a plaintext produced by this algorithm and produces a vector of ciphertexts and a relation  $R$ . The goal of the adversary is to construct  $R$  in such a way that  $R$  holds between the original plaintext and the (element-wise) decryption of the produced ciphertext vector, but not between another plaintext which is sampled independently from the message distribution and the decryption of this same vector. If an adversary can achieve this relation to hold with non-negligible probability, then intuitively the adversary was able to structurally change an encrypted message, which would indicate that the scheme is malleable.

In the existing literature on non-malleability in the quantum setting, the approach taken is quite different from the notion described above. Here, the focus is put on unconditional one-time security notions of symmetric-key non-malleability and authentication. In this setting, a notion of non-malleability was first introduced in [ABW09], which defines non-malleability as a condition on the effective map of an arbitrary attack. The *effective map* of an attack  $\Lambda_A^{CB \rightarrow C\hat{B}}$  is defined as  $\tilde{\Lambda}_A^{MB \rightarrow M\hat{B}} = \mathbb{E}_{k \leftarrow \text{KeyGen}(1^n)} \text{Dec}_k \circ \Lambda^A \circ \text{Enc}_k$ , and can be thought of as the average effect of an attack on the plaintext level.

The main idea of this definition is that a ciphertext cannot be meaningfully transformed into the ciphertext of another message, which means that the effective map of any attack is either identity, in case no transformation is applied, or a  $\langle \rho \rangle$  map, when the ciphertext is fully destroyed and replaced by another. Note that this way of defining non-malleability can also be satisfied by a scheme which has the property that an attacker can transform a ciphertext into another ciphertext of the same message. In other words, the non-malleability is only enforced on the plaintext level, which means it is a form of *plaintext non-malleability*. The classical notions discussed in the previous section do not allow for attacks that map an encrypted message to a different encryption of the same message. This restriction means non-malleability is enforced on the ciphertext level and thus these classical notions define forms of *ciphertext non-malleability*.

This effective-map-based way of describing non-malleability was continued in [AM17], where an insufficiency of the previous definition was demonstrated and a new definition was given. Their definition is given in terms of the mutual information between the plaintext and the side-information collected by the attacker. However, one of the results in their paper is a characterization theorem which we consider as the definition instead.

**Definition 2.5 (Theorem 4.4 in [AM17]).** A SKQES (KeyGen, Enc, Dec) is  $\varepsilon$ -non-malleable ( $\varepsilon$ -NM) if, for any attack  $\Lambda_A^{CB \rightarrow C\hat{B}}$ , its effective map

$\tilde{\Lambda}_A^{MB \rightarrow M\hat{B}}$  is such that

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_1^{B \rightarrow \hat{B}} + \frac{1}{|C|^2 - 1} (|C|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq \varepsilon,$$

where

$$\begin{aligned} \Lambda_1 &= \text{Tr}_{CC'} \left[ \phi^{+CC'} \Lambda_A(\phi^{+CC'} \otimes (\cdot)) \right] \quad \text{and} \\ \Lambda_2 &= \text{Tr}_{CC'} \left[ (\mathbf{1}^{CC'} - \phi^{+CC'}) \Lambda_A(\phi^{+CC'} \otimes (\cdot)) \right]. \end{aligned}$$

A SKQES is non-malleable (NM) if it is  $\varepsilon$ -NM for some  $\varepsilon \leq \text{negl}(n)$ .

In the symmetric-key setting, one can also consider the notion of authentication. A scheme satisfying this notion not only prevents an attacker from meaningfully transforming ciphertexts, but any attempt to do so can also be detected by the receiving party. In [DNS12] a definition is given for this notion, which we adapt slightly to use the diamond norm instead of the trace norm.

**Definition 2.6 (Definition 2.2 in [DNS12]).** A SKQES  $\Pi$  is  $\varepsilon$ -DNS authenticating ( $\varepsilon$ -DNS) if, for any attack  $\Lambda_A^{CB \rightarrow C\hat{B}}$ , its effective map  $\tilde{\Lambda}_A^{MB \rightarrow M\hat{B}}$  is such that

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_{acc}^{B \rightarrow \hat{B}} + \langle |\perp\rangle \langle \perp| \rangle \otimes \Lambda_{rej}^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq \varepsilon,$$

for some CPTNI maps  $\Lambda_{acc}, \Lambda_{rej}$  such that  $\Lambda_{acc} + \Lambda_{rej}$  is CPTP. A SKQES  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  is DNS authenticating (DNS) if it is  $\varepsilon$ -DNS for some  $\varepsilon \leq \text{negl}(n)$ .

It is shown in [AM17] that a NM scheme can be modified to a scheme that is DNS authenticating by appending a tag to the encoded plaintext.

### 3 Non-Malleability for Quantum SKE

While Definition 2.5 of NM presented in [AM17] has many desirable features, it turns out that it is slightly too strong in the sense that it rules out schemes that are clearly non-malleable intuitively. Furthermore, it has not been discussed in [AM17] whether NM actually ensures non-malleability of ciphertexts, or merely plaintext non-malleability. In this section, we will discuss these features of NM in detail. Furthermore, we propose a plaintext and a ciphertext version of NM, shedding light on how these different security properties are expressed in the effective-map formalism.

#### 3.1 Ciphertext Non-Malleability

When inspecting the Definition 2.5 of NM, one can observe that the constraints on  $\Lambda_1$  and  $\Lambda_2$  make NM a type of ciphertext non-malleability: Unless the adversary applies the identity channel, we end up in the case of  $\Lambda_2$ . However, the

use of  $\phi^+$  in defining these constraints can be considered problematic when the ciphertext space is not uniformly used, i.e. when  $\text{Enc}_K(\tau^M) \neq \tau^C$ . We provide an example of how this could be problematic.

*Example 3.1.* Let  $\Pi' = (\text{KeyGen}, \text{Enc}', \text{Dec}')$  be an NM SKQES, with ciphertext space  $\mathcal{H}_{C'}$ . Let  $\mathcal{H}_C = \mathcal{H}_{C'} \otimes \mathcal{H}_T$ , where  $\mathcal{H}_T = \mathbb{C}^2$ , then define  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  as follows, with ciphertext space  $\mathcal{H}_C$ :

- $\text{Enc}_k(X) = \text{Enc}'_k(X) \otimes |0\rangle\langle 0|^T$
- $\text{Dec}_k(Y) = \text{Dec}'_k(\text{Tr}_T[|0\rangle\langle 0|^T Y]) + \text{Tr}[|1\rangle\langle 1|^T Y] |\perp\rangle\langle \perp|$

Consider the attack  $\Lambda(\psi^C) = |0\rangle\langle 0|^T \psi |0\rangle\langle 0|^T + \langle \tau \rangle^{C'} (|1\rangle\langle 1|^T \psi |1\rangle\langle 1|^T)$  (with trivial register  $B$ ), which is the attack of measuring the  $T$  register in the computational basis and replacing the  $C'$  register with the maximally mixed state if the outcome of this measurement is 1 and doing nothing otherwise. As the register  $B$  is trivial,  $\Lambda_1$  is just a probability. We calculate

$$\begin{aligned} \Lambda_1 &= \text{Tr}[\phi^{+CC} \Lambda(\phi^{+CC})] \\ &= \text{Tr}\left[\phi^{+CC} \left(\frac{1}{2}|0\rangle\langle 0| \otimes |0\rangle\langle 0| \otimes \phi^{+C'C'} + \frac{1}{2}|1\rangle\langle 1| \otimes |1\rangle\langle 1| \otimes \tau^{C'} \otimes \tau^{C'}\right)\right] \\ &= \frac{|C'|^2 + 1}{4|C'|^2}. \end{aligned}$$

However the effective map is  $\tilde{\Lambda} = \text{id}$ , which shows that  $\Pi$  is not NM.

What could be considered problematic about this example is that any attack on  $\Pi$  is also an attack on  $\Pi'$ , since the attacker could add and remove the  $T$  register himself. Furthermore, there is a one-to-one correspondence between ciphertexts of  $\Pi$  and ciphertexts of  $\Pi'$ , because the  $T$  register is checked during decryption. This means that if an attacker could perform a malleability attack on  $\Pi$ , i.e. constructively transform a ciphertext into another ciphertext, then the attack obtained by applying the above strategy would be a malleability attack on  $\Pi'$ . Thus one could argue that, intuitively, non-malleability of  $\Pi'$  should imply non-malleability of  $\Pi$ . We suggest the following improved definition that prevents this behavior.

**Definition 3.2.** A SKQES  $(\text{KeyGen}, \text{Enc}, \text{Dec})$  is  $\varepsilon$ -ciphertext non-malleable ( $\varepsilon$ -CiNM) if, for any attack  $\Lambda_A^{CB \rightarrow C\hat{B}}$ , its effective map  $\tilde{\Lambda}_A^{MB \rightarrow M\hat{B}}$  is such that

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_1^{B \rightarrow \hat{B}} + \frac{1}{|C|^2 - 1} (|C|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq \varepsilon,$$

where

$$\begin{aligned} \Lambda_1 &= \mathbb{E}_{k,r} \left[ \text{Tr}_{CM'} \left[ \psi_{k,r}^{CM'} \Lambda_A(\psi_{k,r}^{CM'} \otimes (\cdot)) \right] \right] && \text{and} \\ \Lambda_2 &= \mathbb{E}_{k,r} \left[ \text{Tr}_{CM'} \left[ (\mathbb{1}^{CM'} - \psi_{k,r}^{CM'}) \Lambda_A(\psi_{k,r}^{CM'} \otimes (\cdot)) \right] \right]. \end{aligned}$$

Here  $\text{Enc}_{k;r}$  is as in Theorem 2.3,  $\mathbb{E}_{k,r}$  is taken uniformly over  $k$  and with  $r$  sampled according to  $p_k$  from Theorem 2.3, and  $\psi_{k,r}^{CM'} = \text{Enc}_{k;r}(\phi^{+MM'})$ . A SKQES is ciphertext non-malleable (CiNM) if it is  $\varepsilon$ -CiNM for some  $\varepsilon \leq \text{negl}(n)$ .

### 3.2 Plaintext Non-Malleability

For ciphertext non-malleability, discussed in the last section, the effective map approach seems slightly ill-suited: after all, the effective map is a map on plaintexts! What makes CiNM (and NM, albeit in an overzealous way) definitions of ciphertext non-malleability are the constraints placed on the map which  $\tilde{\Lambda}_A$  is compared with (the *simulator*). These constraints are imposed by the definitions of  $\Lambda_1$  and  $\Lambda_2$  and connect the simulator, which acts on plaintexts, to the attack map, which acts on ciphertexts. In order to construct a definition for plaintext non-malleability from NM, we therefore drop these constraints. In addition, we change the  $|C|^2$  constant for the constant  $|M|^2$ , as the former constant is a direct artifact of the constraints. In other words, plaintext-non-malleability “does not know about ciphertexts”, i.e., in particular, the ciphertext space dimension should be immaterial. We would like to remark that the latter point does not matter when talking about approximate non-malleability in the asymptotic setting, where the plaintext space grows polynomially with the security parameter.

The above considerations lead to the following definition.

**Definition 3.3.** A SKQES (KeyGen, Enc, Dec) is  $\varepsilon$ -plaintext non-malleable ( $\varepsilon$ -PNM) if, for any attack  $\Lambda_A^{CB \rightarrow C\hat{B}}$ , its effective map  $\tilde{\Lambda}_A^{MB \rightarrow M\hat{B}}$  is such that

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_1^{B \rightarrow \hat{B}} + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq \varepsilon,$$

where  $\Lambda_1$  and  $\Lambda_2$  are CPTNI and  $\Lambda_1 + \Lambda_2$  is CPTP. A SKQES is plaintext non-malleable (PNM) if it is  $\varepsilon$ -PNM for some  $\varepsilon \leq \text{negl}(n)$ .

Intuitively, ciphertext non-malleability is a strictly stronger security notion than plaintext non-malleability since the latter is obtained from the former by dropping the constraints on the simulator. This intuition holds true for our proposed PNM definition.

**Lemma 3.4.** Let  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  be an arbitrary SKQES and  $\Lambda_A^{CB \rightarrow C\hat{B}}$  an arbitrary attack on  $\Pi$  with effective map  $\tilde{\Lambda}_A^{MB \rightarrow M\hat{B}}$ . If there exist CPTNI  $\Lambda_1, \Lambda_2$ , such that  $\Lambda_1 + \Lambda_2$  is CPTP and it holds that

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_1^{B \rightarrow \hat{B}} + \frac{1}{|C|^2 - 1} (|C|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq \varepsilon,$$

then for any  $\alpha$  such that  $|M|^2 \leq \alpha \leq |C|^2$  there exist CPTNI  $\Lambda_3, \Lambda_4$  such that  $\Lambda_3 + \Lambda_4$  is CPTP and

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_3^{B \rightarrow \hat{B}} + \frac{1}{\alpha - 1} (\alpha \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_4^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq \varepsilon.$$

*Proof.* For fixed  $\Lambda_1$  and  $\Lambda_2$  one can obtain the statement by defining

$$\begin{aligned} \Lambda_3 &= \Lambda_1 + \left( 1 - \frac{(\alpha - 1)|C|^2}{\alpha(|C|^2 - 1)} \right) \Lambda_2 & \text{and} \\ \Lambda_4 &= \frac{(\alpha - 1)|C|^2}{\alpha(|C|^2 - 1)} \Lambda_2. \end{aligned}$$

The full proof of this lemma is rather technical and can be found in Appendix A.1.  $\square$

Lemma 3.4 shows that the  $|C|^2$  constant present in the NM definition can be decreased down to  $|M|^2$ , obtaining increasingly weaker security notions. This fact immediately implies the following

**Theorem 3.5.** *Any  $\varepsilon$ -NM or  $\varepsilon$ -CiNM SKQES is  $\varepsilon$ -PNM.*

*Proof.* This follows directly from Lemma 3.4 with  $\alpha = |M|^2$ .  $\square$

While PNM does not explicitly restrict the choice of  $\Lambda_1$  and  $\Lambda_2$ , an explicit form for  $\Lambda_i$  can be required without significantly strengthening the definition in the sense that the additional requirement only decreases security by at most a factor of 3.

