Security of Quantum Key Distribution Protocols

Rotem Liss

Security of Quantum Key Distribution Protocols

Research Thesis

In partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Rotem Liss

Submitted to the Senate
of the Technion — Israel Institute of Technology
Sivan 5781 Haifa May 2021

The research thesis was done under the supervision of Assoc. Prof. Tal Mor in the Faculty of Computer Science.

Most results in this thesis have been published as articles by the author and research collaborators in journals and conferences:

1. Michel Boyer, Matty Katz, Rotem Liss, and Tal Mor. Experimentally feasible protocol for semiquantum key distribution. *Physical Review A*, 96:062335, Dec 2017. doi:10.1103/PhysRevA.96.062335. (Chapter 3)

2. Michel Boyer, Rotem Liss, and Tal Mor. Attacks against a simplified experimentally feasible semiquantum key distribution protocol. *Entropy*, 20(7):536, Jul 2018. doi:10.3390/e20070536. (Chapter 4)

3. Walter O. Krawec, Rotem Liss, and Tal Mor. Security proof against collective attacks for an experimentally feasible semi-quantum key distribution protocol. *arXiv preprint arXiv:2012.02127*, Dec 2020. URL: https://arxiv.org/abs/2012.02127. (Chapter 5)

4. Michel Boyer, Rotem Liss, and Tal Mor. Composable security against collective attacks of a modified BB84 QKD protocol with information only in one basis. *Theoretical Computer Science*, 801:96–109, Jan 2020. doi:10.1016/j.tcs.2019.08.014. (Chapter 6)

5. Rotem Liss and Tal Mor. From practice to theory: The "Bright Illumination" attack on quantum key distribution systems. In Carlos Martín-Vide, Miguel A. Vega-Rodríguez, and Miin-Shen Yang, editors, *Theory and Practice of Natural Computing*, pages 82–94, Cham, Dec 2020. Springer International Publishing. doi:10.1007/978-3-030-63000-3_7. (Chapter 8)

Acknowledgements

I would like to thank my advisor, Assoc. Prof. Tal Mor, for his very helpful guidance, discussions, ideas, and advice during this research and for his invaluable help and encouragement throughout my graduate studies. I would also like to thank Assoc. Prof. Michel Boyer and Asst. Prof. Walter Krawec for fruitful research collaborations and many research discussions leading to a variety of results and joint publications.

I would also like to thank Gilles Brassard, Román Orús, Renato Renner, Rotem Arnon-Friedman, Andreas Winter, John Smolin, Charles Bennett, David DiVincenzo, Cica Gustiani, Louis Salvail, Eli Biham, Yossi Weinstein, Roman Shapira, Yair Rezek, and Itay Fayerverker.

My family deserves special thanks.

The generous financial help of Daniel's fund, Jacobs' fund, and the Technion is gratefully acknowledged.

Contents

List of Figures

Abstract

The counter-intuitive features of quantum mechanics make it possible to solve problems and perform tasks that are beyond the abilities of non-quantum (classical) computers and communication devices. The field of *quantum information processing* studies how we can achieve such improvements by representing information as quantum states.

One of the early achievements of quantum information processing is the development of *quantum key distribution* (QKD). QKD protocols allow two participants (Alice and Bob) to achieve the classically-impossible task of generating a secret shared key even if their adversary is computationally unlimited.

Unfortunately, the security promises of QKD are true only in theory; practical implementations of QKD deviate from the theoretical protocols, and many of these deviations give rise to practical attacks. In this research thesis, we study the security properties of various QKD protocols in many practical settings:

First, we study practical security of *semiquantum key distribution* (SQKD) protocols, where either Alice or Bob is non-quantum (classical). Following practical security problems in previous SQKD protocols, we suggest a new SQKD protocol (the "Mirror protocol") which can be securely implemented, and we prove it robust and secure against a wide range of attacks (the "uniform collective attacks").

Then, we study "composable security" of the first QKD protocol created by Bennett and Brassard (BB84). BB84 has its unconditional security proved against adversaries performing the most general attacks in a theoretical (idealized) setting; however, some security approaches do not prove "composable security", which requires the secret key to remain secret even when Alice and Bob actually use it for cryptographic purposes. We generalize an algebraic security approach for BB84, making it prove composable security of BB84 (and many variants of BB84) against the most general attacks.

Finally, we analyze an important practical attack (named "Bright Illumination"), showing how it can be modeled as a theoretical "Reversed-Space" attack.

Overall, all results aim to enhance our understanding on how to bridge the gap between theory and practice in various sub-fields of QKD, and they may help solve a major open problem in the field of QKD: constructing a realistic QKD implementation that can be proved truly and unconditionally secure against any possible attack.

Chapter 1

Introduction to Quantum Information Processing

The field of *quantum information processing* (QIP) uses the laws of quantum physics for performing tasks that are impossible (or hard) in the non-quantum world.

In this chapter, we describe the basic notions of QIP that are used in this thesis. See [NC00, Gru99, RP00, GMD02] for more background and explanations about QIP.

1.1 Quantum States

In QIP, information is represented by *quantum states.* A quantum state is the state of a specific physical system; all possible quantum states of the system belong to a *Hilbert space*, which is defined as a vector space over the field $\mathbb{C}$ (the field of complex numbers) that has an inner product and satisfies the "completeness" property (the exact definition of this property can be found in standard textbooks, and it is satisfied by all finite-dimensional inner product spaces). A *quantum pure state* is represented by $|\psi\rangle$, which denotes a normalized column vector (namely, a column vector of norm 1) in the Hilbert space. In other words, the Hilbert space is the set of *all* possible quantum pure states of a system (including the non-normalized states).

As an important example, the *qubit* Hilbert space is $\mathcal{H}_2 \triangleq \text{Span}\{|0\rangle, |1\rangle\}$, where $|0\rangle$ and $|1\rangle$ are two orthonormal vectors (namely, they are normalized, and their inner product is 0). Two other important orthonormal states in $\mathcal{H}_2$ are $|+\rangle \triangleq \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $|-\rangle \triangleq \frac{|0\rangle-|1\rangle}{\sqrt{2}}$. The most general qubit pure state is $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$ (a normalization condition). The qubit states are sometimes denoted by their vector representations in the $\{|0\rangle, |1\rangle\}$ basis: $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $|-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, and $|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$. We note that $\{|0\rangle, |1\rangle\}$ is an orthonormal basis named "the z basis", "the computational basis", or "the standard basis", and $\{|+\rangle, |-\rangle\}$ is an orthonormal basis named "the x basis" or "the Hadamard basis". We also note various notations of both bases: the states of the z basis are sometimes denoted $\{|0^0\rangle, |1^0\rangle\}$ or $\{|0\rangle_0, |1\rangle_0\}$, and the states of the x basis are sometimes denoted $\{|0^1\rangle, |1^1\rangle\}$ or $\{|0\rangle_1, |1\rangle_1\}$.

We note that multiplying a pure state $|\psi\rangle$ by any global phase $e^{i\phi} \triangleq \cos(\phi) + i\sin(\phi)$ has no physical significance. In other words, two pure states that differ only by a global multiplicative phase $e^{i\phi}$ are identical for all intents and purposes.

The $|\psi\rangle$ notation (the column vector) is named *ket*. A related notation, $\langle\psi|$, is named *bra*, and it is a row vector defined by $\langle\psi| \triangleq [|\psi\rangle]^\dagger$: namely, the bra is the *conjugate transpose* of the ket. For example, if $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, then $\langle\psi| = \alpha^\star\langle 0| + \beta^\star\langle 1|$ (where $\alpha^\star$ is the complex conjugate of α); in vector notations, $\langle\psi| = \begin{pmatrix} \alpha^\star & \beta^\star \end{pmatrix}$.

Given an orthonormal basis $\{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_m\rangle\}$, the *inner product* of two pure states $|\psi\rangle = \sum_{j=1}^m \alpha_j|\psi_j\rangle$ and $|\phi\rangle = \sum_{j=1}^m \beta_j|\psi_j\rangle$ is $\langle\psi|\phi\rangle \triangleq \sum_{j=1}^m \alpha_j^\star\beta_j$. The *norm* of $|\psi\rangle$ is $|||\psi\rangle|| \triangleq \sqrt{\langle\psi|\psi\rangle} = \sqrt{\sum_{j=1}^m |\alpha_j|^2}$, and $|\psi\rangle$ is said to be *normalized* if $\sum_{j=1}^m |\alpha_j|^2 = 1$.

1.1.1 Quantum Measurements

Quantum physics allows us to *measure* a quantum state $|\psi\rangle$ with respect to any orthonormal basis $\{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_m\rangle\}$. The possible measurement outcomes are all states "ψ_j" of this orthonormal basis; each outcome "ψ_j" (corresponding to the quantum state $|\psi_j\rangle$) is obtained with probability $p_j = |\langle\psi_j|\psi\rangle|^2$. Note that $\sum_{j=1}^m p_j = \langle\psi|\psi\rangle = 1$. Also note that a measurement result "ψ_j" is a *classical* (non-quantum) indicator that can be read and used; we have not discussed the resulting *quantum state* after measurement, but we should note that the original quantum state $|\psi\rangle$ may be ruined (or change its state) following the measurement operation.

For example, if the quantum state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is measured with respect to the orthonormal basis $\{|0\rangle, |1\rangle\}$ (namely, it is "measured in the z basis"), the "0" result is obtained with probability $|\langle 0|\psi\rangle|^2 = |\alpha|^2$, and the "1" result is obtained with probability $|\langle 1|\psi\rangle|^2 = |\beta|^2$. Notice that the normalization condition $|\alpha|^2 + |\beta|^2 = 1$ means that these two probabilities sum to 1.

There are also generalized definitions of quantum measurements (see, e.g., in [NC00]), but they can all be reduced to the set of quantum operations described in Section 1.4.

1.1.2 Unitary Operators

Quantum physics allows us to apply any *unitary operator* $U : \mathcal{H} \to \mathcal{H}$ to a quantum state in the Hilbert space $\mathcal{H}$. Unitary operators are linear operators (namely, $U[\alpha|\psi\rangle + \beta|\phi\rangle] = \alpha U|\psi\rangle + \beta U|\phi\rangle$) that satisfy $U^\dagger = U^{-1}$. They preserve inner products and norms.

As an important example, the *Hadamard operator* on the qubit space is defined by $H \triangleq \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$: namely, $H|0\rangle = |+\rangle$ and $H|1\rangle = |-\rangle$. It also satisfies $H|+\rangle = |0\rangle$ and $H|-\rangle = |1\rangle$.

1.2 Bipartite and Multipartite Hilbert Spaces

1.2.1 Tensor Products of Hilbert Spaces

Given two physical systems, A and B, we would like to mathematically represent the *compound* physical system AB (comprised of two subsystems: A and B), given that a quantum state of subsystem A is represented by a vector in the Hilbert space $\mathcal{H}_{\mathrm{A}}$, and a quantum state of subsystem B is represented by a vector in the Hilbert space $\mathcal{H}_{\mathrm{B}}$.

In this case, a quantum state of the compound (*bipartite*) system AB is represented by a vector in the *tensor product* Hilbert space $\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$. An orthonormal basis for this Hilbert space can be a *product* of two orthonormal bases for $\mathcal{H}_{\mathrm{A}}$ and $\mathcal{H}_{\mathrm{B}}$: namely, if $\{|\psi_1\rangle_{\mathrm{A}}, |\psi_2\rangle_{\mathrm{A}}, \ldots, |\psi_k\rangle_{\mathrm{A}}\}$ is an orthonormal basis for $\mathcal{H}_{\mathrm{A}}$, and $\{|\phi_1\rangle_{\mathrm{B}}, |\phi_2\rangle_{\mathrm{B}}, \ldots, |\phi_m\rangle_{\mathrm{B}}\}$ is an orthonormal basis for $\mathcal{H}_{\mathrm{B}}$, then $\{|\psi_j\rangle_{\mathrm{A}} \otimes |\phi_\ell\rangle_{\mathrm{B}} \mid 1 \leq j \leq k \ , \ 1 \leq \ell \leq m\}$ is an orthonormal basis for $\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$.

As an important example, if A and B are both qubit systems (namely, $\mathcal{H}_{\mathrm{A}}$ and $\mathcal{H}_{\mathrm{B}}$ are both $\mathcal{H}_2 \triangleq \mathrm{Span}\{|0\rangle, |1\rangle\}$), the compound *two-qubit system* is represented by $\mathcal{H}_2 \otimes \mathcal{H}_2 = \mathrm{Span}\{|0\rangle \otimes |0\rangle \ , \ |0\rangle \otimes |1\rangle \ , \ |1\rangle \otimes |0\rangle \ , \ |1\rangle \otimes |1\rangle\}$. A shorter notation is $\mathcal{H}_2 \otimes \mathcal{H}_2 = \mathrm{Span}\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

The tensor product of three or more Hilbert spaces (giving a *multipartite* Hilbert space) is similarly defined. For example, $\mathcal{H}_2 \otimes \mathcal{H}_2 \otimes \mathcal{H}_2$ (a tripartite Hilbert space that is the *three-qubit* space) is

$$\mathrm{Span}\{|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle, |101\rangle, |110\rangle, |111\rangle\}. \tag{1.1}$$

1.2.2 Tensor Products of Vectors

Given two Hilbert spaces, $\mathcal{H}_{\mathrm{A}}$ with an orthonormal basis $\{|\psi_1\rangle_{\mathrm{A}}, |\psi_2\rangle_{\mathrm{A}}, \ldots, |\psi_k\rangle_{\mathrm{A}}\}$ and $\mathcal{H}_{\mathrm{B}}$ with an orthonormal basis $\{|\phi_1\rangle_{\mathrm{B}}, |\phi_2\rangle_{\mathrm{B}}, \ldots, |\phi_m\rangle_{\mathrm{B}}\}$, and given two vectors $|\psi\rangle_{\mathrm{A}} = \sum_{j=1}^{k} \alpha_j |\psi_j\rangle_{\mathrm{A}} \in \mathcal{H}_{\mathrm{A}}$ and $|\phi\rangle_{\mathrm{B}} = \sum_{\ell=1}^{m} \beta_\ell |\phi_\ell\rangle_{\mathrm{B}} \in \mathcal{H}_{\mathrm{B}}$, the *tensor product* vector $|\psi\rangle_{\mathrm{A}} \otimes |\phi\rangle_{\mathrm{B}} \in \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$ (or, using a shorter notation, $|\psi\rangle_{\mathrm{A}}|\phi\rangle_{\mathrm{B}}$) is defined as

$$|\psi\rangle_{\mathrm{A}}|\phi\rangle_{\mathrm{B}} \triangleq \sum_{j=1}^{k}\sum_{\ell=1}^{m} (\alpha_j|\psi_j\rangle_{\mathrm{A}}) \otimes (\beta_\ell|\phi_\ell\rangle_{\mathrm{B}}) = \sum_{j=1}^{k}\sum_{\ell=1}^{m} \alpha_j\beta_\ell|\psi_j\rangle_{\mathrm{A}}|\phi_\ell\rangle_{\mathrm{B}}. \tag{1.2}$$

For example, given $|\psi\rangle_{\mathrm{A}} = \alpha|0\rangle_{\mathrm{A}} + \beta|1\rangle_{\mathrm{A}} \in \mathcal{H}_2$ and $|\phi\rangle_{\mathrm{B}} = \gamma|0\rangle_{\mathrm{B}} + \delta|1\rangle_{\mathrm{B}} \in \mathcal{H}_2$, the tensor product vector $|\psi\rangle_{\mathrm{A}}|\phi\rangle_{\mathrm{B}} \in \mathcal{H}_2 \otimes \mathcal{H}_2$ is

$$|\psi\rangle_{\mathrm{A}}|\phi\rangle_{\mathrm{B}} = \alpha\gamma|00\rangle_{\mathrm{AB}} + \alpha\delta|01\rangle_{\mathrm{AB}} + \beta\gamma|10\rangle_{\mathrm{AB}} + \beta\delta|11\rangle_{\mathrm{AB}}. \tag{1.3}$$

An example is

$$|{+}{-}\rangle_{\mathrm{AB}} = \left[\frac{|0\rangle_{\mathrm{A}} + |1\rangle_{\mathrm{A}}}{\sqrt{2}}\right] \otimes \left[\frac{|0\rangle_{\mathrm{B}} - |1\rangle_{\mathrm{B}}}{\sqrt{2}}\right] = \frac{1}{2}[|00\rangle_{\mathrm{AB}} - |01\rangle_{\mathrm{AB}} + |10\rangle_{\mathrm{AB}} - |11\rangle_{\mathrm{AB}}]. \tag{1.4}$$

We note that most states in $\mathcal{H}_{\rm A} \otimes \mathcal{H}_{\rm B}$ are *not* tensor product vectors, and are thus called *entangled*. Four important entangled two-qubit states (that form together an orthonormal basis of $\mathcal{H}_2 \otimes \mathcal{H}_2$, named the *Bell basis* or the *BMR basis*) are:

$$|\Phi_+\rangle \triangleq \frac{|00\rangle + |11\rangle}{\sqrt{2}} \quad , \quad |\Phi_-\rangle \triangleq \frac{|00\rangle - |11\rangle}{\sqrt{2}} \quad , \tag{1.5}$$

$$|\Psi_+\rangle \triangleq \frac{|01\rangle + |10\rangle}{\sqrt{2}} \quad , \quad |\Psi_-\rangle \triangleq \frac{|01\rangle - |10\rangle}{\sqrt{2}} \quad . \tag{1.6}$$

The definition of the tensor product is easily generalized to tensor products of three (or more) vectors: for example,

$$\begin{aligned} |{+}0{-}\rangle_{\rm ABC} &= \left[\frac{|0\rangle_{\rm A} + |1\rangle_{\rm A}}{\sqrt{2}}\right] \otimes |0\rangle_{\rm B} \otimes \left[\frac{|0\rangle_{\rm C} - |1\rangle_{\rm C}}{\sqrt{2}}\right] \\ &= \frac{1}{2}[|000\rangle_{\rm ABC} - |001\rangle_{\rm ABC} + |100\rangle_{\rm ABC} - |101\rangle_{\rm ABC}]. \end{aligned} \tag{1.7}$$

1.2.3 Tensor Products of Operators

Given two linear operators, U operating on the Hilbert space $\mathcal{H}_{\rm A}$ and V operating on the Hilbert space $\mathcal{H}_{\rm B}$, the linear operator $U \otimes V$ operates on the Hilbert space $\mathcal{H}_{\rm A} \otimes \mathcal{H}_{\rm B}$ and is defined as follows:

$$(U \otimes V)(|\psi\rangle_{\rm A} \otimes |\phi\rangle_{\rm B}) \triangleq (U|\psi\rangle_{\rm A}) \otimes (V|\phi\rangle_{\rm B}). \tag{1.8}$$

(It extends by linearity to vectors that are not tensor products, such as $\frac{|00\rangle_{\rm AB}+|11\rangle_{\rm AB}}{\sqrt{2}}$.)

For example, the tensor product of the Hadamard operator H with itself, denoted $H \otimes H$ or $H^{\otimes 2}$, operates as follows:

$$\begin{aligned} H^{\otimes 2}|00\rangle_{\rm AB} &= (H|0\rangle_{\rm A}) \otimes (H|0\rangle_{\rm B}) = |{+}{+}\rangle_{\rm AB} \ , &(1.9)\\ H^{\otimes 2}|01\rangle_{\rm AB} &= (H|0\rangle_{\rm A}) \otimes (H|1\rangle_{\rm B}) = |{+}{-}\rangle_{\rm AB} \ , &(1.10)\\ H^{\otimes 2}|10\rangle_{\rm AB} &= (H|1\rangle_{\rm A}) \otimes (H|0\rangle_{\rm B}) = |{-}{+}\rangle_{\rm AB} \ , &(1.11)\\ H^{\otimes 2}|11\rangle_{\rm AB} &= (H|1\rangle_{\rm A}) \otimes (H|1\rangle_{\rm B}) = |{-}{-}\rangle_{\rm AB} \ . &(1.12) \end{aligned}$$

This definition is generalized to tensor products of three (or more) operators.

1.3 Quantum Mixed States

A *quantum mixed state* is a probability distribution over several pure states: namely, it is a set $\{(|\psi_j\rangle, q_j)\}_j$ consisting of pairs of pure states $|\psi_j\rangle$ and probabilities q_j (where $0 < q_j \le 1$ and $\sum_j q_j = 1$), meaning that each pure state $|\psi_j\rangle$ has probability q_j.

Unlike a pure state, a mixed state is *not* represented by a vector in Hilbert space. It is represented by a density matrix: $\rho = \sum_j q_j |\psi_j\rangle\langle\psi_j|$, where q_j is the probability of the pure state $|\psi_j\rangle$. (This definition of q_j should not be confused with the probabilities of *measurement* results, mentioned in Subsection 1.1.1.) In particular, the pure state $|\psi\rangle$ is represented by the density matrix $\rho = |\psi\rangle\langle\psi|$.

For example, if the system is prepared in state $|0\rangle$ with probability $\frac{1}{3}$ or in state $|+\rangle$ with probability $\frac{2}{3}$, the corresponding quantum mixed state has density matrix $\rho = \frac{1}{3}|0\rangle\langle 0| + \frac{2}{3}|+\rangle\langle +|$. It should be emphasized that these probabilities are of *preparation*, not of any *measurement*. For example, if this state is measured in the z basis $\{|0\rangle, |1\rangle\}$, the probability of measuring "0" is $\frac{1}{3} \cdot 1 + \frac{2}{3} \cdot \frac{1}{2} = \frac{2}{3}$, and the probability of measuring "1" is $\frac{1}{3} \cdot 0 + \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$; and if the state is measured in the x basis $\{|+\rangle, |-\rangle\}$, the probability of measuring "+" is $\frac{1}{3} \cdot \frac{1}{2} + \frac{2}{3} \cdot 1 = \frac{5}{6}$, and the probability of measuring "−" is $\frac{1}{3} \cdot \frac{1}{2} + \frac{2}{3} \cdot 0 = \frac{1}{6}$. Notice that the probability of measuring "0" is *not* $\frac{1}{3}$, and the probability of measuring "+" is *not* $\frac{2}{3}$.

We note that several *different* probability distributions may represent *the same* mixed state: namely, the states they represent are physically identical (e.g., giving exactly the same measurement results in all orthonormal bases). This happens if and only if they are represented by *equal* density matrices. (A similar observation is that a global phase $e^{i\phi}$ for *pure states* has no physical significance; and, indeed, the two pure states $|\psi\rangle$ and $e^{i\phi}|\psi\rangle$ are represented by identical density matrices, $\rho = |\psi\rangle\langle\psi|$.) For example, the *completely mixed state* $\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$ is the same as $\rho = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|$, and these two density matrices are equal.

A density matrix must always satisfy three conditions: (a) it is a *Hermitian* matrix; (b) it is *positive semidefinite*; and (c) it is *normalized* (namely, its trace equals 1). These three conditions are also *sufficient*: any matrix ρ satisfying them is a density matrix. From these conditions it follows that any density matrix ρ can be presented as $\rho = \sum_j \lambda_j |\psi_j\rangle\langle\psi_j|$ (the spectral decomposition), where $\lambda_j \geq 0$, $\sum_j \lambda_j = 1$, and $\{|\psi_j\rangle\}_j$ is an orthonormal set (of eigenvectors); in other words, for any mixed state we can choose a corresponding probability distribution over a set of *orthonormal* states. For example, for $\rho = \frac{1}{3}|0\rangle\langle 0| + \frac{2}{3}|+\rangle\langle +|$, the spectral decomposition is

$$\begin{aligned} \rho &= \frac{3+\sqrt{5}}{6} \left[\frac{2|0\rangle + (\sqrt{5}-1)|1\rangle}{\sqrt{10-2\sqrt{5}}}\right] \left[\frac{2\langle 0| + (\sqrt{5}-1)\langle 1|}{\sqrt{10-2\sqrt{5}}}\right] \\ &+ \frac{3-\sqrt{5}}{6} \left[\frac{2|0\rangle - (\sqrt{5}+1)|1\rangle}{\sqrt{10+2\sqrt{5}}}\right] \left[\frac{2\langle 0| - (\sqrt{5}+1)\langle 1|}{\sqrt{10+2\sqrt{5}}}\right], \end{aligned} \tag{1.13}$$

and it is the *unique* decomposition of ρ as a mixture of *orthonormal* pure states; on the other hand, the completely mixed state $\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|$ has an infinite number of decompositions as a mixture of orthonormal pure states, because its eigenvalue $(\frac{1}{2})$ is degenerate—namely, it has two orthonormal eigenvectors corresponding to the same eigenvalue.

The probability distribution in the definition of mixed states represents the "standard" ("classical") notion of *uncertainty*, and not a quantum phenomenon: it simply represents a lack of knowledge. Nonetheless, mixed states naturally appear in many areas of QIP. For example, if a joint system AB is in the entangled pure state $\sqrt{\frac{1}{3}}|0\rangle_\mathrm{A}|0\rangle_\mathrm{B} + \sqrt{\frac{2}{3}}|1\rangle_\mathrm{A}|+\rangle_\mathrm{B}$, the quantum state of subsystem B is the mixed state $\rho = \frac{1}{3}|0\rangle_\mathrm{B}\langle 0|_\mathrm{B} + \frac{2}{3}|+\rangle_\mathrm{B}\langle +|_\mathrm{B}$ (see Subsection 1.3.2 for details about this computation) that we have seen before. Moreover, we note that the state of the joint system AB can also be represented as $\sqrt{\frac{5}{6}}|+\rangle_\mathrm{A}\frac{2|0\rangle_\mathrm{B}+|1\rangle_\mathrm{B}}{\sqrt{5}} - \sqrt{\frac{1}{6}}|-\rangle_\mathrm{A}|1\rangle_\mathrm{B}$; thus, the state of subsystem B can also be represented as $\rho = \frac{5}{6}\left[\frac{2|0\rangle_\mathrm{B}+|1\rangle_\mathrm{B}}{\sqrt{5}}\right]\left[\frac{2\langle 0|_\mathrm{B}+\langle 1|_\mathrm{B}}{\sqrt{5}}\right] + \frac{1}{6}|1\rangle_\mathrm{B}\langle 1|_\mathrm{B}$. This is another example of multiple probability distributions corresponding to the same mixed state.

We should note an important difference between pure states and mixed states: for a pure state $|\psi\rangle$, there exists an orthonormal basis (consisting of $|\psi\rangle$ itself and states orthonormal to it) such that if we measure $|\psi\rangle$ in this basis, we obtain a specific measurement result ("ψ") *for certain.* This claim is *never* true for a mixed state ρ: if we measure ρ in *any* orthonormal basis, the measurement result is always uncertain.

1.3.1 Quantum Operations on Mixed States

Two important results (that can be mathematically proved) are:

- If we *measure* a mixed state ρ in some orthonormal basis $\{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_m\rangle\}$, we get the measurement result "ψ_j" with probability $p_j = \langle\psi_j|\,\rho\,|\psi_j\rangle$.

- If we apply a *unitary operator* U to a mixed state ρ, the resulting state is the mixed state $U\rho U^\dagger$.

1.3.2 Partial Trace: Removing (Ignoring and Forgetting) a Subsystem

Sometimes, we would like to compute the quantum state of a specific subsystem, while ignoring and forgetting the other subsystems. For example, given a tripartite quantum state ρ_ABE (shared by three parties named Alice (A), Bob (B), and Eve (E)), we may want to ignore the two subsystems A, B and look only at the state of Eve's subsystem E. In other words, we may want to assume that subsystems A, B will never be accessible to Eve (maybe they will be measured by Alice and Bob, who will then forget the measurement results or keep them secret) and find the state ρ_E of subsystem E.

The mathematical operation corresponding to this scenario is the *partial trace.* The partial trace of a bipartite state ρ_XY over subsystem X is defined as follows:

$$\rho_\mathrm{Y} = \mathrm{tr}_\mathrm{X}(\rho_\mathrm{XY}) \triangleq \sum_{|x\rangle\in\mathcal{X}} \langle x|\,\rho_\mathrm{XY}\,|x\rangle\,, \tag{1.14}$$

where $\mathcal{X}$ is an arbitrary orthonormal basis of the Hilbert space $\mathcal{H}_\mathrm{X}$ corresponding to subsystem X. The result of this computation is the quantum state ρ_Y of subsystem Y.

In the above example, the partial trace of ρ_ABE over subsystems A, B is:

$$\rho_\mathrm{E} = \mathrm{tr}_\mathrm{AB}(\rho_\mathrm{ABE}) \triangleq \sum_{|a\rangle\in\mathcal{A},|b\rangle\in\mathcal{B}} \langle a|\,\langle b|\,\rho_\mathrm{ABE}\,|a\rangle\,|b\rangle, \tag{1.15}$$

where $\mathcal{A}, \mathcal{B}$ are some arbitrary orthonormal bases of the Hilbert spaces $\mathcal{H}_{\mathrm{A}}, \mathcal{H}_{\mathrm{B}}$ corresponding to subsystems A, B.

For example, the partial trace of the two-qubit pure state $|\psi\rangle_{\mathrm{XY}} = \sqrt{\frac{1}{3}}|0\rangle_{\mathrm{X}}|-\rangle_{\mathrm{Y}} + \sqrt{\frac{2}{3}}|1\rangle_{\mathrm{X}}|+\rangle_{\mathrm{Y}}$ over subsystem X is

$$\rho_{\mathrm{Y}} = \mathrm{tr}_{\mathrm{X}}(|\psi\rangle_{\mathrm{XY}}\langle\psi|_{\mathrm{XY}}) = \frac{1}{3}|-\rangle_{\mathrm{Y}}\langle-|_{\mathrm{Y}} + \frac{2}{3}|+\rangle_{\mathrm{Y}}\langle+|_{\mathrm{Y}}, \tag{1.16}$$

and the partial trace of the same state $|\psi\rangle_{\mathrm{XY}}$ over subsystem Y is

$$\rho_{\mathrm{X}} = \mathrm{tr}_{\mathrm{Y}}(|\psi\rangle_{\mathrm{XY}}\langle\psi|_{\mathrm{XY}}) = \frac{1}{3}|0\rangle_{\mathrm{X}}\langle 0|_{\mathrm{X}} + \frac{2}{3}|1\rangle_{\mathrm{X}}\langle 1|_{\mathrm{X}}. \tag{1.17}$$

More details about the partial trace are available in standard QIP textbooks (e.g., [NC00]).

1.4 List of Allowed Quantum Operations

1. *measuring* the state with respect to an orthonormal basis (Subsection 1.1.1);
2. applying a *unitary operator* (Subsection 1.1.2);
3. *adding* a new (ancillary) subsystem; and
4. *removing* (ignoring and forgetting) a subsystem (Subsection 1.3.2).

1.5 Trace Distance

The *trace distance* between two quantum states is, informally, a measure of their *distinguishability*. This measure is very useful for security definitions of quantum key distribution protocols (see details in Subsection 2.3.1).

The trace distance of two states ρ and σ is defined as follows:

$$D(\rho, \sigma) \triangleq \frac{1}{2}\,\mathrm{tr}\,|\rho - \sigma| = \frac{1}{2}\sum_j |\lambda_j|, \tag{1.18}$$

where $\{\lambda_j\}_j$ are the eigenvalues of $\rho - \sigma$, all of which are real numbers. (We note that $|A|$ is defined as $\sqrt{A^\dagger A}$.) In other words, the trace distance $D(\rho, \sigma)$ is one half of the *sum of absolute values* of the eigenvalues of $\rho - \sigma$.

1.5.1 The Information-Theoretical Meaning of the Trace Distance

It can be proved [FvdG99, BBBGM02] that the trace distance $D(\rho, \sigma)$ between two quantum states ρ and σ upper-bounds the *Shannon Distinguishability* between ρ and σ. The Shannon Distinguishability is defined as the classical mutual information between the random variable $T \triangleq \begin{cases} 0 & \text{The quantum state is } \rho \\ 1 & \text{The quantum state is } \sigma \end{cases}$ and the random variable X (the result of a measurement), maximized over *all possible quantum measurements* (including measurements consisting of adding an ancillary state, performing a general unitary transformation, and then measuring in some orthonormal basis).

In other words, the trace distance upper-bounds the information that some user, who holds a quantum state and *does not know* whether it is ρ or σ (it can be either ρ or σ, with equal probabilities), can find by using a measurement.

Examples:

- $D(|0\rangle\langle 0|, |1\rangle\langle 1|) = 1$, because the two quantum states $|0\rangle$ and $|1\rangle$ can be distinguished *for certain* by measuring in the z basis $\{|0\rangle, |1\rangle\}$.
- $D(|0\rangle\langle 0|, |0\rangle\langle 0|) = 0$, because the two quantum state $|0\rangle$ and $|0\rangle$ are identical, so they cannot be distinguished from each other at all.

Chapter 2

Introduction to Quantum Key Distribution

The properties of quantum mechanics permit cryptographic protocols that are more secure than standard ("classical" or "non-quantum") protocols. *Quantum key distribution* (QKD) protocols allow two users, conventionally named Alice and Bob, to generate a secret shared key. This thesis is devoted to studying security properties of QKD protocols.

In this chapter, we describe relevant existing knowledge in the research field of QKD. In particular, we discuss security definitions of QKD and the notion of semiquantum key distribution (SQKD) protocols.

2.1 Motivation: Unsolved Encryption Problems in a Non-Quantum World

Cryptography is the science of protecting security and correctness of data against adversaries. One of the most important cryptographic problems is *encryption*—namely, transmitting a secret message from a sender (Alice) to a receiver (Bob), and ensuring the adversary (Eve) cannot read it. Two main encryption methods are used today:

- *Symmetric-key cryptography*, in which the same *secret key* is used by both Alice and Bob: Alice uses the secret key for encrypting her message, and Bob uses the same secret key for decrypting the message. Several examples of symmetric-key ciphers are the Advanced Encryption Standard (AES) [DR13], the older Data Encryption Standard (DES), and one-time pad ("Vernam cipher").

- *Public-key cryptography* [DH76], in which a *public key* (known to everyone) and a *secret key* (known only to Bob) are used: Alice uses the public key for encrypting her message, and Bob uses the secret key for decrypting the message. Several examples of public-key ciphers include RSA [RSA78] and elliptic curve cryptography.

However:

- Current public-key cryptography is not formally proved secure; moreover, its security is only *computational*—namely, it relies on computational hardness of specific problems, such as integer factorization and discrete logarithm. Furthermore, factorization and discrete logarithm can both be efficiently solved on a quantum computer by using Shor's factorization algorithm [Sho94, Sho99]; therefore, if a scalable quantum computer is successfully built in the future, it will break security of many public-key ciphers, including RSA and elliptic curve cryptography.

- Symmetric-key cryptography requires Alice and Bob to share a secret key *in advance*: namely, if Alice and Bob want to share a secret message, they must share a secret key beforehand. Moreover, no *computational* security proofs are known for many current symmetric-key ciphers, including AES and DES; and *unconditional* security against computationally-unlimited adversaries has been proved to require *many resources*: the secret key must be used *only once* and be *at least as long* as the secret message [Sha49].

Nonetheless, there exist ciphers that are fully and unconditionally secure (even against computationally-unlimited adversaries). For example, the *one-time pad* (symmetric-key) cipher is defined as follows: given a message M and a secret key K of the same length, the encrypted message C is computed as the XOR between M and K—namely, $C = M \oplus K$ (decryption can then be performed by computing $M = C \oplus K$). One-time pad has been proved fully and unconditionally secure against any adversary [Sha49]: even if the adversary Eve intercepts the encrypted message C, she gains *no information* on the original message M (assuming she has no information on the secret key K; in particular, assuming K is used only once).

Therefore, for achieving perfectly secure encryption, we only need an efficient way for *sharing a random secret key* between Alice and Bob—a task named "key distribution". Unfortunately, unconditionally secure "classical [non-quantum] key distribution" is *impossible* if the computationally-unlimited adversary can listen to all communication between Alice and Bob. Fortunately, *quantum key distribution* can solve this problem.

2.2 Quantum Key Distribution

Quantum key distribution (QKD) protocols allow Alice and Bob to generate a shared random key. Typically, they require Alice and Bob to use two communication channels: (a) an *insecure quantum channel* (to which Eve may apply any operation allowed by the laws of quantum physics), and (b) an *unjammable classical channel* (to which Eve may listen, but not interfere). Eve listens to both channels and tries obtaining as much information as she can on the final shared key.

For most QKD protocols, the final key is *proved* to be secret even from the most powerful adversaries—adversaries who are limited only by the laws of nature and who are otherwise capable of solving any computational problem and performing any physically-allowed operation. Moreover, the final key is proved to remain secret in the future, even if Eve improves her computational power and other capabilities.

After creating the shared key, Alice and Bob can use it for other cryptographic tasks (e.g., one-time pad encryption). More generally, QKD protocols can be used as a subroutine (secure key distribution) of more complicated cryptographic protocols; in other words, we can integrate QKD into a system to improve its security. See [SML10] for more details about this integration and the practical usability of QKD compared to other methods.

2.2.1 The QKD Protocol of Bennett and Brassard (BB84)

The first and most important QKD protocol, suggested by Bennett and Brassard in 1984, is BB84 [BB84].

In the BB84 protocol, in each round, Alice randomly chooses one of the four possible "BB84 (qubit) states" $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$ and sends it to Bob. Bob randomly chooses one of two orthonormal bases (z or x, both defined in Section 1.1) and measures the state in his chosen basis. If Bob chooses the same basis as Alice (assuming that Eve did not interfere), Bob will get *the same result* as Alice; and if Bob chooses a different basis from Alice (assuming, again, that Eve did not interfere), Bob will get *a random result* (each result with probability $\frac{1}{2}$). For example, if Alice sends $|+\rangle$ and Bob measures it in the x basis, he will get the "+" result for certain; but if Bob measures it in the z basis, he will get a random result (either "0" or "1").

After sending and receiving N qubits (in N rounds), Alice and Bob perform "classical post-processing" (namely, they process their results in a coordinated way, exchanging information via the unjammable classical channel), comprised of the following steps:

1. Alice and Bob expose and compare the bases they chose and discard the qubits for which they chose different bases.

2. Alice and Bob expose and compare a randomly chosen subset of their bits (named "the TEST bits"), check the error rate in this subset, and abort the protocol if the error rate is above a specific threshold. The remaining bits are "the INFO bits".

3. Alice and Bob perform error correction and privacy amplification processes on the INFO bits, so that both of them have the same bit string (the *final key*) and Eve's average information about it is negligible—namely, exponentially small in N.

Full definitions of BB84 and several variants are available in Sections 6.2 and 7.1.

2.2.2 Types of QKD Protocols

Many QKD protocols have been suggested over the years. We should note three important classifications:

1. BB84 and similar protocols are "prepare-and-measure" protocols, because the legitimate parties prepare quantum states, transmit them, and measure them. In contrast, in "entanglement-based" QKD protocols, an untrusted center gives the legitimate parties allegedly-*entangled* quantum states (see Subsection 1.2.2), and the legitimate parties test them and use them for generating a secret key. "Entanglement-based" protocols were first discussed by [Eke91, BBM92], and they are usually almost equivalent to "prepare-and-measure" protocols [BBM92].

2. BB84 and similar protocols are "one-way" protocols, because each quantum state travels *once* from one legitimate party to the other—for example, from Alice to Bob. In contrast, "two-way" protocols [BLMR13] require each quantum state to travel *twice* between the legitimate parties—for example, from Bob to Alice and back to Bob; examples of two-way protocols include the *semiquantum* key distribution protocols discussed in Section 2.4 and Chapters 3–5.

3. All QKD protocols discussed in this thesis are "discrete-variable" protocols, because they use finite-dimensional Hilbert spaces (or, more generally, discrete random variables). In contrast, "continuous-variable" QKD protocols use different techniques; see [SBPCDLP09, XMZLP20, PAB$^+$20] for more details.

2.3 Security and Robustness of QKD

2.3.1 Security Definitions and Composable Security

The main objective of analyzing a QKD protocol is proving its *unconditional security*: proving that even if Eve applies the strongest and most general attacks allowed by the laws of nature (named the "joint attacks"), Eve's average information about the final key is still negligible—namely, exponentially small in the number of rounds.

Originally, a QKD protocol was defined "secure" if the (classical) average *mutual information* between Eve's final measurement result (E) and Alice's and Bob's final shared key (K)[1], maximized over all possible attack strategies and measurements by Eve, was exponentially small in the number of rounds N. Examples of BB84 security proofs based on this security definition include [May01, BBBMR06, SP00]: these security proofs recognized that one cannot analyze the *classical* data held by Eve before privacy amplification (as was done in [BBCM95]), but must analyze Eve's *quantum* state [BMS96]. In other words, they assumed Eve could keep her quantum state until the end of the protocol, and only *then* choose the optimal measurement (based on all the data she observed) and perform this measurement.

Later, it was noticed that this security definition might not be "composable". In other words, although the final key itself is secure if Eve measures the quantum state

[1] More precisely, the *security* definition referred to *Alice*'s final key (A), and a separate condition (*reliability*) required Bob's final key to be identical to Alice's final key, except with negligible probability.

she holds at the end of the QKD protocol, the proof does not apply to *cryptographic applications* of the final key (e.g., encryption): Eve may gain non-negligible information after the key is used, even though her information on the key itself was negligible. This means that the above definition is not sufficient for practical applications: such applications may be insecure if Eve keeps her quantum state until Alice and Bob *use* the final key (thus giving Eve some new information) and only *then* measures.

Therefore, a new notion of "(composable) full security" was defined [BOHLMO05, RGK05, Ren08], following similar definitions of universally composable security in non-quantum cryptography [Can01, PW00], and using the trace distance (see Section 1.5). Intuitively, this notion requires that the final joint quantum state of Alice, Bob, and Eve generated by the QKD protocol is *very close* (namely, the trace distance is exponentially small in N) to the final state generated by an *ideal* key distribution protocol which distributes a *completely random and secret* final key to both Alice and Bob. In other words, if a QKD protocol is (composably) secure, then except with an exponentially small probability, one of the two following events happens: the QKD protocol is aborted, *or* the QKD protocol generates a secret key with the same properties as a perfect key—(a) uniformly distributed (i.e., each possible key has the same probability), (b) identical for Alice and Bob, and (c) independent of Eve's information.

Formally:

- ρ_{ABE} is defined as the final quantum state of Alice, Bob, and Eve at the end of the protocol: Alice's and Bob's quantum states are simply the "classical" states $|k_{\mathrm{A}}\rangle_{\mathrm{A}}, |k_{\mathrm{B}}\rangle_{\mathrm{B}}$, where the bit strings $k_{\mathrm{A}}, k_{\mathrm{B}}$ are the final keys held by Alice and Bob, respectively (ideally, $k_{\mathrm{A}} = k_{\mathrm{B}}$); and Eve's state includes both her quantum ancillary state and the classical information sent over the classical channel.

- ρ_{U} is defined as the complete mixture of all possible final keys that are identical for Alice and Bob. Namely, if the set of possible final keys is K, then:

$$\rho_{\mathrm{U}} \triangleq \frac{1}{|K|} \sum_{k \in K} |k\rangle_{\mathrm{A}} |k\rangle_{\mathrm{B}} \langle k|_{\mathrm{A}} \langle k|_{\mathrm{B}}. \tag{2.1}$$

- ρ_{E} is defined as the partial trace of ρ_{ABE} over the system AB; the definition of the partial trace is available in Subsection 1.3.2.

For the QKD protocol to be fully (composably) secure, the definition requires the following trace distance (see Section 1.5) to satisfy

$$\frac{1}{2} \operatorname{tr} |\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}| \leq \epsilon, \tag{2.2}$$

where ϵ is exponentially small in the number of rounds N. Intuitively, ρ_{ABE} is the *actual* joint state of Alice, Bob, and Eve at the end of the QKD protocol; ρ_{U} is the *ideal* final state of Alice and Bob (an equal mixture of all possible final keys, that is identical for Alice and Bob and completely uncorrelated with Eve); and ρ_{E} is the state of Eve, uncorrelated with the states of Alice and Bob. Note that cases where the protocol is aborted are represented by the zero operator: see [Ren08, Subsection 6.1.2] for details.

We note that non-composable security *does not imply* composable security: in an example found by [KRBM07], the final key satisfied the non-composable security definition, but it could not be securely used even for the one-time pad encryption scheme described in Section 2.1. However, it was shown by [BOHLMO05] that if the mutual information (used in the non-composable security definition) is bounded by μ_2, and the final key is uniformly random and identical for Alice and Bob except with probability $\frac{\mu_1}{2}$, then the trace distance (used in the composable security definition) can be bounded by $2^{\max(m)/2}\sqrt{\mu_2} + \mu_1$ (where $\max(m)$ is the maximal length of the final key given the number of rounds N).

Using this general bound of [BOHLMO05], if μ_2 is exponentially small in the final key length m *and* the exponential decay is sufficiently fast, composable security can sometimes be proved; however, this general bound sometimes *does not imply* composable security, and it is usually non-tight: better bounds can usually be directly found for special cases, similarly to the other bounds suggested by [BOHLMO05] and to the bounds found in Chapters 6 and 7. (We note that the results of [KRBM07] and [BOHLMO05] are consistent with each other: in the example given by [KRBM07], the mutual information is exponentially small in m, while the trace distance is constant, and [BOHLMO05]'s non-tight upper bound on the trace distance *grows* exponentially with m.)

Composable security proofs have been presented for many QKD protocols, including BB84 [RGK05, Ren08].

2.3.2 Collective, "Uniform Collective", and Joint Attacks

Our ultimate objective is proving security of QKD against the most general attacks Eve can possibly apply. However, the most general attacks can be very complicated, so we usually first analyze an important and powerful subclass of attacks named the "collective attacks" [BM97b, BM97a, BBBGM02]. It is sometimes easier to prove security against collective attacks than security against the most general attacks; moreover, security against collective attacks is conjectured (and, in some security notions, proved [Ren08, CKR09]) to imply security against the most general attacks.

Intuitively, in a collective attack, Eve begins by attacking each round *separately*, and she uses a separate probe state (ancillary state) for each round. These probe states cannot be entangled or correlated with each other, but Eve can keep them in a quantum memory. Later, after Alice and Bob have completed classical post-processing, Eve can measure all her probe states together in the optimal way. A formal description of the collective attacks against BB84-like protocols is available in Subsection 6.3.1.

The definition of the "collective attacks" is slightly different in some papers (most notably, [Ren08, RGK05, CKR09]): these papers require Eve not only to attack the rounds *separately and independently*, but they also require her to attack them *identically* (namely, she must apply the same operation in each round). To avoid confusion, we call this specific type of collective attacks "*uniform* collective attacks".

The class of the "joint attacks" includes all theoretical attacks allowed by quantum physics (namely, these are the most general attacks). A formal description of the joint attacks against BB84-like protocols is available in Subsection 7.2.2.

2.3.3 Different Approaches for Security Proofs

We discuss four different approaches for proving unconditional security of QKD protocols:

1. The approach of Mayers [May01] gave the first security proof of a QKD protocol (BB84). This approach proves non-composable security against the most general theoretical attacks.

2. The approach of Biham, Boyer, Boykin, Mor, and Roychowdhury (BBBMR) [BBBMR06] (which follows previous works by similar authors [BM97b, BM97a, BBBGM02]) proves security of BB84 by connecting the *information* Eve obtains and the *disturbance* she induces in the opposite (conjugate) basis (see Subsection 7.2.4). This proof algebraically bounds the trace distance between two possible density matrices held by Eve, and it proves non-composable security against the most general theoretical attacks. In Chapter 7 we adapt this approach to prove *composable* security.

   Security against collective attacks was proved using similar techniques [BBBGM02] that were later improved [BGM09]; in Chapter 6 we make this proof composable.

3. The approach of Shor and Preskill [SP00] proves security of BB84 by analyzing a different, entanglement-based protocol (see Subsection 2.2.2). This protocol uses quantum error correction and entanglement purification, so it requires Alice and Bob to use a quantum computer (unlike BB84); it was earlier proved secure by Lo and Chau [LC99], and then Shor and Preskill [SP00] proved it equivalent to BB84, implying security of BB84. This approach proves non-composable security against the most general theoretical attacks, but later work [BOHLMO05, KRBM07] showed it could be easily modified to prove *composable* security.

4. The approach of Renner [RGK05, Ren08] proves security of various QKD protocols by bounding entropies, min-entropies, and max-entropies appearing in the protocols, and it uses reductions from standard prepare-and-measure QKD protocols to entanglement-based QKD protocols. This approach proves *composable* security against the most general theoretical attacks.

In Chapters 6 and 7 of this thesis, we strengthen BBBMR's security approach [BBBMR06, BGM09] by making it prove *composable* security. This security approach has various advantages and disadvantages compared to other approaches. On the one hand, it is mostly self-contained, while other security approaches require many results from other areas of quantum information (such as various notions of entropy needed for Renner's approach, and entanglement purification and quantum error correction needed for Shor and Preskill's approach); it gives tight finite-key bounds, unlike several other methods (as detailed below); and, at least in some sense, it is simpler than other proof techniques. On the other hand, it is currently limited to BB84-like protocols.

BBBMR's approach gives explicit and tight finite-key bounds. In contrast to this, Shor and Preskill's approach proves only asymptotic security (for infinitely long keys). For Renner's approach, tight finite-key bounds identical to the ones found by BBBMR's approach have been obtained for several protocols, including BB84 [TLGR12]; but at first Renner's approach gave very pessimistic bounds (using de Finetti's theorem [Ren08, Ren07]); later, the bounds were improved for several protocols, including BB84 [SR08]; and finally, tight bounds have been obtained (see [TLGR12] for comparison).

We note that existence of many different proof techniques is important, because some proofs may be more adjustable to various QKD protocols or practical scenarios; some proofs may be clearer to different readers with different backgrounds; analyzing the differences between the proofs and between their obtained results may lead to important insights on the strengths and weaknesses of various techniques; and existence of many proofs makes the security result more certain and less prone to errors.

2.3.4 Robustness Definitions of QKD

A notion much weaker than security is the *robustness* of a QKD protocol [BKM07]. A QKD protocol is *completely robust* if any non-zero information obtained by Eve on the INFO string implies a non-zero probability that Alice and Bob find errors in the TEST bits. In other words, if a protocol is completely robust, Eve cannot obtain any useful information without causing errors that may be noticed by Alice and Bob. Robustness does not imply full security (it does not imply secrecy of Alice and Bob's *final key* after classical post-processing), but it is an important step towards proving security.

To the other extreme, *complete non-robustness* means Eve can get full information without inducing even one error. The two practical attacks described in Subsection 2.5.3 imply their respective protocols to be completely non-robust.

2.4 Semiquantum Key Distribution

Semiquantum key distribution (SQKD) protocols assume either Alice or Bob is a *classical* party [BKM07]. Therefore, these protocols answer a theoretically interesting question: "how quantum" must a QKD protocol be to achieve secure key distribution? We already know *fully classical* key distribution is impossible (see Section 2.1), and we know *fully quantum* key distribution is possible; the existence of SQKD protocols can show that key distribution remains feasible even if one party is classical. Furthermore, SQKD protocols may be easier to implement, so they may also have practical importance.

In SQKD protocols, the classical party can only use "classical" operations—and, in particular, it can prepare states and perform measurements only in the z basis and not in the x basis. For example, consider a practical implementation where $|0\rangle$ and $|1\rangle$ represent two different pulses traveling through the same path one after the other: in this case, the z basis $\{|0\rangle, |1\rangle\}$ represents classical photon pulses, while the x basis $\{|+\rangle, |-\rangle\}$ represents quantum superpositions of classical pulses. The classical party can manipulate classical photon pulses in the z basis, but cannot manipulate superpositions.

The first SQKD protocol was named "QKD with Classical Bob" [BKM07, BGKM09]. In this protocol, in each round, Alice sends to Bob a randomly chosen state of the four BB84 states $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$, but Bob is limited to two "classical" operations:

1. **CTRL:** Bob returns the qubit to Alice undisturbed.
2. **SIFT:** Bob measures the qubit in the z basis and resends to Alice the qubit state he measured.

Bob randomly chooses one of these operations, and Alice measures the returning qubit in the basis she sent it. After N qubits have been sent and received (in N rounds), Alice publicly announces her basis choices for each round, and Bob publicly announces his choices (CTRL or SIFT) for each round. Then, Alice and Bob check the error rates in the CTRL bits and in a random subset of the SIFT bits, aborting if they are too high. Finally, Alice and Bob perform error correction and privacy amplification on the remaining SIFT bits sent by Alice in the z basis, so that they have an identical final key that is completely secret. We note that only SIFT bits are used for generating the final key; CTRL bits are used only for security checks. This protocol was proved completely robust [BKM07] and secure [Kra15b].

Later, [ZQLWL09] suggested a simpler SQKD protocol named "QKD with Classical Alice" (the name is following [BM11]). In this protocol, in each round, Bob sends to Alice the $|+\rangle$ state, and Alice randomly chooses one of the two "classical" operations (CTRL or SIFT) and returns the resulting state to Bob. Bob then measures the received qubit in a randomly chosen basis (z or x), and Alice and Bob proceed almost identically to "QKD with Classical Bob". This protocol was proved completely robust by [BM11], and the proof was extended by [BM10] to include photonic implementations and multi-photon pulses.

Other SQKD protocols have also been suggested, including [LC08, SDL13, YYLH14, Kra15a, ZQZM15]; note that most SQKD protocols are required to be *two-way* protocols (see Subsection 2.2.2) to overcome the limitations of the classical user. Most SQKD protocols have been proven robust, and a few of them also have security analyses [Kra15b, Kra16, ZQM18, Kra18].

2.5 Practical Implementations of QKD Protocols

2.5.1 The Fock Space Notations

Quantum cryptographic protocols are usually implemented with photons. However, standard qubit notations do not describe all possible photon operations and do not properly represent the actual operations of Alice and Bob; as a result, qubit notations do not cover all possible attacks. To correct notations, we must replace the qubit Hilbert space $\mathcal{H}_2 \triangleq \mathrm{Span}\{|\mathtt{0}\rangle, |\mathtt{1}\rangle\}$ by an extended Hilbert space—the "Fock space" $\mathcal{F}$:

- In the simplest case, there are $m \geq 0$ photons, all of them belonging to *one* photonic mode. Here, the Fock state $|m\rangle$ represents m photons in this single mode: $|0\rangle$ is the vacuum state, representing no photons in that mode; $|1\rangle$ represents one photon in that mode; $|2\rangle$ represents two photons in that mode; and so on.

- For describing several different *pulses* of photons (for example, photons traveling through different paths or at different times, or any other external degree of freedom), we need several photonic *modes*. For example, a single photon in one of two pulses (and, thus, in one of two modes) describes *one qubit*, and the z basis states of this qubit are $\{|\mathtt{0}\rangle = |0\rangle \otimes |1\rangle \equiv |0\rangle|1\rangle\ ,\ |\mathtt{1}\rangle = |1\rangle \otimes |0\rangle \equiv |1\rangle|0\rangle\}$. (These two states are mathematically described as tensor products, but we omit the $\otimes$ sign for brevity.) A linear combination describes one photon in a superposition of the two pulses: for example, the x basis states are $\left\{|+\rangle = \frac{|0\rangle|1\rangle+|1\rangle|0\rangle}{\sqrt{2}}\ ,\ |-\rangle = \frac{|0\rangle|1\rangle-|1\rangle|0\rangle}{\sqrt{2}}\right\}$.

- More generally, for describing $m = m_1 + m_0$ photons in two different pulses (two modes), where m_1 photons are in one pulse and m_0 photons are in the other pulse, we write $|m_1\rangle|m_0\rangle$. We add subscripts to specify the type of pulse—for example, $|m_1\rangle_{t_1}|m_0\rangle_{t_0}$ for the two times t_1, t_0, or $|m_1\rangle_{\mathrm{A}}|m_0\rangle_{\mathrm{B}}$ for the two paths A, B.

- For describing more than two pulses (more than two modes), we use generalized notations: for example, $m = m_2 + m_1 + m_0$ photons traveling at times t_2, t_1, t_0 are denoted $|m_2\rangle_{t_2}|m_1\rangle_{t_1}|m_0\rangle_{t_0}$. In particular, the vacuum state (absence of photons) is denoted $|0\rangle$ for one mode, $|0\rangle|0\rangle$ for two modes, $|0\rangle|0\rangle|0\rangle$ for three modes, etc.

All the above notations assume identical photon *polarizations* (which are an *internal* degree of freedom) for all m photons. However, a single photon in a single pulse generally has two orthogonal polarizations: horizontal $\leftrightarrow$ and vertical $\updownarrow$. The two polarizations are described as two modes *for each pulse*, so k pulses mean $2k$ modes.

In this thesis (except Chapters 3–5, for the reasons explained in Section 3.2), polarization modes of $m = m_1 + m_0$ photons are denoted $|m_1, m_0\rangle$ without any subscript, while pulse modes are denoted $|m_1\rangle|m_0\rangle$ with subscripts. Thus:

- If there is exactly one photon in a single pulse, its two polarization modes represent *one qubit*. The z basis states of this qubit are $|\mathtt{0}\rangle = |0, 1\rangle$ (representing one photon in the horizontal polarization mode and zero photons in the vertical polarization mode) and $|\mathtt{1}\rangle = |1, 0\rangle$ (where the single photon is in the vertical mode).

- A linear combination describes one photon in a superposition of those two polarization modes: for example, the x basis states are $\left\{|+\rangle = \frac{|0,1\rangle+|1,0\rangle}{\sqrt{2}}\ ,\ |-\rangle = \frac{|0,1\rangle-|1,0\rangle}{\sqrt{2}}\right\}$.
- The $|m_1, m_0\rangle$ state represents $m = m_1 + m_0$ photons in those two polarization modes: m_1 photons in the vertical mode and m_0 photons in the horizontal mode. In particular, the vacuum state $|0, 0\rangle$ represents an absence of *any* photon.

Formally, for two polarization modes, the entire 2-mode Fock space is:

$$\mathcal{F} \triangleq \mathrm{Span}\{|m_1, m_0\rangle \mid \ m_1 \geq 0\ ,\ m_0 \geq 0\}, \tag{2.3}$$

where the $|m_1, m_0\rangle$ state represents m_1 indistinguishable photons in the $|\mathtt{1}\rangle$ mode and m_0 indistinguishable photons in the $|\mathtt{0}\rangle$ mode.

Similarly, a single photon in the $|+\rangle$ mode may be written as $|0, 1\rangle_{\mathrm{x}}$, and a single photon in the $|-\rangle$ mode may be written as $|1, 0\rangle_{\mathrm{x}}$. The entire 2-mode Fock space can be represented as

$$\mathcal{F} = \mathrm{Span}\{|m_-, m_+\rangle_{\mathrm{x}} \mid \ m_- \geq 0\ ,\ m_+ \geq 0\}, \tag{2.4}$$

where the $|m_-, m_+\rangle_{\mathrm{x}}$ state represents m_- indistinguishable photons in the $|-\rangle$ mode and m_+ indistinguishable photons in the $|+\rangle$ mode.

2.5.2 Experimental Implementations of Polarization-Based QKD

The BB84 protocol may be experimentally implemented in a "polarization-based" implementation, that we can model as follows: the quantum particles sent by Alice to Bob are *single photons* whose polarizations encode the quantum states. The four possible states sent by Alice are $|\mathtt{0}\rangle$, $|\mathtt{1}\rangle$, $|+\rangle$, and $|-\rangle$, where $|\mathtt{0}\rangle = |\leftrightarrow\rangle$ (a single photon in the horizontal polarization) and $|\mathtt{1}\rangle = |\updownarrow\rangle$ (a single photon in the vertical polarization). The $|+\rangle = |{\swarrow\!\!\!\!\!\nearrow}\rangle$ and $|-\rangle = |{\nwarrow\!\!\!\!\!\searrow}\rangle$ states correspond to orthogonal diagonal polarizations.

