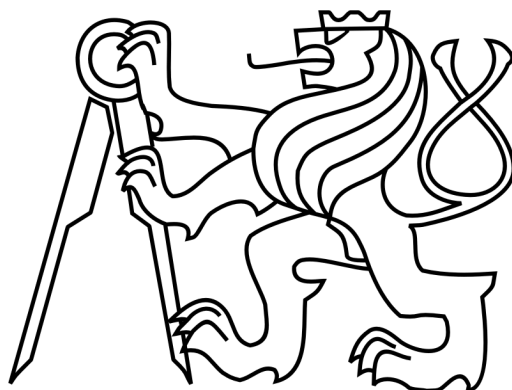


ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE  
FAKULTA JADERNÁ A FYZIKÁLNĚ INŽENÝRSKÁ  
KATEDRA LASEROVÉ FYZIKY A FOTONIKY

Program: Fyzikální elektronika



# Kvantová kryptografie s využitím časového a fázového kódování

Quantum cryptography with time and  
phase encoding

DIPLOMOVÁ PRÁCE

Vypracoval: Bc. Ondřej Čermák  
Vedoucí práce: prof. Ing. Ivan Richter, Dr.  
Konzultant: doc. Mgr. Jan Soubusta, Ph.D.  
Rok: 2023/2024

## I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Čermák** Jméno: **Ondřej** Osobní číslo: **486365**  
Fakulta/ústav: **Fakulta jaderná a fyzikálně inženýrská**  
Zadávající katedra/ústav: **Katedra fyzikální elektroniky**  
Studijní program: **Fyzikální elektronika**  
Specializace: **Fotonika**

## II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

**Kvantová kryptografie s využitím časového a fázového kódování**

Název diplomové práce anglicky:

**Quantum cryptography with time and phase encoding**

Pokyny pro vypracování:

- 1) Na základě rešerše a studia fyzikálních principů kvantové kryptografie se seznamte se základními procesy, metodami a klíčovými komponentami kvantového kryptografického systému, zmapujte a zpracujte praktická a technická hlediska možných realizací, podle jednotlivých přístupů.
- 2) Pozornost dále věnujte studiu metod polarizačního a zejména časového / fázového kódování. V rámci této metody časového / fázového kódování navrhnete vhodný kvantově kryptografický systém, jako synergetickou kombinaci vybraných funkčních komponent.
- 3) Jednotlivé dílčí funkční komponenty tohoto návrhu se pokuste experimentálně realizovat a otestovat, na základě spolupráce obou pracovišť KFE FJFI a SLO UPOL, v korelaci s jejich technickými možnostmi.
- 4) Pro ovládání experimentu (spínání laseru a modulátoru, jednofotonová detekce, QKD protokol) využijte možnosti mikropočítače (např. z rodiny Red Pithaya) v kombinaci s dalšími dostupnými přístroji.
- 5) Z jednotlivých funkčních komponent se následně pokuste sestavit navržený kryptografický systém. Získané výsledky porovnejte s literaturou a diskutujte.

Seznam doporučené literatury:

1. A. Kumar, S. Garhwal, State-of-the-Art Survey of Quantum Cryptography, Archives of Computational Methods in Engineering 28, 3831–3868 (2021).
2. F. Grasselli, Quantum Cryptography - From Key Distribution to Conference Key Agreement, Quantum Science and Technology, Springer, 2021.
3. 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. S. Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, Advances in quantum cryptography, Advances in Optics and Photonics 12, 1012 (2020).
4. A. Shenoy-Hejamadi, A. Pathak, S. Radhakrishna, Quantum Cryptography: Key Distribution and Beyond, Quanta 6, 1 (2017).
5. C. Kollmitzer, M. Pivk (Eds.), Applied Quantum Cryptography, Lect. Notes Phys. 797, Springer, 2010.
6. R. Wolf, Quantum Key Distribution, An Introduction With Exercises, Lect. Notes Phys. 988, Springer, 2021.
7. G. S. Agarwal, Quantum Optics, Cambridge University Press, Cambridge, 2013
8. H. A. Bachor, T. C. Ralph, A Guide to Experiments in Quantum Optics, Wiley-VCH, New York, 2019.
9. Z. J. Ou, Quantum Optics For Experimentalists, World Scientific Publishing, 2017.
10. J. C. Garrison, R. Y. Chiao, Quantum Optics, Oxford University Press, 2008.