**Theorem 3.6.** *Let  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  be an arbitrary  $\varepsilon$ -PNM SKQES for some  $\varepsilon$ , then for any attack  $\Lambda_A^{CB \rightarrow C\hat{B}}$ , its effective map  $\tilde{\Lambda}_A^{MB \rightarrow M\hat{B}}$  is such that*

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_1^{B \rightarrow \hat{B}} + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq 3\varepsilon,$$

where

$$\begin{aligned} \Lambda_1 &= \text{Tr}_{MM'} \left[ \phi^{+MM'} \tilde{\Lambda}_A(\phi^{+MM'} \otimes (\cdot)) \right] && \text{and} \\ \Lambda_2 &= \text{Tr}_{MM'} \left[ (\mathbf{1}^{MM'} - \phi^{+MM'}) \tilde{\Lambda}_A(\phi^{+MM'} \otimes (\cdot)) \right]. \end{aligned}$$

*Proof.* We sketch the proof here, the full proof of this theorem can be found in Appendix A.2. Let  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  be an arbitrary  $\varepsilon$ -PNM SKQES for some  $\varepsilon$  and let  $\Lambda_A^{CB \rightarrow C\hat{B}}$  be an arbitrary attack with effective map  $\tilde{\Lambda}_A^{MB \rightarrow M\hat{B}}$ . Furthermore, let  $\Lambda_1^{B \rightarrow \hat{B}}$  and  $\Lambda_2^{B \rightarrow \hat{B}}$  be such that  $\|\tilde{\Lambda}_A - \tilde{\Lambda}_{ideal}\|_{\diamond} \leq \varepsilon$ , where  $\tilde{\Lambda}_{ideal}^{MB \rightarrow M\hat{B}} = \text{id}^M \otimes \Lambda_1 + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau) \rangle - \text{id})^M \otimes \Lambda_2$ . Lastly, let

$$\begin{aligned} \Lambda_3 &= \text{Tr}_{MM'} \left[ \phi^{+MM'} \tilde{\Lambda}_A(\phi^{+MM'} \otimes (\cdot)) \right], \\ \Lambda_4 &= \text{Tr}_{MM'} \left[ (\mathbf{1}^{MM'} - \phi^{+MM'}) \tilde{\Lambda}_A(\phi^{+MM'} \otimes (\cdot)) \right], \\ \tilde{\Lambda}_{trace}^{MB \rightarrow M\hat{B}} &= \text{id}^M \otimes \Lambda_3 + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau) \rangle - \text{id})^M \otimes \Lambda_4, \\ \Lambda_5 &= \text{Tr}_{MM'} \left[ \phi^{+MM'} \tilde{\Lambda}_{ideal}(\phi^{+MM'} \otimes (\cdot)) \right], \text{ and} \\ \Lambda_6 &= \text{Tr}_{MM'} \left[ (\mathbf{1}^{MM'} - \phi^{+MM'}) \tilde{\Lambda}_{ideal}(\phi^{+MM'} \otimes (\cdot)) \right] \end{aligned}$$

Observe that  $\|\tilde{\Lambda}_{ideal} - \tilde{\Lambda}_{trace}\|_{\diamond} \leq \|(\Lambda_1 - \Lambda_3)\|_{\diamond} + \|(\Lambda_2 - \Lambda_4)\|_{\diamond}$ . Since  $\|(\tilde{\Lambda} - \tilde{\Lambda}_{ideal})(\phi^{+MM'} \otimes (\cdot))\|_{\diamond} \leq \|\tilde{\Lambda} - \tilde{\Lambda}_{ideal}\|_{\diamond} \leq \varepsilon$ , we have  $\|\Lambda_3 - \Lambda_5\|_{\diamond} \leq \varepsilon$

and  $\|A_4 - A_6\|_\diamond \leq \varepsilon$ . Using this we observe that

$$\begin{aligned} \left\| \tilde{A}_{ideal} - \tilde{A}_{trace} \right\|_\diamond &\leq \|A_1 - A_3\|_\diamond + \|A_2 - A_4\|_\diamond \\ &\leq \|A_1 - A_5\|_\diamond + \|A_5 - A_3\|_\diamond + \|A_2 - A_6\|_\diamond + \|A_6 - A_4\|_\diamond \\ &\leq 2\varepsilon + \|A_1 - A_5\|_\diamond + \|A_2 - A_6\|_\diamond. \end{aligned}$$

By substituting the definition of  $\tilde{A}_{ideal}$  we observe that  $A_5 = A_1$  and  $A_6 = A_2$ . From this we conclude

$$\begin{aligned} \left\| \tilde{A} - \tilde{A}_{trace} \right\|_\diamond &\leq \left\| \tilde{A} - \tilde{A}_{ideal} \right\|_\diamond + \left\| \tilde{A}_{ideal} - \tilde{A}_{trace} \right\|_\diamond \\ &\leq 3\varepsilon. \end{aligned}$$

□

**Theorem 3.7.** *There exists a PKQES  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  that is PNM but not NM and not CiNM.*

*Proof.* Let  $\Pi' = (\text{KeyGen}', \text{Enc}', \text{Dec}')$  be an arbitrary PKQES that is NM<sup>7</sup>. Then define  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  as

$$\begin{aligned} \text{KeyGen} &= \text{KeyGen}' \\ \text{Enc}_k &= \text{Enc}'_k \otimes |0\rangle\langle 0|^{R} \\ \text{Dec}_k &= \text{Dec}'_k \circ \text{Tr}_R, \end{aligned}$$

where  $R$  is an auxiliary 1-qubit register. Let  $A$  be an arbitrary attack on  $\Pi$  with effective map  $\tilde{A}$ , then define  $A' = \text{Tr}_R [A((\cdot) \otimes |0\rangle\langle 0|^{R})]$ , which is an attack on  $\Pi'$  with effective map  $\tilde{A}'$ . Observe that  $\tilde{A}' = \tilde{A}$ , since the  $R$  register is only added and then traced out. Because  $\Pi'$  is NM, we have  $A_3^{B \rightarrow \hat{B}}$  and  $A_4^{B \rightarrow \hat{B}}$  such that

$$\left\| \tilde{A}' - \text{id}^M \otimes A_3 + \frac{1}{|C|^2 - 1} (|C|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes A_4 \right\|_\diamond \leq \text{negl}(n)$$

It follows from Lemma 3.4 that  $\Pi$  is PNM.

Now consider the attack  $A_X = \text{id}^C \otimes (X(\cdot)X)^R \otimes \text{Tr}[\cdot]^B$ , where  $X$  is the Pauli  $X$  gate, with  $X|0\rangle = |1\rangle$  and  $X|1\rangle = |0\rangle$ . Let  $f(x_1 \dots x_n) = x_1 \dots x_{n-1}(1 - x_n)$ , i.e. the result of flipping the last bit of some bitstring. Observe that

$$\begin{aligned} A_X(\phi^{+CC'} \otimes (\cdot)^B) &= (\text{id} \otimes (X(\cdot)X)^R \otimes \text{Tr}[\cdot]^B) \left( \sum_{i,j} |ii\rangle\langle jj|^{CC'} \otimes (\cdot)^B \right) \\ &= \text{Tr}[\cdot]^B \sum_{i,j} |f(i)i\rangle\langle f(j)j|. \end{aligned}$$

<sup>7</sup> See [AM17] for such a scheme.

Since this superposition contains no components of the form  $|xx\rangle\langle xx|^{CC'}$  and  $\phi^{+CC'}$  only contains components of this form, we have that  $\phi^{+CC'} \Lambda_X(\phi^{+CC'} \otimes (\cdot)^B) = \mathbf{0}^{BCC' \rightarrow CC'}$ . With  $\psi_{k,r}^{CM'} = \text{Enc}_{k;r}(\phi^{+MM'}) = |0\rangle\langle 0| \otimes \text{Enc}'_{k;r}(\phi^{+MM'})$ , we have  $\psi_{k,r}^{CM'} \Lambda_X(\psi_{k,r}^{CM'} \otimes (\cdot)^B) = \mathbf{0}^{BCC' \rightarrow CC'}$ .

Also note that the effective map of  $\Lambda_X$  is  $\tilde{\Lambda}_X = \text{id}^M \otimes \text{Tr}[\cdot]^B$ , since the attack only acts on  $R$  and  $B$  and thus does not modify the message in  $M$ . Let  $\Lambda_1$  and  $\Lambda_2$  be as in Definition 2.5 or as in Definition 3.2, then  $\text{Tr}[\Lambda_1(\rho)] = 0$  for all  $\rho$ . It follows that  $\Lambda_2 = \text{Tr}_{CC'} \left[ (\mathbb{1} - \phi^{+CC'}) \Lambda_X(\phi^{+CC'} \otimes (\cdot)^B) \right] = \text{Tr}[\cdot]^B$ . Furthermore we have

$$\begin{aligned}
& \left\| \tilde{\Lambda}_X - \left( \text{id}^M \otimes \Lambda_1 + \frac{1}{|C|^2 - 1} (|C|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2 \right) \right\|_{\diamond} \\
&= \left\| \left( \text{id}^M \otimes \text{Tr}[\cdot]^B \right) - \left( \frac{1}{|C|^2 - 1} (|C|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \text{Tr}[\cdot]^B \right) \right\|_{\diamond} \\
&= \left\| \text{id}^M - \frac{1}{|C|^2 - 1} (|C|^2 \langle \text{Dec}_K(\tau) \rangle - \text{id})^M \right\|_{\diamond} \\
&\geq \left\| \phi^{+MM'} - \frac{1}{|C|^2 - 1} (|C|^2 \text{Dec}_K(\tau) \otimes \tau^{M'} - \phi^{+MM'}) \right\|_1 \\
&= 2 \max_{0 \leq P \leq 1} \text{Tr} \left[ P(\phi^{+MM'} - \frac{1}{|C|^2 - 1} (|C|^2 \text{Dec}_K(\tau) \otimes \tau^{M'} - \phi^{+MM'})) \right] \\
&\geq 2 \text{Tr} \left[ \phi^{+MM'} (\phi^{+MM'} - \frac{1}{|C|^2 - 1} (|C|^2 \text{Dec}_K(\tau) \otimes \tau^{M'} - \phi^{+MM'})) \right] \\
&= 2 - \frac{2(|C|^2 - |M|^2)}{|M|^2(|C|^2 - 1)} > 1,
\end{aligned}$$

where we use that  $\text{Tr} \left[ \phi^{+MM'} (\text{Dec}_K(\tau) \otimes \tau^{M'}) \right] = \frac{1}{|M|^2}$ , as is proven in the proof of Theorem A.2, and  $|M| \geq 2$ , which is true when we assume that we are encrypting at least one qubit. This shows that  $\Pi$  is not NM and not CiNM.  $\square$

While the above shows that PNM and CiNM are not the same in general, a special case arises when each plaintext has exactly one ciphertext (per key). Recall that plaintext non-malleability relaxes the constraints of ciphertext non-malleability by allowing the adversary to implement an attack that transforms one ciphertext into another, as long as both decrypt to the same plaintext. Thus in this special case, this relaxation is no relaxation at all. This special case arises in particular when an encryption scheme is *unitary*, meaning that  $\text{Enc}_k(X) = V_k X V_k^\dagger$  for some collection  $\{V_k\}_k$  of unitaries  $V_k \in U(\mathcal{H}_M)$ .