For measuring the incoming photons, Bob uses a polarizing beam splitter (PBS) and two detectors. Bob actively configures the PBS for choosing his random measurement basis (z or x). If the PBS is configured for measurement in the z basis, it sends any *horizontally* polarized photon to one path and any *vertically* polarized photon to the other path. At the end of each path we place a detector, which clicks whenever it detects a photon. Therefore, the detector at the first path clicks *only* if the $|\mathtt{0}\rangle$ mode is detected, and the detector at the second path clicks *only* if the $|\mathtt{1}\rangle$ mode is detected; a *diagonally* polarized photon ($|+\rangle = |{\swarrow\!\!\!\!\!\nearrow}\rangle$ or $|-\rangle = |{\nwarrow\!\!\!\!\!\searrow}\rangle$) would cause exactly one of the detectors (uniformly random) to click. Similarly, if the PBS is configured for measurement in the x basis, it distinguishes $|+\rangle$ from $|-\rangle$. This implementation (using an "active" basis choice) may be slow, because Bob needs to randomly choose a basis for each arriving photon.

A variant of this implementation uses a "passive" basis choice (e.g., [KZH$^+$02]). This variant uses one polarization-independent beam splitter, two PBSs, and *four* detectors. The polarization-independent beam splitter is placed in the front, and it randomly sends each photon to one path or to another. A photon going to the first path is then measured (as described above) in the z basis, while a photon going to the second path is measured (as described above) in the x basis. We note that in this "passive" variant, the basis is chosen "randomly" by the polarization-independent beam splitter, and Bob does not have to actively choose it; however, it is exposed to the "Fixed Apparatus" attack [BGM14] (see Example 3 of Section 8.4).

The above implementations of QKD are further discussed in Chapter 8.

2.5.3 Practical Attacks

The security promises of QKD are true in theory, but its *practical* security is far from being guaranteed: practical implementations of QKD use realistic photons, so they deviate from the theoretical protocols based on ideal qubits. These deviations make possible various attacks (see [LCT14, SBPCDLP09]), similarly to the idea of "side-channel attacks" in classical computer security.

For example, in the "Photon-Number Splitting" attack [BLMS00] (which assumes the QKD system is implemented using photons, and assumes the quantum state sent by Alice should be a single photon), Eve exploits two facts: (a) in most implementations, Alice sometimes sends to Bob more than one photon (e.g., two photons); and (b) Bob usually cannot count the number of photons he measures. Thus, for any pulse consisting of two or more photons, Eve "steals" one of the photons and keeps it in her quantum memory for a later measurement (after Alice and Bob expose the correct bases), obtaining full information without being noticed; and she blocks all single-photon pulses.

Another example is the "Bright Illumination" practical attack [LWWESM10]: this attack uses a weakness of Bob's measurement devices, allowing Eve to "blind" them and fully control Bob's measurement results (full description is available in Section 8.3). Eve can then get full information on the secret key without inducing any error. An extensive discussion of this attack is available in Chapter 8.

Possible solutions to these problems include: (a) a much more careful analysis of practical devices and practical implementations; (b) "Measurement-Device Independent" QKD protocols [BHM96, Ina02, LCQ12, BP12], which may be secure even if the measurement devices are controlled by Eve; and (c) "Fully Device Independent" QKD protocols [MY98, MAP11, VV14], which may be secure even if all quantum devices are untrusted (under certain assumptions).

2.6 Hoeffding's Theorem

The final stages of our security proofs in Chapters 6 and 7 consist mainly of applications of the following Theorem, proven by Hoeffding in [Hoe63, Section 6]:

Theorem 2.1 (Hoeffding's Theorem)**.** *Let $X_1, \ldots, X_n$ be a random sample without replacement taken from a population $c_1, \ldots, c_N$ such that $a \le c_j \le b$ for all $1 \le j \le N$. (That is, each X_i gets the value of a random c_j, such that the same j is never chosen for two different variables $X_i, X_{i'}$.) If $\overline{X} \triangleq \frac{X_1 + \ldots + X_n}{n}$ and $\mu \triangleq E[\overline{X}]$ is the expected value of $\overline{X}$, then:*

1. *For any $\epsilon > 0$,*
$$\Pr\left[\overline{X} - \mu \ge \epsilon\right] \le e^{-\frac{2n\epsilon^2}{(b-a)^2}}. \tag{2.5}$$

2. *$\mu = \frac{1}{N}\sum_{i=1}^{N} c_i$. Namely, the expected value of $\overline{X}$ is the average value of the population.*

The following Corollary of Hoeffding's theorem is useful for proving security:

Corollary 2.2. *Let us be given an $(n+n')$-bit string $\mathbf{c} = c_1 \ldots c_{n+n'}$, and assume that we randomly and uniformly choose a partition of $\mathbf{c}$ into two substrings, $\mathbf{c}_\mathrm{A}$ of length n and $\mathbf{c}_\mathrm{B}$ of length n'. (Formally, this is a random partition of the index set $\{1, \ldots, n+n'\}$ into two disjoint sets, A and B, satisfying $|A| = n$, $|B| = n'$, and $A \cup B = \{1, \ldots, n+n'\}$.) Then, for any $p > 0$ and $\epsilon > 0$,*
$$\Pr\left[\left(\frac{|\mathbf{C}_\mathrm{A}|}{n} > p + \epsilon\right) \wedge \left(\frac{|\mathbf{C}_\mathrm{B}|}{n'} \le p\right)\right] \le e^{-2\left(\frac{n'}{n+n'}\right)^2 n\epsilon^2}, \tag{2.6}$$
where $\mathbf{C}_\mathrm{A}$ and $\mathbf{C}_\mathrm{B}$ are random variables whose values equal to $\mathbf{c}_\mathrm{A}$ and $\mathbf{c}_\mathrm{B}$, respectively.

Proof. The random and uniform partition of $\mathbf{c}$ into two substrings, $\mathbf{c}_\mathrm{A}$ of length n and $\mathbf{c}_\mathrm{B}$ of length n', is actually a sample of size n without replacement from the population $c_1, \ldots, c_{n+n'} \in \{0, 1\}$. (The sampled n bits are the bits of $\mathbf{c}_\mathrm{A}$, while the other n' bits are the bits of $\mathbf{c}_\mathrm{B}$.) Therefore, we can apply Hoeffding's theorem (Theorem 2.1) to this sampling.

Let $\overline{X}$ be the average of the sample, and let μ be the expected value of $\overline{X}$ (so, according to Theorem 2.1, μ is the average value of the population), then
$$\overline{X} = \frac{|\mathbf{C}_\mathrm{A}|}{n}, \tag{2.7}$$
$$\mu = \frac{|\mathbf{C}_\mathrm{A}| + |\mathbf{C}_\mathrm{B}|}{n+n'}. \tag{2.8}$$
□

Then $\frac{|\mathbf{C}_\mathrm{B}|}{n'} \le p$ is equivalent to $(n+n')\mu - n\overline{X} \le n' \cdot p$, and, therefore, to $n \cdot (\overline{X} - \mu) \ge$

$n' \cdot (\mu - p)$. This means that the conditions $\left(\frac{|\mathbf{C}_\mathrm{A}|}{n} > p + \epsilon\right)$ and $\left(\frac{|\mathbf{C}_\mathrm{B}|}{n'} \le p\right)$ rewrite to

$$\left(\overline{X} - \mu > \epsilon + p - \mu\right) \wedge \left(\frac{n}{n'} \cdot (\overline{X} - \mu) \ge \mu - p\right), \tag{2.9}$$

which implies $\left(1 + \frac{n}{n'}\right)(\overline{X} - \mu) > \epsilon$, which is equivalent to $\overline{X} - \mu > \frac{n'}{n+n'}\epsilon$. Using Hoeffding's theorem (Theorem 2.1), we get

$$\Pr\left[\left(\frac{|\mathbf{C}_\mathrm{A}|}{n} > p + \epsilon\right) \wedge \left(\frac{|\mathbf{C}_\mathrm{B}|}{n'} \le p\right)\right] \le \Pr\left[\overline{X} - \mu > \frac{n'}{n+n'}\epsilon\right] \le e^{-2\left(\frac{n'}{n+n'}\right)^2 n\epsilon^2}. \tag{2.10}$$

Using Corollary 2.2 for comparing the error rates in different sets of qubits (e.g., INFO and TEST bits) is allowed, on the condition that the random and uniform sampling occurs only after the qubits are sent by Alice and measured by Bob. In other words, the sampling cannot affect the bases in which the qubits are sent and measured, and it cannot affect Eve's attack.

Similar uses of Hoeffding's theorem for proving security of QKD are available in [BBBMR06, BGM09].

We also use another Theorem, proven by Hoeffding in [Hoe63, Section 2, Theorem 1]:

Theorem 2.3. *Let $X_1, \ldots, X_N$ be independent random variables with finite first and second moments, such that $0 \le X_i \le 1$ for all $1 \le i \le N$. If $\overline{X} \triangleq \frac{X_1 + \ldots + X_N}{N}$ and $\mu \triangleq E[\overline{X}]$ is the expected value of $\overline{X}$, then for any $\epsilon > 0$,*

$$\Pr\left[\overline{X} - \mu \ge \epsilon\right] \le e^{-2N\epsilon^2}, \tag{2.11}$$

and, in a similar way (see [Hoe63, Section 1]),

$$\Pr\left[\mu - \overline{X} \ge \epsilon\right] \le e^{-2N\epsilon^2}. \tag{2.12}$$

We will use the following Corollary of Theorem 2.3 for proving security of the "efficient BB84" protocol in Subsection 7.3.3:

Corollary 2.4. *Let $0 \le p \le 1$ be a parameter, and let $\mathbf{b} = b_1 \ldots b_N$ be an N-bit string, such that each b_i is chosen probabilistically and independently out of $\{0, 1\}$, with $\Pr(b_i = 0) = p$ and $\Pr(b_i = 1) = 1 - p$. Then:*

$$\Pr\left(|\mathbf{b}| \le \frac{(1-p)N}{2}\right) \le e^{-\frac{1}{2}N(1-p)^2}, \tag{2.13}$$

$$\Pr\left(|\overline{\mathbf{b}}| \le \frac{pN}{2}\right) \le e^{-\frac{1}{2}Np^2}. \tag{2.14}$$

Proof. Let us define $X_i = b_i$ for all $1 \le i \le N$. Then X_i are independent random variables with finite first and second moments, such that $0 \le X_i \le 1$ for all $1 \le i \le N$

and $\mu \triangleq E[\overline{X}] = 1 - p$. Therefore, using Theorem 2.3, we get the two following results:

$$\begin{aligned} \Pr\left[(1-p) - \overline{X} \geq \frac{1-p}{2}\right] &\leq e^{-\frac{1}{2}N(1-p)^2}, & (2.15) \\ \Pr\left[\overline{X} - (1-p) \geq \frac{p}{2}\right] &\leq e^{-\frac{1}{2}Np^2}. & (2.16) \end{aligned}$$

We notice that $\overline{X} = \frac{|\mathbf{b}|}{N} = 1 - \frac{|\overline{\mathbf{b}}|}{N}$. Substituting this result, we get

$$\begin{aligned} \Pr\left[-\frac{|\mathbf{b}|}{N} \geq -\frac{1-p}{2}\right] &\leq e^{-\frac{1}{2}N(1-p)^2}, & (2.17) \\ \Pr\left[1 - \frac{|\overline{\mathbf{b}}|}{N} - 1 \geq -\frac{p}{2}\right] &\leq e^{-\frac{1}{2}Np^2}, & (2.18) \end{aligned}$$

and, therefore,

$$\begin{aligned} \Pr\left[|\mathbf{b}| \leq \frac{(1-p)N}{2}\right] &\leq e^{-\frac{1}{2}N(1-p)^2}, & (2.19) \\ \Pr\left[|\overline{\mathbf{b}}| \leq \frac{pN}{2}\right] &\leq e^{-\frac{1}{2}Np^2}. & (2.20) \end{aligned}$$

□

2.7 Notation for Bit Strings

In this thesis, we denote bit strings (of t bits, where $t \geq 0$ is some integer) by a bold letter (e.g., $\mathbf{i} = i_1 \dots i_t$, where $i_1, \dots, i_t \in \{0,1\}$); and we refer to these bit strings as elements of $\mathbf{F}_2^t$—that is, as elements of a t-dimensional vector space over the field $\mathbf{F}_2 = \{0,1\}$, where addition of two vectors corresponds to a XOR operation between them. The number of 1-bits in a bit string $\mathbf{s}$ is denoted by $|\mathbf{s}|$, and the Hamming distance between two strings $\mathbf{s}$ and $\mathbf{s}'$ is $d_{\mathrm{H}}(\mathbf{s}, \mathbf{s}') \triangleq |\mathbf{s} \oplus \mathbf{s}'|$.

2.8 Structure of this Thesis

First, we discuss a new semiquantum key distribution protocol (the "Mirror protocol") that solves a practical security problem:

- In Chapter 3, we present the Mirror protocol and prove it completely robust.
  This chapter is based on a 2017 paper we published in Physical Review A [BKLM17].
- In Chapter 4, we discuss a simplified variant of the Mirror protocol and present several attacks against it, proving this variant to be non-robust.
  This chapter is based on a 2018 paper we published in Entropy [BLM18].
- In Chapter 5, we prove security of the Mirror protocol against "uniform collective" attacks (defined in Subsection 2.3.2).
  This chapter is based on a 2020 preprint we posted to the arXiv [KLM20].

Then, we discuss composable security of generalized BB84 protocols:

- In Chapter 6, we extend [BBBGM02, BGM09] to prove fully composable security of a variant of BB84 (named "BB84-INFO-z") against collective attacks.
  This chapter is based on a 2020 paper we published in Theoretical Computer Science [BLM20].

- In Chapter 7, we extend [BBBMR06] to prove fully composable security of several variants of BB84 against the most general attacks.

Finally, in Chapter 8, we explain how the practical "Bright Illumination" attack [LWWESM10] can be described as a theoretical "Reversed-Space" attack.
This chapter is based on a 2020 paper we published in the TPNC conference [LM20].

Chapter 3

The Mirror Protocol and Robustness Proof

In this chapter, we present an experimental security problem of the currently existing SQKD protocols. To solve this problem, we suggest a new SQKD protocol (the "Mirror protocol") and prove it completely robust.

This chapter is based on a paper published in Physical Review A in 2017 by Michel Boyer, Matty Katz, Rotem Liss, and Tal Mor [BKLM17].

3.1 Experimental Infeasibility of the SIFT Operation in SQKD Protocols

In the currently existing SQKD protocols (see Section 2.4), one of the "classical" operations is SIFT: measuring in the z basis $\{|0\rangle, |1\rangle\}$ and then resending. In practical (photonic) implementations, and especially if limited to the existing technology, the SIFT operation is very hard to securely implement, because the generated photon will probably be at a different timing or frequency, thus leaking information to the eavesdropper; see details in [TLC09] (which is a comment on [BKM07]) and in the reply [BKM09].

For example, let us look at the "QKD with classical Alice" protocol implemented with two *classical* modes, $|0\rangle$ and $|1\rangle$, describing two pulses (two distinct time-bins) on a single arm. The photon can be either in one pulse, in the other, or in a superposition (a non-classical state). In this case, the SIFT operation requires Alice to measure the two pulses, generate a single photon in a state depending on the measurement outcome, and resend it to Bob; on the other hand, Alice can implement the CTRL operation simply by using a mirror (reflecting both pulses). In this case, it is indeed very difficult for Alice to regenerate the SIFT photon exactly at the right timing, so that it is indistinguishable from a CTRL photon.

Furthermore, in [TLC09] it was shown that even if Alice could (somehow) have the machinery to perform SIFT with perfect timing, Eve would still be able to attack the protocol by taking advantage of the fact that Alice's detectors are imperfect: Eve's attack is modifying the *frequency* of the photon generated by Bob. Alice does not notice the change in frequency. If Alice performs SIFT, the photon she generates is in the original frequency; if she performs CTRL, the photon she reflects is in the frequency modified by Eve. Therefore, if Eve is powerful enough, she can measure the frequency and tell whether Alice used SIFT or CTRL. If Eve finds out that Alice used SIFT, she can copy the bit sent by Alice in the z basis; if she finds out that Alice used CTRL, she shifts the frequency back to the original frequency. (A very similar attack works for other implementations, too—e.g., for polarization-based or phase-based implementations.) This "tagging" attack makes it possible for Eve to get full information on the key without inducing noise.

3.2 The Mirror Protocol

We suggest a new SQKD protocol, similar to "QKD with classical Alice", that is experimentally feasible: in the original protocol of "QKD with classical Alice", Alice could choose only between two operations (CTRL and SIFT); in our new protocol, that we name the "Mirror protocol", Alice may choose between four operations (CTRL, SWAP-10, SWAP-01, and SWAP-ALL). This protocol avoids the need of using the infeasible operation SIFT. The two operations SWAP-10 and SWAP-01 correspond to two possible reflections of *pulses* by using a controllable mirror; these operations cannot be described by qubit notations, so below we use 4-level system notations. Our new protocol is based on the Fock space notations, where the $|m_1, m_0\rangle$ state represents m_1 indistinguishable photons in the $|\mathfrak{1}\rangle$ mode and m_0 indistinguishable photons in the $|\mathfrak{0}\rangle$ mode[1]; more details about the Fock space notations are given in Subsection 2.5.1.

This protocol is experimentally feasible and is safe against the "tagging" attack described in [TLC09]. Moreover, in this chapter we prove the protocol to be completely robust against an attacker Eve that can do anything allowed by the laws of quantum physics, including the possibility of sending multi-photon pulses (namely, assuming Eve may use any quantum state consisting of the two modes $|\mathfrak{0}\rangle$ and $|\mathfrak{1}\rangle$—or, equivalently, any superposition of the Fock states $|m_1, m_0\rangle$). In Chapter 5 we also prove it secure against "uniform collective" attacks. An illustration of the protocol is available as Figure 3.1.

We can describe the new protocol in terms of photon pulses that correspond to two distinct time-bins, and of a controllable mirror operated by Alice: in this case, the CTRL operation corresponds to operating the mirror on both pulses (reflecting both pulses back to the originator, Bob); the SWAP-10 operation corresponds to operating

[1] In the three "Mirror" chapters of this thesis (Chapters 3–5), we use the $|m_1, m_0\rangle$ notation to denote two photon pulses (to make notations simpler in case we analyze two or three subsystems, each consisting of several modes), in contrast to the $|m_1\rangle|m_0\rangle$ notation used for this purpose in Subsection 2.5.1.

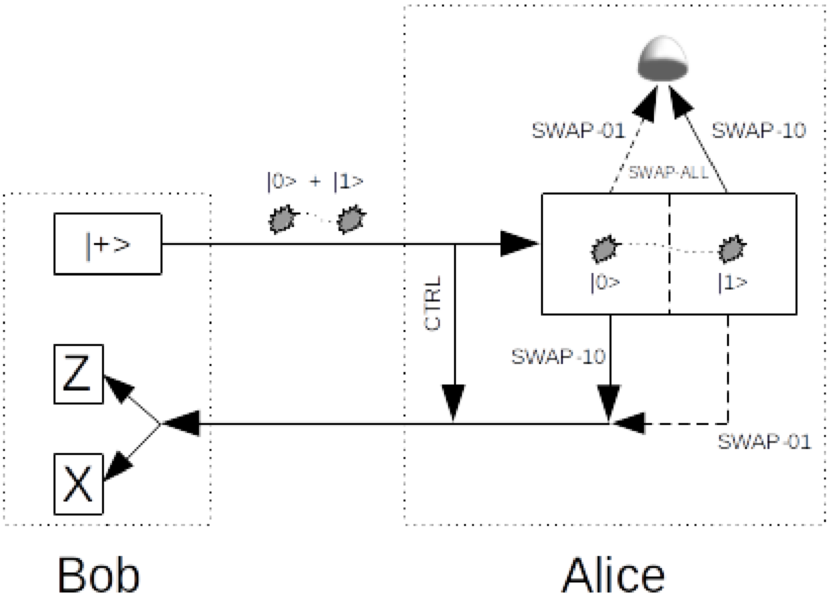

Figure 3.1: **A schematic diagram of the Mirror protocol described in Section 3.2.** This figure was generated by Walter O. Krawec for [KLM20] (Chapter 5).

the mirror only on the $|0\rangle$ pulse while measuring the other pulse (and similarly for the SWAP-01 operation and the $|1\rangle$ pulse); and the SWAP-ALL operation corresponds to measuring all pulses, without reflecting any of them.

For the experimental implementation, we note that a (very slow) mechanically-moved mirror is trivial to implement; a much faster device can be electronically implemented by using standard optical elements (that are commonly used in QKD): a Pockels cell (that can change the polarization of the photon(s) in one of the pulses) and a polarizing beam splitter (that makes it possible to split the two different pulses into two paths, because they are now differently polarized). Like other (fast) QKD experimental settings, implementation is feasible but is not trivial. More details about the experimental implementation of this protocol are available in [Gur13, Tam14].

Let Alice's initial probe be in the vacuum state $|0,0\rangle_{\mathrm{A_{anc}}}$, and let us assume that a single photon is arriving from Bob; thus, the system *as a whole* can be described as a 4-level system (a single photon in four modes). Alice's operations are as follows:

I (CTRL) Do nothing:

$$I|0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}} = |0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}}. \tag{3.1}$$

$\mathbf{S_1}$ (SWAP-10) Swap half of Alice's probe (the left mode) with the $|m_1\rangle_{\mathrm{B}}$ half of Bob's state:

$$S_1|0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}} = |m_1,0\rangle_{\mathrm{A_{anc}}}|0,m_0\rangle_{\mathrm{B}}. \tag{3.2}$$

$\mathbf{S_0}$ (SWAP-01) Swap half of Alice's probe (the right mode) with the $|m_0\rangle_{\mathrm{B}}$ half of

Bob's state:

$$S_0|0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}} = |0,m_0\rangle_{\mathrm{A_{anc}}}|m_1,0\rangle_{\mathrm{B}}. \tag{3.3}$$

S (SWAP-ALL) Swap the entire probe of Alice with the entire state $|m_1,m_0\rangle_{\mathrm{B}}$ of Bob:

$$S|0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}} = |m_1,m_0\rangle_{\mathrm{A_{anc}}}|0,0\rangle_{\mathrm{B}}. \tag{3.4}$$

After each of the three SWAP operations, Alice measures her probe (the $|\cdot\rangle_{\mathrm{A_{anc}}}$ state) in the z basis and sends to Bob the $|\cdot\rangle_{\mathrm{B}}$ state. If there is no noise and no eavesdropping, and if we analyze the "ideal case" (in which exactly one photon is arriving from Bob to Alice), then each round is described by the four-dimensional Hilbert space

$$\mathrm{Span}\{|0,0\rangle_{\mathrm{A_{anc}}}|0,1\rangle_{\mathrm{B}}\ ,\ |0,0\rangle_{\mathrm{A_{anc}}}|1,0\rangle_{\mathrm{B}}\ ,\ |0,1\rangle_{\mathrm{A_{anc}}}|0,0\rangle_{\mathrm{B}}\ ,\ |1,0\rangle_{\mathrm{A_{anc}}}|0,0\rangle_{\mathrm{B}}\}, \tag{3.5}$$

namely, by a four-level system; for our protocol, we use this four-level system instead of the qubit system used by BB84 and by many other QKD schemes. In the most general "theoretical attack" (the attack analyzed by standard QKD security proofs), Eve attacks Alice's and Bob's states using any probe of her choice, but she cannot modify the four-dimensional Hilbert space of the protocol: she can only use these four levels. However, in practical attacks (as analyzed in our robustness analysis), Eve may use an extended Hilbert space (the entire Fock space).

While Eve is fully powerful, it is common to assume that Alice and Bob are limited to use only current technology. In particular, Alice and Bob are limited in the sense that they cannot *count* the number of photons in each mode, but can only check whether a detector corresponding to a specific mode clicks (detects at least one photon in this mode) or not (detects an empty mode). For our protocol to be practical (and for our robustness analysis to be stronger), we assume Alice and Bob are indeed limited in that sense: therefore, when Alice and Bob measure in the z basis, their measurement results are denoted as $\hat{k}_1\hat{k}_0$, where $\hat{k}_0, \hat{k}_1 \in \{0,1\}$. Similarly, when Bob measures in the x basis, his measurement result is $\hat{k}_-\hat{k}_+$, where $\hat{k}_+, \hat{k}_- \in \{0,1\}$.

This limitation leads to the definition of "sum", as follows: let us look at a measurement result of Alice or Bob (that is 00, 01, 10, or 11). The "*sum*" of this measurement result is the number of distinct modes detected to be non-empty during the measurement (namely, the sum of digits in the measurement result). This definition is summarized in Table 3.1.

The protocol consists of the following steps:

1. In each of the N rounds, Bob sends to Alice the $|+\rangle_{\mathrm{B}}$ state; Alice randomly chooses one of her four classical operations (CTRL, SWAP-10, SWAP-01, or SWAP-ALL) and sends the result back to Bob; and Bob measures the state he receives, choosing randomly whether to measure in the z basis or the x basis.

2. Alice reveals her operation choices (CTRL, SWAP-x ($x \in \{01, 10\}$), or SWAP-ALL;

Table 3.1: The four possible measurement results by Alice or Bob (measuring in the z basis), depending on the state obtained by him or her (that is represented in the Fock space notations).

Obtained State	**Measurement Result**	**“Sum”**
$\|0,0\rangle$	00	0
$\|0,m_0\rangle$ $(m_0 \geq 1)$	01	1
$\|m_1,0\rangle$ $(m_1 \geq 1)$	10	1
$\|m_1,m_0\rangle$ $(m_1 \geq 1\ ,\ m_0 \geq 1)$	11	2

Table 3.2: Interpretations of Bob’s measurement results for CTRL states.

Bob’s Result	**Interpretation**
00	a loss
01 (i.e., $\|+\rangle$)	a legal result
10 (i.e., $\|-\rangle$)	an error
11	an error

Alice *does not* reveal her choices between SWAP-10 and SWAP-01, that she keeps as a secret bit string), and Bob reveals his basis choices. They discard all CTRL bits Bob measured in the z basis and all SWAP-x bits he measured in the x basis.

3. For each of the SWAP-x and SWAP-ALL states, Alice and Bob reveal the “sums” of their measurement results.

4. Alice and Bob interpret their measurement results: they consider several types of measurement results as errors, losses, or valid results. See Tables 3.2–3.4 for the details.

5. For all SWAP-x ($x \in \{01, 10\}$) rounds, if Bob’s “sum” is 1 and Alice’s “sum” is 0, then Alice and Bob share a (secret) bit b, because Alice knows (in secret) what operation S_{1-b} she performed, and Bob knows (in secret) what mode $|b\rangle$ he detected. Each one of Alice and Bob keeps this sequence of bits b as his or her secret bit string.

6. Alice and Bob reveal some random subsets of their bit strings, compare them, and estimate the error rate (this is the error rate on the way from Alice back to Bob). They abort the protocol if the error rate in these bits, or any of the error rates measured in Step 4, is above a specified threshold. They discard the revealed bits.

7. Alice and Bob perform error correction and privacy amplification processes on the remaining bit string, yielding a final key that is identical for Alice and Bob and is fully secure from any eavesdropper.

Table 3.3: Interpretations of Alice's and Bob's measurement results for SWAP-x states.

Alice's "Sum"	Bob's "Sum"	Interpretation
0	0	a loss
0	1	Alice and Bob share a bit
1	0	Alice and Bob do not share a bit
1	1	an error
0 or 1	2	an error
2		impossible

Table 3.4: Interpretations of Alice's and Bob's measurement results for SWAP-ALL states.

Alice's Result	Bob's Result	Interpretation
00	00	a loss
01 or 10	00	a legal result
11	00	an error
any	01, 10, or 11	an error

Notice that Bob does not have a special role in the beginning: he always generates the same state, $|+\rangle$. It is even possible that the adversary Eve generates this state instead of him.

3.3 Robustness Analysis

To prove robustness, we will prove that for Eve's attack to be undetectable by Alice and Bob (namely, for Eve's attack not to cause any errors), it must not give Eve any information.

Eve's attack on a state can be performed in both directions: from the source (Bob) to Alice, Eve applies U; from Alice back to Bob, Eve applies V. We may assume, without limiting generality, that Eve uses a fixed probe space $\mathcal{H}_{\mathrm{E}}$ for her attacks.

According to the definition of robustness, we will prove that if, during a run of the protocol, no error can be detected by Alice and Bob, then Eve gets no information on the raw key. According to Tables 3.2–3.4, if Alice and Bob cannot find any error, the following conditions must be true for all measurement results that were not discarded due to basis mismatch:

1. For all CTRL rounds, Bob's measurement result (in the x basis) must not be 10 or 11: namely, Bob must never detect any photon in the $|-\rangle$ mode.
2. For all SWAP-x rounds, Alice's "sum" and Bob's "sum" (in the z basis) must not be both 1.
3. For all SWAP-x rounds, Bob's "sum" (in the z basis) must not be 2: namely, Bob's measurement result must not be 11.

4. For all SWAP-x rounds, no error (that may be detected during the protocol) can exist. In other words:
   (a) For all SWAP-10 rounds, Bob's measurement result (in the z basis) must not be 10.
   (b) For all SWAP-01 rounds, Bob's measurement result (in the z basis) must not be 01.
5. For all SWAP-ALL rounds, Alice's measurement result must not be 11.
6. For all SWAP-ALL rounds, Bob's measurement result must not be 01, 10, or 11.

We now analyze each round of the protocol. After the round begins, the source (Bob) sends to Alice the state $|0,1\rangle_{\mathrm{x,B}} \in \mathcal{H}_{\mathrm{B}}$. Eve can now interfere: she attaches her own probe state (in the Hilbert space $\mathcal{H}_{\mathrm{E}}$) and applies the unitary transformation U. The resulting Bob+Eve state (including Eve's probe) is of the form

$$|\psi_{\mathrm{init}}\rangle \triangleq \sum_{\substack{m_1\geq 0\\ m_0\geq 0}} |m_1, m_0\rangle_{\mathrm{B}} |E_{m_1,m_0}\rangle_{\mathrm{E}}, \tag{3.6}$$

where $|E_{i,j}\rangle_{\mathrm{E}}$ are non-normalized vectors in $\mathcal{H}_{\mathrm{E}}$.

Condition 5 means that $|E_{m_1,m_0}\rangle_{\mathrm{E}} = 0$ for all m_1, m_0 satisfying $m_1 \geq 1$ and $m_0 \geq 1$. Therefore,

$$|\psi_{\mathrm{init}}\rangle = |\phi_{1,0}\rangle + |\phi_{0,1}\rangle + |\phi_{0,0}\rangle, \tag{3.7}$$

where

$$|\phi_{1,0}\rangle \triangleq \sum_{m_1\geq 1} |m_1, 0\rangle_{\mathrm{B}} |E_{m_1,0}\rangle_{\mathrm{E}}, \tag{3.8}$$

$$|\phi_{0,1}\rangle \triangleq \sum_{m_0\geq 1} |0, m_0\rangle_{\mathrm{B}} |E_{0,m_0}\rangle_{\mathrm{E}}, \tag{3.9}$$

$$|\phi_{0,0}\rangle \triangleq |0, 0\rangle_{\mathrm{B}} |E_{0,0}\rangle_{\mathrm{E}}. \tag{3.10}$$

Alice now applies one of the four possible operations (CTRL = I, SWAP-10 = S_1, SWAP-01 = S_0, or SWAP-ALL = S) and destructively measures her probe state. The (non-normalized) state of the Bob+Eve system after Alice's operation and measurement is written in Table 3.5.

Then, Eve applies a second unitary transformation V on the state sent from Alice to Bob (and on her own probe state). According to conditions 2, 3, and 6, the density matrices $V\rho^{(1)}_{\text{S-10}}V^\dagger$, $V\rho^{(1)}_{\text{S-01}}V^\dagger$, and $V\rho_{\text{S-ALL}}V^\dagger$ must only overlap with $|0,0\rangle_{\mathrm{B}}$. It follows that there exists $|H_{0,0}\rangle_{\mathrm{E}} \in \mathcal{H}_{\mathrm{E}}$ such that

$$V|\phi_{0,0}\rangle = |0,0\rangle_{\mathrm{B}} |H_{0,0}\rangle_{\mathrm{E}}. \tag{3.11}$$

Let us denote $V|\phi_{1,0}\rangle = \sum_{\substack{m_1\geq 0\\ m_0\geq 0}} |m_1, m_0\rangle_{\mathrm{B}} |F_{m_1,m_0}\rangle_{\mathrm{E}}$. Let us look at a SWAP-01

Table 3.5: The (non-normalized) state of the Bob+Eve system after Alice's operation, given Alice's "sum". Note that $|\phi_{1,0}\rangle$, $|\phi_{0,1}\rangle$, and $|\phi_{0,0}\rangle$ are defined in Equations (3.8)–(3.10).

Alice's Operation	**Alice's "Sum"**	**Bob+Eve State**
CTRL		$|\psi_{\text{CTRL}}\rangle \triangleq |\phi_{1,0}\rangle + |\phi_{0,1}\rangle + |\phi_{0,0}\rangle$
SWAP-10	0	$|\psi^{(0)}_{\text{S-10}}\rangle \triangleq |\phi_{0,1}\rangle + |\phi_{0,0}\rangle$
SWAP-01	0	$|\psi^{(0)}_{\text{S-01}}\rangle \triangleq |\phi_{1,0}\rangle + |\phi_{0,0}\rangle$
SWAP-10	1	$\rho^{(1)}_{\text{S-10}} \triangleq \sum_{m_1\geq 1} |0,0\rangle_{\text{B}}\langle 0,0|_{\text{B}} \otimes |E_{m_1,0}\rangle_{\text{E}}\langle E_{m_1,0}|_{\text{E}}$
SWAP-01	1	$\rho^{(1)}_{\text{S-01}} \triangleq \sum_{m_0\geq 1} |0,0\rangle_{\text{B}}\langle 0,0|_{\text{B}} \otimes |E_{0,m_0}\rangle_{\text{E}}\langle E_{0,m_0}|_{\text{E}}$
SWAP-ALL		$\rho_{\text{S-ALL}} \triangleq \rho^{(1)}_{\text{S-10}} + \rho^{(1)}_{\text{S-01}} + |\phi_{0,0}\rangle\langle\phi_{0,0}|$

round for which Alice's "sum" is 0: in this round, the state of Bob+Eve after Eve's attack is

$$\begin{aligned} V|\psi^{(0)}_{\text{S-01}}\rangle &= V|\phi_{1,0}\rangle + V|\phi_{0,0}\rangle \\ &= \sum_{\substack{m_1\geq 0\\ m_0\geq 0}} |m_1,m_0\rangle_{\text{B}}|F_{m_1,m_0}\rangle_{\text{E}} + |0,0\rangle_{\text{B}}|H_{0,0}\rangle_{\text{E}}, \end{aligned} \tag{3.12}$$

and following conditions 4b and 3, Bob must not detect a photon in the $|\mathfrak{o}\rangle$ mode (otherwise, the error may be detected during the protocol). Therefore, $|F_{m_1,m_0}\rangle_{\text{E}} = 0$ for all $m_0 \geq 1$. It follows that

$$V|\phi_{1,0}\rangle = \sum_{m_1\geq 1} |m_1,0\rangle_{\text{B}}|F_{m_1,0}\rangle_{\text{E}} + |0,0\rangle_{\text{B}}|F_{0,0}\rangle_{\text{E}}. \tag{3.13}$$

Similarly (following conditions 4a and 3),

$$V|\phi_{0,1}\rangle = \sum_{m_0\geq 1} |0,m_0\rangle_{\text{B}}|G_{0,m_0}\rangle_{\text{E}} + |0,0\rangle_{\text{B}}|G_{0,0}\rangle_{\text{E}}. \tag{3.14}$$

Now, Equations (3.11), (3.13), and (3.14) imply that if Alice applies CTRL, the state of Bob+Eve after Eve's attack is

$$\begin{aligned} V|\psi_{\text{CTRL}}\rangle &= V|\phi_{1,0}\rangle + V|\phi_{0,1}\rangle + V|\phi_{0,0}\rangle \\ &= \sum_{m\geq 1} \left[|m,0\rangle_{\text{B}}|F_{m,0}\rangle_{\text{E}} + |0,m\rangle_{\text{B}}|G_{0,m}\rangle_{\text{E}}\right] + |0,0\rangle_{\text{B}}|H\rangle_{\text{E}}, \end{aligned} \tag{3.15}$$

where $|H\rangle_{\text{E}} \triangleq |F_{0,0}\rangle_{\text{E}} + |G_{0,0}\rangle_{\text{E}} + |H_{0,0}\rangle_{\text{E}}$. Following condition 1, the probability of Bob detecting a photon in the $|-\rangle$ mode must be 0.

We now use the following Lemma:

Lemma 3.1. *If* $|\psi'\rangle = \sum_{m\geq 1}\left[|m,0\rangle_{\text{B}}|F_{m,0}\rangle_{\text{E}} + |0,m\rangle_{\text{B}}|G_{0,m}\rangle_{\text{E}}\right] + |0,0\rangle_{\text{B}}|H\rangle_{\text{E}}$ *is a bi-*

partite state in $\mathcal{H}_\mathrm{B} \otimes \mathcal{H}_\mathrm{E}$, and if there is zero probability that Bob detects a photon in the $|-\rangle$ mode, then $|F_{1,0}\rangle_\mathrm{E} = |G_{0,1}\rangle_\mathrm{E}$, and $|F_{m,0}\rangle_\mathrm{E} = |G_{0,m}\rangle_\mathrm{E} = 0$ for all $m \geq 2$.

Proof. If there is zero probability that Bob detects a photon in the $|-\rangle$ mode, then there is zero probability of measuring any basis state $|m_-, m_+\rangle_{\mathrm{x,B}}$ of $\mathcal{H}_\mathrm{B}$ which satisfies $m_- \geq 1$.

For $m = 1$, since $|0,1\rangle_\mathrm{B} = \frac{|0,1\rangle_{\mathrm{x,B}}+|1,0\rangle_{\mathrm{x,B}}}{\sqrt{2}}$ and $|1,0\rangle_\mathrm{B} = \frac{|0,1\rangle_{\mathrm{x,B}}-|1,0\rangle_{\mathrm{x,B}}}{\sqrt{2}}$, we get the following equation:

$$\begin{aligned} |1,0\rangle_\mathrm{B}|F_{1,0}\rangle_\mathrm{E} + |0,1\rangle_\mathrm{B}|G_{0,1}\rangle_\mathrm{E} &= \frac{|0,1\rangle_{\mathrm{x,B}}}{\sqrt{2}}\left[|G_{0,1}\rangle_\mathrm{E} + |F_{1,0}\rangle_\mathrm{E}\right] \\ &+ \frac{|1,0\rangle_{\mathrm{x,B}}}{\sqrt{2}}\left[|G_{0,1}\rangle_\mathrm{E} - |F_{1,0}\rangle_\mathrm{E}\right]. \qquad (3.16) \end{aligned}$$

Since the probability of detecting a photon in the $|-\rangle$ mode must be 0 (and, in particular, the probability of detecting $|1,0\rangle_{\mathrm{x,B}}$ must be 0), it is necessary that $|F_{1,0}\rangle_\mathrm{E} = |G_{0,1}\rangle_\mathrm{E}$.

For $m \geq 2$, using the ladder operators a_0, a_1, a_+, and a_-, since $a_0 = \frac{a_+ + a_-}{\sqrt{2}}$ and $a_1 = \frac{a_+ - a_-}{\sqrt{2}}$, we get

$$\begin{aligned} |0,m\rangle_\mathrm{B} &= \frac{{a_0^\dagger}^m|0,0\rangle_\mathrm{B}}{\sqrt{m!}} = \left(\frac{a_+^\dagger + a_-^\dagger}{\sqrt{2}}\right)^m \frac{|0,0\rangle_\mathrm{B}}{\sqrt{m!}} \\ &= \frac{1}{\sqrt{2^m \cdot m!}}\sum_{k=0}^{m}\binom{m}{k}{a_-^\dagger}^k{a_+^\dagger}^{m-k}|0,0\rangle_\mathrm{B} \qquad (3.17) \end{aligned}$$

and

$$\begin{aligned} |m,0\rangle_\mathrm{B} &= \frac{{a_1^\dagger}^m|0,0\rangle_\mathrm{B}}{\sqrt{m!}} = \left(\frac{a_+^\dagger - a_-^\dagger}{\sqrt{2}}\right)^m \frac{|0,0\rangle_\mathrm{B}}{\sqrt{m!}} \\ &= \frac{1}{\sqrt{2^m \cdot m!}}\sum_{k=0}^{m}\binom{m}{k}(-1)^k{a_-^\dagger}^k{a_+^\dagger}^{m-k}|0,0\rangle_\mathrm{B}. \qquad (3.18) \end{aligned}$$

From Equations (3.17)–(3.18) it follows that

$$\begin{aligned} |m,0\rangle_\mathrm{B}|F_{m,0}\rangle_\mathrm{E} + |0,m\rangle_\mathrm{B}|G_{0,m}\rangle_\mathrm{E} &= |e^{(m)}\rangle_\mathrm{B}\left[|G_{0,m}\rangle_\mathrm{E} + |F_{m,0}\rangle_\mathrm{E}\right] \\ &+ |o^{(m)}\rangle_\mathrm{B}\left[|G_{0,m}\rangle_\mathrm{E} - |F_{m,0}\rangle_\mathrm{E}\right], \qquad (3.19) \end{aligned}$$

where

$$\begin{aligned} |e^{(m)}\rangle_\mathrm{B} &\triangleq \frac{1}{\sqrt{2^m \cdot m!}}\sum_{k \text{ even}}\binom{m}{k}{a_-^\dagger}^k{a_+^\dagger}^{m-k}|0,0\rangle_\mathrm{B}, \qquad (3.20) \\ |o^{(m)}\rangle_\mathrm{B} &\triangleq \frac{1}{\sqrt{2^m \cdot m!}}\sum_{k \text{ odd}}\binom{m}{k}{a_-^\dagger}^k{a_+^\dagger}^{m-k}|0,0\rangle_\mathrm{B}. \qquad (3.21) \end{aligned}$$

We notice that ${a_-^\dagger}^k{a_+^\dagger}^{m-k}|0,0\rangle_\mathrm{B}$ is, up to a constant factor, the Fock state $|k, m-k\rangle_{\mathrm{x,B}}$.

Because the probability of finding a photon in the $|-\rangle$ mode must be zero, it means that the coefficient of ${a_-^\dagger}^k {a_+^\dagger}^{m-k} |0,0\rangle_\mathrm{B}$ must be zero for all $k \geq 1$.

Substituting $|e^{(m)}\rangle_\mathrm{B}$ and $|o^{(m)}\rangle_\mathrm{B}$ by their values in Equation (3.19), the coefficient of ${a_-^\dagger}^k {a_+^\dagger}^{m-k} |0,0\rangle_\mathrm{B}$ (up to a non-zero constant factor) is $|G_{0,m}\rangle_\mathrm{E} + |F_{m,0}\rangle_\mathrm{E}$ for even values of k and $|G_{0,m}\rangle_\mathrm{E} - |F_{m,0}\rangle_\mathrm{E}$ for odd values of k. Since $k = m \geq 1$ and $k = m-1 \geq 1$ have different parities, this implies both $|G_{0,m}\rangle_\mathrm{E} + |F_{m,0}\rangle_\mathrm{E}$ and $|G_{0,m}\rangle_\mathrm{E} - |F_{m,0}\rangle_\mathrm{E}$ must be 0, and thus $|F_{m,0}\rangle_\mathrm{E} = |G_{0,m}\rangle_\mathrm{E} = 0$. □

Applying Lemma 3.1, we deduce that $|F_{m,0}\rangle_\mathrm{E} = |G_{0,m}\rangle_\mathrm{E} = 0$ for all $m \geq 2$, and that $|F_{1,0}\rangle_\mathrm{E} = |G_{0,1}\rangle_\mathrm{E} \triangleq |F\rangle_\mathrm{E}$.

It follows that the joint states of Bob+Eve after Eve's attack, when Alice performed SWAP-x and her "sum" was 0 (these are the only rounds in which Alice and Bob may share a secret bit), are: (using Table 3.5 and Equations (3.11), (3.13), and (3.14))

$$V|\psi^{(0)}_{\text{S-10}}\rangle = V|\phi_{0,1}\rangle + V|\phi_{0,0}\rangle = |0,1\rangle_\mathrm{B}|F\rangle_\mathrm{E} + |0,0\rangle_\mathrm{B}\left[|G_{0,0}\rangle_\mathrm{E} + |H_{0,0}\rangle_\mathrm{E}\right], \quad (3.22)$$

$$V|\psi^{(0)}_{\text{S-01}}\rangle = V|\phi_{1,0}\rangle + V|\phi_{0,0}\rangle = |1,0\rangle_\mathrm{B}|F\rangle_\mathrm{E} + |0,0\rangle_\mathrm{B}\left[|F_{0,0}\rangle_\mathrm{E} + |H_{0,0}\rangle_\mathrm{E}\right]. \quad (3.23)$$

Therefore, the state of Eve's probe is independent of all Alice's and Bob's shared bits, and is equal to $|F\rangle_\mathrm{E}$ whenever Alice and Bob share a bit. Eve can thus get no information on the bits shared by Alice and Bob without causing errors that may be noticed by Alice and Bob.

3.4 Conclusion

In this chapter, we have suggested a solution for a practical security problem of SQKD protocols, that was discussed in Section 3.1 and [TLC09]: we have presented a new semiquantum key distribution protocol and proved it robust (see Chapter 5 for full security analysis against "uniform collective" attacks). Unlike all previous SQKD protocols, our new protocol can be experimentally implemented in a secure way.

Chapter 4

Attacks Against a Simplified Variant of the Mirror Protocol

In this chapter, we present a simpler variant of the Mirror protocol (the "simplified Mirror protocol") which is easier to implement. Our variant allows the classical party, Alice, to choose one of three operations, while the Mirror protocol allows her to choose one of four operations. We then present two attacks against this variant, proving it non-robust. Our results show the four classical operations allowed by the Mirror protocol are probably necessary for robustness.

This chapter is based on a paper published in Entropy in 2018 by Michel Boyer, Rotem Liss, and Tal Mor [BLM18].

4.1 The Simplified Mirror Protocol

The simplified Mirror protocol we present in this chapter is identical to the Mirror protocol described in Section 3.2, except that it does not include the SWAP-ALL operation. In other words, in the simplified protocol, Alice chooses at random one of the three classical operations CTRL, SWAP-10, and SWAP-01.

The simplified protocol is easier to implement, because the SWAP-ALL operation poses some experimental challenges to the electronic implementation discussed in Section 3.2: for implementing SWAP-ALL, the Pockels cell should either remain working for a long time (changing polarization for both pulses) or be operated twice (changing polarization for each pulse separately). In more details, for the two pulses representing the $|0\rangle$ mode and the $|1\rangle$ mode: if we assume the duration of each pulse is t and the time difference between the two pulses is T (where $t \ll T$), the first solution means keeping the Pockels cell operating during the time period $[0, T + 2t]$, and the second solution means operating the Pockels cell during the two time periods $[0, t]$ and $[T + t, T + 2t]$. The first solution may be problematic for some models of the Pockels cell, and the second solution may be problematic because of the recovery time needed for the Pockels cell. Therefore, at least in some implementations, the simplified Mirror protocol is much easier to implement than the standard Mirror protocol.

Moreover, analyzing the simplified protocol gives a better understanding of the properties required for an SQKD protocol to be robust. In particular, this analysis explains why the structure and complexity of the Mirror protocol are necessary for robustness.

For completeness, we provide below the full description of the simplified Mirror protocol. We emphasize, however, that this protocol is almost identical to the Mirror protocol described in Section 3.2, and the *only* difference is removing the SWAP-ALL operation.

In the simplified Mirror protocol, in each round, Bob sends to Alice the initial state $|+\rangle_{\mathrm{B}}$, which is equivalent to $|0,1\rangle_{\mathrm{x,B}} \triangleq \frac{|0,1\rangle_{\mathrm{B}}+|1,0\rangle_{\mathrm{B}}}{\sqrt{2}}$. Then, Alice prepares an ancillary state in the initial vacuum state $|0,0\rangle_{\mathrm{A_{anc}}}$ and chooses *at random* one of the following three classical operations (defined on any Fock state she may possibly get, due to Eve's attack):

I (CTRL) Reflect all photons towards Bob, without measuring any photon. The mathematical description is:

$$I|0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}} = |0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}}. \tag{4.1}$$

$\mathbf{S_1}$ (SWAP-10) Reflect all photons in the $|\mathtt{0}\rangle$ mode towards Bob, and measure all photons in the $|\mathtt{1}\rangle$ mode. The mathematical description is:

$$S_1|0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}} = |m_1,0\rangle_{\mathrm{A_{anc}}}|0,m_0\rangle_{\mathrm{B}}. \tag{4.2}$$

$\mathbf{S_0}$ (SWAP-01) Reflect all photons in the $|\mathtt{1}\rangle$ mode towards Bob, and measure all photons in the $|\mathtt{0}\rangle$ mode. The mathematical description is:

$$S_0|0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}} = |0,m_0\rangle_{\mathrm{A_{anc}}}|m_1,0\rangle_{\mathrm{B}}. \tag{4.3}$$

We note that in the above mathematical description, Alice measures her ancillary state $|\cdot\rangle_{\mathrm{A_{anc}}}$ in the z basis and sends back to Bob the $|\cdot\rangle_{\mathrm{B}}$ state. The states sent from Alice to Bob (without any error, loss, or eavesdropping) are detailed in Table 4.1. Then, Bob measures the incoming state in a random basis (either the z basis or the x basis).

After completing all rounds, Alice sends over the classical channel her operation choices (CTRL or SWAP-x; she keeps $x \in \{01, 10\}$ in secret), Bob sends over the classical channel his basis choices, and both of them reveal some non-secret information on their measurement results (as elaborated in Section 3.2). Then, Alice and Bob reveal and compute the error rate on test bits for which Alice used SWAP-10 or SWAP-01 and Bob measured in the z basis, and the error rate on test bits for which Alice used CTRL and Bob measured in the x basis. They also check whether other errors exist (for example, it must never happen that *both* Alice and Bob detect a photon). Alice

Table 4.1: The state sent from Alice to Bob in the simplified Mirror protocol without errors or losses, depending on Alice's classical operation and on whether Alice detected a photon or not.

Alice's Operation	**Did Alice Detect a Photon?**	**State Sent from Alice to Bob**
CTRL	no (happens with certainty)	$\lvert 0,1\rangle_{\mathrm{x,B}} = \frac{1}{\sqrt{2}}\left[\lvert 0,1\rangle_{\mathrm{B}} + \lvert 1,0\rangle_{\mathrm{B}}\right]$
SWAP-10	no (happens with probability $\frac{1}{2}$)	$\lvert 0,1\rangle_{\mathrm{B}}$
SWAP-10	yes (happens with probability $\frac{1}{2}$)	$\lvert 0,0\rangle_{\mathrm{B}}$
SWAP-01	no (happens with probability $\frac{1}{2}$)	$\lvert 1,0\rangle_{\mathrm{B}}$
SWAP-01	yes (happens with probability $\frac{1}{2}$)	$\lvert 0,0\rangle_{\mathrm{B}}$

and Bob also discard mismatched rounds, such as rounds in which Alice used SWAP-10 and Bob used the x basis.

In the non-testing rounds, Alice and Bob share the secret bit 0 if Alice uses SWAP-10 and detects no photon while Bob measures in the z basis and detects a photon in the $\lvert\mathtt{0}\rangle$ mode; similarly, they share the secret bit 1 if Alice uses SWAP-01 and detects no photon while Bob measures in the z basis and detects a photon in the $\lvert\mathtt{1}\rangle$ mode.

Finally, Alice and Bob verify that the error rates are below some thresholds, and they perform error correction and privacy amplification in the standard way for QKD protocols. At the end of the protocol, Alice and Bob hold an identical final key that is supposed to be completely secure against any eavesdropper.

4.2 Attacks Against the Simplified Mirror Protocol

We prove the simplified protocol to be non-robust by presenting two attacks: a "full attack" described in Subsection 4.2.1, which gives Eve full information but causes full loss of the CTRL bits, and a "weaker attack" described in Subsection 4.2.2, which gives Eve less information but causes fewer losses of CTRL bits.

4.2.1 A Full Attack on the Simplified Protocol

In this attack, Eve gets full information of all secret bits: namely, she gets full information on the SWAP-10 and SWAP-01 bits that were measured by Bob in the z basis.

Eve applies her attack in two stages: the first stage is on the way from Bob to Alice, and the second stage is on the way from Alice to Bob. In both stages she uses her own probe space (namely, ancillary space) $\mathcal{H}_{\mathrm{E}} = \mathcal{H}_3$ spanned by the orthonormal basis $\{\lvert 0\rangle_{\mathrm{E}}, \lvert 1\rangle_{\mathrm{E}}, \lvert 2\rangle_{\mathrm{E}}\}$. We assume that Eve fully controls the environment, the errors, and the losses (this is a standard assumption when analyzing the security of QKD): namely, no losses and no errors exist between Bob and Eve or between Alice and Eve.

In the first stage of the attack (on the way from Bob to Alice), Eve intercepts the

Table 4.2: The state of Bob+Eve after Alice's classical operation for the attacks described in Subsections 4.2.1 and 4.2.2, depending on Alice's classical operation and on whether Alice detected a photon or not.

Alice's Operation	**Did Alice Detect a Photon?**	**Bob+Eve State**
CTRL	no (happens with certainty)	$\frac{1}{\sqrt{3}}\left[\|0,1\rangle_{\mathrm{B}}\|1\rangle_{\mathrm{E}}+\|1,0\rangle_{\mathrm{B}}\|1\rangle_{\mathrm{E}}\right.$ $\left.+\|0,0\rangle_{\mathrm{B}}\|0\rangle_{\mathrm{E}}\right]$
SWAP-10	no (happens with probability $\frac{2}{3}$)	$\frac{1}{\sqrt{2}}\left[\|0,1\rangle_{\mathrm{B}}\|1\rangle_{\mathrm{E}}+\|0,0\rangle_{\mathrm{B}}\|0\rangle_{\mathrm{E}}\right]$
SWAP-10	yes (happens with probability $\frac{1}{3}$)	$\|0,0\rangle_{\mathrm{B}}\|1\rangle_{\mathrm{E}}$
SWAP-01	no (happens with probability $\frac{2}{3}$)	$\frac{1}{\sqrt{2}}\left[\|1,0\rangle_{\mathrm{B}}\|1\rangle_{\mathrm{E}}+\|0,0\rangle_{\mathrm{B}}\|0\rangle_{\mathrm{E}}\right]$
SWAP-01	yes (happens with probability $\frac{1}{3}$)	$\|0,0\rangle_{\mathrm{B}}\|1\rangle_{\mathrm{E}}$

state $|+\rangle_{\mathrm{B}} = |0,1\rangle_{\mathrm{x,B}}$ sent by Bob, generates instead the state

$$\frac{1}{\sqrt{3}}\left[|0,1\rangle_{\mathrm{B}}|1\rangle_{\mathrm{E}}+|1,0\rangle_{\mathrm{B}}|1\rangle_{\mathrm{E}}+|0,0\rangle_{\mathrm{B}}|0\rangle_{\mathrm{E}}\right]=\sqrt{\frac{2}{3}}|0,1\rangle_{\mathrm{x,B}}|1\rangle_{\mathrm{E}}+\sqrt{\frac{1}{3}}|0,0\rangle_{\mathrm{B}}|0\rangle_{\mathrm{E}}, \quad (4.4)$$

and sends to Alice the B part of the state. This state causes Alice to get no photons with probability $\frac{1}{3}$ and get the expected $|+\rangle_{\mathrm{B}}$ state with probability $\frac{2}{3}$. Alice then performs at random one of the three classical operations CTRL, SWAP-10, or SWAP-01; the resulting possible states of Bob+Eve are described in Table 4.2.

In the second stage of the attack (on the way from Alice to Bob), Eve applies the unitary operator V on the joint Bob+Eve state, where V is defined as follows:

$$\begin{aligned}
V|0,1\rangle_{\mathrm{B}}|1\rangle_{\mathrm{E}} &= -\sqrt{\frac{1}{3}}|1,0\rangle_{\mathrm{B}}|1\rangle_{\mathrm{E}}+\sqrt{\frac{2}{3}}|0,0\rangle_{\mathrm{B}}|0\rangle_{\mathrm{E}}, && (4.5)\\
V|1,0\rangle_{\mathrm{B}}|1\rangle_{\mathrm{E}} &= -\sqrt{\frac{1}{3}}|0,1\rangle_{\mathrm{B}}|0\rangle_{\mathrm{E}}+\sqrt{\frac{2}{3}}|0,0\rangle_{\mathrm{B}}|1\rangle_{\mathrm{E}}, && (4.6)\\
V|0,0\rangle_{\mathrm{B}}|0\rangle_{\mathrm{E}} &= \sqrt{\frac{1}{3}}|0,1\rangle_{\mathrm{B}}|0\rangle_{\mathrm{E}}+\sqrt{\frac{1}{3}}|1,0\rangle_{\mathrm{B}}|1\rangle_{\mathrm{E}}+\sqrt{\frac{1}{3}}|0,0\rangle_{\mathrm{B}}|+\rangle_{\mathrm{E}}, && (4.7)\\
V|0,0\rangle_{\mathrm{B}}|1\rangle_{\mathrm{E}} &= |0,0\rangle_{\mathrm{B}}|2\rangle_{\mathrm{E}}. && (4.8)
\end{aligned}$$

V is indeed a unitary operator, because we can prove the right-hand sides to be orthonormal: all right-hand sides are normalized vectors; the first two vectors are clearly orthogonal; the third vector is orthogonal to the first two, because $\langle 0|+\rangle_{\mathrm{E}} = \langle 1|+\rangle_{\mathrm{E}} = \frac{1}{\sqrt{2}}$; and the fourth vector is orthogonal to the three others. Thus, V defines (or, more precisely, can be extended to) a unitary operator on $\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{E}}$.

Applying the unitary operator V to Table 4.2 gives the states listed in Table 4.3. Comparing it with Table 4.1, we conclude that this attack never causes Alice and Bob to detect an error. Moreover, Eve detects the entire secret key: Eve measures "0" in her probe if Alice and Bob agree on the "secret" bit 0, and she measures "1" in her probe if Alice and Bob agree on the "secret" bit 1. However, Eve causes several kinds of losses; in particular, all CTRL bits are lost.

Table 4.3: The state of Bob+Eve after completing Eve's attack described in Subsection 4.2.1, depending on Alice's classical operation and on whether Alice detected a photon or not.

