Quantum Plaintext Non-Malleability

MSc Thesis *(Afstudeerscriptie)*

written by

Jeroen van Wier
(born August 13, 1995 in Amsterdam, Netherlands)

under the supervision of **Dr Christian Majenz** and **Dr Christian Schaffner**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
June 19, 2018	Prof Dr Yde Venema
	Dr Christian Schaffner
	Dr Christian Majenz
	Dr Alexandru Baltag
	Dr Serge Fehr
	Dr Maris Ozols

Institute for Logic, Language and Computation

Abstract

When two parties communicate they often require a certain level of security. In particular one might desire that a message cannot be altered in transit. This is called non-malleability and has been studied extensively in the non-quantum setting, but was only recently introduced in the field of quantum computing. This thesis provides two ways of defining non-malleability in the quantum setting. One of these definitions is based on previous quantum notions, and the other is an extension of a well-researched classical notion. These definitions capture different forms of non-malleability, and we provide an argument why either can be considered correct.

Contents

CHAPTER 1

Introduction

Quantum computers are closer to becoming a reality than ever before and many researchers have become interested in exploring their capabilities. Many amazing applications of quantum computers exist, but they also pose a new threat when used by those with less than honest intents. A well-known example of such an application is Shor's factoring algorithm [Sho99], which can be used against modern-day cryptographic protocols such as RSA. To ensure safe communications in the future, it is important to look for or construct notions of security that remain secure in the face of this new threat.

In this thesis, we focus on the notion of non-malleability, which captures the idea that an encrypted message cannot be altered by a third party. This form of security does not inherently prevent a third party from learning the message that was encrypted and in the world of classical (non-quantum) computers, these notions are considered separate. When considering quantum computers we will see that an inherent connection between these two notions exists, but it is still valuable to consider both. We will approach the problem of defining this notion from two different directions.

First, we will modify the quantum notion of non-malleability given in [AM17], weakening it to only require security on the plaintext level, meaning that any structural change to the encryption of a message cannot alter the message itself. The basic idea of the resulting notion is that an attacker performs some attack, and we analyze the effect of this attack averaged over all keys, which represents the attacker's ignorance of the key. The security of the scheme ensures that the impact of this attack on the message is either negligible or completely destructive. The shortcomings of this approach include that it only simulates the case where an attacker receives a single encrypted message, whereas an attack might be able to intercept multiple messages in a real-world application. The advantage, however, is that it makes no assumptions about the computational power of an adversary.

Second, we will approach the problem from the classical public-key setting, where we will provide a quantum translation of one of the notions presented in [BS99]. This classical notion, and by extension its quantum variant, differ significantly from the definitions that have already been analyzed in the quantum setting. The major disadvantage of this approach is that the resulting security notion is difficult to relate to previous work. However, the advantage this approach offers is that the resulting notion can be well compared to the corresponding classical notion, which has been analyzed in great detail.

1.1 Social and Scientific Relevance

Non-malleability has been researched extensively in the non-quantum setting, for example in [BS99], however it has been introduced in the field of quantum computing only recently, in [AM17] and [ABW09]. The research done in the quantum setting focuses on symmetric-key encryption, where both the sender and receiver must share the same key. In this thesis, we will attempt to provide some insight into the public-key case, where only the receiver holds a secret key and the sender holds a different key, which is publicly known. Furthermore, we provide a weaker version of the non-malleability defined in [AM17], which intuitively provides a similar level of security but might be easier to satisfy.

The relevance of non-malleability was recently demonstrated with the attack on the PGP protocol, used to securely authenticate e-mail [Pod+18]. The attack demonstrates a flaw in the PGP protocol which allows a possible attacker to insert text of her own choosing into an encrypted message, which in turn exploits the behavior of the program used to receive the e-mail. 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. Besides e-mail, there are many more settings where one would like a message to not change during transmission, for example when communicating with a bank.

1.2 Overview of this Thesis

As discussed in the previous section this thesis provides two possible ways of looking at quantum non-malleability. In Chapter 2 we will give a brief overview of quantum computing. Afterward, we will introduce the notions of non-malleability discussed in previous work in Chapter 3. Chapter 4 will introduce the notion of plaintext non-malleability, an extension of non-malleability as described in [AM17]. The approach of starting from the classical point of view is presented in Chapter 5. Lastly, we will briefly summarize our results in Chapter 6.

CHAPTER 2

Quantum Computation

In this chapter, we define the basic notions of quantum computation and introduce the mathematical tools used throughout this thesis. We first introduce some notation in Section 2.1, then cover pure and mixed states in Section 2.2. Afterward, we look at possible operations on quantum states in Section 2.3. Lastly, we discuss a number of norms and distances in Section 2.4, which will be used to analyze states and operations.

While we strive to provide a solid understanding of quantum computing, this chapter should not be seen as a complete overview. For a more thorough understanding of the basics of quantum computing, one can look at, for example, [Wat18]. We will assume that the reader is familiar with the linear algebra used in this chapter.

2.1 Notation

In this section, we introduce a part of the notation used. More notation will be introduced throughout this thesis when additional concepts are defined and an overview of all notation used can be found in Appendix A.

For any complex matrix M, we denote its conjugate transpose as $M^\dagger$ and its trace as $\mathrm{Tr}[M]$. Throughout this thesis we will only consider finite-dimensional Hilbert spaces. For Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$ we write $|A| := \dim(\mathcal{H}_A)$ for the dimension of a Hilbert space, $\mathbb{I}^A$ for the identity matrix of dimension $|A|$ and $\mathbf{0}^{A\to B}$ or $\mathbf{0}^A$ for the zero matrix of dimension $|A| \times |B|$ or $|A| \times |A|$ respectively. We will denote the set of square matrices that act on $\mathcal{H}_A$ as $\mathcal{B}(\mathcal{H}_A)$. We may omit the superscripts if the spaces are clear from context. For any vector v we denote its Euclidean norm as $\|v\| := \sqrt{v^\dagger v}$.

We call a function $\varepsilon(n)$ negligible if for every polynomial p there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ it holds that $\varepsilon(n) < \frac{1}{p(n)}$. We write $\varepsilon \leq \mathrm{negl}(n)$ to state that the function $\varepsilon(n)$ is negligible. Furthermore we use $\log(x)$ to denote the base 2 logarithm of x.

2.2 Quantum States

In quantum computing one uses quantum bits, or **qubits**, to perform computations. One qubit can be 0, 1, or in a superposition of these two values.

2.2.1 Qubits

To mathematically describe these superpositions, a qubit is a vector of two values. The classical bits 0 and 1 are represented as

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Any **superposition** of these two vectors is represented as $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, with $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. To combine qubits the tensor product is used as follows:

$$|\phi_1\rangle \otimes |\phi_2\rangle = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \otimes \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = \begin{pmatrix} \alpha_1\beta_1 \\ \vdots \\ \alpha_1\beta_m \\ \alpha_2\beta_1 \\ \vdots \\ \alpha_n\beta_m \end{pmatrix}.$$

We will often denote $|\phi_1\rangle \otimes |\phi_2\rangle$ as $|\phi_1\rangle|\phi_2\rangle$ or even $|\phi_1\phi_2\rangle$. The vector resulting from combining multiple qubits has norm 1 and a combination of n qubits is represented by a vector of dimension 2^n. In general one can perform quantum computation with norm-1 vectors of any dimension, however throughout this thesis, we will consider only the case where the dimensions are powers of 2, or in other words, where vectors represent some number of qubits.

If a vector $|\phi\rangle$ is part of a Hilbert space $\mathcal{H}_A$ we may denote it $|\phi\rangle^A$ for clarity, although this superscript is often omitted when the Hilbert space is clear from context. 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**.

The notation used above, where a vector is denoted as $|\phi\rangle$, is known as **bra-ket** notation, where $|\phi\rangle$ is pronounced as "ket phi". The conjugate transpose of this vector is denoted as $\langle\phi| := |\phi\rangle^\dagger$ and is pronounced as "bra phi". A bra and a ket together form a bracket, $\langle\phi|\psi\rangle$, which is the inner product between $|\phi\rangle$ and $|\psi\rangle$.

2.2.2 Pure and mixed states

Quantum states are described by **density matrices**, which are positive semi-definite Hermitian matrices with trace 1. We will write $\mathcal{D}(\mathcal{H})$ to denote the set of all density matrices on a Hilbert space $\mathcal{H}$. A 'ket' and a 'bra' together form a **pure state** $|\phi\rangle\langle\phi|$. A pure state ϕ is any density matrix such that $\phi = |\phi\rangle\langle\phi|$ for some $|\phi\rangle$. Because of this property we often denote a pure state $|\psi\rangle\langle\psi|$ as simply ψ. Equivalently any rank-1 density matrix is a pure state. We will sometimes also refer to a norm-1 vector $|\phi\rangle$ as a (pure) state.

When a density matrix is not a pure state, it is called a **mixed state**. These mixed states can be seen as the equivalent of a random variable over pure states, and are often denoted ρ or σ. As before these states may be accompanied by a superscript ρ^A to denote that ρ is a matrix in $\mathcal{D}(\mathcal{H}_A)$. The uniform distribution over all states is known as the **maximally mixed state**, and is defined as $\tau^A = \frac{\mathbb{I}}{|A|}$.

Theorem 2.1 (Spectral Decomposition, Corollary 1.4 in [Wat18]). *Any mixed state $\rho \in \mathcal{D}(\mathcal{H}_A)$ can be written as*

$$\rho = \sum_i p(i)\phi_i,$$

for some probability distribution p and some set of pure states $\{\phi_i\}$.

Not all states can be written as a tensor product of qubits. An example of such a state is

$$|\phi^+\rangle^{AA'} = \frac{1}{\sqrt{|A|}} \sum_{x \in \{0,1\}^{\log(|A|)}} |xx\rangle,$$

for any $\mathcal{H}_{AA'} = \mathcal{H}_A \otimes \mathcal{H}_{A'}$ with $\mathcal{H}_A = \mathcal{H}_{A'}$. When a pure state on a composite space cannot be written as a tensor product of two states it is called **entangled**, and the state $\phi^{+AA'} = |\phi^+\rangle\langle\phi^+|^{AA'}$ is known as the **maximally entangled state**. A mixed state is entangled when at least one of the pure states in its spectral decomposition is entangled.

2.2.3 Registers

In order to store quantum states and refer to them in the description of an algorithm, we use the notion of registers. A **register** A can store a density matrix on a Hilbert space $\mathcal{H}_A$. One can think of registers as the quantum version of a variable. Given two registers A and B we will denote the combined register as AB, which stores a density matrix on $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ and can be seen as a variable that contains both the contents of A and B. For this thesis, it is enough to think about registers as variables as is done above, however, a more formal definition is given in [Wat18].

2.3 Operations on States

The previous section introduced the quantum analogs of variables and values for these variables, but of course, these are of little use without ways to manipulate them. In this section, we discuss how one can manipulate and observe quantum states.

2.3.1 Gates

On the most basic level, one can manipulate quantum data by the use of **gates**. When thinking of a pure state as a vector, manipulations of this state are represented by matrices. Since states are represented by norm-1 vectors, this norm should not be changed by the gates and thus the gates are represented by unitary matrices. Unitary matrices satisfy $UU^\dagger = U^\dagger U = \mathbb{I}$ and are norm and inner-product preserving. If $|\psi\rangle^A$ is some state and $U_1, U_2 \in \mathcal{B}(\mathcal{H}_A)$ are unitary matrices then U_1U_2 is also unitary and $U_1|\psi\rangle$ is a valid pure state. To apply a unitary U to a density matrix one performs the map

$$\rho \mapsto U\rho U^\dagger.$$

An example of a gate is the Pauli X gate, which performs the classical 'NOT' operation:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

This gate satisfies $X|0\rangle = |1\rangle$ and $X|1\rangle = |0\rangle$.

We will assume that, when a unitary U^A is applied to a state on a space $\mathcal{H}_{AB}$ rather than $\mathcal{H}_A$, it acts as identity on the B space. This means we write $U^A \rho^{AB} U^{\dagger A}$ to mean $(U^A \otimes \mathbb{I}^B)\rho^{AB}(U^{\dagger A} \otimes \mathbb{I}^B)$.

2.3.2 Measurements

After applying some gates to a state, one might want to know the result. In classical computation, the outcome of a computation is straightforward to observe, but in quantum computation, this is not the case. In general, it is not possible to observe the value of a quantum state, but it is possible to deduce some information about a given state with the help of a **measurement**. The simplest type of measurement is a **projective measurement**, which is described by a set of non-zero projection matrices $\{\Pi_1, \ldots, \Pi_n\}$. The outcome of measuring ρ with such a measurement is i with probability $\mathrm{Tr}\left[\Pi_i \rho\right]$ for all i. After measuring outcome i the state ρ **collapses** to $\frac{\Pi_i \rho \Pi_i}{\mathrm{Tr}[\Pi_i \rho]}$, which means that, if ρ was in register A and this register is measured yielding outcome i, the register now contains $\frac{\Pi_i \rho \Pi_i}{\mathrm{Tr}[\Pi_i \rho]}$.

On a register of n qubits, the **computational basis measurement** is the measurement $\{\Pi_x = |x\rangle\langle x| \mid x \in \{0,1\}^n\}$. This measurement has the effect of reducing any state to a classical pure state and has no effect on classical pure states. Whenever a quantum state is given to a classical algorithm, which cannot accept quantum states, it is assumed that instead the state is measured in the computational basis and the outcome is given to the classical algorithm.