Jméno a pracoviště vedoucí(ho) diplomové práce:

**prof. Dr. Ing. Ivan Richter katedra fyzikální elektroniky FJFI**

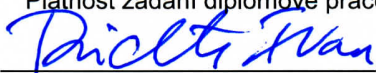
Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) diplomové práce:

**doc.Mgr. Jan Soubusta, Ph.D. Univerzita Palackého Olomouc - Přírodovědecká fak.**

Datum zadání diplomové práce: **12.10.2023**

Termín odevzdání diplomové práce: **10.05.2024**

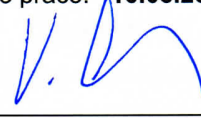
Platnost zadání diplomové práce: **12.10.2025**



prof. Dr. Ing. Ivan Richter  
podpis vedoucí(ho) práce



prof. Dr. Ing. Ivan Richter  
podpis vedoucí(ho) ústavu/katedry



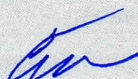
doc. Ing. Václav Čuba, Ph.D.  
podpis děkana(ky)

### III. PŘEVZETÍ ZADÁNÍ

Diplomant bere na vědomí, že je povinen vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v diplomové práci.

29. 11. 2023

Datum převzetí zadání



Podpis studenta

## Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne 10.5.2024

  
.....

Bc. Ondřej Čermák

## **Poděkování**

Děkuji panu prof. Ing. Ivani Richterovi, Dr. za odborné rady a čas, který mi věnoval. Poté bych chtěl velmi poděkovat svým konzultantům z Univerzity Palackého v Olomouci, především doc. Mgr. Janu Soubustovi, Ph.D., Mgr. Antonínu Černochovi Ph.D. a dále doc. Mgr. Karlu Lemrovi, Ph.D., za ochotu a velkou pomoc při sestavování experimentu v laboratoři na jejich univerzitě. Také děkuji podpoře v rámci SGS projektu SGS22/185/OHK4/3T/14.

Bc. Ondřej Čermák

*Název práce:*

## **Kvantová kryptografie s využitím časového a fázového kódování**

*Autor:* Bc. Ondřej Čermák

*Program:* Fyzikální elektronika

*Druh práce:* Diplomová práce

*Vedoucí práce:* prof. Ing. Ivan Richter, Dr.

*Konzultanti:* doc. Mgr. Jan Soubusta, Ph.D.

*Abstrakt:* Limitujícím faktorem dnešního šifrování je neschopnost bezchybně detekovat odposlouchávání, a proto všechny šifrovací algoritmy s tím musí počítat. Pro distribuci šifrovacího klíče se dnes používá asymetrické šifrování založené na problému diskretního logaritmu nebo faktorizaci prvočísel. Nové technologie, jako je nově se objevující oblast kvantových počítačů, však nabízejí algoritmy, které mohou efektivně vyřešit tyto úlohy, a tedy případně prolomit dnešní šifrování. Navíc by to mohlo ovlivnit všechna šifrovaná data, tedy nejen po vynálezu kvantového počítače schopného spouštět Shorův algoritmus, ale také všechna dříve uložená šifrovaná data. Kvantová distribuce klíčů umožňuje detekci odposlechu za pomoci zákonů kvantové mechaniky a umožňuje tak zcela bezpečně distribuovat symetrický klíč mezi dvě strany.

Tato práce prezentuje simulaci, experimentální realizaci a testování systému kvantové distribuce klíče (QKD) založeného na protokolu BB84 s využitím časového a fázového kódování. Hlavním cílem bylo sestavit praktické optické zařízení schopné demonstrovat principy QKD na úrovni jednotlivých fotonů.

Bylo vyvinuto plně integrované zařízení, které se skládá ze dvou asymetrických Mach-Zehnderových interferometrů rozdělených do sekcí Alice a Bob, propojených optickým vláknem, kvantovým kanálem. Mikropočítač Red Pitaya STEMLab 125-14 byl využit pro ovládání systému, generování laserových pulsů, aplikaci fázové modulace a provádění detekce fotonů. Na jednotlivých komponentách bylo provedeno testování a kalibrace pro optimalizaci systému. Klasická interference dosáhla viditelnosti 45,66%, což se blíží teoretickému maximu 50%. Kvantové měření na úrovni jednotlivých fotonů dosáhlo viditelnosti 32,60%, což demonstruje úspěšnou interferenci a schopnost kódovat informace na jednotlivé fotony.