Alice's Operation	**Did Alice Detect a Photon?**	**Bob+Eve State**
CTRL	no (happens with certainty)	$\lvert 0,0\rangle_{\mathrm{B}}\lvert +\rangle_{\mathrm{E}}$
SWAP-10	no (happens with probability $\frac{2}{3}$)	$\frac{1}{\sqrt{6}}\lvert 0,1\rangle_{\mathrm{B}}\lvert 0\rangle_{\mathrm{E}} + \lvert 0,0\rangle_{\mathrm{B}}\frac{3\lvert 0\rangle_{\mathrm{E}} + \lvert 1\rangle_{\mathrm{E}}}{\sqrt{12}}$
SWAP-10	yes (happens with probability $\frac{1}{3}$)	$\lvert 0,0\rangle_{\mathrm{B}}\lvert 2\rangle_{\mathrm{E}}$
SWAP-01	no (happens with probability $\frac{2}{3}$)	$\frac{1}{\sqrt{6}}\lvert 1,0\rangle_{\mathrm{B}}\lvert 1\rangle_{\mathrm{E}} + \lvert 0,0\rangle_{\mathrm{B}}\frac{\lvert 0\rangle_{\mathrm{E}} + 3\lvert 1\rangle_{\mathrm{E}}}{\sqrt{12}}$
SWAP-01	yes (happens with probability $\frac{1}{3}$)	$\lvert 0,0\rangle_{\mathrm{B}}\lvert 2\rangle_{\mathrm{E}}$

Therefore, this attack makes it possible for Eve to get full information without inducing any error. However, Eve causes many losses, including full loss of the CTRL bits.

4.2.2 A Weaker Attack on the Simplified Protocol

The full attack described in Subsection 4.2.1 makes it impossible for Bob to ever detect a CTRL bit, which may look suspicious. We now present a weaker attack that lets Bob detect some CTRL bits but gives Eve less information.

The first stage of the attack (on the way from Bob to Alice) remains the same: that is, the state Eve sends to Alice is still given by Equation (4.4), and the resulting Bob+Eve state after Alice's classical operation is still shown in Table 4.2. Eve's probe space is, too, the same as before: $\mathcal{H}_{\mathrm{E}} = \mathcal{H}_3 \triangleq \mathrm{Span}\{\lvert 0\rangle_{\mathrm{E}}, \lvert 1\rangle_{\mathrm{E}}, \lvert 2\rangle_{\mathrm{E}}\}$.

This attack is characterized by the parameter $0 \le \epsilon \le 1$. We will see that $\epsilon = 0$ gives the full attack described in Subsection 4.2.1, while $\epsilon = 1$ gives Eve no information at all.

Another important parameter used by the attack is

$$\kappa \triangleq \sqrt{\frac{1-\epsilon^2}{3-2\epsilon^2}}. \tag{4.9}$$

We notice that for small values of ϵ, the value of κ is close to $\sqrt{\frac{1}{3}}$. Moreover, for all $0 \le \epsilon \le 1$, it holds that $0 < \epsilon^2 + \kappa^2 \le 1$ and $2\kappa^2 < 1$.

In the second stage of the attack (on the way from Alice to Bob), Eve applies the unitary operator V on the joint Bob+Eve state, where V is defined as follows:

$$V\lvert 0,1\rangle_{\mathrm{B}}\lvert 1\rangle_{\mathrm{E}} = \epsilon\lvert 0,1\rangle_{\mathrm{B}}\lvert 2\rangle_{\mathrm{E}} - \kappa\lvert 1,0\rangle_{\mathrm{B}}\lvert 1\rangle_{\mathrm{E}} + \sqrt{1-\kappa^2-\epsilon^2}\lvert 0,0\rangle_{\mathrm{B}}\lvert 0\rangle_{\mathrm{E}}, \tag{4.10}$$

$$V\lvert 1,0\rangle_{\mathrm{B}}\lvert 1\rangle_{\mathrm{E}} = -\kappa\lvert 0,1\rangle_{\mathrm{B}}\lvert 0\rangle_{\mathrm{E}} + \epsilon\lvert 1,0\rangle_{\mathrm{B}}\lvert 2\rangle_{\mathrm{E}} + \sqrt{1-\kappa^2-\epsilon^2}\lvert 0,0\rangle_{\mathrm{B}}\lvert 1\rangle_{\mathrm{E}}, \tag{4.11}$$

$$V\lvert 0,0\rangle_{\mathrm{B}}\lvert 0\rangle_{\mathrm{E}} = \kappa\lvert 0,1\rangle_{\mathrm{B}}\lvert 0\rangle_{\mathrm{E}} + \kappa\lvert 1,0\rangle_{\mathrm{B}}\lvert 1\rangle_{\mathrm{E}} + \sqrt{1-2\kappa^2}\lvert 0,0\rangle_{\mathrm{B}}\lvert +\rangle_{\mathrm{E}}, \tag{4.12}$$

$$V\lvert 0,0\rangle_{\mathrm{B}}\lvert 1\rangle_{\mathrm{E}} = \lvert 0,0\rangle_{\mathrm{B}}\lvert 2\rangle_{\mathrm{E}}. \tag{4.13}$$

Table 4.4: The state of Bob+Eve after completing Eve's attack described in Subsection 4.2.2, depending on Alice's classical operation and on whether Alice detected a photon or not. The parameters a and b are defined in Equations (4.16)–(4.17).

Alice's Operation	**Did Alice Detect a Photon?**	**Bob+Eve State**
CTRL	no (happens with certainty)	$\sqrt{\frac{2\epsilon^2}{3}}\|0,1\rangle_{\mathrm{x,B}}\|2\rangle_{\mathrm{E}}$ $+\sqrt{1-\frac{2\epsilon^2}{3}}\|0,0\rangle_{\mathrm{B}}\|+\rangle_{\mathrm{E}}$
SWAP-10	no (happens with probability $\frac{2}{3}$)	$\frac{1}{\sqrt{2}}\left[\|0,1\rangle_{\mathrm{B}}\left(\epsilon\|2\rangle_{\mathrm{E}}+\kappa\|0\rangle_{\mathrm{E}}\right)\right.$ $\left.+\|0,0\rangle_{\mathrm{B}}\left(a\|0\rangle_{\mathrm{E}}+b\|1\rangle_{\mathrm{E}}\right)\right]$
SWAP-10	yes (happens with probability $\frac{1}{3}$)	$\|0,0\rangle_{\mathrm{B}}\|2\rangle_{\mathrm{E}}$
SWAP-01	no (happens with probability $\frac{2}{3}$)	$\frac{1}{\sqrt{2}}\left[\|1,0\rangle_{\mathrm{B}}\left(\epsilon\|2\rangle_{\mathrm{E}}+\kappa\|1\rangle_{\mathrm{E}}\right)\right.$ $\left.+\|0,0\rangle_{\mathrm{B}}\left(b\|0\rangle_{\mathrm{E}}+a\|1\rangle_{\mathrm{E}}\right)\right]$
SWAP-01	yes (happens with probability $\frac{1}{3}$)	$\|0,0\rangle_{\mathrm{B}}\|2\rangle_{\mathrm{E}}$

V is indeed a unitary operator, because we can prove the right-hand sides to be orthonormal: all right-hand sides are clearly normalized; the first two vectors are orthogonal; the fourth vector is orthogonal to the three others; and the third vector is orthogonal to the first and to the second, because

$$1-2\kappa^2 = \frac{3-2\epsilon^2-2(1-\epsilon^2)}{3-2\epsilon^2} = \frac{1}{3-2\epsilon^2}, \tag{4.14}$$

$$1-\kappa^2-\epsilon^2 = \frac{(3-2\epsilon^2)-(1-\epsilon^2)-(3\epsilon^2-2\epsilon^4)}{3-2\epsilon^2} = \frac{2(1-\epsilon^2)^2}{3-2\epsilon^2}, \tag{4.15}$$

and thus $\frac{\sqrt{1-\kappa^2-\epsilon^2}\sqrt{1-2\kappa^2}}{\sqrt{2}} = \kappa^2$. Therefore, V extends to a unitary operator on $\mathcal{H}_{\mathrm{B}}\otimes\mathcal{H}_{\mathrm{E}}$.

The final global state after Eve's attack is described in Table 4.4 (computed by applying the operator V to Table 4.2), given the following definitions:

$$a \triangleq \sqrt{1-\kappa^2-\epsilon^2}+\frac{\sqrt{1-2\kappa^2}}{\sqrt{2}}, \tag{4.16}$$

$$b \triangleq \frac{\sqrt{1-2\kappa^2}}{\sqrt{2}}. \tag{4.17}$$

We notice that for $\epsilon = 0$, the attack is the same as in Subsection 4.2.1. If $\epsilon = 1$, the loss rate of CTRL bits is $\frac{1}{3}$, and Eve gets no information at all on the information bits (because $\kappa = 0$).

In general, if Alice and Bob share a "secret" bit $b \in \{0,1\}$, Eve's probe state is in the (normalized) state

$$\frac{\epsilon|2\rangle_{\mathrm{E}}+\kappa|b\rangle_{\mathrm{E}}}{\sqrt{\epsilon^2+\kappa^2}}. \tag{4.18}$$

When Eve measures her probe state in the computational basis $\{|0\rangle_{\mathrm{E}}, |1\rangle_{\mathrm{E}}, |2\rangle_{\mathrm{E}}\}$,

Table 4.5: The probability p of Eve obtaining an information bit, and the loss rates R_{CTRL} and $R_{\text{SWAP-x}}$ of CTRL and SWAP-x bits (where $x \in \{01, 10\}$), respectively, for several values of the attack's parameter ϵ.

$\boldsymbol{\epsilon}$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
p	1	0.97	0.89	0.78	0.66	0.55	0.44	0.34	0.25	0.15	0
$\mathbf{R}_{\text{CTRL}}$	1	0.99	0.97	0.94	0.89	0.83	0.76	0.67	0.57	0.46	0.33
$\mathbf{R}_{\text{SWAP-x}}$	0.83	0.83	0.82	0.79	0.76	0.73	0.68	0.63	0.58	0.53	0.5

she gets the information bit b with probability

$$p = \frac{\kappa^2}{\epsilon^2 + \kappa^2} = \frac{1 - \epsilon^2}{1 + 2\epsilon^2 - 2\epsilon^4}, \tag{4.19}$$

and the loss rates of CTRL and SWAP-x bits (where $x \in \{01, 10\}$) are

$$\begin{aligned} R_{\text{CTRL}} &= 1 - \frac{2\epsilon^2}{3}, & (4.20) \\ R_{\text{SWAP-x}} &= 1 - \frac{\epsilon^2 + \kappa^2}{2}, & (4.21) \end{aligned}$$

respectively.

Table 4.5 shows the probabilities p and the loss rates $R_{\text{CTRL}}, R_{\text{SWAP-x}}$ for various values of ϵ. For example, for $\epsilon = 0.5$, Eve still gets the information bit with probability $p \approx 0.55$, Bob's loss rate for the CTRL bits is $R_{\text{CTRL}} \approx 0.83$, and his loss rate for the SWAP-x bits is $R_{\text{SWAP-x}} \approx 0.73$.

For all values of ϵ, the attack causes no errors. However, in principle, it can be detected because it causes different loss rates to different types of bits: the loss rate experienced by Bob in the CTRL bits, R_{CTRL}, is usually different from the loss rate in the SWAP-x bits, $R_{\text{SWAP-x}}$ (see Table 4.5 for details). Therefore, in principle, the attack can be detected by a statistical test for most values of ϵ.

The loss rates become equal only for the value $\epsilon = \epsilon_0 \triangleq \sqrt{\frac{3-\sqrt{3}}{2}} \approx 0.796$ (which gives $\kappa^2 = \frac{\epsilon^2}{3}$). It seems that this specific attack *cannot* be detected, even in principle: it causes no errors, and it causes the same loss rate for all qubits. For this attack, Eve gets the information bit with probability $p = \frac{1}{4}$, and the loss rates are $R_{\text{CTRL}} = R_{\text{SWAP-x}} = \frac{1}{\sqrt{3}} \approx 0.577$. Therefore, this attack gives Eve a reasonable amount of information, and it is not detectable by looking at errors or comparing loss rates. (We can slightly modify the attack to make the loss rates identical in both directions of the quantum channel, too.)

We conclude that this weaker attack gives Eve partial information, causes no errors, and causes several loss rates. We also conclude that since the loss rates caused by the attack are usually different for different types of bits, the attack can be detected, in principle, for any value of ϵ except ϵ_0. However, for $\epsilon = \epsilon_0$, the attack seems undetectable.

4.3 Conclusion

We have discussed a simpler and natural variant of the Mirror protocol (the "simplified Mirror protocol") which is easier to implement. We have found the simplified Mirror protocol to be completely non-robust, actually making it an "over-simplified" Mirror protocol. We have presented in Subsection 4.2.1 an attack giving Eve full information without causing any error; in addition, since this attack also causes full loss of the CTRL bits, we have presented in Subsection 4.2.2 weaker attacks giving Eve partial information, causing no errors, and causing fewer losses. In particular, we have presented a specific attack (characterized by the parameter $\epsilon = \epsilon_0 \triangleq \sqrt{\frac{3-\sqrt{3}}{2}} \approx 0.796$) that seems undetectable and gives Eve one quarter ($\frac{1}{4}$) of all information bits.

These attacks prove the simplified Mirror protocol, which allows Alice to use only three classical operations (CTRL, SWAP-10, and SWAP-01), to be completely non-robust. On the other hand, the Mirror protocol is proved completely robust (see Section 3.3). As explained in Section 4.1, the only difference between the simplified Mirror protocol and the Mirror protocol is that the Mirror protocol allows a fourth classical operation, SWAP-ALL; therefore, allowing the SWAP-ALL operation is necessary for robustness. More generally, the Mirror protocol probably cannot be made much simpler while keeping it robust: its complexity is crucial for robustness. Therefore, we have seen that if we need an SQKD protocol that is experimentally feasible in a secure way, we may have to use a relatively complicated protocol.

In this chapter, we have not checked the experimental feasibility of Eve's attacks, because Eve is usually assumed to be all-powerful. Nonetheless, it can be interesting to check in the future the experimental feasibility of those attacks and discover whether the simplified Mirror protocol is flawed also in practice and not "only" in theory. Other interesting directions for future research include trying to find experimentally feasible SQKD protocols that are simpler than the Mirror protocol, and trying to find similar attacks against other QKD and SQKD protocols that have not been proved completely robust.

Chapter 5

Security of the Mirror Protocol Against Uniform Collective Attacks

In this chapter, we prove security of the Mirror protocol against “uniform collective” attacks (defined in Subsection 2.3.2) and evaluate the resulting key rate.

This chapter is based on a preprint posted to the arXiv in 2020 by Walter O. Krawec, Rotem Liss, and Tal Mor [KLM20].

5.1 Introduction

This chapter proves security of the Mirror protocol under a large class of uniform collective attacks. The class of the “uniform collective attacks” is an important and powerful subclass of possible attacks (see Subsection 2.3.2 for details); some existing security proofs of SQKD protocols against general attacks may in fact be limited to uniform collective attacks, because they use de Finetti’s theorem and similar techniques (see [Ren08, CKR09]) that can directly be applied only to entanglement-based protocols[1]. Therefore, in this chapter we restrict our analysis to uniform collective attacks.

The uniform collective attacks analyzed in this chapter allow Eve to inject *multiple photons* into the classical user’s lab, but not into the quantum user’s lab (attacks of the later kind are left for future analysis, but we briefly discuss them in the beginning of Section 5.3). In addition, we limit our analysis to two-mode quantum communication, leaving more complicated attacks for future research. We assume Alice’s and Bob’s devices precisely implement the needed operations (most notably, Alice’s classical operations described in Equations (5.1)–(5.4)), and without loss of generality, we assume an all-powerful Eve controlling all errors and losses in the quantum channel.

[1] Applying de Finetti’s theorem and similar techniques to prepare-and-measure protocols (including SQKD protocols) is usually easy for one-way QKD protocols, but it does not necessarily work for two-way protocols. See Subsection 2.2.2 for details about the different types of QKD protocols.

We derive an information-theoretic proof of security against these attacks and simulate the performance of the protocol in a variety of realistic scenarios, including lossy quantum channels, compared to the BB84 protocol. Ultimately, this chapter shows that SQKD protocols hold the potential to be secure and feasible in practice, and not just “secure in ideal conditions”. The methods and techniques we present in this work may also be applicable to security proofs of other SQKD protocols or even other two-way QKD protocols where users are limited in some manner in their quantum capabilities.

5.2 The Mirror Protocol: a Concise Description

In this section we present a concise description of the Mirror protocol, which should be useful for the security proof. A full description of the protocol is available in Section 3.2;

In the Mirror protocol, in each round, Bob sends to Alice the initial state $|+\rangle_{\mathrm{B}}$, which is equivalent to $|0,1\rangle_{\mathrm{x,B}} \triangleq \frac{|0,1\rangle_{\mathrm{B}}+|1,0\rangle_{\mathrm{B}}}{\sqrt{2}}$. Then, Alice prepares an ancillary state in the initial vacuum state $|0,0\rangle_{\mathrm{A_{anc}}}$ and chooses *at random* one of the following four classical operations (defined on any Fock state she may possibly get, due to Eve’s attack):

I (CTRL) Reflect all photons towards Bob, without measuring any photon. The mathematical description is:

$$I|0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}} = |0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}}. \tag{5.1}$$

$\mathbf{S_1}$ (SWAP-10) Reflect all photons in the $|0\rangle$ mode towards Bob, and measure all photons in the $|1\rangle$ mode. The mathematical description is:

$$S_1|0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}} = |m_1,0\rangle_{\mathrm{A_{anc}}}|0,m_0\rangle_{\mathrm{B}}. \tag{5.2}$$

$\mathbf{S_0}$ (SWAP-01) Reflect all photons in the $|1\rangle$ mode towards Bob, and measure all photons in the $|0\rangle$ mode. The mathematical description is:

$$S_0|0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}} = |0,m_0\rangle_{\mathrm{A_{anc}}}|m_1,0\rangle_{\mathrm{B}}. \tag{5.3}$$

S (SWAP-ALL) Measure all photons, without reflecting any photon towards Bob. The mathematical description is:

$$S|0,0\rangle_{\mathrm{A_{anc}}}|m_1,m_0\rangle_{\mathrm{B}} = |m_1,m_0\rangle_{\mathrm{A_{anc}}}|0,0\rangle_{\mathrm{B}}. \tag{5.4}$$

We note that in the above mathematical description, Alice measures her ancillary state $|\cdot\rangle_{\mathrm{A_{anc}}}$ in the z basis and sends back to Bob the $|\cdot\rangle_{\mathrm{B}}$ state.

The states sent from Alice to Bob (without any error, loss, or eavesdropping) and their interpretations, depending on Alice’s random choice of a classical operation and on whether Alice detected a photon or not, are detailed in Table 5.1.

Table 5.1: The state sent from Alice to Bob in the Mirror protocol without errors or losses, and its interpretation, depending on Alice's random choice of a classical operation and on whether Alice detected a photon or not.

Alice's Op.	**Did Alice Detect a Photon?**	**State to Bob**	**Round Type**	**Raw Key**
CTRL	no (happens with certainty)	$\lvert 0,1\rangle_{\mathrm{x,B}}$	"test"	none
SWAP-10	no (happens with probability $\frac{1}{2}$)	$\lvert 0,1\rangle_{\mathrm{B}}$	"raw key"	0
SWAP-10	yes (happens with probability $\frac{1}{2}$)	$\lvert 0,0\rangle_{\mathrm{B}}$	"raw key"	none
SWAP-01	no (happens with probability $\frac{1}{2}$)	$\lvert 1,0\rangle_{\mathrm{B}}$	"raw key"	1
SWAP-01	yes (happens with probability $\frac{1}{2}$)	$\lvert 0,0\rangle_{\mathrm{B}}$	"raw key"	none
SWAP-ALL	yes (happens with certainty)	$\lvert 0,0\rangle_{\mathrm{B}}$	"SWAP-ALL"	none

Then, Bob measures the incoming state in a random basis (either the z basis or the x basis). We assume here, as is true in most experimental setups, that Alice and Bob use detectors and not counters: namely, their detectors cannot *count* the number of incoming photons. Therefore, when a detector clicks, Alice and Bob cannot know whether it detected a single-photon pulse (a single photon in its measured mode) or a multi-photon pulse (more than one photon in its measured mode).

After completing all rounds, Alice and Bob perform *classical post-processing*: Alice sends over the classical channel her operation choices (CTRL, SWAP-x, or SWAP-ALL; she keeps $x \in \{01, 10\}$ in secret); Bob sends over the classical channel his basis choices; and both of them reveal all rounds where they got a loss, and all measurement results each of them got in all testing rounds (CTRL, SWAP-ALL, and a random subset of the SWAP-x rounds, for which Alice also reveals her values of $x \in \{01, 10\}$) and in all mismatched rounds (such as rounds in which Alice used SWAP-10 and Bob used the x basis). In the non-testing rounds, as detailed in Table 5.1, Alice and Bob share the raw key bit 0 if Alice uses SWAP-10 and detects no photon while Bob measures in the z basis and detects a photon (or photons) in the $\lvert\mathtt{0}\rangle$ mode; similarly, they share the raw key bit 1 if Alice uses SWAP-01 and detects no photon while Bob measures in the z basis and detects a photon (or photons) in the $\lvert\mathtt{1}\rangle$ mode.

Now, Alice and Bob have enough information for computing all the probabilities they need for finding the key rate (that are detailed later, in Table 5.3), so they compute all these probabilities and deduce the final key rate according to the algorithm in Subsection 5.3.7. If the final key rate is negative, they abort the protocol; otherwise, they perform error correction and privacy amplification in the standard way for QKD protocols. At the end of the protocol, Alice and Bob hold an identical final key that is completely secure against any eavesdropper.

5.3 Security Proof of the Mirror Protocol Against Uniform Collective Attacks

We now prove security of the Mirror protocol. For our security proof, we assume that the adversary Eve is restricted to uniform collective attacks—namely, that Eve attacks each round in an independent and identical manner, but she is allowed to postpone the measurement of her private quantum ancilla until any future point in time. Beyond this, we will also assume in our security analysis that Eve is allowed to inject *any* signal into the forward channel (linking quantum Bob to classical Alice); in the reverse channel, she is free to perform any quantum unitary probe, but we will assume that the number of photons returning to Bob is at most one. That is, Eve is allowed to inject multiple photons into the channel going to Alice, but on the way back, only a single photon or no photons at all will be returned to Bob. This assumption means that Eve may need to remove photons on the way from Alice to Bob, if she sent multiple photons towards Alice; in Subsection 5.3.1 we explain how Eve can perform this attack.

The above assumption (that at most one photon is sent towards Bob) is made to simplify the analysis of the return channel. We point out that in Chapter 3 we proved the Mirror protocol to be completely *robust* even without this assumption—namely, proved it robust against *all* multi-photon attacks and *all* kinds of losses and dark counts (see Section 3.3); however, full *security* analysis of the multi-photon case, including both losses and dark counts, is very difficult even in the simplest one-way standard QKD, and even more so in any standard two-way QKD protocol such as "Plug & Play" [MHHTZG97], "Ping Pong" [BF02], and LM05 [LM05] (see also [BLMR13]). Furthermore, this case has not been analyzed in security proofs of many other SQKD protocols (e.g., [Kra15b, Kra16, ZQM18, Kra18]. Therefore, we do not aim to solve this major issue here in the specific case of the Mirror protocol: extending the full security proof to this most general case is left for future research.

5.3.1 Eve's Attacks

Eve's first attack: We first analyze the forward-channel attack—namely, the attack on the way from Bob to Alice. Here, we note that it is to Eve's advantage to simply discard the signal coming from Bob (which should be the same each round and carries no information at this point) and inject a signal of her own, possibly consisting of multiple photons and entangled with her private quantum ancilla.

Specifically, in each round, Bob sends to Alice the same quantum state: $|0,1\rangle_{\mathrm{x,B}} \triangleq \frac{|0,1\rangle_{\mathrm{B}}+|1,0\rangle_{\mathrm{B}}}{\sqrt{2}}$. At this point, Eve performs her *first* attack: she replaces Bob's original state by her own state. Without loss of generality, Eve's state is of the form:

$$|\psi_0\rangle \triangleq \sum_{\substack{m_1\geq 0\\ m_0\geq 0}} |m_1, m_0\rangle_{\mathrm{B}} |e_{m_1,m_0}\rangle_{\mathrm{E}}. \tag{5.5}$$

Eve's second attack: Then, Alice performs her classical operation (CTRL, SWAP-10, SWAP-01, or SWAP-ALL) and sends the resulting state back to Bob. Now, Eve performs her *second* attack, described as the unitary operator U_{R}. As explained above, for the second attack we make the simplifying assumption that Eve always sends *at most one photon*—namely, she sends a superposition of $|0,1\rangle_{\mathrm{B}}$, $|1,0\rangle_{\mathrm{B}}$, and $|0,0\rangle_{\mathrm{B}}$ with her corresponding ancillary states $|g^{0,1}_{m_1,m_0}\rangle_{\mathrm{E}}$, $|g^{1,0}_{m_1,m_0}\rangle_{\mathrm{E}}$, and $|g^{0,0}_{m_1,m_0}\rangle_{\mathrm{E}}$. We emphasize that this simplifying assumption applies only to the second attack, and *not* to the first attack.

Thus, Eve's second attack is of the form:

$$\begin{aligned} U_{\mathrm{R}}|m'_1,m'_0\rangle_{\mathrm{B}}|e_{m_1,m_0}\rangle_{\mathrm{E}} &= |0,1\rangle_{\mathrm{B}}|f^{0,1}_{m'_1,m'_0,m_1,m_0}\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}}|f^{1,0}_{m'_1,m'_0,m_1,m_0}\rangle_{\mathrm{E}} \\ &+ |0,0\rangle_{\mathrm{B}}|f^{0,0}_{m'_1,m'_0,m_1,m_0}\rangle_{\mathrm{E}}. \end{aligned} \tag{5.6}$$

However, in our security proof we use terms of the following simplified notations:

$$U_{\mathrm{R}}|m_1,m_0\rangle_{\mathrm{B}}|e_{m_1,m_0}\rangle_{\mathrm{E}} = |0,1\rangle_{\mathrm{B}}|g^{0,1}_{m_1,m_0}\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}}|g^{1,0}_{m_1,m_0}\rangle_{\mathrm{E}} + |0,0\rangle_{\mathrm{B}}|g^{0,0}_{m_1,m_0}\rangle_{\mathrm{E}}. \tag{5.7}$$

where we denote $|g^{j,k}_{m_1,m_0}\rangle_{\mathrm{E}} \triangleq |f^{j,k}_{m_1,m_0,m_1,m_0}\rangle_{\mathrm{E}}$. We note that the operation of U_{R} on states $|m'_1,m'_0\rangle_{\mathrm{B}}|e_{m_1,m_0}\rangle_{\mathrm{E}}$ where $m'_1 \neq m_1$ or $m'_0 \neq m_0$ will not appear in our security proof, because these states do not give us meaningful statistics[2] and thus do not contribute to the probabilities in Table 5.3. We also note that since Eve is all-powerful, she will have no trouble performing any unitary operation, even if it includes a complicated operation for reducing the number of photons.

In both attacks, subsystem B is sent to a legitimate user, while subsystem E is kept as Eve's ancilla.

5.3.2 Analyzing all Types of Rounds

In Table 5.2 we classify all rounds into six types, that Alice and Bob need to analyze. The rounds are classified according to Alice's random choice of a classical operation and Bob's random choice of a measurement basis.

Notice the use of basis-mismatched rounds. Technically, we could have used only the "standard" (basis-matching) rounds for completing the security proof, by using the Cauchy-Schwarz inequality for finding worst-case bounds. However, using the technique of analyzing "mismatched measurements" [BHP93, WMU08], we can derive a significantly improved formula for the final key rate.

Alice and Bob have to find relevant statistics for each type of round and compute all

[2] States of the form $U_{\mathrm{R}}|0,m_0\rangle_{\mathrm{B}}|e_{m_1,m_0}\rangle_{\mathrm{E}}$ and $U_{\mathrm{R}}|m_1,0\rangle_{\mathrm{B}}|e_{m_1,m_0}\rangle_{\mathrm{E}}$ may appear in "raw key" rounds analyzed in Subsection 5.3.3, but we analyze only rounds which contribute to the raw key, where Alice detects no photon—namely, $m_1 = 0$ or $m_0 = 0$, respectively. In addition, states of the form $U_{\mathrm{R}}|0,0\rangle_{\mathrm{B}}|e_{m_1,m_0}\rangle_{\mathrm{E}}$ may appear in "SWAP-ALL" rounds analyzed in Subsection 5.3.5, but we analyze only "double-clicks" of Alice (where Eve's attack U_{R} is irrelevant, although we use it algebraically to prove Lemma 5.1) and "creation" events (where Alice detects no photon, so $m_1 = m_0 = 0$).

Table 5.2: All types of rounds, according to Alice's random choice of a classical operation [CTRL, SWAP-x ($x \in \{01, 10\}$), or SWAP-ALL] and Bob's random choice of a measurement basis (z or x).

Round Type	**Alice's Operation**	**Bob's Basis**
"raw key"	SWAP-x	computational (z)
mismatched "raw key"	SWAP-x	Hadamard (x)
"test"	CTRL	Hadamard (x)
mismatched "test"	CTRL	computational (z)
"SWAP-ALL"	SWAP-ALL	computational (z)
mismatched "SWAP-ALL"	SWAP-ALL	Hadamard (x)

probabilities listed in Table 5.3. In Subsections 5.3.3–5.3.5 we relate these probabilities to the quantum states appearing in our security proof, and in Subsection 5.3.6 we derive the resulting final key rate formula.

Table 5.3: All the probabilities Alice and Bob need to compute, and the formulas relating them to quantum states in our security proof. All formulas are proved in Subsections 5.3.3–5.3.5.

Prob.	**Round**	**Definition**	**Formula**
$\langle E_0\|E_0\rangle_{\mathrm{E}}$	"raw key"	Alice, Bob get raw key bits $0, 0$, respectively	
$\langle E_1\|E_1\rangle_{\mathrm{E}}$	"raw key"	Alice, Bob get raw key bits $0, 1$, respectively	
$\langle E_2\|E_2\rangle_{\mathrm{E}}$	"raw key"	Alice, Bob get raw key bits $1, 0$, respectively	
$\langle E_3\|E_3\rangle_{\mathrm{E}}$	"raw key"	Alice, Bob get raw key bits $1, 1$, respectively	
M	"raw key"	both Alice and Bob get raw key bits	$= \sum_{i=0}^{3} \langle E_i\|E_i\rangle_{\mathrm{E}}$
$p_{0,+}$	mismatched "raw key"	Alice gets raw key bit 0; Bob observes $\|+\rangle$	$2\Re\langle E_0\|E_1\rangle_{\mathrm{E}} = 2p_{0,+} - (\langle E_0\|E_0\rangle_{\mathrm{E}} + \langle E_1\|E_1\rangle_{\mathrm{E}})$
$p_{1,+}$	mismatched "raw key"	Alice gets raw key bit 1; Bob observes $\|+\rangle$	$2\Re\langle E_2\|E_3\rangle_{\mathrm{E}} = 2p_{1,+} - (\langle E_2\|E_2\rangle_{\mathrm{E}} + \langle E_3\|E_3\rangle_{\mathrm{E}})$
$p_{+,+}$	"test"	Bob observes $\|+\rangle$	$= \left\| \sum_{i=0}^{3} \|E_i\rangle_{\mathrm{E}} - \sum_{j=0}^{1} (\|g_j\rangle_{\mathrm{E}} - \|h_j\rangle_{\mathrm{E}}) \right\|^2$
$p_{\mathtt{CTRL:0}}$	mismatched "test"	Bob observes $\|0, 1\rangle$	$= 2 \left\| \|E_0\rangle_{\mathrm{E}} + \|E_2\rangle_{\mathrm{E}} - \|g_0\rangle_{\mathrm{E}} + \|h_0\rangle_{\mathrm{E}} \right\|^2$
$p_{\mathtt{CTRL:1}}$	mismatched "test"	Bob observes $\|1, 0\rangle$	$= 2 \left\| \|E_1\rangle_{\mathrm{E}} + \|E_3\rangle_{\mathrm{E}} - \|g_1\rangle_{\mathrm{E}} + \|h_1\rangle_{\mathrm{E}} \right\|^2$
$p_{\mathtt{double}}$	"SWAP-ALL"	Alice observes a "double-click" event ($\|1, 1\rangle$)	$\langle h_0\|h_0\rangle_{\mathrm{E}} + \langle h_1\|h_1\rangle_{\mathrm{E}} \leq \frac{1}{2} p_{\mathtt{double}}$
$p_{\mathtt{create:0}}$	"SWAP-ALL"	Alice observes $\|0, 0\rangle$; Bob observes $\|0, 1\rangle$	$= 2\langle g_0\|g_0\rangle_{\mathrm{E}}$
$p_{\mathtt{create:1}}$	"SWAP-ALL"	Alice observes $\|0, 0\rangle$; Bob observes $\|1, 0\rangle$	$= 2\langle g_1\|g_1\rangle_{\mathrm{E}}$

In *all* types of rounds, Bob begins by sending $|0, 1\rangle_{\mathrm{x,B}} \triangleq \frac{|0,1\rangle_{\mathrm{B}} + |1,0\rangle_{\mathrm{B}}}{\sqrt{2}}$, which Eve immediately replaces by her own state $|\psi_0\rangle \triangleq \sum_{\substack{m_1 \geq 0 \\ m_0 \geq 0}} |m_1, m_0\rangle_{\mathrm{B}} |e_{m_1,m_0}\rangle_{\mathrm{E}}$ (see Equation (5.5)). Then, Alice chooses her classical operation, as detailed below.

5.3.3 "Raw Key" Rounds: Alice Chooses the SWAP-x Operation

In "raw key" rounds, Alice chooses either SWAP-10 or SWAP-01 (each with probability $\frac{1}{2}$), that are defined in Equations (5.2)–(5.3). Then, the non-normalized state of the joint system, conditioning on Alice detecting *no photon*[3], is:

$$\rho_{\mathrm{ABE}}^{\text{(after Alice)}} = \frac{1}{2}|0\rangle\langle 0|_{\mathrm{A}} \otimes P\left(\sum_{m_0 \geq 0} |0, m_0\rangle_{\mathrm{B}} |e_{0,m_0}\rangle_{\mathrm{E}}\right) + \frac{1}{2}|1\rangle\langle 1|_{\mathrm{A}} \otimes P\left(\sum_{m_1 \geq 0} |m_1, 0\rangle_{\mathrm{B}} |e_{m_1,0}\rangle_{\mathrm{E}}\right), \tag{5.8}$$

where we define:

$$P(|\psi\rangle) \triangleq |\psi\rangle\langle\psi|. \tag{5.9}$$

We note that $|0\rangle_{\mathrm{A}}$ and $|1\rangle_{\mathrm{A}}$ denote the raw key bit of Alice: Alice deduces it from her own choice of SWAP-10 (which corresponds to $|0\rangle_{\mathrm{A}}$) or SWAP-01 (which corresponds to $|1\rangle_{\mathrm{A}}$), as explained in Table 5.1.

After Eve's second attack (namely, after Eve applies the U_{R} operator defined in Equation (5.7)), the joint non-normalized state becomes:

$$\begin{aligned}
&U_{\mathrm{R}} \rho_{\mathrm{ABE}}^{\text{(after Alice)}} U_{\mathrm{R}}^\dagger \\
&= \frac{1}{2}|0\rangle\langle 0|_{\mathrm{A}} \otimes P\left(|0,1\rangle_{\mathrm{B}} \sum_{m_0 \geq 0} |g_{0,m_0}^{0,1}\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}} \sum_{m_0 \geq 0} |g_{0,m_0}^{1,0}\rangle_{\mathrm{E}} + |0,0\rangle_{\mathrm{B}} \sum_{m_0 \geq 0} |g_{0,m_0}^{0,0}\rangle_{\mathrm{E}}\right) \\
&+ \frac{1}{2}|1\rangle\langle 1|_{\mathrm{A}} \otimes P\left(|0,1\rangle_{\mathrm{B}} \sum_{m_1 \geq 0} |g_{m_1,0}^{0,1}\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}} \sum_{m_1 \geq 0} |g_{m_1,0}^{1,0}\rangle_{\mathrm{E}} + |0,0\rangle_{\mathrm{B}} \sum_{m_1 \geq 0} |g_{m_1,0}^{0,0}\rangle_{\mathrm{E}}\right).
\end{aligned} \tag{5.10}$$

To simplify notation, we define the following states in subsystem E:

$$\begin{aligned}
|E_0\rangle_{\mathrm{E}} &\triangleq \frac{1}{\sqrt{2}} \sum_{m_0 \geq 0} |g_{0,m_0}^{0,1}\rangle_{\mathrm{E}}, \\
|E_1\rangle_{\mathrm{E}} &\triangleq \frac{1}{\sqrt{2}} \sum_{m_0 \geq 0} |g_{0,m_0}^{1,0}\rangle_{\mathrm{E}}, \\
|E_2\rangle_{\mathrm{E}} &\triangleq \frac{1}{\sqrt{2}} \sum_{m_1 \geq 0} |g_{m_1,0}^{0,1}\rangle_{\mathrm{E}}, \\
|E_3\rangle_{\mathrm{E}} &\triangleq \frac{1}{\sqrt{2}} \sum_{m_1 \geq 0} |g_{m_1,0}^{1,0}\rangle_{\mathrm{E}},
\end{aligned} \tag{5.11}$$

[3] Notice that according to Table 5.1, raw key bits are shared by Alice and Bob only in "raw key" rounds where Alice detects *no photon* and Bob *does* detect a photon.

so Equation (5.10) becomes:

$$\begin{aligned}
&U_{\mathrm{R}}\rho_{\mathrm{ABE}}^{(\text{after Alice})}U_{\mathrm{R}}^{\dagger} \\
&= |0\rangle\langle 0|_{\mathrm{A}} \otimes P\left(|0,1\rangle_{\mathrm{B}}|E_0\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}}|E_1\rangle_{\mathrm{E}} + |0,0\rangle_{\mathrm{B}}\frac{1}{\sqrt{2}}\sum_{m_0\geq 0}|g_{0,m_0}^{0,0}\rangle_{\mathrm{E}}\right) \\
&+ |1\rangle\langle 1|_{\mathrm{A}} \otimes P\left(|0,1\rangle_{\mathrm{B}}|E_2\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}}|E_3\rangle_{\mathrm{E}} + |0,0\rangle_{\mathrm{B}}\frac{1}{\sqrt{2}}\sum_{m_1\geq 0}|g_{m_1,0}^{0,0}\rangle_{\mathrm{E}}\right). \qquad (5.12)
\end{aligned}$$

(a) Standard "Raw Key" Rounds: Bob Chooses the z Basis

Now, Bob measures his subsystem in the z basis, and his raw key bit is simply his measurement result ("0" or "1"). Conditioning on Bob detecting a photon (namely, measuring $|0,1\rangle_{\mathrm{B}}$ or $|1,0\rangle_{\mathrm{B}}$), the final *normalized* state of the joint system after Bob's measurement is:

$$\begin{aligned}
\rho_{\mathrm{ABE}} = \frac{1}{M}(&|00\rangle\langle 00|_{\mathrm{AB}} \otimes |E_0\rangle\langle E_0|_{\mathrm{E}} + |01\rangle\langle 01|_{\mathrm{AB}} \otimes |E_1\rangle\langle E_1|_{\mathrm{E}} \\
&+|10\rangle\langle 10|_{\mathrm{AB}} \otimes |E_2\rangle\langle E_2|_{\mathrm{E}} + |11\rangle\langle 11|_{\mathrm{AB}} \otimes |E_3\rangle\langle E_3|_{\mathrm{E}}), \qquad (5.13)
\end{aligned}$$

where M is a normalization term (which will be computed soon).

Equation (5.13) confirms that, as written in Table 5.3:

$$\begin{aligned}
\langle E_0|E_0\rangle_{\mathrm{E}} &= \Pr\left(\text{Alice gets raw key bit 0, and Bob gets raw key bit 0}\right), && (5.14)\\
\langle E_1|E_1\rangle_{\mathrm{E}} &= \Pr\left(\text{Alice gets raw key bit 0, and Bob gets raw key bit 1}\right), && (5.15)\\
\langle E_2|E_2\rangle_{\mathrm{E}} &= \Pr\left(\text{Alice gets raw key bit 1, and Bob gets raw key bit 0}\right), && (5.16)\\
\langle E_3|E_3\rangle_{\mathrm{E}} &= \Pr\left(\text{Alice gets raw key bit 1, and Bob gets raw key bit 1}\right). && (5.17)
\end{aligned}$$

In addition, we can compute the normalization term M:

$$\begin{aligned}
M &= \sum_{i=0}^{3}\langle E_i|E_i\rangle_{\mathrm{E}} = \Pr(\text{both Alice and Bob get raw key bits}) && (5.18)\\
&= \Pr\left(\text{Alice observes no photon, and Bob observes a photon}\right).
\end{aligned}$$

Notice that all these probabilities are *observable* quantities: Alice and Bob estimate $\langle E_0|E_0\rangle_{\mathrm{E}}$, $\langle E_1|E_1\rangle_{\mathrm{E}}$, $\langle E_2|E_2\rangle_{\mathrm{E}}$, $\langle E_3|E_3\rangle_{\mathrm{E}}$, and M during the classical post-processing stage by testing a random subset of raw key bits.

(b) Mismatched "Raw Key" Rounds: Bob Chooses the x Basis

In this case, Bob measures his subsystem in the x basis. Let us rewrite the state he measures, provided in Equation (5.12), by substituting $|0,1\rangle_{\mathrm{B}} = \frac{|+\rangle_{\mathrm{B}}+|-\rangle_{\mathrm{B}}}{\sqrt{2}}$ and

$|1,0\rangle_{\mathrm{B}} = \frac{|+\rangle_{\mathrm{B}} - |-\rangle_{\mathrm{B}}}{\sqrt{2}}$. We get:

$$\begin{aligned}
&U_{\mathrm{R}} \rho_{\mathrm{ABE}}^{(\text{after Alice})} U_{\mathrm{R}}^{\dagger} \\
&= |\mathsf{0}\rangle\langle\mathsf{0}|_{\mathrm{A}} \otimes P\left(|0,1\rangle_{\mathrm{B}}|E_0\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}}|E_1\rangle_{\mathrm{E}} + |0,0\rangle_{\mathrm{B}}\frac{1}{\sqrt{2}}\sum_{m_0\geq 0}|g_{0,m_0}^{0,0}\rangle_{\mathrm{E}}\right) \\
&+ |\mathsf{1}\rangle\langle\mathsf{1}|_{\mathrm{A}} \otimes P\left(|0,1\rangle_{\mathrm{B}}|E_2\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}}|E_3\rangle_{\mathrm{E}} + |0,0\rangle_{\mathrm{B}}\frac{1}{\sqrt{2}}\sum_{m_1\geq 0}|g_{m_1,0}^{0,0}\rangle_{\mathrm{E}}\right) \\
&= |\mathsf{0}\rangle\langle\mathsf{0}|_{\mathrm{A}} \otimes P\left(\frac{|+\rangle_{\mathrm{B}}}{\sqrt{2}}(|E_0\rangle_{\mathrm{E}} + |E_1\rangle_{\mathrm{E}}) + \cdots\right) \\
&+ |\mathsf{1}\rangle\langle\mathsf{1}|_{\mathrm{A}} \otimes P\left(\frac{|+\rangle_{\mathrm{B}}}{\sqrt{2}}(|E_2\rangle_{\mathrm{E}} + |E_3\rangle_{\mathrm{E}}) + \cdots\right),
\end{aligned} \tag{5.19}$$

where the remainders of the above terms (the "$\cdots$") are irrelevant to our discussion.

We denote by $p_{0,+}$ the probability that Alice gets the raw key bit 0 and Bob observes $|+\rangle_{\mathrm{B}}$ (see Table 5.3). Similarly, we denote by $p_{1,+}$ the probability that Alice gets the raw key bit 1 and Bob observes $|+\rangle_{\mathrm{B}}$. These probabilities are:

$$p_{0,+} = \left|\frac{|E_0\rangle_{\mathrm{E}} + |E_1\rangle_{\mathrm{E}}}{\sqrt{2}}\right|^2 = \frac{1}{2}\left(\langle E_0|E_0\rangle_{\mathrm{E}} + \langle E_1|E_1\rangle_{\mathrm{E}} + 2\Re\langle E_0|E_1\rangle_{\mathrm{E}}\right), \tag{5.20}$$

$$p_{1,+} = \left|\frac{|E_2\rangle_{\mathrm{E}} + |E_3\rangle_{\mathrm{E}}}{\sqrt{2}}\right|^2 = \frac{1}{2}\left(\langle E_2|E_2\rangle_{\mathrm{E}} + \langle E_3|E_3\rangle_{\mathrm{E}} + 2\Re\langle E_2|E_3\rangle_{\mathrm{E}}\right). \tag{5.21}$$

Therefore, we find:

$$2\Re\langle E_0|E_1\rangle_{\mathrm{E}} = 2p_{0,+} - \left(\langle E_0|E_0\rangle_{\mathrm{E}} + \langle E_1|E_1\rangle_{\mathrm{E}}\right), \tag{5.22}$$

$$2\Re\langle E_2|E_3\rangle_{\mathrm{E}} = 2p_{1,+} - \left(\langle E_2|E_2\rangle_{\mathrm{E}} + \langle E_3|E_3\rangle_{\mathrm{E}}\right). \tag{5.23}$$

5.3.4 "Test" Rounds: Alice Chooses the CTRL Operation

In "test" rounds, Eve sends to Alice her state $|\psi_0\rangle \triangleq \sum_{\substack{m_1\geq 0\\ m_0\geq 0}}|m_1,m_0\rangle_{\mathrm{B}}|e_{m_1,m_0}\rangle_{\mathrm{E}}$ (see Equation (5.5)), and Alice chooses the CTRL operation—namely, Alice does nothing (see Equation (5.1)). Then, Eve applies her second attack U_{R} (see Equation (5.7)), and the resulting quantum state is:

$$U_{\mathrm{R}}|\psi_0\rangle = |0,1\rangle_{\mathrm{B}}\sum_{\substack{m_1\geq 0\\ m_0\geq 0}}|g_{m_1,m_0}^{0,1}\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}}\sum_{\substack{m_1\geq 0\\ m_0\geq 0}}|g_{m_1,m_0}^{1,0}\rangle_{\mathrm{E}} + |0,0\rangle_{\mathrm{B}}\sum_{\substack{m_1\geq 0\\ m_0\geq 0}}|g_{m_1,m_0}^{0,0}\rangle_{\mathrm{E}}. \tag{5.24}$$

(a) Standard "Test" Rounds: Bob Chooses the x Basis

Changing basis, whereby $|0,1\rangle_{\text{B}} = \frac{|+\rangle_{\text{B}}+|-\rangle_{\text{B}}}{\sqrt{2}}$ and $|1,0\rangle_{\text{B}} = \frac{|+\rangle_{\text{B}}-|-\rangle_{\text{B}}}{\sqrt{2}}$, we find:

$$U_{\text{R}}|\psi_0\rangle = \frac{|+\rangle_{\text{B}}}{\sqrt{2}}\left(\sum_{\substack{m_1\geq 0\\ m_0\geq 0}} |g^{0,1}_{m_1,m_0}\rangle_{\text{E}} + \sum_{\substack{m_1\geq 0\\ m_0\geq 0}} |g^{1,0}_{m_1,m_0}\rangle_{\text{E}}\right) + \cdots, \tag{5.25}$$

where the extra $\cdots$ term is irrelevant to our discussion.

Let $p_{+,+}$ be the probability that Bob observes $|+\rangle_{\text{B}}$ (see Table 5.3). From Equation (5.25) we deduce:

$$\begin{aligned} p_{+,+} &= \left|\frac{1}{\sqrt{2}}\sum_{\substack{m_1\geq 0\\ m_0\geq 0}} |g^{0,1}_{m_1,m_0}\rangle_{\text{E}} + \frac{1}{\sqrt{2}}\sum_{\substack{m_1\geq 0\\ m_0\geq 0}} |g^{1,0}_{m_1,m_0}\rangle_{\text{E}}\right|^2 \qquad (5.26)\\ &= |(|E_0\rangle_{\text{E}} + |E_2\rangle_{\text{E}} - |g_0\rangle_{\text{E}} + |h_0\rangle_{\text{E}}) + (|E_1\rangle_{\text{E}} + |E_3\rangle_{\text{E}} - |g_1\rangle_{\text{E}} + |h_1\rangle_{\text{E}})|^2 \\ &= ||E_0\rangle_{\text{E}} + |E_2\rangle_{\text{E}} - |g_0\rangle_{\text{E}} + |h_0\rangle_{\text{E}}|^2 + ||E_1\rangle_{\text{E}} + |E_3\rangle_{\text{E}} - |g_1\rangle_{\text{E}} + |h_1\rangle_{\text{E}}|^2 \\ &+ 2\Re\left[(\langle E_0|_{\text{E}} + \langle E_2|_{\text{E}} - \langle g_0|_{\text{E}} + \langle h_0|_{\text{E}}) \cdot (|E_1\rangle_{\text{E}} + |E_3\rangle_{\text{E}} - |g_1\rangle_{\text{E}} + |h_1\rangle_{\text{E}})\right], \end{aligned}$$

where we define:

$$\begin{aligned} |g_0\rangle_{\text{E}} &\triangleq \frac{1}{\sqrt{2}}|g^{0,1}_{0,0}\rangle_{\text{E}}, \\ |g_1\rangle_{\text{E}} &\triangleq \frac{1}{\sqrt{2}}|g^{1,0}_{0,0}\rangle_{\text{E}}, \\ |h_0\rangle_{\text{E}} &\triangleq \frac{1}{\sqrt{2}}\sum_{\substack{m_1\geq 1\\ m_0\geq 1}} |g^{0,1}_{m_1,m_0}\rangle_{\text{E}}, \\ |h_1\rangle_{\text{E}} &\triangleq \frac{1}{\sqrt{2}}\sum_{\substack{m_1\geq 1\\ m_0\geq 1}} |g^{1,0}_{m_1,m_0}\rangle_{\text{E}}, \end{aligned} \tag{5.27}$$

and we remember from Equation (5.11) that:

$$\begin{aligned} |E_0\rangle_{\text{E}} &\triangleq \frac{1}{\sqrt{2}}\sum_{m_0\geq 0} |g^{0,1}_{0,m_0}\rangle_{\text{E}}, \\ |E_1\rangle_{\text{E}} &\triangleq \frac{1}{\sqrt{2}}\sum_{m_0\geq 0} |g^{1,0}_{0,m_0}\rangle_{\text{E}}, \\ |E_2\rangle_{\text{E}} &\triangleq \frac{1}{\sqrt{2}}\sum_{m_1\geq 0} |g^{0,1}_{m_1,0}\rangle_{\text{E}}, \\ |E_3\rangle_{\text{E}} &\triangleq \frac{1}{\sqrt{2}}\sum_{m_1\geq 0} |g^{1,0}_{m_1,0}\rangle_{\text{E}}. \end{aligned} \tag{5.28}$$

(b) Mismatched "Test" Rounds: Bob Chooses the z Basis

In this case, we denote by $p_{\texttt{CTRL:0}}$ the probability of Bob observing $|0,1\rangle_{\mathrm{B}}$ (see Table 5.3). From Equation (5.24), we find (similarly to the computation of $p_{+,+}$):

$$p_{\texttt{CTRL:0}} = \left| \sum_{\substack{m_1 \geq 0 \\ m_0 \geq 0}} |g^{0,1}_{m_1,m_0}\rangle_{\mathrm{E}} \right|^2 = 2 \left| |E_0\rangle_{\mathrm{E}} + |E_2\rangle_{\mathrm{E}} - |g_0\rangle_{\mathrm{E}} + |h_0\rangle_{\mathrm{E}} \right|^2 . \tag{5.29}$$

Similarly, denoting by $p_{\texttt{CTRL:1}}$ the probability of Bob observing $|1,0\rangle_{\mathrm{B}}$, we find:

$$p_{\texttt{CTRL:1}} = \left| \sum_{\substack{m_1 \geq 0 \\ m_0 \geq 0}} |g^{1,0}_{m_1,m_0}\rangle_{\mathrm{E}} \right|^2 = 2 \left| |E_1\rangle_{\mathrm{E}} + |E_3\rangle_{\mathrm{E}} - |g_1\rangle_{\mathrm{E}} + |h_1\rangle_{\mathrm{E}} \right|^2 . \tag{5.30}$$

5.3.5 "SWAP-ALL" Rounds: Alice Chooses the SWAP-ALL Operation, and Bob Chooses the z Basis

(a) The Probability of a "Double-Click" Event: Used for Upper-Bounding $\langle h_0|h_0\rangle_{\mathrm{E}}$ and $\langle h_1|h_1\rangle_{\mathrm{E}}$

In "SWAP-ALL" rounds, Eve sends to Alice the initial state $|\psi_0\rangle \triangleq \sum_{\substack{m_1 \geq 0 \\ m_0 \geq 0}} |m_1, m_0\rangle_{\mathrm{B}} |e_{m_1,m_0}\rangle_{\mathrm{E}}$ described in Equation (5.5), and Alice chooses the SWAP-ALL operation defined in Equation (5.4), which essentially means that Alice measures subsystem B and sends a vacuum state towards Bob.

Let us denote by $p_{\texttt{double}}$ the probability that Alice observes a "double-click" event (detecting a photon in *both* modes $|\mathfrak{o}\rangle$ and $|\imath\rangle$)—namely, that she measures a state $|m_1, m_0\rangle_{\mathrm{A_{anc}}}$ where $m_1, m_0 \geq 1$ (see Table 5.3). This probability is easily found to be:

$$p_{\texttt{double}} = \sum_{\substack{m_1 \geq 1 \\ m_0 \geq 1}} \langle e_{m_1,m_0} | e_{m_1,m_0}\rangle_{\mathrm{E}}. \tag{5.31}$$

We can thus prove the following Lemma:

Lemma 5.1. *$\langle h_0|h_0\rangle_{\mathrm{E}} \leq \frac{1}{2} p_{\texttt{double}}$ and $\langle h_1|h_1\rangle_{\mathrm{E}} \leq \frac{1}{2} p_{\texttt{double}}$, where $|h_0\rangle_{\mathrm{E}}, |h_1\rangle_{\mathrm{E}}$ were defined in Equation* (5.27)*.*

Proof. Let us define the non-normalized state $|\zeta\rangle$ as:

$$|\zeta\rangle \triangleq \frac{1}{\sqrt{2}} \sum_{\substack{m_1 \geq 1 \\ m_0 \geq 1}} |m_1, m_0\rangle_{\mathrm{B}} |e_{m_1,m_0}\rangle_{\mathrm{E}}. \tag{5.32}$$

(We use the state $|\zeta\rangle$ only for this algebraic proof; it does not appear in the protocol.)

Clearly:

$$\langle\zeta|\zeta\rangle = \frac{1}{2}\sum_{\substack{m_1\geq 1\\ m_0\geq 1}} \langle e_{m_1,m_0}|e_{m_1,m_0}\rangle_{\mathrm{E}} = \frac{1}{2}p_{\texttt{double}}. \tag{5.33}$$

Applying U_{R} (see Equation (5.7)), the state $|\zeta\rangle$ evolves to:

$$\begin{aligned} U_{\mathrm{R}}|\zeta\rangle &= \frac{1}{\sqrt{2}}\sum_{\substack{m_1\geq 1\\ m_0\geq 1}} \left(|0,1\rangle_{\mathrm{B}}|g^{0,1}_{m_1,m_0}\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}}|g^{1,0}_{m_1,m_0}\rangle_{\mathrm{E}} + |0,0\rangle_{\mathrm{B}}|g^{0,0}_{m_1,m_0}\rangle_{\mathrm{E}}\right) \\ &= |0,1\rangle_{\mathrm{B}}|h_0\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}}|h_1\rangle_{\mathrm{E}} + |0,0\rangle_{\mathrm{B}}|h_{\mathrm{vac}}\rangle_{\mathrm{E}} \end{aligned} \tag{5.34}$$

(where $|h_0\rangle_{\mathrm{E}}, |h_1\rangle_{\mathrm{E}}$ were defined in Equation (5.27), and $|h_{\mathrm{vac}}\rangle_{\mathrm{E}} \triangleq \frac{1}{\sqrt{2}}\sum_{\substack{m_1\geq 1\\ m_0\geq 1}}|g^{0,0}_{m_1,m_0}\rangle_{\mathrm{E}}$).

By unitarity of U_{R}, we have:

$$\frac{1}{2}p_{\texttt{double}} = \langle\zeta|\zeta\rangle = \langle h_0|h_0\rangle_{\mathrm{E}} + \langle h_1|h_1\rangle_{\mathrm{E}} + \langle h_{\mathrm{vac}}|h_{\mathrm{vac}}\rangle_{\mathrm{E}}, \tag{5.35}$$

which implies that $\langle h_0|h_0\rangle_{\mathrm{E}} + \langle h_1|h_1\rangle_{\mathrm{E}} \leq \frac{1}{2}p_{\texttt{double}}$. Since both $\langle h_0|h_0\rangle_{\mathrm{E}}$ and $\langle h_1|h_1\rangle_{\mathrm{E}}$ are non-negative, this implies $\langle h_0|h_0\rangle_{\mathrm{E}} \leq \frac{1}{2}p_{\texttt{double}}$ and $\langle h_1|h_1\rangle_{\mathrm{E}} \leq \frac{1}{2}p_{\texttt{double}}$, as we wanted.□

(b) The Probability of a "Creation" Event: Used for Computing $\langle g_0|g_0\rangle_{\mathrm{E}}$ and $\langle g_1|g_1\rangle_{\mathrm{E}}$

Let $p_{\texttt{create:0}}$ denote the probability that Alice observes $|0,0\rangle_{\mathrm{A_{anc}}}$ (namely, a vacuum state) and Bob observes $|0,1\rangle_{\mathrm{B}}$ (see Table 5.3). In this event, Eve "creates" (on the way from Alice to Bob) a photon in the $|\mathfrak{o}\rangle$ mode that should not have existed. (See Section 4.2 for examples of such attacks.) Similarly, let $p_{\texttt{create:1}}$ denote the probability that Alice observes $|0,0\rangle_{\mathrm{A_{anc}}}$ and Bob observes $|1,0\rangle_{\mathrm{B}}$.