Sometimes a projective measurement is not fine-grained enough for the goal at hand, and a different measurement is needed. This type of measurement is called a **positive-operator valued measure (POVM) measurement**, and is described by a set $\{M_i \mid i \in \mathbb{N}\}$, such that all M_i are positive semi-definite and $\sum_i M_i = \mathbb{I}$. One can think of these measurements as appending extra qubits to the system and then performing a projective measurement. Similar to the projective measurements, the probability of an outcome M_i on a state ρ is $\mathrm{Tr}\left[M_i \rho\right]$, although the post-measurement state depends on the implementation of the POVM, which is often not specified. If no specific implementation is given for the POVM, then it is assumed the quantum state is destroyed in the process of measuring and the register that it was in cannot be used in further computation.

2.3.3 Partial trace

Because of entanglement it is not always possible to write a state ρ on $\mathcal{H}_{AB}$ as a tensor product of two states on $\mathcal{H}_A$ and $\mathcal{H}_B$. However, in some settings an agent might only have access to one of these registers and we would still like to argue about the content of the register from his point of view. For this reason we have the **partial trace**. For any space $\mathcal{H}_{X_1,\ldots,X_n} = \mathcal{H}_{X_1} \otimes \cdots \otimes \mathcal{H}_{X_n}$ and for all i, the partial trace over $\mathcal{H}_{X_i}$ is defined as the unique linear operator that satisfies

$$\mathrm{Tr}_{X_i}\left[\rho_1^{X_1} \otimes \cdots \otimes \rho_n^{X_n}\right] = \mathrm{Tr}\left[\rho_i^{X_i}\right] \rho_1^{X_1} \otimes \cdots \otimes \rho_{i-1}^{X_{i-1}} \otimes \rho_{i+1}^{X_{i+1}} \otimes \cdots \otimes \rho_n^{X_n}.$$

Note that the partial trace can transform pure states into mixed states, for example $\mathrm{Tr}_A\left[\phi^{+AA'}\right] = \tau^{A'}$. If a pure state ψ^{AB} is such that $\mathrm{Tr}_B\left[\psi\right] = \rho^A$, then ψ is a **purification** of ρ. It follows straightforward from Theorem 2.1 that any mixed state has a canonical purification.

Corollary 2.2. *For any state ρ^A, there exists a pure state ψ_ρ that is a purification of ρ.*

Proof. Observe that $\mathrm{Tr}\left[|i\rangle\langle j|\right] = \delta_{i,j}$, where

$$\delta_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

By Theorem 2.1 we have $\rho = \sum_i p(i)\phi_i$. Let

$$|\psi_\rho\rangle^{AB} = \sum_i \sqrt{p(i)}|\phi_i\rangle|i\rangle,$$

then

$$\begin{aligned}
\mathrm{Tr}_B\left[\psi_\rho\right] &= \mathrm{Tr}_B\left[\sum_{i,j}\sqrt{p(i)}\sqrt{p(j)}|\phi_i\rangle|i\rangle\langle\phi_j|\langle j|\right] \\
&= \mathrm{Tr}_B\left[\sum_{i,j}\sqrt{p(i)}\sqrt{p(j)}|\phi_i\rangle\langle\phi_j|^A \otimes |i\rangle\langle j|^B\right] \\
&= \sum_{i,j}\sqrt{p(i)}\sqrt{p(j)}|\phi_i\rangle\langle\phi_j|^A \otimes \mathrm{Tr}\left[|i\rangle\langle j|^B\right] \\
&= \sum_{i,j}\sqrt{p(i)}\sqrt{p(j)}|\phi_i\rangle\langle\phi_j|^A \otimes \delta_{i,j} \\
&= \rho.
\end{aligned}$$

□

2.3.4 Channels

Using the methods described above, one can transform a quantum state in many different ways. In order to reason about these transformations without knowing their exact nature, we make use of quantum channels, which can represent any quantum operation that maps one state to another. A **quantum channel** is any completely positive trace-preserving (CPTP) map $\Lambda^{A\to B} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$, which means that for any positive semi-definite matrix M^A, $\Lambda(M)$ is also positive semi-definite and $\mathrm{Tr}\left[M\right] = \mathrm{Tr}\left[\Lambda(M)\right]$. Furthermore, for any Hilbert space $\mathcal{H}_C$ and any positive semi-definite matrix M^{AC}, $(\Lambda\otimes\mathbb{I}^C)(M)$ is also positive semi-definite. Note that we again use gray superscripts to indicate the space or register that Λ acts on. we consider a quantum channel Λ^A as $\Lambda^A\otimes\mathrm{id}^B$ when it is applied to a state on $\mathcal{H}_{AB}$ or composed with a channel Λ^B.

These quantum channels can take many forms. An important example of a quantum channel is the constant-ρ channel, which is defined for any density matrix ρ^A as

$$\langle\rho\rangle^A(X) = \mathrm{Tr}\left[X\right]\rho^A.$$

When describing some quantum channels we will make use of completely positive trace non-increasing (CPTNI) maps $\Lambda^{A\to B} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$, which means that for any positive semi-definite matrix M, $\Lambda(M)$ is also positive semi-definite and $\mathrm{Tr}\left[M\right] \geq \mathrm{Tr}\left[\Lambda(M)\right]$. These can be used as the description of a part of a quantum channel and often we will consider a set of such maps whose sum is CPTP.

2.4 Norms on States and Channels

In order to compare quantum states and channels, we introduce some norms in this section and show how they can be used to measure similarity.

2.4.1 Trace norm and trace distance

An effective tool for measuring distinguishability between two quantum states is the **trace norm**. The trace norm, or Schatten 1-norm, is defined as

$$\|M\|_1 = \mathrm{Tr}\left[\sqrt{M^\dagger M}\right].$$

This norm is used to define the **trace distance** between density matrices, which is

$$D(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\|_1 .$$

A well-known property of this distance measure, proven in [NC02], is that it is directly related to the maximum probability of distinguishing two density matrices with a POVM:

$$D(\rho, \sigma) = \max_{0 \leq P \leq \mathbb{I}} \mathrm{Tr}\left[P(\rho - \sigma)\right],$$

where the condition $0 \leq P \leq \mathbb{I}$ means that P is both positive semi-definite (≥ 0) and that $\mathbb{I} - P$ is also positive semi-definite ($\leq \mathbb{I}$). This is equivalent to the condition that $\{P, \mathbb{I} - P\}$ is a valid POVM. If one is given an unknown state ρ, which is either ρ_1 or ρ_2 each with probability $\frac{1}{2}$ and uses the POVM $\{P, \mathbb{I} - P\}$ to determine which of the two was given, where P is the matrix that obtains the maximum in $D(\rho_1, \rho_2)$, then the probability of successfully guessing is

$$\begin{aligned}
\Pr\left[\text{Outcome correct}\right] &= \frac{1}{2}\Pr\left[\text{Guess } \rho_1 | \rho = \rho_1\right] + \frac{1}{2}\Pr\left[\text{Guess } \rho_2 | \rho = \rho_2\right] \\
&= \frac{1}{2}\left(\mathrm{Tr}\left[P\rho_1\right] + \mathrm{Tr}\left[(\mathbb{I} - P)\rho_2\right]\right) \\
&= \frac{1}{2}\left(\mathrm{Tr}\left[P(\rho_1 - \rho_2)\right] + \mathrm{Tr}\left[\rho_2\right]\right) \\
&= \frac{1}{2}\left(1 + D(\rho_1, \rho_2)\right).
\end{aligned}$$

Intuitively this makes sense because when $\rho_1 = \rho_2$ then one can only guess and thus the probability of guessing correct is $\frac{1}{2}$. However, when ρ_1 and ρ_2 can be perfectly distinguished and are thus far away in trace distance, then one always guesses correctly, thus this probability is 1.

2.4.2 Induced trace norm

We have seen that the trace norm is an effective tool to quantify distinguishability between states. Motivated by this fact one might be convinced that it can also be used to effectively measure distinguishability between quantum channels. We define the **induced trace norm** of any operator $\Gamma : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ as

$$\|\Gamma\|_1 = \max_\rho \|\Gamma(\rho)\|_1 ,$$

where the maximum is taken over all density matrices ρ. For any two quantum channels $\Lambda_1^{A\to B}, \Lambda_2^{A\to B}$, the distance $\frac{1}{2}\|\Lambda_1 - \Lambda_2\|_1$ is related to the maximum probability of distinguishing between these two channels in the same way as the trace distance could distinguish between states, however with the assumption that no side information, in the form of additional registers unaffected by Λ_1 or Λ_2, is used to distinguish between these two channels.

In the following example we show that disregarding side information can in some cases significantly reduce the probability of distinguishing two channels. Consider the case where the channel Λ_1^A or Λ_2^A is applied with probability λ or $1-\lambda$ respectively, where

$$\begin{aligned} n &:= |A| \geq 2 \\ \lambda &:= \frac{n+1}{2n} \\ \Lambda_1(X) &:= \frac{1}{n+1}\left(\mathrm{Tr}\left[X\right]\mathbb{I} + X^T\right) \\ \Lambda_2(X) &:= \frac{1}{n-1}\left(\mathrm{Tr}\left[X\right]\mathbb{I} - X^T\right) \end{aligned}$$

and X^T is the transpose of X with respect to the computational basis, that is, $|i\rangle\langle j|^T = |j\rangle\langle i|$ for all i, j. The channels Λ_1 and Λ_2 are known as Holevo-Werner channels, and in [Wat18] it is proven that

$$\|\lambda\Lambda_1 - (1-\lambda)\Lambda_2\|_1 = \frac{1}{n}.$$

Intuitively one might argue that it makes sense that the distance between these channels is small, because both channels are very similar to the completely depolarizing channel $\langle\tau^A\rangle$. However, if one allows side information in the form of an additional register A' with $|A| = |A'|$ then it is also shown in [Wat18] that

$$\left\|\lambda(\Lambda_1 \otimes \mathbb{I}^{A'}) - (1-\lambda)(\Lambda_2 \otimes \mathbb{I}^{A'})\right\|_1 = 1,$$

which shows that it is possible that channels can be distinguished better with the use of side information.

2.4.3 Diamond norm

In order to allow side information, we make use of the **diamond norm**, also known as the completely bounded trace norm. The diamond norm of an operator $\Gamma : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ is defined as

$$\|\Gamma\|_\diamond = \left\|\Gamma \otimes \mathbb{I}^{A'}\right\|_1,$$

where $\mathcal{H}_{A'}$ is some Hilbert space such that $|A| = |A'|$. This requirement on the size of the side information register A' is necessary and sufficient in the sense that for any smaller register there are channels with a strictly smaller norm than the diamond norm and with any larger register one cannot achieve a larger norm as is shown in [Wat18].

The diamond norm is widely used as the tool to measure the similarity between channels. It satisfies a number of useful properties, such as

(Triangle Inequality) $\|\Lambda_1 + \Lambda_2\|_\diamond \leq \|\Lambda_1\|_\diamond + \|\Lambda_2\|_\diamond$

(Submultiplicativity) $\|\Lambda_1 \circ \Lambda_2\|_\diamond \leq \|\Lambda_1\|_\diamond \|\Lambda_2\|_\diamond$

(Distribution over $\otimes$) $\|\Lambda_1 \otimes \Lambda_2\|_\diamond = \|\Lambda_1\|_\diamond \|\Lambda_2\|_\diamond$.

Furthermore the diamond norm satisfies $\|\Lambda'\|_\diamond \leq 1$ for all CPTNI Λ', where equality holds if Λ' is CPTP. Lastly $\frac{1}{2}\|\Lambda_1 - \Lambda_2\|_\diamond$ is a distance measure between channels Λ_1 and Λ_2, thus $0 \leq \|\Lambda_1 - \Lambda_2\|_\diamond \leq 2$.

CHAPTER 3

Non-Malleability and Authentication

In this chapter, we explain the essence of non-malleability and provide context on non-malleability from the existing literature. In Section 3.1 we explain what the concept behind non-malleability is and why it is important. Afterward, we define encryption schemes for both the classical and the quantum setting in Section 3.2. In Sections 3.3 and 3.4 we discuss a number of different notions of non-malleability in both the classical and the quantum setting.

3.1 Importance of Non-Malleability

When communicating confidential information one often puts a high priority on the privacy of the communication channel used. This prioritization means that when a message is encrypted and then transmitted no intermediate party can read this message or even partially deduce its contents. This property where an unbounded attacker is given an encrypted message and can deduce nothing about the contents of this message is known as **information-theoretic security**. A well-known example of a scheme that is secure in this way is the one-time pad scheme, defined as

$$\text{Encrypt}(m, k) = \text{Decrypt}(m, k) = k \oplus m.$$

Here $\oplus$ is the bitwise xor function, which is defined as $a \oplus b = (a+b) \mod 2$ for $a, b \in \{0,1\}$ and $x_1 \dots x_n \oplus y_1 \dots y_n = (x_1 \oplus y_1) \dots (x_n \oplus y_n)$, with $x = x_1 \dots x_n$, $y = y_1 \dots y_n$ and $x, y \in \{0,1\}^n$. If a key $k \in \{0,1\}^n$ is chosen uniformly at random, then one cannot deduce any information about an encrypted message $c = m \oplus k$, since any other message m' can be encrypted to the same c with key $k' = m' \oplus c$ and thus without knowing the key all possible messages are equally likely to be the encoded message.

However, despite not being able to deduce any information about the content, an attacker that intercepts the message can still change its contents in a structural way, by choosing some a and performing $c \oplus a$. If some message m is encrypted, attacked and then decrypted in this way the result is $k \oplus k \oplus m \oplus a = m \oplus a$. Informally an encryption scheme is non-malleable if no attack can be performed that meaningfully translates one encrypted message into another.

In many applications, non-malleability can be a desired property. For example, consider the setting where person A is communicating with his bank, and sends the message

"send $100 to person B", then it would be bad if a person C can transform this into "send $500 to person B" or even "send $100 to person C".