**Theorem 3.8.** *For any unitary SKQES  $\Pi$ ,  $\Pi$  is PNM iff  $\Pi$  is CiNM iff  $\Pi$  is NM.*

*Proof.* Since CiNM, NM  $\Rightarrow$  PNM for all SKQES, we only need to show the converse direction. Let  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  be a PNM unitary SKQES and  $\Lambda_A$  an

arbitrary attack on this scheme. By Theorem 3.6, we have, for some  $\varepsilon \leq \text{negl}(n)$ , that

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_1^{B \rightarrow \hat{B}} + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq 3\varepsilon,$$

where

$$\begin{aligned} \Lambda_1 &= \text{Tr}_{MM'} \left[ \phi^{+MM'} \tilde{\Lambda}_A(\phi^{+MM'} \otimes (\cdot)) \right] & \text{and} \\ \Lambda_2 &= \text{Tr}_{MM'} \left[ (\mathbf{1}^{MM'} - \phi^{+MM'}) \tilde{\Lambda}_A(\phi^{+MM'} \otimes (\cdot)) \right]. \end{aligned}$$

Let  $\{V_k^M\}_k$  be the collection such that  $\text{Enc}_k(X) = V_k X V_k^\dagger$  and note that  $\text{Enc}_{k;r} = \text{Enc}_k$  and  $C = M$ , where  $\text{Enc}_{k;r}$  is as in Theorem 2.3. Observe that

$$\begin{aligned} \Lambda_1 &= \text{Tr}_{MM'} \left[ \phi^{+MM'} \mathbb{E}_k [\text{Dec}_k(\Lambda_A(\text{Enc}_k(\phi^{+MM'} \otimes (\cdot))))] \right] \\ &= \text{Tr}_{MM'} \left[ \phi^{+MM'} \mathbb{E}_k \left[ V_k^\dagger (\Lambda_A(V_k(\phi^{+MM'} \otimes (\cdot)) V_k^\dagger)) V_k \right] \right] \\ &= \mathbb{E}_k \left[ \text{Tr}_{MM'} \left[ V_k \phi^{+MM'} V_k^\dagger (\Lambda_A(V_k(\phi^{+MM'} \otimes (\cdot)) V_k^\dagger)) \right] \right] \\ &= \mathbb{E}_{k,r} \left[ \text{Tr}_{CM'} \left[ \psi_{k,r}^{CM'} \Lambda_A(\psi_{k,r}^{CM'} \otimes (\cdot)) \right] \right], \end{aligned}$$

where  $\psi_{k,r} = V_k \phi^{+MM'} V_k^\dagger = \text{Enc}_{k;r}(\phi^{+MM'})$ . In the same way one can deduce that  $\Lambda_2 = \mathbb{E}_{k,r} \left[ \text{Tr}_{CM'} \left[ (\mathbf{1}^{CM'} - \psi_{k,r}^{CM'}) \Lambda_A(\psi_{k,r}^{CM'} \otimes (\cdot)) \right] \right]$ , and thus  $\Pi$  is CiNM. Similarly, we have

$$\begin{aligned} \Lambda_1 &= \text{Tr}_{MM'} \left[ \phi^{+MM'} \mathbb{E}_k \left[ V_k^\dagger (\Lambda_A(V_k(\phi^{+MM'} \otimes (\cdot)) V_k^\dagger)) V_k \right] \right] \\ &= \text{Tr}_{MM'} \left[ \phi^{+MM'} \mathbb{E}_k \left[ V_k^\dagger (\Lambda_A(V_k^{TM'}(\phi^{+MM'} \otimes (\cdot)) \bar{V}_k^{M'})) V_k \right] \right] \\ &= \text{Tr}_{MM'} \left[ \phi^{+MM'} \mathbb{E}_k \left[ (V_k^\dagger \otimes V_k^{TM'}) (\Lambda_A(\phi^{+MM'} \otimes (\cdot))) (V_k \otimes \bar{V}_k^{M'}) \right] \right] \\ &= \mathbb{E}_k \left[ \text{Tr}_{MM'} \left[ (V_k \otimes \bar{V}_k^{M'}) \phi^{+MM'} (V_k^\dagger \otimes V_k^{TM'}) (\Lambda_A(\phi^{+MM'} \otimes (\cdot))) \right] \right] \\ &= \text{Tr}_{MM'} \left[ \phi^{+MM'} (\Lambda_A(\phi^{+MM'} \otimes (\cdot))) \right], \end{aligned}$$

where we have used the ‘‘mirror lemma,’’  $A^M |\phi\rangle^{+MM'} = A^{TM'} |\phi\rangle^{+MM'}$  in the first and third equality, and  $(\cdot)^T$  is the transpose with respect to the computational basis and  $(\bar{\cdot})$  is the complex conjugate. In the same way one can deduce that

$$\Lambda_2 = \text{Tr}_{MM'} \left[ (\mathbf{1} - \phi^{+MM'}) (\Lambda_A(\phi^{+MM'} \otimes (\cdot))) \right] \text{ and thus } \Pi \text{ is NM.} \quad \square$$

We can use this equivalence to adopt results proven for NM in [AM17], particularly that the unitaries in a unitary encryption scheme form a unitary 2-design.



**Definition 3.9.** A family of unitary matrices  $D$  is an  $\varepsilon$ -approximate 2-design if

$$\left\| \frac{1}{|D|} \sum_{U \in D} (U \otimes U)(\cdot)(U^\dagger \otimes U^\dagger) - \int (U \otimes U)(\cdot)(U^\dagger \otimes U^\dagger) dU \right\|_\diamond \leq \varepsilon.$$

**Corollary 3.10.** Let  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  be a unitary SKQES such that  $\text{Enc}_k(\rho) = V_k \rho V_k^\dagger$  for some family of unitaries  $D = \{V_k\}_k$  and  $|M| = |C| = 2^n$ , then  $\Pi$  being PNM or CiNM is equivalent to  $D$  to being an approximate 2-design, in the sense that, for a sufficiently large constant  $r$ <sup>8</sup>,

1. If  $D$  is a  $\Omega(2^{-rn})$ -approximate 2-design then  $\Pi$  is  $2^{-\Omega(n)}$ -PNM and  $2^{-\Omega(n)}$ -CiNM.
2. If  $\Pi$  is  $\Omega(2^{-rn})$ -PNM or  $\Omega(2^{-rn})$ -CiNM, then  $D$  is a  $2^{-\Omega(n)}$ -approximate 2-design.

To provide additional evidence that PNM captures plaintext non-malleability for SKQES in a satisfactory way, we show that any PNM-secure scheme can be used to construct a plaintext-authenticating scheme in the sense of [DNS12], see Definition 2.6. The intuition behind DNS-authentication is that, after a possible attack, one can determine from a received plaintext whether or not an attack was performed, unless the attack did not change the underlying plaintext. For this reason, DNS-authentication is a notion of plaintext authentication. We use the fact that a PNM scheme protects a plaintext from modification to protect a tag register, which we then use to detect whether an attack was attempted. With this in mind, we first determine what state makes a good tag.

**Lemma 3.11.** For any SKQES  $(\text{KeyGen}, \text{Enc}, \text{Dec})$  and any  $m \in \mathbb{N}$  such that  $M = M'R$  for some registers  $M'$  and  $R$  with  $\log |R| = m$  there exists an  $x \in \{0, 1\}^m$  such that  $\text{Tr}[\langle x |^R \text{Dec}_K(\tau^C) |x \rangle^R] \leq \frac{1}{|R|}$ .

*Proof.* Observe that

$$\begin{aligned} \mathbb{E}_{x \in \{0, 1\}^m} [\text{Tr}[\langle x | \text{Dec}_K(\tau^C) |x \rangle]] &= \sum_{x \in \{0, 1\}^m} \frac{1}{2^m} \text{Tr}[\langle x | \text{Dec}_K(\tau^C) |x \rangle] \\ &= \frac{1}{2^m} \text{Tr}[\text{Dec}_K(\tau^C)] = \frac{1}{|R|} \end{aligned}$$

Since the expected value of  $\text{Tr}[\langle x | \text{Dec}_K(\tau^C) |x \rangle]$  is  $\frac{1}{|R|}$ , there must be at least one  $x$  such that  $\text{Tr}[\langle x | \text{Dec}_K(\tau^C) |x \rangle] \leq \frac{1}{|R|}$ .  $\square$

Lemma 3.11 allows us to find tags that have little overlap with  $\text{Dec}_K(\tau^C)$ , which means one can distinguish well between the case where the tag was left unharmed and the case where the ciphertext was depolarized. We use this property to build a scheme that is DNS authenticating.

<sup>8</sup> For the exact value of  $r$  and the constants hidden by the  $\Omega$ -s we refer to Theorem C.3 in [AM17] and Lemma 2.2.14 in [Low10]

**Theorem 3.12.** For any  $\varepsilon$ -PNM SKQES  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$ , there exists some  $x$  such that the scheme  $\Pi' = (\text{KeyGen}, \text{Enc}', \text{Dec}')$  is  $\left(\frac{3}{|R|} + \varepsilon\right)$ -DNS-authenticating, where

$$\begin{aligned}\text{Enc}'_k &= \text{Enc}_k((\cdot)^{M'} \otimes |x\rangle\langle x|^R) \\ \text{Dec}'_k &= \langle x|^R \text{Dec}_k(\cdot) |x\rangle^R + \text{Tr}[(\mathbb{1}^R - |x\rangle\langle x|^R) \text{Dec}_k(\cdot)] |\perp\rangle\langle\perp|\end{aligned}$$

*Proof.* We sketch the proof here, the full proof of this theorem can be found in Appendix A.3. Take  $x$  as in Lemma 3.11. Let  $\Lambda_A$  be an arbitrary attack map on  $\Pi'$ , then its effective map is

$$\tilde{\Lambda}'_A = \text{Dec}_{check} \circ \tilde{\Lambda}_A \circ \text{Enc}_{append},$$

where  $\tilde{\Lambda}_A$  is the effective map of  $\Lambda_A$  as an attack on  $\Pi$  and  $\text{Enc}_{append}$  and  $\text{Dec}_{check}$  are the channels that perform adding  $|x\rangle\langle x|$  to the plaintext during encryption and removing and checking of  $|x\rangle\langle x|$  during decryption respectively. Since  $\Pi$  is  $\varepsilon$ -PNM, there exist  $\Lambda_1, \Lambda_2$  such that

$$\left\| \tilde{\Lambda}_A - \text{id} \otimes \Lambda_1 + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2 \right\|_{\diamond} \leq \varepsilon.$$

Define  $\Lambda_{acc} = \Lambda_1$ ,  $\Lambda_{rej} = \Lambda_2$ , then

$$\left\| \tilde{\Lambda}'_A - \text{id} \otimes \Lambda_{acc} - \langle |\perp\rangle\langle\perp| \rangle \otimes \Lambda_{rej} \right\|_{\diamond} \leq \varepsilon + \frac{3}{|R|},$$

which means that  $\Pi'$  is  $\left(\frac{3}{|R|} + \varepsilon\right)$ -DNS authenticating.  $\square$

Note that  $|R|$  is a parameter of the scheme and any PNM scheme (with negligible  $\varepsilon$ ) can be made into a DNS scheme (with negligible  $\varepsilon$ ) by taking  $|R| = 2^n$ , i.e. taking  $R$  as  $n$  qubits.

## 4 Non-Malleability for Quantum PKE

### 4.1 Quantum Comparison-Based Non-Malleability

In this section, we will define a notion of many-time non-malleability for quantum public-key encryption, quantum comparison-based non-malleability (QCNM), as a quantum analog of the classical notion of comparison-based non-malleability (CNM, see Section 2) introduced in [BS99]. We first analyze CNM with the goal of finding appropriate quantum analogs of each of its components.

The message distribution  $M$  in the CNM definition allows an adversary to select messages that she thinks might produce ciphertexts that can be modified in a structural way. This choice is given because the total plaintext space is exponentially large, thus if one picks a message completely at random and only a few of them can be modified into related ciphertexts, then the winning probability is negligible despite the scheme being insecure. In the quantum representation of this message space we consider the following requirements:

1. As mentioned earlier, the quantum no-cloning theorem prevents copying the plaintext after sampling it for future reference. In order to check the relation in the last step of CNM, we require that two related states are produced, one of which will be kept by the challenger and the other encrypted and used by the adversary.
2. It must not be possible for the adversary to correlate herself with either of the produced messages. This is to prevent the adversary from influencing the second copy of the state later on. For example, consider the case where the adversary produces the state  $\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$ , where the first two qubits are the two copies of the message and the last is kept by the adversary. The adversary can then measure her qubit, collapsing the superposition and informing her in which (classical) state the second copy now is. This allows her to trivially construct a relation between her output and the second copy.

In order to satisfy requirement (1), we have chosen to represent  $M$  by a unitary  $U^{MRP}$  such that  $U|0\rangle$  is a purification of the message distribution, where the message resides in  $M$ , the second (reference) state in  $R$ , and  $P$  is used for the purification<sup>9</sup>. This purification register allows the adversary to implement any quantum channel on  $MR$ , with  $U$  being a Stinespring dilation of that channel. The first part of the quantum adversary,  $\mathcal{A}_1$ , produces this unitary in the form of a circuit along with some side information  $S$  to be passed on to the next stage. We denote this process by  $(U, S) \leftarrow \mathcal{A}_1(pk)$ .