After Eve sends the initial state $|\psi_0\rangle \triangleq \sum_{\substack{m_1\geq 0\\ m_0\geq 0}}|m_1,m_0\rangle_{\mathrm{B}}|e_{m_1,m_0}\rangle_{\mathrm{E}}$ described in Equation (5.5), and after Alice applies the SWAP-ALL operation defined in Equation (5.4), the resulting state is:

$$\sum_{\substack{m_1\geq 0\\ m_0\geq 0}} |m_1,m_0\rangle_{\mathrm{A_{anc}}}|0,0\rangle_{\mathrm{B}}|e_{m_1,m_0}\rangle_{\mathrm{E}}. \tag{5.36}$$

For computing the probabilities $p_{\texttt{create:0}}$ and $p_{\texttt{create:1}}$, we need to analyze the term where Alice observes $|0,0\rangle_{\mathrm{A_{anc}}}$—namely, the term $|0,0\rangle_{\mathrm{A_{anc}}}|0,0\rangle_{\mathrm{B}}|e_{0,0}\rangle_{\mathrm{E}}$. Now, Eve's second attack applies the unitary operator U_{R} (described in Equation (5.7)) to this non-normalized term, which gives the following final result:

$$|0,0\rangle_{\mathrm{A_{anc}}}\otimes U_{\mathrm{R}}|0,0\rangle_{\mathrm{B}}|e_{0,0}\rangle_{\mathrm{E}} = |0,0\rangle_{\mathrm{A_{anc}}}\otimes\left[|0,1\rangle_{\mathrm{B}}|g^{0,1}_{0,0}\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}}|g^{1,0}_{0,0}\rangle_{\mathrm{E}} + |0,0\rangle_{\mathrm{B}}|g^{0,0}_{0,0}\rangle_{\mathrm{E}}\right]. \tag{5.37}$$

Since $p_{\texttt{create:0}}$ is the probability that Alice observes $|0,0\rangle_{\mathrm{A_{anc}}}$ and Bob observes $|0,1\rangle_{\mathrm{B}}$ (and similarly for $p_{\texttt{create:1}}$), we get, according to the definitions of $|g_0\rangle_{\mathrm{E}}, |g_1\rangle_{\mathrm{E}}$ in

Equation (5.27):

$$\begin{aligned} p_{\texttt{create:0}} &= \langle g^{0,1}_{0,0} | g^{0,1}_{0,0} \rangle_{\mathrm{E}} = 2\langle g_0 | g_0 \rangle_{\mathrm{E}}, && (5.38) \\ p_{\texttt{create:1}} &= \langle g^{1,0}_{0,0} | g^{1,0}_{0,0} \rangle_{\mathrm{E}} = 2\langle g_1 | g_1 \rangle_{\mathrm{E}}. && (5.39) \end{aligned}$$

5.3.6 Deriving the Final Key Rate

We remember that the final normalized state of the joint system after Bob's measurement, in standard "raw key" rounds where raw key bits *are* generated, is, according to Equation (5.13):

$$\begin{aligned} \rho_{\mathrm{ABE}} = \frac{1}{M}(&|00\rangle\langle 00|_{\mathrm{AB}} \otimes |E_0\rangle\langle E_0|_{\mathrm{E}} + |01\rangle\langle 01|_{\mathrm{AB}} \otimes |E_1\rangle\langle E_1|_{\mathrm{E}} \\ &+|10\rangle\langle 10|_{\mathrm{AB}} \otimes |E_2\rangle\langle E_2|_{\mathrm{E}} + |11\rangle\langle 11|_{\mathrm{AB}} \otimes |E_3\rangle\langle E_3|_{\mathrm{E}}). \qquad (5.40) \end{aligned}$$

Theorem 1 from [Kra17] allows us to mathematically compute a bound on the conditional von Neumann entropy $S(A|E)$ of ρ_{ABE}, as follows:

$$\begin{aligned} S(A|E) \geq\ & \frac{\langle E_0|E_0\rangle_{\mathrm{E}} + \langle E_3|E_3\rangle_{\mathrm{E}}}{M} \cdot \left[H_2\left(\frac{\langle E_0|E_0\rangle_{\mathrm{E}}}{\langle E_0|E_0\rangle_{\mathrm{E}} + \langle E_3|E_3\rangle_{\mathrm{E}}} \right) - H_2(\lambda_1) \right] (5.41) \\ +\ & \frac{\langle E_1|E_1\rangle_{\mathrm{E}} + \langle E_2|E_2\rangle_{\mathrm{E}}}{M} \cdot \left[H_2\left(\frac{\langle E_1|E_1\rangle_{\mathrm{E}}}{\langle E_1|E_1\rangle_{\mathrm{E}} + \langle E_2|E_2\rangle_{\mathrm{E}}} \right) - H_2(\lambda_2) \right], \end{aligned}$$

where:

$$\begin{aligned} \lambda_1 &\triangleq \frac{1}{2} + \frac{\sqrt{\left(\langle E_0|E_0\rangle_{\mathrm{E}} - \langle E_3|E_3\rangle_{\mathrm{E}}\right)^2 + 4\Re^2\langle E_0|E_3\rangle_{\mathrm{E}}}}{2\left(\langle E_0|E_0\rangle_{\mathrm{E}} + \langle E_3|E_3\rangle_{\mathrm{E}}\right)}, && (5.42) \\ \lambda_2 &\triangleq \frac{1}{2} + \frac{\sqrt{\left(\langle E_1|E_1\rangle_{\mathrm{E}} - \langle E_2|E_2\rangle_{\mathrm{E}}\right)^2 + 4\Re^2\langle E_1|E_2\rangle_{\mathrm{E}}}}{2\left(\langle E_1|E_1\rangle_{\mathrm{E}} + \langle E_2|E_2\rangle_{\mathrm{E}}\right)}, && (5.43) \\ H_2(x) &\triangleq -x\log_2(x) - (1-x)\log_2(1-x). && (5.44) \end{aligned}$$

Thus, to complete our proof of security, we only need bounds on the quantities $\Re\langle E_0|E_3\rangle_{\mathrm{E}}$ and $\Re\langle E_1|E_2\rangle_{\mathrm{E}}$; all the other parameters in the above expressions ($\langle E_0|E_0\rangle_{\mathrm{E}}$, $\langle E_1|E_1\rangle_{\mathrm{E}}$, $\langle E_2|E_2\rangle_{\mathrm{E}}$, $\langle E_3|E_3\rangle_{\mathrm{E}}$, and M) are observable probabilities that appear in Table 5.3 and can be directly computed by Alice and Bob.

We thus expand Equation (5.26) and substitute Equations (5.22)–(5.23) and (5.29)–(5.30) (all appearing in Table 5.3):

$$\begin{aligned} p_{+,+} &= \left| |E_0\rangle_{\mathrm{E}} + |E_2\rangle_{\mathrm{E}} - |g_0\rangle_{\mathrm{E}} + |h_0\rangle_{\mathrm{E}} \right|^2 + \left| |E_1\rangle_{\mathrm{E}} + |E_3\rangle_{\mathrm{E}} - |g_1\rangle_{\mathrm{E}} + |h_1\rangle_{\mathrm{E}} \right|^2 \\ &+ 2\Re\left[\left(\langle E_0|_{\mathrm{E}} + \langle E_2|_{\mathrm{E}} - \langle g_0|_{\mathrm{E}} + \langle h_0|_{\mathrm{E}} \right) \cdot \left(|E_1\rangle_{\mathrm{E}} + |E_3\rangle_{\mathrm{E}} - |g_1\rangle_{\mathrm{E}} + |h_1\rangle_{\mathrm{E}} \right) \right] \\ &= \frac{1}{2}(p_{\texttt{CTRL:0}} + p_{\texttt{CTRL:1}}) \\ &+ 2\Re\left(\langle E_0|_{\mathrm{E}} + \langle E_2|_{\mathrm{E}} \right)\left(|E_1\rangle_{\mathrm{E}} + |E_3\rangle_{\mathrm{E}} \right) - 2\Re\left(\langle E_0|_{\mathrm{E}} + \langle E_2|_{\mathrm{E}} \right)\left(|g_1\rangle_{\mathrm{E}} - |h_1\rangle_{\mathrm{E}} \right) \\ &- 2\Re\left(\langle g_0|_{\mathrm{E}} - \langle h_0|_{\mathrm{E}} \right)\left(|E_1\rangle_{\mathrm{E}} + |E_3\rangle_{\mathrm{E}} \right) + 2\Re\left(\langle g_0|_{\mathrm{E}} - \langle h_0|_{\mathrm{E}} \right)\left(|g_1\rangle_{\mathrm{E}} - |h_1\rangle_{\mathrm{E}} \right) \end{aligned}$$

$$\begin{aligned} &= \frac{1}{2}(p_{\texttt{CTRL:0}} + p_{\texttt{CTRL:1}}) \\ &+ 2p_{0,+} - (\langle E_0|E_0\rangle_{\mathrm{E}} + \langle E_1|E_1\rangle_{\mathrm{E}}) + 2\Re\langle E_0|E_3\rangle_{\mathrm{E}} \\ &+ 2p_{1,+} - (\langle E_2|E_2\rangle_{\mathrm{E}} + \langle E_3|E_3\rangle_{\mathrm{E}}) + 2\Re\langle E_1|E_2\rangle_{\mathrm{E}} \\ &- 2\Re\left(\langle E_0|_{\mathrm{E}} + \langle E_2|_{\mathrm{E}}\right)\left(|g_1\rangle_{\mathrm{E}} - |h_1\rangle_{\mathrm{E}}\right) - 2\Re\left(\langle g_0|_{\mathrm{E}} - \langle h_0|_{\mathrm{E}}\right)\left(|E_1\rangle_{\mathrm{E}} + |E_3\rangle_{\mathrm{E}}\right) \\ &+ 2\Re\left(\langle g_0|_{\mathrm{E}} - \langle h_0|_{\mathrm{E}}\right)\left(|g_1\rangle_{\mathrm{E}} - |h_1\rangle_{\mathrm{E}}\right). \end{aligned} \tag{5.45}$$

The resulting equation is: (substituting Equation (5.18), which appears in Table 5.3)

$$\begin{aligned} \Re\left(\langle E_0|E_3\rangle_{\mathrm{E}} + \langle E_1|E_2\rangle_{\mathrm{E}}\right) &= \frac{1}{2}p_{+,+} - p_{0,+} - p_{1,+} - \frac{1}{4}(p_{\texttt{CTRL:0}} + p_{\texttt{CTRL:1}}) + \frac{1}{2}M \\ &+ \Re\left(\langle g_1|_{\mathrm{E}} - \langle h_1|_{\mathrm{E}}\right)\left(|E_0\rangle_{\mathrm{E}} + |E_2\rangle_{\mathrm{E}}\right) \\ &+ \Re\left(\langle g_0|_{\mathrm{E}} - \langle h_0|_{\mathrm{E}}\right)\left(|E_1\rangle_{\mathrm{E}} + |E_3\rangle_{\mathrm{E}}\right) \\ &- \Re\left(\langle g_0|_{\mathrm{E}} - \langle h_0|_{\mathrm{E}}\right)\left(|g_1\rangle_{\mathrm{E}} - |h_1\rangle_{\mathrm{E}}\right). \end{aligned} \tag{5.46}$$

By the Cauchy-Schwarz inequality, Lemma 5.1, and Equations (5.38)–(5.39) (all appearing in Table 5.3), we determine the following bound:

$$\begin{aligned} \Re\left(\langle E_0|E_3\rangle_{\mathrm{E}} + \langle E_1|E_2\rangle_{\mathrm{E}}\right) &\geq \frac{1}{2}p_{+,+} - p_{0,+} - p_{1,+} - \frac{1}{4}(p_{\texttt{CTRL:0}} + p_{\texttt{CTRL:1}}) + \frac{1}{2}M \\ &- \frac{1}{\sqrt{2}}\left(\sqrt{p_{\texttt{create:1}}} + \sqrt{p_{\texttt{double}}}\right)\left(\sqrt{\langle E_0|E_0\rangle_{\mathrm{E}}} + \sqrt{\langle E_2|E_2\rangle_{\mathrm{E}}}\right) \\ &- \frac{1}{\sqrt{2}}\left(\sqrt{p_{\texttt{create:0}}} + \sqrt{p_{\texttt{double}}}\right)\left(\sqrt{\langle E_1|E_1\rangle_{\mathrm{E}}} + \sqrt{\langle E_3|E_3\rangle_{\mathrm{E}}}\right) \\ &- \frac{1}{2}\left(\sqrt{p_{\texttt{create:0}}} + \sqrt{p_{\texttt{double}}}\right)\left(\sqrt{p_{\texttt{create:1}}} + \sqrt{p_{\texttt{double}}}\right). \end{aligned} \tag{5.47}$$

To compute $S(A|E)$, we will simply minimize Equation (5.41) with respect to the condition outlined above and the following conditions (resulting from the Cauchy-Schwarz inequality):

$$|\Re\langle E_0|E_3\rangle_{\mathrm{E}}| \leq \sqrt{\langle E_0|E_0\rangle_{\mathrm{E}} \cdot \langle E_3|E_3\rangle_{\mathrm{E}}}, \tag{5.48}$$

$$|\Re\langle E_1|E_2\rangle_{\mathrm{E}}| \leq \sqrt{\langle E_1|E_1\rangle_{\mathrm{E}} \cdot \langle E_2|E_2\rangle_{\mathrm{E}}}. \tag{5.49}$$

In addition, we need to compute the expression $H(A|B)$:

$$H(A|B) = H(AB) - H(B), \tag{5.50}$$

where:

$$H(AB) = H\left(\frac{\langle E_0|E_0\rangle_{\mathrm{E}}}{M}, \frac{\langle E_1|E_1\rangle_{\mathrm{E}}}{M}, \frac{\langle E_2|E_2\rangle_{\mathrm{E}}}{M}, \frac{\langle E_3|E_3\rangle_{\mathrm{E}}}{M}\right), \tag{5.51}$$

$$H(B) = H\left(\frac{\langle E_0|E_0\rangle_{\mathrm{E}} + \langle E_2|E_2\rangle_{\mathrm{E}}}{M}, \frac{\langle E_1|E_1\rangle_{\mathrm{E}} + \langle E_3|E_3\rangle_{\mathrm{E}}}{M}\right). \tag{5.52}$$

The final key rate expression is given by the Devetak-Winter key rate formula [DW05]:

$$r = S(A|E) - H(A|B), \tag{5.53}$$

using $S(A|E)$ and $H(A|B)$ computed above.

5.3.7 Algorithm for Computing the Key Rate

The following algorithm allows us to compute the key rate for any noise model and experimental data:

1. Estimate all probabilities and inner products listed in Table 5.3. (All these probabilities can be computed by Alice and Bob in the classical post-processing stage.)

2. Compute the minimal value of the lower bound for $S(A|E)$ presented in Equation (5.41), which is copied here:

$$\begin{aligned} S(A|E) \geq {} & \frac{\langle E_0|E_0\rangle_\mathrm{E} + \langle E_3|E_3\rangle_\mathrm{E}}{M} \cdot \left[H_2\left(\frac{\langle E_0|E_0\rangle_\mathrm{E}}{\langle E_0|E_0\rangle_\mathrm{E} + \langle E_3|E_3\rangle_\mathrm{E}}\right) - H_2(\lambda_1)\right] \\ & + \frac{\langle E_1|E_1\rangle_\mathrm{E} + \langle E_2|E_2\rangle_\mathrm{E}}{M} \cdot \left[H_2\left(\frac{\langle E_1|E_1\rangle_\mathrm{E}}{\langle E_1|E_1\rangle_\mathrm{E} + \langle E_2|E_2\rangle_\mathrm{E}}\right) - H_2(\lambda_2)\right], \end{aligned} \tag{5.54}$$

   where

$$\lambda_1 \triangleq \frac{1}{2} + \frac{\sqrt{\left(\langle E_0|E_0\rangle_\mathrm{E} - \langle E_3|E_3\rangle_\mathrm{E}\right)^2 + 4\Re^2\langle E_0|E_3\rangle_\mathrm{E}}}{2\left(\langle E_0|E_0\rangle_\mathrm{E} + \langle E_3|E_3\rangle_\mathrm{E}\right)}, \tag{5.55}$$

$$\lambda_2 \triangleq \frac{1}{2} + \frac{\sqrt{\left(\langle E_1|E_1\rangle_\mathrm{E} - \langle E_2|E_2\rangle_\mathrm{E}\right)^2 + 4\Re^2\langle E_1|E_2\rangle_\mathrm{E}}}{2\left(\langle E_1|E_1\rangle_\mathrm{E} + \langle E_2|E_2\rangle_\mathrm{E}\right)}, \tag{5.56}$$

$$H_2(x) \triangleq -x\log_2(x) - (1-x)\log_2(1-x), \tag{5.57}$$

   where the minimum is taken over $\Re\langle E_0|E_3\rangle_\mathrm{E}$ and $\Re\langle E_1|E_2\rangle_\mathrm{E}$, subject to the three following constraints:

$$\begin{aligned} \Re\left(\langle E_0|E_3\rangle_\mathrm{E} + \langle E_1|E_2\rangle_\mathrm{E}\right) \geq {} & \frac{1}{2}p_{+,+} - p_{0,+} - p_{1,+} - \frac{1}{4}(p_{\mathtt{CTRL:0}} + p_{\mathtt{CTRL:1}}) + \frac{1}{2}M \\ & - \frac{1}{\sqrt{2}}\left(\sqrt{p_{\mathtt{create:1}}} + \sqrt{p_{\mathtt{double}}}\right)\left(\sqrt{\langle E_0|E_0\rangle_\mathrm{E}} + \sqrt{\langle E_2|E_2\rangle_\mathrm{E}}\right) \\ & - \frac{1}{\sqrt{2}}\left(\sqrt{p_{\mathtt{create:0}}} + \sqrt{p_{\mathtt{double}}}\right)\left(\sqrt{\langle E_1|E_1\rangle_\mathrm{E}} + \sqrt{\langle E_3|E_3\rangle_\mathrm{E}}\right) \\ & - \frac{1}{2}\left(\sqrt{p_{\mathtt{create:0}}} + \sqrt{p_{\mathtt{double}}}\right)\left(\sqrt{p_{\mathtt{create:1}}} + \sqrt{p_{\mathtt{double}}}\right), \end{aligned} \tag{5.58}$$

$$|\Re\langle E_0|E_3\rangle_\mathrm{E}| \leq \sqrt{\langle E_0|E_0\rangle_\mathrm{E} \cdot \langle E_3|E_3\rangle_\mathrm{E}}, \tag{5.59}$$

$$|\Re\langle E_1|E_2\rangle_\mathrm{E}| \leq \sqrt{\langle E_1|E_1\rangle_\mathrm{E} \cdot \langle E_2|E_2\rangle_\mathrm{E}}. \tag{5.60}$$

Note that we evaluate the minimum because we assume the worst-case scenario—namely, that Eve chooses her attack so as to minimize $S(A|E)$ (and, thus, minimize the key rate r).

In practice, we can minimize over a single parameter (say, $\Re\langle E_1|E_2\rangle_{\mathrm{E}}$), and take the other one ($\Re\langle E_0|E_3\rangle_{\mathrm{E}}$) as the right-hand-side of Equation (5.58), minus the free parameter $\Re\langle E_1|E_2\rangle_{\mathrm{E}}$ (but not less than 0). This will give us the minimum, because for any given value of $\Re\langle E_1|E_2\rangle_{\mathrm{E}}$, it is beneficial for Eve to have the smallest possible (non-negative) value of $\Re\langle E_0|E_3\rangle_{\mathrm{E}}$.

3. Compute $H(A|B)$ using the observed parameters:

$$\begin{aligned} H(A|B) &= H(AB) - H(B) \\ &= H\left(\frac{\langle E_0|E_0\rangle_{\mathrm{E}}}{M}, \frac{\langle E_1|E_1\rangle_{\mathrm{E}}}{M}, \frac{\langle E_2|E_2\rangle_{\mathrm{E}}}{M}, \frac{\langle E_3|E_3\rangle_{\mathrm{E}}}{M}\right) \\ &- H\left(\frac{\langle E_0|E_0\rangle_{\mathrm{E}} + \langle E_2|E_2\rangle_{\mathrm{E}}}{M}, \frac{\langle E_1|E_1\rangle_{\mathrm{E}} + \langle E_3|E_3\rangle_{\mathrm{E}}}{M}\right). \end{aligned} \tag{5.61}$$

4. Find the final key rate expression, using the Devetak-Winter key rate formula [DW05]:

$$r = S(A|E) - H(A|B). \tag{5.62}$$

5.4 Examples

The key rate bounds we found in Section 5.3 work in a wide range of scenarios, and they can be evaluated for all the possible values of all probabilities in Table 5.3. We would now like to evaluate our bounds for two concrete scenarios, that are easily comparable with attacks on other QKD and SQKD protocols.

5.4.1 First Scenario: Single-Photon Attacks without Losses

In the first scenario, let us assume that Bob has a perfect qubit source (no multi-photon pulses) and there are no photon losses. Furthermore, let us assume that Eve does not perform a multi-qubit attack at all (not even in her *first* attack). In this scenario, the only free parameters are the noises $Q_{\mathrm{Z}}, Q_{\mathrm{X}}$ in the channel: Q_{Z} is the probability that a $|0,1\rangle_{\mathrm{B}}$ state is flipped into $|1,0\rangle_{\mathrm{B}}$ (and vice versa) in "raw key" rounds, and Q_{X} is the probability that a $|+\rangle_{\mathrm{B}}$ state is flipped into $|-\rangle_{\mathrm{B}}$ in "test" rounds.

We consider the following noise model:

- In the "raw key" rounds, we consider that *both* the forward channel (from Bob to Alice) and the reverse channel (from Alice to Bob) are depolarizing channels with error Q_{Z}, as follows:

$$\mathcal{E}_{Q_{\mathrm{Z}}}(\rho) = (1 - 2Q_{\mathrm{Z}})\rho + 2Q_{\mathrm{Z}} \cdot \frac{I_2}{2}. \tag{5.63}$$

- In the "test" rounds, we consider that the whole channel (from Bob to Alice and back to Bob; notice that Alice does nothing in such rounds) is a depolarizing channel with error Q_{X}, as follows:

$$\mathcal{E}_{Q_{\mathrm{X}}}(\rho) = (1 - 2Q_{\mathrm{X}})\rho + 2Q_{\mathrm{X}} \cdot \frac{I_2}{2}. \tag{5.64}$$

Here, in the forward attack, Eve always replaces Bob's original state $|0,1\rangle_{\mathrm{x,B}} \triangleq \frac{|0,1\rangle_{\mathrm{B}}+|1,0\rangle_{\mathrm{B}}}{\sqrt{2}}$ by the following state (a special case of Equation (5.5)):

$$|\psi_0\rangle = |0,1\rangle_{\mathrm{B}}|e_{0,1}\rangle_{\mathrm{E}} + |1,0\rangle_{\mathrm{B}}|e_{1,0}\rangle_{\mathrm{E}}, \tag{5.65}$$

with $\langle e_{0,1}|e_{0,1}\rangle_{\mathrm{E}} = \langle e_{1,0}|e_{1,0}\rangle_{\mathrm{E}} = \frac{1}{2}$.

5.4.2 Second Scenario: Single-Photon Attacks *with* Losses

In the second scenario, our noise model remains identical to the first scenario, except two modifications:

- In the forward channel (from Bob to Alice), a loss occurs with probability p_ℓ^{F}; if it *does not* occur, the original noise model is applied.

- In the reverse channel (from Alice to Bob), a loss occurs with probability p_ℓ^{R}; if it *does not* occur, the original noise model is applied.

We assume, in particular, that a loss is *final*: if a loss occurs in the forward channel, no photon will ever be observed in this round by either Alice or Bob.

5.4.3 Evaluation Results

In Table 5.4 we evaluate all probabilities in both scenarios.

Table 5.4: Computing all probabilities in Table 5.3 for both examples (both scenarios).

Probability	**Single-Photon; no Losses**	**Single-Photon + Losses**
$\langle E_0\|E_0\rangle_{\mathrm{E}} = \langle E_3\|E_3\rangle_{\mathrm{E}} =$	$\frac{1}{4}(1 - Q_{\mathrm{Z}})$	$\frac{1}{4}(1 - p_\ell^{\mathrm{F}})(1 - p_\ell^{\mathrm{R}})(1 - Q_{\mathrm{Z}})$
$\langle E_1\|E_1\rangle_{\mathrm{E}} = \langle E_2\|E_2\rangle_{\mathrm{E}} =$	$\frac{1}{4}Q_{\mathrm{Z}}$	$\frac{1}{4}(1 - p_\ell^{\mathrm{F}})(1 - p_\ell^{\mathrm{R}})Q_{\mathrm{Z}}$
$M =$	$\frac{1}{2}$	$\frac{1}{2}(1 - p_\ell^{\mathrm{F}})(1 - p_\ell^{\mathrm{R}})$
$p_{0,+} = p_{1,+} =$	$\frac{1}{8}$	$\frac{1}{8}(1 - p_\ell^{\mathrm{F}})(1 - p_\ell^{\mathrm{R}})$
$p_{+,+} =$	$1 - Q_{\mathrm{X}}$	$(1 - p_\ell^{\mathrm{F}})(1 - p_\ell^{\mathrm{R}})(1 - Q_{\mathrm{X}})$
$p_{\mathtt{CTRL:0}} = p_{\mathtt{CTRL:1}} =$	$\frac{1}{2}$	$\frac{1}{2}(1 - p_\ell^{\mathrm{F}})(1 - p_\ell^{\mathrm{R}})$
$p_{\mathtt{double}} =$	0	0
$p_{\mathtt{create:0}} = p_{\mathtt{create:1}} =$	0	0

First scenario—single-photon attacks without losses: Substituting the probabilities from Table 5.4 in Equations (5.58)–(5.60), we find the three constraints to be:

$$\begin{aligned} \Re\left(\langle E_0|E_3\rangle_{\mathrm{E}} + \langle E_1|E_2\rangle_{\mathrm{E}}\right) &\geq \frac{1}{4} - \frac{1}{2}Q_{\mathrm{X}}, && (5.66) \\ |\Re\langle E_0|E_3\rangle_{\mathrm{E}}| &\leq \frac{1}{4}(1 - Q_{\mathrm{Z}}), && (5.67) \\ |\Re\langle E_1|E_2\rangle_{\mathrm{E}}| &\leq \frac{1}{4}Q_{\mathrm{Z}}. && (5.68) \end{aligned}$$

As explained in Subsection 5.3.7, we numerically find the minimal value of the key-rate expression $r = S(A|E) - H(A|B)$ for various values of $Q_{\mathrm{Z,X}}$ by using the lower bound on $S(A|E)$ presented in Equation (5.54), which is evaluated under the three above constraints on the values of $\Re\langle E_0|E_3\rangle_{\mathrm{E}}$ and $\Re\langle E_1|E_2\rangle_{\mathrm{E}}$. This numerical optimization yields the graph shown in Figure 5.1, presenting two cases:

- In the *dependent* noise model, where the error rates Q_{X} and Q_{Z} are identical (namely, $Q_{\mathrm{X}} = Q_{\mathrm{Z}}$), we recover the asymptotic BB84 noise tolerance of 11%.
- In the *independent* noise model, where the two-way channel is modeled as two independent depolarizing channels (namely, $Q_{\mathrm{X}} = 2Q_{\mathrm{Z}}(1 - Q_{\mathrm{Z}})$), the maximal (asymptotic) noise tolerance is 7.9%.

Interestingly, both values agree with the values found in [Kra17] for the original "QKD with Classical Bob" SQKD protocol [BKM07].

In both scenarios, because the Mirror protocol is two-way, we compare it to *two* copies of BB84 performed from Alice to Bob; this is a common comparison for two-way protocols (see, for example, [BLMR13]). The key rate of two copies of BB84 is $2(1 - 2H_2(p))$—namely, twice the original key rate of BB84.

Second scenario—single-photon attacks with losses: Substituting the probabilities from Table 5.4 in Equations (5.58)–(5.60), we find the three constraints to be:

$$\begin{aligned} \Re\left(\langle E_0|E_3\rangle_{\mathrm{E}} + \langle E_1|E_2\rangle_{\mathrm{E}}\right) &\geq (1 - p_\ell^{\mathrm{F}})(1 - p_\ell^{\mathrm{R}})\left(\frac{1}{4} - \frac{1}{2}Q_{\mathrm{X}}\right), && (5.69) \\ |\Re\langle E_0|E_3\rangle_{\mathrm{E}}| &\leq \frac{1}{4}(1 - p_\ell^{\mathrm{F}})(1 - p_\ell^{\mathrm{R}})(1 - Q_{\mathrm{Z}}), && (5.70) \\ |\Re\langle E_1|E_2\rangle_{\mathrm{E}}| &\leq \frac{1}{4}(1 - p_\ell^{\mathrm{F}})(1 - p_\ell^{\mathrm{R}})Q_{\mathrm{Z}}. && (5.71) \end{aligned}$$

The numerical analysis for this scenario is similar to the previous one. However, here we must also model the loss rates, so we consider a fiber channel with loss rates $p_\ell^{\mathrm{F,R}} = 1 - 10^{-\alpha\ell}$ (where $\alpha = 0.15\frac{\mathrm{dB}}{\mathrm{km}}$ is the loss coefficient, and ℓ is measured in kilometers). We consider two examples of fiber lengths: $\ell = 10$km and $\ell = 50$km. Results are presented in Figure 5.2.

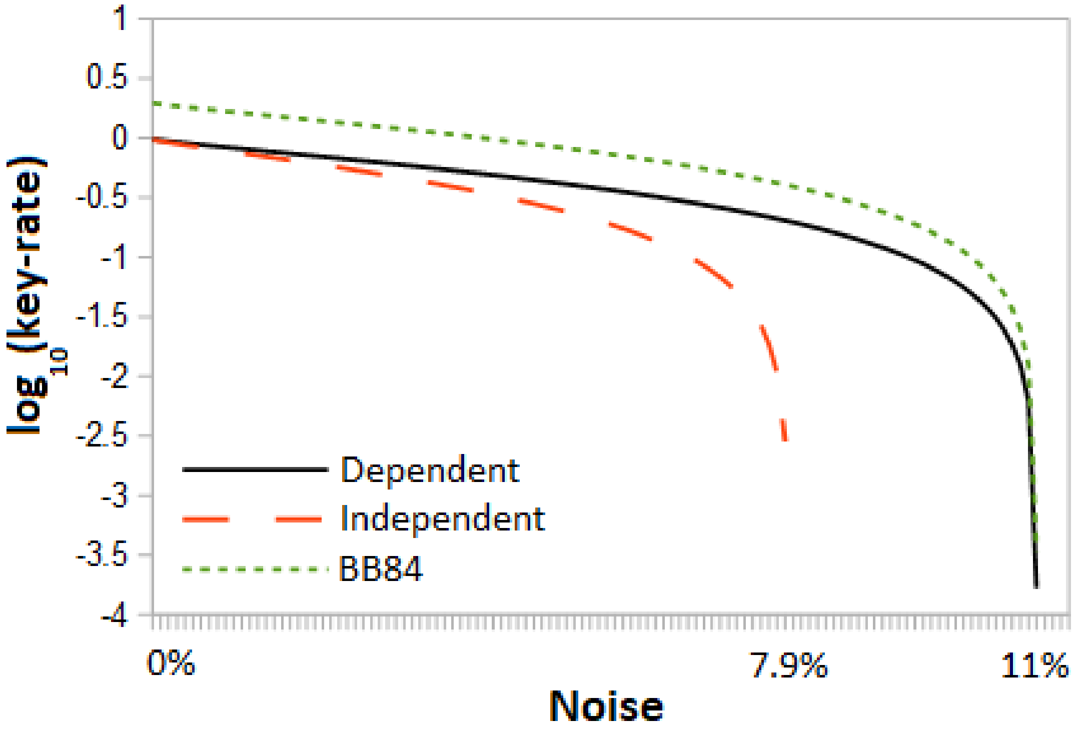

Figure 5.1: **A graph of the final key rate versus the noise level of the Mirror protocol in the first scenario** (single-photon attacks without losses), for dependent ($Q_\mathrm{X} = Q_\mathrm{Z}$) and independent ($Q_\mathrm{X} = 2Q_\mathrm{Z}(1 - Q_\mathrm{Z})$) noise models, compared to two copies of BB84.

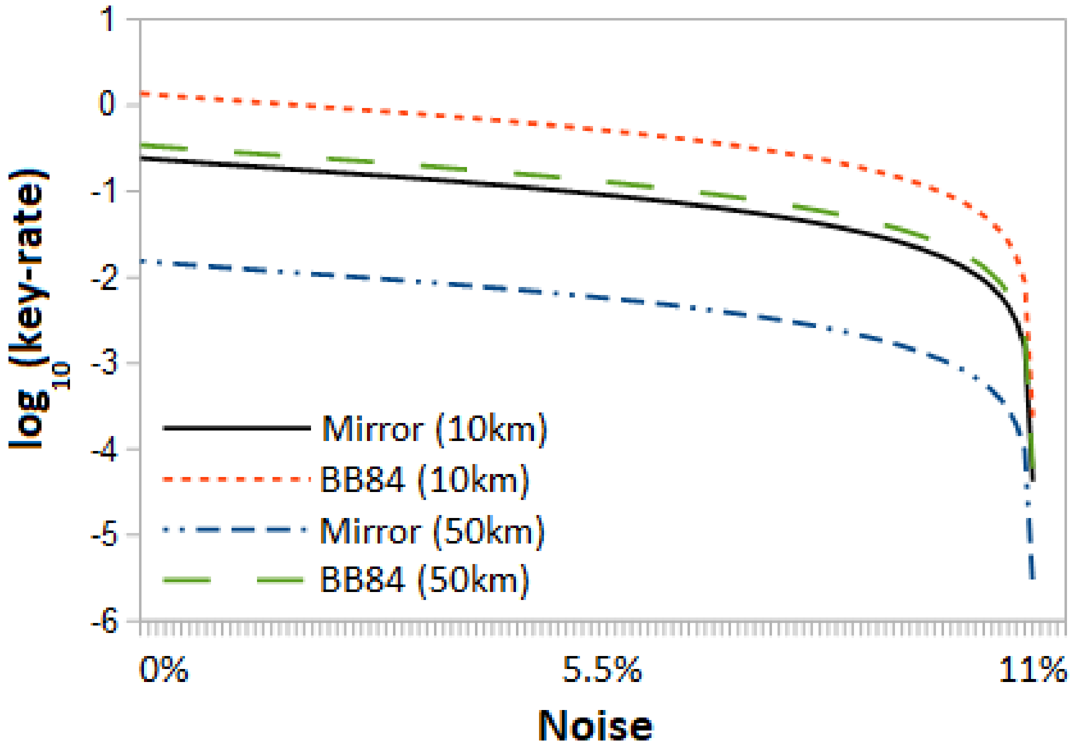

Figure 5.2: **A graph of the final key rate versus the noise level of the Mirror protocol in the second scenario** (single-photon attacks *with* losses), compared to two copies of BB84, for two possible lengths of fiber channels ($\ell = 10$km and $\ell = 50$km).

5.5 Conclusion

We have proved security of the Mirror protocol against uniform collective attacks, including attacks where the adversary Eve sends multiple photons towards the classical user (Alice). Our analysis shows that the asymptotic noise tolerance of the Mirror protocol is comparable, in the single-photon scenario, to the "QKD with Classical Bob" protocol [BKM07, Kra17] and even to the BB84 protocol. Moreover, we have suggested a general framework for analyzing multi-photon attacks; this framework may be useful for other QKD and SQKD protocols, too.

We conclude the Mirror protocol is theoretically secure against uniform collective attacks, and we suspect similar security results can be achieved for general attacks. Extensions of our results, such as security against general attacks, security against multi-photon attacks on both channels, and evaluation of our key-rate formula in the multi-photon case, are left for future research. Our extension to multi-photon attacks also suggests the intriguing possibility of analyzing SQKD protocols employing decoy states and similar counter-measures against practical attacks.

Our results show that SQKD protocols can potentially be implemented in a secure way, overcoming the practical attacks suggested by [TLC09, BKM09]. They therefore hold the potential to transform the SQKD protocols, making them not only theoretically fascinating, but also practically secure.

Chapter 6

Composable Security of the "BB84-INFO-z" Protocol Against Collective Attacks

In this chapter, we present a fully composable security proof of a new QKD protocol, that we name "BB84-INFO-z", against collective attacks (described in Subsection 2.3.2). The proof uses BBBMR's security approach, that is described in Subsection 2.3.3.

This chapter is based on a paper published in Theoretical Computer Science in 2020 by Michel Boyer, Rotem Liss, and Tal Mor [BLM20].

This is an extended (journal) version; the conference version was presented in the COMPLEXIS conference in 2017 by the same authors [BLM17] and was part of my M.Sc. thesis [Lis17], but its security proof was not fully composable. This journal version is extended to make the security proof (against collective attacks) fully composable.

6.1 Introduction

In this chapter, we extend the security proof of BB84 against collective attacks given in [BGM09], and we prove security of a QKD protocol we shall name "*BB84-INFO-z*" against collective attacks. This protocol is almost identical to BB84, except that all its INFO bits are in the z basis; in other words, the x basis is used only for testing. The bits are thus partitioned into three disjoint sets: INFO, TEST-Z, and TEST-X, of arbitrary sizes (n INFO bits, n_z TEST-Z bits, and n_x TEST-X bits).

Unlike the other papers that discussed BBBMR's security approach [BM97b, BM97a, BBBGM02, BBBMR06, BGM09] (see Subsection 2.3.3 for details), here we prove *fully composable* security of BB84-INFO-z against collective attacks. The method implemented in this chapter also directly applies to the BB84 security proof of [BGM09] against collective attacks, proving the fully composable security of BB84 against collective attacks. In Chapter 7 we further extend this method to show that the BB84 security proof of [BBBMR06] proves the fully composable security of BB84 (and, furthermore, of many BB84-like protocols) against joint attacks. (We note that in the conference version of this chapter [BLM17], we used a weaker security definition: it was not sufficient for proving fully composable security, but it was more composable than in previous papers.)

6.2 Full Definition of the "BB84-INFO-z" Protocol

Below we formally define all steps of the BB84-INFO-z protocol, as used in this chapter. See Section 2.7 for an explanation of the notation of bit strings ($\mathbf{s}$, $\mathbf{b}$, etc.), and see Section 1.1 for an explanation of the notations $|0^0\rangle, |1^0\rangle, |0^1\rangle, |1^1\rangle$.

1. Before the protocol, Alice and Bob choose some shared (and public) parameters: numbers n, n_z, and n_x (we denote $N \triangleq n + n_z + n_x$), error thresholds $p_{a,z}$ and $p_{a,x}$, an $r \times n$ parity check matrix P_{C} (corresponding to a linear error-correcting code C), and an $m \times n$ privacy amplification matrix P_{K} (representing a linear key-generation function). It is required that *all* $r + m$ rows of the matrices P_{C} and P_{K} put together are linearly independent.

2. Alice randomly chooses a partition $\mathcal{P} = (\mathbf{s}, \mathbf{z}, \mathbf{b})$ of the N bits by randomly choosing three N-bit strings $\mathbf{s}, \mathbf{z}, \mathbf{b} \in \mathbf{F}_2^N$ that satisfy $|\mathbf{s}| = n, |\mathbf{z}| = n_z, |\mathbf{b}| = n_x$, and $|\mathbf{s} + \mathbf{z} + \mathbf{b}| = N$. Thus, $\mathcal{P}$ partitions the set of indexes $\{1, 2, ..., N\}$ into three disjoint sets:

   - I (INFO bits, where $s_j = 1$) of size n;
   - $\mathrm{T_Z}$ (TEST-Z bits, where $z_j = 1$) of size n_z; and
   - $\mathrm{T_X}$ (TEST-X bits, where $b_j = 1$) of size n_x.

3. Alice randomly chooses an N-bit string $\mathbf{i} \in \mathbf{F}_2^N$ and sends the N qubit states $|i_1^{b_1}\rangle, |i_2^{b_2}\rangle, \ldots, |i_N^{b_N}\rangle$, one after the other, to Bob using the quantum channel. Notice that Alice uses the z basis for sending the INFO and TEST-Z bits, and that she uses the x basis for sending the TEST-X bits. Bob keeps each received qubit in quantum memory, not measuring it yet[1].

4. Alice sends to Bob over the classical channel the bit string $\mathbf{b} = b_1 \ldots b_N$. Bob measures each of the qubits he saved in the correct basis (namely, when measuring the i-th qubit, he measures it in the z basis if $b_i = 0$, and he measures it in the x basis if $b_i = 1$).

   The bit string measured by Bob is denoted by $\mathbf{i}^{\mathrm{B}}$. If there is no noise and no eavesdropping, then $\mathbf{i}^{\mathrm{B}} = \mathbf{i}$.

[1] Here we assume that Bob has a quantum memory and can delay his measurement. In practical implementations, Bob usually cannot do that, but is assumed to measure in a randomly-chosen basis (z or x), so that Alice and Bob later discard the qubits measured in the wrong basis. In that case, we need to assume that Alice sends more than N qubits, so that N qubits are finally detected by Bob and measured in the correct basis. In the original scheme, the probability of choosing each basis (z or x) was $\frac{1}{2}$, which caused half of the sent qubits to be lost; in the improved scheme suggested by [LCA05], the probability of choosing the z basis can be much higher, which means that fewer qubits get lost.

5. Alice sends to Bob over the classical channel the bit string $\mathbf{s}$. The INFO bits (that will be used for creating the final key) are the n bits with $s_j = 1$, while the TEST-Z and TEST-X bits (that will be used for testing) are the $n_z + n_x$ bits with $s_j = 0$. We denote the substrings of $\mathbf{i}, \mathbf{b}$ that correspond to the INFO bits by $\mathbf{i_s}$ and $\mathbf{b_s}$, respectively.

6. Alice and Bob both publish the bit values they have for all the TEST-Z and TEST-X bits, and they compare the bit values. If more than $n_z \cdot p_{a,z}$ TEST-Z bits are different between Alice and Bob *or* more than $n_x \cdot p_{a,x}$ TEST-X bits are different between them, they abort the protocol. We note that $p_{a,z}$ and $p_{a,x}$ (the pre-agreed error thresholds) are the maximal allowed error rates on the TEST-Z and TEST-X bits, respectively—namely, in each basis (z and x) separately.

7. The values of the remaining n bits (the INFO bits, with s_j =1) are kept in secret by Alice and Bob. The bit string of Alice is denoted $\mathbf{x} = \mathbf{i_s}$, and the bit string of Bob is denoted $\mathbf{x}^{\mathrm{B}}$.

8. Alice sends to Bob the *syndrome* of $\mathbf{x}$ (with respect to the error-correcting code C and to its corresponding parity check matrix P_{C}), that consists of r bits and is defined as $\boldsymbol{\xi} = \mathbf{x}P_{\mathrm{C}}^{\mathrm{T}}$. By using $\boldsymbol{\xi}$, Bob corrects the errors in his $\mathbf{x}^{\mathrm{B}}$ string (so that it is the same as $\mathbf{x}$).

9. The final key consists of m bits and is defined as $\mathbf{k} = \mathbf{x}P_{\mathrm{K}}^{\mathrm{T}}$. Both Alice and Bob compute it.

The protocol is defined similarly to BB84 (and to its description in [BGM09]), except that it uses the generalized bit numbers n, n_z, and n_x (numbers of INFO, TEST-Z, and TEST-X bits, respectively); that it uses the partition $\mathcal{P} = (\mathbf{s}, \mathbf{z}, \mathbf{b})$ for dividing the N-bit string $\mathbf{i}$ into three disjoint sets of indexes (I, $\mathrm{T_Z}$, and $\mathrm{T_X}$); and that it uses two separate thresholds ($p_{a,z}$ and $p_{a,x}$) instead of one (p_a).

6.3 Security Proof for the BB84-INFO-z Protocol Against Collective Attacks

6.3.1 The General Collective Attack of Eve

Before the beginning of the QKD protocol (and, thus, independently of $\mathbf{i}$ and $\mathcal{P}$), Eve chooses some collective attack to perform. A *collective attack* is bitwise: it attacks each qubit separately, by using a separate probe (ancillary state). Each probe is attached by Eve to the quantum state, and Eve saves it in a quantum memory. Eve can keep her quantum probes indefinitely, even after the final key is used by Alice and Bob; and she can perform, at any time of her choice, an optimal measurement of all her probes together, chosen based on all the information she has at the time of the measurement (including the classical information sent during the protocol, and including the information she acquires when Alice and Bob use the key).

Given the j-th qubit $|i_j^{b_j}\rangle_{\mathrm{T}_j}$ sent from Alice to Bob ($1 \leq j \leq N$), Eve attaches a probe state $|0^{\mathrm{E}}\rangle_{\mathrm{E}_j}$ and applies some unitary operator U_j of her choice to the compound system $|0^{\mathrm{E}}\rangle_{\mathrm{E}_j}|i_j^{b_j}\rangle_{\mathrm{T}_j}$. Then, Eve keeps to herself (in a quantum memory) the subsystem E_j, which is her probe state; and sends to Bob the subsystem T_j, which is the qubit sent from Alice to Bob (which may have been modified by her attack U_j).

The most general collective attack U_j of Eve on the j-th qubit, represented in the orthonormal basis $\{|0^{b_j}\rangle_{\mathrm{T}_j}, |1^{b_j}\rangle_{\mathrm{T}_j}\}$, is

$$U_j|0^{\mathrm{E}}\rangle_{\mathrm{E}_j}|0^{b_j}\rangle_{\mathrm{T}_j} = |E_{00}^{b_j}\rangle_{\mathrm{E}_j}|0^{b_j}\rangle_{\mathrm{T}_j} + |E_{01}^{b_j}\rangle_{\mathrm{E}_j}|1^{b_j}\rangle_{\mathrm{T}_j}, \tag{6.1}$$

$$U_j|0^{\mathrm{E}}\rangle_{\mathrm{E}_j}|1^{b_j}\rangle_{\mathrm{T}_j} = |E_{10}^{b_j}\rangle_{\mathrm{E}_j}|0^{b_j}\rangle_{\mathrm{T}_j} + |E_{11}^{b_j}\rangle_{\mathrm{E}_j}|1^{b_j}\rangle_{\mathrm{T}_j}, \tag{6.2}$$

where $|E_{00}^{b_j}\rangle_{\mathrm{E}_j}$, $|E_{01}^{b_j}\rangle_{\mathrm{E}_j}$, $|E_{10}^{b_j}\rangle_{\mathrm{E}_j}$, and $|E_{11}^{b_j}\rangle_{\mathrm{E}_j}$ are non-normalized states in Eve's probe system E_j attached to the j-th qubit.

We thus notice that Eve can modify the original *product state* of the compound system, $|0^{\mathrm{E}}\rangle_{\mathrm{E}_j}|i_j^{b_j}\rangle_{\mathrm{T}_j}$, into an *entangled state* (e.g., $|E_{00}^{b_j}\rangle_{\mathrm{E}_j}|0^{b_j}\rangle_{\mathrm{T}_j} + |E_{01}^{b_j}\rangle_{\mathrm{E}_j}|1^{b_j}\rangle_{\mathrm{T}_j}$). Eve's attack may thus cause Bob's state to become entangled with her probe. On the one hand, this may give Eve some information on Bob's state; on the other hand, this causes disturbance that may be detected by Bob. Our security proof shows that the information obtained by Eve and the disturbance caused by Eve are inherently correlated: this correlation is the basic reason QKD protocols are secure.

6.3.2 Results from [BGM09]

The security proof of BB84-INFO-z against collective attacks is very similar to the security proof of BB84 itself against collective attacks, that was detailed in [BGM09]. Most parts of the proof are not affected at all by the changes made to BB84 to get the BB84-INFO-z protocol (changes detailed in Section 6.2 of this chapter), because these parts assume fixed strings $\mathbf{s}$ and $\mathbf{b}$, and because the attack is collective (so the analysis is restricted to the INFO bits).

Therefore, the reader is referred to the proof in Section 2 and Subsections 3.1–3.5 of [BGM09], that applies to BB84-INFO-z without any changes (except changing the total number of bits, $2n$, to N, which does not affect the proof at all), and that will not be repeated here.

We denote the rows of the error-correction parity check matrix P_{C} as the vectors $v_1, \ldots, v_r$ in $\mathbf{F}_2^n$, and the rows of the privacy amplification matrix P_{K} as the vectors $v_{r+1}, \ldots, v_{r+m}$. We also define, for every r', $V_{r'} \triangleq \mathrm{Span}\{v_1, ..., v_{r'}\}$; and we define

$$d_{r,m} \triangleq \min_{r \leq r' < r+m} d_{\mathrm{H}}(v_{r'+1}, V_{r'}) = \min_{r \leq r' < r+m} d_{r',1}. \tag{6.3}$$

For a 1-bit final key $k \in \{0, 1\}$, we define $\widehat{\rho}_k$ to be the state of Eve corresponding to the final key k, given that she knows $\boldsymbol{\xi}$. Thus,

$$\widehat{\rho}_k = \frac{1}{2^{n-r-1}} \sum_{\mathbf{x} \,\Big|\, \substack{\mathbf{x} P_{\mathrm{C}}^{\mathrm{T}} = \boldsymbol{\xi} \\ \mathbf{x} \cdot v_{r+1} = k}} \rho_{\mathbf{x}}^{\mathbf{b}'}, \tag{6.4}$$

where $\rho_{\mathbf{x}}^{\mathbf{b}'}$ is Eve's state after the attack, given that Alice sent the INFO bit string $\mathbf{x}$ encoded in the bases $\mathbf{b}' = \mathbf{b_s}$. In [BGM09], the state $\widetilde{\rho}_k$ was also defined: it is a lift-up of $\widehat{\rho}_k$ (which means that $\widehat{\rho}_k$ is a partial trace of $\widetilde{\rho}_k$), in which the states $\rho_{\mathbf{x}}^{\mathbf{b}'}$ appearing in $\widehat{\rho}_k$ are replaced by their purifications (see full definition in Subsection 3.4 of [BGM09]).

In the end of Subsection 3.5 of [BGM09], it was found that (in the case of a 1-bit final key, i.e., $m = 1$)

$$\frac{1}{2} \operatorname{tr} |\widetilde{\rho}_0 - \widetilde{\rho}_1| \le 2 \sqrt{\Pr \left[|\mathbf{C}_{\mathrm{I}}| \ge \frac{d_{r,1}}{2} \ \mid \ \mathbf{B}_{\mathrm{I}} = \overline{\mathbf{b}'}, \mathbf{s} \right]}, \tag{6.5}$$

where $\mathbf{C}_{\mathrm{I}}$ is a random variable whose value is the n-bit string of errors on the n INFO bits; $\mathbf{B}_{\mathrm{I}}$ is a random variable whose value is the n-bit string of bases of the n INFO bits; $\overline{\mathbf{b}'}$ is the bit-flipped string of $\mathbf{b}' = \mathbf{b_s}$; and $d_{r,1}$ (and, in general, $d_{r,m}$) was defined above.

Now, according to [NC00, Theorem 9.2 and page 407], and using the fact that $\widehat{\rho}_k$ is a partial trace of $\widetilde{\rho}_k$, we find that $\frac{1}{2} \operatorname{tr} |\widehat{\rho}_0 - \widehat{\rho}_1| \le \frac{1}{2} \operatorname{tr} |\widetilde{\rho}_0 - \widetilde{\rho}_1|$. From this result and from inequality (6.5) we deduce that

$$\frac{1}{2} \operatorname{tr} |\widehat{\rho}_0 - \widehat{\rho}_1| \le 2 \sqrt{\Pr \left[|\mathbf{C}_{\mathrm{I}}| \ge \frac{d_{r,1}}{2} \ \mid \ \mathbf{B}_{\mathrm{I}} = \overline{\mathbf{b}'}, \mathbf{s} \right]}. \tag{6.6}$$

6.3.3 Bounding the Differences Between Eve's States

We define $\mathbf{c} \triangleq \mathbf{i} + \mathbf{i}^{\mathrm{B}}$: namely, $\mathbf{c}$ is the XOR of the N-bit string $\mathbf{i}$ sent by Alice and of the N-bit string $\mathbf{i}^{\mathrm{B}}$ measured by Bob. For all indexes $1 \le \ell \le N$, $c_\ell = 1$ if and only if Bob's ℓ-th bit value is different from the ℓ-th bit sent by Alice. The partition $\mathcal{P}$ divides the N bits into n INFO bits, n_z TEST-Z bits, and n_x TEST-X bits. The corresponding substrings of the error string $\mathbf{c}$ are $\mathbf{c_s}$ (the string of errors on the INFO bits), $\mathbf{c_z}$ (the string of errors on the TEST-Z bits), and $\mathbf{c_b}$ (the string of errors on the TEST-X bits). The random variables that correspond to $\mathbf{c_s}$, $\mathbf{c_z}$, and $\mathbf{c_b}$ are denoted by $\mathbf{C}_{\mathrm{I}}$, $\mathbf{C}_{\mathrm{T_Z}}$, and $\mathbf{C}_{\mathrm{T_X}}$, respectively.

We define $\widetilde{\mathbf{C}_{\mathrm{I}}}$ to be a random variable whose value is the string of errors on the INFO bits *if Alice had encoded and sent the INFO bits in the x basis* (instead of the z basis dictated by the protocol). In these notations, Equation (6.6) reads as

$$\frac{1}{2} \operatorname{tr} |\widehat{\rho}_0 - \widehat{\rho}_1| \le 2 \sqrt{\Pr \left[|\widetilde{\mathbf{C}_{\mathrm{I}}}| \ge \frac{d_{r,1}}{2} \ \mid \ \mathcal{P} \right]} = 2 \sqrt{\Pr \left[|\widetilde{\mathbf{C}_{\mathrm{I}}}| \ge \frac{d_{r,1}}{2} \ \mid \ \mathbf{c_z}, \mathbf{c_b}, \mathcal{P} \right]}, \tag{6.7}$$

using the fact that Eve's attack is collective, so the qubits are attacked independently, and, therefore, the errors on the INFO bits are independent of the errors on the TEST-Z and TEST-X bits (namely, of $\mathbf{c_z}$ and $\mathbf{c_b}$).

As explained in [BGM09], Equation (6.7) was not derived for the actual attack $U = U_1 \otimes \ldots \otimes U_N$ applied by Eve, but for a virtual flat attack (that depends on $\mathbf{b}$ and therefore could not have been applied by Eve). That flat attack gives the same states $\widehat{\rho}_0$ and $\widehat{\rho}_1$ as given by the original attack U, and it gives a lower (or the same) error rate in the conjugate basis. Therefore, Equation (6.7) holds for the original attack U, too. This means that, starting from this point, all our results apply to the original attack U rather than to the flat attack.

So far, we have discussed a 1-bit key. We will now discuss a general m-bit key $\mathbf{k}$. We define $\widehat{\rho}_{\mathbf{k}}$ to be the state of Eve corresponding to the final key $\mathbf{k}$, given that she knows $\boldsymbol{\xi}$:

$$\widehat{\rho}_{\mathbf{k}} = \frac{1}{2^{n-r-m}} \sum_{\mathbf{x} \left| \substack{\mathbf{x}P_{\mathrm{C}}^{\mathrm{T}} = \boldsymbol{\xi} \\ \mathbf{x}P_{\mathrm{K}}^{\mathrm{T}} = \mathbf{k}} \right.} \rho_{\mathbf{x}}^{\mathbf{b}'}. \tag{6.8}$$

Proposition 6.1. *For any two keys $\mathbf{k}, \mathbf{k}'$ of m bits,*

$$\frac{1}{2} \operatorname{tr} |\widehat{\rho}_{\mathbf{k}} - \widehat{\rho}_{\mathbf{k}'}| \le 2m \sqrt{\Pr\left[|\widetilde{\mathbf{C}}_{\mathrm{I}}| \ge \frac{d_{r,m}}{2} \mid \mathbf{c_z}, \mathbf{c_b}, \mathcal{P}\right]}. \tag{6.9}$$

Proof. We define the key $\mathbf{k}_j$, for $0 \le j \le m$, to consist of the first j bits of $\mathbf{k}'$ and the last $m - j$ bits of $\mathbf{k}$. This means that $\mathbf{k}_0 = \mathbf{k}$, $\mathbf{k}_m = \mathbf{k}'$, and $\mathbf{k}_{j-1}$ differs from $\mathbf{k}_j$ at most on a single bit (the j-th bit).

First, we find a bound on $\frac{1}{2} \operatorname{tr} |\widehat{\rho}_{\mathbf{k}_{j-1}} - \widehat{\rho}_{\mathbf{k}_j}|$: since $\mathbf{k}_{j-1}$ differs from $\mathbf{k}_j$ at most on a single bit (the j-th bit, given by the formula $\mathbf{x} \cdot v_{r+j}$), we can use the same proof that gave us Equation (6.7), attaching the other (identical) key bits to $\boldsymbol{\xi}$ of the original proof; and we find that

$$\frac{1}{2} \operatorname{tr} |\widehat{\rho}_{\mathbf{k}_{j-1}} - \widehat{\rho}_{\mathbf{k}_j}| \le 2 \sqrt{\Pr\left[|\widetilde{\mathbf{C}}_{\mathrm{I}}| \ge \frac{d_j}{2} \mid \mathbf{c_z}, \mathbf{c_b}, \mathcal{P}\right]}, \tag{6.10}$$

where we define d_j as $d_{\mathrm{H}}(v_{r+j}, V_j')$, and $V_j' \triangleq \operatorname{Span}\{v_1, v_2, \ldots, v_{r+j-1}, v_{r+j+1}, \ldots, v_{r+m}\}$.

Now we notice that d_j is the Hamming distance between v_{r+j} and some vector in V_j', which means that $d_j = \left|\sum_{i=1}^{r+m} a_i v_i\right|$ with $a_i \in \mathbf{F}_2$ and $a_{r+j} \neq 0$. The properties of Hamming distance assure us that d_j is at least $d_{\mathrm{H}}(v_{r'+1}, V_{r'})$ for some $r \le r' < r + m$. Therefore, we find that $d_{r,m} = \min_{r \le r' < r+m} d_{\mathrm{H}}(v_{r'+1}, V_{r'}) \le d_j$.

The result $d_{r,m} \le d_j$ implies that if $|\widetilde{\mathbf{C}}_{\mathrm{I}}| \ge \frac{d_j}{2}$ then $|\widetilde{\mathbf{C}}_{\mathrm{I}}| \ge \frac{d_{r,m}}{2}$. Therefore, Equation (6.10) implies

$$\frac{1}{2} \operatorname{tr} |\widehat{\rho}_{\mathbf{k}_{j-1}} - \widehat{\rho}_{\mathbf{k}_j}| \le 2 \sqrt{\Pr\left[|\widetilde{\mathbf{C}}_{\mathrm{I}}| \ge \frac{d_{r,m}}{2} \mid \mathbf{c_z}, \mathbf{c_b}, \mathcal{P}\right]}. \tag{6.11}$$

Now we use the triangle inequality for norms to find

$$\begin{aligned}\frac{1}{2}\operatorname{tr}|\widehat{\rho}_{\mathbf{k}}-\widehat{\rho}_{\mathbf{k}'}| = \frac{1}{2}\operatorname{tr}|\widehat{\rho}_{\mathbf{k}_0}-\widehat{\rho}_{\mathbf{k}_m}| &\le \sum_{j=1}^{m}\frac{1}{2}\operatorname{tr}|\widehat{\rho}_{\mathbf{k}_{j-1}}-\widehat{\rho}_{\mathbf{k}_j}| \\ &\le 2m\sqrt{\Pr\left[|\widetilde{\mathbf{C}_{\mathrm{I}}}| \ge \frac{d_{r,m}}{2} \ \mid\ \mathbf{c_z},\mathbf{c_b},\mathcal{P}\right]},\end{aligned} \tag{6.12}$$

as we wanted. □

We would now like to bound the expected value (namely, the average value) of the trace distance between two states of Eve corresponding to two final keys. However, we should take into account that if the test fails, no final key is generated, in which case we define the distance to be 0. We thus define the random variable $\Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}')$ for any two final keys $\mathbf{k},\mathbf{k}'$:

$$\Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}'|\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b}) \triangleq \begin{cases}\frac{1}{2}\operatorname{tr}|\widehat{\rho}_{\mathbf{k}}-\widehat{\rho}_{\mathbf{k}'}| & \text{if } \dfrac{|\mathbf{c_z}|}{n_z}\le p_{a,z} \text{ and } \dfrac{|\mathbf{c_b}|}{n_x}\le p_{a,x} \\ 0 & \text{otherwise}\end{cases}. \tag{6.13}$$

We need to bound the expected value $\langle\Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}')\rangle$, that is given by:

$$\langle\Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}')\rangle = \sum_{\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b}} \Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}'|\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b})\cdot\Pr(\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b}). \tag{6.14}$$

(In Subsection 6.3.6 we prove that this expected value is indeed the quantity we need to bound for proving fully composable security, defined in Subsection 2.3.1.)

Theorem 6.2.

$$\langle\Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}')\rangle \le 2m\sqrt{\Pr\left[\left(\frac{|\widetilde{\mathbf{C}_{\mathrm{I}}}|}{n}\ge\frac{d_{r,m}}{2n}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z}\le p_{a,z}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x}\le p_{a,x}\right)\right]}, \tag{6.15}$$

where $\frac{|\widetilde{\mathbf{C}_{\mathrm{I}}}|}{n}$ is a random variable whose value is the error rate on the INFO bits if they had been encoded in the x basis, $\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z}$ is a random variable whose value is the error rate on the TEST-Z bits, and $\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x}$ is a random variable whose value is the error rate on the TEST-X bits.