3.2 Encryption Schemes

In order to reason about encryption schemes we first give a formal definition. In this section we give four variants of this definition, distinguishing between symmetric- and public-key encryption and between the classical and quantum setting. We assume that the reader is familiar with basic notions of computational complexity, such as the ones discussed in [Pap03].

3.2.1 Algorithms

In a classical setting we describe the encryption and decryption procedure in the form of an algorithm. An **algorithm** in this setting is a partial function computed by a Turing Machine. We write $y \leftarrow A(x_1, \ldots, x_n)$ to mean that y is the result of running an algorithm A on inputs $x_1, \ldots, x_n$. A **probabilistic algorithm** is an algorithm A that has input arguments $x_1, \ldots x_n$ and an input argument r representing the randomness. For some probabilistic algorithm A we write $y \leftarrow A(x_1, \ldots, x_n)$ to denote the action of picking r uniformly random and then performing $y \leftarrow A(x_1, \ldots, x_n, r)$. If for a probabilistic algorithm A and some values $x_1, \ldots, x_n$ there exists an r such that $y \leftarrow A(x_1, \ldots, x_n, r)$, then we say y **can be output by** $A(x_1, \ldots, x_n)$, which we denote $y \leftarrow A(x_1, \ldots, x_n)$. We call an algorithm **deterministic** to explicitly state that it is not probabilistic. If we write $y \leftarrow \alpha$ and α is not an algorithm then this simply means that y is assigned the value α.

To describe the efficiency of an algorithm, we consider the amount of steps a Turing Machine has to take to compute the outcome of this algorithm. We say an algorithm $A(x_1, \ldots, x_n)$, computed by a Turing Machine T, **runs in time** t if T terminates on input $x_1, \ldots, x_n$ in less than or equal to t steps.

For simplicity we write $x \xleftarrow{\$} S$ to mean that x is chosen uniformly at random from some finite set S and $x \xleftarrow{p} S$ if x is chosen from some set S according to the probability distribution p.

3.2.2 Classical encryption schemes

An encryption scheme describes how to transform a message (**plaintext**) into an encoded message (**ciphertext**). As stated before we distinguish four variants of encryption schemes, two of which are discussed in this section. The simplest form of encryption is symmetric-key encryption, which means that encryption and decryption are done with the same key. For simplicity we use Enc_k to mean $\text{Enc}(k, \cdot)$ and Dec_k to mean $\text{Dec}(k, \cdot)$.

Definition 3.1. *A **symmetric-key encryption scheme (SKES)** is a triple* $(\text{KeyGen}, \text{Enc}, \text{Dec})$ *such that:*

- KeyGen *is a probabilistic algorithm that takes as input* $n \in \mathbb{N}$ *in unary and outputs a key* $k \in \{0,1\}^*$*. Here* n *is known as the **security parameter** of the scheme.*
- Enc *is a probabilistic algorithm that takes as inputs* $x, k \in \{0,1\}^*$*, where* x *is the message to encode and* k *the key, and outputs* $y \in \{0,1\}^*$*.*

- Dec *is a deterministic algorithm that takes as inputs* $y, k \in \{0,1\}^*$*, where* y *is the encrypted message to decode and* k *the key, and outputs either* $x \in \{0,1\}^*$ *or a symbol* $\perp$ *if* y *is not a valid ciphertext.*

- *The scheme is* ***correct****, which means it satisfies* $\mathrm{Dec}_k(\mathrm{Enc}_k(x)) = x$ *for all* $x \in \{0,1\}^*$ *,* $n \in \mathbb{N}$ *and* $k \leftarrow \mathrm{KeyGen}(1^n)$ *such that* $\mathrm{Enc}_k(x)$ *is defined.*

- *The scheme is* ***efficient****, which means that for all* $n \in \mathbb{N}$*,* $\mathrm{KeyGen}(1^n), \mathrm{Enc}_k(x)$ *and* $\mathrm{Dec}_k(y)$ *run in time* $p(n)$ *for some polynomial* p *and all* $x, y \in \{0,1\}^*$ *and* $k \leftarrow \mathrm{KeyGen}(1^n)$*.*

Note that the efficiency requirement also limits the maximum size of x and y for which Enc and Dec are defined to $p(n)$, since any Turing Machine that runs in time $p(n)$ can read at most $p(n)$ input bits. We mostly use **fixed-length** encryption schemes, which means that Enc is only defined for x, k of a length that only scales with the security parameter. This means that when the security parameter is fixed we can assume that all x and k are of the same length.

Symmetric-key encryption is widely used because it is often simple to construct a scheme that is both efficient and secure, however, it has the shortcoming that both the sender and receiver need to have access to the same key beforehand. If the sender and receiver do not already share such a key then this setup can be a problem, for which public-key encryption is a possible solution. The idea of public-key encryption is that a key consists of two parts, the public and the private part. The public part is publicly announced and known to all parties, including possible attackers, but the private part is only known to the receiver of a message. In such a setting one can use the public key to encode a message, send this message to the receiver, and the receiver can use his private key to decode the message.

Definition 3.2 ([BS99])**.** *A* ***public-key encryption scheme (PKES)*** *is a triple* $(\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ *such that:*

- KeyGen *is a probabilistic algorithm that takes as input* $n \in \mathbb{N}$ *in unary and outputs a pair* $(sk, pk) \in (\{0,1\}^*)^2$*. Here* n *is known as the* ***security parameter*** *of the scheme.*

- Enc *is a probabilistic algorithm that takes as inputs* $x, pk \in \{0,1\}^*$*, where* x *is the message to encode and* pk *the public key, and outputs* $y \in \{0,1\}^*$*.*

- Dec *is a deterministic algorithm that takes as inputs* $y, sk \in \{0,1\}^*$*, where* y *is the encrypted message to decode and* sk *the secret key, and outputs either* $x \in \{0,1\}^*$ *or a symbol* $\perp$ *if* y *is not a valid ciphertext.*

- *The scheme is* ***correct****, which means it satisfies* $\mathrm{Dec}_{sk}(\mathrm{Enc}_{pk}(x)) = x$ *for all* $x \in \{0,1\}^*$*,* $n \in \mathbb{N}$ *and* $(pk, sk) \leftarrow \mathrm{KeyGen}(1^n)$ *such that* $\mathrm{Enc}_{pk}(x)$ *is defined.*

- *The scheme is* ***efficient****, which means that for all* $n \in \mathbb{N}$*,* $\mathrm{KeyGen}(1^n), \mathrm{Enc}_{pk}(x)$ *and* $\mathrm{Dec}_{sk}(y)$ *run in time* $p(n)$ *for some polynomial* p *and all* $x, y \in \{0,1\}^*$ *and* $(pk, sk) \leftarrow \mathrm{KeyGen}(1^n)$*.*

In the private-key case, an attacker does not know the key and it is possible to create a scheme, for example one-time pad, where an attacker also has no hope of obtaining this key from an encrypted message without additional information. In the public-key case, however, it is not possible to construct such a scheme. Since the attacker has access to

the public key and thus the encryption procedure, he can simply encrypt his own message and then attempt to decrypt it with all possible secret keys until he succeeds, in which case he probably obtained the secret key corresponding to the public key. For this reason, the KeyGen algorithm is equipped with a security parameter, which is often used as a parameter for the minimal amount of time an attacker needs to perform a meaningful attack.

3.2.3 Quantum algorithms

Similar to the classical case, we also describe quantum encryption schemes by means of quantum algorithms. Instead of using a Turing Machine, a quantum algorithm is computed by a family of quantum circuits, since each quantum circuit has a fixed input space. A **quantum algorithm** A is a family of quantum channels $\{\Lambda_n \mid n \in \mathbb{N}\}$, where each Λ_n in this family has a n-qubit input space and is implemented by some quantum circuit Q_n. This implementation means that running Q_n on some state ρ produces $\Lambda_n(\rho)$. Each quantum circuit is built from unitary gates and measurements, and which circuit is executed is implicitly decided by the size of the state given to A.

We write $\sigma \leftarrow A(\rho)$ to denote that σ is the result of applying Λ_n to some state $\rho \in \mathcal{D}(\mathcal{H}_B)$ with $|B| = 2^n$. To deal more efficiently with registers we write $R \leftarrow A(S)$ to denote that Λ_n is applied to the contents of register S with $|S| = 2^n$ and the result is stored in register R. We call an algorithm a **classical algorithm** to explicitly state it is not a quantum algorithm, but an algorithm as defined in Section 3.2.1.

Often a quantum algorithm is run on a register from an infinite family of registers, because the size of the contents of these registers scales with some parameter, such as the security parameter. In this case, we refer to a family of registers $\{M_i\}_{i\in\mathbb{N}}$ as the register M, where it is understood that the register with the correct size for contents is chosen implicitly. If a quantum algorithm A takes a classical argument x then this argument is simply converted to the classical state $|x\rangle\langle x|$ before the corresponding circuit is applied.

Quantum algorithms also have an execution time, similar to classical algorithms. For any $\rho \in \mathcal{H}_B$, with $|B| = 2^n$, a quantum algorithm $A(\rho)$, computed by the family of quantum circuits $\{Q_n \mid n \in \mathbb{N}\}$, **runs in time** t if the size, or number of gates and measurements, of Q_n is less than t. Because there is no inherent relation between each of the Q_n it is in some cases possible to store information in the definition of Q_n, which effectively circumvents a significant part of the computation and thus significantly reduces the time needed for the calculation. To prevent this we only allow quantum algorithms to be computed by **uniform** circuit families, which means that there exists a classical algorithm P_A and a polynomial p such that for all n, $P_A(n)$ outputs a description of Q_n and $P_A(n)$ runs in time $p(n)$.

3.2.4 Quantum encryption schemes

When dealing with quantum computation, one might want to send a quantum state over a public channel in a secure way. In this setting one can choose from a number of different options when defining encryption schemes, for example in [BB14] a definition is considered where the key and ciphertext are quantum and the plaintext is classical and in [OTU00] a notion is defined where these three are all classical. In this thesis, we discuss schemes where the key is classical, but the plaintext and ciphertext are quantum states, which means one can send a quantum state using only a classical key. Using classical keys is useful because the setup for such a scheme, for example the key distribution, can be done

classically.

Throughout this thesis, we use M as a register for the plaintext state and C as a register for the ciphertext.

Definition 3.3 (Definition 1 in [AM17])**.** *A **symmetric-key quantum encryption scheme (SKQES)** is a triple* $(\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ *such that:*

- KeyGen *is a classical probabilistic algorithm that given a security parameter* $n \in \mathbb{N}$ *outputs a key* $k \in \{0,1\}^*$*.*
- Enc *is a quantum algorithm which takes as input a classical key* k *and a quantum state in register* M *and outputs a quantum state in register* C*.*
- Dec *is a quantum 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* $|\bot\rangle\langle\bot|^{\bot}$*.*
- *The scheme is **correct**, which means* $\left\|\mathrm{Dec}_k \circ \mathrm{Enc}_k - \mathrm{id}^{M \to M \oplus \bot}\right\|_\diamond \leq \mathrm{negl}(n)$ *for all* $k \leftarrow \mathrm{KeyGen}(1^n)$*.*
- *The scheme is **efficient**, which means that for all* $n \in \mathbb{N}$*,* $\mathrm{KeyGen}(1^n)$*,* $\mathrm{Enc}_k(M)$ *and* $\mathrm{Dec}_k(C)$ *run in time* $p(n)$ *for some polynomial* p *and all* $k \leftarrow \mathrm{KeyGen}(1^n)$*. Furthermore* $|M| \leq |C| \leq 2^{q(n)}$ *for some polynomial* q*.*

We only consider schemes that are **fixed-length**, meaning that $|M|$ and $|C|$ are a function of the security parameter n. Note that KeyGen runs in time polynomial in n and can thus only output keys of a size polynomial in n.

We adopt the convention that every honest party applies the measurement $\{|\bot\rangle\langle\bot|, \mathbb{I} - |\bot\rangle\langle\bot|\}$ after running Dec, and denote with $\mathrm{Dec}_k(C) \neq \bot$ the event that this measurement did not measure $|\bot\rangle\langle\bot|$ and thus produced a valid plaintext. Because of this convention we often state that the output space of Dec is $\mathcal{D}(\mathcal{H}_M)$ although it is technically $\mathcal{D}(\mathcal{H}_M \oplus \mathcal{H}_\bot)$, where $\mathcal{H}_\bot = \mathbb{C}|\bot\rangle$.

For public-key quantum encryption, we can make similar adaptations to the definition as are made in the classical case.

Definition 3.4. *A **public-key quantum encryption scheme (PKQES)** is a triple* $(\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ *such that:*

- KeyGen *is a classical probabilistic algorithm that given a **security parameter*** $n \in \mathbb{N}$ *outputs a pair of keys* $(pk, sk) \in (\{0,1\}^*)^2$*.*
- Enc *is a quantum 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*.*
- Dec *is a quantum 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* $|\bot\rangle\langle\bot|^{\bot}$*.*
- *The scheme is **correct**, which means* $\left\|\mathrm{Dec}_{sk} \circ \mathrm{Enc}_{pk} - \mathrm{id}^{M \to M \oplus \bot}\right\|_\diamond \leq \mathrm{negl}(n)$ *for all* $(pk, sk) \leftarrow \mathrm{KeyGen}(1^n)$*.*
- *The scheme is **efficient**, which means that for all* $n \in \mathbb{N}$*,* $\mathrm{KeyGen}(1^n)$*,* $\mathrm{Enc}_{pk}(M)$ *and* $\mathrm{Dec}_{sk}(C)$ *run in time* $p(n)$ *for some polynomial* p *and all* $(pk, sk) \leftarrow \mathrm{KeyGen}(1^n)$*. Furthermore* $|M| \leq |C| \leq 2^{q(n)}$ *for some polynomial* q*.*

Here we again have that we only consider fixed-length schemes, that all keys are of polynomial size and that we follow the same convention of measuring $\mathrm{Dec}_{sk}(C) \neq \bot$.