For the QCNM definition we define two experiments, similar to the CNM definition. In the following we describe how the different elements of the CNM experiment are instantiated in the quantum case. The appropriate quantum notion of a relation  $R$  on plaintexts is given by a POVM element  $0 \leq E^{MR} \leq \mathbb{1}$ . The two registers  $MR$  are considered to contain related states if an application of the measurement  $\{E, \mathbb{1} - E\}$  returns the outcome corresponding to  $E$ . Of course, this POVM is provided by the adversary in form of a circuit and must hence be efficient. The quantum analogue of the vector  $\mathbf{y}$  is given by a collection of registers  $\mathbf{C} = C_1 \dots C_m$ , where  $m$  is at most polynomial in  $n$ , the security parameter of the considered scheme, and each  $C_i$  satisfies  $M_i T_i = C_i \cong C = MT$ . The quantum analogue of the vector  $\mathbf{x}$  is similarly given as  $\mathbf{M} = M_1 \dots M_m$ . Observe that any PKQES can also be seen as a SKQES, with keys of the form  $k = (pk, sk)$ , which allows us to use Theorem 2.3<sup>10</sup>. For any PKQES  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  with security parameter  $n$ , let  $\{V_k \mid k = (pk, sk) \leftarrow \text{KeyGen}(1^n)\}$ ,  $t = \log |T|$ ,  $\{\psi_{k,r} \mid k = (pk, sk) \leftarrow \text{KeyGen}(1^n), r \in \{0, 1\}^t\}$  and  $\{p_k \mid k = (pk, sk) \leftarrow \text{KeyGen}(1^n)\}$  be as in Theorem 2.3 in the QCNM experiments.

Lastly, we define the unitary  $U_{prep}$  combining the preparation of the message state and the encryption of its part in register  $M$ . This means a check similar to the  $y \notin \mathbf{y}$  check in the CNM experiments can be implemented by sequentially undoing  $U_{prep}$  on all combinations  $C_i RP$  and then measuring whether the result

<sup>9</sup> A purification is a quantum register that is similar to the “garbage” register in reversible computation.

<sup>10</sup> This characterization could also be invoked with  $k = pk$ , however the resulting encryption unitary is then (likely) not efficiently implementable.

is  $|0\rangle\langle 0|$ , which is only the case if  $C_i$  contained part of  $U_{prep}|0\rangle$ , which is the original ciphertext given to the adversary.

We are now ready to define the real and ideal experiments for quantum comparison-based non-malleability.

---

**Experiment 3: QCNM-Real**

---

**Input** :  $\Pi, \mathcal{A}, n$   
**Output**:  $b \in \{0, 1\}$

- 1  $k = (pk, sk) \leftarrow \text{KeyGen}(1^n)$
- 2  $(U^{MRP}, S) \leftarrow \mathcal{A}_1(pk)$
- 3  $r \xleftarrow{pk} \{0, 1\}^t$
- 4 Construct  $U_\psi^T$  such that  $U_\psi^T|0\rangle^T = |\psi_{k,r}\rangle^T$
- 5 Construct  $U_{prep}^{MTRP} = V_k^{MT}(U^{MRP} \otimes U_\psi^T)$
- 6 Prepare  $U_{prep}|0\rangle\langle 0|U_{prep}^\dagger$  in  $MTRP$
- 7  $(\mathbf{C}, E) \leftarrow \mathcal{A}_2(MT, S)$
- 8 **for**  $i = 1, \dots, |\mathbf{C}|$  **do**
- 9     Perform  $U_{prep}^\dagger$  on  $C_iRP$
- 10    Measure  $\{|0\rangle\langle 0|, \mathbb{1} - |0\rangle\langle 0|\}$  on  $C_iRP$  with outcome  $b$
- 11    **if**  $b = 0$  **then**
- 12     Output 0
- 13    Perform  $U_{prep}$  on  $C_iRP$
- 14  $\mathbf{M} \leftarrow \text{Dec}_{sk}(\mathbf{C})$
- 15  $\{E, \mathbb{1} - E\}$  on  $R\mathbf{M}$  with outcome  $e$
- 16 Output  $e$

---



---

**Experiment 4: QCNM-Ideal**

---

**Input** :  $\Pi, \mathcal{A}, n$   
**Output**:  $b \in \{0, 1\}$

- 1 Run lines 1-14 of Experiment QCNM-Real
- 15 Prepare  $U|0\rangle\langle 0|U^\dagger$  in  $\tilde{M}\tilde{R}\tilde{P}$
- 16  $\{E, \mathbb{1} - E\}$  on  $\tilde{R}\mathbf{M}$  with outcome  $e$
- 17 Output  $e$

---

A PKQES is now defined to be QCNM-secure, if no adversary can achieve higher success probability in the experiment QCNM-Real than in QCNM-Ideal.

**Definition 4.1.** A PKQES  $\Pi$  is quantum comparison-based non-malleable (QCNM) if for any QPT adversary  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$  it holds that

$$\Pr[\text{QCNM-Real}(\Pi, \mathcal{A}, n) = 1] - \Pr[\text{QCNM-Ideal}(\Pi, \mathcal{A}, n) = 1] \leq \text{negl}(n),$$

if  $\mathcal{A}$  such that:

- $\mathcal{A}_1$  outputs a valid unitary  $U$  which can be implemented by a QPT algorithm,
- $\mathcal{A}_2$  outputs a POVM element  $E$  which can be implemented by a QPT algorithm,
- $\mathcal{A}_2$  outputs a vector of registers  $\mathbf{C}$  such that  $\perp \notin \text{Dec}_{sk}(\mathbf{C})$ .

## 4.2 Relation Between QCNM and CNM

In order to compare QCNM to CNM, we consider both definitions modified for quantum adversaries and encryption schemes that have classical input and output but can perform quantum computation. In the case that a quantum state is sent to such a post-quantum algorithm, it is first measured in the computational basis to obtain a classical input.

---

### Experiment 5: QCNM-Real<sub>PQ</sub>

---

**Input** :  $\Pi, \mathcal{A}, n$   
**Output**:  $b \in \{0, 1\}$

- 1  $k = (pk, sk) \leftarrow \text{KeyGen}(1^n)$
- 2  $(U^{MRP}, S) \leftarrow \mathcal{A}_1(pk)$
- 3  $r \xleftarrow{pk} \{0, 1\}^t$
- 4 Construct  $U_\psi^T$  such that  $U_\psi^T |0\rangle^T = |\psi_{k,r}\rangle^T$
- 5 Construct  $U_{prep}^{MTRP} = V_k^{MT}(U^{MRP} \otimes U_\psi^T)$
- 6 Prepare  $U_{prep}|0\rangle\langle 0|U_{prep}^\dagger$  in  $MTRP$
- 7 Measure  $MTR$  in the computational basis with outcome  $y^{MT} z^R$
- 8  $(\mathbf{C}, E) \leftarrow \mathcal{A}_2(MT, S)$
- 9 **for**  $i = 1, \dots, |\mathbf{C}|$  **do**
- 10     Measure  $\{|y\rangle\langle y|, \mathbb{1} - |y\rangle\langle y|\}$  on  $C_i$  with outcome  $b$
- 11     **if**  $b = y$  **then**
- 12         Output 0
- 13  $\mathbf{M} \leftarrow \text{Dec}_{sk}(\mathbf{C})$
- 14  $\{E, \mathbb{1} - E\}$  on  $R\mathbf{M}$  with outcome  $e$
- 15 Output  $e$

---



---

### Experiment 6: QCNM-Ideal<sub>PQ</sub>

---

**Input** :  $\Pi, \mathcal{A}, n$   
**Output**:  $b \in \{0, 1\}$

- 1 Run lines 1-13 of Experiment QCNM-Ideal<sub>PQ</sub>
- 2 Prepare  $U|0\rangle\langle 0|U^\dagger$  in  $\tilde{M}\tilde{R}\tilde{P}$
- 3 Measure  $\tilde{M}\tilde{R}$  in the computational basis
- 4  $\{E, \mathbb{1} - E\}$  on  $\tilde{R}\tilde{\mathbf{M}}$  with outcome  $e$
- 5 Output  $e$

---

We consider the above experiments to be the post-quantum version of the QCNM experiments. The main modification is the measurement in Step 7, which enforces the requirement that  $\mathcal{A}_2$  only takes classical input. The modification of Steps 9 through 12 is made because the measurement in Step 7 disturbs the state in an irreversible fashion, thus performing  $U_{prep}^\dagger$  no longer inverts the sampling/encryption process. Lastly, in the Ideal setting Step 15 is added to mimic the effect that Step 7 would have on  $M'$ .

**Definition 4.2.** A PKQES  $\Pi$  is post-quantum comparison-based non-malleable (QCNM<sub>PQ</sub>) if for any adversary  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$  it holds that

$$\Pr [\text{QCNM-Real}_{PQ}(\Pi, \mathcal{A}, n) = 1] - \Pr [\text{QCNM-Ideal}_{PQ}(\Pi, \mathcal{A}, n) = 1] \leq \text{negl}(n),$$

if  $\mathcal{A}$  and  $\Pi$  are such that:

- $\mathcal{A}_1$  and  $\mathcal{A}_2$  are QPT and output only classical states,
- $\mathcal{A}_1$  outputs a valid unitary  $U$  which can be implemented by a QPT algorithm,
- $\mathcal{A}_2$  outputs a POVM element  $E$  which implementable by a QPT algorithm, and
- $\mathcal{A}_2$  outputs a vector of registers  $\mathbf{C}$  such that  $\perp \notin \text{Dec}_{sk}(\mathbf{C})$ .

Sampling of the message by the challenger is now done by not only applying  $U$  to  $|0\rangle$ , but in addition also measuring in the computational basis. Similarly, we define a post-quantum version of CNM.

**Definition 4.3.** A PKQES  $\Pi$  is comparison-based non-malleable against post-quantum adversaries (CNM<sub>PQ</sub>) if for any QPT adversary  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$  it holds that

$$\Pr [\text{CNM-Real}(\Pi, \mathcal{A}, n) = 1] - \Pr [\text{CNM-Ideal}(\Pi, \mathcal{A}, n) = 1] \leq \text{negl}(n),$$

if  $\Pi$  and  $\mathcal{A}$  are such that:

- $\mathcal{A}_1$  and  $\mathcal{A}_2$  are QPT and output classical strings,
- $\mathcal{A}_1$  outputs a valid QPT algorithm  $M$  which produces classical strings,
- $\mathcal{A}_2$  outputs a QPT algorithm  $R$ ,
- $\mathcal{A}_2$  outputs a vector  $\mathbf{y}$  such that  $\perp \notin \text{Dec}_{sk}(\mathbf{y})$ .

The only difference between CNM and CNM<sub>PQ</sub> is that the latter allows the encryption scheme, adversary and any algorithms produced by the adversary to use a quantum computer. Furthermore, the relation  $R$  has become probabilistic, but since it is used only once there is no difference between using a probabilistic relation or picking a deterministic relation at random. Observe that CNM<sub>PQ</sub> is simply a stronger requirement than CNM since it requires security against a strict superset of adversaries, and thus trivially implies CNM.

**Theorem 4.4.** A PKQES  $\Pi$  is QCNM<sub>PQ</sub> if and only if  $\Pi$  is CNM<sub>PQ</sub>.

*Proof.* For the  $\Rightarrow$  direction, let  $\Pi$  be an arbitrary PKQES which is QCNM<sub>PQ</sub>-secure and let  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$  be an arbitrary quantum adversary intended to perform the CNM<sub>PQ</sub> experiments. Assume that  $\Pi$  is such that Enc and Dec take only classical input and produce only classical output. Define  $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$  as follows:

$\mathcal{B}_1(pk)$ :

- 
- 1  $(M, s) \leftarrow \mathcal{A}_1(pk)$
  - 2 Let  $p_M(x)$  be the probability that  $x \leftarrow M$ , then construct  $U$  such that  $U|0\rangle^{MM'P} = \frac{1}{|R|} \sum_{r \in R} |M(r)M(r)r\rangle = \sum_{x \leftarrow M} \sqrt{p_M(x)} |xx\phi_x\rangle^{MM'P}$ , where  $M' \cong M$  is the reference register,  $R$  is the set of possible input for  $M$  and  $\phi_x$  is the uniform superposition over all  $|r\rangle$  such that  $x \leftarrow M(r)$ .
  - 3 Output  $(U, |s\rangle\langle s|)$
-

$\mathcal{B}_2(|s\rangle\langle s|^S, |y\rangle\langle y|^{MT})$ :

- 
- 1  $(R, \mathbf{y}) \leftarrow \mathcal{A}_2(y, s)$
  - 2 Construct  $E = \sum_{i, \mathbf{j}} R(i, \mathbf{j}) |i\mathbf{j}\rangle\langle i\mathbf{j}|$
  - 3 Output  $(E, |\mathbf{y}\rangle\langle \mathbf{y}|^{C_1 \dots C_m})$
- 

Observe that the definition of  $\text{QCNM-Real}_{PQ}(II, \mathcal{B}, n)$ , after some simplification, yields