Simulace pokrývá celý proces QKD, od generování kvantových stavů Alice, přenosu přes simulovaný kvantový kanál s šumem a ztrátami, měření Bobem a následnými postprocesingovými fázemi, včetně filtrování klíčů, odhadu chyb, korekce chyb, zesílení soukromí klíče a autentizace klíče. Provedeny byly rozsáhlé analýzy pro hodnocení dopadu různých zdrojů šumu a scénářů odposlechu na kvantovou chybovost bitů (QBER) a finální délku klíče.

*Klíčová slova:* Kvantová distribuce klíče, kvantová informace, kvantová kryptografie, kvantová komunikace, protokol BB84

*Thesis title:*

## **Quantum cryptography with time and phase encoding**

*Author:* Bc. Ondřej Čermák

*Branch of study:* Physical Electronics, spec. Photonics

*Kind of thesis:* Master thesis

*Supervisor:* prof. Ing. Ivan Richter, Dr.

*Consultants:* doc. Mgr. Jan Soubusta, Ph.D.

*Abstract:* The limiting factor for today's encryption is the inability to flawlessly detect eavesdropping and thus all encryption algorithms have to take this into account. To distribute a cryptographic key, asymmetric encryption based on discrete logarithm problem or prime factorization is being used. However, novel technologies, like the emerging field of quantum computing, offer algorithms, which can efficiently solve these tasks and hence possibly break today's encryption. Moreover, this could possibly affect all encrypted data, not just after the invention of the quantum computer capable of running Shor algorithm, but also all eavesdropped encrypted data from before are vulnerable. Quantum key distribution allows detection of eavesdropping by relying on the laws of quantum mechanics, which ensures unconditionally secure distribution of a symmetric key between two parties.

This thesis presents a simulation, an experimental implementation and testing of a quantum key distribution (QKD) system based on the BB84 protocol using time-phase encoding. The primary objective was to construct a practical optical setup capable of demonstrating the principles of QKD at the single-photon level.

A fully integrated setup was developed, consisting of two asymmetric Mach-Zehnder interferometers divided into Alice and Bob sections connected via an optical fiber quantum channel. The Red Pitaya STEMLab 125-14 microcomputer was utilized for system control, generating trigger pulses, applying phase modulation, and performing photon detection. Extensive testing and calibration were conducted on the individual components to optimize the system performance. Classical interference measurements achieved a visibility of 45.66%, closely approaching the theoretical maximum of 50%. Quantum measurements at the single-photon level realized a visibility of 32.60%, demonstrating successful interference and the ability to encode information on single photons.

The simulation covers the entire QKD process, from the generation of quantum states by Alice, transmission through a simulated quantum channel with noise and losses, measurement by Bob, and subsequent post-processing stages, including key sifting, error estimation, error correction, privacy amplification, and key authentication. Extensive analyses were conducted to evaluate the impact of different noise sources and eavesdropping scenarios on the quantum bit error rate (QBER) and the final key length.

*Key words:* Quantum key distribution, quantum information, quantum cryptography, quantum communication, BB84 protocol

# Contents

<b>1</b>	<b>Introduction and motivation</b>	<b>1</b>
1.1	Introduction . . . . .	1
1.2	Goals of the master thesis . . . . .	3
1.3	Thesis structure and description . . . . .	4
<b>2</b>	<b>Quantum information theory</b>	<b>6</b>
2.1	Introductory definitions . . . . .	6
2.1.1	Probability . . . . .	6
2.1.2	Bipartite systems and entanglement . . . . .	6
2.2	Quantum operators . . . . .	8
2.3	Measurements . . . . .	10
2.3.1	Projective measurement . . . . .	10
2.3.2	POVM measurement . . . . .	12
2.4	Communication channels . . . . .	13
2.4.1	Classical channels . . . . .	13
2.4.2	Quantum channels . . . . .	15
2.5	Entropy . . . . .	17
2.5.1	Shannon entropy . . . . .	19
2.5.2	Entropies of two variables . . . . .	21
2.5.3	Von Neumann entropy . . . . .	25
2.5.4	One-Shot entropies . . . . .	29
2.5.5	Entropic uncertainty principle . . . . .	32
<b>3</b>	<b>Quantum Key Distribution Protocols</b>	<b>35</b>
3.1	Fundamental principles . . . . .	36
3.1.1	No-cloning theorem . . . . .	36
3.1.2	Quantum description of Mach-Zehnder interferometer . . . . .	38
3.2	Discrete variable QKD . . . . .	41
3.2.1	BB84 protocol . . . . .	42
3.2.2	B92 protocol . . . . .	46
3.2.3	BB84 and B92 protocols with time and phase encoding . . . . .	47
3.2.4	E91 protocol . . . . .	52
3.2.5	BBM92 protocol . . . . .	53
3.3	Post-processing . . . . .	54
3.3.1	Parameter estimation . . . . .	54
3.3.2	Error correction . . . . .	58
3.3.3	Privacy amplification . . . . .	62
<b>4</b>	<b>Security of QKD</b>	<b>65</b>
4.1	One-time pad . . . . .	65
4.2	Practical limitations . . . . .	66