3.2.5 Attacks and effective maps

An encryption scheme as defined above is not inherently secure in any way. To show that a scheme is secure, or not secure, one looks at possible attacks that can be performed on the ciphertext. We consider attacks in the setting where the objective of an attack is to modify or learn about one given ciphertext. These attacks are represented by algorithms that act between encryption and decryption. A **classical attack** on a classical encryption scheme Π is a probabilistic algorithm A, which takes a ciphertext of Π and possibly some side information as input and outputs a string, which is possibly a different ciphertext. Note that attacks are always defined with a target encryption scheme, and we call an attack **efficient** if it runs in time $p(n)$, where p is some polynomial and n is the security parameter of the target encryption scheme.

In the quantum setting the concept of an attack is similar. A **quantum attack** on a quantum encryption scheme Π is a quantum algorithm A, which has input registers BC and output registers $\hat{B}C$. Here B and $\hat{B}$ contain the attacker's side information before and after the attack respectively. An attack A on Π is **efficient** if A runs in time $p(n)$, where p is some polynomial and n is the security parameter of Π.

In an information-theoretic setting, where the efficiency of an attack is disregarded, we may also think of an attack A as a quantum channel $\Lambda_A^{BC\to\hat{B}C}$, which is $\Lambda_{|BC|}$ from the family of channels defining A.

For all versions of encryption schemes, it is the case that the effect that an attack has on the plaintext depends not only on the plaintext that is chosen but also on the key. When the objective of an attack is to change the plaintext, it is not useful to look what the maximum impact is of an attack, because one can always guess the key and perform some attack on the plaintext under the assumption that the guessed key is correct. Of course, the probability of guessing the correct key is negligible, but if guessed correctly the plaintext can be changed in any way, thus if one only considers the maximal effect of an attack no scheme is secure. For this reason, we look at the average effect, or **effective map**, of an attack, where the average is taken over all keys. For some (classical or quantum) symmetric-key encryption scheme $\Pi = (\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ and some attack A on this scheme, the effective map of A is defined as

$$\tilde{A} = \mathop{\mathbb{E}}_{k\leftarrow\mathrm{KeyGen}(1^n)}[\mathrm{Dec}_k \circ A \circ \mathrm{Enc}_k].$$

For public-key encryption schemes, the definition is identical, but the expected value is taken over all $(pk, sk) \leftarrow \mathrm{KeyGen}(1^n)$. Note that the effective map is different for each value of the security parameter n, and describes the average effect an attack has on a given plaintext combined with possible side information. It will prove useful to also define

$$\mathrm{Enc}_K = \mathop{\mathbb{E}}_{k\leftarrow\mathrm{KeyGen}(1^n)}[\mathrm{Enc}_k]$$
$$\mathrm{Dec}_K = \mathop{\mathbb{E}}_{k\leftarrow\mathrm{KeyGen}(1^n)}[\mathrm{Dec}_k],$$

which are encryption and decryption operations averaged over all keys.

3.2.6 Characterization of quantum encryption schemes

Quantum encryption schemes as defined above can be difficult to analyze because very little information about the structure of the encryption and decryption algorithms is

specified in the definition. In [AM17], a possible solution to this problem is presented in the form of a characterization of the encryption and decryption algorithms.

Theorem 3.1 (Lemma B.9 in [AM17]). *Let* $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$ *be a* SKQES*, then* Enc *and* Dec *have the following form:*

$$\text{Enc}_k(X^M) = V_k(X^M \otimes \sigma_k^T)V_k^\dagger$$
$$\text{Dec}_k(Y^C) = \text{Tr}_T\left[P_{\sigma_k}^T(V_k^\dagger Y^C V_k)P_{\sigma_k}^T\right] + \hat{D}_k\left[\bar{P}_{\sigma_k}^T(V_k^\dagger Y^C V_k)\bar{P}_{\sigma_k}^T\right].$$

Here σ_k *is a state on register* T*,* V_k *is a unitary and* $\hat{D}_k$ *is a quantum channel. Furthermore* P_{σ_k} *is the projector onto the support of* σ_k*, that is, if* $\sigma_k = \sum_i \alpha_i |\phi_i\rangle\langle\phi_i|$*, then* $P_{\sigma_k} = \sum_i |\phi_i\rangle\langle\phi_i|$ *and* $\bar{P}_{\sigma_k} = \mathbb{I} - P_{\sigma_k}$*.*

Since each σ_k can be seen as a probability distribution over pure states, this result can be refined further, as is done in [AGM18].

Corollary 3.2 (Corollary 1 in [AGM18]). *Let* $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$ *be a* SKQES*, then for every* k *there exists a probability distribution* $p_k : \{0,1\}^t \to [0,1]$ *and a family of quantum states* $|\psi_{k,r}\rangle^T$ *such that* Enc_k *is equivalent to the following algorithm:*

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

Here V_k *and* T *are defined as in Theorem 3.1, and* $t = \log(|T|)$ *is the number of qubits in* T*.*

These results are proven using an information-theoretic approach, and thus disregard the efficiency of the encryption scheme when implemented in this form. Whether this characterization can always be implemented efficiently is still an open question. In case such an efficient implementation is required for a result, we will explicitly state this using Condition 1.

Condition 1 (Condition 1 in [AGM18]). *Let* Π *be a* SKQES *with security parameter* n*, and let* p_k*,* $|\psi_{k,r}\rangle$*, and* V_k *be as defined in Corollary 3.2. We say that* Π *satisfies Condition 1 if, for all but a negligible fraction of* k *and* r*, there exist a polynomial* p *and quantum algorithms that run in time* $p(n)$ *for (i.) sampling from* p_k*, (ii.) preparing* $\psi_{k,r}$*, and (iii.) implementing* V_k*.*

3.3 Classical Non-Malleability

In the classical setting the notion of non-malleability has been studied in great detail, for example in [KPT11] and [BS99]. In this thesis, we mostly consider comparison-based non-malleability as defined in [BS99], although we also briefly touch on simulation-based non-malleability.

Comparison-based non-malleability can be viewed as a promise that an adversary, given some ciphertext, cannot construct any meaningful relation between the original plaintext and the decryption of the ciphertext after the attack. To test for this form of non-malleability, an adversary is challenged to distinguish between two settings.

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 sampled from the message distribution and the decryption of this same vector. If an adversary can achieve this with non-negligible probability, then intuitively the adversary was able to structurally change an encrypted message, which would indicate that the scheme is malleable.

For any PKES $\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})$ with security parameter n and any pair of algorithms, or **adversary**, $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ we define the following **experiments**, which produce random variables:

Experiment 3.1.
The CNM-Real$(\Pi, \mathcal{A}, n)$ experiment:

1: $(pk, sk) \leftarrow \text{KeyGen}(1^n)$
2: $(M, s) \leftarrow \mathcal{A}_1(pk)$
3: $x \leftarrow M$
4: $y \leftarrow \text{Enc}_{pk}(x)$
5: $(R, \mathbf{y}) \leftarrow \mathcal{A}_2(s, y)$
6: $\mathbf{x} \leftarrow \text{Dec}_{sk}(\mathbf{y})$
7: **return** 1 iff $(y \notin \mathbf{y}) \wedge R(x, \mathbf{x})$

Experiment 3.2.
The CNM-Ideal$(\Pi, \mathcal{A}, n)$ experiment:

1: $(pk, sk) \leftarrow \text{KeyGen}(1^n)$
2: $(M, s) \leftarrow \mathcal{A}_1(pk)$
3: $x, \tilde{x} \leftarrow M$
4: $\tilde{y} \leftarrow \text{Enc}_{pk}(\tilde{x})$
5: $(R, \tilde{\mathbf{y}}) \leftarrow \mathcal{A}_2(s, \tilde{y})$
6: $\tilde{\mathbf{x}} \leftarrow \text{Dec}_{sk}(\tilde{\mathbf{y}})$
7: **return** 1 iff $(\tilde{y} \notin \tilde{\mathbf{y}}) \wedge R(x, \tilde{\mathbf{x}})$

Using these experiments we can define comparison-based non-malleability.

Definition 3.5 (Definition 2 in [BS99] (CNM-CPA))**.** *A* PKES Π *is* ***comparison-based non-malleable for chosen-plaintext attacks (*CNM-*CPA or* 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 there exists a polynomial* p *such that for all* n*:*

- $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *run in time* $p(n)$
- $\mathcal{A}_1$ *outputs a valid message space* M *which can be sampled in time* $p(n)$
- $\mathcal{A}_2$ *outputs a relation* R *computable in time* $p(n)$
- $\mathcal{A}_2$ *outputs a vector* $\mathbf{y}$ *such that* $\perp \notin \text{Dec}_{sk}(\mathbf{y})$

Another notion of non-malleability is discussed in [BS99], called simulation-based non-malleability (SNM). SNM differs from CNM in two key ways, but in the classical case is an equivalent notion to CNM. In the experiments of CNM an adversary is challenged to build a meaningful relation given a ciphertext, but in SNM the relation is fixed and the adversary is simply challenged to find a ciphertext such that its decryption is related to the original plaintext. A scheme is called secure in the sense of SNM if, for any fixed relation, no adversary can construct such a ciphertext more than negligibly better than a simulator could, which is given no input. A third notion is discussed in [BS99] which is also equivalent in the sense of security but is based on indistinguishability instead of non-malleability. Our goal is to compare a possible quantum version of these security notions to other quantum notions of non-malleability. For this reason, we have chosen to study CNM in more detail, since the relation in the definition of SNM depends on not only

plaintext vectors, but also the message distribution, which may be difficult to implement in the quantum setting.

3.4 Quantum Non-Malleability and Authentication

In the quantum setting, the approach taken by previous research is quite different from the notions described in the previous section. Here, the focus is put on symmetric-key non-malleability and authentication, which we discuss in this section.

3.4.1 Evolution of non-malleability

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 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 that is such 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 a flaw in 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. Note that we here consider attacks on the scheme as quantum channels since the efficiency is disregarded.

Definition 3.6 (Theorem 4.4 in [AM17])**.** *A* SKQES $\Pi = (\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ *is* ε***-non-malleable*** *(*ε**-NM**$)$ *if, for any attack* $\Lambda_A^{CB\to C\hat{B}}$*, its effective map* $\tilde{\Lambda}_A^{MB\to M\hat{B}}$ *is such that*

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

where

$$\Lambda_1 = \mathrm{Tr}_{CC'} \left[\phi^{+CC'} \Lambda_A(\phi^{+CC'} \otimes (\cdot)) \right] \qquad \textit{and}$$
$$\Lambda_2 = \mathrm{Tr}_{CC'} \left[(\mathbb{I}^{CC'} - \phi^{+CC'}) \Lambda_A(\phi^{+CC'} \otimes (\cdot)) \right].$$

At first glance, this definition looks very similar to the one described in [ABW09], since it is also based on the fact that the effective map is a combination of identity and a constant map. Note however that the requirements imposed on Λ_1 and Λ_2 ensure that no attack exists which maps a ciphertext into another encryption of the same plaintext, which means this is a notion of ciphertext non-malleability. In the next chapter, we build upon this definition and use it to define a plaintext non-malleability notion.

3.4.2 Authentication

In the symmetric-key setting, another notion can be realized, called authentication. In this setting, it is not only impossible to meaningfully transform 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 3.7 (Definition 2.2 in [DNS12])**.** *A* SKQES II *is* ε***-DNS authenticating*** *(*ε**-DNS***) if, for any attack* $\Lambda_A^{CB\to C\hat{B}}$*, its effective map* $\tilde{\Lambda}_A^{MB\to M\hat{B}}$ *is such that*

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

for some CPTNI *maps* $\Lambda_{acc}, \Lambda_{rej}$ *such that* $\Lambda_{acc} + \Lambda_{rej}$ *is* CPTP*.*

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.

CHAPTER 4

Plaintext Non-Malleability

In this chapter we will build upon the work discussed in Section 3.4 and define non-malleability based on the possible effective maps. In the previous chapter, we touched upon the concepts of ciphertext and plaintext non-malleability, which we will discuss in more detail in Section 4.1. After this we will present a possible definition of plaintext non-malleability in Section 4.2, and show that it differs from NM. In Section 4.3 we will show that plaintext non-malleability is strong enough to build a DNS authenticating scheme. Lastly, we will briefly look at the public key setting in Section 4.4 and motivate a possible definition for plaintext non-malleability in this setting.

4.1 Difference between Ciphertext and Plaintext Non-Malleability

In the previous chapter, we explained that the difference between plaintext and ciphertext non-malleability is that a ciphertext non-malleable scheme does not allow an attacker to transform one ciphertext into another, while a plaintext non-malleable scheme does allow this transformation but only if both the transformed and the original ciphertext decrypt to the same plaintext.

If we look at what this additional freedom for attackers means for the honest parties of an encryption scheme, we observe that this distinction makes little difference. When using a plaintext non-malleable scheme a possible attacker can still only implement, on the plaintext level, the same attacks as it could if the scheme was ciphertext non-malleable, namely a combination of the identity map and the map that results from depolarizing the ciphertext.