- 
- 1  $k = (pk, sk) \leftarrow \text{KeyGen}(1^n)$
  - 2  $(M, s) \leftarrow \mathcal{A}_1(pk)$
  - 3 Let  $p_M(x)$  be the probability that  $x \leftarrow M$ , then construct  $U$  such that
 
$$U|0\rangle^{MM'P} = \sum_{x \leftarrow M} \sqrt{p_M(x)} |xx\phi_x\rangle^{MM'P}$$
  - 4  $r \xleftarrow{pk} \{0, 1\}^t$
  - 5 Construct  $U_\psi^T$  such that  $U_\psi^T |0\rangle^T = |\psi_{k,r}\rangle^T$
  - 6 Construct  $U_{prep}^{MTM'P} = V_k^{MT}(U^{MM'P} \otimes U_\psi^T)$
  - 7 Prepare  $U_{prep}|0\rangle\langle 0|U_{prep}^\dagger$  in  $MTM'P$
  - 8 Measure  $MTM'$  in the computational basis with outcome  $yz$
  - 9  $(R, \mathbf{y}) \leftarrow \mathcal{A}_2(y, s)$
  - 10 Construct  $E = \sum_{i, \mathbf{j}} R(i, \mathbf{j}) |i\mathbf{j}\rangle\langle i\mathbf{j}|$
  - 11 Prepare  $|\mathbf{y}\rangle\langle \mathbf{y}|$  in  $\mathbf{C}$
  - 12 **for**  $i = 1, \dots, |\mathbf{C}|$  **do**
  - 13     Measure  $\{|y\rangle\langle y|, \mathbb{1} - |y\rangle\langle y|\}$  on  $C_i$  with outcome  $b$
  - 14     **if**  $b = y$  **then**
  - 15         Output 0
  - 16  $\mathbf{M} \leftarrow \text{Dec}_{sk}(\mathbf{C})$
  - 17  $\{E, \mathbb{1} - E\}$  on  $M'M$  with outcome  $e$
  - 18 Output  $e$
- 

Here Steps 3,5,6,7 and 8 together simply execute  $x \leftarrow M; y \leftarrow \text{Enc}_{k;r}(x)$ . Furthermore, if  $y \in \mathbf{y}$  then some  $C_i$  contains  $|y\rangle\langle y|$ , which will guarantee the output to be  $y$  in Step 13. Conversely if  $y \notin \mathbf{y}$ , then all  $C_i$  contain some state orthogonal to  $|y\rangle\langle y|$  and thus Step 13 has 0 probability of outputting  $y$  in this case, thus Step 13 effectively implements the  $y \notin \mathbf{y}$  check. Lastly, note that  $E$  is a projective measurement which projects onto the space spanned by all  $|i\mathbf{j}\rangle$  such that  $R(i, \mathbf{j})$ , which means that Step 17 outputs 1 iff  $R(x, \mathbf{x})$ , where  $x$  is stored in  $M'$  and  $\mathbf{x}$  in  $\mathbf{M}$ . We conclude that  $\text{QCNM-Real}_{PQ}(II, \mathcal{B}, n)$  produces the same random variable as  $\text{CNM-Real}(II, \mathcal{A}, n)$ . By similar reasoning the same is true for the **Ideal** case, with the additional observation that preparing  $U|0\rangle$  in  $\tilde{M}\tilde{M}'\tilde{P}$  and measuring  $\tilde{M}$  in the computational basis with result  $\tilde{x}$  is equivalent to  $\tilde{x} \leftarrow M$  and collapses  $\tilde{M}'$  to  $\tilde{x}$ . It follows that  $II$  is  $\text{CNM}_{PQ}$ .

For the  $\Leftarrow$  direction, let  $II$  be an arbitrary PKQES fulfilling  $\text{CNM}_{PQ}$  and let  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$  be an arbitrary classical adversary on this scheme intended to perform the  $\text{QCNM}_{PQ}$  experiments. Define  $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$  as follows:

$\mathcal{B}_1(pk)$ :

- 
- 1  $(U, |s\rangle\langle s|^S) \leftarrow \mathcal{A}_1(pk)$
  - 2 Prepare  $U|0\rangle$  twice, in  $M_0R_0P_0$  and  $M_1R_1P_1$
  - 3 Measure  $M_0R_0M_1R_1$  in the computational basis with outcome  $m_0z_0m_1z_1$
  - 4 Construct  $M$  to be the uniform distribution over  $\{m_0, m_1\}$
  - 5 Output  $(M, sm_0z_0m_1z_1)$
- 

$\mathcal{B}_2(s', y)$ :

- 
- 1 Parse  $s'$  as  $sm_0z_0m_1z_1$
  - 2  $(E, \mathbf{y}) \leftarrow \mathcal{A}_2(|s\rangle\langle s|^S, |y\rangle\langle y|^{MT})$
  - 3 Construct  $R(x, \mathbf{x})$  to be
  - 4     Find  $i$  such that  $m_i = x$
  - 5     **prepare**  $|z_i\mathbf{x}\rangle\langle z_i\mathbf{x}|$  in  $R'\mathbf{M}$
  - 6     **measure**  $\{E, \mathbb{1} - E\}$  on  $R'\mathbf{M}$ , **output** 1 iff the outcome is  $E$
  - 7 Output  $(R, |\mathbf{y}\rangle\langle \mathbf{y}|)$
- 

Observe that the definition of  $\text{CNM-Real}_{PQ}(\Pi, \mathcal{B}, n)$ , after some simplification, yields

- 
- 1  $k = (pk, sk) \leftarrow \text{KeyGen}(1^n)$
  - 2  $(U, |s\rangle\langle s|^S) \leftarrow \mathcal{A}_1(pk)$
  - 3 Prepare  $U|0\rangle$  in  $M_0R_0P_0$
  - 4 Prepare  $U|0\rangle$  in  $M_1R_1P_1$
  - 5 Measure  $M_0R_0M_1R_1$  in the computational basis with outcome  $m_0z_0m_1z_1$
  - 6 Pick  $i \leftarrow \{0, 1\}$
  - 7  $y \leftarrow \text{Enc}_{pk}(m_i)$
  - 8  $(E, |\mathbf{y}\rangle\langle \mathbf{y}|) \leftarrow \mathcal{A}_2(|s\rangle\langle s|^S, |y\rangle\langle y|^{MT})$
  - 9  $\mathbf{x} \leftarrow \text{Dec}_{sk}(\mathbf{y})$
  - 10 **if**  $y \in \mathbf{y}$  **then**
  - 11     Output 0
  - 12 Find  $j$  such that  $m_j = m_i$
  - 13 Prepare  $|z_j\mathbf{x}\rangle\langle z_j\mathbf{x}|$  in  $R\mathbf{M}$
  - 14  $\{E, \mathbb{1} - E\}$  on  $R\mathbf{M}$  with outcome  $e$
  - 15 Output  $e$
- 

Similarly, the  $\text{CNM-Ideal}_{PQ}(\Pi, \mathcal{B}, n)$  yields the same Experiment except with line 12 replaced with “Pick  $j \leftarrow \{0, 1\}$ ”. Note that Step 7, the encrypting, is not performed by  $U_{prep}$  but simply by  $\text{Enc}$  and that Step 10 simply checks  $y \in \mathbf{y}$  instead of loop that we earlier argued to be equivalent. Additionally the measurement in Step 3 and 4 are equivalent to measuring the ciphertext after encryption (as is done in  $\text{QCNM}$ ), because it is assumed that encryption, and thus  $V_k$ , maps classical states to classical states.

Note that w.l.o.g. we can assume that  $m_0 \neq m_1$ , since if this is not the case then the Real and Ideal case are equivalent and thus the adversary has no hope of winning. This makes that the  $\text{CNM-Real}_{PQ}(\Pi, \mathcal{B}, n)$  and  $\text{QCNM-Real}_{PQ}(\Pi, \mathcal{A}, n)$  are equivalent given the observations in the previous paragraph. Furthermore,



when  $i = j$  in the CNM-Ideal case then it is equivalent to the CNM-Real case. When  $i \neq j$ , the  $\text{CNM-Ideal}_{PQ}(II, \mathcal{B}, n)$  and  $\text{QCNM-Ideal}_{PQ}(II, \mathcal{A}, n)$  experiments are equivalent. Thus the advantage of  $\mathcal{B}$  in CNM is half the advantage of  $\mathcal{A}$  in QCNM, which implies that  $II$  is  $\text{QCNM}_{PQ}$ .  $\square$

Note that we argued earlier that, for any PKES, being  $\text{CNM}_{PQ}$  trivially implies being CNM, thus we derive the following corollary.

**Corollary 4.5.** *Any  $\text{QCNM}_{PQ}$  PKES is CNM.*

### 4.3 A QCNM Secure Scheme

In this section we show how QCNM-security can be achieved using a quantum-classical hybrid construction like the ones used in [AGM18b; AGM18a]. The idea is similar to the classical technique of hybrid encryption. We construct a quantum-non-malleable PKQES by encrypting each plaintext with a quantum one-time non-malleable scheme and encrypting the key using a classical non-malleable PKES. We begin by defining the general quantum-classical hybrid construction.

**Construction 1** *Let  $II^{\text{Qu}} = (\text{KeyGen}^{\text{Qu}}, \text{Enc}^{\text{Qu}}, \text{Dec}^{\text{Qu}})$  be a SKQES and  $II^{\text{Cl}} = (\text{KeyGen}^{\text{Cl}}, \text{Enc}^{\text{Cl}}, \text{Dec}^{\text{Cl}})$  a PKES. We define the hybrid scheme  $II^{\text{Hyb}}[II^{\text{Qu}}, II^{\text{Cl}}] = (\text{KeyGen}^{\text{Hyb}}, \text{Enc}^{\text{Hyb}}, \text{Dec}^{\text{Hyb}})$  as follows. We set  $\text{KeyGen}^{\text{Hyb}} = \text{KeyGen}^{\text{Cl}}$ . The encryption algorithm  $\text{Enc}_{\text{pk}}^{\text{Hyb}}$ , on input  $X$ ,*

1. *generates a key  $k \leftarrow \text{KeyGen}^{\text{Qu}}(1^{n(\text{pk})})$ , and*
2. *outputs the pair  $(\text{Enc}_k^{\text{Qu}}(X), \text{Enc}_{\text{pk}}^{\text{Cl}}(k))$ .*

*Decryption is done in the obvious way, by first decrypting the second part of the ciphertext using  $\text{Dec}^{\text{Cl}}$  to obtain the one-time key  $k'$ , and then decrypting the first part using  $\text{Dec}_{k'}^{\text{Qu}}$ .*

We continue by proving that if  $II^{\text{Qu}}$  is unitary and secure according to NM, CiNM or PNM (they are all equivalent for unitary SKQES according to Theorem 3.8), and  $II^{\text{Cl}}$  to be CNM, then  $II^{\text{Hyb}}[II^{\text{Qu}}, II^{\text{Cl}}]$  is QCNM.

**Theorem 4.6.** *Let  $II^{\text{Qu}} = (\text{KeyGen}^{\text{Qu}}, \text{Enc}^{\text{Qu}}, \text{Dec}^{\text{Qu}})$  be a NM secure SKQES with unitary encryption and decryption map, and  $II^{\text{Cl}} = (\text{KeyGen}^{\text{Cl}}, \text{Enc}^{\text{Cl}}, \text{Dec}^{\text{Cl}})$  a postquantum-CNM secure PKES. Then  $II^{\text{Hyb}}[II^{\text{Qu}}, II^{\text{Cl}}]$  is QCNM.*

*Proof.* We begin by defining modified versions of the two experiments used in defining QCNM, sckQCNM-Real and sckQCNM-Ideal (for spoofed classical key). These two experiments are defined exactly as the experiments QCNM-Real and QCNM-Ideal, except for the following modifications:

1. When creating the ciphertext register  $C$  that is handed to the adversary, its classical part  $c$  is produced by encrypting a fresh, independently sampled one-time key  $k' \leftarrow \text{KeyGen}^{\text{Qu}}$ . The pair  $(c, k)$  is stored ( $k$  being the key used for encryption with  $\text{Enc}^{\text{Qu}}$ .)

2. The test whether the ciphertext was modified by the adversary is done by first checking whether the classical part  $c'$  is equal to  $c$ . If it is not, the ciphertext was modified and no further test of the quantum part is necessary. If  $c' = c$ , the modification check from the games QCNM-Real and QCNM-Ideal is applied, using the stored one-time key  $k$ . Note that this is equivalent to the check mandated for the QCNM experiments.
3. Before decrypting any ciphertext, the challenger checks whether its classical part is equal to  $c$ . If not, he proceeds with decryption, otherwise, he just decrypts the quantum ciphertext with  $\text{Dec}_k^{\text{Qu}}$ .

Let  $\mathcal{A}$  be a QCNM-adversary against  $\Pi^{\text{Hyb}}$ . Recall that it was proven in [BS99] that CNM is equivalent to IND-parCCA2, indistinguishability under parallel chosen ciphertext attacks. In this attack model, after receiving the challenge ciphertext, the adversary is allowed to submit one tuple of ciphertexts that is decrypted in case none of them is equal to the challenge ciphertext. Define the following IND-parCCA2 adversary  $\mathcal{A}'$  against  $\Pi^{\text{Cl}}$ .  $\mathcal{A}'$  simulates the QCNM-Real( $\Pi^{\text{Hyb}}$ ,  $\mathcal{A}$ ,  $n$ )-experiment. When the QCNM-Real challenger is supposed to encrypt a plaintext to be sent to  $\mathcal{A}$ ,  $\mathcal{A}'$  sends  $m_0 = k$  and  $m_1 = k'$  as challenge plaintexts to the IND-parCCA2 challenger, where  $k, k' \leftarrow \text{KeyGen}^{\text{Qu}}$ , and  $k$  is used to encrypt the quantum plaintext. After storing a copy of the resulting classical ciphertext  $c$  and the one-time key  $k$ ,  $\mathcal{A}'$  continues to simulate QCNM-Real( $\Pi^{\text{Hyb}}$ ,  $\mathcal{A}$ ,  $n$ ) but using the mixed quantum-classical modification check from the spoofed classical key experiments defined above. Decryption is done using the parCCA2 oracle, except for the ciphertexts with classical part  $c$ , which are just decrypted using the stored one-time key  $k$ . Now  $\mathcal{A}'$  outputs the result of the simulated experiment QCNM-Real( $\Pi^{\text{Hyb}}$ ,  $\mathcal{A}$ ,  $n$ ).