4.3	Security model of QKD . . . . .	67
4.4	Attacks on QKD . . . . .	68
4.4.1	Eavesdropping . . . . .	69
4.4.2	Photon number splitting attack . . . . .	71
4.5	Detector control attack . . . . .	73
4.6	Modified QKD protocols . . . . .	74
4.6.1	SARG04 protocol . . . . .	74
4.6.2	Decoy states protocol . . . . .	75
4.6.3	Differential phase shift . . . . .	76
4.6.4	Coherent one-way protocol . . . . .	77
<b>5</b>	<b>Simulation of phase-time BB84 protocol</b>	<b>80</b>
5.1	Quantum transmission . . . . .	81
5.2	Post-processing . . . . .	83
5.3	Simulation results . . . . .	84
5.3.1	QBER analysis for different scenarios . . . . .	85
5.3.2	Key length analysis . . . . .	87
<b>6</b>	<b>Experimental part</b>	<b>91</b>
6.1	Experimental setup description . . . . .	91
6.1.1	Polarization controller . . . . .	94
6.1.2	Phase modulator . . . . .	94
6.1.3	Single-photon avalanche diode . . . . .	95
6.1.4	Microcomputer and programming environment . . . . .	96
6.2	Stages of the experiment . . . . .	96
6.2.1	Equipment testing . . . . .	97
6.2.2	Constructing Alice . . . . .	99
6.2.3	Connecting Alice and Bob . . . . .	101
6.2.4	Quantum measurements . . . . .	103
<b>7</b>	<b>Results and conclusions</b>	<b>109</b>
	<b>Bibliography</b>	<b>112</b>

# 1 Introduction and motivation

## 1.1 Introduction

Cryptography, one of the oldest branches of mathematics, has been practiced since ancient times. It was used for sending secret messages containing sensitive information about armies or about intel gained from spying on the enemy. Probably the most famous example from history is Caesar's cipher used in ancient Rome. In the past, solely symmetric ciphers were used, where each party has the same secret key, which is used for decryption. "It was not until the latter half of the 20th century when the modern digital era of computers and internet communication has given rise to asymmetric cryptography, where each party has different part of the key, called public and private key. With the public key, data can be encrypted by anyone, but only the holder of the corresponding private key can decrypt it back to plain text. That is possible only using private key. Key exchange protocols based on discrete logarithm problem or prime factorization are now an indispensable part of the modern world. We use these protocols daily. During connecting to WiFi, during communication over the internet, in cryptocurrencies, and many more. All in all, today's world is dependent on cryptography.

A persistent challenge in classical cryptography is the key distribution (key exchange) problem. All classical communication can be eavesdropped, hence symmetric cryptography can't be directly used. Instead, secured communication is realized using a combination of symmetric and asymmetric cryptography. First, asymmetric cryptography is used to distribute a secret key between Alice and Bob. Then, a combination of symmetric cryptography and special one-way function called hash functions, is used to ensure secret communication between the two parties.

Classical cryptography relies on computational tasks which are difficult to solve using today's computers. While these mathematical problems are now computationally demanding, with enough time and powerful enough computers, they can eventually be cracked.

Furthermore, the ongoing efforts to build a quantum computer that can execute Shor's algorithm pose a significant threat. This algorithm can decrypt exponentially faster the currently used asymmetric cryptographic systems, that rely on discrete logarithm problem or prime factorization. Thus we cannot be sure that our cryptography is strong enough for future use. Additionally, classical digital signatures use the same principles of asymmetrical cryptography, thus they are also vulnerable to Shor's algorithm. Unfortunately, digital signatures are used every time we want to verify someone's identity over the internet, for example during a bank transfer or when verifying authenticity of Hypertext Transfer Protocol Secure (HTTPS) digital certificates.