If one instead looks at what information a possible attacker might obtain regarding the plaintext, then one again observes a similarity. This similarity is because of the connection between non-malleability and indistinguishability in the quantum setting. Consider the following example, where Π is a SKQES such that an attacker is capable of gaining some information about the plaintext without knowing the key. For demonstrative purposes we assume that some attacker can perfectly distinguish, without knowing k, between $\rho_0 = \mathrm{Enc}_k(|0\rangle\langle 0|)$ and $\rho_1 = \mathrm{Enc}_k(|1\rangle\langle 1|)$, thus the attack has the projector onto the support of ρ_0, P, which is such that $\mathrm{Tr}\left[P\rho_0\right] = 1$ and $\mathrm{Tr}\left[P\rho_1\right] = 0$. We observe that P is Hermitian and $P^2 = P$, thus $(2P - \mathbb{I})(2P - \mathbb{I})^\dagger = (2P - \mathbb{I})^2 = \mathbb{I}$, which means $2P - \mathbb{I}$ is a unitary matrix. Furthermore we have, by correctness of Π, that $\mathrm{Dec}_k \circ$

$P(\cdot)P \circ \mathrm{Enc}_k$ encrypts a state, projects it onto the part that is an encryption of $|0\rangle\langle 0|$ and then decrypts that part to $|0\rangle\langle 0|$, thus acts as $|0\rangle\langle 0|(\cdot)|0\rangle\langle 0|$. By linearity it follows that $\mathrm{Dec}_k \circ (2P - \mathbb{I})(\cdot)(2P - \mathbb{I}) \circ \mathrm{Enc}_k = (2|0\rangle\langle 0| - \mathbb{I})(\cdot)(2|0\rangle\langle 0| - \mathbb{I}) = Z(\cdot)Z$, where Z is the Pauli Z-gate, which satisfies $Z|0\rangle = |0\rangle$ and $Z|1\rangle = -|1\rangle$. Since the attacker could apply this attack without knowledge of k, the effective map of this attack is also $Z(\cdot)Z$, which maps $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ to $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ and $|-\rangle$ to $|+\rangle$. Since this map is self-inverse and not the identity map, it cannot be written as a combination of id and $\langle \mathrm{Dec}_K(\tau)\rangle$, thus the scheme is malleable. This example shows that plaintext and ciphertext non-malleable schemes both disallow the attacker to gain information about the plaintext from the ciphertext, meaning that both notions imply information-theoretic security. In [AM17] it is proven that NM implies information-theoretic security, but for the definition given in the next section this is still an open question.

The difference between plaintext and ciphertext non-malleability lies in what part of the effect of an attack acts as identity. In the ciphertext non-malleability setting, this is fully determined by how much the attack behaves like identity on the ciphertext, but for plaintext non-malleability, this is a larger part of the effective map, because an attacker is allowed to map a ciphertext to another ciphertext of the same plaintext.

4.2 An Effective-Map-based Definition

To properly define plaintext non-malleability, we take a look at the NM definition. Observe that the constraints placed on Λ_1 and Λ_2 in Definition 3.6 enforce the effect described at the end of the previous section. For any state ρ^B, the trace of the output of $\Lambda_1(\rho^B)$ is determined by what part of the C register of $\phi^{+CC'}$ is not altered by the attack, or in other words, how much the attack acts as identity. Since this restriction is enforced on the ciphertext level, the first step in defining plaintext non-malleability is to remove the constraints placed on Λ_1 and Λ_2. After this removal we observe some extra freedom in the choice of the ideal effective map.

Lemma 4.1. *Let* $\Pi = (\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ *be an arbitrary* SKQES *and* $\Lambda_A^{CB\to C\hat{B}}$ *an arbitrary attack on* Π *with effective map* $\tilde{\Lambda}_A^{MB\to M\hat{B}}$*. If there exist* CPTNI Λ_1, Λ_2*, such that* $\Lambda_1 + \Lambda_2$ *is* CPTP*and it holds that*

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

then for any α *such that* $|M|^2 \leq \alpha \leq |C|^2$ *there exist* CPTNI Λ_3, Λ_4 *such that* $\Lambda_3 + \Lambda_4$ *is* CPTP *and*

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

Proof. Assume that for some CPTNI Λ_1, Λ_2 such that $\Lambda_1 + \Lambda_2$ is CPTP it holds that

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

Define

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

Note that $0 < \frac{(\alpha-1)|C|^2}{\alpha(|C|^2-1)} \leq 1$ as long as $1 < \alpha \leq |C|^2$ and thus Λ_3 and Λ_4 are CPTNI. Furthermore $\Lambda_3 + \Lambda_4 = \Lambda_1 + \Lambda_2$, thus $\Lambda_3 + \Lambda_4$ is CPTP. Observe that

$$\begin{aligned}
&\mathrm{id}^M \otimes \Lambda_3^{B\to\hat{B}} + \frac{1}{\alpha-1}\left(\alpha\langle \mathrm{Dec}_K(\tau^C)\rangle - \mathrm{id}\right)^M \otimes \Lambda_4^{B\to\hat{B}} \\
&\quad = \mathrm{id}^M \otimes \left(\Lambda_1 + \left(1 - \frac{(\alpha-1)|C|^2}{\alpha(|C|^2-1)}\right)\Lambda_2\right) + \frac{1}{\alpha-1}(\alpha\langle \mathrm{Dec}_K(\tau^C)\rangle - \mathrm{id})^M \otimes \left(\frac{(\alpha-1)|C|^2}{\alpha(|C|^2-1)}\Lambda_2\right) \\
&\quad = \mathrm{id}^M \otimes \Lambda_1 + \frac{|C|^2-\alpha}{\alpha(|C|^2-1)}\,\mathrm{id}^M \otimes \Lambda_2 + \frac{|C|^2}{\alpha(|C|^2-1)}(\alpha\langle \mathrm{Dec}_K(\tau^C)\rangle - \mathrm{id})^M \otimes \Lambda_2 \\
&\quad = \mathrm{id}^M \otimes \Lambda_1 + \frac{|C|^2-\alpha}{\alpha(|C|^2-1)}\,\mathrm{id}^M \otimes \Lambda_2 + \frac{|C|^2}{|C|^2-1}\langle \mathrm{Dec}_K(\tau)\rangle \otimes \Lambda_2 - \frac{|C|^2}{\alpha(|C|^2-1)}\,\mathrm{id}^M \otimes \Lambda_2 \\
&\quad = \mathrm{id}^M \otimes \Lambda_1 + \frac{|C|^2}{|C|^2-1}\langle \mathrm{Dec}_K(\tau)\rangle \otimes \Lambda_2 - \frac{1}{|C|^2-1}\,\mathrm{id}^M \otimes \Lambda_2 \\
&\quad = \mathrm{id}^M \otimes \Lambda_1 + \frac{1}{|C|^2-1}(|C|^2\langle \mathrm{Dec}_K(\tau^C)\rangle - \mathrm{id})^M \otimes \Lambda_2.
\end{aligned}$$

From this it follows that

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

□

We choose $\alpha = |M|^2$ in the lemma above to obtain a definition of plaintext non-malleability that has similar properties to NM.

Definition 4.1. *A* SKQES $\Pi = (\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ *is* ***ε-plaintext non-malleable (ε-PNM)*** *if, for any attack $\Lambda_A^{CB\to C\hat{B}}$, its effective map $\tilde{\Lambda}_A^{MB\to M\hat{B}}$ is such that*

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

where Λ_1 and Λ_2 are CPTNI *and $\Lambda_1 + \Lambda_2$ is* CPTP.

While this definition does not explicitly restrict the choice of Λ_1 and Λ_2, a restriction similar to the one in Definition 3.6 can be made while still obtaining a definition equivalent to PNM.

Theorem 4.1. *Let* $\Pi = (\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ *be an arbitrary* ε-PNM SKQES *for some ε, then for any attack $\Lambda_A^{CB\to C\hat{B}}$, its effective map $\tilde{\Lambda}_A^{MB\to M\hat{B}}$ is such that*

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

where

$$\Lambda_1 = \mathrm{Tr}_{MM'}\left[\phi^{+MM'}\tilde{\Lambda}_A(\phi^{+MM'}\otimes(\cdot))\right] \qquad \textit{and}$$
$$\Lambda_2 = \mathrm{Tr}_{MM'}\left[(\mathbb{I}^{MM'}-\phi^{+MM'})\tilde{\Lambda}_A(\phi^{+MM'}\otimes(\cdot))\right].$$

In the case of $\varepsilon = 0$, the proof of this statement follows directly from substituting $\tilde{\Lambda}_A$ in the definition of Λ_1 and Λ_2. The proof of the general case can be found in Appendix B. It follows directly from Lemma 4.1 (with $\alpha = |M|^2$) that any NM PKQES is also PNM. Intuitively one could say this makes sense, since if one cannot meaningfully transform any ciphertext then one can also not meaningfully transform any plaintext. This implication does not hold in the other direction. In Theorem 4.2 we show that 0-PNM does not imply 1-NM, which one can think of as PNM does not imply NM. Note that ε-NM is meaningfully defined for $\varepsilon \leq 2$ since the diamond norm of the difference between two quantum channels lies between 0 and 2[1]. Nevertheless not being 1-NM already indicates that a scheme does not capture the notion of non-malleability anymore. For example, if one has a scheme where an attacker can implement any transformation for negligibly less than half the keys, but no meaningful transformation for the other keys, then this scheme is 1-NM.

We prove Theorem 4.2 by taking a 0-NM scheme and appending a $|0\rangle$ to the ciphertext, which is traced out during decryption. The result is that an attacker can freely modify this appended qubit, which allows for meaningful transformations on the ciphertext level, for example the NOT transformation, which maps this appended qubit to a $|1\rangle$. On the plaintext level, however, no meaningful transformation is possible since the original scheme was 0-NM. Note that a similar construction would also work if we modify a 0-PNM scheme in the same way.

Theorem 4.2. *There exists a* PKQES $\Pi = (\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ *that is* 0-PNM *but not* 1-NM.

Proof. Let $\Pi' = (\mathrm{KeyGen}', \mathrm{Enc}', \mathrm{Dec}')$ be an arbitrary PKQES that is 0-NM[2]. Then define $\Pi = (\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ as

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

where R is an auxiliary 1-qubit register. Let Λ be an arbitrary attack on Π with effective map $\tilde{\Lambda}$, then define $\Lambda' = \mathrm{Tr}_R\left[\Lambda((\cdot)\otimes|0\rangle\langle 0|^R)\right]$, which is an attack on Π' with effective map $\tilde{\Lambda}'$. Observe that

$$\begin{aligned}\tilde{\Lambda}' &= \mathop{\mathbb{E}}_{k\leftarrow\mathrm{KeyGen}(1^n)}[\mathrm{Dec}_k \circ \Lambda' \circ \mathrm{Enc}_k]\\ &= \mathop{\mathbb{E}}_{k\leftarrow\mathrm{KeyGen}(1^n)}[\mathrm{Dec}_k \circ \mathrm{Tr}_R \circ \Lambda \circ ((\cdot)\otimes|0\rangle\langle 0|^R) \circ \mathrm{Enc}_k]\\ &= \mathop{\mathbb{E}}_{k\leftarrow\mathrm{KeyGen}(1^n)}[\mathrm{Dec}'_k \circ \Lambda \circ \mathrm{Enc}'_k]\\ &= \tilde{\Lambda}.\end{aligned}$$

[1] One might be inclined to scale this to $[0,1]$, as is done in some literature, but we choose not to for consistency with [AM17].

[2] See [AM17] for such a scheme

Because Π' is 0-NM, we have that

$$\tilde{\Lambda}' = \mathrm{id}^M \otimes \Lambda_3^{B\to\hat{B}} + \frac{1}{|C|^2-1}\left(|C|^2\langle \mathrm{Dec}_K(\tau^C)\rangle - \mathrm{id}\right)^M \otimes \Lambda_4^{B\to\hat{B}}.$$

It follows from Lemma 4.1 that Π is 0-PNM.

Now consider the attack $\Lambda_X = \mathrm{id}\otimes(X(\cdot)X)^R \otimes \mathrm{Tr}[\cdot]^B$, 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

$$\begin{aligned}\Lambda_X(\phi^{+CC'}\otimes(\cdot)^B) &= (\mathrm{id}\otimes(X(\cdot)X)^R\otimes\mathrm{Tr}[\cdot]^B)\left(\sum_{i,j\in\{0,1\}^{\log|C|}}|ii\rangle\langle jj|^{CC'}\otimes(\cdot)^B\right)\\ &= \mathrm{Tr}\,[\cdot]^B\sum_{i,j\in\{0,1\}^{\log|C|}}|f(i)i\rangle\langle f(j)j|.\end{aligned}$$

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 $\phi^{+CC'}\Lambda_X(\phi^{+CC'}\otimes(\cdot)^B) = \mathbf{0}^{BCC'\to CC'}$. Also note that the effective map of Λ_X is $\tilde{\Lambda}_X = \mathrm{id}^M\otimes\mathrm{Tr}[\cdot]^B$, since the attack only acts on R and B and thus does not modify the message in M. Let Λ_1 and Λ_2 be as in Definition 3.6, then $\mathrm{Tr}\,[\Lambda_1(\rho)] = 0$ for all ρ. It follows that $\Lambda_2 = \mathrm{Tr}_{CC'}\left[(\mathbb{I}-\phi^{+CC'})\Lambda_X(\phi^{+CC'}\otimes(\cdot)^B)\right] = \mathrm{Tr}[\cdot]^B$. Furthermore we have

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

where we use that $\mathrm{Tr}\left[\phi^{+MM'}(\mathrm{Dec}_K(\tau)\otimes\tau^{M'})\right] = \frac{1}{|M|^2}$, as is proven in the proof of Theorem B.1, and $|M|\geq 2$, which is true since we assume that we are encrypting at least one qubit. □

4.3 DNS Authentication from PNM

Now that we have observed that PNM and NM do not syntactically capture the same notion, we use DNS authentication to show that PNM still captures the essence of plaintext

non-malleability. The idea of DNS authentication is that, after a possible attack, one can determine from a received plaintext whether or not an attack was performed. For this reason, DNS authentication is also called 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 4.2. *For any* SKQES $(\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ *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* $\mathrm{Tr}[\langle x|^R \mathrm{Dec}_K(\tau^C)|x\rangle^R] \leq \frac{1}{|R|}$.