Now observe that if the IND-parCCA2 challenger's bit comes up  $b = 0$ ,  $\mathcal{A}'$  faithfully simulated the experiment QCNM-Real( $\Pi^{\text{Hyb}}$ ,  $\mathcal{A}$ ,  $n$ ), while the case  $b = 1$  results in a simulation of sckQCNM-Real( $\Pi^{\text{Hyb}}$ ,  $\mathcal{A}$ ,  $n$ ). Therefore, the IND-parCCA2 security of  $\Pi^{\text{Cl}}$  implies that the games QCNM-Real and sckQCNM-Real have the same result, up to negligible difference.

We can also define an IND-parCCA2 adversary  $\mathcal{A}''$  against  $\Pi^{\text{Cl}}$  in the same way as  $\mathcal{A}'$ , but this time using the QCNM-Ideal experiments. This implies analogously that the experiments QCNM-Ideal and sckQCNM-Ideal also have the same result, up to negligible difference.

Finally, what is left to prove is that the experiments sckQCNM-Real and sckQCNM-Ideal have the same outcome due to the NM security of  $\Pi^{\text{Qu}}$ . If the classical part of the ciphertext has been modified,  $\text{Dec}_k$  is never applied. By the fact that the scheme  $\Pi^{\text{Qu}}$  is IND secure [AM17],  $\mathbf{M}$  is independent of (i.e. in a product state with)  $R$ , i.e.  $\mathbf{M}R$  and  $\mathbf{M}\bar{R}$  have the same state. Therefore, sckQCNM-Real and sckQCNM-Ideal have the same outcome. For the remaining case of  $c' = c$ , note that the modification test in lines 8 through 13 of Experiments 3 and 4 are identical, and that the application of  $(V^M)^\dagger$  ( $T$  is trivial for unitary encryption) is equal to decryption. We can hence decrypt all ciphertexts before the modification test in the experiments sckQCNM-Real and sckQCNM-Ideal (line 9 in experiments QCNM-Real and QCNM-Ideal), and replace  $U_{\text{prep}}$  by  $U$ . It follows

that the rest of the experiment after decryption does not depend on the one-time key  $k$  anymore. Hence the experiment has the form of a multi-decryption attack on the scheme  $\Pi^{\text{Qu}}$ , i.e. where one ciphertext (the one that  $\mathcal{A}_2$  receives as input) is mapped to many ciphertexts (the ones in  $\mathbf{C}$ ) and are subsequently decrypted. We can therefore apply Lemma A.4 to conclude that the modification test outputs 0 unless  $\mathbf{M}$  is in product with  $R$ , in which case  $\mathbf{M}R$  and  $\mathbf{M}\tilde{R}$  have the same state. sckQCNM-Real and sckQCNM-Ideal therefore have the same outcome.  $\square$

## 5 Open Questions

After providing the first definition of non-malleability for quantum public-key encryption and showing how to fulfill it, and providing a comprehensive taxonomy of one-time security notions in the symmetric-key case, our work leaves a number of interesting open questions.

First, one might wonder what other connections PNM and CiNM have to other established security notions, such as the suggestion made in [AM17] that NM, CiNM or PNM might be used to construct a totally authenticating scheme as defined in [GYZ17].

Second, many interesting problems remain in the computational setting. While our proposed definition of QCNM provides a natural extension of CNM to the quantum setting, a number of alternative but equivalent definitions of classical non-malleability exist, such as simulation-based non-malleability as defined in [BS99]. Besides the natural question whether QCNM truly captures non-malleability, one might want to consider quantum versions of other classical notions of non-malleability and the relations between them. Furthermore, a symmetric-key version of QCNM could be explored, which we conjecture to be distinct from a computational version of CiNM due to the mismatch of the way side information is handled (the  $S$  register in QCNM and the  $B$  register in CiNM).

### Acknowledgements

The authors thank the anonymous referees for the encouragement to attempt a better informal explanation of our results. CM would like to thank Gorjan Alagic for stimulating discussions. CM and CS were funded by a NWO VIDI grant (Project No. 639.022.519). CM was supported by a NWO VENI grant (Project No. VI.Veni.192.159). JvW acknowledges the support from the Luxembourg National Research Fund via the CORE project Q-CoDe (Project No. 11689058).

### References

- [ABW09] Andris Ambainis, Jan Bouda, and Andreas Winter. “Nonmalleable encryption of quantum information”. In: *Journal of Mathematical Physics* 50.4 (2009), p. 042106.

- [AGM18a] Gorjan Alagic, Tommaso Gagliardoni, and Christian Majenz. “Can you sign a quantum state?” In: *arXiv preprint arXiv:1811.11858* (2018).
- [AGM18b] Gorjan Alagic, Tommaso Gagliardoni, and Christian Majenz. “Unforgeable Quantum Encryption”. In: *Advances in Cryptology – EUROCRYPT 2018*. Ed. by Jesper Buus Nielsen and Vincent Rijmen. Cham: Springer International Publishing, 2018, pp. 489–519. ISBN: 978-3-319-78372-7.
- [Ala+16] Gorjan Alagic, Anne Broadbent, Bill Fefferman, Tommaso Gagliardoni, Christian Schaffner, and Michael St. Jules. “Computational Security of Quantum Encryption”. In: *Information Theoretic Security*. Ed. by Anderson C.A. Nascimento and Paulo Barreto. Cham: Springer International Publishing, 2016, pp. 47–71. ISBN: 978-3-319-49175-2.
- [Ala+18] Gorjan Alagic, Christian Majenz, Alexander Russell, and Fang Song. “Quantum-secure message authentication via blind-unforgeability”. In: *arXiv preprint arXiv:1803.03761* (2018).
- [AM17] Gorjan Alagic and Christian Majenz. “Quantum Non-malleability and Authentication”. In: *Advances in Cryptology – CRYPTO 2017*. Ed. by Jonathan Katz and Hovav Shacham. Cham: Springer International Publishing, 2017, pp. 310–341. ISBN: 978-3-319-63715-0.
- [BJ15] Anne Broadbent and Stacey Jeffery. “Quantum Homomorphic Encryption for Circuits of Low T-gate Complexity”. In: *Advances in Cryptology – CRYPTO 2015*. Ed. by Rosario Gennaro and Matthew Robshaw. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 609–629. ISBN: 978-3-662-48000-7.
- [BS99] Mihir Bellare and Amit Sahai. “Non-malleable Encryption: Equivalence between Two Notions, and an Indistinguishability-Based Characterization”. In: *Advances in Cryptology – CRYPTO ’99*. Ed. by Michael Wiener. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 519–536. ISBN: 978-3-540-48405-9.
- [BW16] Anne Broadbent and Evelyn Wainwright. “Efficient simulation for quantum message authentication”. In: *International Conference on Information Theoretic Security*. Springer. 2016, pp. 72–91.
- [BZ13a] Dan Boneh and Mark Zhandry. “Quantum-Secure Message Authentication Codes”. In: *Advances in Cryptology – EUROCRYPT 2013*. Ed. by Thomas Johansson and Phong Q. Nguyen. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 592–608. ISBN: 978-3-642-38348-9.
- [BZ13b] Dan Boneh and Mark Zhandry. “Secure Signatures and Chosen Ciphertext Security in a Quantum Computing World”. In: *Advances in Cryptology – CRYPTO 2013*. Ed. by Ran Canetti and Juan A. Garay. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 361–379. ISBN: 978-3-642-40084-1.

- [Cho75] Man-Duen Choi. “Completely positive linear maps on complex matrices”. In: *Linear algebra and its applications* 10.3 (1975), pp. 285–290.
- [DDN03] D. Dolev, C. Dwork, and M. Naor. “Nonmalleable Cryptography”. In: *SIAM Review* 45.4 (2003), pp. 727–784. DOI: [10.1137/S0036144503429856](https://doi.org/10.1137/S0036144503429856).
- [DNS12] Frédéric Dupuis, Jesper Buus Nielsen, and Louis Salvail. “Actively Secure Two-Party Evaluation of Any Quantum Operation”. In: *Advances in Cryptology – CRYPTO 2012*. Ed. by Reihaneh Safavi-Naini and Ran Canetti. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 794–811. ISBN: 978-3-642-32009-5.
- [Don+19] Jelle Don, Serge Fehr, Christian Majenz, and Christian Schaffner. *Security of the Fiat-Shamir Transformation in the Quantum Random-Oracle Model*. Cryptology ePrint Archive, Report 2019/190. <https://eprint.iacr.org/2019/190>. 2019.
- [GYZ17] Sumegha Garg, Henry Yuen, and Mark Zhandry. “New Security Notions and Feasibility Results for Authentication of Quantum Data”. In: *Advances in Cryptology – CRYPTO 2017*. Ed. by Jonathan Katz and Hovav Shacham. Cham: Springer International Publishing, 2017, pp. 342–371. ISBN: 978-3-319-63715-0.
- [Jam72] Andrzej Jamiołkowski. “Linear transformations which preserve trace and positive semidefiniteness of operators”. In: *Reports on Mathematical Physics* 3.4 (1972), pp. 275–278.
- [Low10] Richard A Low. “Pseudo-randomness and learning in quantum computation”. In: *arXiv preprint arXiv:1006.5227* (2010).
- [Pod+18] Damian Poddebniak, Christian Dresen, Jens Müller, Fabian Ising, Sebastian Schinzel, Simon Friedberger, Juraj Somorovsky, and Jörg Schwenk. “Efail: Breaking S/MIME and OpenPGP Email Encryption using Exfiltration Channels”. In: *27th USENIX Security Symposium (USENIX Security 18)*. Baltimore, MD: USENIX Association, 2018, pp. 549–566. ISBN: 978-1-931971-46-1. URL: <https://www.usenix.org/conference/usenixsecurity18/presentation/poddebniak>.
- [PV06] Rafael Pass, abhi shelat abhi, and Vinod Vaikuntanathan. “Construction of a Non-malleable Encryption Scheme from Any Semantically Secure One”. In: *Advances in Cryptology – CRYPTO 2006*. Ed. by Cynthia Dwork. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 271–289. ISBN: 978-3-540-37433-6.
- [Unr12] Dominique Unruh. “Quantum Proofs of Knowledge”. In: *Advances in Cryptology – EUROCRYPT 2012*. Ed. by David Pointcheval and Thomas Johansson. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 135–152. ISBN: 978-3-642-29011-4.
- [Wal18] Nicholas Wallace. “EU runs to catch up as governments pledge more cash for quantum computing”. Ed. by Richard L. Hudson. Nov. 8, 2018. URL: <https://web.archive.org/web/20181108205334/>

<https://sciencebusiness.net/news/eu-runs-catch-governments-pledge-more-cash-quantum-computing>.

- [Wat18] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018. URL: <https://cs.uwaterloo.ca/~watrous/TQI/TQI.pdf>.
- [WEH18] Stephanie Wehner, David Elkouss, and Ronald Hanson. “Quantum internet: A vision for the road ahead”. In: *Science* 362.6412 (2018). ISSN: 0036-8075. DOI: [10.1126/science.aam9288](https://doi.org/10.1126/science.aam9288).