Proof. We use the convexity of x^2, namely, the fact that for all $\{p_i\}_i$ satisfying $p_i\ge 0$ and $\sum_i p_i = 1$, it holds that $\left(\sum_i p_i x_i\right)^2 \le \sum_i p_i x_i^2$. We find that:

$$\begin{aligned}&\langle\Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}')\rangle^2 \\ &= \left[\sum_{\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b}} \Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}'|\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b})\cdot\Pr(\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b})\right]^2 \qquad \text{(by (6.14))}\end{aligned}$$

$$
\begin{aligned}
&\le \sum_{\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b}} \left(\Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}'|\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b})\right)^2 \cdot \Pr(\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b}) && \text{(by convexity of } x^2\text{)} \\
&= \sum_{\mathcal{P},\boldsymbol{\xi},\frac{|\mathbf{c_z}|}{n_z}\le p_{a,z},\frac{|\mathbf{c_b}|}{n_x}\le p_{a,x}} \left(\tfrac{1}{2}\operatorname{tr}|\widehat{\rho}_{\mathbf{k}}-\widehat{\rho}_{\mathbf{k}'}|\right)^2 \cdot \Pr(\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b}) && \text{(by (6.13))} \\
&\le 4m^2 \cdot \sum_{\mathcal{P},\boldsymbol{\xi},\frac{|\mathbf{c_z}|}{n_z}\le p_{a,z},\frac{|\mathbf{c_b}|}{n_x}\le p_{a,x}} \Pr\left[|\widetilde{\mathbf{C}}_{\mathrm{I}}| \ge \tfrac{d_{r,m}}{2} \mid \mathbf{c_z},\mathbf{c_b},\mathcal{P}\right] \cdot \Pr(\mathcal{P},\boldsymbol{\xi},\mathbf{c_z},\mathbf{c_b}) && \text{(by (6.9))} \\
&= 4m^2 \cdot \sum_{\mathcal{P},\frac{|\mathbf{c_z}|}{n_z}\le p_{a,z},\frac{|\mathbf{c_b}|}{n_x}\le p_{a,x}} \Pr\left[|\widetilde{\mathbf{C}}_{\mathrm{I}}| \ge \tfrac{d_{r,m}}{2} \mid \mathbf{c_z},\mathbf{c_b},\mathcal{P}\right] \cdot \Pr(\mathcal{P},\mathbf{c_z},\mathbf{c_b}) \\
&= 4m^2 \cdot \sum_{\mathcal{P}} \Pr\left[\left(|\widetilde{\mathbf{C}}_{\mathrm{I}}| \ge \tfrac{d_{r,m}}{2}\right) \wedge \left(\tfrac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z} \le p_{a,z}\right) \wedge \left(\tfrac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \le p_{a,x}\right) \mid \mathcal{P}\right] \cdot \Pr(\mathcal{P}) \\
&= 4m^2 \cdot \Pr\left[\left(|\widetilde{\mathbf{C}}_{\mathrm{I}}| \ge \tfrac{d_{r,m}}{2}\right) \wedge \left(\tfrac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z} \le p_{a,z}\right) \wedge \left(\tfrac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \le p_{a,x}\right)\right], && (6.16)
\end{aligned}
$$

as we wanted. □

6.3.4 Proof of Security

Following [BGM09] and [BBBMR06], we choose matrices P_{C} and P_{K} such that the inequality $\frac{d_{r,m}}{2n} > p_{a,x} + \epsilon$ is satisfied for some ϵ (we will explain in Subsection 6.3.7 why this is possible). This means that

$$
\begin{aligned}
&\Pr\left[\left(\tfrac{|\widetilde{\mathbf{C}}_{\mathrm{I}}|}{n} \ge \tfrac{d_{r,m}}{2n}\right) \wedge \left(\tfrac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z} \le p_{a,z}\right) \wedge \left(\tfrac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \le p_{a,x}\right)\right] \\
&\quad \le \Pr\left[\left(\tfrac{|\widetilde{\mathbf{C}}_{\mathrm{I}}|}{n} > p_{a,x} + \epsilon\right) \wedge \left(\tfrac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \le p_{a,x}\right)\right]. \qquad (6.17)
\end{aligned}
$$

We will now prove the right-hand-side of (6.17) to be exponentially small in n.

As said earlier, the random variable $\widetilde{\mathbf{C}}_{\mathrm{I}}$ corresponds to the bit string of errors on the INFO bits if they had been encoded in the x basis. The TEST-X bits are also encoded in the x basis, and the random variable $\mathbf{C}_{\mathrm{T_X}}$ corresponds to the bit string of errors on these bits. Therefore, we can treat the selection of the indexes of the n INFO bits and the n_x TEST-X bits as a random sampling (after the numbers n, n_z, and n_x *and* the indexes of the TEST-Z bits have all already been chosen) and use Hoeffding's theorem and Corollary 2.2 (that are described in Section 2.6).

Applying Corollary 2.2, we get:

$$
\Pr\left[\left(\frac{|\widetilde{\mathbf{C}}_{\mathrm{I}}|}{n} > p_{a,x} + \epsilon\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \le p_{a,x}\right)\right] \le e^{-2\left(\frac{n_x}{n+n_x}\right)^2 n\epsilon^2}. \qquad (6.18)
$$

In the above discussion, we have actually proved the following Theorem:

Theorem 6.3. *Let us be given $\delta > 0$, $R > 0$, and, for infinitely many values of n, a family $\{v_1^n, \dots, v_{r_n+m_n}^n\}$ of linearly independent vectors in $\mathbf{F}_2^n$ such that $\delta < \frac{d_{r_n,m_n}}{n}$ and $\frac{m_n}{n} \le R$. Then for any $p_{a,z}, p_{a,x} > 0$ and $\epsilon_{\mathrm{sec}} > 0$ such that $p_{a,x} + \epsilon_{\mathrm{sec}} \le \frac{\delta}{2}$, and for any $n, n_z, n_x > 0$ and two m_n-bit final keys $\mathbf{k}, \mathbf{k}'$, the distance between Eve's states* *corresponding to* $\mathbf{k}$ *and* $\mathbf{k}'$ *satisfies the following bound:*

$$\langle \Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}')\rangle \leq 2R\,n e^{-\left(\frac{n_x}{n+n_x}\right)^2 n\epsilon_{\mathrm{sec}}^2}. \tag{6.19}$$

In Subsection 6.3.7 we explain why the vectors required by this Theorem exist.

We note that the quantity $\langle \Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}')\rangle$ bounds the expected values of the Shannon Distinguishability and of the mutual information between Eve and the final key, as done in [BGM09] and [BBBMR06], which is sufficient for proving non-composable security; but it also avoids composability problems: Eve is not required to measure immediately after the protocol ends, but she is allowed to wait until she gets more information. In Subsection 6.3.6 we use this bound for proving a fully composable security.

6.3.5 Reliability

Security itself is not sufficient; we also need the key to be reliable (namely, to be the same for Alice and Bob). This means that we should make sure that the number of errors on the INFO bits is less than the maximal number of errors that can be corrected by the error-correcting code. We demand that our error-correcting code can correct $n(p_{a,z}+\epsilon_{\mathrm{rel}})$ errors (we explain in Subsection 6.3.7 why this demand is satisfied). Therefore, reliability of the final key with exponentially small probability of failure is guaranteed by the following inequality: (as said, $\mathbf{C}_{\mathrm{I}}$ corresponds to the actual bit string of errors on the INFO bits in the protocol, when they are encoded in the z basis)

$$\Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} > p_{a,z}+\epsilon_{\mathrm{rel}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z} \leq p_{a,z}\right)\right] \leq e^{-2\left(\frac{n_z}{n+n_z}\right)^2 n\epsilon_{\mathrm{rel}}^2}. \tag{6.20}$$

This inequality is proved by an argument similar to the one used in Subsection 6.3.4: the selection of the indexes of the INFO bits and the TEST-Z bits is a random partition of $n+n_z$ bits into two subsets of sizes n and n_z, respectively (assuming that the indexes of the TEST-X bits have already been chosen), and thus it corresponds to Hoeffding's sampling used for Corollary 2.2.

6.3.6 Proof of Fully Composable Security

We now prove that the BB84-INFO-z protocol satisfies the definition of composable security for a QKD protocol: namely, that it satisfies Equation (2.2) presented in Subsection 2.3.1. We prove that the expression $\frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}|$ is exponentially small in n, where ρ_{ABE} is the actual joint state of Alice, Bob, and Eve; ρ_{U} is an ideal (random, secret, and shared) key distributed to Alice and Bob; and ρ_{E} is the partial trace of ρ_{ABE} over the system AB (see Subsection 1.3.2).

To make reading easier, we use the following notations, where $\mathbf{i}$ is the bit string sent by Alice, $\mathbf{i}^{\mathrm{B}}$ is the bit string received by Bob, and $\mathbf{c} = \mathbf{i} \oplus \mathbf{i}^{\mathrm{B}}$ is the string of errors:

$$\begin{aligned}
\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}} &\triangleq \left(\mathbf{i}_{\mathbf{z}}, \mathbf{i}_{\mathbf{b}}, \mathbf{i}_{\mathbf{z}}^{\mathrm{B}}, \mathbf{i}_{\mathbf{b}}^{\mathrm{B}}\right), && (6.21)\\
\mathbf{T} &\triangleq \begin{cases} 1 & \text{if } \frac{|\mathbf{c}_{\mathbf{z}}|}{n_z} \le p_{a,z} \text{ and } \frac{|\mathbf{c}_{\mathbf{b}}|}{n_x} \le p_{a,x} \\ 0 & \text{otherwise} \end{cases}. && (6.22)
\end{aligned}$$

In other words, $\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}$ consists of all TEST-Z and TEST-X bits of Alice and Bob; and $\mathbf{T}$ is the random variable representing the result of the test.

According to the above definitions, the states ρ_{ABE} and ρ_{U} are

$$\begin{aligned}
\rho_{\mathrm{ABE}} &= \sum_{\mathbf{i}, \mathbf{i}^{\mathrm{B}}, \mathcal{P} | \mathbf{T}=1} \Pr\left(\mathbf{i}, \mathbf{i}^{\mathrm{B}}, \mathcal{P}\right) \cdot |\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}|_{\mathrm{A}} \otimes |\mathbf{k}'\rangle_{\mathrm{B}}\langle\mathbf{k}'|_{\mathrm{B}} \\
&\otimes \left(\rho_{\mathbf{x}, \mathbf{x}^{\mathrm{B}}}^{\mathbf{b}'}\right)_{\mathrm{E}} \otimes |\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}|_{\mathrm{C}}, && (6.23)\\
\rho_{\mathrm{U}} &= \frac{1}{2^m} \sum_{\mathbf{k}} |\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}|_{\mathrm{A}} \otimes |\mathbf{k}\rangle_{\mathrm{B}}\langle\mathbf{k}|_{\mathrm{B}}, && (6.24)
\end{aligned}$$

where $\left(\rho_{\mathbf{x}, \mathbf{x}^{\mathrm{B}}}^{\mathbf{b}'}\right)_{\mathrm{E}}$ is defined to be Eve's quantum state if Alice sends the INFO string $\mathbf{x} = \mathbf{i}_{\mathbf{s}}$ in the bases $\mathbf{b}' = \mathbf{b}_{\mathbf{s}}$ and Bob gets the INFO string $\mathbf{x}^{\mathrm{B}} = \mathbf{i}_{\mathbf{s}}^{\mathrm{B}}$. All the other states actually represent classical information: subsystems A and B represent the final keys held by Alice ($\mathbf{k} = \mathbf{x} P_{\mathrm{K}}^{\mathrm{T}}$) and Bob ($\mathbf{k}'$, that is obtained from $\mathbf{x}^{\mathrm{B}}$, $\boldsymbol{\xi} = \mathbf{x} P_{\mathrm{C}}^{\mathrm{T}}$, and P_{K}), and subsystem C represents the information published in the unjammable classical channel during the protocol (this information is known to Alice, Bob, and Eve)—namely, $\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}$ (all the test bits), $\mathcal{P}$ (the partition), and $\boldsymbol{\xi} = \mathbf{x} P_{\mathrm{C}}^{\mathrm{T}}$ (the syndrome).

We note that in the definition of ρ_{ABE}, we sum only over events in which the test is *passed* (namely, in which the protocol is not aborted by Alice and Bob): in such cases, an m-bit key is generated. The cases in which the protocol aborts do not exist in the sum—namely, they are represented by the zero operator, as required by the definition of composable security (see Subsection 2.3.1 and [Ren08, Subsection 6.1.2]). Thus, ρ_{ABE} is a non-normalized state, and $\mathrm{tr}(\rho_{\mathrm{ABE}})$ is the probability that the test is passed.

To help us bound the trace distance, we define the following intermediate state:

$$\begin{aligned}
\rho'_{\mathrm{ABE}} &\triangleq \sum_{\mathbf{i}, \mathbf{i}^{\mathrm{B}}, \mathcal{P} | \mathbf{T}=1} \Pr\left(\mathbf{i}, \mathbf{i}^{\mathrm{B}}, \mathcal{P}\right) \cdot |\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}|_{\mathrm{A}} \otimes |\mathbf{k}\rangle_{\mathrm{B}}\langle\mathbf{k}|_{\mathrm{B}} \\
&\otimes \left(\rho_{\mathbf{x}, \mathbf{x}^{\mathrm{B}}}^{\mathbf{b}'}\right)_{\mathrm{E}} \otimes |\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}|_{\mathrm{C}}. && (6.25)
\end{aligned}$$

This state is identical to ρ_{ABE}, except that Bob holds the Alice's final key ($\mathbf{k}$) instead of his own calculated final key ($\mathbf{k}'$). In particular, the similarity between ρ_{ABE} and ρ'_{ABE} means, by definition, that $\rho_{\mathrm{E}} \triangleq \mathrm{tr}_{\mathrm{AB}}\left(\rho_{\mathrm{ABE}}\right)$ and $\rho'_{\mathrm{E}} \triangleq \mathrm{tr}_{\mathrm{AB}}\left(\rho'_{\mathrm{ABE}}\right)$ are the same state: namely, $\rho_{\mathrm{E}} = \rho'_{\mathrm{E}}$.

Proposition 6.4. *Under the same conditions as Theorem 6.3, it holds that*

$$\frac{1}{2}\operatorname{tr}\left|\rho'_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}\right| \leq 2Rne^{-\left(\frac{n_x}{n+n_x}\right)^2 n\epsilon_{\mathrm{sec}}^2}, \tag{6.26}$$

for ρ'_{ABE} and ρ_{U} defined above and for the partial trace $\rho_{\mathrm{E}} \triangleq \operatorname{tr}_{\mathrm{AB}}\left(\rho_{\mathrm{ABE}}\right)$.

Proof. We notice that in ρ'_{ABE}, the only factors depending directly on $\mathbf{x}$ and $\mathbf{x}^{\mathrm{B}}$ (and not only on $\mathbf{k}$ and $\boldsymbol{\xi}$) are the probability $\Pr\left(\mathbf{i}, \mathbf{i}^{\mathrm{B}}, \mathcal{P}\right)$ and Eve's state $\left(\rho^{\mathbf{b}'}_{\mathbf{x},\mathbf{x}^{\mathrm{B}}}\right)_{\mathrm{E}}$. The probability can be reformulated as:

$$\begin{aligned}
\Pr\left(\mathbf{i}, \mathbf{i}^{\mathrm{B}}, \mathcal{P}\right) &= \Pr\left(\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right) \cdot \Pr\left(\mathbf{k} \mid \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right) \\
&\cdot \Pr\left(\mathbf{x} \mid \mathbf{k}, \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right) \cdot \Pr\left(\mathbf{x}^{\mathrm{B}} \mid \mathbf{x}, \mathbf{k}, \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right) \\
&= \Pr\left(\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right) \cdot \frac{1}{2^m} \cdot \frac{1}{2^{n-r-m}} \\
&\cdot \Pr\left(\mathbf{x}^{\mathrm{B}} \mid \mathbf{x}, \mathbf{k}, \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right).
\end{aligned} \tag{6.27}$$

(Because all the possible n-bit values of $\mathbf{x}$ have the same probability, $\frac{1}{2^n}$; and because all the $r+m$ rows of the matrices P_{C} and P_{K} are linearly independent, so there are exactly 2^{n-r-m} values of $\mathbf{x}$ corresponding to each specific pair $(\boldsymbol{\xi}, \mathbf{k})$.)

Therefore, the state ρ'_{ABE} takes the following form:

$$\begin{aligned}
\rho'_{\mathrm{ABE}} &= \frac{1}{2^m} \sum_{\mathbf{k}, \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi} \mid \mathbf{T}=1} \Pr\left(\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right) \cdot |\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}|_{\mathrm{A}} \otimes |\mathbf{k}\rangle_{\mathrm{B}}\langle\mathbf{k}|_{\mathrm{B}} \\
&\otimes \left[\frac{1}{2^{n-r-m}} \sum_{\mathbf{x}, \mathbf{x}^{\mathrm{B}} \left| \substack{\mathbf{x}P_{\mathrm{C}}^{\mathrm{T}} = \boldsymbol{\xi} \\ \mathbf{x}P_{\mathrm{K}}^{\mathrm{T}} = \mathbf{k}} \right.} \Pr\left(\mathbf{x}^{\mathrm{B}} \mid \mathbf{x}, \mathbf{k}, \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right) \cdot \left(\rho^{\mathbf{b}'}_{\mathbf{x},\mathbf{x}^{\mathrm{B}}}\right)_{\mathrm{E}}\right] \\
&\otimes |\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}|_{\mathrm{C}} \\
&= \frac{1}{2^m} \sum_{\mathbf{k}, \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi} \mid \mathbf{T}=1} \Pr\left(\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right) \cdot |\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}|_{\mathrm{A}} \otimes |\mathbf{k}\rangle_{\mathrm{B}}\langle\mathbf{k}|_{\mathrm{B}} \\
&\otimes (\widehat{\rho}_{\mathbf{k}})_{\mathrm{E}} \otimes |\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}|_{\mathrm{C}}.
\end{aligned} \tag{6.28}$$

($\widehat{\rho}_{\mathbf{k}}$ was defined in Equation (6.8).)

The partial trace $\rho'_{\mathrm{E}} = \operatorname{tr}_{\mathrm{AB}}\left(\rho'_{\mathrm{ABE}}\right)$, that (as proved above) is the same as ρ_{E}, is

$$\rho_{\mathrm{E}} = \rho'_{\mathrm{E}} = \frac{1}{2^m} \sum_{\mathbf{k}, \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi} \mid \mathbf{T}=1} \Pr\left(\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right) \cdot (\widehat{\rho}_{\mathbf{k}})_{\mathrm{E}} \otimes |\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}|_{\mathrm{C}}, \tag{6.29}$$

and the state $\rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}$ is

$$\rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}} = \frac{1}{2^{2m}} \sum_{\mathbf{k}, \mathbf{k}'', \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi} \mid \mathbf{T}=1} \Pr\left(\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right) \cdot |\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}|_{\mathrm{A}} \otimes |\mathbf{k}\rangle_{\mathrm{B}}\langle\mathbf{k}|_{\mathrm{B}}$$

$$\otimes \left(\widehat{\rho}_{\mathbf{k}''}\right)_{\mathrm{E}} \otimes |\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\rangle_{\mathrm{C}}\langle \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}|_{\mathrm{C}}.$$

By using the triangle inequality for norms, since ρ'_{ABE} and $\rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}$ are the same (except the difference between Eve's states, $(\widehat{\rho}_{\mathbf{k}})_{\mathrm{E}}$ and $(\widehat{\rho}_{\mathbf{k}''})_{\mathrm{E}}$), we get, by using the definition of $\langle \Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}'')\rangle$ (Equation (6.14)) and Theorem 6.3:

$$\begin{aligned}
\frac{1}{2}\operatorname{tr}\left|\rho'_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}\right| &\leq \frac{1}{2^{2m}} \sum_{\mathbf{k},\mathbf{k}'',\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}},\mathcal{P},\boldsymbol{\xi}|\mathbf{T}=1} \Pr\left(\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\right) \cdot \frac{1}{2}\operatorname{tr}\left|(\widehat{\rho}_{\mathbf{k}})_{\mathrm{E}} - (\widehat{\rho}_{\mathbf{k}''})_{\mathrm{E}}\right| \\
&= \frac{1}{2^{2m}} \sum_{\mathbf{k},\mathbf{k}''} \langle \Delta_{\mathrm{Eve}}^{(p_{a,z},p_{a,x})}(\mathbf{k},\mathbf{k}'')\rangle \\
&\leq 2R\, n e^{-\left(\frac{n_x}{n+n_x}\right)^2 n\epsilon_{\mathrm{sec}}^2}, \qquad (6.30)
\end{aligned}$$

as we wanted. □

We still have to bound the following difference:

$$\begin{aligned}
\rho_{\mathrm{ABE}} - \rho'_{\mathrm{ABE}} &= \sum_{\mathbf{i},\mathbf{i}^{\mathrm{B}},\mathcal{P}|\mathbf{T}=1} \Pr\left(\mathbf{i}, \mathbf{i}^{\mathrm{B}}, \mathcal{P}\right) \\
&\cdot\ |\mathbf{k}\rangle_{\mathrm{A}}\langle \mathbf{k}|_{\mathrm{A}} \otimes \left[|\mathbf{k}'\rangle_{\mathrm{B}}\langle \mathbf{k}'|_{\mathrm{B}} - |\mathbf{k}\rangle_{\mathrm{B}}\langle \mathbf{k}|_{\mathrm{B}}\right] \\
&\otimes \left(\rho_{\mathbf{x},\mathbf{x}^{\mathrm{B}}}^{\mathbf{b}'}\right)_{\mathrm{E}} \otimes |\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\rangle_{\mathrm{C}}\langle \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}|_{\mathrm{C}} \\
&= \Pr\left(\left(\mathbf{k} \neq \mathbf{k}'\right) \wedge (\mathbf{T} = 1)\right) \\
&\cdot \sum_{\mathbf{i},\mathbf{i}^{\mathrm{B}},\mathcal{P}} \Pr\left(\mathbf{i}, \mathbf{i}^{\mathrm{B}}, \mathcal{P} \mid \left(\mathbf{k} \neq \mathbf{k}'\right) \wedge (\mathbf{T} = 1)\right) \\
&\cdot\ |\mathbf{k}\rangle_{\mathrm{A}}\langle \mathbf{k}|_{\mathrm{A}} \otimes \left[|\mathbf{k}'\rangle_{\mathrm{B}}\langle \mathbf{k}'|_{\mathrm{B}} - |\mathbf{k}\rangle_{\mathrm{B}}\langle \mathbf{k}|_{\mathrm{B}}\right] \\
&\otimes \left(\rho_{\mathbf{x},\mathbf{x}^{\mathrm{B}}}^{\mathbf{b}'}\right)_{\mathrm{E}} \otimes |\mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}\rangle_{\mathrm{C}}\langle \mathbf{i}_{\mathcal{T}}^{\mathrm{AB}}, \mathcal{P}, \boldsymbol{\xi}|_{\mathrm{C}}. \qquad (6.31)
\end{aligned}$$

Because the trace distance between every two normalized states is bounded by 1, and because of the reliability proof in Subsection 6.3.5, we get:

$$\frac{1}{2}\operatorname{tr}\left|\rho_{\mathrm{ABE}} - \rho'_{\mathrm{ABE}}\right| \leq \Pr\left(\left(\mathbf{k} \neq \mathbf{k}'\right) \wedge (\mathbf{T} = 1)\right) \leq e^{-2\left(\frac{n_z}{n+n_z}\right)^2 n\epsilon_{\mathrm{rel}}^2}. \qquad (6.32)$$

(Because if $\mathbf{k} \neq \mathbf{k}'$, Alice and Bob have different final keys, and this means that the error correction stage did not succeed. According to the discussion in Subsection 6.3.5, this can happen only if there are too many errors in the information string—namely, if $\frac{|\mathbf{C}_{\mathrm{I}}|}{n} > p_{a,z} + \epsilon_{\mathrm{rel}}$.)

To sum up, we get the following bound:

$$\begin{aligned}
\frac{1}{2}\operatorname{tr}\left|\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}\right| &\leq \frac{1}{2}\operatorname{tr}\left|\rho_{\mathrm{ABE}} - \rho'_{\mathrm{ABE}}\right| + \frac{1}{2}\operatorname{tr}\left|\rho'_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}\right| \\
&\leq e^{-2\left(\frac{n_z}{n+n_z}\right)^2 n\epsilon_{\mathrm{rel}}^2} + 2R\, n e^{-\left(\frac{n_x}{n+n_x}\right)^2 n\epsilon_{\mathrm{sec}}^2}. \qquad (6.33)
\end{aligned}$$

This bound is exponentially small in n. Thus, we have proved composable security of

BB84-INFO-z.

6.3.7 Security, Reliability, and Error Rate Threshold

According to Theorem 6.3 and to the discussion in Subsection 6.3.5, to get both security and reliability we only need vectors $\{v_1^n, \ldots, v_{r_n+m_n}^n\}$ satisfying both the conditions of the Theorem (distance $\frac{d_{r_n,m_n}}{2n} > \frac{\delta}{2} \geq p_{a,x} + \epsilon_{\text{sec}}$) and the reliability condition (the ability to correct $n(p_{a,z} + \epsilon_{\text{rel}})$ errors). Such families were proven to exist in Appendix E of [BBBMR06], giving the following upper bound on the bit-rate:

$$R_{\text{secret}} \triangleq \frac{m}{n} < 1 - H_2(2p_{a,x} + 2\epsilon_{\text{sec}}) - H_2\left(p_{a,z} + \epsilon_{\text{rel}} + \frac{1}{n}\right), \tag{6.34}$$

where $H_2(x) \triangleq -x\log_2(x) - (1-x)\log_2(1-x)$.

Note that we use here the error thresholds $p_{a,x}$ for security and $p_{a,z}$ for reliability. This is possible, because in [BBBMR06] these conditions (security and reliability) on the codes are discussed separately.

To get the asymptotic error rate thresholds, we require $R_{\text{secret}} > 0$, and we get the condition:

$$H_2(2p_{a,x} + 2\epsilon_{\text{sec}}) + H_2\left(p_{a,z} + \epsilon_{\text{rel}} + \frac{1}{n}\right) < 1. \tag{6.35}$$

The secure asymptotic error rate thresholds zone is shown in Figure 6.1 (it is below the curve), assuming that $\frac{1}{n}$ is negligible. Note the trade-off between the error rates $p_{a,z}$ and $p_{a,x}$. Also note that in the case $p_{a,z} = p_{a,x}$, we get the same threshold as BB84 ([BBBMR06] and [BGM09]), which is 7.56%.

6.4 Conclusion

In this chapter, we have proved the BB84-INFO-z protocol to be fully secure against collective attacks. We have discovered that the results of BB84 hold very similarly for BB84-INFO-z, with only two exceptions:

1. The error rates must be *separately* checked to be below the thresholds $p_{a,z}$ and $p_{a,x}$ for the TEST-Z and TEST-X bits, respectively, while in BB84 the error rate threshold p_a applies to all the TEST bits together.

2. The exponents of Eve's information (security) and of the failure probability of the error-correcting code (reliability) are different than in [BGM09], because different numbers of test bits are now allowed (n_z and n_x are arbitrary). This implies that the exponents may decrease more slowly (or more quickly) as a function of n. However, if we choose $n_z = n_x = n$ (thus sending $N = 3n$ qubits from Alice to Bob), then we get exactly the same exponents as in [BGM09].

The asymptotic error rate thresholds found in this chapter allow us to tolerate a higher threshold for a specific basis (say, the x basis) if we demand a lower threshold

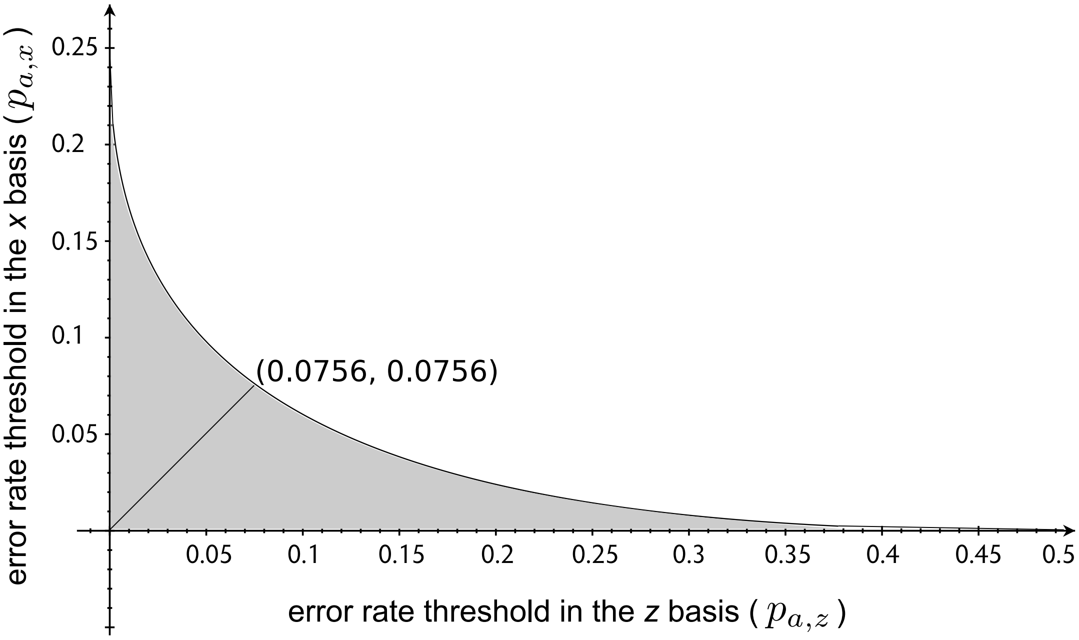

Figure 6.1: **The secure asymptotic error rates zone for BB84-INFO-z** (below the curve)

for the other basis (z). If we choose the same error rate threshold for both bases, then the asymptotic bound is 7.56%, exactly the bound found for BB84 in [BBBMR06] and [BGM09].

We conclude that even if we change the BB84 protocol to have INFO bits only in the z basis, this does not harm its security and reliability (at least against collective attacks). This does not even change the asymptotic error rate threshold. The only drawbacks of this change are the need to check the error rate for the two bases separately, and the need to either send more qubits ($3n$ qubits in total, rather than $2n$) or get a slower exponential decrease of the exponents required for security and reliability.

We thus find that the feature of BB84, that both bases are used for information, is not very important for security and reliability, and that BB84-INFO-z (that lacks this feature) is almost as useful as BB84. This may have important implications on security and reliability of other protocols that, too, use only one basis for information qubits, such as the "three-state protocol" [Mor98] and some two-way protocols [BKM07, ZQLWL09].

We also present a better approach for the proof, that uses the quantum distance between two states rather than the classical information. In [BGM09, BBBGM02, BBBMR06], the classical mutual information between Eve's information (after an optimal measurement) and the final key was calculated (by using the trace distance between two quantum states); although we should note that in [BGM09, BBBMR06], the trace distance was used for the proof of security of a single bit of the final key even when all other bits are given to Eve, and only the last stages of the proof discussed bounding the classical mutual information. In this chapter, on the other hand, we use the trace distance between the two quantum states until the end of the proof, which allows us to prove fully composable security.

Therefore, our proof shows the fully composable security of BB84-INFO-z against collective attacks (and, in particular, security even if Eve keeps her quantum states until she gets more information when Alice and Bob use the key, rather than measuring them at the end of the protocol); and a very similar approach can be applied to [BGM09], immediately proving the composable security of BB84 against collective attacks. Our proof also makes a step towards making the security proof in [BBBMR06] (security proof of BB84 against joint attacks) prove the *composable* security of BB84 against joint attacks, a proof fully achieved in Chapter 7.

Our results show that the BB84-INFO-z protocol can securely be used for distributing a secret key; the security is of an ideal implementation and against an adversary limited to collective attacks (a generalization to the most general attacks (joint attacks), by using the methods of [BBBMR06], is proposed in Chapter 7). Moreover, security of the final key is universally composable, which means that the key may be used for any cryptographic purpose without harming security, even if Eve keeps her quantum states and makes optimal use of any information she gets in the future.

The techniques described in our proof may be applied in the future for proving security of other protocols by using similar methods, and, in particular, for proving security of other QKD protocols that use only one basis for the information bits, such as [Mor98, BKM07, ZQLWL09] mentioned above.

Chapter 7

Composable Security of Generalized BB84 Protocols Against General (Joint) Attacks

In this chapter, we present a fully composable security proof of "generalized BB84" QKD protocols against joint attacks (namely, against the most general theoretical attacks, as described in Subsection 2.3.2). The protocols for which we prove security are the BB84-INFO-z protocol (Subsection 7.3.1), the standard BB84 protocol (Subsection 7.3.2), the "efficient BB84" protocol (Subsection 7.3.3), and the "modified efficient BB84" protocol (Subsection 7.3.4). The proof uses BBBMR's security approach, that is described in Subsection 2.3.3.

This chapter is based on a paper being prepared by Michel Boyer and Rotem Liss[1].

7.1 Full Definition of the Generalized BB84 Protocols

The protocols for which we prove security in this chapter belong to a generalized class of BB84-like protocols. Below we formally define this general class of protocols. Some of the details in this definition are decided by each specific protocol, but most of the details are shared by all the protocols. See Section 2.7 for an explanation of the notation of bit strings ($\mathbf{s}$, $\mathbf{b}$, etc.), and see Section 1.1 for an explanation of the notations $|0\rangle_0, |1\rangle_0, |0\rangle_1, |1\rangle_1$.

1. Before the protocol begins, Alice and Bob choose some shared (and public) parameters: the numbers N and n, the sets B and $\{S_\mathbf{b}\}_{\mathbf{b} \in B}$ and probability distributions over them (decided by the specific protocol) that will control the choice of the bit strings $\mathbf{b}, \mathbf{s} \in \mathbf{F}_2^N$, the testing function T (decided by the specific protocol), the $r \times n$ parity check matrix P_C (corresponding to a linear error-correcting code C), and the $m \times n$ privacy amplification matrix P_K (representing a

[1] This paper is in preparation.

linear key-generation function). It is required that *all* $r+m$ rows of the matrices P_{C} and P_{K} put together are linearly independent.

Formally, for choosing the sets B and $\{S_{\mathbf{b}}\}_{\mathbf{b}\in B}$ and the corresponding probability distributions, Alice and Bob should choose the set $B \subseteq \mathbf{F}_2^N$ of basis strings, the probabilities $\Pr(\mathbf{b})$ for all $\mathbf{b} \in B$, the sets $S_{\mathbf{b}} \subseteq \mathbf{F}_2^N$ of $\mathbf{s}$ strings for all $\mathbf{b} \in B$, and the probabilities $\Pr(\mathbf{s} \mid \mathbf{b})$ for all $\mathbf{b} \in B$ and $\mathbf{s} \in S_{\mathbf{b}}$. We require that $|\mathbf{s}| = n$ for all $\mathbf{s} \in S_{\mathbf{b}}$. The testing function $T : \mathbf{F}_2^{N-n} \times \mathbf{F}_2^{N-n} \times \mathbf{F}_2^N \to \{0,1\}$ must get $(\mathbf{i}_{\mathrm{T}} \oplus \mathbf{j}_{\mathrm{T}}, \mathbf{b}_{\mathrm{T}}, \mathbf{s})$ as inputs and give 0 or 1 as an output. In Section 7.3 we give examples of protocols and their formal definitions using these notations.

2. Alice randomly chooses an N-bit string $\mathbf{i} \in \mathbf{F}_2^N$, an N-bit string $\mathbf{b} \in B$, and an N-bit string $\mathbf{s} \in S_{\mathbf{b}}$ (that must satisfy $|\mathbf{s}| = n$), and sends the N qubit states $|i_1\rangle_{b_1}, |i_2\rangle_{b_2}, \ldots, |i_N\rangle_{b_N}$, one after the other, to Bob using the quantum channel. Bob keeps each received qubit in a quantum memory, not measuring it yet[2].

3. Alice sends to Bob over the classical channel the bit string $\mathbf{b} = b_1 \ldots b_N$. Bob measures each of the qubits he saved in the correct basis (namely, when measuring the i-th qubit, he measures it in the z basis if $b_i = 0$, and he measures it in the x basis if $b_i = 1$).

   The bit string measured by Bob is denoted by $\mathbf{j}$. The XOR of $\mathbf{i}$ and $\mathbf{j}$ is denoted $\mathbf{c} \triangleq \mathbf{i} \oplus \mathbf{j}$. If there is no noise and no eavesdropping, then $\mathbf{i} = \mathbf{j}$ (that is, $\mathbf{c} = \mathbf{0}$).

4. Alice sends $\mathbf{s}$ to Bob over the classical channel. The INFO bits (that will be used for creating the final key) are the n bits with $s_j = 1$, while the TEST bits (that will be used for testing) are the $N-n$ bits with $s_j = 0$. We denote the substrings of $\mathbf{i}, \mathbf{j}, \mathbf{c}, \mathbf{b}$ that correspond to the INFO bits by $\mathbf{i}_{\mathrm{I}}$, $\mathbf{j}_{\mathrm{I}}$, $\mathbf{c}_{\mathrm{I}}$, and $\mathbf{b}_{\mathrm{I}}$, respectively; and we denote the substrings of $\mathbf{i}, \mathbf{j}, \mathbf{c}, \mathbf{b}$ that correspond to the TEST bits by $\mathbf{i}_{\mathrm{T}}$, $\mathbf{j}_{\mathrm{T}}$, $\mathbf{c}_{\mathrm{T}}$, and $\mathbf{b}_{\mathrm{T}}$, respectively.

5. Alice and Bob both publish the bit values they have for all the TEST bits ($\mathbf{i}_{\mathrm{T}}$ and $\mathbf{j}_{\mathrm{T}}$, respectively), and they compute their XOR $\mathbf{c}_{\mathrm{T}} = \mathbf{i}_{\mathrm{T}} \oplus \mathbf{j}_{\mathrm{T}}$. They compute $T(\mathbf{c}_{\mathrm{T}}, \mathbf{b}_{\mathrm{T}}, \mathbf{s})$: if it is 0, they abort the protocol; if it is 1, they continue the run of the protocol.

6. The values of the remaining n bits (the INFO bits, with s_j =1) are kept in secret by Alice and Bob. The bit string of Alice is $\mathbf{i}_{\mathrm{I}}$, the bit string of Bob is $\mathbf{j}_{\mathrm{I}}$, and their XOR is $\mathbf{c}_{\mathrm{I}}$.

[2] Here we assume that Bob has a quantum memory and can delay his measurement. In practical implementations, Bob usually cannot do that, but he is assumed to choose his own random basis string $\mathbf{b}'' \in B$ and measure in the bases it dictates. Later, Alice and Bob discard the qubits measured in the wrong basis. In that case, we need to assume that Alice sends more than N qubits, so that N qubits are finally detected by Bob and measured in the correct basis. In Appendix A of [BBBMR06] it is explained why this change of the protocol does not hurt security.

7. Alice sends to Bob the *syndrome* of $\mathbf{i}_\mathrm{I}$ (with respect to the error-correcting code C and to its corresponding parity check matrix P_C), that consists of r bits and is defined as $\boldsymbol{\xi} \triangleq \mathbf{i}_\mathrm{I} P_\mathrm{C}^\mathrm{T}$. By using $\boldsymbol{\xi}$, Bob corrects the errors in his $\mathbf{j}_\mathrm{I}$ string (so that it is the same as $\mathbf{i}_\mathrm{I}$).

8. The final key consists of m bits and is defined as $\mathbf{k} \triangleq \mathbf{i}_\mathrm{I} P_\mathrm{K}^\mathrm{T}$. Both Alice and Bob compute it.

7.2 Bound on the Security Definition for the Generalized BB84 Protocols

7.2.1 The Hypothetical "Inverted-INFO-Basis" Protocol

For the security proof, we use an alternative, hypothetical protocol, in which Alice sends to Bob the qubits after inverting the bases of the INFO bits (without changing the bases of the TEST bits). We call this protocol "hypothetical" because it is never actually used by Alice and Bob, and we do not perform any reduction to it (or from it), but we compute probabilities of certain events in the hypothetical protocol for use in our security bound. In particular, we use the error rate in the hypothetical protocol for bounding the trace distance in the security definition of the real protocol.

In the hypothetical protocol, Alice, Bob, and Eve do everything exactly as they would do in the real protocol, except that Alice and Bob use (and publish) the basis string $\mathbf{b}^0 \triangleq \mathbf{b} \oplus \mathbf{s}$ instead of $\mathbf{b}$: namely, they use the basis string $\mathbf{b}_\mathrm{T}$ for the TEST bits and the basis string $\overline{\mathbf{b}_\mathrm{I}}$ (the bitwise NOT of $\mathbf{b}_\mathrm{I}$) for the INFO bits.

Formally, this hypothetical protocol is defined by replacing Steps 2–3 of the original protocol (as described in Section 7.1) by the following steps:

2. Alice randomly chooses an N-bit string $\mathbf{i} \in \mathbf{F}_2^N$, an N-bit string $\mathbf{b} \in B$, and an N-bit string $\mathbf{s} \in S_\mathbf{b}$ (that must satisfy $|\mathbf{s}| = n$). Then, she computes the N-bit string $\mathbf{b}^0 \triangleq \mathbf{b} \oplus \mathbf{s}$, and sends the N qubit states $|i_1\rangle_{b_1^0}, |i_2\rangle_{b_2^0}, \ldots, |i_N\rangle_{b_N^0}$, one after the other, to Bob using the quantum channel. Bob keeps each received qubit in a quantum memory, not measuring it yet.

3. Alice sends to Bob over the classical channel the bit string $\mathbf{b}^0 = b_1^0 \ldots b_N^0$. Bob measures each of the qubits he saved in the correct basis (namely, when measuring the i-th qubit, he measures it in the z basis if $b_i^0 = 0$, and he measures it in the x basis if $b_i^0 = 1$).

We notice that in this protocol, Alice *chooses* $\mathbf{b}$ and $\mathbf{s}$ in the same way as she would choose them in the real protocol, but *uses* (and sends to Bob for his use) $\mathbf{b}^0$ and $\mathbf{s}$ instead.

In the security proof, we will use the notation of $\Pr_{\text{inverted-INFO-basis}}$ for calculating the probability of a certain event assuming that Alice and Bob use the hypothetical protocol.

In particular, we note that $\Pr_{\text{inverted-INFO-basis}}(\cdot \mid \mathbf{b}, \mathbf{s})$ is a conditional probability on Alice *choosing* the bit strings $\mathbf{b}, \mathbf{s}$ (while she actually *uses* the basis string $\mathbf{b}^0$).

It should be noted that $\Pr_{\text{inverted-INFO-basis}}(\cdot \mid \mathbf{b}, \mathbf{s})$ is usually the same as $\Pr(\cdot \mid \mathbf{b}^0, \mathbf{s})$: namely, the hypothetical protocol given that Alice chooses $\mathbf{b}, \mathbf{s}$ (and thus uses $\mathbf{b}^0, \mathbf{s}$) is the same as the real protocol given that Alice chooses $\mathbf{b}^0, \mathbf{s}$. However, the second notation is not always well-defined, because it may be the case that $\mathbf{b} \in B$ while $\mathbf{b}^0 \notin B$, or that $\mathbf{s} \in S_{\mathbf{b}}$ while $\mathbf{s} \notin S_{\mathbf{b}^0}$; therefore, it may be the case that $\mathbf{b}^0$ is not an allowed basis string for the real protocol. In the standard BB84 protocol (see Subsection 7.3.2), such problems are impossible, and this is why [BBBMR06] uses the notation of $\Pr(\cdot \mid \mathbf{b}^0, \mathbf{s})$ instead of $\Pr_{\text{inverted-INFO-basis}}(\cdot \mid \mathbf{b}, \mathbf{s})$. However, in our chapter, we discuss generalized BB84 protocols, and we must use the notation of $\Pr_{\text{inverted-INFO-basis}}(\cdot \mid \mathbf{b}, \mathbf{s})$.[3]

7.2.2 The General Joint Attack of Eve

Before the beginning of the QKD protocol (and, thus, independently of $\mathbf{i}$, $\mathbf{b}$, and $\mathbf{s}$), Eve chooses some joint attack to perform. In a *joint attack*, all the qubits are attacked by using a shared giant probe (ancillary state) kept by Eve. Eve saves her probe in a quantum memory and can keep it indefinitely, even after the final key is used by Alice and Bob; and she can perform, at any time of her choice, an optimal measurement of her giant probe, chosen based on all the information she has at the time of the measurement (including the classical information sent during the protocol, and including the information she acquires when Alice and Bob use the key).

Given that Alice sends to Bob the state $|\mathbf{i}\rangle_{\mathbf{b}} \triangleq \otimes_{j=1}^{N} |i_j\rangle_{b_j}$ (namely, the N-bit string is $\mathbf{i}$ and the N-bit basis string is $\mathbf{b}$), Eve attaches a probe state $|0\rangle_{\mathrm{E}}$ and applies some unitary operator U of her choice to the compound system $|0\rangle_{\mathrm{E}}|\mathbf{i}\rangle_{\mathbf{b}}$. Then, Eve keeps to herself (in a quantum memory) her probe state, and she sends to Bob the N-qubit quantum state sent from Alice to Bob (which may have been modified due to her attack U).

The most general joint attack U of Eve is

$$U|0\rangle_{\mathrm{E}}|\mathbf{i}\rangle_{\mathbf{b}} = \sum_{\mathbf{j} \in \mathbf{F}_2^N} |E'_{\mathbf{i},\mathbf{j}}\rangle_{\mathbf{b}} |\mathbf{j}\rangle_{\mathbf{b}}, \tag{7.1}$$

where $|E'_{\mathbf{i},\mathbf{j}}\rangle_{\mathbf{b}}$ are non-normalized states in Eve's probe system. We note that

$$\langle E'_{\mathbf{i},\mathbf{j}} | E'_{\mathbf{i},\mathbf{j}}\rangle_{\mathbf{b}} = \Pr(\mathbf{j} \mid \mathbf{i}, \mathbf{b}, \mathbf{s}). \tag{7.2}$$

Writing the INFO and TEST bits of Alice and Bob separately ($\mathbf{i}_{\mathrm{T}}, \mathbf{i}_{\mathrm{I}}$ instead of $\mathbf{i}$, and $\mathbf{j}_{\mathrm{T}}, \mathbf{j}_{\mathrm{I}}$ instead of $\mathbf{j}$), we can denote $|E'_{\mathbf{i},\mathbf{j}}\rangle_{\mathbf{b}}$ by $|E'_{\mathbf{i}_{\mathrm{T}},\mathbf{i}_{\mathrm{I}},\mathbf{j}_{\mathrm{T}},\mathbf{j}_{\mathrm{I}}}\rangle_{\mathbf{b}}$.

In Subsection 3.4 of [BBBMR06], the notation of $|E_{\mathbf{i}_{\mathrm{I}},\mathbf{j}_{\mathrm{I}}}\rangle_{\mathbf{b},\mathbf{s}}$ is introduced. This notation is useful, because it treats $\mathbf{i}_{\mathrm{T}}$ and $\mathbf{j}_{\mathrm{T}}$ as constants (since they are ultimately

[3] It is also possible that $\Pr(\mathbf{b}, \mathbf{s}) \neq \Pr(\mathbf{b}^0, \mathbf{s})$, in which case the use of $\mathbf{b}^0, \mathbf{s}$ in the real protocol does not happen with the same probability as the use of $\mathbf{b}^0, \mathbf{s}$ in the hypothetical protocol.

published by Alice and Bob, and then they are known to Eve), assuming their values to be known. It is defined as

$$|E_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}\rangle_{\mathbf{b},\mathbf{s}} \triangleq \frac{1}{\sqrt{\Pr(\mathbf{j}_\mathrm{T} \mid \mathbf{i}_\mathrm{T},\mathbf{i}_\mathrm{I},\mathbf{b},\mathbf{s})}}|E'_{\mathbf{i}_\mathrm{T},\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{T},\mathbf{j}_\mathrm{I}}\rangle_\mathbf{b}. \tag{7.3}$$

We note that $|E_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}\rangle_{\mathbf{b},\mathbf{s}}$ also depends on the constants $\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T}$ (and not only on $\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I},\mathbf{b},\mathbf{s}$). According to Equations (3.22)–(3.23) of [BBBMR06], given that Alice sends $\mathbf{i}_\mathrm{I},\mathbf{i}_\mathrm{T},\mathbf{b},\mathbf{s}$ and that Bob measures $\mathbf{j}_\mathrm{T}$, the normalized state of Eve and Bob is

$$|\psi_{\mathbf{i}_\mathrm{I}}\rangle = \sum_{\mathbf{j}_\mathrm{I}\in\mathbf{F}_2^n} |E_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}\rangle_{\mathbf{b},\mathbf{s}}|\mathbf{j}_\mathrm{I}\rangle_\mathbf{b}, \tag{7.4}$$

and it also holds that

$$\langle E_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}|E_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}\rangle_{\mathbf{b},\mathbf{s}} = \Pr(\mathbf{j}_\mathrm{I} \mid \mathbf{i}_\mathrm{I},\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s}). \tag{7.5}$$

Let us define $\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}$ (which also depends on $\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T}$) to be the normalized state of Eve if Alice sends $\mathbf{i}_\mathrm{I},\mathbf{i}_\mathrm{T},\mathbf{b},\mathbf{s}$ and Bob measures $\mathbf{j}_\mathrm{I},\mathbf{j}_\mathrm{T}$. That is, $\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}$ is the normalization of $|E'_{\mathbf{i},\mathbf{j}}\rangle_\mathbf{b}$ and of $|E_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}\rangle_{\mathbf{b},\mathbf{s}}$, so

$$\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}} \triangleq \frac{|E_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}\rangle_{\mathbf{b},\mathbf{s}}\langle E_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}|}{\Pr(\mathbf{j}_\mathrm{I} \mid \mathbf{i}_\mathrm{I},\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s})} = \frac{|E'_{\mathbf{i},\mathbf{j}}\rangle_\mathbf{b}\langle E'_{\mathbf{i},\mathbf{j}}|}{\Pr(\mathbf{j} \mid \mathbf{i},\mathbf{b},\mathbf{s})}. \tag{7.6}$$

The state of Eve after her attack (tracing out Bob) is

$$\rho^{\mathbf{i}_\mathrm{I}} \triangleq \mathrm{tr}_\mathrm{Bob}(|\psi_{\mathbf{i}_\mathrm{I}}\rangle\langle\psi_{\mathbf{i}_\mathrm{I}}|) = \sum_{\mathbf{j}_\mathrm{I}\in\mathbf{F}_2^n} |E_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}\rangle_{\mathbf{b},\mathbf{s}}\langle E_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}| = \sum_{\mathbf{j}_\mathrm{I}\in\mathbf{F}_2^n} \Pr(\mathbf{j}_\mathrm{I} \mid \mathbf{i}_\mathrm{I},\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s})\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}, \tag{7.7}$$

and we define its purification $|\varphi_{\mathbf{i}_\mathrm{I}}\rangle$ (so that $\rho^{\mathbf{i}_\mathrm{I}}$ is a partial trace of $|\varphi_{\mathbf{i}_\mathrm{I}}\rangle$) as

$$|\varphi_{\mathbf{i}_\mathrm{I}}\rangle \triangleq \sum_{\mathbf{j}_\mathrm{I}\in\mathbf{F}_2^n} |E_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}\rangle_{\mathbf{b},\mathbf{s}}|\mathbf{i}_\mathrm{I}\oplus\mathbf{j}_\mathrm{I}\rangle. \tag{7.8}$$

7.2.3 The Symmetrized Attack of Eve

In [BBBMR06], the most general joint attack is not directly analyzed: for simplicity, it is assumed that Eve applies a process called *symmetrization*, resulting in a *symmetrized* attack. The process of symmetrization is always beneficial for Eve (it does not change the error rate, and we prove in Proposition 7.4 that it does not decrease Eve's information), so a security proof against all symmetrized attacks implies a security proof against all the possible joint attacks.

In Eve's original attack, she has her own probe subsystem E. In the symmetrization process, Eve adds another probe subsystem M, in the initial state of $|0_x\rangle_\mathrm{M} \triangleq \frac{1}{\sqrt{2^N}}\sum_{\mathbf{m}\in\mathbf{F}_2^N}|\mathbf{m}\rangle_\mathrm{M}$. Given the original attack U (applied to Alice's qubits and to the probe E), the symmetrized attack U^sym (applied to Alice's qubits and to both probes E and M) is defined by

$$U^{\text{sym}} \triangleq (\mathbf{I}_{\text{E}} \otimes S^\dagger)(U \otimes \mathbf{I}_{\text{M}})(\mathbf{I}_{\text{E}} \otimes S), \tag{7.9}$$

where S is a unitary operation applied to Alice's qubits and to the probe M, and it operates as follows:

$$S|\mathbf{i}\rangle_{\mathbf{b}}|\mathbf{m}\rangle_{\text{M}} = (-1)^{(\mathbf{i}\oplus\mathbf{b})\cdot\mathbf{m}}|\mathbf{i} \oplus \mathbf{m}\rangle_{\mathbf{b}}|\mathbf{m}\rangle_{\text{M}}. \tag{7.10}$$

Intuitively, Eve first XORs Alice's bit values with a random string $\mathbf{m}$ (kept by her); then she applies her original attack; and then she reverses the XOR with $\mathbf{m}$. Full definition and explanations are available in Subsection 3.1 of [BBBMR06].

In this chapter, we use several properties of the symmetrized attack. First of all, the "Basic Lemma of Symmetrization" (Lemma 3.1 of [BBBMR06]) gives the expression for $|E^{\text{sym}\prime}_{\mathbf{i},\mathbf{j}}\rangle_{\mathbf{b}}$ (of the symmetrized attack) as a function of $|E'_{\mathbf{i},\mathbf{j}}\rangle_{\mathbf{b}}$ (of the original attack):

$$|E^{\text{sym}\prime}_{\mathbf{i},\mathbf{j}}\rangle_{\mathbf{b}} = \frac{1}{\sqrt{2^N}} \sum_{\mathbf{m}\in\mathbf{F}_2^N} (-1)^{(\mathbf{i}\oplus\mathbf{j})\cdot\mathbf{m}}|E'_{\mathbf{i}\oplus\mathbf{m},\mathbf{j}\oplus\mathbf{m}}\rangle_{\mathbf{b}}|\mathbf{m}\rangle_{\text{M}}. \tag{7.11}$$

The second property we use, proved in Corollary 3.3 of [BBBMR06], is the fact that the probabilities of the error strings $\mathbf{c}_{\text{I}}$ and $\mathbf{c}_{\text{T}}$ (if *not conditioning* on $\mathbf{i}$) are not affected by the symmetrization. Namely,

$$\Pr{}^{\text{sym}}(\mathbf{c}_{\text{I}}, \mathbf{c}_{\text{T}} \mid \mathbf{b}, \mathbf{s}) = \Pr(\mathbf{c}_{\text{I}}, \mathbf{c}_{\text{T}} \mid \mathbf{b}, \mathbf{s}). \tag{7.12}$$

This is true for all the basis strings $\mathbf{b}$; in particular, this is true for the basis string $\mathbf{b}^0 \triangleq \mathbf{b}\oplus\mathbf{s}$ used in the hypothetical "inverted-INFO-basis" protocol defined in Subsection 7.2.1, so

$$\Pr{}^{\text{sym}}_{\text{inverted-INFO-basis}}(\mathbf{c}_{\text{I}}, \mathbf{c}_{\text{T}} \mid \mathbf{b}, \mathbf{s}) = \Pr{}_{\text{inverted-INFO-basis}}(\mathbf{c}_{\text{I}}, \mathbf{c}_{\text{T}} \mid \mathbf{b}, \mathbf{s}). \tag{7.13}$$

The third property we use, proved in Lemma 3.8 of [BBBMR06], is the fact that the probabilities for errors in the TEST bits are not affected by the bases used for the INFO bits:

$$\Pr{}^{\text{sym}}(\mathbf{j}_{\text{T}} \mid \mathbf{i}_{\text{T}}, \mathbf{b}, \mathbf{s}) = \Pr{}^{\text{sym}}(\mathbf{j}_{\text{T}} \mid \mathbf{i}_{\text{T}}, \mathbf{b}_{\text{T}}, \mathbf{s}). \tag{7.14}$$

In particular, since the only difference between the hypothetical "inverted-INFO-basis" protocol and the real protocol is the basis string used for the *INFO bits* ($\overline{\mathbf{b}_{\text{I}}}$ and $\mathbf{b}_{\text{I}}$, respectively), this means that the probabilities of errors in the TEST bits are the same for both of these protocols:

$$\Pr{}^{\text{sym}}_{\text{inverted-INFO-basis}}(\mathbf{j}_{\text{T}} \mid \mathbf{i}_{\text{T}}, \mathbf{b}, \mathbf{s}) = \Pr{}^{\text{sym}}(\mathbf{j}_{\text{T}} \mid \mathbf{i}_{\text{T}}, \mathbf{b}, \mathbf{s}). \tag{7.15}$$

The fourth property we use, proved in Corollary 3.6 of [BBBMR06], is the fact that the probability of any string of INFO bits $\mathbf{i}_{\text{I}}$ is uniform (that is, $\frac{1}{2^n}$) even when conditioning on the four parameters $\mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}$, that are ultimately known to Eve (we note that $\mathbf{j}_\mathrm{T}$ is affected by Eve's attack). Namely,

$$\Pr^{\mathrm{sym}}(\mathbf{i}_\mathrm{I} \mid \mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}) = \frac{1}{2^n}. \tag{7.16}$$

7.2.4 Results from [BBBMR06]

The security proof of the generalized BB84 protocols is very similar to the security proof of BB84 itself, that was detailed in [BBBMR06]. Most parts of the proof are not affected at all by the changes made to BB84 to get the generalized BB84 protocols (changes detailed in Section 7.1 of this chapter), because these parts assume fixed strings $\mathbf{s}$ and $\mathbf{b}$.

Therefore, the reader is referred to the proof in Section 3 (except Subsection 3.3.2) and Subsections 4.1–4.4 of [BBBMR06], that applies to all the generalizations of BB84 without any changes (except changing the total number of bits, $2n$, to N, which does not affect the proof at all), and that will not be repeated here.

We denote the rows of the error-correction parity check matrix P_C as the vectors $v_1, \ldots, v_r$ in $\mathbf{F}_2^n$, and the rows of the privacy amplification matrix P_K as the vectors $v_{r+1}, \ldots, v_{r+m}$. We also denote, for any $1 \le r' \le r + m$,

$$V_{r'}^{\mathrm{exc}} \triangleq \mathrm{Span}\{v_1, \ldots, v_{r'-1}, v_{r'+1}, \ldots, v_{r+m}\}, \tag{7.17}$$

namely, $V_{r'}^{\mathrm{exc}}$ is the $(r + m - 1)$-dimensional vector space that spans all the error correction and privacy amplification vectors, except $v_{r'}$; and we also define

$$\hat{v} \triangleq \min_{r+1 \le r' \le r+m} d_\mathrm{H}(v_{r'}, V_{r'}^{\mathrm{exc}}). \tag{7.18}$$

For a 1-bit final key $k \in \{0, 1\}$ (that is, for $m = 1$), and given a symmetrized attack of Eve, we define $\widehat{\rho}_k^{\mathrm{sym}}$ to be the state of Eve corresponding to the final key k, given that she knows $\boldsymbol{\xi}$. Thus,

$$\widehat{\rho}_k^{\mathrm{sym}} = \frac{1}{2^{n-r-1}} \sum_{\mathbf{i}_\mathrm{I} \,\middle|\, \substack{\mathbf{i}_\mathrm{I} P_\mathrm{C}^\mathrm{T} = \boldsymbol{\xi} \\ \mathbf{i}_\mathrm{I} \cdot v_{r+1} = k}} (\rho^{\mathbf{i}_\mathrm{I}})^{\mathrm{sym}}, \tag{7.19}$$

where $(\rho^{\mathbf{i}_\mathrm{I}})^{\mathrm{sym}}$, as defined in Equation (7.7), is Eve's state after the (symmetrized) attack, given that Alice sent the INFO bit string $\mathbf{i}_\mathrm{I}$ (and given the bit strings $\mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}$, that are ultimately known to Eve).