Proof. Observe that

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

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

Lemma 4.2 allows us to find tags that have little overlap with $\mathrm{Dec}_K(\tau^C)$, which means one can distinguish well between the case were the tag was left unharmed, and thus still is equal to $|x\rangle\langle x|$, and the case where the ciphertext was depolarized and thus also the tag. We use this property to build a scheme that is DNS authenticating.

Theorem 4.3. *For any* $0 \leq \varepsilon \leq 2$ *and any* ε-PNM SKQES $\Pi = (\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$*, there exists some* x *such that the scheme* $\Pi' = (\mathrm{KeyGen}', \mathrm{Enc}', \mathrm{Dec}')$ *is* $\left(\frac{3}{|R|} + \varepsilon\right)$-DNS *authenticating, where*

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

Proof. By Lemma 4.2, there exists an $x \in \{0,1\}^{\log|R|}$ such that $\mathrm{Tr}\left[\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle\right] \leq \frac{1}{|R|}$. Fix this x and define Π' as above. Define $\mathrm{Enc}_{append}(X) = X \otimes |x\rangle\langle x|$ and $\mathrm{Dec}_{check}(Y) = \langle x|Y|x\rangle + \mathrm{Tr}\left[(\mathbb{I} - |x\rangle\langle x|)Y\right]|\perp\rangle\langle\perp|$ and observe that $\mathrm{Enc}' = \mathrm{Enc} \circ \mathrm{Enc}_{append}$ and $\mathrm{Dec}' = \mathrm{Dec}_{check} \circ \mathrm{Dec}$. Let Λ_A be an arbitrary attack map on Π', then its effective map is

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

Since Enc_{append} and Dec_{check} do not change with k and are linear, we have

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

where $\tilde{\Lambda}_A = \mathop{\mathbb{E}}_{k\leftarrow \mathrm{KeyGen}(1^n)}[(\mathrm{Dec}_k \circ \Lambda_A \circ \mathrm{Enc}_k)]$. Since Π is ε-PNM, there exist Λ_1, Λ_2 such that

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

Since Enc_{append} and Dec_{check} are both CPTP, by submultiplicativity we have that

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

which is equivalent to

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

Observe that

$$\mathrm{Dec}_{check} \circ \mathrm{id} \circ \mathrm{Enc}_{append} = \langle x|((\cdot) \otimes |x\rangle\langle x|)|x\rangle + \mathrm{Tr}[(\mathbb{I} - |x\rangle\langle x|)((\cdot) \otimes |x\rangle\langle x|)]|\perp\rangle\langle\perp| = \mathrm{id}\,.$$

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

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

then we have

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

Here the second equality uses the fact that Enc_{append} is trace preserving and $\langle \mathrm{Dec}_K(\tau)\rangle$ is a constant channel, which only uses the trace of the input. Since every term now ends in $\otimes\Lambda_2$, we can remove this term and multiply with $\|\Lambda_2\|_\diamond$, which is less than 1 since Λ_2 is CPTNI. We continue by expanding Dec_{check}, where we note that $\langle x|\,\mathrm{Tr}\left[(\cdot)^{M'}\right] \mathrm{Dec}_K(\tau^C)|x\rangle = \langle\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle\rangle$.

$$\begin{aligned}
&\left\| \frac{1}{|M|^2 - 1}(|M|^2(\mathrm{Dec}_{check} \circ \mathrm{Tr}\left[(\cdot)^{M'}\right] \mathrm{Dec}_K(\tau^C)) - \mathrm{id}) \otimes \Lambda_2 - \langle|\perp\rangle\langle\perp|\rangle \otimes \Lambda_2 \right\|_\diamond \\
\leq&\left\| \frac{1}{|M|^2 - 1}(|M|^2(\mathrm{Dec}_{check} \circ \mathrm{Tr}\left[(\cdot)^{M'}\right] \mathrm{Dec}_K(\tau^C)) - \mathrm{id}) - \langle|\perp\rangle\langle\perp|\rangle \right\|_\diamond \\
=&\left\| \frac{1}{|M|^2 - 1}(|M|^2 \left(\langle\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle\rangle + \mathrm{Tr}[(\mathbb{I} - |x\rangle\langle x|)\mathrm{Dec}_K(\tau^C)]\langle|\perp\rangle\langle\perp|\rangle\right) - \mathrm{id}) - \langle|\perp\rangle\langle\perp|\rangle \right\|_\diamond .
\end{aligned}$$

We can rewrite this expression by first collecting all multipliers of $\langle|\perp\rangle\langle\perp|\rangle$, then distributing the $|M|^2$ term and lastly rewriting $\mathrm{Tr}[(\mathbb{I} - |x\rangle\langle x|)\mathrm{Dec}_K(\tau^C)]$ as $1 - \mathrm{Tr}[\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle]$

and simplifying the resulting term.

$$
\begin{aligned}
&\left\| \frac{1}{|M|^2-1}(|M|^2\left(\langle\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle\rangle + \mathrm{Tr}[(\mathbb{I}-|x\rangle\langle x|)\mathrm{Dec}_K(\tau^C)]\langle|\perp\rangle\langle\perp|\rangle\right) - \mathrm{id}) - \langle|\perp\rangle\langle\perp|\rangle \right\|_\diamond \\
=&\left\| \frac{1}{|M|^2-1}(|M|^2\left(\langle\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle\rangle + \left(\mathrm{Tr}[(\mathbb{I}-|x\rangle\langle x|)\mathrm{Dec}_K(\tau^C)] - \frac{|M|^2-1}{|M|^2}\right)\langle|\perp\rangle\langle\perp|\rangle\right) - \mathrm{id}) \right\|_\diamond \\
=&\left\| \frac{1}{|M|^2-1}(|M|^2\langle\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle\rangle + \left(|M|^2\,\mathrm{Tr}[(\mathbb{I}-|x\rangle\langle x|)\mathrm{Dec}_K(\tau^C)] - (|M|^2-1)\right)\langle|\perp\rangle\langle\perp|\rangle - \mathrm{id}) \right\|_\diamond \\
=&\left\| \frac{1}{|M|^2-1}(|M|^2\langle\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle\rangle + \left(|M|^2(1-\mathrm{Tr}[\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle]) - (|M|^2-1)\right)\langle|\perp\rangle\langle\perp|\rangle - \mathrm{id}) \right\|_\diamond \\
=&\left\| \frac{1}{|M|^2-1}(|M|^2\langle\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle\rangle + \left(1-|M|^2\,\mathrm{Tr}[\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle]\right)\right)\langle|\perp\rangle\langle\perp|\rangle - \mathrm{id}) \right\|_\diamond \\
\leq&\frac{1}{|M|^2-1}\left(|M|^2\left\|\langle\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle\rangle\right\|_\diamond + \left\|(1-|M|^2\,\mathrm{Tr}[\langle x|\mathrm{Dec}_K(\tau^C)|x\rangle])\langle|\perp\rangle\langle\perp|\rangle\right\|_\diamond + \|\mathrm{id}\|_\diamond\right) \\
\leq&\frac{1}{|M|^2-1}\left(\frac{|M|^2}{|R|} + \left(\frac{|M|^2}{|R|}-1\right)+1\right) \leq \frac{3}{|R|}.
\end{aligned}
$$

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

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

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

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

4.4 PNM in the Public-Key Setting

In the public-key setting, no definition has yet been set for plaintext non-malleability. In this section, we provide a number of details that one has to consider when defining plaintext non-malleability and give a possible effective-map-based definition.

In the public-key setting, it is usually assumed that an attacker has access to the public key, and thus access to the encryption map. Thus, no PNM public-key scheme can exist since on any scheme there is an attack $\Lambda_A = \langle\mathrm{Enc}_{pk}(\rho)\rangle$ with effective map $\langle\rho\rangle$, where ρ is chosen arbitrarily. Furthermore, we have to consider that in the public-key setting an attacker can always decode the ciphertext with the public key if enough time is given to the adversary, thus we only consider efficient attacks. Combining these properties we give the following suggestion for public-key PNM.

Definition 4.2. *A* PKQES $\Pi = (\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ *is* ε***-plaintext non-malleable (***ε***-PNM)** *if, for any attack A that runs in time $p(n)$ for some polynomial p and implements $\Lambda_A^{CB\to C\hat{B}}$, its effective map $\tilde{\Lambda}_A^{MB\to M\hat{B}}$ is such that*

$$\left\|\tilde{\Lambda}_A - \left(\mathrm{id}^M\otimes\Lambda_1^{B\to\hat{B}} + \frac{1}{|M|^2-1}\left(|M|^2(\Lambda_2^{B\to M\hat{B}}\circ\mathrm{Tr}_M) - \mathrm{id}^M\otimes(\mathrm{Tr}_M\circ\Lambda_2^{B\to M\hat{B}})\right)\right)\right\|_\diamond \leq \varepsilon,$$

where Λ_1 and Λ_2 are CPTNI *and $\Lambda_1 + (\mathrm{Tr}_M\circ\Lambda_2)$ is* CPTP.

The main difference in this definition is that the attacker is allowed to implement any map into M, but since Λ_2 is not given M as input, it is still not possible to produce a meaningfully related plaintext.

CHAPTER 5

Quantum Comparison-based Non-Malleability

In this chapter, we consider a different approach for defining non-malleability. In Section 3.3 we discussed CNM, a classical definition for non-malleability. In Section 5.1 we will define QCNM, a notion of non-malleability in the quantum setting, in a way that is similar to CNM. In Section 5.2 we will argue that QCNM is a possible quantum translation of CNM. Lastly, we will discuss how QCNM relates to PNM and NM in Section 5.3.

5.1 A CNM-based definition

In this section, we will define QCNM as a quantum analogue of CNM. We first analyze CNM and decide how to implement each of its components in the quantum setting. We consider an adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ to be a pair of quantum algorithms, where $\mathcal{A}_1$ is responsible for producing the quantum alternative to the message distribution M and $\mathcal{A}_2$ will produce the relation R and the vector of candidate ciphertexts $\mathbf{y}$.

5.1.1 The message distribution M

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 must consider the following requirements:

1. In order to check the relation in the last step of CNM, we require that two copies of the same message are produced, one of which will be kept by the challenger and the other encoded and used by $\mathcal{A}_2$.

2. In order to prevent the adversary from cheating, it must not be possible for the adversary to entangle herself with the produced message.

We first considered a mixed state as a representation of M, but this way of representing M makes it impossible to enforce requirement (2). As such we have chosen to represent M by a unitary $U^{MM'P}$ such that $U|0\rangle$ is a purification of the message distribution, where the message resides in M and its copy in M' and P is used for the purification. The first part of the quantum adversary, $\mathcal{A}_1$, produces this unitary in the form of a circuit, which we simply denote by $(U, S) \leftarrow \mathcal{A}_1(pk)$. In order to ensure that M' indeed contains a copy of M we require that MM' resides in the **symmetric subspace** of M, which means that if $\mathrm{Tr}_P\left[U|0\rangle\langle 0|U^\dagger\right] = \rho$, then we have that $\rho = W^{MM'}\rho$, where W is the swap operator, which performs the operation $W|ij\rangle = |ji\rangle$. Note that this restriction does not disallow U to produce a state where M and M' are entangled, for example, the state $\phi^{+MM'}$ is part of the symmetric subspace. For more information on the symmetric subspace, we refer to Chapter 7 of [Wat18].

5.1.2 The QCNM experiments

For the QCNM definition we define two experiments, similar to the CNM definition. Note that here the relation R is modeled as a matrix E, which is assumed to be a POVM element. The vector $\mathbf{y}$ is modeled by a vector 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 vector $\mathbf{x}$ is modelled similarly 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 Corollary 3.2. For any PKQES $\Pi = (\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ with security parameter n, let $\{V_k \mid k = (pk, sk) \leftarrow \mathrm{KeyGen}(1^n)\}$, $t = \log|T|$, $\{\psi_{k,r} \mid k = (pk, sk) \leftarrow \mathrm{KeyGen}(1^n), r \in \{0,1\}^t\}$ and $\{p_k \mid k = (pk, sk) \leftarrow \mathrm{KeyGen}(1^n)\}$ be as in Corollary 3.2, then the QCNM experiments are defined as follows.

Experiment 5.1.
The QCNM-Real$(\Pi, \mathcal{A}, n)$ experiment:

1: $k = (pk, sk) \leftarrow \mathrm{KeyGen}(1^n)$
2: $(U^{MM'P}, S) \leftarrow \mathcal{A}_1(pk)$
3: $r \xleftarrow{p_k} \{0,1\}^t$
4: **construct** U_ψ^T such that $U_\psi^T|0\rangle^T = |\psi_{k,r}\rangle^T$
5: **construct** $U_{prep}^{MTM'P} = V_k^{MT}(U^{MM'P} \otimes U_\psi^T)$
6: **prepare** $U_{prep}|0\rangle\langle 0|U_{prep}^\dagger$ in $MTM'P$
7: $(\mathbf{C}, E) \leftarrow \mathcal{A}_2(MT, S)$
8: **for each** i such that $1 \le i \le m$:
    1: **perform** $U_{prep}^\dagger$ on $C_i M'P$
    2: **measure** $\{|0\rangle\langle 0|, \mathbb{I} - |0\rangle\langle 0|\}$ on $C_i M'P$, **if** the outcome is 0: **output** 0
    3: **perform** U_{prep} on $C_i M'P$
9: $\mathbf{M} \leftarrow \mathrm{Dec}_{sk}(\mathbf{C})$
10: **measure** $\{E, \mathbb{I} - E\}$ on $M'\mathbf{M}$, **output** 1 iff the outcome is E.