## A Proofs

**Lemma A.1 (Lemma 3.4).** *Let  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  be an arbitrary SKQES and  $\Lambda_A^{CB \rightarrow C\hat{B}}$  an arbitrary attack on  $\Pi$  with effective map  $\tilde{\Lambda}_A^{MB \rightarrow M\hat{B}}$ . If there exist CPTNI  $\Lambda_1, \Lambda_2$ , such that  $\Lambda_1 + \Lambda_2$  is CPTP and it holds that*

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_1^{B \rightarrow \hat{B}} + \frac{1}{|C|^2 - 1} (|C|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq \varepsilon,$$

then for any  $\alpha$  such that  $|M|^2 \leq \alpha \leq |C|^2$  there exist CPTNI  $\Lambda_3, \Lambda_4$  such that  $\Lambda_3 + \Lambda_4$  is CPTP and

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_3^{B \rightarrow \hat{B}} + \frac{1}{\alpha - 1} (\alpha \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_4^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq \varepsilon.$$

*Proof.* Assume that for some CPTNI  $\Lambda_1, \Lambda_2$  such that  $\Lambda_1 + \Lambda_2$  is CPTP it holds that

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_1^{B \rightarrow \hat{B}} + \frac{1}{|C|^2 - 1} (|C|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq \varepsilon.$$

Define  $\gamma = \frac{(\alpha - 1)|C|^2}{\alpha(|C|^2 - 1)}$ ,  $\Lambda_3 = \Lambda_1 + (1 - \gamma)\Lambda_2$ , and  $\Lambda_4 = \gamma\Lambda_2$ . Note that  $0 < \gamma \leq 1$  as long as  $1 < \alpha \leq |C|^2$  and thus  $\Lambda_3$  and  $\Lambda_4$  are CPTNI. Furthermore  $\Lambda_3 + \Lambda_4 = \Lambda_1 + \Lambda_2$ , thus  $\Lambda_3 + \Lambda_4$  is CPTP. Observe that

$$\begin{aligned} & \text{id}^M \otimes \Lambda_3^{B \rightarrow \hat{B}} + \frac{1}{\alpha - 1} (\alpha \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_4^{B \rightarrow \hat{B}} \\ &= \text{id}^M \otimes (\Lambda_1 + (1 - \gamma)\Lambda_2) + \frac{1}{\alpha - 1} (\alpha \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \gamma\Lambda_2 \\ &= \text{id}^M \otimes \Lambda_1 + (1 - \gamma)\text{id}^M \otimes \Lambda_2 + \frac{\gamma}{\alpha - 1} (\alpha \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2 \\ &= \text{id}^M \otimes \Lambda_1 + \frac{|C|^2}{|C|^2 - 1} \langle \text{Dec}_K(\tau) \rangle \otimes \Lambda_2 - \frac{1}{|C|^2 - 1} \text{id}^M \otimes \Lambda_2 \\ &= \text{id}^M \otimes \Lambda_1 + \frac{1}{|C|^2 - 1} (|C|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2. \end{aligned}$$

From this it follows that

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_3^{B \rightarrow \hat{B}} + \frac{1}{\alpha - 1} (\alpha \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_4^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq \varepsilon.$$

□

**Theorem A.2 (Theorem 3.6).** *Let  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  be an arbitrary  $\varepsilon$ -PNM SKQES for some  $\varepsilon$ , then for any attack  $\Lambda_A^{CB \rightarrow C\hat{B}}$ , its effective map  $\tilde{\Lambda}_A^{MB \rightarrow M\hat{B}}$  is such that*

$$\left\| \tilde{\Lambda}_A - \left( \text{id}^M \otimes \Lambda_1^{B \rightarrow \hat{B}} + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2^{B \rightarrow \hat{B}} \right) \right\|_{\diamond} \leq 3\varepsilon,$$

where

$$\begin{aligned} \Lambda_1 &= \text{Tr}_{MM'} \left[ \phi^{+MM'} \tilde{\Lambda}_A(\phi^{+MM'} \otimes (\cdot)) \right] & \text{and} \\ \Lambda_2 &= \text{Tr}_{MM'} \left[ (\mathbf{1}^{MM'} - \phi^{+MM'}) \tilde{\Lambda}_A(\phi^{+MM'} \otimes (\cdot)) \right]. \end{aligned}$$

*Proof.* Let  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  be an arbitrary  $\varepsilon$ -PNM SKQES for some  $\varepsilon$  and let  $\Lambda_A^{CB \rightarrow C\hat{B}}$  be an arbitrary attack with effective map  $\tilde{\Lambda}_A^{MB \rightarrow M\hat{B}}$ . Furthermore, let  $\Lambda_1^{B \rightarrow \hat{B}}$  and  $\Lambda_2^{B \rightarrow \hat{B}}$  be such that

$$\left\| \tilde{\Lambda}_A - \tilde{\Lambda}_{ideal} \right\|_{\diamond} \leq \varepsilon,$$

where  $\tilde{\Lambda}_{ideal}^{MB \rightarrow M\hat{B}} = \text{id}^M \otimes \Lambda_1 + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau) \rangle - \text{id})^M \otimes \Lambda_2$ . Lastly, let

$$\begin{aligned} \Lambda_3 &= \text{Tr}_{MM'} \left[ \phi^{+MM'} \tilde{\Lambda}_A(\phi^{+MM'} \otimes (\cdot)) \right], \\ \Lambda_4 &= \text{Tr}_{MM'} \left[ (\mathbf{1}^{MM'} - \phi^{+MM'}) \tilde{\Lambda}_A(\phi^{+MM'} \otimes (\cdot)) \right], \text{ and} \\ \tilde{\Lambda}_{trace}^{MB \rightarrow M\hat{B}} &= \text{id}^M \otimes \Lambda_3 + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau) \rangle - \text{id})^M \otimes \Lambda_4. \end{aligned}$$

Observe that, by the triangle inequality,  $\left\| \tilde{\Lambda} - \tilde{\Lambda}_{trace} \right\|_{\diamond} \leq \left\| \tilde{\Lambda} - \tilde{\Lambda}_{ideal} \right\|_{\diamond} + \left\| \tilde{\Lambda}_{ideal} - \tilde{\Lambda}_{trace} \right\|_{\diamond}$ . Furthermore,

$$\begin{aligned} & \left\| \tilde{\Lambda}_{ideal} - \tilde{\Lambda}_{trace} \right\|_{\diamond} \\ &= \left\| \text{id}^M \otimes (\Lambda_1 - \Lambda_3) + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau) \rangle - \text{id})^M \otimes (\Lambda_2 - \Lambda_4) \right\|_{\diamond} \\ &\leq \left\| \text{id}^M \otimes (\Lambda_1 - \Lambda_3) \right\|_{\diamond} + \left\| \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau) \rangle - \text{id})^M \otimes (\Lambda_2 - \Lambda_4) \right\|_{\diamond} \\ &= \|\text{id}\|_{\diamond} \|(\Lambda_1 - \Lambda_3)\|_{\diamond} + \left\| \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau) \rangle - \text{id})^M \right\|_{\diamond} \|(\Lambda_2 - \Lambda_4)\|_{\diamond} \\ &\leq \|(\Lambda_1 - \Lambda_3)\|_{\diamond} + \|(\Lambda_2 - \Lambda_4)\|_{\diamond} \end{aligned}$$

Let  $\Lambda_5 = \text{Tr}_{MM'} \left[ \phi^{+MM'} \tilde{\Lambda}_{ideal}(\phi^{+MM'} \otimes (\cdot)) \right]$  and  $\Lambda_6 = \text{Tr}_{MM'} \left[ (\mathbb{1}^{MM'} - \phi^{+MM'}) \tilde{\Lambda}_{ideal}(\phi^{+MM'} \otimes (\cdot)) \right]$ . Observe that the mapping

$$\rho \mapsto |0\rangle\langle 0| \otimes \text{Tr}_{MM'}[\phi^{+MM'} \rho] + |1\rangle\langle 1| \otimes \text{Tr}_{MM'}[(\mathbb{1}^{MM'} - \phi^{+MM'})\rho]$$

is CPTP. Since  $\left\| (\tilde{\Lambda} - \tilde{\Lambda}_{ideal})(\phi^{+MM'} \otimes (\cdot)) \right\|_{\diamond} \leq \left\| \tilde{\Lambda} - \tilde{\Lambda}_{ideal} \right\|_{\diamond} \leq \varepsilon$  and the diamond norm is non-increasing under CPTP maps<sup>11</sup>, we have  $\left\| |0\rangle\langle 0| \otimes (\Lambda_3 - \Lambda_5) + |1\rangle\langle 1| \otimes (\Lambda_4 - \Lambda_6) \right\|_{\diamond} \leq \varepsilon$  and thus  $\|\Lambda_3 - \Lambda_5\|_{\diamond} \leq \varepsilon$  and  $\|\Lambda_4 - \Lambda_6\|_{\diamond} \leq \varepsilon$ . Using this we observe that

$$\begin{aligned} \left\| \tilde{\Lambda}_{ideal} - \tilde{\Lambda}_{trace} \right\|_{\diamond} &\leq \|\Lambda_1 - \Lambda_3\|_{\diamond} + \|\Lambda_2 - \Lambda_4\|_{\diamond} \\ &\leq \|\Lambda_1 - \Lambda_5\|_{\diamond} + \|\Lambda_5 - \Lambda_3\|_{\diamond} + \|\Lambda_2 - \Lambda_6\|_{\diamond} + \|\Lambda_6 - \Lambda_4\|_{\diamond} \\ &\leq 2\varepsilon + \|\Lambda_1 - \Lambda_5\|_{\diamond} + \|\Lambda_2 - \Lambda_6\|_{\diamond}. \end{aligned}$$

Furthermore we have

$$\begin{aligned} \Lambda_5 &= \text{Tr}_{MM'}[\phi^+ \tilde{\Lambda}_{ideal}(\phi^+ \otimes (\cdot))] \\ &= \text{Tr}_{MM'} \left[ \phi^+ \left( \phi^+ \otimes \Lambda_1 + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau) \rangle - \text{id})(\phi^+) \otimes \Lambda_2 \right) \right] \\ &= \text{Tr}_{MM'} \left[ \phi^+ \left( \phi^+ \otimes \Lambda_1 + \frac{1}{|M|^2 - 1} (|M|^2 \text{Dec}_K(\tau) \otimes \tau^{M'} - \phi^+) \otimes \Lambda_2 \right) \right] \\ &= \Lambda_1 + \text{Tr} \left[ \frac{1}{|M|^2 - 1} (|M|^2 \phi^+ (\text{Dec}_K(\tau) \otimes \tau^{M'}) - \phi^+) \right] \Lambda_2 \\ &= \Lambda_1, \end{aligned}$$

where  $\phi^+ = \phi^{+MM'}$  and the last equality holds because

$$\begin{aligned} \text{Tr} \left[ \phi^{+MM'} (\text{Dec}_K(\tau) \otimes \tau^{M'}) \right] &= \frac{1}{|M|} \text{Tr} \left[ \sum_{i,j=0}^{|M|-1} |ii\rangle\langle jj| (\text{Dec}_K(\tau) \otimes \tau^{M'}) \right] \\ &= \frac{1}{|M|} \text{Tr} \left[ \sum_{i,j=0}^{|M|-1} |i\rangle\langle j| \text{Dec}_K(\tau) \otimes |i\rangle\langle j| \tau^{M'} \right] \\ &= \frac{1}{|M|^2} \text{Tr} \left[ \sum_{i=0}^{|M|-1} |i\rangle\langle i| \text{Dec}_K(\tau) \right] \\ &= \frac{1}{|M|^2} \end{aligned}$$

<sup>11</sup> See [Wat18], Proposition 3.48(1)



Similarly

$$\begin{aligned}
\Lambda_6 &= \text{Tr}_{MM'} \left[ (\mathbb{I}^{MM'} - \phi^{+MM'}) \tilde{\Lambda}^{ideal}(\phi^{+MM'} \otimes (\cdot)) \right] \\
&= \text{Tr}_{MM'} \left[ \tilde{\Lambda}^{ideal}(\phi^{+MM'} \otimes (\cdot)) \right] - \Lambda_5 \\
&= \text{Tr}_{MM'} \left[ \left( \phi^+ \otimes \Lambda_1 + \frac{1}{|M|^2 - 1} (|M|^2 \text{Dec}_K(\tau) \otimes \tau^{M'} - \phi^+) \otimes \Lambda_2 \right) \right] \\
&= \Lambda_1 + \Lambda_2 - \Lambda_5 \\
&= \Lambda_2.
\end{aligned}$$

From this we conclude

$$\begin{aligned}
\left\| \tilde{\Lambda} - \tilde{\Lambda}_{trace} \right\|_{\diamond} &\leq \left\| \tilde{\Lambda} - \tilde{\Lambda}^{ideal} \right\|_{\diamond} + \left\| \tilde{\Lambda}^{ideal} - \tilde{\Lambda}_{trace} \right\|_{\diamond} \\
&\leq \varepsilon + \left\| \tilde{\Lambda}^{ideal} - \tilde{\Lambda}_{trace} \right\|_{\diamond} \\
&\leq 3\varepsilon + \|\Lambda_1 - \Lambda_5\|_{\diamond} + \|\Lambda_2 - \Lambda_6\|_{\diamond} \\
&= 3\varepsilon.
\end{aligned}$$

□

**Theorem A.3 (Theorem 3.12).** *For any  $0 \leq \varepsilon \leq 2$  and any  $\varepsilon$ -PNM SKQES  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$ , there exists some  $x$  such that the scheme  $\Pi' = (\text{KeyGen}, \text{Enc}', \text{Dec}')$  is  $\left(\frac{3}{|R|} + \varepsilon\right)$ -DNS-authenticating, where*

$$\begin{aligned}
\text{Enc}'_k &= \text{Enc}_k((\cdot)^{M'} \otimes |x\rangle\langle x|^R) \\
\text{Dec}'_k &= \langle x|^R \text{Dec}_k(\cdot) |x\rangle^R + \text{Tr}[(\mathbb{1}^R - |x\rangle\langle x|^R) \text{Dec}_k(\cdot)] |\perp\rangle\langle\perp|
\end{aligned}$$