There are two branches of solutions to this problem. The first is to continue using computationally demanding tasks based on mathematical proofs which cannot be efficiently solved even with a quantum computer. This approach is called *post-quantum cryptography* and is currently in development. As of today, the National Institute of Standards and Technology (NIST) has chosen 2 and is working on their standardisation. NIST allowed to receive public comments until November 22, 2023 after which the standardisation should happen in the upcoming months [1], [2]. The proposed algorithms KYBER and DILITHIUM are based on lattice-based cryptography with learning with errors. This method is not based on one difficult mathematical problem, but two. The first one is the Shortest Vector Problem. Given a lattice defined by lattice basis, one must find the shortest vector from an arbitrary location in the lattice to the nearest lattice point. Additionally, the algorithms incorporate the Learning With Errors (LWE) principle, where a system of equations is intentionally overspecified—having more equations than variables and the right hand side being polluted with small errors. Each column of coefficients from the system of equations is a vector in the lattice basis. In KYBER, the genuine solution to these equations is kept as the private key, while the manipulated system with the errors, is disclosed as the public key. For encryption using this public key, an additional error is added by the sender to represent the encrypted message, positioning it as a point in the lattice; a minor error represents a bit value of 0 by keeping the point close to the lattice point, while a major error indicates a bit value of 1 by placing the point farther away. Decrypting with the private key is straightforward and allows for pinpointing the nearest lattice point. In contrast, trying to decrypt without the private key is like attempting to find that specific point without knowledge of its original placement. DILITHIUM is using the same concepts but for a digital signatures. Third proposed protocol is Hash-Based, where the hard problem is These protocols are in public drafts FIPS 203, 204 205.

There are two main fields being explored to address these cryptographic challenges. The first continues to leverage computationally intensive problems that, according to current mathematical proofs, cannot be efficiently solved by quantum computers. This strategy is known as post-quantum cryptography and is the subject of ongoing research and development efforts. The National Institute of Standards and Technology (NIST) has been the main actor in this area, having selected several algorithms for potential standardization. As of the latest updates, NIST opened public comment window on these candidates until November 22, 2023, with the formal standardization process expected to happen in the subsequent months [1], [2]. Two of the algorithms under consideration are KYBER, designated for key exchange, and DILITHIUM, intended for digital signatures. Both are founded on lattice-based cryptography. This approach relies on the hardness of the Shortest Vector Problem, which involves finding the minimal distance vector within a lattice defined by its basis points. Third candidate is the Stateless Hash-Based Digital Signature protocol, which is proposed only as a digital signatures scheme rather than key exchange. All

these protocols are currently documented as public drafts FIPS 203, 204, and 205.

The second field of research is called *quantum cryptography*, particularly *quantum key distribution* (QKD). QKD introduces a different method for distributing a symmetrical key between two parties, typically referred to as Alice and Bob, via the transmission of qubits, or quantum bits of information. Distinctively, QKD does not depend on the computational hardness of certain mathematical problems. Instead, it leverages the fundamental principles of quantum mechanics to provide secure communication. This approach guarantees unconditional security; any attempt at eavesdropping can be detected, resulting in aborting the current protocol and initiating a new session by Alice and Bob. Thus, unlike conventional methods that increasingly rely on the escalating complexity of computational challenges, QKD offers a robust solution solely on the unbreakable laws of quantum mechanics.

## 1.2 Goals of the master thesis

The aim of this work will be to get familiar with the methods of quantum cryptography, particularly with the physical description, information theory, and technical and practical aspects. Based on this knowledge, we propose our own quantum key distribution experiment and try to implement parts of the experiment. To list the assignments, in particular:

- Based on the research and study of the physical principles of quantum cryptography, familiarize yourself with the basic processes, and key components of the cryptographic system, map and process the practical and technical aspects of possible implementations, according to individual approaches.
- By comparing the properties of the individual approaches and the possibilities of implementation, both from the point of view of the whole and individual components, choose a suitable strategy with which to design a suitable quantum cryptographic system.
- Study the possibilities of practical experimental implementation for each functional part of the QKD system and, in cooperation with consultants and the trainer, choose an appropriate method, depending on the possibilities of both collaborating universities, Czech technical university and Palacky university in Olomouc.
- Try to experimentally implement and test selected components of this proposal, based on the cooperation of both workplaces. Discuss the obtained results.