In addition, we define the state $\widetilde{\widehat{\rho}}_k^{\mathrm{sym}}$, that is a lift-up of $\widehat{\rho}_k^{\mathrm{sym}}$ (which means that $\widehat{\rho}_k^{\mathrm{sym}}$ is a partial trace of $\widetilde{\widehat{\rho}}_k^{\mathrm{sym}}$), by assuming that Eve knows the purification $|\varphi_{\mathbf{i}_\mathrm{I}}^{\mathrm{sym}}\rangle$

defined in Equation (7.8):

$$\widetilde{\rho}_k^{\mathrm{sym}} = \frac{1}{2^{n-r-1}} \sum_{\mathbf{i}_{\mathrm{I}} \,\middle|\, \substack{\mathbf{i}_{\mathrm{I}} P_{\mathrm{C}}^{\mathrm{T}} = \boldsymbol{\xi} \\ \mathbf{i}_{\mathrm{I}} \cdot v_{r+1} = k}} |\varphi_{\mathbf{i}_{\mathrm{I}}}^{\mathrm{sym}}\rangle\langle\varphi_{\mathbf{i}_{\mathrm{I}}}^{\mathrm{sym}}|. \tag{7.20}$$

(This state was defined in Equation (4.10) of [BBBMR06], but was denoted there as $\rho_k(v_{r+1}, \boldsymbol{\xi})$.)

In the end of Subsection 4.4 of [BBBMR06] (in its Proposition 4.6, and according to the proof of Lemma 4.5, which appears in Appendix D.2 of [BBBMR06]), it was found that (in the case of a 1-bit final key, i.e., $m = 1$), for any *symmetrized* attack,

$$\frac{1}{2} \operatorname{tr} |\widetilde{\rho}_0^{\mathrm{sym}} - \widetilde{\rho}_1^{\mathrm{sym}}| \le 2\sqrt{\Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}\left[|\mathbf{C}_{\mathrm{I}}| \ge \frac{\hat{v}}{2} \;\mid\; \mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}\right]}, \tag{7.21}$$

where $\mathbf{C}_{\mathrm{I}}$ is the random variable whose value equals to $\mathbf{c}_{\mathrm{I}} \triangleq \mathbf{i}_{\mathrm{I}} \oplus \mathbf{j}_{\mathrm{I}}$, and $\Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}$ means that the probability is taken over the hypothetical "inverted-INFO-basis" protocol defined in Subsection 7.2.1 (to which Eve applies the same symmetrized attack that she applies to the real protocol). We also note that $\hat{v}$ was defined above, and that in the current case ($m = 1$), its definition is simplified to $\hat{v} = d_{\mathrm{H}}(v_{r+1}, V_{r+1}^{\mathrm{exc}})$ (and V_{r+1}^{exc} is simply $\mathrm{Span}\{v_1, \ldots, v_r\}$).

Now, according to [NC00, Theorem 9.2 and page 407], and using the fact that $\widehat{\rho}_k^{\mathrm{sym}}$ is a partial trace of $\widetilde{\rho}_k^{\mathrm{sym}}$, we find out that

$$\frac{1}{2} \operatorname{tr} |\widehat{\rho}_0^{\mathrm{sym}} - \widehat{\rho}_1^{\mathrm{sym}}| \le \frac{1}{2} \operatorname{tr} |\widetilde{\rho}_0^{\mathrm{sym}} - \widetilde{\rho}_1^{\mathrm{sym}}|. \tag{7.22}$$

From this result and from Equation (7.21) we deduce that

$$\frac{1}{2} \operatorname{tr} |\widehat{\rho}_0^{\mathrm{sym}} - \widehat{\rho}_1^{\mathrm{sym}}| \le 2\sqrt{\Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}\left[|\mathbf{C}_{\mathrm{I}}| \ge \frac{\hat{v}}{2} \;\mid\; \mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}\right]}. \tag{7.23}$$

7.2.5 Bounding the Differences Between Eve's States

So far, we have discussed a 1-bit key. We will now discuss a general m-bit key $\mathbf{k}$. We define $\widehat{\rho}_{\mathbf{k}}^{\mathrm{sym}}$ to be the state of Eve corresponding to the final key $\mathbf{k}$, given that she knows $\boldsymbol{\xi}$:

$$\widehat{\rho}_{\mathbf{k}}^{\mathrm{sym}} = \frac{1}{2^{n-r-m}} \sum_{\mathbf{i}_{\mathrm{I}} \,\middle|\, \substack{\mathbf{i}_{\mathrm{I}} P_{\mathrm{C}}^{\mathrm{T}} = \boldsymbol{\xi} \\ \mathbf{i}_{\mathrm{I}} P_{\mathrm{K}}^{\mathrm{T}} = \mathbf{k}}} (\rho^{\mathbf{i}_{\mathrm{I}}})^{\mathrm{sym}}. \tag{7.24}$$

We note (for use in Subsection 7.2.6) that if we substitute $(\rho^{\mathbf{i}_{\mathrm{I}}})^{\mathrm{sym}}$ from Equation (7.7), we get

$$\widehat{\rho}_{\mathbf{k}}^{\mathrm{sym}} = \frac{1}{2^{n-r-m}} \sum_{\mathbf{i}_{\mathrm{I}},\mathbf{j}_{\mathrm{I}} \,\middle|\, \substack{\mathbf{i}_{\mathrm{I}} P_{\mathrm{C}}^{\mathrm{T}} = \boldsymbol{\xi} \\ \mathbf{i}_{\mathrm{I}} P_{\mathrm{K}}^{\mathrm{T}} = \mathbf{k}}} \mathrm{Pr}^{\mathrm{sym}}(\mathbf{j}_{\mathrm{I}} \mid \mathbf{i}_{\mathrm{I}}, \mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}) \cdot \left(\rho_{\mathbf{i}_{\mathrm{I}},\mathbf{j}_{\mathrm{I}}}^{\mathbf{b},\mathbf{s}}\right)^{\mathrm{sym}}. \tag{7.25}$$

Proposition 7.1. *For any two keys $\mathbf{k}, \mathbf{k}'$ of m bits, and for any symmetrized attack,*

$$\frac{1}{2}\,\mathrm{tr}\,|\widehat{\rho}_{\mathbf{k}}^{\mathrm{sym}} - \widehat{\rho}_{\mathbf{k}'}^{\mathrm{sym}}| \le 2m\sqrt{\mathrm{Pr}_{\text{inverted-INFO-basis}}^{\mathrm{sym}}\left[|\mathbf{C}_{\mathrm{I}}| \ge \frac{\hat{v}}{2} \mid \mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}\right]}, \tag{7.26}$$

where $\mathbf{C}_{\mathrm{I}}$ is the random variable whose value equals to $\mathbf{c}_{\mathrm{I}} \triangleq \mathbf{i}_{\mathrm{I}} \oplus \mathbf{j}_{\mathrm{I}}$, and, in addition, $\hat{v} \triangleq \min_{r+1 \le r' \le r+m} d_{\mathrm{H}}(v_{r'}, V_{r'}^{\mathrm{exc}})$.

Proof. We define the key $\mathbf{k}_j$, for $0 \le j \le m$, to consist of the first j bits of $\mathbf{k}'$ and the last $m - j$ bits of $\mathbf{k}$. This means that $\mathbf{k}_0 = \mathbf{k}$, $\mathbf{k}_m = \mathbf{k}'$, and $\mathbf{k}_{j-1}$ differs from $\mathbf{k}_j$ at most on a single bit (the j-th bit).

First, we find a bound on $\frac{1}{2}\,\mathrm{tr}\,|\widehat{\rho}_{\mathbf{k}_{j-1}}^{\mathrm{sym}} - \widehat{\rho}_{\mathbf{k}_j}^{\mathrm{sym}}|$: since $\mathbf{k}_{j-1}$ differs from $\mathbf{k}_j$ at most on a single bit (the j-th bit, given by the formula $\mathbf{i}_{\mathrm{I}} \cdot v_{r+j}$), we can use the same proof that gave us Equation (7.23), attaching the other (identical) key bits to $\boldsymbol{\xi}$ of the original proof; and we find out that

$$\frac{1}{2}\,\mathrm{tr}\,|\widehat{\rho}_{\mathbf{k}_{j-1}}^{\mathrm{sym}} - \widehat{\rho}_{\mathbf{k}_j}^{\mathrm{sym}}| \le 2\sqrt{\mathrm{Pr}_{\text{inverted-INFO-basis}}^{\mathrm{sym}}\left[|\mathbf{C}_{\mathrm{I}}| \ge \frac{\widehat{v_j}}{2} \mid \mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}\right]}, \tag{7.27}$$

where we define $\widehat{v_j}$ to be $d_{\mathrm{H}}(v_{r+j}, V_{r+j}^{\mathrm{exc}})$, and, therefore, $\hat{v} = \min_{1 \le j' \le m} \widehat{v_{j'}}$.

In particular, $\hat{v} \le \widehat{v_j}$. Therefore, if $|\mathbf{C}_{\mathrm{I}}| \ge \frac{\widehat{v_j}}{2}$, then $|\mathbf{C}_{\mathrm{I}}| \ge \frac{\hat{v}}{2}$. Therefore, Equation (7.27) implies

$$\frac{1}{2}\,\mathrm{tr}\,|\widehat{\rho}_{\mathbf{k}_{j-1}}^{\mathrm{sym}} - \widehat{\rho}_{\mathbf{k}_j}^{\mathrm{sym}}| \le 2\sqrt{\mathrm{Pr}_{\text{inverted-INFO-basis}}^{\mathrm{sym}}\left[|\mathbf{C}_{\mathrm{I}}| \ge \frac{\hat{v}}{2} \mid \mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}\right]}. \tag{7.28}$$

Now we use the triangle inequality for norms to find

$$\begin{aligned} \frac{1}{2}\,\mathrm{tr}\,|\widehat{\rho}_{\mathbf{k}}^{\mathrm{sym}} - \widehat{\rho}_{\mathbf{k}'}^{\mathrm{sym}}| = \frac{1}{2}\,\mathrm{tr}\,|\widehat{\rho}_{\mathbf{k}_0}^{\mathrm{sym}} - \widehat{\rho}_{\mathbf{k}_m}^{\mathrm{sym}}| &\le \sum_{j=1}^{m} \frac{1}{2}\,\mathrm{tr}\,|\widehat{\rho}_{\mathbf{k}_{j-1}}^{\mathrm{sym}} - \widehat{\rho}_{\mathbf{k}_j}^{\mathrm{sym}}| \\ &\le 2m\sqrt{\mathrm{Pr}_{\text{inverted-INFO-basis}}^{\mathrm{sym}}\left[|\mathbf{C}_{\mathrm{I}}| \ge \frac{\hat{v}}{2} \mid \mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}\right]}. \end{aligned} \tag{7.29}$$

We would now like to bound the expected value (namely, the average value) of the trace distance between two states of Eve corresponding to two final keys. However, we should take into account that if the test fails, no final key is generated, in which case we define the distance to be 0. We thus define the random variable $\Delta_{\mathrm{Eve}}^{\mathrm{sym}}(\mathbf{k}, \mathbf{k}')$ for any two final keys $\mathbf{k}, \mathbf{k}'$:

$$\Delta_{\mathrm{Eve}}^{\mathrm{sym}}(\mathbf{k}, \mathbf{k}' \mid \mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}) \triangleq \begin{cases} \frac{1}{2} \operatorname{tr} |\widehat{\rho}_{\mathbf{k}}^{\mathrm{sym}} - \widehat{\rho}_{\mathbf{k}'}^{\mathrm{sym}}| & \text{if } T(\mathbf{i}_{\mathrm{T}} \oplus \mathbf{j}_{\mathrm{T}}, \mathbf{b}_{\mathrm{T}}, \mathbf{s}) = 1 \\ 0 & \text{otherwise} \end{cases}. \tag{7.30}$$

We need to bound the expected value $\langle \Delta_{\mathrm{Eve}}^{\mathrm{sym}}(\mathbf{k}, \mathbf{k}') \rangle$, that is given by:

$$\langle \Delta_{\mathrm{Eve}}^{\mathrm{sym}}(\mathbf{k}, \mathbf{k}') \rangle = \sum_{\substack{\mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}} \in \mathbf{F}_2^{N-n}, \\ \mathbf{b} \in B, \mathbf{s} \in S_{\mathbf{b}}, \boldsymbol{\xi} \in \mathbf{F}_2^n}} \Delta_{\mathrm{Eve}}^{\mathrm{sym}}(\mathbf{k}, \mathbf{k}' \mid \mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}) \cdot \Pr^{\mathrm{sym}}(\mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}). \tag{7.31}$$

(In Subsection 7.2.6 we prove that this expected value is indeed the quantity we need to bound for proving fully composable security, defined in Subsection 2.3.1.)

Theorem 7.2. *For any two final keys $\mathbf{k}, \mathbf{k}'$,*

$$\langle \Delta_{\mathrm{Eve}}^{\mathrm{sym}}(\mathbf{k}, \mathbf{k}') \rangle \leq 2m \sqrt{\Pr_{\text{inverted-INFO-basis}} \left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n} \right) \wedge (\mathbf{T} = 1) \right]}, \tag{7.32}$$

where $\frac{|\mathbf{C}_{\mathrm{I}}|}{n}$ is a random variable whose value is the error rate on the INFO bits, and $\mathbf{T}$ is a random variable whose value is 1 if the test passes and 0 otherwise. We note that the protocol considered for the probability in the right-hand-side is the hypothetical "inverted-INFO-basis" protocol defined in Subsection 7.2.1, in which Alice and Bob use the basis string $\mathbf{b}^0 \triangleq \mathbf{b} \oplus \mathbf{s}$ instead of $\mathbf{b}$. We note that the probability in the right-hand-side is the probability for the original (non-symmetrized) attack.

Proof. We use the convexity of x^2, namely, the fact that for all $\{p_i\}_i$ satisfying $p_i \geq 0$ and $\sum_i p_i = 1$, it holds that $(\sum_i p_i x_i)^2 \leq \sum_i p_i x_i^2$. We also use the fact that

$$\begin{aligned} & \Pr_{\text{inverted-INFO-basis}}^{\mathrm{sym}}(\mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}) \\ = \; & \Pr_{\text{inverted-INFO-basis}}^{\mathrm{sym}}(\mathbf{i}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}) \cdot \Pr_{\text{inverted-INFO-basis}}^{\mathrm{sym}}(\mathbf{j}_{\mathrm{T}} \mid \mathbf{i}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}) \\ = \; & \Pr^{\mathrm{sym}}(\mathbf{i}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}) \cdot \Pr^{\mathrm{sym}}(\mathbf{j}_{\mathrm{T}} \mid \mathbf{i}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}) \\ = \; & \Pr^{\mathrm{sym}}(\mathbf{i}_{\mathrm{T}}, \mathbf{j}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}), \end{aligned} \tag{7.33}$$

which is correct because $\mathbf{i}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}$ are all *chosen* in the same way both in the hypothetical "inverted-INFO-basis" protocol and in the real protocol (even though different basis strings are *used* in these protocols), and because according to the third property of the symmetrized attack (Equation (7.15)), $\Pr_{\text{inverted-INFO-basis}}^{\mathrm{sym}}(\mathbf{j}_{\mathrm{T}} \mid \mathbf{i}_{\mathrm{T}}, \mathbf{b}, \mathbf{s}) = \Pr^{\mathrm{sym}}(\mathbf{j}_{\mathrm{T}} \mid \mathbf{i}_{\mathrm{T}}, \mathbf{b}, \mathbf{s})$.

In addition, we use the result

$$\begin{aligned}&\Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}\left[\left(|\mathbf{C}_\mathrm{I}| \geq \frac{\hat{v}}{2}\right) \wedge (\mathbf{T}=1) \mid \mathbf{b},\mathbf{s}\right]\\ =\ &\Pr_{\text{inverted-INFO-basis}}\left[\left(|\mathbf{C}_\mathrm{I}| \geq \frac{\hat{v}}{2}\right) \wedge (\mathbf{T}=1) \mid \mathbf{b},\mathbf{s}\right], \qquad (7.34)\end{aligned}$$

□

which is correct because according to the second property of the symmetrized attack (Equation (7.13)), $\Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}(\mathbf{c}_\mathrm{I},\mathbf{c}_\mathrm{T} \mid \mathbf{b},\mathbf{s}) = \Pr_{\text{inverted-INFO-basis}}(\mathbf{c}_\mathrm{I},\mathbf{c}_\mathrm{T} \mid \mathbf{b},\mathbf{s})$, and because the random variable $\mathbf{T}$ depends only on the random variable $\mathbf{C}_\mathrm{T}$ and on the parameters $\mathbf{b}_\mathrm{T},\mathbf{s}$.

We also use the result

$$\Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}(\mathbf{b},\mathbf{s}) = \Pr_{\text{inverted-INFO-basis}}(\mathbf{b},\mathbf{s}), \qquad (7.35)$$

which is correct because Alice's random choice of $\mathbf{b},\mathbf{s}$ is independent of Eve's attack.

We find out that:

$$\begin{aligned}&\langle\Delta^{\mathrm{sym}}_{\mathrm{Eve}}(\mathbf{k},\mathbf{k}')\rangle^2\\
&=\left[\sum_{\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi}} \Delta^{\mathrm{sym}}_{\mathrm{Eve}}(\mathbf{k},\mathbf{k}' \mid \mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi}) \cdot \Pr^{\mathrm{sym}}(\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi})\right]^2 && \text{(by (7.31))}\\
&\leq \sum_{\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi}} \left(\Delta^{\mathrm{sym}}_{\mathrm{Eve}}(\mathbf{k},\mathbf{k}' \mid \mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi})\right)^2 \cdot \Pr^{\mathrm{sym}}(\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi}) && \text{(by convexity of } x^2)\\
&= \sum_{\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\mid\mathbf{T}=1} \left(\tfrac{1}{2}\operatorname{tr}|\widehat{\rho}^{\mathrm{sym}}_{\mathbf{k}} - \widehat{\rho}^{\mathrm{sym}}_{\mathbf{k}'}|\right)^2 \cdot \Pr^{\mathrm{sym}}(\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi}) && \text{(by (7.30))}\\
&\leq 4m^2 \cdot \sum_{\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\mid\mathbf{T}=1} \Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}\left[|\mathbf{C}_\mathrm{I}| \geq \tfrac{\hat{v}}{2} \mid \mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s}\right]\\
&\cdot\ \Pr^{\mathrm{sym}}(\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi}) && \text{(by (7.26))}\\
&= 4m^2 \cdot \sum_{\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s}\mid\mathbf{T}=1} \Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}\left[|\mathbf{C}_\mathrm{I}| \geq \tfrac{\hat{v}}{2} \mid \mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s}\right]\\
&\cdot\ \Pr^{\mathrm{sym}}(\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s})\\
&= 4m^2 \cdot \sum_{\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s}} \Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}\left[\left(|\mathbf{C}_\mathrm{I}| \geq \tfrac{\hat{v}}{2}\right) \wedge (\mathbf{T}=1) \mid \mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s}\right]\\
&\cdot\ \Pr^{\mathrm{sym}}(\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s})\\
&= 4m^2 \cdot \sum_{\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s}} \Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}\left[\left(|\mathbf{C}_\mathrm{I}| \geq \tfrac{\hat{v}}{2}\right) \wedge (\mathbf{T}=1) \mid \mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s}\right]\\
&\cdot\ \Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}(\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s}) && \text{(by (7.33))}\\
&= 4m^2 \cdot \sum_{\mathbf{b},\mathbf{s}} \Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}\left[\left(|\mathbf{C}_\mathrm{I}| \geq \tfrac{\hat{v}}{2}\right) \wedge (\mathbf{T}=1) \mid \mathbf{b},\mathbf{s}\right]\\
&\cdot\ \Pr^{\mathrm{sym}}_{\text{inverted-INFO-basis}}(\mathbf{b},\mathbf{s})\end{aligned}$$

$$= 4m^2 \cdot \sum_{\mathbf{b},\mathbf{s}} \Pr_{\text{inverted-INFO-basis}} \left[\left(|\mathbf{C}_\text{I}| \geq \tfrac{\hat{v}}{2} \right) \wedge (\mathbf{T} = 1) \mid \mathbf{b}, \mathbf{s} \right]$$

$$\cdot \quad \Pr_{\text{inverted-INFO-basis}}(\mathbf{b}, \mathbf{s}) \qquad \text{(by (7.34)–(7.35))}$$

$$= 4m^2 \cdot \Pr_{\text{inverted-INFO-basis}} \left[\left(|\mathbf{C}_\text{I}| \geq \tfrac{\hat{v}}{2} \right) \wedge (\mathbf{T} = 1) \right] \tag{7.36}$$

7.2.6 Bound for Fully Composable Security

We now prove a crucial part of the claim that generalized BB84 protocols satisfy the definition of composable security for a QKD protocol: namely, that they satisfy Equation (2.2) presented in Subsection 2.3.1. We derive an upper bound for the expression $\frac{1}{2} \operatorname{tr} |\rho_\text{ABE} - \rho_\text{U} \otimes \rho_\text{E}|$, where ρ_ABE is the actual joint state of Alice, Bob, and Eve at the end of the protocol; ρ_U is an ideal (random, secret, and shared) key distributed to Alice and Bob; and ρ_E is the partial trace of ρ_ABE over the system AB (see Subsection 1.3.2). In other words, we upper-bound the trace distance between the system after the real QKD protocol and the system after an ideal key distribution protocol (which first performs the real QKD protocol and then magically distributes to Alice and Bob a random, secret, and shared key).

The states ρ_ABE and ρ_U are

$$\begin{aligned} \rho_\text{ABE} &= \sum_{\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s} | \mathbf{T}=1} \Pr(\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s}) \cdot |\mathbf{k}\rangle_\text{A}\langle\mathbf{k}| \otimes |\mathbf{k}^\text{B}\rangle_\text{B}\langle\mathbf{k}^\text{B}| \\ &\otimes \left(\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_\text{I},\mathbf{j}_\text{I}} \right)_\text{E} \otimes |\mathbf{i}_\text{T}, \mathbf{j}_\text{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}\rangle_\text{C}\langle\mathbf{i}_\text{T}, \mathbf{j}_\text{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}|, \end{aligned} \tag{7.37}$$

$$\rho_\text{U} = \frac{1}{2^m} \sum_{\mathbf{k}} |\mathbf{k}\rangle_\text{A}\langle\mathbf{k}| \otimes |\mathbf{k}\rangle_\text{B}\langle\mathbf{k}|, \tag{7.38}$$

where $\left(\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_\text{I},\mathbf{j}_\text{I}} \right)_\text{E}$ is defined in Equation (7.6) to be Eve's quantum normalized state if Alice sends the bit strings $\mathbf{i}_\text{I}, \mathbf{i}_\text{T}, \mathbf{b}, \mathbf{s}$ and Bob measures the bit strings $\mathbf{j}_\text{I}, \mathbf{j}_\text{T}$. All the other states actually represent classical information: subsystems A and B represent the final keys held by Alice ($\mathbf{k} \triangleq \mathbf{i}_\text{I} P_\text{K}^\text{T}$) and Bob (his key $\mathbf{k}^\text{B}$ is obtained from $\mathbf{j}_\text{I}$, $\boldsymbol{\xi} \triangleq \mathbf{i}_\text{I} P_\text{C}^\text{T}$, and P_K), and subsystem C represents the information published in the unjammable classical channel during the protocol (this information is known to Alice, Bob, and Eve)—namely, $\mathbf{i}_\text{T}, \mathbf{j}_\text{T}$ (all the TEST bits), $\mathbf{b}$ (the basis string), $\mathbf{s}$ (the string representing the partition into INFO and TEST bits), and $\boldsymbol{\xi} \triangleq \mathbf{i}_\text{I} P_\text{C}^\text{T}$ (the syndrome).

We note that in the definition of ρ_ABE, we sum only over the events in which the test is *passed* (namely, in which the protocol is not aborted by Alice and Bob): in such cases, an m-bit key is generated. The cases in which the protocol aborts do not exist in the sum—namely, they are represented by the zero operator, as required by the definition of composable security (see [Ren08, Subsection 6.1.2]). Thus, ρ_ABE is a non-normalized state, and $\operatorname{tr}(\rho_\text{ABE})$ is the probability that the test is passed.

To help us bound the trace distance, we define the following intermediate state:

$$\begin{aligned}\sigma_{\mathrm{ABE}} &\triangleq \sum_{\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s}|\mathbf{T}=1} \Pr\left(\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s}\right)\cdot|\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}|\otimes|\mathbf{k}\rangle_{\mathrm{B}}\langle\mathbf{k}| \\ &\otimes \left(\rho_{\mathbf{i}_{\mathrm{I}},\mathbf{j}_{\mathrm{I}}}^{\mathbf{b},\mathbf{s}}\right)_{\mathrm{E}}\otimes|\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}|. \qquad (7.39)\end{aligned}$$

This state is identical to ρ_{ABE}, except that Bob holds Alice's final key ($\mathbf{k}$) instead of his own calculated final key ($\mathbf{k}^{\mathrm{B}}$). In particular, the similarity between ρ_{ABE} and σ_{ABE} means, by definition, that $\rho_{\mathrm{E}}\triangleq\mathrm{tr}_{\mathrm{AB}}\left(\rho_{\mathrm{ABE}}\right)$ and $\sigma_{\mathrm{E}}\triangleq\mathrm{tr}_{\mathrm{AB}}\left(\sigma_{\mathrm{ABE}}\right)$ are the same state: that is, $\rho_{\mathrm{E}}=\sigma_{\mathrm{E}}$.

Proposition 7.3. *For any symmetrized attack, it holds that*

$$\begin{aligned}&\frac{1}{2}\,\mathrm{tr}\left|\sigma_{\mathrm{ABE}}^{\mathrm{sym}}-\rho_{\mathrm{U}}\otimes\sigma_{\mathrm{E}}^{\mathrm{sym}}\right| \\ \leq\ & 2m\sqrt{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n}\geq\frac{\hat{v}}{2n}\right)\wedge\left(\mathbf{T}=1\right)\right]}, \qquad (7.40)\end{aligned}$$

for $\sigma_{\mathrm{ABE}}^{\mathrm{sym}}$ and ρ_{U} defined above (but for the symmetrized attack) and for the partial trace $\sigma_{\mathrm{E}}^{\mathrm{sym}}\triangleq\mathrm{tr}_{\mathrm{AB}}\left(\sigma_{\mathrm{ABE}}^{\mathrm{sym}}\right)$. We note that the probability in the right-hand-side is the probability for the original (non-symmetrized) attack.

Proof. We notice that in $\sigma_{\mathrm{ABE}}^{\mathrm{sym}}$, the only factors depending directly on $\mathbf{i}_{\mathrm{I}}$ and $\mathbf{j}_{\mathrm{I}}$ (and not only on $\mathbf{k}$ and $\boldsymbol{\xi}$) are the probability $\Pr^{\mathrm{sym}}\left(\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s}\right)$ and Eve's state $\left(\rho_{\mathbf{i}_{\mathrm{I}},\mathbf{j}_{\mathrm{I}}}^{\mathbf{b},\mathbf{s}}\right)_{\mathrm{E}}^{\mathrm{sym}}$. The probability can be reformulated as

$$\begin{aligned}\Pr^{\mathrm{sym}}\left(\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s}\right) &= \Pr^{\mathrm{sym}}\left(\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\right)\cdot\Pr^{\mathrm{sym}}\left(\mathbf{k}\mid\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\right) \\ &\cdot\ \Pr^{\mathrm{sym}}\left(\mathbf{i}_{\mathrm{I}}\mid\mathbf{k},\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\right)\cdot\Pr^{\mathrm{sym}}\left(\mathbf{j}_{\mathrm{I}}\mid\mathbf{i}_{\mathrm{I}},\mathbf{k},\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\right) \\ &= \Pr^{\mathrm{sym}}\left(\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\right)\cdot\frac{1}{2^m}\cdot\frac{1}{2^{n-r-m}} \\ &\cdot\ \Pr^{\mathrm{sym}}\left(\mathbf{j}_{\mathrm{I}}\mid\mathbf{i}_{\mathrm{I}},\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s}\right). \qquad (7.41)\end{aligned}$$

(This is correct because all the possible n-bit values of $\mathbf{i}_{\mathrm{I}}$ have the same probability, $\frac{1}{2^n}$, conditioned on $\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s}$, according to the fourth property of the symmetrized attack (Equation (7.16)); and because all the $r+m$ rows of the matrices P_{C} and P_{K} are linearly independent, so there are exactly 2^{n-r-m} values of $\mathbf{i}_{\mathrm{I}}$ corresponding to each specific pair $(\boldsymbol{\xi},\mathbf{k})$.)

Therefore, the state $\sigma_{\mathrm{ABE}}^{\mathrm{sym}}$ takes the following form:

$$\sigma_{\mathrm{ABE}}^{\mathrm{sym}} = \frac{1}{2^m}\sum_{\mathbf{k},\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}|\mathbf{T}=1}\Pr^{\mathrm{sym}}\left(\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\right)\cdot|\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}|\otimes|\mathbf{k}\rangle_{\mathrm{B}}\langle\mathbf{k}|$$

$$
\begin{aligned}
&\otimes \left[\frac{1}{2^{n-r-m}} \sum_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I} \big| \substack{\mathbf{i}_\mathrm{I} P_\mathrm{C}^\mathrm{T} = \boldsymbol{\xi} \\ \mathbf{i}_\mathrm{I} P_\mathrm{K}^\mathrm{T} = \mathbf{k}}} \Pr{}^{\mathrm{sym}}\left(\mathbf{j}_\mathrm{I} \mid \mathbf{i}_\mathrm{I}, \mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}\right) \cdot \left(\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}\right)^{\mathrm{sym}}_\mathrm{E}\right] \\
&\otimes |\mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}\rangle_\mathrm{C}\langle \mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}| \\
&= \frac{1}{2^m} \sum_{\mathbf{k},\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi} | \mathbf{T}=1} \Pr{}^{\mathrm{sym}}\left(\mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}\right) \cdot |\mathbf{k}\rangle_\mathrm{A}\langle\mathbf{k}| \otimes |\mathbf{k}\rangle_\mathrm{B}\langle\mathbf{k}| \\
&\otimes \left(\widehat{\rho}^{\,\mathrm{sym}}_{\mathbf{k}}\right)_\mathrm{E} \otimes |\mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}\rangle_\mathrm{C}\langle \mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}|. \qquad (7.42)
\end{aligned}
$$

(This expression for $\widehat{\rho}^{\,\mathrm{sym}}_{\mathbf{k}}$ was found in Equation (7.25).)

The partial trace $\sigma^{\mathrm{sym}}_\mathrm{E} \triangleq \mathrm{tr}_\mathrm{AB}\left(\sigma^{\mathrm{sym}}_\mathrm{ABE}\right)$ is

$$
\sigma^{\mathrm{sym}}_\mathrm{E} = \frac{1}{2^m} \sum_{\mathbf{k},\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi} | \mathbf{T}=1} \Pr{}^{\mathrm{sym}}\left(\mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}\right) \cdot \left(\widehat{\rho}^{\,\mathrm{sym}}_{\mathbf{k}}\right)_\mathrm{E} \otimes |\mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}\rangle_\mathrm{C}\langle \mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}|, \qquad (7.43)
$$

and the state $\rho_\mathrm{U} \otimes \sigma^{\mathrm{sym}}_\mathrm{E}$ is

$$
\begin{aligned}
\rho_\mathrm{U} \otimes \sigma^{\mathrm{sym}}_\mathrm{E} &= \frac{1}{2^{2m}} \sum_{\mathbf{k},\mathbf{k}',\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi} | \mathbf{T}=1} \Pr{}^{\mathrm{sym}}\left(\mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}\right) \cdot |\mathbf{k}\rangle_\mathrm{A}\langle\mathbf{k}| \otimes |\mathbf{k}\rangle_\mathrm{B}\langle\mathbf{k}| \\
&\otimes \left(\widehat{\rho}^{\,\mathrm{sym}}_{\mathbf{k}'}\right)_\mathrm{E} \otimes |\mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}\rangle_\mathrm{C}\langle \mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}|. \qquad (7.44)
\end{aligned}
$$

Since $\sigma^{\mathrm{sym}}_\mathrm{ABE}$ and $\rho_\mathrm{U} \otimes \sigma^{\mathrm{sym}}_\mathrm{E}$ are the same (except the difference between Eve's states, $\left(\widehat{\rho}^{\,\mathrm{sym}}_{\mathbf{k}}\right)_\mathrm{E}$ and $\left(\widehat{\rho}^{\,\mathrm{sym}}_{\mathbf{k}'}\right)_\mathrm{E}$), we get, by using the triangle inequality for norms, the definition of $\langle \Delta^{\mathrm{sym}}_{\mathrm{Eve}}(\mathbf{k},\mathbf{k}')\rangle$ (Equation (7.31)), and Theorem 7.2:

$$
\begin{aligned}
&\frac{1}{2}\,\mathrm{tr}\left|\sigma^{\mathrm{sym}}_\mathrm{ABE} - \rho_\mathrm{U} \otimes \sigma^{\mathrm{sym}}_\mathrm{E}\right| \\
\leq\; & \frac{1}{2^{2m}} \sum_{\mathbf{k},\mathbf{k}',\mathbf{i}_\mathrm{T},\mathbf{j}_\mathrm{T},\mathbf{b},\mathbf{s},\boldsymbol{\xi} | \mathbf{T}=1} \Pr{}^{\mathrm{sym}}\left(\mathbf{i}_\mathrm{T}, \mathbf{j}_\mathrm{T}, \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}\right) \cdot \frac{1}{2}\,\mathrm{tr}\left|\left(\widehat{\rho}^{\,\mathrm{sym}}_{\mathbf{k}}\right)_\mathrm{E} - \left(\widehat{\rho}^{\,\mathrm{sym}}_{\mathbf{k}'}\right)_\mathrm{E}\right| \\
=\; & \frac{1}{2^{2m}} \sum_{\mathbf{k},\mathbf{k}'} \langle \Delta^{\mathrm{sym}}_{\mathrm{Eve}}(\mathbf{k},\mathbf{k}')\rangle \\
\leq\; & 2m \sqrt{\Pr{}_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_\mathrm{I}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T}=1)\right]}. \qquad (7.45)
\end{aligned}
$$

$\square$

Proposition 7.4. *For any attack, it holds that*

$$
\frac{1}{2}\,\mathrm{tr}\left|\sigma_\mathrm{ABE} - \rho_\mathrm{U} \otimes \sigma_\mathrm{E}\right| \leq \frac{1}{2}\,\mathrm{tr}\left|\sigma^{\mathrm{sym}}_\mathrm{ABE} - \rho_\mathrm{U} \otimes \sigma^{\mathrm{sym}}_\mathrm{E}\right|, \qquad (7.46)
$$

for σ_ABE, $\sigma^{\mathrm{sym}}_\mathrm{ABE}$, and ρ_U defined above and for the partial traces $\sigma_\mathrm{E} \triangleq \mathrm{tr}_\mathrm{AB}\left(\sigma_\mathrm{ABE}\right)$ and $\sigma^{\mathrm{sym}}_\mathrm{E} \triangleq \mathrm{tr}_\mathrm{AB}\left(\sigma^{\mathrm{sym}}_\mathrm{ABE}\right)$.

Proof. First, we have to find an expression for $\left(\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_\mathrm{I},\mathbf{j}_\mathrm{I}}\right)^{\mathrm{sym}}_\mathrm{E}$. According to Equation (7.6),

$$\left(\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_{\mathrm{I}},\mathbf{j}_{\mathrm{I}}}\right)^{\mathrm{sym}}_{\mathrm{E}} = \frac{\left[|E^{\mathrm{sym}\prime}_{\mathbf{i},\mathbf{j}}\rangle_{\mathbf{b}}\langle E^{\mathrm{sym}\prime}_{\mathbf{i},\mathbf{j}}|\right]_{\mathrm{E}}}{\Pr^{\mathrm{sym}}(\mathbf{j} \mid \mathbf{i},\mathbf{b},\mathbf{s})}, \tag{7.47}$$

and according to the "Basic Lemma of Symmetrization" (see Equation (7.11)),

$$|E^{\mathrm{sym}\prime}_{\mathbf{i},\mathbf{j}}\rangle_{\mathbf{b}} = \frac{1}{\sqrt{2^N}} \sum_{\mathbf{m}\in\mathbf{F}_2^N} (-1)^{(\mathbf{i}\oplus\mathbf{j})\cdot\mathbf{m}} |E'_{\mathbf{i}\oplus\mathbf{m},\mathbf{j}\oplus\mathbf{m}}\rangle_{\mathbf{b}}|\mathbf{m}\rangle_{\mathrm{M}}. \tag{7.48}$$

Therefore,

$$\left(\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_{\mathrm{I}},\mathbf{j}_{\mathrm{I}}}\right)^{\mathrm{sym}}_{\mathrm{E}} = \frac{1}{2^N} \sum_{\mathbf{m},\mathbf{m}'\in\mathbf{F}_2^N} \frac{(-1)^{(\mathbf{i}\oplus\mathbf{j})\cdot(\mathbf{m}\oplus\mathbf{m}')}\left[|E'_{\mathbf{i}\oplus\mathbf{m},\mathbf{j}\oplus\mathbf{m}}\rangle_{\mathbf{b}}\langle E'_{\mathbf{i}\oplus\mathbf{m}',\mathbf{j}\oplus\mathbf{m}'}| \otimes |\mathbf{m}\rangle_{\mathrm{M}}\langle\mathbf{m}'|\right]_{\mathrm{E}}}{\Pr^{\mathrm{sym}}(\mathbf{j} \mid \mathbf{i},\mathbf{b},\mathbf{s})}. \tag{7.49}$$

The state $\sigma^{\mathrm{sym}}_{\mathrm{ABE}}$ now takes the following form:

$$\begin{aligned}
\sigma^{\mathrm{sym}}_{\mathrm{ABE}} &= \frac{1}{2^N} \sum_{\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s},\mathbf{m},\mathbf{m}'|\mathbf{T}=1} \Pr^{\mathrm{sym}}(\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s}) \cdot |\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}| \otimes |\mathbf{k}\rangle_{\mathrm{B}}\langle\mathbf{k}| \\
&\otimes \frac{(-1)^{(\mathbf{i}\oplus\mathbf{j})\cdot(\mathbf{m}\oplus\mathbf{m}')}\left[|E'_{\mathbf{i}\oplus\mathbf{m},\mathbf{j}\oplus\mathbf{m}}\rangle_{\mathbf{b}}\langle E'_{\mathbf{i}\oplus\mathbf{m}',\mathbf{j}\oplus\mathbf{m}'}| \otimes |\mathbf{m}\rangle_{\mathrm{M}}\langle\mathbf{m}'|\right]_{\mathrm{E}}}{\Pr^{\mathrm{sym}}(\mathbf{j} \mid \mathbf{i},\mathbf{b},\mathbf{s})} \\
&\otimes |\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}| \\
&= \frac{1}{2^N} \sum_{\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s},\mathbf{m},\mathbf{m}'|\mathbf{T}=1} \Pr^{\mathrm{sym}}(\mathbf{i},\mathbf{b},\mathbf{s}) \cdot |\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}| \otimes |\mathbf{k}\rangle_{\mathrm{B}}\langle\mathbf{k}| \\
&\otimes (-1)^{(\mathbf{i}\oplus\mathbf{j})\cdot(\mathbf{m}\oplus\mathbf{m}')}\left[|E'_{\mathbf{i}\oplus\mathbf{m},\mathbf{j}\oplus\mathbf{m}}\rangle_{\mathbf{b}}\langle E'_{\mathbf{i}\oplus\mathbf{m}',\mathbf{j}\oplus\mathbf{m}'}| \otimes |\mathbf{m}\rangle_{\mathrm{M}}\langle\mathbf{m}'|\right]_{\mathrm{E}} \\
&\otimes |\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}|.
\end{aligned} \tag{7.50}$$

We define a unitary operator V: given the state $|\mathbf{m}\rangle_{\mathrm{M}}$ (held by Eve), the unitary operator V takes a XOR of all the states in the subsystems A, B, and C with the relevant parts of $\mathbf{m}$. Namely, if we define $\mathbf{m}_{\mathrm{I}}$ and $\mathbf{m}_{\mathrm{T}}$ as the INFO bits and the TEST bits (respectively) of $\mathbf{m}$ (of course, they depend on $\mathbf{s}$), and if we define $\mathbf{k}_{\mathbf{m}} \triangleq \mathbf{m}_{\mathrm{I}}P_{\mathrm{K}}^{\mathrm{T}}$ and $\boldsymbol{\xi}_{\mathbf{m}} \triangleq \mathbf{m}_{\mathrm{I}}P_{\mathrm{C}}^{\mathrm{T}}$, then

$$\begin{aligned}
V &\quad |\mathbf{k}\rangle_{\mathrm{A}}|\mathbf{k}\rangle_{\mathrm{B}}\left[|E\rangle|\mathbf{m}\rangle_{\mathrm{M}}\right]_{\mathrm{E}}|\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\rangle_{\mathrm{C}} \\
= &\quad |\mathbf{k}\oplus\mathbf{k}_{\mathbf{m}}\rangle_{\mathrm{A}}|\mathbf{k}\oplus\mathbf{k}_{\mathbf{m}}\rangle_{\mathrm{B}}\left[|E\rangle|\mathbf{m}\rangle_{\mathrm{M}}\right]_{\mathrm{E}} \\
&\quad |\mathbf{i}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\oplus\boldsymbol{\xi}_{\mathbf{m}}\rangle_{\mathrm{C}}.
\end{aligned} \tag{7.51}$$

Therefore (also using the fact that $\Pr^{\mathrm{sym}}(\mathbf{i},\mathbf{b},\mathbf{s}) = \Pr(\mathbf{i},\mathbf{b},\mathbf{s}) = \Pr(\mathbf{i}\oplus\mathbf{m},\mathbf{b},\mathbf{s})$),

$$\begin{aligned}
V\sigma^{\mathrm{sym}}_{\mathrm{ABE}}V^{\dagger} &= \frac{1}{2^N} \sum_{\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s},\mathbf{m},\mathbf{m}'|\mathbf{T}=1} \Pr(\mathbf{i}\oplus\mathbf{m},\mathbf{b},\mathbf{s}) \cdot |\mathbf{k}\oplus\mathbf{k}_{\mathbf{m}}\rangle_{\mathrm{A}}\langle\mathbf{k}\oplus\mathbf{k}_{\mathbf{m}'}| \otimes |\mathbf{k}\oplus\mathbf{k}_{\mathbf{m}}\rangle_{\mathrm{B}}\langle\mathbf{k}\oplus\mathbf{k}_{\mathbf{m}'}| \\
&\otimes (-1)^{(\mathbf{i}\oplus\mathbf{j})\cdot(\mathbf{m}\oplus\mathbf{m}')}\left[|E'_{\mathbf{i}\oplus\mathbf{m},\mathbf{j}\oplus\mathbf{m}}\rangle_{\mathbf{b}}\langle E'_{\mathbf{i}\oplus\mathbf{m}',\mathbf{j}\oplus\mathbf{m}'}| \otimes |\mathbf{m}\rangle_{\mathrm{M}}\langle\mathbf{m}'|\right]_{\mathrm{E}} \\
&\otimes |\mathbf{i}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\oplus\boldsymbol{\xi}_{\mathbf{m}}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathrm{T}}\oplus\mathbf{m}'_{\mathrm{T}},\mathbf{j}_{\mathrm{T}}\oplus\mathbf{m}'_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\oplus\boldsymbol{\xi}_{\mathbf{m}'}|.
\end{aligned} \tag{7.52}$$

Tracing out the subsystem M (which is a part of Eve's probe), we get

$$\begin{aligned}\mathrm{tr}_{\mathrm{M}}\left[V\sigma_{\mathrm{ABE}}^{\mathrm{sym}}V^{\dagger}\right] &= \frac{1}{2^N}\sum_{\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s},\mathbf{m}|\mathbf{T}=1}\Pr\left(\mathbf{i}\oplus\mathbf{m},\mathbf{b},\mathbf{s}\right)\cdot|\mathbf{k}\oplus\mathbf{k_m}\rangle_{\mathrm{A}}\langle\mathbf{k}\oplus\mathbf{k_m}|\otimes|\mathbf{k}\oplus\mathbf{k_m}\rangle_{\mathrm{B}}\langle\mathbf{k}\oplus\mathbf{k_m}|\\ &\otimes \left[|E'_{\mathbf{i}\oplus\mathbf{m},\mathbf{j}\oplus\mathbf{m}}\rangle_{\mathbf{b}}\langle E'_{\mathbf{i}\oplus\mathbf{m},\mathbf{j}\oplus\mathbf{m}}|\right]_{\mathrm{E}}\\ &\otimes |\mathbf{i}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\oplus\boldsymbol{\xi}_{\mathbf{m}}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\oplus\boldsymbol{\xi}_{\mathbf{m}}|.\end{aligned} \quad (7.53)$$

Now we can change the indexes of the sum, in the following way: we denote $\mathbf{i}'\triangleq\mathbf{i}\oplus\mathbf{m}$ and $\mathbf{j}'\triangleq\mathbf{j}\oplus\mathbf{m}$ (for a fixed $\mathbf{m}$), and we immediately get, according to the definitions, the results $\mathbf{i}'_{\mathrm{T}}=\mathbf{i}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}}$, $\mathbf{j}'_{\mathrm{T}}=\mathbf{j}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}}$, $\mathbf{k}'\triangleq\mathbf{i}'_{\mathrm{I}}P_{\mathrm{K}}^{\mathrm{T}}=(\mathbf{i}_{\mathrm{I}}\oplus\mathbf{m}_{\mathrm{I}})P_{\mathrm{K}}^{\mathrm{T}}=\mathbf{k}\oplus\mathbf{k_m}$, and similarly $\boldsymbol{\xi}'\triangleq\mathbf{i}'_{\mathrm{I}}P_{\mathrm{C}}^{\mathrm{T}}=\boldsymbol{\xi}\oplus\boldsymbol{\xi}_{\mathbf{m}}$. We also notice that $\mathbf{T}$ gets $(\mathbf{i}_{\mathrm{T}}\oplus\mathbf{j}_{\mathrm{T}},\mathbf{b}_{\mathrm{T}},\mathbf{s})$ as inputs, and that they all stay the same (because $\mathbf{i}'_{\mathrm{T}}\oplus\mathbf{j}'_{\mathrm{T}}=(\mathbf{i}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}})\oplus(\mathbf{j}_{\mathrm{T}}\oplus\mathbf{m}_{\mathrm{T}})=\mathbf{i}_{\mathrm{T}}\oplus\mathbf{j}_{\mathrm{T}}$), and therefore the change of indexes does not impact the condition $\mathbf{T}=1$. Therefore,

$$\begin{aligned}\mathrm{tr}_{\mathrm{M}}\left[V\sigma_{\mathrm{ABE}}^{\mathrm{sym}}V^{\dagger}\right] &= \frac{1}{2^N}\sum_{\mathbf{i}',\mathbf{j}',\mathbf{b},\mathbf{s},\mathbf{m}|\mathbf{T}=1}\Pr\left(\mathbf{i}',\mathbf{b},\mathbf{s}\right)\cdot|\mathbf{k}'\rangle_{\mathrm{A}}\langle\mathbf{k}'|\otimes|\mathbf{k}'\rangle_{\mathrm{B}}\langle\mathbf{k}'|\\ &\otimes \left[|E'_{\mathbf{i}',\mathbf{j}'}\rangle_{\mathbf{b}}\langle E'_{\mathbf{i}',\mathbf{j}'}|\right]_{\mathrm{E}}\\ &\otimes |\mathbf{i}'_{\mathrm{T}},\mathbf{j}'_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}'\rangle_{\mathrm{C}}\langle\mathbf{i}'_{\mathrm{T}},\mathbf{j}'_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}'|.\end{aligned} \quad (7.54)$$

Using the relation $\left(\rho_{\mathbf{i}_{\mathrm{I}},\mathbf{j}_{\mathrm{I}}}^{\mathbf{b},\mathbf{s}}\right)_{\mathrm{E}}=\frac{\left[|E'_{\mathbf{i},\mathbf{j}}\rangle_{\mathbf{b}}\langle E'_{\mathbf{i},\mathbf{j}}|\right]_{\mathrm{E}}}{\Pr(\mathbf{j}|\mathbf{i},\mathbf{b},\mathbf{s})}$ from Equation (7.6), we get

$$\begin{aligned}\mathrm{tr}_{\mathrm{M}}\left[V\sigma_{\mathrm{ABE}}^{\mathrm{sym}}V^{\dagger}\right] &= \sum_{\mathbf{i}',\mathbf{j}',\mathbf{b},\mathbf{s}|\mathbf{T}=1}\Pr\left(\mathbf{i}',\mathbf{b},\mathbf{s}\right)\cdot|\mathbf{k}'\rangle_{\mathrm{A}}\langle\mathbf{k}'|\otimes|\mathbf{k}'\rangle_{\mathrm{B}}\langle\mathbf{k}'|\\ &\otimes \Pr(\mathbf{j}'\mid\mathbf{i}',\mathbf{b},\mathbf{s})\left(\rho_{\mathbf{i}'_{\mathrm{I}},\mathbf{j}'_{\mathrm{I}}}^{\mathbf{b},\mathbf{s}}\right)_{\mathrm{E}}\\ &\otimes |\mathbf{i}'_{\mathrm{T}},\mathbf{j}'_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}'\rangle_{\mathrm{C}}\langle\mathbf{i}'_{\mathrm{T}},\mathbf{j}'_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}'|\\ &= \sum_{\mathbf{i}',\mathbf{j}',\mathbf{b},\mathbf{s}|\mathbf{T}=1}\Pr\left(\mathbf{i}',\mathbf{j}',\mathbf{b},\mathbf{s}\right)\cdot|\mathbf{k}'\rangle_{\mathrm{A}}\langle\mathbf{k}'|\otimes|\mathbf{k}'\rangle_{\mathrm{B}}\langle\mathbf{k}'|\\ &\otimes \left(\rho_{\mathbf{i}'_{\mathrm{I}},\mathbf{j}'_{\mathrm{I}}}^{\mathbf{b},\mathbf{s}}\right)_{\mathrm{E}}\otimes|\mathbf{i}'_{\mathrm{T}},\mathbf{j}'_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}'\rangle_{\mathrm{C}}\langle\mathbf{i}'_{\mathrm{T}},\mathbf{j}'_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}'|\\ &= \sigma_{\mathrm{ABE}}.\end{aligned} \quad (7.55)$$

To sum up, we get the result $\sigma_{\mathrm{ABE}}=\mathrm{tr}_{\mathrm{M}}\left[V\sigma_{\mathrm{ABE}}^{\mathrm{sym}}V^{\dagger}\right]$. A very similar proof gives us the result $\rho_{\mathrm{U}}\otimes\sigma_{\mathrm{E}}=\mathrm{tr}_{\mathrm{M}}\left[V\left(\rho_{\mathrm{U}}\otimes\sigma_{\mathrm{E}}^{\mathrm{sym}}\right)V^{\dagger}\right]$. Since the trace distance is preserved under unitary operators and does not increase under partial trace, we get

$$\begin{aligned}\frac{1}{2}\mathrm{tr}\left|\sigma_{\mathrm{ABE}}-\rho_{\mathrm{U}}\otimes\sigma_{\mathrm{E}}\right| &= \frac{1}{2}\mathrm{tr}\left|\mathrm{tr}_{\mathrm{M}}\left[V\left(\sigma_{\mathrm{ABE}}^{\mathrm{sym}}-\rho_{\mathrm{U}}\otimes\sigma_{\mathrm{E}}^{\mathrm{sym}}\right)V^{\dagger}\right]\right|\\ &\leq \frac{1}{2}\mathrm{tr}\left|V\left(\sigma_{\mathrm{ABE}}^{\mathrm{sym}}-\rho_{\mathrm{U}}\otimes\sigma_{\mathrm{E}}^{\mathrm{sym}}\right)V^{\dagger}\right|\\ &= \frac{1}{2}\mathrm{tr}\left|\sigma_{\mathrm{ABE}}^{\mathrm{sym}}-\rho_{\mathrm{U}}\otimes\sigma_{\mathrm{E}}^{\mathrm{sym}}\right|.\end{aligned} \quad (7.56)$$

□

Proposition 7.5. *For any attack,*

$$\frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \sigma_{\mathrm{ABE}}| \le \Pr\left[\left(\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}\right) \wedge (\mathbf{T} = 1)\right], \tag{7.57}$$

for ρ_{ABE} and σ_{ABE} defined above, and for $\mathbf{k}$ being the final key computed by Alice and $\mathbf{k}^{\mathrm{B}}$ being the final key computed by Bob.

Proof.

$$\begin{aligned}
\rho_{\mathrm{ABE}} - \sigma_{\mathrm{ABE}} &= \sum_{\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s}|\mathbf{T}=1} \Pr(\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s}) \\
&\cdot\ |\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}| \otimes \left[|\mathbf{k}^{\mathrm{B}}\rangle_{\mathrm{B}}\langle\mathbf{k}^{\mathrm{B}}| - |\mathbf{k}\rangle_{\mathrm{B}}\langle\mathbf{k}|\right] \\
&\otimes\ \left(\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_{\mathrm{I}},\mathbf{j}_{\mathrm{I}}}\right)_{\mathrm{E}} \otimes |\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}| \\
&= \Pr\left[\left(\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}\right) \wedge (\mathbf{T} = 1)\right] \\
&\cdot\ \sum_{\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s}} \Pr\left[\mathbf{i},\mathbf{j},\mathbf{b},\mathbf{s} \mid \left(\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}\right) \wedge (\mathbf{T} = 1)\right] \\
&\cdot\ |\mathbf{k}\rangle_{\mathrm{A}}\langle\mathbf{k}| \otimes \left[|\mathbf{k}^{\mathrm{B}}\rangle_{\mathrm{B}}\langle\mathbf{k}^{\mathrm{B}}| - |\mathbf{k}\rangle_{\mathrm{B}}\langle\mathbf{k}|\right] \\
&\otimes\ \left(\rho^{\mathbf{b},\mathbf{s}}_{\mathbf{i}_{\mathrm{I}},\mathbf{j}_{\mathrm{I}}}\right)_{\mathrm{E}} \otimes |\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}\rangle_{\mathrm{C}}\langle\mathbf{i}_{\mathrm{T}},\mathbf{j}_{\mathrm{T}},\mathbf{b},\mathbf{s},\boldsymbol{\xi}|.
\end{aligned} \tag{7.58}$$

□

The trace distance between any two normalized states is bounded by 1. Therefore,

$$\frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \sigma_{\mathrm{ABE}}| \le \Pr\left[\left(\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}\right) \wedge (\mathbf{T} = 1)\right]. \tag{7.59}$$

Corollary 7.6. *For any attack,*

$$\begin{aligned}
&\frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}| \\
\le\ &\Pr\left[\left(\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}\right) \wedge (\mathbf{T} = 1)\right] \\
+\ &2m\sqrt{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \ge \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T} = 1)\right]},
\end{aligned} \tag{7.60}$$

for ρ_{ABE} and ρ_{U} defined above and for the partial trace $\rho_{\mathrm{E}} \triangleq \operatorname{tr}_{\mathrm{AB}}(\rho_{\mathrm{ABE}})$.

Proof. Using Propositions 7.3, 7.4, and 7.5, and also the fact that $\rho_{\mathrm{E}} = \sigma_{\mathrm{E}}$, we get:

$$\begin{aligned}
&\frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}| \\
\le\ &\frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \sigma_{\mathrm{ABE}}| + \frac{1}{2}\operatorname{tr}|\sigma_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \sigma_{\mathrm{E}}| \\
\le\ &\frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \sigma_{\mathrm{ABE}}| + \frac{1}{2}\operatorname{tr}|\sigma^{\mathrm{sym}}_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \sigma^{\mathrm{sym}}_{\mathrm{E}}| \\
\le\ &\Pr\left[\left(\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}\right) \wedge (\mathbf{T} = 1)\right] \\
+\ &2m\sqrt{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \ge \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T} = 1)\right]}.
\end{aligned} \tag{7.61}$$

□

We have thus found an upper bound for the expression $\frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}|$. In Section 7.3 we prove this upper bound to be exponentially small in n for specific protocols.

7.3 Full Security Proofs for Specific Protocols

Below we prove full security for specific important examples of generalized BB84 protocols.

In this section we use Hoeffding's theorem, as described in Section 2.6; in particular, we use Corollary 2.2.

7.3.1 The BB84-INFO-z Protocol

In the BB84-INFO-z protocol, all the INFO bits are sent by Alice in the z basis, while the TEST bits are sent in both the z and the x bases. This means that $\mathbf{b}$ and $\mathbf{s}$ together define a random partition of the set of indexes $\{1, 2, \dots, N\}$ into three disjoint sets:

- I (INFO bits, where $s_j = 1$ and $b_j = 0$) of size n;
- $\mathrm{T_Z}$ (TEST-Z bits, where $s_j = 0$ and $b_j = 0$) of size n_z; and
- $\mathrm{T_X}$ (TEST-X bits, where $s_j = 0$ and $b_j = 1$) of size n_x.

Formally, Alice and Bob agree on parameters n, n_z, n_x (such that $N = n + n_z + n_x$), and we choose $B = \{\mathbf{b} \in \mathbf{F}_2^N \mid |\mathbf{b}| = n_x\}$ and $S_{\mathbf{b}} = \{\mathbf{s} \in \mathbf{F}_2^N \mid (|\mathbf{s}| = n) \wedge (|\mathbf{s} \oplus \mathbf{b}| = n + n_x)\}$ (namely, $\mathbf{s} \in S_{\mathbf{b}}$ if it consists of n 1-bits that do not overlap with the n_x 1-bits of $\mathbf{b}$) for all $\mathbf{b} \in B$. The probability distributions $\Pr(\mathbf{b})$ and $\Pr(\mathbf{s} \mid \mathbf{b})$ are all uniform—namely, $\Pr(\mathbf{b}, \mathbf{s})$ is identical for all $\mathbf{b} \in B$ and $\mathbf{s} \in S_{\mathbf{b}}$.

Alice and Bob also agree on error rate thresholds, $p_{a,z}$ and $p_{a,x}$ (for the TEST-Z and TEST-X bits, respectively). The testing function T is defined as follows:

$$T(\mathbf{i}_{\mathrm{T}} \oplus \mathbf{j}_{\mathrm{T}}, \mathbf{b}_{\mathrm{T}}, \mathbf{s}) = 1 \;\Leftrightarrow\; (|\mathbf{i}_{\mathrm{T_Z}} \oplus \mathbf{j}_{\mathrm{T_Z}}| \le n_z \cdot p_{a,z}) \wedge (|\mathbf{i}_{\mathrm{T_X}} \oplus \mathbf{j}_{\mathrm{T_X}}| \le n_x \cdot p_{a,x})\,. \quad (7.62)$$

Namely, the test passes if and only if the error rate on the TEST-Z bits is at most $p_{a,z}$ *and* the error rate on the TEST-X bits is at most $p_{a,x}$.