*Proof.* By Lemma 3.11, there exists an  $x \in \{0, 1\}^{\log |R|}$  such that  $\text{Tr}[\langle x | \text{Dec}_K(\tau^C) |x\rangle] \leq \frac{1}{|R|}$ . Fix this  $x$  and define  $\Pi'$  as above. Define  $\text{Enc}_{ap}(X) = X \otimes |x\rangle\langle x|$  and  $\text{Dec}_{ch}(Y) = \langle x | Y |x\rangle + \text{Tr}[(\mathbb{1} - |x\rangle\langle x|) Y] |\perp\rangle\langle\perp|$  and observe that  $\text{Enc}' = \text{Enc} \circ \text{Enc}_{ap}$  and  $\text{Dec}' = \text{Dec}_{ch} \circ \text{Dec}$ . Let  $\Lambda_A$  be an arbitrary attack map on  $\Pi'$ , then its effective map is

$$\tilde{\Lambda}'_A = \mathbb{E}_{k \leftarrow \text{KeyGen}(1^n)} [\text{Dec}'_k \circ \Lambda_A \circ \text{Enc}'_k].$$

Since  $\text{Enc}_{ap}$  and  $\text{Dec}_{ch}$  do not change with  $k$  and are linear, we have

$$\tilde{\Lambda}'_A = \text{Dec}_{ch} \circ \tilde{\Lambda}_A \circ \text{Enc}_{ap},$$

where  $\tilde{\Lambda}_A = \mathbb{E}_{k \leftarrow \text{KeyGen}(1^n)} [(\text{Dec}_k \circ \Lambda_A \circ \text{Enc}_k)]$ . Since  $\Pi$  is  $\varepsilon$ -PNM, there exist  $\Lambda_1, \Lambda_2$  such that

$$\left\| \tilde{\Lambda}_A - \text{id} \otimes \Lambda_1 + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2 \right\|_{\diamond} \leq \varepsilon.$$

Since  $\text{Enc}_{ap}$  and  $\text{Dec}_{ch}$  are both CPTP, by submultiplicativity we have that

$$\left\| \text{Dec}_{ch} \circ \left( \tilde{A}_A - \text{id} \otimes \Lambda_1 + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau) \rangle - \text{id}) \otimes \Lambda_2 \right) \circ \text{Enc}_{ap} \right\|_{\diamond} \leq \varepsilon,$$

which is equivalent to

$$\left\| \tilde{A}'_A - \text{Dec}_{ch} \circ \left( \text{id} \otimes \Lambda_1 + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau) \rangle - \text{id}) \otimes \Lambda_2 \right) \circ \text{Enc}_{ap} \right\|_{\diamond} \leq \varepsilon.$$

Observe that

$$\text{Dec}_{ch} \circ \text{id} \circ \text{Enc}_{ap} = \langle x | ((\cdot) \otimes |x\rangle\langle x|) |x\rangle + \text{Tr}[(\mathbb{1} - |x\rangle\langle x|)((\cdot) \otimes |x\rangle\langle x|)] |\perp\rangle\langle\perp| = \text{id}.$$

Define  $\Lambda_{acc} = \Lambda_1$ ,  $\Lambda_{rej} = \Lambda_2$  and

$$\tilde{A}_{ideal} = \text{Dec}_{ch} \circ \left( \text{id} \otimes \Lambda_1 + \frac{1}{|M|^2 - 1} (|M|^2 \langle \text{Dec}_K(\tau^C) \rangle - \text{id})^M \otimes \Lambda_2 \right) \circ \text{Enc}_{ap},$$

then we have

$$\begin{aligned} & \left\| \tilde{A}_{ideal} - \text{id} \otimes \Lambda_{acc} - \langle |\perp\rangle\langle\perp| \rangle \otimes \Lambda_{rej} \right\|_{\diamond} \\ &= \left\| \left( \frac{1}{|M|^2 - 1} (|M|^2 (\text{Dec}_{ch} \circ \langle \text{Dec}_K(\tau) \rangle) \circ \text{Enc}_{ap}) - \text{id} \right) - \langle |\perp\rangle\langle\perp| \rangle \right\|_{\diamond} \otimes \Lambda_2 \left\|_{\diamond} \right. \\ &\leq \left\| \frac{1}{|M|^2 - 1} (|M|^2 (\text{Dec}_{ch} \circ \text{Tr}[(\cdot)^{M'}] \text{Dec}_K(\tau)) - \text{id}) - \langle |\perp\rangle\langle\perp| \rangle \right\|_{\diamond} \end{aligned}$$

Here the inequality uses the fact that  $\text{Enc}_{ap}$  is trace preserving and  $\langle \text{Dec}_K(\tau) \rangle$  is a constant channel, which only uses the trace of the input. Since every term ended in  $\otimes \Lambda_2$ , we removed this term and multiplied with  $\|\Lambda_2\|_{\diamond}$ , which is less than 1 since  $\Lambda_2$  is CPTNI. We continue by expanding  $\text{Dec}_{ch}$ , where we use that  $\langle x | \text{Tr}[(\cdot)^{M'}] \text{Dec}_K(\tau^C) |x\rangle = \langle\langle x | \text{Dec}_K(\tau^C) |x\rangle\rangle$  and we abbreviate  $\psi = \langle x | \text{Dec}_K(\tau^C) |x\rangle$  and  $[\perp] = \langle |\perp\rangle\langle\perp| \rangle$ .

$$\begin{aligned} & \left\| \frac{1}{|M|^2 - 1} (|M|^2 (\text{Dec}_{ch} \circ \text{Tr}[(\cdot)^{M'}] \text{Dec}_K(\tau^C)) - \text{id}) - [\perp] \right\|_{\diamond} \\ &= \left\| \frac{1}{|M|^2 - 1} (|M|^2 (\langle \psi \rangle + \text{Tr}[(\mathbb{1} - |x\rangle\langle x|) \text{Dec}_K(\tau^C)] [\perp]) - \text{id}) - [\perp] \right\|_{\diamond}. \end{aligned}$$

We can rewrite this expression by first rewriting  $\text{Tr}[(\mathbb{1} - |x\rangle\langle x|) \text{Dec}_K(\tau^C)]$  as  $1 - \text{Tr}[\psi]$ , then collecting all multipliers of  $\langle |\perp\rangle\langle\perp| \rangle$ , and lastly distributing the

$|M|^2$  term and simplifying the resulting term.

$$\begin{aligned}
& \left\| \frac{1}{|M|^2 - 1} (|M|^2 \langle \psi \rangle + (1 - \text{Tr}[\psi]) \langle |\perp\rangle \langle \perp| \rangle) - \text{id} - \langle |\perp\rangle \langle \perp| \rangle \right\|_{\diamond} \\
&= \left\| \frac{1}{|M|^2 - 1} (|M|^2 \left( \langle \psi \rangle + \left( (1 - \text{Tr}[\psi]) - \frac{|M|^2 - 1}{|M|^2} \right) \langle |\perp\rangle \langle \perp| \rangle \right) - \text{id}) \right\|_{\diamond} \\
&= \left\| \frac{1}{|M|^2 - 1} (|M|^2 \langle \psi \rangle + (|M|^2(1 - \text{Tr}[\psi]) - (|M|^2 - 1)) \langle |\perp\rangle \langle \perp| \rangle - \text{id}) \right\|_{\diamond} \\
&= \left\| \frac{1}{|M|^2 - 1} (|M|^2 \langle \psi \rangle + (1 - |M|^2 \text{Tr}[\psi]) \langle |\perp\rangle \langle \perp| \rangle - \text{id}) \right\|_{\diamond} \\
&\leq \frac{1}{|M|^2 - 1} (|M|^2 \|\langle \psi \rangle\|_{\diamond} + \|(1 - |M|^2 \text{Tr}[\psi]) \langle |\perp\rangle \langle \perp| \rangle\|_{\diamond} + \|\text{id}\|_{\diamond}) \\
&\leq \frac{1}{|M|^2 - 1} \left( \frac{|M|^2}{|R|} + \left( \frac{|M|^2}{|R|} - 1 \right) + 1 \right) \leq \frac{3}{|R|}.
\end{aligned}$$

Here the first inequality is an application of the triangle inequality. The second inequality uses the fact that  $\|\text{id}\|_{\diamond} = \|\langle |\perp\rangle \langle \perp| \rangle\|_{\diamond} = 1$  and that  $|R|\langle \psi \rangle$  is CPTNI because  $\text{Tr}[\langle x | \text{Dec}_K(\tau^C) | x \rangle] \leq \frac{1}{|R|}$  and thus  $\|\langle \psi \rangle\|_{\diamond} \leq \frac{1}{|R|}$ .

Since  $\|\tilde{A}'_A - \tilde{A}_{ideal}\|_{\diamond} \leq \varepsilon$  and  $\|\tilde{A}_{ideal} - \text{id} \otimes A_{acc} - \langle |\perp\rangle \langle \perp| \rangle \otimes A_{rej}\|_{\diamond} \leq \frac{3}{|R|}$ , we have by the triangle inequality that

$$\left\| \tilde{A}'_A - \text{id} \otimes A_{acc} - \langle |\perp\rangle \langle \perp| \rangle \otimes A_{rej} \right\|_{\diamond} \leq \varepsilon + \frac{3}{|R|},$$

which means that  $\Pi'$  is  $\left(\frac{3}{|R|} + \varepsilon\right)$ -DNS authenticating.  $\square$

To prove QCNM security of the classical-quantum hybrid scheme, we need the following lemma.

**Lemma A.4.** *Let  $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$  be a SKQES, let  $\ell \in \mathbb{N}$ , let  $\mathbf{C} = C_1 \dots C_{\ell} \cong C^{\ell}$  and  $\mathbf{M} = M_1 \dots M_{\ell} \cong M^{\ell}$  be vectors of registers, let  $\Lambda^{C \rightarrow C}$  be a CPTP map, and set*

$$\tilde{\Lambda}^{M \rightarrow M} = \mathbb{E}_{k \leftarrow \text{KeyGen}(1^n)} \left[ (\text{Dec}_k)^{\otimes \ell} \circ \Lambda \circ \text{Enc}_k \right].$$

*If SKQES is CiNM secure, then for some  $p_0$  and  $\{\sigma_i\}_i$  we have that*

$$\tilde{\Lambda}^{M \rightarrow M} = \sum_{i=1}^{\ell} p_i \text{id}^{M \rightarrow M_i} \otimes \sigma_i^{M_{-i}} + p_0 \langle \sigma_0^M \rangle,$$

*where  $q_i$  is the probability that is equal to  $\Lambda_1$  from Definition 3.2 applied to the attack map  $\text{Tr}_{\mathbf{C}_{-i}} \circ \Lambda$  and  $\mathbf{M}_{-i} = M_1 \dots M_{i-1} M_{i+1} \dots M_{\ell}$  and  $p_i = q_i - \frac{1}{|C|^{\ell-1}}(1 - q_i)$ .*

*Proof.* For fixed  $i$ , consider the attack  $\Lambda' = \text{Tr}_{\mathbf{C}_{-i}} \circ \Lambda$  on  $\Pi$  and observe that its effective map satisfies  $\tilde{\Lambda}' = \mathbb{E}_k [\text{Dec}_k \circ \text{Tr}_{\mathbf{C}_{-i}} \circ \Lambda \circ \text{Enc}_k] = \text{Tr}_{\mathbf{M}_{-i}} \circ \tilde{\Lambda}$ . Because  $\Pi$  is CiNM, we have  $\tilde{\Lambda}' = p_i \text{id} + (1 - p_i) \langle \text{Dec}_K(\tau) \rangle$  and thus

$$\text{Tr}_{\mathbf{M}_{-i}} \circ \tilde{\Lambda}(\phi^{+MM'}) = p_i \phi^{+M_i M'} + (1 - p_i) \text{Dec}_K(\tau) \otimes \tau^{M'}. \quad (1)$$

Consider the state  $\tilde{\Lambda}(\phi^{+MM'})$ . Because  $\phi^{+M_i M'}$  is a pure state, we know that  $\tilde{\Lambda}(\phi^{+MM'})$  is a convex combination of terms of the form  $p_i \phi^{+M_i M'} \otimes \sigma_{\mathbf{M}_{-i}}$  and a term  $p_0 \sigma_0^{\mathbf{M}} \otimes \tau^{M'}$ , i.e.

$$\tilde{\Lambda}(\phi^{+MM'}) = \sum_{i=1}^{\ell} p_i \phi^{+M_i M'} \otimes \sigma_i^{\mathbf{M}_{-i}} + p_0 \sigma_0^{\mathbf{M}} \otimes \tau^{M'}. \quad (2)$$

By the Choi-Jamiolkowski isomorphism[[Jam72](#); [Cho75](#)] this means that

$$\tilde{\Lambda}^{M \rightarrow \mathbf{M}} = \sum_{i=1}^{\ell} p_i \text{id}^{M \rightarrow M_i} \otimes \sigma_i^{\mathbf{M}_{-i}} + p_0 \langle \sigma_0^{\mathbf{M}} \rangle.$$

Using Equation (1), we get in addition that all single-system marginals of  $\sigma_i$ ,  $i = 0, \dots, \ell$  are equal to  $\text{Dec}_K(\tau)$ .  $\square$

Note that a similar statement can be proven for attack maps  $\Lambda^{MB \rightarrow \mathbf{M}\tilde{B}}$  with side information, but we only need the above statement in [Theorem 4.6](#).