Experiment 5.2.
The QCNM-Ideal$(\Pi, \mathcal{A}, n)$ experiment:

1: $k = (pk, sk) \leftarrow \text{KeyGen}(1^n)$
2: $(U^{MM'P}, S) \leftarrow \mathcal{A}_1(pk)$
3: $r \xleftarrow{p_k} \{0,1\}^t$
4: **construct** U_ψ^T such that $U_\psi^T|0\rangle^T = |\psi_{k,r}\rangle^T$
5: **construct** $U_{prep}^{MTM'P} = V_k^{MT}(U^{MM'P} \otimes U_\psi^T)$
6: **prepare** $U_{prep}|0\rangle\langle 0|U_{prep}^\dagger$ in $MTM'P$
7: $(\mathbf{C}, E) \leftarrow \mathcal{A}_2(MT, S)$
8: **for each** i such that $1 \leq i \leq m$:
    1: **perform** $U_{prep}^\dagger$ on $C_iM'P$
    2: **measure** $\{|0\rangle\langle 0|, \mathbb{I} - |0\rangle\langle 0|\}$ on $C_iM'P$, **if** the outcome is 0: **output** 0
    3: **perform** U_{prep} on $C_iM'P$
9: $\mathbf{M} \leftarrow \text{Dec}_{sk}(\mathbf{C})$
10: **prepare** $U|0\rangle\langle 0|U^\dagger$ in $\tilde{M}\tilde{M'}\tilde{P}$
11: **measure** $\{E, \mathbb{I} - E\}$ on $\tilde{M'}\mathbf{M}$, **output** 1 iff the outcome is E.

Note that here the preparation of the message state and the encoding is both done by U_{prep}, which means the $y \notin \mathbf{y}$ check in the CNM experiments can be implemented by undoing U_{prep} on all C_i and then measuring whether the result 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.

Definition 5.1. *A* PKQES Π *is* ***quantum comparison-based non-malleable (*QCNM*)*** *if for any quantum adversary* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ *it holds that*

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

if $\mathcal{A}$ *is such that there exists a polynomial* p *such that for all* n*:*

- $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *run in time* $p(n)$
- $\mathcal{A}_1$ *outputs a valid unitary* U *which can be implemented in time* $p(n)$ *and* $\rho = \text{Tr}_P\left[U|0\rangle\langle 0|U^\dagger\right] = \rho$ *is such that* $\rho = W^{MM'}\rho$.
- $\mathcal{A}_2$ *outputs a POVM element* E *which can be implemented in time* $p(n)$
- $\mathcal{A}_2$ *outputs a vector of registers* $\mathbf{C}$ *such that* $\perp \notin \text{Dec}_{sk}(\mathbf{C})$

5.2 Relation between QCNM and CNM

In this section, we compare QCNM to CNM, by considering both in a post-quantum setting. 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. For this reason, we also restrict QCNM by only

allowing $\mathcal{A}_1$ to output a U such that $U|0\rangle$, when measured in the computational basis, always yields $|xx\rangle$ according to some probability distribution p_X.

Experiment 5.3.
The QCNM-Real$_{PQ}(\Pi, \mathcal{A}, n)$ experiment:

1: $k = (pk, sk) \leftarrow \text{KeyGen}(1^n)$
2: $(U^{MM'P}, S) \leftarrow \mathcal{A}_1(pk)$
3: $r \xleftarrow{p_k} \{0,1\}^t$
4: **construct** U_ψ^T such that $U_\psi^T|0\rangle^T = |\psi_{k,r}\rangle^T$
5: **prepare** $U|0\rangle$ in $MM'P$
6: **measure** $MM'P$ in the computational basis, store the result of M in x
7: **construct** $U_{prep}^{MT} = V_k^{MT}(\mathbb{I}^M \otimes U_\psi^T)$
8: **prepare** $|0\rangle\langle 0|$ in T
9: **perform** U_{prep} on MT
10: $(\mathbf{C}, E) \leftarrow \mathcal{A}_2(MT, S)$
11: **for each** i such that $1 \leq i \leq m$:
    1: **perform** $U_{prep}^\dagger$ on C_i
    2: **measure** $\{M = |x\rangle\langle x|^M \otimes |0\rangle\langle 0|^T, \mathbb{I} - M\}$ on C_i, **if** the outcome is M: **output** 0
    3: **perform** U_{prep} on C_i
12: $\mathbf{M} \leftarrow \text{Dec}_{sk}(\mathbf{C})$
13: **measure** $\{E, \mathbb{I} - E\}$ on $M'\mathbf{M}$, **output** 1 iff the outcome is E.

Experiment 5.4.
The QCNM-Ideal$_{PQ}(\Pi, \mathcal{A}, n)$ experiment:

1: $k = (pk, sk) \leftarrow \text{KeyGen}(1^n)$
2: $(U^{MM'P}, S) \leftarrow \mathcal{A}_1(pk)$
3: $r \xleftarrow{p_k} \{0,1\}^t$
4: **construct** U_ψ^T such that $U_\psi^T|0\rangle^T = |\psi_{k,r}\rangle^T$
5: **prepare** $U|0\rangle$ in $MM'P$
6: **measure** $MM'P$ in the computational basis, store the result of M in x
7: **construct** $U_{prep}^{MT} = V_k^{MT}(\mathbb{I}^M \otimes U_\psi^T)$
8: **prepare** $|0\rangle\langle 0|$ in T
9: **perform** U_{prep} on MT
10: $(\mathbf{C}, E) \leftarrow \mathcal{A}_2(MT, S)$
11: **for each** i such that $1 \leq i \leq m$:
    1: **perform** $U_{prep}^\dagger$ on C_i
    2: **measure** $\{M = |x\rangle\langle x|^M \otimes |0\rangle\langle 0|^T, \mathbb{I} - M\}$ on C_i, **if** the outcome is M: **output** 0
    3: **perform** U_{prep} on C_i
12: $\mathbf{M} \leftarrow \text{Dec}_{sk}(\mathbf{C})$
13: **prepare** $U|0\rangle$ in $\tilde{M}\tilde{M}'\tilde{P}$
14: **measure** $\tilde{M}\tilde{M}'\tilde{P}$ in the computational basis
15: **measure** $\{E, \mathbb{I} - E\}$ on $\tilde{M}'\mathbf{M}$, **output** 1 iff the outcome is E.

We consider the above experiments to be the post-quantum version of the QCNM

experiments. The main modification is the measurement in Step 6, which is essential to ensure that the register passed to $\mathcal{A}_2$ contains a classical state. The rest of the modifications are required for the algorithm to still work. For example, the modification in Step 11.2 ensures that the outcome of the measurement still represents the $y \notin \mathbf{y}$ statement, but now by undoing only the encryption since the measurement in Step 6 is not reversible.

Definition 5.2. *A* PKQES Π *is* ***post-quantum comparison-based non-malleable (QCNM$_{PQ}$)*** *if for any quantum adversary* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ *it holds that*

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

if $\mathcal{A}$ *and* Π *are such that there exists a polynomial* p *such that for all* n*:*

- Enc *and* Dec *take only classical input and produce only classical output*
- $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *run in time* $p(n)$
- $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *output only classical states*
- $\mathcal{A}_1$ *outputs a valid unitary* U *which can be implemented in time* $p(n)$ *and* $\rho = \text{Tr}_P\left[U|0\rangle\langle 0|U^\dagger\right] = \rho$ *is such that* ρ *yields* $|xx\rangle$ *for some* x *when measured in the computational basis.*
- $\mathcal{A}_2$ *outputs a POVM element* E *which can be implemented in time* $p(n)$
- $\mathcal{A}_2$ *outputs a vector of registers* $\mathbf{C}$ *such that* $\perp \notin \text{Dec}_{sk}(\mathbf{C})$

Here the extra constraint placed on U is required but can be considered equivalent to the statement that $\text{Tr}_P\left[U|0\rangle\langle 0|U^\dagger\right]$ should yield a state in the symmetric subspace when measured in the computational basis. Thus the new requirement is practically the same but takes into consideration the measurement in Step 6. Similarly, we define a post-quantum version of CNM.

Definition 5.3. *A* PKQES Π *is* ***comparison-based non-malleable for post-quantum adversaries (CNM$_{PQ}$)*** *if for any quantum adversary* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ *it holds that*

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

if Π *and* $\mathcal{A}$ *are such that there exists a polynomial* p *such that for all* n*:*

- Enc *and* Dec *take only classical input and produce only classical output*
- $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *run in time* $p(n)$
- $\mathcal{A}_1$ *outputs a valid quantum algorithm* M *which runs in time* $p(n)$ *and produces classical strings*
- $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *output only classical states*
- $\mathcal{A}_2$ *outputs a quantum algorithm* R *computable in time* $p(n)$
- $\mathcal{A}_2$ *outputs a vector* $\mathbf{y}$ *such that* $\perp \notin \text{Dec}_{sk}(\mathbf{y})$

The only difference between CNM and CNM_{PQ} is that the latter assumes the encryption scheme, adversary and any algorithms produced by the adversary may require a quantum computer to compute. 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_{PQ} is simply a stronger requirement than CNM since it requires security against a strict superset of adversaries, and thus trivially implies CNM.

Theorem 5.1. *A* PKQES Π *is* QCNM_{PQ} *if and only if* Π *is* CNM_{PQ}*.*

Proof. $\Rightarrow$ Let Π be an arbitrary QCNM_{PQ} PKQES and let $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ be an arbitrary quantum adversary intended to perform the CNM_{PQ} experiments. Assume that Π 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 R is the set of possible input for M and ϕ_x is the uniform superposition over all $|r\rangle$ such that $x \leftarrow M(r)$.
3: **output** (U, S)

Note that Step 2 here is always possible. Given a classical deterministic algorithm $M(r)$, one can always define M' such that $(x, r) \leftarrow M'(r)$ when $x \leftarrow M(r)$ and since M' is reversible, one can implement it as a quantum gate that performs $U_M : |0\rangle|r\rangle \mapsto |M(r)\rangle|r\rangle$. We can use this U_M to construct U, which first prepares a uniform superposition over all r by applying Hadamard gates to the $\log|R|$ qubits in P, then performs U_M twice, first on MP and then on $M'P$.

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

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** $(E, |\mathbf{y}\rangle\langle\mathbf{y}|^{C_1\ldots C_m})$

Observe that the definition of $\text{QCNM-Real}_{PQ}(\Pi, \mathcal{B}, n)$, 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 $U|0\rangle^{MM'P} = \sum_{x\leftarrow M}\sqrt{p_M(x)}|xxx\rangle^{MM'P}$
4: $r \xleftarrow{p_k} \{0,1\}^t$
5: **construct** U_ψ^T such that $U_\psi^T|0\rangle^T = |\psi_{k,r}\rangle^T$
6: **prepare** $U|0\rangle$ in $MM'P$
7: **measure** $MM'P$ in the computational basis, store the result of M in x
8: **construct** $U_{prep}^{MT} = V_k^{MT}(\mathbb{I}^M \otimes U_\psi^T)$

9: **prepare** $|0\rangle\langle 0|$ in T
10: **perform** U_{prep} on MT
11: $(R, \mathbf{y}) \leftarrow \mathcal{A}_2(y, s)$
12: **construct** $E = \sum_{i,\mathbf{j}} R(i,\mathbf{j})|i\mathbf{j}\rangle\langle i\mathbf{j}|$
13: **prepare** $|\mathbf{y}\rangle\langle\mathbf{y}|$ in $\mathbf{C}$
14: **for each** i such that $1 \leq i \leq m$:
    1: **perform** $U^\dagger_{prep}$ on C_i
    2: **measure** $\{M = |x\rangle\langle x|^M \otimes |0\rangle\langle 0|^T, \mathbb{I} - M\}$ on C_i, **if** the outcome is M: **output** 0
    3: **perform** U_{prep} on C_i
15: $\mathbf{M} \leftarrow \mathrm{Dec}_{sk}(\mathbf{C})$
16: **measure** $\{E, \mathbb{I} - E\}$ on $M'\mathbf{M}$, **output** 1 iff the outcome is E.

Here Steps 3,5,6 and 7 together simply execute $x \leftarrow M$. Furthermore, if $y \in \mathbf{y}$ then some C_i contains $|y\rangle\langle y|$, which will guarantee the output to be 0 in Step 14. Conversely if $y \notin \mathbf{y}$, then all C_i contain some state orthogonal to $|y\rangle\langle y|$ and thus Step 4 has 0 probability of outputting 0 in this case, thus Step 14 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 16 outputs 1 iff $R(x,\mathbf{x})$, where x is stored in M' and $\mathbf{x}$ in $\mathbf{M}$. Combined with the fact that U_{prep} performs $\mathrm{Enc}_{k;r}$ we conclude that QCNM-Real$_{PQ}(\Pi, \mathcal{B}, n)$ produces the same random variable as CNM-Real$(\Pi, \mathcal{A}, n)$. 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 $\tilde{x} \leftarrow M$. Since the random variables produced are identical, so is their difference in probability of being 1, and thus if follows that Π is CNM$_{PQ}$.