Following Corollary 7.6, we know the following bound:

$$\begin{aligned} & \frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}| \\ \le\ & \Pr\left[\left(\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}\right) \wedge (\mathbf{T} = 1)\right] \\ +\ & 2m\sqrt{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \ge \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T} = 1)\right]}. \end{aligned} \quad (7.63)$$

Below we prove the two probabilities in the right-hand-side to be exponentially small in n:

Theorem 7.7. *Let us be given $\delta_{\mathrm{sec}} > 0$, and, for infinitely many values of n, a family $\{v_1^n, \ldots, v_{r_n+m_n}^n\}$ of linearly independent vectors in $\mathbf{F}_2^n$ such that $\delta_{\mathrm{sec}} < \frac{\hat{v}}{n}$. Then for any $p_{a,z}, p_{a,x} > 0$ and $\epsilon_{\mathrm{sec}} > 0$ such that $p_{a,x} + \epsilon_{\mathrm{sec}} \leq \frac{\delta_{\mathrm{sec}}}{2}$, and for any $n_z, n_x > 0$, it holds for the BB84-INFO-z protocol that*

$$\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T} = 1)\right] \leq e^{-2\left(\frac{n_x}{n+n_x}\right)^2 n\epsilon_{\mathrm{sec}}^2}. \tag{7.64}$$

Proof. Because $\frac{\hat{v}}{2n} > \frac{\delta_{\mathrm{sec}}}{2} \geq p_{a,x} + \epsilon_{\mathrm{sec}}$, it holds that

$$\begin{aligned}
& \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T} = 1)\right] \\
= \; & \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z} \leq p_{a,z}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \leq p_{a,x}\right)\right] \\
\leq \; & \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} > p_{a,x} + \epsilon_{\mathrm{sec}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \leq p_{a,x}\right)\right].
\end{aligned} \tag{7.65}$$

□

In the hypothetical "inverted-INFO-basis" protocol, the INFO and TEST-X bits are sent and measured in the x basis, while the TEST-Z bits are sent and measured in the z basis. Therefore, the random and uniform sampling of the n INFO bits out of the $n + n_x$ bits sent in the x basis (assuming that the TEST-Z bits have already been chosen) does not affect the bases in the hypothetical protocol. This means that we can apply Corollary 2.2 to this sampling, and we get

$$\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} > p_{a,x} + \epsilon_{\mathrm{sec}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \leq p_{a,x}\right)\right] \leq e^{-2\left(\frac{n_x}{n+n_x}\right)^2 n\epsilon_{\mathrm{sec}}^2}. \tag{7.66}$$

Theorem 7.8. *Let us be given $\delta_{\mathrm{rel}} > 0$, and, for infinitely many values of n, a family $\{v_1^n, \ldots, v_{r_n+m_n}^n\}$ of linearly independent vectors in $\mathbf{F}_2^n$ such that the parity-check matrix P_{C}, whose rows are $\{v_1^n, \ldots, v_{r_n}^n\}$, defines an error-correcting code that can correct up to $n \cdot \delta_{\mathrm{rel}}$ errors on an n-bit string. Then for any $p_{a,z}, p_{a,x} > 0$ and $\epsilon_{\mathrm{rel}} > 0$ such that $p_{a,z} + \epsilon_{\mathrm{rel}} \leq \delta_{\mathrm{rel}}$, and for any $n_z, n_x > 0$, it holds for the BB84-INFO-z protocol that*

$$\Pr\left[\left(\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}\right) \wedge (\mathbf{T} = 1)\right] \leq e^{-2\left(\frac{n_z}{n+n_z}\right)^2 n\epsilon_{\mathrm{rel}}^2}. \tag{7.67}$$

Proof. If $\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}$, Alice and Bob have different final keys, and this means that the error correction stage did not succeed. The error-correcting code can correct up to $n \cdot \delta_{\mathrm{rel}}$ errors, and, therefore, it can correct up to $n \cdot (p_{a,z} + \epsilon_{\mathrm{rel}})$ errors (since $p_{a,z} + \epsilon_{\mathrm{rel}} \leq \delta_{\mathrm{rel}}$). Therefore, a failure of the error correction stage must mean that there are more than $n \cdot (p_{a,z} + \epsilon_{\mathrm{rel}})$ errors in the INFO bits: namely, if $\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}$, then necessarily $\frac{|\mathbf{C}_{\mathrm{I}}|}{n} > p_{a,z} + \epsilon_{\mathrm{rel}}$.

Therefore,

$$\begin{aligned}\Pr\left[\left(\mathbf{k}\neq\mathbf{k}^{\mathrm{B}}\right)\wedge(\mathbf{T}=1)\right] &= \Pr\left[\left(\mathbf{k}\neq\mathbf{k}^{\mathrm{B}}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z}\leq p_{a,z}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x}\leq p_{a,x}\right)\right]\\ &\leq \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n}>p_{a,z}+\epsilon_{\mathrm{rel}}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z}\leq p_{a,z}\right)\right]. \qquad (7.68)\end{aligned}$$

□

In the real protocol, the INFO and TEST-Z bits are sent and measured in the z basis, while the TEST-X bits are sent and measured in the x basis. Therefore, the random and uniform sampling of the n INFO bits out of the $n+n_z$ bits sent in the z basis (assuming that the TEST-X bits have already been chosen) does not affect the bases in the real protocol. This means that we can apply Corollary 2.2 to this sampling, and we get

$$\Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n}>p_{a,z}+\epsilon_{\mathrm{rel}}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z}\leq p_{a,z}\right)\right]\leq e^{-2\left(\frac{n_z}{n+n_z}\right)^2 n\epsilon_{\mathrm{rel}}^2}. \qquad (7.69)$$

If we combine the conditions and the results of Corollary 7.6, Theorem 7.7, and Theorem 7.8, we get the following result:

Corollary 7.9. *Let us be given $\delta_{\mathrm{sec}}>0$, $\delta_{\mathrm{rel}}>0$, and, for infinitely many values of n, a family $\{v_1^n,\ldots,v_{r_n+m_n}^n\}$ of linearly independent vectors in $\mathbf{F}_2^n$ such that $\delta_{\mathrm{sec}}<\frac{\hat{v}}{n}$* ***and*** *such that the parity-check matrix P_{C}, whose rows are $\{v_1^n,\ldots,v_{r_n}^n\}$, defines an error-correcting code that can correct up to $n\cdot\delta_{\mathrm{rel}}$ errors on an n-bit string. Then for any $p_{a,z},p_{a,x}>0$ and $\epsilon_{\mathrm{sec}},\epsilon_{\mathrm{rel}}>0$ such that $p_{a,x}+\epsilon_{\mathrm{sec}}\leq\frac{\delta_{\mathrm{sec}}}{2}$ and $p_{a,z}+\epsilon_{\mathrm{rel}}\leq\delta_{\mathrm{rel}}$, and for any $n_z,n_x>0$, it holds for the BB84-INFO-z protocol that*

$$\frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}}-\rho_{\mathrm{U}}\otimes\rho_{\mathrm{E}}|\leq e^{-2\left(\frac{n_z}{n+n_z}\right)^2 n\epsilon_{\mathrm{rel}}^2}+2m_n e^{-\left(\frac{n_x}{n+n_x}\right)^2 n\epsilon_{\mathrm{sec}}^2}. \qquad (7.70)$$

This bound is exponentially small in n.

All that is left to be explained is why the vectors required by Corollary 7.9 exist. We need a family of vectors $\{v_1^n,\ldots,v_{r_n+m_n}^n\}$ satisfying both the condition $\frac{\hat{v}}{2n}>\frac{\delta_{\mathrm{sec}}}{2}\geq p_{a,x}+\epsilon_{\mathrm{sec}}$ and the ability to correct up to $n(p_{a,z}+\epsilon_{\mathrm{rel}})$ errors. Such families were proven to exist in Appendix E of [BBBMR06], giving the following upper bound on the bit-rate:

$$R_{\mathrm{secret}}\triangleq\frac{m_n}{n}<1-H_2(2p_{a,x}+2\epsilon_{\mathrm{sec}})-H_2\left(p_{a,z}+\epsilon_{\mathrm{rel}}+\frac{1}{n}\right), \qquad (7.71)$$

where $H_2(x)\triangleq -x\log_2(x)-(1-x)\log_2(1-x)$.

Note that we use here the error thresholds $p_{a,x}$ for the condition on $\hat{v}$ and $p_{a,z}$ for error correction. This is possible, because in [BBBMR06] these conditions on the codes are discussed separately.

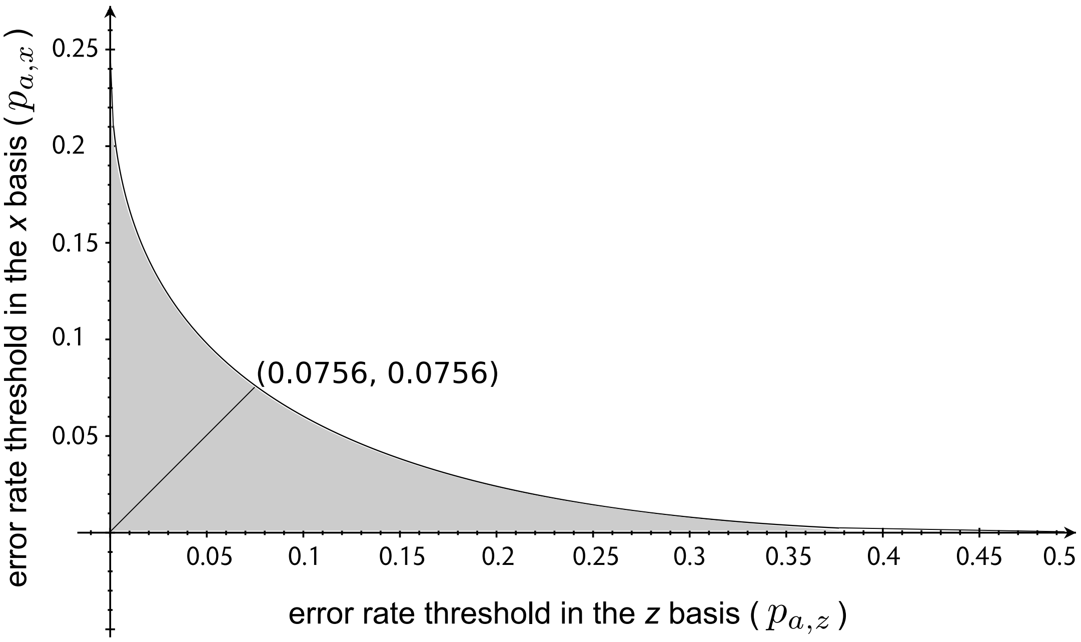

Figure 7.1: **The secure asymptotic error rates zone for BB84-INFO-z** (below the curve)

To get the asymptotic error rate thresholds, we require $R_{\text{secret}} > 0$, and we get the condition

$$H_2(2p_{a,x} + 2\epsilon_{\text{sec}}) + H_2\left(p_{a,z} + \epsilon_{\text{rel}} + \frac{1}{n}\right) < 1. \tag{7.72}$$

The secure asymptotic error rate thresholds zone is shown in Figure 7.1 (it is below the curve), assuming that $\frac{1}{n}$ is negligible. Note the trade-off between the error rate thresholds $p_{a,z}$ and $p_{a,x}$. Also note that in the case of $p_{a,z} = p_{a,x}$, we get the same threshold as in similar security proofs of BB84 [BBBMR06, BGM09], which is 7.56%.

7.3.2 The Standard BB84 Protocol

In the standard BB84 protocol, the strings $\mathbf{b}$ and $\mathbf{s}$ are chosen randomly (except that we demand $|\mathbf{s}| = n$) and independently, and $N = 2n$. In other words, there are n INFO bits and n TEST bits (chosen randomly), and for each one of them, the basis (z or x) is chosen randomly and independently.

Formally, in BB84, we choose $N = 2n$, $B = \mathbf{F}_2^N$, and $S_{\mathbf{b}} = \{\mathbf{s} \in \mathbf{F}_2^N \mid |\mathbf{s}| = n\}$ for all $\mathbf{b} \in B$. The probability distributions $\Pr(\mathbf{b})$ and $\Pr(\mathbf{s} \mid \mathbf{b}) = \Pr(\mathbf{s})$ are all uniform—namely, $\Pr(\mathbf{b}, \mathbf{s})$ is identical for all $\mathbf{b} \in B$ and $\mathbf{s} \in S_{\mathbf{b}}$.

Given the parameter p_a agreed by Alice and Bob, the testing function T is

$$T(\mathbf{i}_{\text{T}} \oplus \mathbf{j}_{\text{T}}, \mathbf{b}_{\text{T}}, \mathbf{s}) = 1 \ \Leftrightarrow\ |\mathbf{i}_{\text{T}} \oplus \mathbf{j}_{\text{T}}| \leq n \cdot p_a. \tag{7.73}$$

Namely, the test passes if and only if the error rate on the TEST bits is at most p_a.

Proposition 7.10. *In the standard BB84 protocol,*

$$\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T}=1)\right] = \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T}=1)\right]. \tag{7.74}$$

Proof.

$$\begin{aligned}
& \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T}=1)\right] \\
= & \sum_{\mathbf{b},\mathbf{s}} \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T}=1) \mid \mathbf{b},\mathbf{s}\right] \cdot \Pr(\mathbf{b},\mathbf{s}) \\
= & \sum_{\mathbf{b},\mathbf{s}} \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T}=1) \mid \mathbf{b}^0,\mathbf{s}\right] \cdot \Pr(\mathbf{b}^0,\mathbf{s}) \\
= & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T}=1)\right]
\end{aligned} \tag{7.75}$$

(where $\mathbf{b}^0 \triangleq \mathbf{b} \oplus \mathbf{s}$). □

The security of the standard BB84 protocol is now easily obtained:

Theorem 7.11. *Let us be given $\delta_{\mathrm{sec}} > 0$, $\delta_{\mathrm{rel}} > 0$, and, for infinitely many values of n, a family $\{v_1^n, \ldots, v_{r_n+m_n}^n\}$ of linearly independent vectors in $\mathbf{F}_2^n$ such that $\delta_{\mathrm{sec}} < \frac{\hat{v}}{n}$* ***and*** *such that the parity-check matrix P_{C}, whose rows are $\{v_1^n, \ldots, v_{r_n}^n\}$, defines an error-correcting code that can correct up to $n \cdot \delta_{\mathrm{rel}}$ errors on an n-bit string. Then for any $p_a > 0$ and $\epsilon_{\mathrm{sec}}, \epsilon_{\mathrm{rel}} > 0$ such that $p_a + \epsilon_{\mathrm{sec}} \leq \frac{\delta_{\mathrm{sec}}}{2}$ and $p_a + \epsilon_{\mathrm{rel}} \leq \delta_{\mathrm{rel}}$, it holds for the standard BB84 protocol that*

$$\frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}| \leq e^{-\frac{1}{2}n\epsilon_{\mathrm{rel}}^2} + 2m_n e^{-\frac{1}{4}n\epsilon_{\mathrm{sec}}^2}. \tag{7.76}$$

Proof. By using Corollary 7.6 and Proposition 7.10, we get the following bound for BB84:

$$\begin{aligned}
& \frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}| \\
\leq & \Pr\left[(\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}) \wedge (\mathbf{T}=1)\right] \\
+ & 2m_n\sqrt{\Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T}=1)\right]} \\
= & \Pr\left[(\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T}}|}{n} \leq p_a\right)\right] \\
+ & 2m_n\sqrt{\Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T}}|}{n} \leq p_a\right)\right]}.
\end{aligned} \tag{7.77}$$

Because $\frac{\hat{v}}{2n} > \frac{\delta_{\mathrm{sec}}}{2} \geq p_a + \epsilon_{\mathrm{sec}}$, and also because the event $\mathbf{k} \neq \mathbf{k}^{\mathrm{B}}$ implies that the error rate on the INFO bits is higher than $\delta_{\text{rel}} \ge p_a + \epsilon_{\text{rel}}$, we get:

$$\begin{aligned}
& \Pr\left[(\mathbf{k} \ne \mathbf{k}^{\text{B}}) \wedge \left(\frac{|\mathbf{C}_{\text{T}}|}{n} \le p_a\right)\right] \\
+ \quad & 2m_n\sqrt{\Pr\left[\left(\frac{|\mathbf{C}_{\text{I}}|}{n} \ge \frac{\hat{v}}{2n}\right) \wedge \left(\frac{|\mathbf{C}_{\text{T}}|}{n} \le p_a\right)\right]} \\
\le \quad & \Pr\left[\left(\frac{|\mathbf{C}_{\text{I}}|}{n} > p_a + \epsilon_{\text{rel}}\right) \wedge \left(\frac{|\mathbf{C}_{\text{T}}|}{n} \le p_a\right)\right] \\
+ \quad & 2m_n\sqrt{\Pr\left[\left(\frac{|\mathbf{C}_{\text{I}}|}{n} > p_a + \epsilon_{\text{sec}}\right) \wedge \left(\frac{|\mathbf{C}_{\text{T}}|}{n} \le p_a\right)\right]}. \qquad (7.78)
\end{aligned}$$

All the bits in the protocol are sent in random and independent bases. Therefore, the random and uniform sampling of the n INFO bits out of the $N = 2n$ bits does not affect the bases (in the real protocol). This means that we can apply Corollary 2.2 to this sampling, and we get

$$\Pr\left[\left(\frac{|\mathbf{C}_{\text{I}}|}{n} > p_a + \epsilon_{\text{rel}}\right) \wedge \left(\frac{|\mathbf{C}_{\text{T}}|}{n} \le p_a\right)\right] \le e^{-\frac{1}{2}n\epsilon_{\text{rel}}^2}, \qquad (7.79)$$

$$\Pr\left[\left(\frac{|\mathbf{C}_{\text{I}}|}{n} > p_a + \epsilon_{\text{sec}}\right) \wedge \left(\frac{|\mathbf{C}_{\text{T}}|}{n} \le p_a\right)\right] \le e^{-\frac{1}{2}n\epsilon_{\text{sec}}^2}. \qquad (7.80)$$

□

Combining Equations (7.77)–(7.80), we get

$$\frac{1}{2}\operatorname{tr}|\rho_{\text{ABE}} - \rho_{\text{U}} \otimes \rho_{\text{E}}| \le e^{-\frac{1}{2}n\epsilon_{\text{rel}}^2} + 2m_n e^{-\frac{1}{4}n\epsilon_{\text{sec}}^2}. \qquad (7.81)$$

In Appendix E of [BBBMR06] we get the following results on vector families satisfying the requirements of Theorem 7.11: the bit-rate satisfies

$$R_{\text{secret}} \triangleq \frac{m_n}{n} < 1 - H_2(2p_a + 2\epsilon_{\text{sec}}) - H_2\left(p_a + \epsilon_{\text{rel}} + \frac{1}{n}\right), \qquad (7.82)$$

and the condition on the asymptotic error rate threshold is

$$H_2(2p_a + 2\epsilon_{\text{sec}}) + H_2\left(p_a + \epsilon_{\text{rel}} + \frac{1}{n}\right) < 1. \qquad (7.83)$$

This gives an asymptotic error rate threshold of 7.56%.

7.3.3 The "Efficient BB84" Protocol

In the "efficient BB84" protocol (suggested by [LCA05]), the bit string $\mathbf{b}$ is chosen probabilistically, but *not uniformly*: each qubit is sent in the z basis with probability p (and in the x basis with probability $1 - p$), where $0 < p \le \frac{1}{2}$. Then, the bit string $\mathbf{s}$ is chosen such that there are n_z TEST-Z bits and n_x TEST-X bits. In other words, as in BB84-INFO-z, the strings $\mathbf{b}$ and $\mathbf{s}$ together define a random partition of the set of indexes $\{1, 2, \ldots, N\}$ into three disjoint sets:

- I (INFO bits, where $s_j = 1$) of size n. However, *unlike* BB84-INFO-z, this set consists of both z qubits and x qubits; therefore, it can be divided to two disjoint subsets:
  - $\mathrm{I_Z}$ (INFO-Z bits, where $s_j = 1$ and $b_j = 0$); and
  - $\mathrm{I_X}$ (INFO-X bits, where $s_j = 1$ and $b_j = 1$).
- $\mathrm{T_Z}$ (TEST-Z bits, where $s_j = 0$ and $b_j = 0$) of size n_z; and
- $\mathrm{T_X}$ (TEST-X bits, where $s_j = 0$ and $b_j = 1$) of size n_x.

Formally, in "efficient BB84", Alice and Bob agree on parameters n, n_z, n_x (such that $N = n + n_z + n_x$) and on a parameter $0 < p \leq \frac{1}{2}$, and we choose $B = \mathbf{F}_2^N$ and $S_{\mathbf{b}} = \{\mathbf{s} \in \mathbf{F}_2^N \mid (|\mathbf{s}| = n) \wedge (|\overline{\mathbf{s}} \wedge \mathbf{b}| = n_x)\}$ for all $\mathbf{b} \in B$ (namely, it is required that there are n INFO bits, n_z TEST-Z bits, and n_x TEST-X bits). This time, the probability distribution $\Pr(\mathbf{b})$ is *not* uniform: it holds that $\Pr(\mathbf{b}) = (1-p)^{|\mathbf{b}|} \cdot p^{N-|\mathbf{b}|}$, because the probability of each bit to be in the x basis is $1 - p$. On the other hand, the probability distribution $\Pr(\mathbf{s} \mid \mathbf{b})$ is uniform.

Remark. A subtle point is that for some values $\mathbf{b} \in \mathbf{F}_2^N$ (for example, for $\mathbf{b} = 00\ldots0$), the set $S_{\mathbf{b}}$ is empty: no $\mathbf{s}$ can be agreed by Alice and Bob for such values of $\mathbf{b}$. In that case, as assumed in [LCA05, Section 4.3], the protocol aborts, and other values of $\mathbf{b}$ and $\mathbf{s}$ are randomly chosen; this is equivalent to assuming Alice is not allowed to choose these values of $\mathbf{b}$. Therefore, to be more precise, we must re-define

$$B = \{\mathbf{b} \in \mathbf{F}_2^N \mid S_{\mathbf{b}} \neq \emptyset\} = \{\mathbf{b} \in \mathbf{F}_2^N \mid (|\mathbf{b}| \geq n_x) \wedge (|\overline{\mathbf{b}}| \geq n_z)\}, \tag{7.84}$$

and we must normalize the probabilities by defining $\Pr_0(\mathbf{b}) \triangleq (1-p)^{|\mathbf{b}|} \cdot p^{N-|\mathbf{b}|}$ (the original probability of each $\mathbf{b}$), $N_p \triangleq \sum_{\mathbf{b} \in B} \Pr_0(\mathbf{b})$ (the sum of all the original probabilities for all the allowed values of $\mathbf{b} \in B$), and then the real probability of each $\mathbf{b} \in B$ is

$$\Pr(\mathbf{b}) = \frac{\Pr_0(\mathbf{b})}{N_p} = \frac{(1-p)^{|\mathbf{b}|} \cdot p^{N-|\mathbf{b}|}}{N_p}. \tag{7.85}$$

This guarantees that the sum of probabilities of all the allowed values $\mathbf{b} \in B$ is 1.

Alice and Bob also agree on an error rate threshold, p_a (applied *both* to the TEST-Z bits and to the TEST-X bits). The testing function T is defined as follows:

$$T(\mathbf{i}_{\mathrm{T}} \oplus \mathbf{j}_{\mathrm{T}}, \mathbf{b}_{\mathrm{T}}, \mathbf{s}) = 1 \ \Leftrightarrow\ (|\mathbf{i}_{\mathrm{T_Z}} \oplus \mathbf{j}_{\mathrm{T_Z}}| \leq n_z \cdot p_a) \wedge (|\mathbf{i}_{\mathrm{T_X}} \oplus \mathbf{j}_{\mathrm{T_X}}| \leq n_x \cdot p_a). \tag{7.86}$$

Namely, the test passes if and only if the error rate on the TEST-Z bits is at most p_a *and* the error rate on the TEST-X bits is at most p_a.

In this security proof, instead of analyzing all the INFO bits together, we analyze the INFO-Z and the INFO-X bits separately. We define the following random variables:

- $\mathbf{C}_{\mathrm{I_Z}}$ and $\mathbf{C}_{\mathrm{I_X}}$ are the random variables corresponding to the error strings on the INFO-Z bits and on the INFO-X bits, respectively.
- $N_{\mathrm{I_Z}}$ and $N_{\mathrm{I_X}}$ are random variables equal to the numbers of INFO-Z and INFO-X bits, respectively. (We note that the parameters n, n_z, n_x are deterministically chosen by Alice and Bob, while $N_{\mathrm{I_Z}}$ and $N_{\mathrm{I_X}}$ are determined by the probabilistic choice of $\mathbf{b}$. We also note that, necessarily, $n = N_{\mathrm{I_Z}} + N_{\mathrm{I_X}}$.)

Proposition 7.12. *For any $\epsilon > 0$,*

$$\begin{aligned} & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right] \\ \le \; & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{N_{\mathrm{I_Z}}} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right] \\ + \; & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{N_{\mathrm{I_X}}} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right]. \end{aligned} \tag{7.87}$$

Equation (7.87) *similarly applies to the hypothetical "inverted-INFO-basis" protocol, too (namely, it applies even if* $\Pr$ *is replaced by* $\Pr_{\text{inverted-INFO-basis}}$*).*

Proof. We observe that if the error rate on all the INFO bits together is larger than $p_a + \epsilon$, then at least one of the error rates (on the INFO-Z bits or on the INFO-X bits) must be larger than $p_a + \epsilon$. (Equivalently, if both error rates on the INFO-Z bits and on the INFO-X bits are less than $p_a + \epsilon$, then the error rate on the INFO bits is less than $p_a + \epsilon$.) Namely,

$$\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} > p_a + \epsilon\right) \;\Rightarrow\; \left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{N_{\mathrm{I_Z}}} > p_a + \epsilon\right) \vee \left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{N_{\mathrm{I_X}}} > p_a + \epsilon\right). \tag{7.88}$$

In particular, the corresponding probabilities satisfy

$$\begin{aligned} & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right] \\ \le \; & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{N_{\mathrm{I_Z}}} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right] \\ + \; & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{N_{\mathrm{I_X}}} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right]. \end{aligned} \tag{7.89}$$

This result applies both to the real protocol and to the hypothetical "inverted-INFO-basis" protocol. □

Proposition 7.13. *For any $\epsilon > 0$ and $\delta > 0$,*

$$\begin{aligned} & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{N_{\mathrm{I_Z}}} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right] \\ \le \; & \Pr\left(N_{\mathrm{I_Z}} \le \delta\right) \\ + \; & \max_{\delta \le t_z \le n} \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1) \mid N_{\mathrm{I_Z}} = t_z\right], \end{aligned} \tag{7.90}$$

and

$$\begin{aligned}
& \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{N_{\mathrm{I_X}}} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right] \\
\leq \;& \Pr\left(N_{\mathrm{I_X}} \leq \delta\right) \\
+ \;& \max_{\delta \leq t_x \leq n} \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1) \mid N_{\mathrm{I_X}} = t_x\right]. \qquad (7.91)
\end{aligned}$$

Equations (7.90)–(7.91) *similarly apply to the hypothetical "inverted-INFO-basis" protocol, too (namely, they apply even if* Pr *is replaced by* $\Pr_{\text{inverted-INFO-basis}}$*).*

Proof. First, we prove Equation (7.90):

$$\begin{aligned}
& \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{N_{\mathrm{I_Z}}} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right] \\
= \;& \sum_{t_z} \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1) \mid N_{\mathrm{I_Z}} = t_z\right] \cdot \Pr\left(N_{\mathrm{I_Z}} = t_z\right) \\
= \;& \sum_{t_z < \delta} \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1) \mid N_{\mathrm{I_Z}} = t_z\right] \cdot \Pr\left(N_{\mathrm{I_Z}} = t_z\right) \\
+ \;& \sum_{\delta \leq t_z \leq n} \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1) \mid N_{\mathrm{I_Z}} = t_z\right] \cdot \Pr\left(N_{\mathrm{I_Z}} = t_z\right) \\
\leq \;& \sum_{t_z \leq \delta} \Pr\left(N_{\mathrm{I_Z}} = t_z\right) \\
+ \;& \max_{\delta \leq t_z \leq n} \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1) \mid N_{\mathrm{I_Z}} = t_z\right] \cdot \sum_{\delta \leq t_z \leq n} \Pr\left(N_{\mathrm{I_Z}} = t_z\right) \\
\leq \;& \Pr\left(N_{\mathrm{I_Z}} \leq \delta\right) + \max_{\delta \leq t_z \leq n} \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1) \mid N_{\mathrm{I_Z}} = t_z\right]. \qquad (7.92)
\end{aligned}$$

The proof of Equation (7.91) is similar. Both proofs apply both to the real protocol and to the "inverted-INFO-basis" protocol. □

Theorem 7.14. *Let us be given* $\delta_{\mathrm{sec}} > 0$*,* $\delta_{\mathrm{rel}} > 0$*, and, for infinitely many values of* n*, a family* $\{v_1^n, \ldots, v_{r_n+m_n}^n\}$ *of linearly independent vectors in* $\mathbf{F}_2^n$ *such that* $\delta_{\mathrm{sec}} < \frac{\hat{v}}{n}$ ***and** such that the parity-check matrix* P_{C}*, whose rows are* $\{v_1^n, \ldots, v_{r_n}^n\}$*, defines an error-correcting code that can correct up to* $n \cdot \delta_{\mathrm{rel}}$ *errors on an* n*-bit string. Then for any* $p_a > 0$ *and* $\epsilon_{\mathrm{sec}}, \epsilon_{\mathrm{rel}} > 0$ *such that* $p_a + \epsilon_{\mathrm{sec}} \leq \frac{\delta_{\mathrm{sec}}}{2}$ *and* $p_a + \epsilon_{\mathrm{rel}} \leq \delta_{\mathrm{rel}}$*, for any* $0 < p \leq \frac{1}{2}$*, and for any* $0 < n_z < \frac{pN}{2}$ *and* $0 < n_x < \frac{(1-p)N}{2}$*, it holds for the "efficient BB84" protocol that*

$$\begin{aligned}
\frac{1}{2}\operatorname{tr}\left|\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}\right| \;\leq\;& e^{-\frac{1}{2}Np^2} + e^{-2\left(\frac{n_z}{n+n_z}\right)^2\left(\frac{pN}{2} - n_z\right)\epsilon_{\mathrm{rel}}^2} \\
+\;& e^{-\frac{1}{2}N(1-p)^2} + e^{-2\left(\frac{n_x}{n+n_x}\right)^2\left(\frac{(1-p)N}{2} - n_x\right)\epsilon_{\mathrm{rel}}^2}
\end{aligned}$$

$$
+ \quad \frac{2m_n\sqrt{e^{-\frac{1}{2}Np^2} + e^{-2\left(\frac{n_x}{n+n_x}\right)^2\left(\frac{pN}{2}-n_z\right)\epsilon_{\text{sec}}^2} +}{e^{-\frac{1}{2}N(1-p)^2} + e^{-2\left(\frac{n_z}{n+n_z}\right)^2\left(\frac{(1-p)N}{2}-n_x\right)\epsilon_{\text{sec}}^2}}. \tag{7.93}
$$

Proof. By using Corollary 7.6 and Proposition 7.12, we get the following bound:

$$
\begin{aligned}
& \frac{1}{2}\operatorname{tr}|\rho_{\text{ABE}} - \rho_{\text{U}} \otimes \rho_{\text{E}}| \\
\leq \quad & \Pr\left[\left(\mathbf{k} \neq \mathbf{k}^{\text{B}}\right) \wedge (\mathbf{T} = 1)\right] \\
+ \quad & 2m_n\sqrt{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\text{I}}|}{n} \geq \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T} = 1)\right]} \\
\leq \quad & \Pr\left[\left(\frac{|\mathbf{C}_{\text{I}}|}{n} > p_a + \epsilon_{\text{rel}}\right) \wedge (\mathbf{T} = 1)\right] \\
+ \quad & 2m_n\sqrt{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\text{I}}|}{n} > p_a + \epsilon_{\text{sec}}\right) \wedge (\mathbf{T} = 1)\right]} \\
\leq \quad & \Pr\left[\left(\frac{|\mathbf{C}_{\text{I}_\text{Z}}|}{N_{\text{I}_\text{Z}}} > p_a + \epsilon_{\text{rel}}\right) \wedge (\mathbf{T} = 1)\right] \\
+ \quad & \Pr\left[\left(\frac{|\mathbf{C}_{\text{I}_\text{X}}|}{N_{\text{I}_\text{X}}} > p_a + \epsilon_{\text{rel}}\right) \wedge (\mathbf{T} = 1)\right] \\
+ \quad & \frac{2m_n\sqrt{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\text{I}_\text{Z}}|}{N_{\text{I}_\text{Z}}} > p_a + \epsilon_{\text{sec}}\right) \wedge (\mathbf{T} = 1)\right] +}}{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\text{I}_\text{X}}|}{N_{\text{I}_\text{X}}} > p_a + \epsilon_{\text{sec}}\right) \wedge (\mathbf{T} = 1)\right]}.
\end{aligned} \tag{7.94}
$$

Proposition 7.13 and the definition of T give us the following bounds:

$$
\begin{aligned}
& \Pr\left[\left(\frac{|\mathbf{C}_{\text{I}_\text{Z}}|}{N_{\text{I}_\text{Z}}} > p_a + \epsilon_{\text{rel}}\right) \wedge (\mathbf{T} = 1)\right] \leq \Pr\left(N_{\text{I}_\text{Z}} \leq \frac{pN}{2} - n_z\right) \\
+ \quad & \max_{\frac{pN}{2}-n_z \leq t_z \leq n} \Pr\left[\left(\tfrac{|\mathbf{C}_{\text{I}_\text{Z}}|}{t_z} > p_a + \epsilon_{\text{rel}}\right) \wedge \left(\tfrac{|\mathbf{C}_{\text{T}_\text{Z}}|}{n_z} \leq p_a\right) | N_{\text{I}_\text{Z}} = t_z\right],
\end{aligned} \tag{7.95}
$$

$$
\begin{aligned}
& \Pr\left[\left(\frac{|\mathbf{C}_{\text{I}_\text{X}}|}{N_{\text{I}_\text{X}}} > p_a + \epsilon_{\text{rel}}\right) \wedge (\mathbf{T} = 1)\right] \leq \Pr\left(N_{\text{I}_\text{X}} \leq \frac{(1-p)N}{2} - n_x\right) \\
+ \quad & \max_{\frac{(1-p)N}{2}-n_x \leq t_x \leq n} \Pr\left[\left(\tfrac{|\mathbf{C}_{\text{I}_\text{X}}|}{t_x} > p_a + \epsilon_{\text{rel}}\right) \wedge \left(\tfrac{|\mathbf{C}_{\text{T}_\text{X}}|}{n_x} \leq p_a\right) | N_{\text{I}_\text{X}} = t_x\right],
\end{aligned} \tag{7.96}
$$

$$
\begin{aligned}
& \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\text{I}_\text{Z}}|}{N_{\text{I}_\text{Z}}} > p_a + \epsilon_{\text{sec}}\right) \wedge (\mathbf{T} = 1)\right] \\
\leq \quad & \Pr_{\text{inverted-INFO-basis}}\left(N_{\text{I}_\text{Z}} \leq \frac{pN}{2} - n_z\right) + \max_{\frac{pN}{2}-n_z \leq t_z \leq n} \\
& \Pr_{\text{inverted-INFO-basis}}\left[\left(\tfrac{|\mathbf{C}_{\text{I}_\text{Z}}|}{t_z} > p_a + \epsilon_{\text{sec}}\right) \wedge \left(\tfrac{|\mathbf{C}_{\text{T}_\text{X}}|}{n_x} \leq p_a\right) | N_{\text{I}_\text{Z}} = t_z\right],
\end{aligned} \tag{7.97}
$$

$$
\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\text{I}_\text{X}}|}{N_{\text{I}_\text{X}}} > p_a + \epsilon_{\text{sec}}\right) \wedge (\mathbf{T} = 1)\right]
$$

$$\leq \Pr_{\text{inverted-INFO-basis}}\left(N_{\mathrm{I_X}} \leq \frac{(1-p)N}{2} - n_x\right) + \max_{\frac{(1-p)N}{2} - n_x \leq t_x \leq n} \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a + \epsilon_{\mathrm{sec}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z} \leq p_a\right) | N_{\mathrm{I_X}} = t_x\right]. \quad (7.98)$$

For each one of Equations (7.95)–(7.98), we need to upper-bound two probabilities. For bounding the first set of probabilities, we use the results of Corollary 2.4:

$$\Pr\left(|\mathbf{b}| \leq \frac{(1-p)N}{2}\right) \leq e^{-\frac{1}{2}N(1-p)^2}, \quad (7.99)$$

$$\Pr\left(|\overline{\mathbf{b}}| \leq \frac{pN}{2}\right) \leq e^{-\frac{1}{2}Np^2}. \quad (7.100)$$

We notice that $|\mathbf{b}| = N_{\mathrm{I_X}} + n_x$ and $|\overline{\mathbf{b}}| = N_{\mathrm{I_Z}} + n_z$; therefore,

$$\Pr\left(N_{\mathrm{I_X}} \leq \frac{(1-p)N}{2} - n_x\right) \leq e^{-\frac{1}{2}N(1-p)^2}, \quad (7.101)$$

$$\Pr\left(N_{\mathrm{I_Z}} \leq \frac{pN}{2} - n_z\right) \leq e^{-\frac{1}{2}Np^2}, \quad (7.102)$$

$$\Pr_{\text{inverted-INFO-basis}}\left(N_{\mathrm{I_X}} \leq \frac{(1-p)N}{2} - n_x\right) \leq e^{-\frac{1}{2}N(1-p)^2}, \quad (7.103)$$

$$\Pr_{\text{inverted-INFO-basis}}\left(N_{\mathrm{I_Z}} \leq \frac{pN}{2} - n_z\right) \leq e^{-\frac{1}{2}Np^2}. \quad (7.104)$$

For bounding the second set of probabilities, *given* specific values of $N_{\mathrm{I_Z}} = t_z$ and $N_{\mathrm{I_X}} = t_x$, we use Corollary 2.2:

In the real protocol, the INFO-Z and TEST-Z bits are sent and measured in the z basis, while the INFO-X and TEST-X bits are sent and measured in the x basis. Therefore, the random and uniform sampling of the t_z INFO-Z bits out of the $t_z + n_z$ bits sent in the z basis (assuming that the INFO-X and TEST-X bits have already been chosen) does not affect the bases in the real protocol; similarly, the random and uniform sampling of the t_x INFO-X bits out of the $t_x + n_x$ bits sent in the x basis (assuming that the INFO-Z and TEST-Z bits have already been chosen) does not affect the bases in the real protocol. We note that these samplings *are* uniform, because the probability $\Pr(\mathbf{s} \mid \mathbf{b})$ is uniform for all the allowed values of $\mathbf{b}$ and $\mathbf{s}$. This means that we can apply Corollary 2.2 to both of these samplings, and we get

$$\Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon_{\mathrm{rel}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z} \leq p_a\right) \mid N_{\mathrm{I_Z}} = t_z\right] \leq e^{-2\left(\frac{n_z}{t_z+n_z}\right)^2 t_z \epsilon_{\mathrm{rel}}^2}, \quad (7.105)$$

$$\Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a + \epsilon_{\mathrm{rel}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \leq p_a\right) \mid N_{\mathrm{I_X}} = t_x\right] \leq e^{-2\left(\frac{n_x}{t_x+n_x}\right)^2 t_x \epsilon_{\mathrm{rel}}^2}. \quad (7.106)$$

Maximizing over t_z and t_x, we get:

$$\begin{aligned}
&\max_{\frac{pN}{2}-n_z\le t_z\le n} \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a+\epsilon_{\mathrm{rel}}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z}\le p_a\right)\mid N_{\mathrm{I_Z}}=t_z\right] \\
\le\ & e^{-2\left(\frac{n_z}{n+n_z}\right)^2\left(\frac{pN}{2}-n_z\right)\epsilon_{\mathrm{rel}}^2}, \qquad (7.107)\\
&\max_{\frac{(1-p)N}{2}-n_x\le t_x\le n} \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a+\epsilon_{\mathrm{rel}}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x}\le p_a\right)\mid N_{\mathrm{I_X}}=t_x\right] \\
\le\ & e^{-2\left(\frac{n_x}{n+n_x}\right)^2\left(\frac{(1-p)N}{2}-n_x\right)\epsilon_{\mathrm{rel}}^2}. \qquad (7.108)
\end{aligned}$$

In the hypothetical “inverted-INFO-basis” protocol, the INFO-X and TEST-Z bits are sent and measured in the z basis, while the INFO-Z and TEST-X bits are sent and measured in the x basis. Therefore, the random and uniform sampling of the t_x INFO-X bits out of the t_x+n_z bits sent in the z basis (assuming that the INFO-Z and TEST-X bits have already been chosen) does not affect the bases in the hypothetical protocol; similarly, the random and uniform sampling of the t_z INFO-Z bits out of the t_z+n_x bits sent in the x basis (assuming that the INFO-X and TEST-Z bits have already been chosen) does not affect the bases in the hypothetical protocol. We note that these samplings *are* uniform, because the probability $\Pr(\mathbf{b})$ depends only on $|\mathbf{b}|$ and is invariant to permutations. This means that we can apply Corollary 2.2 to both of these samplings, and we get

$$\begin{aligned}
&\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a+\epsilon_{\mathrm{sec}}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x}\le p_a\right)\mid N_{\mathrm{I_Z}}=t_z\right] \\
\le\ & e^{-2\left(\frac{n_x}{t_z+n_x}\right)^2 t_z\epsilon_{\mathrm{sec}}^2}, \qquad (7.109)\\
&\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a+\epsilon_{\mathrm{sec}}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z}\le p_a\right)\mid N_{\mathrm{I_X}}=t_x\right] \\
\le\ & e^{-2\left(\frac{n_z}{t_x+n_z}\right)^2 t_x\epsilon_{\mathrm{sec}}^2}. \qquad (7.110)
\end{aligned}$$

Maximizing over t_z and t_x, we get:

$$\begin{aligned}
&\max_{\frac{pN}{2}-n_z\le t_z\le n} \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a+\epsilon_{\mathrm{sec}}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x}\le p_a\right)\mid N_{\mathrm{I_Z}}=t_z\right] \\
\le\ & e^{-2\left(\frac{n_x}{n+n_x}\right)^2\left(\frac{pN}{2}-n_z\right)\epsilon_{\mathrm{sec}}^2}, \qquad (7.111)\\
&\max_{\frac{(1-p)N}{2}-n_x\le t_x\le n} \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a+\epsilon_{\mathrm{sec}}\right)\wedge\left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z}\le p_a\right)\mid N_{\mathrm{I_X}}=t_x\right] \\
\le\ & e^{-2\left(\frac{n_z}{n+n_z}\right)^2\left(\frac{(1-p)N}{2}-n_x\right)\epsilon_{\mathrm{sec}}^2}. \qquad (7.112)
\end{aligned}$$

To sum up, we get the following bound:

$$\frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}}-\rho_{\mathrm{U}}\otimes\rho_{\mathrm{E}}|$$

$$
\begin{aligned}
\le \quad & e^{-\frac{1}{2}Np^2} + e^{-2\left(\frac{n_z}{n+n_z}\right)^2\left(\frac{pN}{2}-n_z\right)\epsilon_{\mathrm{rel}}^2} \\
+ \quad & e^{-\frac{1}{2}N(1-p)^2} + e^{-2\left(\frac{n_x}{n+n_x}\right)^2\left(\frac{(1-p)N}{2}-n_x\right)\epsilon_{\mathrm{rel}}^2} \\
+ \quad & 2m_n\sqrt{e^{-\frac{1}{2}Np^2} + e^{-2\left(\frac{n_x}{n+n_x}\right)^2\left(\frac{pN}{2}-n_z\right)\epsilon_{\mathrm{sec}}^2}+} \\
& \overline{e^{-\frac{1}{2}N(1-p)^2} + e^{-2\left(\frac{n_z}{n+n_z}\right)^2\left(\frac{(1-p)N}{2}-n_x\right)\epsilon_{\mathrm{sec}}^2}}.
\end{aligned}
\tag{7.113}
$$

□

Similarly to the standard BB84, we get the following results on vector families satisfying the requirements of Theorem 7.14: the bit-rate satisfies

$$R_{\mathrm{secret}} \triangleq \frac{m_n}{n} < 1 - H_2(2p_a + 2\epsilon_{\mathrm{sec}}) - H_2\left(p_a + \epsilon_{\mathrm{rel}} + \frac{1}{n}\right), \tag{7.114}$$

and the condition on the asymptotic error rate threshold is

$$H_2(2p_a + 2\epsilon_{\mathrm{sec}}) + H_2\left(p_a + \epsilon_{\mathrm{rel}} + \frac{1}{n}\right) < 1. \tag{7.115}$$

This gives an asymptotic error rate threshold of 7.56%.

7.3.4 The "Modified Efficient BB84" Protocol

A relatively minor property of the definition of the "efficient BB84" protocol in [LCA05] (and in Subsection 7.3.3) makes both the security bound and the security proof pretty complicated. In this subsection, we describe a modified protocol that has an easier security proof. The only modification in this protocol is setting the number of INFO-Z and INFO-X bits to be fixed, rather than letting them vary probabilistically. This change simplifies the description of the protocol, because it is no longer needed to set the probability p and to treat illegal choices of $\mathbf{b}, \mathbf{s}$ (see Remark 1); and it also simplifies the security proof, because it is no longer needed to probabilistically analyze the numbers of INFO-Z and INFO-X bits (as done in Subsection 7.3.3).

In the "modified efficient BB84" protocol, the strings $\mathbf{b}$ and $\mathbf{s}$ together define a random partition of the set of indexes $\{1, 2, \ldots, N\}$ into four disjoint sets:

- $\mathrm{I_Z}$ (INFO-Z bits, where $s_j = 1$ and $b_j = 0$) of size t_z;
- $\mathrm{I_X}$ (INFO-X bits, where $s_j = 1$ and $b_j = 1$) of size t_x;
- $\mathrm{T_Z}$ (TEST-Z bits, where $s_j = 0$ and $b_j = 0$) of size n_z; and
- $\mathrm{T_X}$ (TEST-X bits, where $s_j = 0$ and $b_j = 1$) of size n_x.

Formally, in "modified efficient BB84", Alice and Bob agree on parameters t_z, t_x, n_z, n_x (such that $N = n+n_z+n_x$ and $n = t_z+t_x$), and we choose $B = \{\mathbf{b} \in \mathbf{F}_2^N \mid |\mathbf{b}| = t_x+n_x\}$ and $S_{\mathbf{b}} = \{\mathbf{s} \in \mathbf{F}_2^N \mid (|\mathbf{s}| = n) \wedge (|\overline{\mathbf{s}} \wedge \mathbf{b}| = n_x)\}$ for all $\mathbf{b} \in B$ (namely, it is required that there are t_z INFO-Z bits, t_x INFO-X bits, n_z TEST-Z bits, and n_x TEST-X bits). The probability distributions $\Pr(\mathbf{b})$ and $\Pr(\mathbf{s} \mid \mathbf{b})$ are uniform (because $|\mathbf{b}|$, which is the only parameter that affects $\Pr(\mathbf{b})$ in Subsection 7.3.3, is fixed in the modified protocol).

Alice and Bob also agree on an error rate threshold, p_a (applied *both* to the TEST-Z bits and to the TEST-X bits). The testing function T is defined as follows:

$$T(\mathbf{i}_{\mathrm{T}} \oplus \mathbf{j}_{\mathrm{T}}, \mathbf{b}_{\mathrm{T}}, \mathbf{s}) = 1 \ \Leftrightarrow\ (|\mathbf{i}_{\mathrm{T_Z}} \oplus \mathbf{j}_{\mathrm{T_Z}}| \le n_z \cdot p_a) \wedge (|\mathbf{i}_{\mathrm{T_X}} \oplus \mathbf{j}_{\mathrm{T_X}}| \le n_x \cdot p_a)\,. \tag{7.116}$$

Namely, the test passes if and only if the error rate on the TEST-Z bits is at most p_a *and* the error rate on the TEST-X bits is at most p_a.

Proposition 7.15. *For any $\epsilon > 0$,*

$$\begin{aligned} & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right] \\ \le\ & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right] \\ +\ & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a + \epsilon\right) \wedge (\mathbf{T} = 1)\right]. \end{aligned} \tag{7.117}$$

Equation (7.117) *similarly applies to the hypothetical "inverted-INFO-basis" protocol, too (namely, it applies even if* $\Pr$ *is replaced by* $\Pr_{\text{inverted-INFO-basis}}$*).*

Proof. The same proof as Proposition 7.12. □

Theorem 7.16. *Let us be given $\delta_{\mathrm{sec}} > 0$, $\delta_{\mathrm{rel}} > 0$, and, for infinitely many values of n, a family $\{v_1^n, \ldots, v_{r_n+m_n}^n\}$ of linearly independent vectors in $\mathbf{F}_2^n$ such that $\delta_{\mathrm{sec}} < \frac{\hat{v}}{n}$* ***and*** *such that the parity-check matrix P_{C}, whose rows are $\{v_1^n, \ldots, v_{r_n}^n\}$, defines an error-correcting code that can correct up to $n \cdot \delta_{\mathrm{rel}}$ errors on an n-bit string. Then for any $p_a > 0$ and $\epsilon_{\mathrm{sec}}, \epsilon_{\mathrm{rel}} > 0$ such that $p_a + \epsilon_{\mathrm{sec}} \le \frac{\delta_{\mathrm{sec}}}{2}$ and $p_a + \epsilon_{\mathrm{rel}} \le \delta_{\mathrm{rel}}$, and for any $t_z, t_x, n_z, n_x > 0$ such that $n = t_z + t_x$, it holds for the "modified efficient BB84" protocol that*

$$\begin{aligned} & \frac{1}{2} \operatorname{tr} |\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}| \\ \le\ & e^{-2\left(\frac{n_z}{t_z+n_z}\right)^2 t_z \epsilon_{\mathrm{rel}}^2} + e^{-2\left(\frac{n_x}{t_x+n_x}\right)^2 t_x \epsilon_{\mathrm{rel}}^2} \\ +\ & 2m_n \sqrt{e^{-2\left(\frac{n_x}{t_z+n_x}\right)^2 t_z \epsilon_{\mathrm{sec}}^2} + e^{-2\left(\frac{n_z}{t_x+n_z}\right)^2 t_x \epsilon_{\mathrm{sec}}^2}}. \end{aligned} \tag{7.118}$$

Proof. By using Corollary 7.6 and Proposition 7.15, we get the following bound:

$$\begin{aligned} & \frac{1}{2} \operatorname{tr} |\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}| \\ \le\ & \Pr\left[(\mathbf{k} \ne \mathbf{k}^{\mathrm{B}}) \wedge (\mathbf{T} = 1)\right] \\ +\ & 2m_n \sqrt{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} \ge \frac{\hat{v}}{2n}\right) \wedge (\mathbf{T} = 1)\right]} \end{aligned}$$

$$
\begin{aligned}
\leq\ & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} > p_a + \epsilon_{\mathrm{rel}}\right) \wedge (\mathbf{T} = 1)\right] \\
+\ & 2m_n\sqrt{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I}}|}{n} > p_a + \epsilon_{\mathrm{sec}}\right) \wedge (\mathbf{T} = 1)\right]} \\
\leq\ & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon_{\mathrm{rel}}\right) \wedge (\mathbf{T} = 1)\right] \\
+\ & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a + \epsilon_{\mathrm{rel}}\right) \wedge (\mathbf{T} = 1)\right] \\
+\ & 2m_n\sqrt{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon_{\mathrm{sec}}\right) \wedge (\mathbf{T} = 1)\right] + \atop \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a + \epsilon_{\mathrm{sec}}\right) \wedge (\mathbf{T} = 1)\right]} \\
\leq\ & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon_{\mathrm{rel}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z} \leq p_a\right)\right] \\
+\ & \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a + \epsilon_{\mathrm{rel}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \leq p_a\right)\right] \\
+\ & 2m_n\sqrt{\Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon_{\mathrm{sec}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \leq p_a\right)\right] + \atop \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a + \epsilon_{\mathrm{sec}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z} \leq p_a\right)\right]}. \qquad (7.119)
\end{aligned}
$$

For bounding these probabilities, we use Corollary 2.2:

In the real protocol, the INFO-Z and TEST-Z bits are sent and measured in the z basis, while the INFO-X and TEST-X bits are sent and measured in the x basis. Therefore, the random and uniform sampling of the t_z INFO-Z bits out of the $t_z + n_z$ bits sent in the z basis (assuming that the INFO-X and TEST-X bits have already been chosen) does not affect the bases in the real protocol; similarly, the random and uniform sampling of the t_x INFO-X bits out of the $t_x + n_x$ bits sent in the x basis (assuming that the INFO-Z and TEST-Z bits have already been chosen) does not affect the bases in the real protocol. This means that we can apply Corollary 2.2 to both of these samplings, and we get

$$
\begin{aligned}
& \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon_{\mathrm{rel}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z} \leq p_a\right)\right] \\
\leq\ & e^{-2\left(\frac{n_z}{t_z+n_z}\right)^2 t_z \epsilon_{\mathrm{rel}}^2}, \qquad (7.120) \\
& \Pr\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a + \epsilon_{\mathrm{rel}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \leq p_a\right)\right] \\
\leq\ & e^{-2\left(\frac{n_x}{t_x+n_x}\right)^2 t_x \epsilon_{\mathrm{rel}}^2}. \qquad (7.121)
\end{aligned}
$$

In the hypothetical "inverted-INFO-basis" protocol, the INFO-X and TEST-Z bits are sent and measured in the z basis, while the INFO-Z and TEST-X bits are sent and measured in the x basis. Therefore, the random and uniform sampling of the t_x INFO-X bits out of the $t_x + n_z$ bits sent in the z basis (assuming that the INFO-Z and TEST-X bits have already been chosen) does not affect the bases in the hypothetical protocol; similarly, the random and uniform sampling of the t_z INFO-Z bits out of the $t_z + n_x$ bits sent in the x basis (assuming that the INFO-X and TEST-Z bits have already been chosen) does not affect the bases in the hypothetical protocol. This means that we can apply Corollary 2.2 to both of these samplings, and we get

$$\begin{aligned} & \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I_Z}}|}{t_z} > p_a + \epsilon_{\mathrm{sec}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_X}}|}{n_x} \leq p_a\right)\right] \\ \leq\ & e^{-2\left(\frac{n_x}{t_z+n_x}\right)^2 t_z \epsilon_{\mathrm{sec}}^2}, \qquad (7.122) \\ & \Pr_{\text{inverted-INFO-basis}}\left[\left(\frac{|\mathbf{C}_{\mathrm{I_X}}|}{t_x} > p_a + \epsilon_{\mathrm{sec}}\right) \wedge \left(\frac{|\mathbf{C}_{\mathrm{T_Z}}|}{n_z} \leq p_a\right)\right] \\ \leq\ & e^{-2\left(\frac{n_z}{t_x+n_z}\right)^2 t_x \epsilon_{\mathrm{sec}}^2}. \qquad (7.123) \end{aligned}$$

To sum up, we get the following bound:

$$\begin{aligned} & \frac{1}{2}\operatorname{tr}|\rho_{\mathrm{ABE}} - \rho_{\mathrm{U}} \otimes \rho_{\mathrm{E}}| \\ \leq\ & e^{-2\left(\frac{n_z}{t_z+n_z}\right)^2 t_z \epsilon_{\mathrm{rel}}^2} + e^{-2\left(\frac{n_x}{t_x+n_x}\right)^2 t_x \epsilon_{\mathrm{rel}}^2} \\ +\ & 2m_n \sqrt{e^{-2\left(\frac{n_x}{t_z+n_x}\right)^2 t_z \epsilon_{\mathrm{sec}}^2} + e^{-2\left(\frac{n_z}{t_x+n_z}\right)^2 t_x \epsilon_{\mathrm{sec}}^2}}. \qquad (7.124) \end{aligned}$$

□

Similarly to the standard BB84 and to the "efficient BB84" protocols, we get the following results on vector families satisfying the requirements of Theorem 7.16: the bit-rate satisfies

$$R_{\mathrm{secret}} \triangleq \frac{m_n}{n} < 1 - H_2(2p_a + 2\epsilon_{\mathrm{sec}}) - H_2\left(p_a + \epsilon_{\mathrm{rel}} + \frac{1}{n}\right), \qquad (7.125)$$

and the condition on the asymptotic error rate threshold is

$$H_2(2p_a + 2\epsilon_{\mathrm{sec}}) + H_2\left(p_a + \epsilon_{\mathrm{rel}} + \frac{1}{n}\right) < 1. \qquad (7.126)$$

This gives an asymptotic error rate threshold of 7.56%.

Chapter 8

From Practice to Theory: the "Bright Illumination" Attack on Quantum Key Distribution Systems

In this chapter, we explain how the practical "Bright Illumination" attack on QKD systems can be described as a theoretical "Reversed-Space" attack.

This chapter is based on a paper published in the 9th International Conference on the Theory and Practice of Natural Computing (TPNC) in 2020 by Rotem Liss and Tal Mor [LM20].

8.1 Introduction

In the area of quantum information processing, theory usually precedes experiment. For example, the BB84 protocol was suggested in 1984 [BB84], five years before it was implemented [BBBSS92], and it still cannot be implemented in a perfectly secure way even today [LCT14, SK14]. The "Photon-Number Splitting" attack was suggested in 2000 [BLMS00], but it is not implementable today. Quantum computing was suggested in the 1980s [Deu85, Fey82, Ben80], but no useful and universal quantum computer (with a large number of clean qubits) has been implemented until today [Pre18]. The same applies to Shor's factorization algorithm [Sho94, Sho99], quantum teleportation [BBCJPW93] (at least to some extent; see also [PHB$^+$14]), and many other examples.

In contrast to the above examples, the "Bright Illumination" attack against practical QKD systems was presented and fully implemented in 2010 [LWWESM10], *prior* to any theoretical prediction of the possibility of such an attack.

In this chapter, we ask the question: could the "Bright Illumination" attack have been theoretically predicted? How can the "Bright Illumination" attack be theoretically modeled (even approximately) by using the Fock space notations? We show that the "Bright Illumination" attack can be modeled as a "Reversed-Space" attack [GM12] (or, more generally, as a "Quantum Space" attack [GM07, Gel08, GM12]) and that this attack and similar attacks could and should have been suggested by theoreticians.

8.2 Imperfections in Experimental Implementation of QKD

In this chapter, we usually consider the polarization-based implementations of BB84 discussed in Subsection 2.5.2, in which $|\mathtt{0}\rangle = |\leftrightarrow\rangle$, $|\mathtt{1}\rangle = |\updownarrow\rangle$, $|+\rangle = |\swarrow\!\!\!\!\!\nearrow\rangle$, and $|-\rangle = |\nwarrow\!\!\!\!\!\searrow\rangle$. For describing the practical system, we use the Fock space notations described in Subsection 2.5.1, in which the $|m_1, m_0\rangle$ state represents m_1 indistinguishable photons in the $|\mathtt{1}\rangle$ mode and m_0 indistinguishable photons in the $|\mathtt{0}\rangle$ mode.

Two important examples of imperfections (see [GM16]) are highly relevant to various "Reversed-Space" attacks. As we show in this chapter, these two imperfections must be *combined* for understanding the "Bright Illumination" attack.