### 1.3 Thesis structure and description

Firstly, information theory and quantum information theory will be discussed. Specifically, in the beginning, we start with introductory definitions, such as bipartite systems of two quantum states or a hybrid state composed of classical-quantum states. An example of the classical-quantum state can be the state immediately after the quantum part of a QKD protocol. We then discuss two important topics: Quantum operators and the representation of measurements. This allows us to describe the process of data transmission and manipulation during the protocol. Finally, we discuss classical and quantum entropy. Both are substantial for a complete understanding of a QKD protocol and its security.

In the next chapter, we introduce the most important QKD protocols. We begin with the no-cloning theorem - the fundamental theorem, essential for the proof of QKD unconditional security. Then we proceed to describe particular protocols with a focus on the BB84. We also describe the protocol in general and then a specific realization of the BB84 protocol with information encoded in the phase of a photon. Lastly, we discuss a classical part of the QKD protocol called post-processing.

Furthermore, the chapter *Security of QKD* introduces basic security definitions together with the most popular attack on QKD called *Photon number splitting attack*. We then introduce three modified protocols which prevent performing the attack.

In Chapter 5, we present a comprehensive simulation of the BB84 protocol with time and phase encoding, incorporating various customizable parameters and features essential for practical quantum key distribution systems. The simulation covers the entire QKD process, from the generation of quantum states by Alice to the transmission through a simulated quantum channel, measurement by Bob, and subsequent post-processing stages. It includes techniques such as decoy state protocols, error estimation, error correction, privacy amplification, and key authentication. Additionally, the simulation allows for the investigation of two different eavesdropping attack scenarios: the intercept-resend attack and the photon-number splitting attack.

In Chapter 6, we describe the experimental implementation of the BB84 protocol using phase and time encoding. The experimental setup is divided into Alice and Bob components, connected via an optical fiber, with a microcomputer (Red Pitaya STEMLab 125-14) serving as the control unit for laser triggering, phase modulation, and single-photon detection. We introduce the individual components of the setup, including polarization controllers, phase modulators, and single-photon avalanche diodes. The experimental procedure is outlined in stages, starting with equipment testing, followed by the construction of a symmetrical Mach-Zehnder interferometer for Alice, and then integrating Bob's asymmetrical interferometer. We present the results of classical and quantum visibility measurements, demonstrating the successful operation of the system at the single-photon level. Finally, we discuss the utilization

of the Red Pitaya microcomputer for controlling the experiment and highlight the challenges and limitations encountered during the implementation.

In the last chapter, we conclude all the goal achieved in this thesis and discuss our results.

## 2 Quantum information theory

The goal of information theory is to analyze and quantify the processes involved in acquiring, transmitting, and storing information. For example, when describing a communication channel, one may ask at which rate data can be transmitted reliably and most compressed, given that they are still recoverable. We also want to be able to describe noisy quantum channel and predict errors in the data transmission.

### 2.1 Introductory definitions

#### 2.1.1 Probability

Suppose we have a random variable  $X$  with possible outcomes  $x$ , the probability distribution  $P$  is a mapping

$$P : X \rightarrow \mathbb{R}_+. \quad (2.1)$$

The probability distribution assigns each possible outcome a probability  $p_x$  indicating how likely it is to occur. For two random variables  $X$  and  $Y$ , we can define joint probability where we want to know the probability that  $X$  and  $Y$  occurs simultaneously. The joint probability is defined as

$$P_{XY} : X \times Y \rightarrow \mathbb{R}_+. \quad (2.2)$$

When we know the joint probability and we want to find out marginal distribution of one of the probabilities, for example  $y$ , we can sum the probability over this variable as

$$P_X(x) = \sum_y P_{XY}(x,y). \quad (2.3)$$

Finally, conditional probability is defined as

$$P_{X|Y}(x|y) = \frac{P_{XY}(x,y)}{P_Y(y)}, \quad (2.4)$$

which we can interpret as what is the probability that  $X$  will happen given that  $Y$  happened. If the two random variables are independent, then it holds