$\Leftarrow$ Let Π be an arbitrary CNM$_{PQ}$ PKQES and let $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ be an arbitrary classical adversary on this scheme intended to perform the 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: **construct** M to be
    1: **prepare** $U|0\rangle$ in $M_a M'_a P'$
    2: **measure** M_a in the computational basis and **output** the result
3: **output** (M, s)

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

1: $(E, \mathbf{y}) \leftarrow \mathcal{A}_2(|s\rangle\langle s|^S, |y\rangle\langle y|^{MT})$
2: **construct** $R(i,\mathbf{j})$ to be:
    (1) **prepare** $|i\mathbf{j}\rangle\langle i\mathbf{j}|$ in $M'\mathbf{M}$
    (2) **measure** $\{E, \mathbb{I} - E\}$ on $M'\mathbf{M}$, **output** 1 iff the outcome is E
3: **output** $(R, |\mathbf{y}\rangle\langle\mathbf{y}|)$

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

1: $(pk, sk) \leftarrow \mathrm{KeyGen}(1^n)$
2: $(U, |s\rangle\langle s|^S) \leftarrow \mathcal{A}_1(pk)$

3: **prepare** $U|0\rangle$ in $M_a M'_a P'$
4: **measure** M_a in the computational basis and let x be the result
5: $y \leftarrow \mathrm{Enc}_{pk}(x)$
6: $(E, |\mathbf{y}\rangle\langle\mathbf{y}|) \leftarrow \mathcal{A}_2(|s\rangle\langle s|^S), |y\rangle\langle y|^{MT}$
7: $\mathbf{x} \leftarrow \mathrm{Dec}_{sk}(\mathbf{y})$
8: **return** 0 if $(y \in \mathbf{y})$
9: **prepare** $|x\mathbf{x}\rangle\langle x\mathbf{x}|$ in $M'\mathbf{M}$
10: **measure** $\{E, \mathbb{I} - E\}$ on $M'\mathbf{M}$, **output** 1 iff the outcome is E.

Observe that the resulting experiment is identical to QCNM-Real$_{PQ}(\Pi, \mathcal{A}, n)$, with the exception that Step 5, the encrypting, is not performed by U_{prep} but simply by Enc and that Step 8 simply checks $y \in \mathbf{y}$ instead of loop that we earlier argued to be equivalent. The Ideal case has the exact same differences, and thus by the same argument as before it is the case that Π is QCNM$_{PQ}$.

□

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

Corollary 5.2. *Any* QCNM$_{PQ}$ PKES *is* CNM.

5.3 Relation between QCNM, PNM, and NM

In this Section, we will briefly touch upon how QCNM relates to PNM and NM. Firstly we have already seen that CNM, and by extension QCNM, are notions designed to enforce ciphertext non-malleability. This means that it should not be the case that PNM implies QCNM in general, which it indeed does not. Consider the scheme used in the proof of Theorem 4.2, where a $|0\rangle\langle 0|$ is appended to the ciphertext of a NM scheme. While the original NM scheme might already not be QCNM, the appended scheme is definitely not since one can trivially implement the identity map on M by having $\mathcal{A}_1(pk)$ output U such that $U|0\rangle = |\phi^+\rangle^{MM'}$ and having $\mathcal{A}_2$ implement the Pauli X on the appended qubit and output $E = \phi^{+MM'}$. The resulting adversary will always output 1 in the Real case but only with probability $\frac{1}{|M|}$ in the Ideal case.

The rest of the possible relations are less obvious and remain open questions. If one considers a restriction of QCNM where $\mathbf{y}$ is not a vector but a single ciphertext, then one can consider steps 3-8 of QCNM as an effective map and it is almost clear then NM implies this restriction of QCNM for SKQES, although more work remains to be done in finding out the exact portion of this effective map that is id, as this part might be used to construct a measurement E. Furthermore, it is our belief that for SKQES QCNM will imply PNM, although no definitive argument for this has yet been found.

CHAPTER 6

Conclusion and Discussion

We have presented two new definitions for non-malleability in the quantum setting, one in the symmetric-key setting and one in the public-key setting. We have shown the first, PNM, to be a weaker version of the notion defined in [AM17], but argued that it nevertheless captures a similar level of security. We have shown that PNM is DNS authenticating and have separated PNM from NM. We also presented a possible definition for the public-key setting, based on this approach. Furthermore we presented QCNM, which is intended to capture public-key non-malleability in the quantum setting. We have shown that this notion, when restricted to a post-quantum setting, is equivalent to CNM.

6.1 Future work

At the end of Chapters 4 and 5 we briefly noted some open questions, which we briefly summarize here. In the context of PNM, one might wonder whether the presented idea for a public-key definition is correct, or whether a different one can be defined based on effective maps. In the context of QCNM, open questions include what the effect is of restricting QCNM to the symmetric-key case, where $\mathcal{A}_1$ is not given any input. Furthermore, it is also important to look at the connections between NM, PNM and QCNM. Lastly, we note that, in the quantum setting, non-malleability and indistinguishability are inherently connected, and as such one might wonder whether any of the presented definitions are equivalent to some form of indistinguishability.

6.2 Acknowledgements

I thank Christian Majenz, who has inspired many of the concepts presented in this thesis and spent many hours reviewing this thesis and explaining critical concepts. I also thank Christian Schaffner, who introduced me to the field of quantum cryptography, helped set up this project and provided many insights through critical analysis of the concepts in this thesis. I also thank the QuSoft team for providing an excellent scientific environment and including me in many discussions, presentations and other activities. Lastly I thank Yde Venema, Alexandru Baltag, Serge Fehr, and Maris Ozols for their time spent reviewing this thesis as part of the committee.

Bibliography

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

[AGM18] Gorjan Alagic, Tommaso Gagliardoni, and Christian Majenz. "Unforgeable quantum encryption". In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2018, pp. 489–519.

[AM17] Gorjan Alagic and Christian Majenz. *Quantum non-malleability and authentication*. Version 3. Oct. 13, 2017. arXiv: `1610.04214v3`.

[BB14] Charles H Bennett and Gilles Brassard. "Quantum cryptography: Public key distribution and coin tossing". In: *Theor. Comput. Sci.* 560.P1 (2014), pp. 7–11.

[BS99] Mihir Bellare and Amit Sahai. "Non-malleable encryption: Equivalence between two notions, and an indistinguishability-based characterization". In: *Annual International Cryptology Conference*. Springer. 1999, pp. 519–536.

[DNS12] Frédéric Dupuis, Jesper Buus Nielsen, and Louis Salvail. "Actively secure two-party evaluation of any quantum operation". In: *Advances in Cryptology–CRYPTO 2012*. Springer, 2012, pp. 794–811.

[KPT11] Akinori Kawachi, Christopher Portmann, and Keisuke Tanaka. "Characterization of the relations between information-theoretic non-malleability, secrecy, and authenticity". In: *International Conference on Information Theoretic Security*. Springer. 2011, pp. 6–24.

[NC02] Michael A Nielsen and Isaac Chuang. *Quantum computation and quantum information*. AAPT, 2002.

[OTU00] Tatsuaki Okamoto, Keisuke Tanaka, and Shigenori Uchiyama. "Quantum public-key cryptosystems". In: *Annual International Cryptology Conference*. Springer. 2000, pp. 147–165.

[Pap03] Christos H Papadimitriou. *Computational complexity*. John Wiley and Sons Ltd., 2003.

[Pod+18] Damian Poddebniak et al. "Efail: Breaking S/MIME and OpenPGP Email Encryption using Exfiltration Channels". In: *27th USENIX Security Symposium* (2018).

[Sho99] Peter W Shor. "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer". In: *SIAM review* 41.2 (1999), pp. 303–332.

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

APPENDIX A

List of Notation

$\mathbb{N}$	the set of natural numbers	5
$\mathbb{C}$	the set of complex numbers	5
$\mathcal{H}$	a finite dimensional Hilbert space	5
$\lvert A\rvert$	the dimension of $\mathcal{H}_A$	5
$\lVert\cdot\rVert$	the Euclidean norm	5
$\lVert\cdot\rVert_1$	the (induced) trace norm	10
$\lVert\cdot\rVert_\diamond$	the diamond norm	10
$\mathrm{Tr}\left[\cdot\right]$	the trace of a matrix	5
$\mathrm{Tr}_P\left[\cdot\right]$	the partial trace over $\mathcal{H}_P$	7
$D(\cdot,\cdot)$	the trace distance	10
$(\cdot)^\dagger$	the conjugate transpose of a matrix	5
$\mathcal{B}(\cdot)$	the set of square matrices that act on a Hilbert space	5
$\mathcal{D}(\cdot)$	the set of density matrices that act on a Hilbert space	5
$(\cdot)^{A\to B}$	an operator that maps from $\mathcal{H}_A$ to $\mathcal{H}_B$ or $\mathcal{B}(\mathcal{H}_A)$ to $\mathcal{B}(\mathcal{H}_B)$	5
$(\cdot)^A$	abbreviation of $(\cdot)^{A\to A}$	5
$p(n), q(n)$	polynomial functions	5
$(\cdot)\leq \mathrm{negl}(n)$	a negligible function	5
$\mathbb{I}$	the identity matrix	5
id	the identity channel	21
$\langle\rho\rangle$	the $\mathrm{Tr}\left[\cdot\right]\rho$ map	7
$\mathbf{0}$	the zero matrix	5
$\lvert\phi\rangle\langle\phi\rvert, \lvert\psi\rangle\langle\psi\rvert$	pure states	5
ϕ, ψ	abbreviation of $\lvert\phi\rangle\langle\phi\rvert$, $\lvert\psi\rangle\langle\psi\rvert$	5
ρ, σ	mixed states	5
τ	the maximally mixed state	5
ϕ^+	the maximally entangled state	5
$\lvert x\rangle$	a classical state	5
$\otimes$	the tensor product	5
$\oplus$	the bitwise xor of bitstrings or the direct sum of Hilbert spaces	5
U	a unitary matrix	7
X, Z	the Pauli X and Z gates	7, 23
$\leftarrow$	(possible) output or assignment	14
$\xleftarrow{\$}, \xleftarrow{p}$	randomly picking uniformly or according to p respectively	14

SK(Q)ES	symmetric-key (quantum) encryption scheme	14
PK(Q)ES	public-key (quantum) encryption scheme	14
n	security parameter	14
k, pk, sk	(symmetric), public or secret/private key respectively	14
$\mathrm{Enc}, \mathrm{Enc}_k, \mathrm{Enc}_{pk}$	encryption algorithm (with key k or pk)	14
$\mathrm{Dec}, \mathrm{Dec}_k, \mathrm{Dec}_{sk}$	decryption algorithm (with key k or sk)	14
KeyGen	key generation algorithm	14
$\mathrm{Enc}_K, \mathrm{Dec}_K$	average encryption/decryption over all keys	14
$\perp, \lvert\perp\rangle\langle\perp\rvert$	a value or state output by an algorithm to indicate failure	14
Π	an encryption scheme	14
Q_n	a quantum circuit	14
M, T, C	the registers for plaintext, tag, and ciphertext, $MT = C$	14
$B, \hat{B}$	the register containing an attackers side information	14
A, Λ_A	an attack, as (quantum) algorithm or channel	14
$\tilde{A}, \tilde{\Lambda}_A$	the effective map of an attack	14
CNM, QCNM	(quantum) comparison-based non-malleability	19,32
$\mathrm{CNM}_{PQ}, \mathrm{QCNM}_{PQ}$	post-quantum (Q)CNM	34
NM, PNM	(plaintext) non-malleability	21,24
DNS	DNS/plaintext authentication	21
$\mathcal{A}, \mathcal{B}$	a (Q)CNM adversary	19,32

APPENDIX B

Proof of PNM characterization theorem

Theorem B.1 (Theorem 4.1)**.** *Let* $\Pi = (\mathrm{KeyGen}, \mathrm{Enc}, \mathrm{Dec})$ *be an arbitrary* ε-PNM SKQES *for some* ε*, then for any attack* $\Lambda_A^{CB\to C\hat{B}}$*, its effective map* $\tilde{\Lambda}_A^{MB\to M\hat{B}}$ *is such that*

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

where

$$\Lambda_1 = \mathrm{Tr}_{MM'}\left[\phi^{+MM'}\tilde{\Lambda}_A(\phi^{+MM'}\otimes(\cdot))\right] \qquad \textit{and}$$
$$\Lambda_2 = \mathrm{Tr}_{MM'}\left[(\mathbb{I}^{MM'} - \phi^{+MM'})\tilde{\Lambda}_A(\phi^{+MM'}\otimes(\cdot))\right].$$

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

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

where $\tilde{\Lambda}_{ideal}^{MB\to M\hat{B}} = \mathrm{id}^M \otimes \Lambda_1 + \frac{1}{|M|^2-1}(|M|^2\langle \mathrm{Dec}_K(\tau)\rangle - \mathrm{id})^M \otimes \Lambda_2$. Lastly, let $\Lambda_3 = \mathrm{Tr}_{MM'}\left[\phi^{+MM'}\tilde{\Lambda}_A(\phi^{+MM'}\otimes(\cdot))\right]$, $\Lambda_4 = \mathrm{Tr}_{MM'}\left[(\mathbb{I}^{MM'} - \phi^{+MM'})\tilde{\Lambda}_A(\phi^{+MM'}\otimes(\cdot))\right]$, and $\tilde{\Lambda}_{trace}^{MB\to M\hat{B}} = \mathrm{id}^M \otimes \Lambda_3 + \frac{1}{|M|^2-1}(|M|^2\langle \mathrm{Dec}_K(\tau)\rangle - \mathrm{id})^M \otimes \Lambda_4$.

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,

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

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

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

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[1], we have $\||0\rangle\langle 0|\otimes(\Lambda_3-\Lambda_5)+|1\rangle\langle 1|\otimes(\Lambda_4-\Lambda_6)\|_\diamond \leq \varepsilon$ 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

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

where the last equality holds because

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

[1] See [Wat18], Proposition 3.48(1)

Similarly

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

From this we conclude

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

which means that Π is 3ε-PNM. □