Imperfection 1: Our realistic assumption, which is true for standard detectors in QKD implementations, is that Bob's detectors cannot *count* the number of photons in a pulse. Thus, they cannot distinguish *all* Fock states $|m\rangle$ from one another, but can only distinguish the Fock state $|0\rangle$ (a lack of photons) from the Fock states $\{|m\rangle : m \geq 1\}$. Namely, standard detectors can only decide whether the mode is empty ($m = 0$) or has at least one photon ($m > 0$). In contrast, we assume that Eve can (in principle) do anything allowed by the laws of quantum physics; in particular, Eve may have such "photon counters".

In particular, let us assume that there are *two* pulses, each of them consisting of a single mode. Bob cannot know whether a pulse contains one photon or two photons; therefore, he cannot distinguish between $|1\rangle|0\rangle$ and $|2\rangle|0\rangle$ (and, similarly, he cannot distinguish between $|0\rangle|1\rangle$ and $|0\rangle|2\rangle$). For example, assume that Alice sends the $|1\rangle|0\rangle$ state (a qubit) to Bob, and Eve replaces Alice's state by $|2\rangle|0\rangle$ and sends it to Bob instead (or, similarly, assume that Eve replaces $|0\rangle|1\rangle$ by $|0\rangle|2\rangle$). In this case, Bob cannot notice the change, and no error can occur; still, Bob got a state he had not expected to get. It may be possible for Eve to take advantage of this fact in a fully-designed attack.

Imperfection 2: Our realistic assumption is that Bob cannot know exactly *when* the photon he measured arrived. For example (in a polarization-based implementation):

- Alice's ideal qubit arrives at time t (states denoted $|0,1\rangle_t|0,0\rangle_{t+\delta}$, $|1,0\rangle_t|0,0\rangle_{t+\delta}$).

- Eve's photon may arrive at time $t+\delta$ (states denoted $|0,0\rangle_t|0,1\rangle_{t+\delta}$, $|0,0\rangle_t|1,0\rangle_{t+\delta}$).

Again, Eve may take advantage of this fact in a fully-designed attack.

Similar imperfections can be found if Bob cannot know exactly what the *wavelength* of the photon is, or *where* the photon arrives.

The conceptual difference between the two imperfections is in whether Bob can (ideally) avoid measuring the extra states sent by Eve, or not:

- In Imperfection 1, Eve may send more than one photon, and Bob must measure the state (while he cannot count the number of photons using current technology).
- In Imperfection 2, Eve sends states in two separate subsystems. Bob can, in principle, ignore the "wrong" subsystem in case he knows for sure it has not been sent by Alice.

8.3 The "Bright Illumination" Attack

The "Bright Illumination" blinding attack [LWWESM10] works against QKD systems that use Avalanche Photodiodes (APDs) as Bob's detectors. As an example, we describe below the implementation of this attack against a system implementing the BB84 protocol in a polarization-based scheme, but it is important to note that the attack can be adapted to most QKD protocols and implementations that use APDs [LWWESM10].

The APDs can be operated in two "modes of operation": the "linear mode" that detects only a light beam above a specific power threshold, and "Geiger mode" that detects even a single photon (but cannot count the number of photons). In this attack, the adversary Eve sends a continuous strong light beam towards Bob's detectors, causing them to operate *only* in the linear mode (thus "blinding" the detectors).

After Bob's detectors have been blinded (and in parallel to sending the continuous strong beam, making sure they are kept blind), Eve performs a "measure-resend" attack: she detects the qubit (single photon) sent by Alice, measures it in one of the two bases (exactly as Bob would do), and sends to Bob a *strong* light beam depending on the state she measured, a little above the power threshold of the detectors. For example, if Eve measures the state $|1,0\rangle$, she sends to Bob the state $|m,0\rangle$ for $m \gg 1$. Now, if Bob chooses the same basis as Eve, he will measure the same result as Eve; and if Bob chooses a different basis, he will measure nothing, because the strong light beam will get split between the two detectors. This means that Bob will always either measure the same result as Eve or lose the bit.

In the end, Bob and Eve have exactly the same information, so Eve can copy Bob's classical post-processing and get the same final key as Alice and Bob do. Moreover, Eve's attack causes no detectable disturbance, because Bob does not know that his detectors have operated in the wrong mode of operation; the only effect is a loss rate of 50% (that is not problematic: the loss rate for the single photons sent by Alice is usually much higher, so Eve can cause Bob to get the same loss rate he expects to get).

This attack was both developed and experimentally demonstrated against commercial QKD systems by [LWWESM10]. See [LWWESM10] for more details and for diagrams.

8.4 "Reversed-Space" Attacks

The "Reversed-Space" methodology, described in [Gel08, GM16, GM12], is a theoretical framework of attacks exploiting the imperfections of Bob. This methodology is a special case (easier to analyze) of the more general methodology of "Quantum Space" attacks [GM07, Gel08], that exploits the imperfections of *both* Alice and Bob; the "Reversed-Space" methodology assumes Alice to be ideal and only exploits Bob's imperfections [Gel08, GM12, BGM14, GM16]. (Another special case of a "Quantum Space" attack is the "Photon-Number Splitting" attack described in Subsection 2.5.3.)

In the ideal QKD protocol, Bob expects to get from Alice a state in the Hilbert space $\mathcal{H}^{\mathrm{A}}$; however, in the "Reversed-Space" attack, Bob gets from Eve an unexpected state, residing in a larger Hilbert space called the "space of the protocol" and denoted by $\mathcal{H}^{\mathrm{P}}$. In principle, Eve could have used a huge space $\mathcal{H}'$ such that $\mathcal{H}^{\mathrm{A}} \subseteq \mathcal{H}^{\mathrm{P}} \subseteq \mathcal{H}'$: the huge Hilbert space $\mathcal{H}'$ consists of *all* the quantum states that Eve *can possibly* send to Bob, but it is too large, and most of it is irrelevant.

Because "Reversed-Space" attacks assume a "perfect Alice" (sending prefect qubits), it is usually easy to find the *relevant* subspace $\mathcal{H}^{\mathrm{P}}$, as we demonstrate by three examples below; $\mathcal{H}^{\mathrm{P}}$ is only enlarged (relative to the ideal space $\mathcal{H}^{\mathrm{A}}$) by Bob's imperfections. Therefore, $\mathcal{H}^{\mathrm{P}}$ is the space that includes all the states that may be useful for Eve to send to Bob. The space $\mathcal{H}^{\mathrm{P}}$ is defined by taking all the possible measurement results of Bob and reversing them in time; more precisely, it is the span of all the states in $\mathcal{H}^{\mathrm{A}}$ *and* all the states that Eve can send to Bob so that he gets the measurement results she desires.

Whether Bob is aware of it or not, his experimental setting treats not only the states in $\mathcal{H}^{\mathrm{A}}$, but all the possible inputs in the "space of the protocol" $\mathcal{H}^{\mathrm{P}}$. Bob then classifies them into three classes: (1) valid states from Alice, (2) losses, and (3) invalid states. *Valid states* are always treated in conventional security analysis: a random subset is compared with Alice for estimating the error rate, and then the final key is obtained using the error correction and privacy amplification processes. *Losses* are expected, and they are not counted as noise. *Invalid states* are usually counted as errors (noise), but they do not appear in ideal analyses of ideal protocols. We note that loss rate and error rate are computed separately: the error rate must be small (e.g., around 10%) for the protocol not to be aborted by Alice and Bob, while the loss rate can be much higher (even higher than 99%). Any "Reversed-Space" attack takes advantage of the possibility that Bob treats some states in $\mathcal{H}^{\mathrm{P}}$ in the wrong way, because he does not expect to get these states.

Eve's attack is called "Reversed-Space" because Eve can devise her attack by looking at Bob's possible measurement results: Eve finds a measurement result she wants to be obtained by Bob (because he interprets it in a way desired by her) and reverses the measurement result in time for finding the state in $\mathcal{H}^{\mathrm{P}}$ she should send to Bob. In particular, if Bob applies the unitary operation $\mathcal{U}_{\mathrm{B}}$ on his state prior to his measurement, Eve should apply the inverted operation $\mathcal{U}_{\mathrm{B}}^{-1} = \mathcal{U}_{\mathrm{B}}^{\dagger}$ to each state corresponding to each possible measurement outcome of Bob.

We present three examples of "Reversed-Space" attacks. For simplicity, we only consider BB84 implemented in a polarization-based scheme (as described in Subsection 2.5.2 and Section 8.2), but the attacks may be generalized to other implementations, too. We emphasize that all three examples have been chosen to satisfy two conditions, also satisfied by the "Bright Illumination" attack: (a) Eve performs a "measure-resend" attack in a basis she chooses randomly, and (b) it is possible for Eve to get full information without inducing noise.

Example 1 (a special case of the "Trojan Pony" attack [GLLP04]): This example exploits Imperfection 1 described in Section 8.2, and it assumes Bob uses an "active" basis choice (see Subsection 2.5.2).

In this attack, Eve performs a "measure-resend" attack—namely, she measures each qubit state sent from Alice to Bob in a random basis, and resends "it" towards Bob. However, instead of resending it as a single photon, she resends a huge number of photons towards Bob: she sends many *identical* photons, all with the same polarization as the state she measured ($|\mathtt{0}\rangle$, $|\mathtt{1}\rangle$, $|+\rangle$, or $|-\rangle$). If Bob chooses the same basis as Eve, he will get the same result as her, because Imperfection 1 causes his system to treat the incoming states $|0, m\rangle$ and $|m, 0\rangle$ (for any $m \geq 1$) as if they were $|0, 1\rangle$ and $|1, 0\rangle$, respectively; but if he chooses a different basis from Eve, both of his detectors will (almost surely) click. If Bob decides to treat this *invalid* event (a two-detector click) as an "error", the error rate will be around 50%, so Alice and Bob will abort the protocol; but if he naively decides to treat this event as a "loss", Eve can get full information without inducing errors.

Alice sends an ideal qubit (a single photon), while Eve may send any number of photons. Therefore, using the Fock space notations, $\mathcal{H}^{\mathrm{A}} = \mathcal{H}_2 \triangleq \mathrm{Span}\{|0, 1\rangle, |1, 0\rangle\}$ and $\mathcal{H}^{\mathrm{P}} = \mathrm{Span}\{|m_1, m_0\rangle : m_1, m_0 \geq 0\}$.

Example 2 (a special case of the "Faked States" attack [MH05, MAS06, Gel08]): This attack exploits Imperfection 2 described in Section 8.2. We assume that Bob has four detectors (namely, that he uses the "passive" basis choice variant of the polarization-based encoding: see Subsection 2.5.2), and that his detectors have different (but overlapping) *time gates* during which they are sensitive: given the three different times $t_0 < t_{1/2} < t_1$, the detectors for the computational basis are sensitive only to pulses sent at t_0 or $t_{1/2}$ (or in between), and the detectors for the Hadamard basis are sensitive only to pulses sent at $t_{1/2}$ or t_1 (or in between). Alice normally sends her pulses at $t_{1/2}$ (when both detectors are sensitive), but Eve may send her pulses at

t_0, $t_{1/2}$, or t_1.

Eve performs a "measure-resend" attack by measuring Alice's state in a random basis, and resending it towards Bob as follows: if Eve measures in the computational basis, she resends the state at time t_0; and if she measures in the Hadamard basis, she resends the state at time t_1. Therefore, Bob gets the same result as Eve if he measures in the same basis as hers, but he gets a loss otherwise (because Bob's detectors for the other basis are not sensitive at that timing). This means that Eve gets full information without inducing any error.

Using the same notations as in Imperfection 2, the state $|m_1, m_0\rangle_{t_0}|n_1, n_0\rangle_{t_{1/2}}|o_1, o_0\rangle_{t_1}$ consists of the Fock states $|m_1, m_0\rangle$ sent at time t_0, $|n_1, n_0\rangle$ sent at time $t_{1/2}$, and $|o_1, o_0\rangle$ sent at time t_1. Alice sends an ideal qubit (a single photon at time $t_{1/2}$), while Eve may send a single photon at any of the times t_0, $t_{1/2}$, or t_1, or a superposition.

Therefore, $\mathcal{H}^{\mathrm{A}} = \mathcal{H}_2 \triangleq \mathrm{Span}\{|0,0\rangle_{t_0}|0,1\rangle_{t_{1/2}}|0,0\rangle_{t_1}\ ,\ |0,0\rangle_{t_0}|1,0\rangle_{t_{1/2}}|0,0\rangle_{t_1}\}$ and $\mathcal{H}^{\mathrm{P}} = \mathrm{Span}\{|0,1\rangle_{t_0}|0,0\rangle_{t_{1/2}}|0,0\rangle_{t_1}\ ,\ |1,0\rangle_{t_0}|0,0\rangle_{t_{1/2}}|0,0\rangle_{t_1}\ ,\ |0,0\rangle_{t_0}|0,1\rangle_{t_{1/2}}|0,0\rangle_{t_1}\ ,\ |0,0\rangle_{t_0}|1,0\rangle_{t_{1/2}}|0,0\rangle_{t_1}\ ,\ |0,0\rangle_{t_0}|0,0\rangle_{t_{1/2}}|0,1\rangle_{t_1}\ ,\ |0,0\rangle_{t_0}|0,0\rangle_{t_{1/2}}|1,0\rangle_{t_1}\}$.

Example 3 (the "Fixed Apparatus" attack [BGM14]) can be applied by Eve if Bob uses a "passive" basis choice (see Subsection 2.5.2). In this attack, Eve sends to Bob an unexpected state, and this state "forces" Bob to obtain the basis Eve wants. This attack makes it possible for Eve to force Bob choose the same basis as her (and, therefore, get the same outcome as her), thus stealing the whole key without inducing any errors or losses. The attack is only possible if Eve has a one-time access to Bob's laboratory, because it requires Eve to first compromise Bob's device (otherwise, she cannot send him that unexpected state).

Assume that Bob uses a polarization-independent beam splitter that splits the incoming beam into two different output arms (as described in Subsection 2.5.2). This beam splitter has two input arms: a *regular arm*, through which the standard incoming beam comes, and a *blocked arm*, where the incoming state is always assumed to be the zero-photon beam $|0,0\rangle$ (the vacuum state of two polarizations). If Eve can drill a small hole in Bob's device, exactly where the blocked arm gets its input from, then she can send a beam to the blocked arm and not only to the standard arm. It is proved [BGM14] that Eve can then cause the beam splitter to choose an output arm to her desire, instead of choosing a "random" arm. The state $|m_1, m_0\rangle_r|n_1, n_0\rangle_b$ consists of the Fock state $|m_1, m_0\rangle$ sent through the *regular arm* of the beam splitter and the Fock state $|n_1, n_0\rangle$ sent through the *blocked arm*. Alice sends an ideal qubit (a single photon through the regular arm), while Eve may send a single photon through any of the two arms or a superposition. Therefore, $\mathcal{H}^{\mathrm{A}} = \mathcal{H}_2 \triangleq \mathrm{Span}\{|0,1\rangle_r|0,0\rangle_b\ ,\ |1,0\rangle_r|0,0\rangle_b\}$ and $\mathcal{H}^{\mathrm{P}} = \mathrm{Span}\{|0,1\rangle_r|0,0\rangle_b\ ,\ |1,0\rangle_r|0,0\rangle_b\ ,\ |0,0\rangle_r|0,1\rangle_b\ ,\ |0,0\rangle_r|1,0\rangle_b\}$.

8.5 Quantum Side-Channel Attacks

Shamir's "Quantum Side-Channel Attack" on Polarization-Based QKD: The following attack was proposed by Adi Shamir in a meeting with Tal Mor (one of the authors of [LM20], on which this chapter is based) around 1996–1999 [Sha], and it may have never been published (but see similar attacks below). Shamir's attack only applies to QKD implementations that use "*active*" basis choice (as opposed to the "passive" basis choice, which leads to the "Fixed Apparatus" attack described in Example 3 of Section 8.4). The attack is related to Imperfection 2 described in Section 8.2: Bob's apparatus must be fully or partially ready to receive Alice's photon before it arrives. For example, if the photon is supposed to arrive at time t, then Bob's setup is already partially ready at time $t - \delta$; in particular, Bob decides the *basis choice* and configures the polarizing beam splitter accordingly before time $t - \delta$. The attack also assumes that the detectors themselves are still inactive (blocked) at time $t - \delta$, and are activated just before time t. Therefore, at time $t - \delta$, the polarizing beam splitter is already configured to match the required basis (the computational basis or the Hadamard basis), while the detectors are still blocked.

Eve's attack is sending a strong pulse at time $t - \delta$, that hits the polarizing beam splitter (but not the blocked detectors) and gets reflected back to Eve, containing full or partial information on the direction of the polarizing beam splitter—and, thus, on the basis choice. Assuming Eve gets the information on Bob's basis choice *before* she receives Alice's pulse, Eve could employ the following full attack: Eve measures the photon coming from Alice *in the same basis chosen by Bob*, learns the qubit's value, and resends to Bob the resulting state (in the same basis), obtaining full information while inducing no errors and no losses.

One can suggest two ways to possibly prevent the attack: (a) opening the detection window (activating the detectors) *shortly* after the polarizing beam splitter is configured according to the basis choice (if the time difference is sufficiently short, Eve cannot find Bob's basis choice on time for employing the full attack); or (b) blocking access to the polarizing beam splitter until the detectors are activated (although this solution may be hard to implement).

As we explain in Section 8.6, the "Bright Illumination" attack could have been predicted by adding Imperfection 1 described in Section 8.2 (namely, detection of multi-photon pulses) to the above idea of a strong pulse sent at time $t - \delta$ towards Bob (i.e., Imperfection 2, as already discussed here) and using the Fock space notations.

"Conventional Optical Eavesdropping" and "Quantum Side-Channel Attacks": Other attacks, similar to Shamir's attack, have been independently developed—for example, the "Large Pulse" attack [VMH01], which attacks both Alice's and Bob's set-ups. As written in [VMH01]: "This [large pulse] attack is one of the possible methods of conventional optical eavesdropping, a new strategy of eavesdropping on quantum cryptosystems, which eliminates the need of immediate interaction with transmitted quantum states. It allows the eavesdropper to avoid inducing transmission errors that disclose her presence to the legal users."

Instead of restricting ourselves to "conventional optical eavesdropping on quantum cryptosystems", we make use of a different sentence from [VMH01]—"eavesdropping on quantum cryptosystems which eliminates the need of immediate interaction with transmitted quantum states"—and we define "quantum side-channel attacks" as follows:

A *quantum side-channel attack* is any eavesdropping strategy which eliminates the need of any immediate interaction with the transmitted quantum states.

According to the above definition, both Shamir's attack and the "Large Pulse" attack are "quantum side-channel attacks", because they attack the devices and not the quantum states themselves. On the other hand, the "Reversed-Space" attacks and the "Quantum Space" attacks (see Section 8.4) can be fully described using a proper description of the QKD protocol, which uses the Fock space notations; therefore, they should *not* be considered as "quantum side-channel attacks". In fact, we can say they are *complementary* to "quantum side-channel attacks", and we name them "*state*-channel attacks".

In a classical communication world, the notion of "side-channel attacks" makes use of any information leaked by the *physical* execution of the algorithm (see, for example, [KB07]). Accordingly, other researchers (e.g., [SBPCDLP09]) have chosen to adopt a wide definition of "quantum side-channels", which also includes the "Photon-Number Splitting" attack and many other practical attacks. However, we prefer to take a narrower view of "quantum side-channel attacks", as explained above.

8.6 From Practice to Theory: The Possibility of Predicting the "Bright Illumination" Attack

The "Bright Illumination" attack could have been predicted, because it simply combines Imperfections 1 and 2 that were described in Section 8.2: namely, detecting many photons at time $t - \delta$, while the single "information" photon should have arrived at time t. In some sense, it seems to merge a "Reversed-Space" attack and a "quantum side-channel attack", because it attacks both the transmitted quantum states and the detectors themselves. However, because Bob's detectors are fully exposed to Eve at both times t and $t - \delta$ (unlike the "Large Pulse" attack [VMH01], where the detectors are not exposed at time $t - \delta$), we see the "Bright Illumination" attack as a special (and fascinating) case of "Reversed-Space" attack, and not as a "quantum side-channel attack".

The "Bright Illumination" attack is made possible by a *lack of information* on the "space of the protocol" $\mathcal{H}^{\mathrm{P}}$: Eve sends many photons (as in Imperfection 1) at time $t - \delta$

(as in Imperfection 2), and Bob does not notice her disruption because he cannot *count* the number of photons and cannot *block* the detectors at time $t - \delta$.

For preventing all the possible attacks and proving full security, it must be known how Bob's detectors treat *any* number m of photons sent to him by Eve, and it must also be known how Bob's detectors treat multiple pulses. In particular, a detector definitely cannot operate properly in the hypothetical scenario where an infinite number of photons (with infinite energy) arrives as its input. A potentially secure system must have an estimated threshold N, such that if $m \lesssim N$ photons arrive, they are correctly measured by the detectors (treated as one photon), and if $m \gtrsim N$ photons arrive, the measurement result is clearly invalid and is known to Bob (for example, smoke comes out of the detectors, or the detectors are burned). N is estimated, so there is a small unknown range near it.

Prior to the "Bright Illumination" attack, it seems that no systematic effort has been invested in finding or approximating the threshold N and characterizing the detectors' behavior on *all* possible inputs (any number of photons m). A proper "Reversed-Space" analysis would have suggested that experimentalists *must* check what N is and fully analyze the behavior of Bob's detectors on each quantum state; such an analysis would then have found the "space of the protocol" $\mathcal{H}^{\mathrm{P}}$ which is available for Eve's attack.

A careful "Reversed-Space" analysis—if it had been carried out—would then have found that instead of *one* estimated threshold N (with some small unknown range around it), there are *two* estimated thresholds N_1, N_2, such that $N_1 < N_2$, with a some small unknown range around each of them, and a *large* difference between them. Therefore, there are three main ranges of the numbers of photons m: (a) for $m \lesssim N_1$ photons, Bob's detectors work well (and click if at least one photon arrives); (b) for $N_1 \lesssim m \lesssim N_2$ photons, it would have become *known* that some strange phenomena happen—for example, that Bob's detectors switch to the "linear mode"; and (c) for $m \gtrsim N_2$ photons, Bob's detectors malfunction (e.g., the detectors are burned).

Thus, surprisingly, even if the experimentalist had not known about the two modes of operation ("Geiger mode" and the "linear mode") existing for each detector, he or she could still have discovered the two different thresholds N_1, N_2 and then investigated the detectors' behavior in the middle range $N_1 \lesssim m \lesssim N_2$. This would have allowed him or her to discover the "linear mode" and realize that there is also a need to check *multiple* pulses for finding the correct "space of the protocol" and for analyzing the security against "Reversed-Space" attacks. Namely, the "Reversed-Space" approach makes it possible to discover attacks even if the detectors are treated as *a black box* whose internal behavior is unknown. By theoretically trying to prove security against any theoretical "Reversed-Space" attack, it would have been possible to find the practical "Bright Illumination" attack; it would have even been possible to study the operation of a "*black-box*" detector and discover, for example, that it has a "linear mode" of operation (even if this mode of operation had not been already known for realistic detectors).

8.7 Conclusion

We have seen a rare example (in quantum information processing) where experiment preceded theory. We can see now that this experimental attack could have been theoretically predicted: for a system to be secure, Bob must be sure that Eve cannot attack by sending an unexpected number of photons, and he must know what happens to his detectors for any number of photons. Otherwise—Eve can attack; and we could have known that this may be possible.

We have also defined the general notion of "quantum side-channel attacks": we have distinguished "state-channel attacks" (including "Reversed-Space" and "Quantum Space" attacks) that interact with the transmitted (prepared or measured) quantum states, from "quantum side-channel attacks" that *do not interact* with the transmitted quantum states.

Chapter 9

Summary

In this research, we have answered several important questions about security of QKD:

1. In Chapters 3, 4, and 5, we have discussed practical security of semiquantum key distribution (SQKD) protocols. Unlike previous SQKD protocols, our newly suggested “Mirror protocol” is experimentally feasible, and we have proved it secure against “uniform collective” attacks.

   Notice that these chapters analyze security of a *two-way* protocol (see Subsection 2.2.2) which is harder to analyze than one-way protocols; thus, as explained in Section 5.1, its security analysis is limited to uniform collective attacks.

2. In Chapters 6 and 7, we have improved and generalized the security approach of [BBBMR06] to prove fully composable security of the BB84 protocol and many of its variants against the most general attacks.

3. In Chapter 8, we have shown how a practical attack (the “Bright Illumination” attack) can be theoretically modeled as a “Reversed-Space” attack.

All three directions share the motivation of *bridging the gap* between theory and practice, and all of them are aimed (in different ways) to answer one ultimate question: can we experimentally implement a QKD protocol with full and unconditional security against any possible attack (including all attacks that use practical imperfections)? This general question is one of the most important open problems in the field of QKD.

More specific open problems include: analyzing experimental implementations of the Mirror protocol; proving unconditional security of SQKD protocols against the most general attacks (and not only collective or “uniform collective” attacks) and against all multi-photon attacks; generalizing [BBBMR06]’s security approach to various QKD protocols that are not similar to BB84; and systematically mapping practical attacks to theoretical models.

Much work remains to be done on the general problem of obtaining full security proofs for realistic QKD systems, but we believe our research has improved the understanding of practical security in a variety of important sub-fields of QKD.

Bibliography

[BB84] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *International Conference on Computers, Systems & Signal Processing, IEEE, 1984*, pages 175–179, Dec 1984.

[BBBGM02] Eli Biham, Michel Boyer, Gilles Brassard, Jeroen van de Graaf, and Tal Mor. Security of quantum key distribution against all collective attacks. *Algorithmica*, 34(4):372–388, Nov 2002. `doi:10.1007/s00453-002-0973-6`.

[BBBMR06] Eli Biham, Michel Boyer, Oscar P. Boykin, Tal Mor, and Vwani Roychowdhury. A proof of the security of quantum key distribution. *Journal of Cryptology*, 19(4):381–439, Apr 2006. `doi:10.1007/s00145-005-0011-3`.

[BBBSS92] Charles H. Bennett, François Bessette, Gilles Brassard, Louis Salvail, and John Smolin. Experimental quantum cryptography. *Journal of Cryptology*, 5(1):3–28, Jan 1992. `doi:10.1007/BF00191318`.

[BBCJPW93] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70:1895–1899, Mar 1993. `doi:10.1103/PhysRevLett.70.1895`.

[BBCM95] Charles H. Bennett, Gilles Brassard, Claude Crépeau, and Ueli M. Maurer. Generalized privacy amplification. *IEEE Transactions on Information Theory*, 41(6):1915–1923, Nov 1995. `doi:10.1109/18.476316`.

[BBM92] Charles H. Bennett, Gilles Brassard, and N. David Mermin. Quantum cryptography without Bell's theorem. *Physical Review Letters*, 68:557–559, Feb 1992. `doi:10.1103/PhysRevLett.68.557`.

[Ben80] Paul Benioff. The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. *Journal of Statistical Physics*, 22(5):563–591, May 1980. doi:10.1007/BF01011339.

[BF02] Kim Boström and Timo Felbinger. Deterministic secure direct communication using entanglement. *Physical Review Letters*, 89:187902, Oct 2002. doi:10.1103/PhysRevLett.89.187902.

[BGKM09] Michel Boyer, Ran Gelles, Dan Kenigsberg, and Tal Mor. Semiquantum key distribution. *Physical Review A*, 79:032341, Mar 2009. doi:10.1103/PhysRevA.79.032341.

[BGM09] Michel Boyer, Ran Gelles, and Tal Mor. Security of the Bennett-Brassard quantum key distribution protocol against collective attacks. *Algorithms*, 2(2):790–807, Jun 2009. doi:10.3390/a2020790.

[BGM14] Michel Boyer, Ran Gelles, and Tal Mor. Attacks on fixed-apparatus quantum-key-distribution schemes. *Physical Review A*, 90:012329, Jul 2014. doi:10.1103/PhysRevA.90.012329.

[BHM96] Eli Biham, Bruno Huttner, and Tal Mor. Quantum cryptographic network based on quantum memories. *Physical Review A*, 54:2651–2658, Oct 1996. doi:10.1103/PhysRevA.54.2651.

[BHP93] Stephen M. Barnett, Bruno Huttner, and Simon J.D. Phoenix. Eavesdropping strategies and rejected-data protocols in quantum cryptography. *Journal of Modern Optics*, 40(12):2501–2513, 1993. doi:10.1080/09500349314552491.

[BKLM17] Michel Boyer, Matty Katz, Rotem Liss, and Tal Mor. Experimentally feasible protocol for semiquantum key distribution. *Physical Review A*, 96:062335, Dec 2017. doi:10.1103/PhysRevA.96.062335.

[BKM07] Michel Boyer, Dan Kenigsberg, and Tal Mor. Quantum key distribution with classical Bob. *Physical Review Letters*, 99:140501, Oct 2007. doi:10.1103/PhysRevLett.99.140501.

[BKM09] Michel Boyer, Dan Kenigsberg, and Tal Mor. Boyer, Kenigsberg, and Mor reply:. *Physical Review Letters*, 102:098902, Mar 2009. doi:10.1103/PhysRevLett.102.098902.

[BLM17] Michel Boyer, Rotem Liss, and Tal Mor. Security against collective attacks of a modified BB84 QKD protocol with information only in one basis. In *Proceedings of the 2nd International Conference on*

Complexity, Future Information Systems and Risk—COMPLEXIS, 24–26 April, 2017, Porto, Portugal, pages 23–29. INSTICC, Apr 2017. doi:10.5220/0006241000230029.

[BLM18] Michel Boyer, Rotem Liss, and Tal Mor. Attacks against a simplified experimentally feasible semiquantum key distribution protocol. *Entropy*, 20(7):536, Jul 2018. doi:10.3390/e20070536.

[BLM20] Michel Boyer, Rotem Liss, and Tal Mor. Composable security against collective attacks of a modified BB84 QKD protocol with information only in one basis. *Theoretical Computer Science*, 801:96–109, Jan 2020. doi:10.1016/j.tcs.2019.08.014.

[BLMR13] Normand J. Beaudry, Marco Lucamarini, Stefano Mancini, and Renato Renner. Security of two-way quantum key distribution. *Physical Review A*, 88:062302, Dec 2013. doi:10.1103/PhysRevA.88.062302.

[BLMS00] Gilles Brassard, Norbert Lütkenhaus, Tal Mor, and Barry C. Sanders. Limitations on practical quantum cryptography. *Physical Review Letters*, 85:1330–1333, Aug 2000. doi:10.1103/PhysRevLett.85.1330.

[BM97a] Eli Biham and Tal Mor. Bounds on information and the security of quantum cryptography. *Physical Review Letters*, 79:4034–4037, Nov 1997. doi:10.1103/PhysRevLett.79.4034.

[BM97b] Eli Biham and Tal Mor. Security of quantum cryptography against collective attacks. *Physical Review Letters*, 78:2256–2259, Mar 1997. doi:10.1103/PhysRevLett.78.2256.

[BM10] Michel Boyer and Tal Mor. On the robustness of (photonic) quantum key distribution with classical Alice. *arXiv preprint arXiv:1012.2418*, Dec 2010. URL: https://arxiv.org/abs/1012.2418.

[BM11] Michel Boyer and Tal Mor. Comment on "semiquantum-key distribution using less than four quantum states". *Physical Review A*, 83:046301, Apr 2011. doi:10.1103/PhysRevA.83.046301.

[BMS96] Charles H. Bennett, Tal Mor, and John A. Smolin. Parity bit in quantum cryptography. *Physical Review A*, 54:2675–2684, Oct 1996. doi:10.1103/PhysRevA.54.2675.

[BOHLMO05] Michael Ben-Or, Michał Horodecki, Debbie W. Leung, Dominic Mayers, and Jonathan Oppenheim. The universal composable security of quantum key distribution. In Joe Kilian, editor, *Theory of Cryptography: Second Theory of Cryptography Conference, TCC 2005, Cambridge, MA, USA, February 10-12, 2005. Proceedings*, pages 386–406, Berlin, Heidelberg, Feb 2005. Springer Berlin Heidelberg. doi:10.1007/978-3-540-30576-7_21.

[BP12] Samuel L. Braunstein and Stefano Pirandola. Side-channel-free quantum key distribution. *Physical Review Letters*, 108:130502, Mar 2012. doi:10.1103/PhysRevLett.108.130502.

[Can01] Ran Canetti. Universally composable security: a new paradigm for cryptographic protocols. In *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*, pages 136–145, Oct 2001. doi:10.1109/SFCS.2001.959888.

[CKR09] Matthias Christandl, Robert König, and Renato Renner. Post-selection technique for quantum channels with applications to quantum cryptography. *Physical Review Letters*, 102:020504, Jan 2009. doi:10.1103/PhysRevLett.102.020504.

[Deu85] David Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London A: Mathematical and Physical Sciences*, 400(1818):97–117, Jul 1985. doi:10.1098/rspa.1985.0070.

[DH76] Whitfield Diffie and Martin Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, Nov 1976. doi:10.1109/TIT.1976.1055638.

[DR13] Joan Daemen and Vincent Rijmen. *The design of Rijndael: AES—the advanced encryption standard.* Springer Science & Business Media, 2013.

[DW05] Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. *Proceedings of the Royal Society A*, 461(2053):207–235, Jan 2005. doi:10.1098/rspa.2004.1372.

[Eke91] Artur K. Ekert. Quantum cryptography based on Bell's theorem. *Physical Review Letters*, 67:661–663, Aug 1991. doi:10.1103/PhysRevLett.67.661.

[Fey82] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6):467–488, Jun 1982. doi:10.1007/BF02650179.

[FvdG99] Christopher A. Fuchs and Jeroen van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, May 1999. doi:10.1109/18.761271.

[Gel08] Ran Gelles. On the security of theoretical and realistic quantum key distribution schemes. Master's thesis, Technion—Israel Institute of Technology, Haifa, Sep 2008. URL: https://www.graduate.technion.ac.il/theses/Abstracts.asp?Id=24946.

[GLLP04] Daniel Gottesman, Hoi-Kwong Lo, Norbert Lütkenhaus, and John Preskill. Security of quantum key distribution with imperfect devices. *Quantum Information & Computation*, 4(5):325–360, Sep 2004. URL: http://www.rintonpress.com/journals/qiconline.html#v4n5.

[GM07] Ran Gelles and Tal Mor. Quantum-space attacks. *arXiv preprint arXiv:0711.3019*, Nov 2007. URL: https://arxiv.org/abs/0711.3019.

[GM12] Ran Gelles and Tal Mor. On the security of interferometric quantum key distribution. In Adrian-Horia Dediu, Carlos Martín-Vide, and Bianca Truthe, editors, *Theory and Practice of Natural Computing: First International Conference, TPNC 2012, Tarragona, Spain, October 2-4, 2012. Proceedings*, pages 133–146, Berlin, Heidelberg, Oct 2012. Springer Berlin Heidelberg. doi:10.1007/978-3-642-33860-1_12.

[GM16] Ran Gelles and Tal Mor. Reversed space attacks. *arXiv preprint arXiv:1110.6573*, May 2016. URL: https://arxiv.org/abs/1110.6573.

[GMD02] A. Galindo and M. A. Martín-Delgado. Information and computation: Classical and quantum aspects. *Reviews of Modern Physics*, 74:347–423, May 2002. doi:10.1103/RevModPhys.74.347.

[Gru99] Jozef Gruska. *Quantum computing*. McGraw-Hill London, 1999.

[Gur13] Pavel Gurevich. Experimental quantum key distribution with classical Alice. Master's thesis, Technion—Israel Institute of Technology, Haifa, May 2013. URL: https://www.graduate.technion.ac.il/Theses/Abstracts.asp?Id=26105.

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Associ-*

ation, 58(301):13–30, Mar 1963. doi:10.1080/01621459.1963.10500830.

[Ina02] Hitoshi Inamori. Security of practical time-reversed EPR quantum key distribution. *Algorithmica*, 34(4):340–365, Nov 2002. doi:10.1007/s00453-002-0983-4.

[KB07] Boris Köpf and David Basin. An information-theoretic model for adaptive side-channel attacks. In *Proceedings of the 14th ACM Conference on Computer and Communications Security*, CCS '07, pages 286–296, New York, NY, USA, Oct 2007. Association for Computing Machinery. doi:10.1145/1315245.1315282.

[KLM20] Walter O. Krawec, Rotem Liss, and Tal Mor. Security proof against collective attacks for an experimentally feasible semi-quantum key distribution protocol. *arXiv preprint arXiv:2012.02127*, Dec 2020. URL: https://arxiv.org/abs/2012.02127.

[Kra15a] Walter O. Krawec. Mediated semiquantum key distribution. *Physical Review A*, 91:032323, Mar 2015. doi:10.1103/PhysRevA.91.032323.

[Kra15b] Walter O. Krawec. Security proof of a semi-quantum key distribution protocol. In *2015 IEEE International Symposium on Information Theory (ISIT)*, pages 686–690. IEEE, Jun 2015. doi:10.1109/ISIT.2015.7282542.

[Kra16] Walter O. Krawec. Security of a semi-quantum protocol where reflections contribute to the secret key. *Quantum Information Processing*, 15(5):2067–2090, Feb 2016. doi:10.1007/s11128-016-1266-3.

[Kra17] Walter O. Krawec. Quantum key distribution with mismatched measurements over arbitrary channels. *Quantum Information & Computation*, 17(3&4):209–241, Mar 2017. doi:10.26421/QIC17.3-4-2.

[Kra18] Walter O. Krawec. Practical security of semi-quantum key distribution. In Eric Donkor, editor, *Proceedings of SPIE, Quantum Information Science, Sensing, and Computation X*, volume 10660, page 1066009, May 2018. doi:10.1117/12.2303759.

[KRBM07] Robert König, Renato Renner, Andor Bariska, and Ueli Maurer. Small accessible quantum information does not imply security. *Physical Review Letters*, 98:140502, Apr 2007. doi:10.1103/PhysRevLett.98.140502.

[KZH$^+$02] C. Kurtsiefer, P. Zarda, M. Halder, H. Weinfurter, P. M. Gorman, P. R. Tapster, and J. G. Rarity. A step towards global key distribution. *Nature*, 419(6906):450–450, Oct 2002. `doi:10.1038/419450a`.

[LC99] Hoi-Kwong Lo and H. F. Chau. Unconditional security of quantum key distribution over arbitrarily long distances. *Science*, 283(5410):2050–2056, Mar 1999. `doi:10.1126/science.283.5410.2050`.

[LC08] Hua Lu and Qing-Yu Cai. Quantum key distribution with classical Alice. *International Journal of Quantum Information*, 06(06):1195–1202, Dec 2008. `doi:10.1142/S0219749908004353`.

[LCA05] Hoi-Kwong Lo, H.F. Chau, and M. Ardehali. Efficient quantum key distribution scheme and a proof of its unconditional security. *Journal of Cryptology*, 18(2):133–165, Apr 2005. `doi:10.1007/s00145-004-0142-y`.

[LCQ12] Hoi-Kwong Lo, Marcos Curty, and Bing Qi. Measurement-device-independent quantum key distribution. *Physical Review Letters*, 108:130503, Mar 2012. `doi:10.1103/PhysRevLett.108.130503`.

[LCT14] Hoi-Kwong Lo, Marcos Curty, and Kiyoshi Tamaki. Secure quantum key distribution. *Nature Photonics*, 8(8):595–604, Jul 2014. `doi:10.1038/nphoton.2014.149`.

[Lis17] Rotem Liss. Entanglement and geometrical distances in quantum information and quantum cryptography. Master's thesis, Technion—Israel Institute of Technology, Haifa, May 2017. URL: `https://www.graduate.technion.ac.il/Theses/Abstracts.asp?Id=30246`.

[LM05] Marco Lucamarini and Stefano Mancini. Secure deterministic communication without entanglement. *Physical Review Letters*, 94:140501, Apr 2005. `doi:10.1103/PhysRevLett.94.140501`.

[LM20] Rotem Liss and Tal Mor. From practice to theory: The "Bright Illumination" attack on quantum key distribution systems. In Carlos Martín-Vide, Miguel A. Vega-Rodríguez, and Miin-Shen Yang, editors, *Theory and Practice of Natural Computing*, pages 82–94, Cham, Dec 2020. Springer International Publishing. `doi:10.1007/978-3-030-63000-3_7`.

[LWWESM10] Lars Lydersen, Carlos Wiechers, Christoffer Wittmann, Dominique Elser, Johannes Skaar, and Vadim Makarov. Hacking commercial quantum cryptography systems by tailored bright illumination. *Nature Photonics*, 4(10):686–689, Aug 2010. doi:10.1038/nphoton.2010.214.

[MAP11] Lluís Masanes, Antonio Acín, and Stefano Pironio. Secure device-independent quantum key distribution with causally independent measurement devices. *Nature Communications*, 2(238):1–7, Mar 2011. doi:10.1038/ncomms1244.

[MAS06] Vadim Makarov, Andrey Anisimov, and Johannes Skaar. Effects of detector efficiency mismatch on security of quantum cryptosystems. *Physical Review A*, 74:022313, Aug 2006. doi:10.1103/PhysRevA.74.022313.

[May01] Dominic Mayers. Unconditional security in quantum cryptography. *Journal of the ACM*, 48(3):351–406, May 2001. doi:10.1145/382780.382781.

[MH05] Vadim Makarov and Dag R. Hjelme. Faked states attack on quantum cryptosystems. *Journal of Modern Optics*, 52(5):691–705, Mar 2005. doi:10.1080/09500340410001730986.

[MHHTZG97] A. Muller, T. Herzog, B. Huttner, W. Tittel, H. Zbinden, and N. Gisin. "plug and play" systems for quantum cryptography. *Applied Physics Letters*, 70(7):793–795, Feb 1997. doi:10.1063/1.118224.

[Mor98] Tal Mor. No cloning of orthogonal states in composite systems. *Physical Review Letters*, 80:3137–3140, Apr 1998. doi:10.1103/PhysRevLett.80.3137.

[MY98] Dominic Mayers and Andrew Yao. Quantum cryptography with imperfect apparatus. In *Proceedings 39th Annual Symposium on Foundations of Computer Science*, pages 503–509, Nov 1998. doi:10.1109/SFCS.1998.743501.

[NC00] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information.* Cambridge university press (10th anniversary edition, 2010), 2000.

[PAB$^+$20] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. Shamsul Shaari, M. Tomamichel,

V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden. Advances in quantum cryptography. *Advances in Optics and Photonics*, 12(4):1012–1236, Dec 2020. doi:10.1364/AOP.361502.

[PHB+14] W. Pfaff, B. J. Hensen, H. Bernien, S. B. van Dam, M. S. Blok, T. H. Taminiau, M. J. Tiggelman, R. N. Schouten, M. Markham, D. J. Twitchen, and R. Hanson. Unconditional quantum teleportation between distant solid-state quantum bits. *Science*, 345(6196):532–535, Aug 2014. doi:10.1126/science.1253512.

[Pre18] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, Aug 2018. doi:10.22331/q-2018-08-06-79.

[PW00] Birgit Pfitzmann and Michael Waidner. Composition and integrity preservation of secure reactive systems. In *Proceedings of the 7th ACM Conference on Computer and Communications Security*, CCS '00, pages 245–254, New York, NY, USA, Nov 2000. ACM. doi:10.1145/352600.352639.

[Ren07] Renato Renner. Symmetry of large physical systems implies independence of subsystems. *Nature Physics*, 3(9):645–649, Jul 2007. doi:10.1038/nphys684.

[Ren08] Renato Renner. Security of quantum key distribution. *International Journal of Quantum Information*, 6(01):1–127, Feb 2008. doi:10.1142/S0219749908003256.

[RGK05] Renato Renner, Nicolas Gisin, and Barbara Kraus. Information-theoretic security proof for quantum-key-distribution protocols. *Physical Review A*, 72:012332, Jul 2005. doi:10.1103/PhysRevA.72.012332.

[RP00] Eleanor Rieffel and Wolfgang Polak. An introduction to quantum computing for non-physicists. *ACM Computing Surveys*, 32(3):300–335, Sep 2000. doi:10.1145/367701.367709.

[RSA78] Ronald L. Rivest, Adi Shamir, and Leonard Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, Feb 1978. doi:10.1145/359340.359342.

[SBPCDLP09] Valerio Scarani, Helle Bechmann-Pasquinucci, Nicolas J. Cerf, Miloslav Dušek, Norbert Lütkenhaus, and Momtchil Peev. The security of practical quantum key distribution. *Reviews of Modern Physics*, 81:1301–1350, Sep 2009. doi:10.1103/RevModPhys.81.1301.

[SDL13] Zhi-Wei Sun, Rui-Gang Du, and Dong-Yang Long. Quantum key distribution with limited classical Bob. *International Journal of Quantum Information*, 11(01):1350005, Apr 2013. doi:10.1142/S0219749913500056.

[Sha] Adi Shamir. personal communication.

[Sha49] Claude E. Shannon. Communication theory of secrecy systems. *The Bell System Technical Journal*, 28(4):656–715, Oct 1949. doi:10.1002/j.1538-7305.1949.tb00928.x.

[Sho94] Peter W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, Nov 1994. doi:10.1109/SFCS.1994.365700.

[Sho99] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Review*, 41(2):303–332, Apr 1999. doi:10.1137/S0036144598347011.

[SK14] Valerio Scarani and Christian Kurtsiefer. The black paper of quantum cryptography: Real implementation problems. *Theoretical Computer Science*, 560:27–32, Dec 2014. Theoretical Aspects of Quantum Cryptography—celebrating 30 years of BB84. doi:10.1016/j.tcs.2014.09.015.

[SML10] Douglas Stebila, Michele Mosca, and Norbert Lütkenhaus. The case for quantum key distribution. In Alexander Sergienko, Saverio Pascazio, and Paolo Villoresi, editors, *Quantum Communication and Quantum Networking: First International Conference, QuantumComm 2009, Naples, Italy, October 26-30, 2009, Revised Selected Papers*, pages 283–296, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. doi:10.1007/978-3-642-11731-2_35.

[SP00] Peter W. Shor and John Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Physical Review Letters*, 85:441–444, Jul 2000. doi:10.1103/PhysRevLett.85.441.

[SR08] Valerio Scarani and Renato Renner. Quantum cryptography with finite resources: Unconditional security bound for discrete-variable protocols with one-way postprocessing. *Physical Review Letters*, 100:200501, May 2008. doi:10.1103/PhysRevLett.100.200501.

[Tam14] Natan Tamari. Experimental semiquantum key distribution: Classical Alice with mirror. Master's thesis, Technion—Israel Institute of Technology, Haifa, Nov 2014. URL: `https://www.graduate.technion.ac.il/Theses/Abstracts.asp?Id=28660`.

[TLC09] Yong-gang Tan, Hua Lu, and Qing-yu Cai. Comment on "quantum key distribution with classical Bob". *Physical Review Letters*, 102:098901, Mar 2009. `doi:10.1103/PhysRevLett.102.098901`.

[TLGR12] Marco Tomamichel, Charles Ci Wen Lim, Nicolas Gisin, and Renato Renner. Tight finite-key analysis for quantum cryptography. *Nature Communications*, 3(634):1–6, Jan 2012. `doi:10.1038/ncomms1631`.

[VMH01] Artem Vakhitov, Vadim Makarov, and Dag R. Hjelme. Large pulse attack as a method of conventional optical eavesdropping in quantum cryptography. *Journal of Modern Optics*, 48(13):2023–2038, 2001. `doi:10.1080/09500340108240904`.

[VV14] Umesh Vazirani and Thomas Vidick. Fully device-independent quantum key distribution. *Physical Review Letters*, 113:140501, Sep 2014. `doi:10.1103/PhysRevLett.113.140501`.

[WMU08] Shun Watanabe, Ryutaroh Matsumoto, and Tomohiko Uyematsu. Tomography increases key rates of quantum-key-distribution protocols. *Physical Review A*, 78:042316, Oct 2008. `doi:10.1103/PhysRevA.78.042316`.

[XMZLP20] Feihu Xu, Xiongfeng Ma, Qiang Zhang, Hoi-Kwong Lo, and Jian-Wei Pan. Secure quantum key distribution with realistic devices. *Reviews of Modern Physics*, 92:025002, May 2020. `doi:10.1103/RevModPhys.92.025002`.

[YYLH14] Kun-Fei Yu, Chun-Wei Yang, Ci-Hong Liao, and Tzonelih Hwang. Authenticated semi-quantum key distribution protocol using Bell states. *Quantum Information Processing*, 13(6):1457–1465, Mar 2014. `doi:10.1007/s11128-014-0740-z`.

[ZQLWL09] Xiangfu Zou, Daowen Qiu, Lvzhou Li, Lihua Wu, and Lvjun Li. Semiquantum-key distribution using less than four quantum states. *Physical Review A*, 79:052312, May 2009. `doi:10.1103/PhysRevA.79.052312`.

[ZQM18] Wei Zhang, Daowen Qiu, and Paulo Mateus. Security of a single-state semi-quantum key distribution protocol. *Quantum Information Processing*, 17(6):135, Apr 2018. `doi:10.1007/s11128-018-1904-z`.

[ZQZM15] Xiangfu Zou, Daowen Qiu, Shengyu Zhang, and Paulo Mateus. Semiquantum key distribution without invoking the classical party's measurement capability. *Quantum Information Processing*, 14(8):2981–2996, Aug 2015. `doi:10.1007/s11128-015-1015-z`.

או uniform collective attacks). התוצאה היא פרוטוקול SQKD חדש, יעיל ובטוח למימוש פרקטי כנגד מגוון רחב של התקפות.

שנית, חקרנו את הבטיחות ה"ניתנת להרכבה" של פרוטוקול ה־QKD הראשון והחשוב ביותר, הקרוי BB84 (על שם בנט וברסר שהמציאו אותו ופרסמו אותו בשנת 1984). לפרוטוקול BB84, כמו לפרוטוקולים רבים נוספים, קיימות הוכחות בטיחות נגד יריבים חזקים מאוד המשתמשים בהתקפות הכלליות ביותר האפשריות על מימוש תאורטי (אידאלי) של הפרוטוקול; יחד עם זאת, חלק מההוכחות אינן מראות בטיחות "ניתנת להרכבה" – כלומר, הן אינן מוכיחות שהמפתח נשאר סודי גם כאשר אליס ובוב משתמשים בו בפועל כחלק מפרוטוקול קריפטוגרפי (למשל, לצורך הצפנה). בפרק 6 דנו בפרוטוקול מעט שונה, הקרוי "BB84-INFO-z", והוכחנו את בטיחותו ה"ניתנת להרכבה" כנגד מחלקת ה"התקפות הקיבוציות" שהזכרנו קודם. בפרק 7 הרחבנו גישת בטיחות אלגברית מסוימת שבעבר השתמשו בה כדי להוכיח בטיחות (שאינה "ניתנת להרכבה") עבור BB84, הראינו שהיא עובדת עבור מגוון פרוטוקולים המבוססים על BB84 (בין היתר: BB84 עצמו, BB84-INFO-z, וכן וריאנטים של BB84 המאפשרים מימוש יעיל יותר), ושינינו אותה כדי שהיא תוכיח בטיחות "ניתנת להרכבה" כנגד כל ההתקפות האפשריות: כלומר, שיפרנו את גישת הבטיחות הזו והראינו שהיא יכולה להשיג תוצאות טובות הדומות לתוצאות שהושגו בגישות אחרות ומורכבות יותר.

לבסוף, בפרק 8 חקרנו התקפה פרקטית חשובה הנקראת "Bright Illumination" על מערכות QKD, והראינו שניתן למדל אותה באמצעות מודל תאורטי של התקפות ("Reversed-Space Attacks"). תוצאה זו מראה שבאופן עקרוני, ניתן לחזות התקפות פרקטיות מסוג זה בעזרת ניתוח תאורטי.

כל התוצאות שמצאנו מיועדות לגשר על הפער הקיים בין תאוריה לניסוי בתחום ה־QKD, והן עשויות לעזור בפתרון אחת הבעיות הפתוחות החשובות בתחום זה: כיצד לבנות מימוש פרקטי ויעיל של QKD בעולם האמיתי, שניתן להוכיח (ללא סייגים) שהוא בטוח לחלוטין כנגד כל ההתקפות האפשריות.

תקציר

חוקי הפיסיקה הקוונטית מאפשרים ליצור מצבים פיסיקליים מנוגדים לאינטואיציה שניתן היה לחשוב שהם בלתי אפשריים: למשל, חלקיק עשוי להיות בסופרפוזיציה – כגון סכום או הפרש – של כמה מיקומים שונים, כמה זמנים שונים, או כמה מצבים שונים. תחום המחקר הנקרא עיבוד אינפורמציה קוונטית חוקר דרכים לנצל את החוקים האלה לצורך ייצוג אינפורמציה ועיבודה, ולכן הוא מאפשר לנו לפתור בעיות ולבצע משימות שאינן אפשריות (או שהן קשות) למחשבים ולמכשירי תקשורת קלאסיים – כלומר, סטנדרטיים ולא־קוונטיים.

אחד ההישגים הראשונים שהושגו בתחום האינפורמציה הקוונטית הוא פיתוח שיטת הפצת המפתחות הקוונטית (Quantum Key Distribution, או בקיצור QKD). פרוטוקולי QKD נועדו לאפשר לשני משתמשים (המכונים בדרך כלל ״אליס״ ו״בוב״) ליצור מפתח סודי לחלוטין, אקראי ומשותף לשניהם. סודיות מלאה כזו היא בלתי אפשרית בעולם קלאסי (לא־קוונטי), שבו אין שום דרך למנוע מהיריבה ״איב״ להעתיק את כל המידע המשודר בין אליס לבוב; לעומת זאת, בעולם הקוונטי הסודיות נשמרת אפילו אם לאיב יש כוח חישוב בלתי מוגבל, ואפילו אם היא יכולה לבצע כל פעולה שאינה מנוגדת לחוקי הפיסיקה. אליס ובוב משתמשים בערוץ קוונטי לא־בטוח ובערוץ קלאסי מאומת: איב יכולה ליירט ולשנות כרצונה את כל המצבים הקוונטיים הנשלחים בערוץ הקוונטי, אבל מותר לה רק להאזין לכל המידע הקלאסי הנשלח בערוץ הקלאסי (היא אינה יכולה לשנותו).

למרבה הצער, הבטחת ה״בטיחות המושלמת״ של QKD נכונה רק בתאוריה: קיים מגוון רחב של בעיות בטיחות במימושי QKD בעולם האמיתי, כי הם אינם מממשים במדויק את הפרוטוקולים התאורטיים אלא משתמשים ברכיבים קוונטיים אמיתיים. (למשל, פרוטוקולי QKD תאורטיים מניחים בדרך כלל שאליס שולחת לבוב חלקיק אור (פוטון) בודד, אבל במציאות היא שולחת לפעמים שני פוטונים או יותר.) לכן חקרנו את בטיחותם של מגוון פרוטוקולי QKD בכמה סביבות פרקטיות שונות:

ראשית, בדקנו את הבטיחות של פרוטוקולי הפצת מפתחות קוונטית־למחצה (Semiquantum Key Distribution, או בקיצור SQKD), שבהם אליס או בוב הם ״קלאסיים״ – כלומר, אינם יכולים לבצע פעולות קוונטיות. בפרק 3 ניתחנו בעיות בטיחות פרקטיות בפרוטוקולי SQKD קיימים, הצענו פרוטוקול SQKD חדש ומעט מורכב יותר (״פרוטוקול Mirror״, המבוסס על שימוש במראה מתכווננת) שפותר את בעיות הבטיחות הנ״ל וניתן למימוש פרקטי מאובטח, והראינו שהפרוטוקול החדש מקיים דרישות בטיחות בסיסיות (complete robustness). בפרק 4 הוכחנו שאם מפשטים מעט את פרוטוקול Mirror, מקבלים פרוטוקול שאינו בטוח, והסקנו שמורכבותו של פרוטוקול Mirror המקורי היא כנראה חיונית. לבסוף, בפרק 5 הוכחנו את הבטיחות של פרוטוקול Mirror כנגד מחלקה נרחבת של התקפות (הקרויות ״התקפות קיבוציות״, או collective attacks; ליתר דיוק, הוכחנו בטיחות כנגד תת־קבוצה חשובה של התקפות אלה, שנקראות ״התקפות קיבוציות אחידות״

המחקר נעשה בהנחיית פרופ׳ חבר טל מור, בפקולטה למדעי המחשב.

רוב התוצאות בחיבור זה פורסמו כמאמרים מאת המחבר ושותפיו למחקר בכתבי־עת ובכנסים:

1. Michel Boyer, Matty Katz, Rotem Liss, and Tal Mor. Experimentally feasible protocol for semiquantum key distribution. *Physical Review A*, 96:062335, Dec 2017. `doi:10.1103/PhysRevA.96.062335`. (Chapter 3)

2. Michel Boyer, Rotem Liss, and Tal Mor. Attacks against a simplified experimentally feasible semiquantum key distribution protocol. *Entropy*, 20(7):536, Jul 2018. `doi:10.3390/e20070536`. (Chapter 4)

3. Walter O. Krawec, Rotem Liss, and Tal Mor. Security proof against collective attacks for an experimentally feasible semi-quantum key distribution protocol. *arXiv preprint arXiv:2012.02127*, Dec 2020. URL: `https://arxiv.org/abs/2012.02127`. (Chapter 5)

4. Michel Boyer, Rotem Liss, and Tal Mor. Composable security against collective attacks of a modified BB84 QKD protocol with information only in one basis. *Theoretical Computer Science*, 801:96–109, Jan 2020. `doi:10.1016/j.tcs.2019.08.014`. (Chapter 6)

5. Rotem Liss and Tal Mor. From practice to theory: The “Bright Illumination” attack on quantum key distribution systems. In Carlos Martín-Vide, Miguel A. Vega-Rodríguez, and Miin-Shen Yang, editors, *Theory and Practice of Natural Computing*, pages 82–94, Cham, Dec 2020. Springer International Publishing. `doi:10.1007/978-3-030-63000-3_7`. (Chapter 8)

תודות

ברצוני להודות למנחה שלי, פרופ׳ חבר טל מור, על הדרכתו המועילה ועל הדיונים, הרעיונות והעצות במהלך המחקר, ובמיוחד על העזרה והתמיכה שנתן לי לכל אורך לימודיי לתארים מתקדמים. ברצוני להודות גם לפרופ׳ חבר מישל בואייה ולפרופ׳ משנה וולטר קראווק על שיתוף פעולה פורה במחקר ועל דיונים מחקריים רבים שהובילו למגוון תוצאות ופרסומים משותפים.

ברצוני להודות גם לז׳יל ברסר, לרומן אורוס, לרנאטו רנר, לרותם ארנון־פרידמן, לאנדריאס וינטר, לג׳ון סמולין, לצ׳רלס בנט, לדיוויד דיווינצנזו, לצ׳יצ׳ה גסטיאני, ללואי סלביי, לאלי ביהם, ליוסי וינשטיין, לרומן שפירא, ליאיר רזק, ולאיתי פיירוורקר.

תודה מיוחדת מגיעה למשפחתי.

אני מודה לקרן דניאל, לקרן ג׳ייקובס ולטכניון על התמיכה הכספית הנדיבה בהשתלמותי.

הבטיחות של פרוטוקולי
הפצת מפתחות קוונטית

חיבור על מחקר

לשם מילוי חלקי של הדרישות לקבלת התואר
דוקטור לפילוסופיה

רותם ליס

הוגש לסנט הטכניון – מכון טכנולוגי לישראל
סיוון התשפ"א חיפה מאי 2021

הבטיחות של פרוטוקולי
הפצת מפתחות קוונטית

רותם ליס