$$P_{XY}(x,y) = P_X(x)P_Y(y). \quad (2.5)$$

#### 2.1.2 Bipartite systems and entanglement

A composite system formed from two Hilbert spaces,  $\mathcal{H}_A$  and  $\mathcal{H}_B$  is represented by their tensor product  $\mathcal{H}_A \otimes \mathcal{H}_B$  with basis  $\{|e_i\rangle \otimes |f_j\rangle\}_{i,j}$ , where  $\{|e_i\rangle\}$  is the basis of  $\mathcal{H}_A$  and  $\{|f_j\rangle\}$  is the basis of  $\mathcal{H}_B$ . Such composite system is described using a density matrix,  $\rho_{AB}$ , which provides the complete statistical description of the

state of systems  $A$  and  $B$ . To obtain information about the individual subsystems within the composite system, we utilize the partial density matrices  $\rho_A$  and  $\rho_B$ , derived through the operation known as the partial trace. The partial trace of the composite system  $\rho_{AB}$  with respect to system  $B$  gives the reduced state of  $\rho_{AB}$  on system  $A$ :

$$\rho_A = \text{Tr}_B \rho_{AB} = \sum_i (\mathbb{I} \otimes \langle f_i |) \rho_{AB} (\mathbb{I} \otimes |f_i\rangle), \quad (2.6)$$

where  $\mathbb{I}$  is the identity matrix on  $\mathcal{H}_A$ . This operation provides a description of an isolated system  $A$ , irrespective of the state of a system  $B$ . Similarly, the partial trace over the subsystem  $A$  results in the reduced density matrix for the subsystem  $B$ ,  $\rho_B$ . The dimension of a composite system is the product of the dimensions of its subsystems:

$$\dim(\mathcal{H}_A \otimes \mathcal{H}_B) = \dim(\mathcal{H}_A)\dim(\mathcal{H}_B). \quad (2.7)$$

A bipartite system is called entangled, if the subsystems cannot be described separately, i.e., they are not separable. Separability means that we can write the two states as a simple tensor product  $|\psi\rangle \otimes |\varphi\rangle$ . To determine if a given state is entangled, we may use the *Schmidt decomposition*, which states that a pure entangled bipartite state  $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$  can be written as

$$|\psi\rangle_{AB} = \sum_i \lambda_i |e_i\rangle_A |f_i\rangle_B, \quad (2.8)$$

where  $\lambda_i \in \mathbb{R}$ ,  $\lambda_i \geq 0$ , and satisfy  $\sum_i \lambda_i^2 = 1$ . States  $|e_i\rangle_A$  form an orthonormal basis for the system  $A$ , and states  $|f_i\rangle_B$  are orthonormal basis of the system  $B$ . Systems  $A$  and  $B$  can be thus described by partial density matrices  $\rho_A$ , and  $\rho_B$ , respectively, as

$$\rho_A = \sum_i \lambda_i^2 |e_i\rangle_A \langle e_i|_A, \quad (2.9a)$$

$$\rho_B = \sum_i \lambda_i^2 |f_i\rangle_B \langle f_i|_B. \quad (2.9b)$$

One of the most important consequence is that Schmidt decomposition tells us that the two subsystems  $A$ ,  $B$  share their eigenvalues  $\lambda_i^2$ . Moreover, the presence of more than one term in the Schmidt decomposition signifies entanglement, with the number of terms, or the Schmidt rank, serving as a quantifier of the state's entanglement.

To illustrate the concept of entanglement with a practical example, consider the Bell states, which are a specific group of maximally entangled quantum states of two qubits. Bell states are particularly significant in the study of quantum information, serving as a foundational element in protocols utilizing quantum entanglement, such as quantum teleportation or superdense coding[3]. The four Bell states are defined as follows:



$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad (2.10a)$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \quad (2.10b)$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad (2.10c)$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \quad (2.10d)$$

Each Bell state is a superposition of product states where the coefficients have equal amplitude, demonstrating perfect quantum correlations between the two qubits. For instance, measuring one qubit in the state  $|\Phi^+\rangle$  instantly determines the state of the other qubit, irrespective of the distance separating them. This non-local behavior illustrates the essence of entanglement.

## 2.2 Quantum operators

We will begin with the definition of a bounded linear operator:  $L : \mathcal{H}_A \rightarrow \mathcal{H}_B$  is a bounded linear operator such that for every state  $\forall |\psi\rangle \in \mathcal{H}_A$ , the following condition holds

$$\|L|\psi\rangle\|_{\mathcal{H}_B} \leq \alpha \|\psi\|_{\mathcal{H}_A}, \quad (2.11)$$

where  $\alpha \in \mathbb{R}$  is the smallest possible constant and is called the norm of the operator  $L$ ,  $\min(\alpha) = \|L\|$ .

A quantum operator, or superoperator, denoted as  $\mathbb{L} : \mathcal{H}_A \rightarrow \mathcal{H}_B$  is a linear map from Hilbert space  $\mathcal{H}_A$  to a Hilbert space  $\mathcal{H}_B$ . For infinite-dimensional vector spaces, quantum operator is defined as a linear map  $\mathbb{L} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ , where  $\mathcal{B}(\mathcal{H})$  is a set of all bounded linear operators acting on  $\mathcal{H}$ . In infinite-dimensional spaces, all operators must be bounded to maintain physicality. In spaces with finite dimension, all linear operators are naturally bounded.

Now we will present four key properties of a quantum operator. Firstly, we will start with definition of a linear map. A map  $\mathbb{L} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  is said to be linear if

$$\mathbb{L}(\alpha\rho_A + \sigma_A) = \alpha\mathbb{L}(\rho_A) + \mathbb{L}(\sigma_A), \quad (2.12)$$

where  $\rho_A, \sigma_A$  are density matrices in  $\mathcal{H}_A$  and  $\alpha \in \mathbb{C}$ . The principle of linearity in quantum mechanics ensures that the superposition principle holds when operators act on quantum states.

A linear map  $\mathbb{L} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  is *positive* if

$$\mathbb{L}(\rho_A) \geq 0, \quad \forall \rho_A \geq 0, \quad (2.13)$$

where  $\rho_A \in \mathcal{B}(\mathcal{H}_A)$ , i. e., a linear map  $\mathbb{L}$  is positive if it is positive semi-definitive (PSD) for all positive semi-definitive density matrices from  $\mathcal{B}(\mathcal{H}_A)$ . PSD density matrix means that all of its eigenvalues are  $\geq 0$ . A positive quantum operator means that after applying this operator to the density matrix  $\rho_A$ , all the eigenvalues of transformation  $\mathbb{L}(\rho_A)$  remain greater or equal to zero. This is, however, not sufficient for cases when we have more systems, but here we act only on a single one. Therefore we need complete positivity. Complete positivity is critical for maintaining the physical validity of quantum operations, particularly when considering subsystems of entangled systems. An operation is completely positive if, when extended to act on part of an entangled system (while the rest is left untouched), it still yields physically valid states.

A linear map  $\mathbb{L} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  is *completely positive* (CP) if  $\mathbb{L} \otimes \text{id}_n : \mathcal{B}(\mathcal{H}_A) \otimes \mathcal{B}(\mathbb{C}^n) \rightarrow \mathcal{B}(\mathcal{H}_B) \otimes \mathcal{B}(\mathbb{C}^n)$  is positive  $\forall n \in \mathbb{N}$ , where  $\text{id}_n$  is the identity map in  $\mathbb{C}^n$ . That is,  $\mathbb{L}$  maps arbitrary semi-positive operator while including all other systems which are part of the one composite system. These systems are unaffected by  $\mathbb{L}$ . Suppose we have a quantum mechanical system consisting of a initial state  $\rho = \rho_A \otimes \rho_B = 1/\sqrt{2}(|00\rangle + |11\rangle)$ , from two Hilbert spaces  $\mathcal{H}_A$  and  $\mathcal{H}_B$ , respectively. There are two quantum channels for this system. We apply transpose operator  $T$  to the quantum channel that belongs to system  $\mathcal{H}_A$  and apply no operator, i. e., identity operator to the second channel. However, our state  $\rho$  is an entangled state. That means that applying  $T \otimes \mathbb{I}$  on the state  $\rho$  will result in

$$\rho = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{T(\rho_A) \otimes \mathbb{I}(\rho_B)} \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2.14)$$

The eigenvalues of the transformed state  $(T \otimes \mathbb{I})(\rho)$  are now  $\{1/2, 1/2, 1/2, -1/2\}$ . We obtain a density matrix with a negative eigenvalue, which is not a valid density matrix. Therefore, we require for the map to be completely positive. Note that  $T$  is positive, but not completely positive. If we applied  $T$  to the whole system, the final density matrix would be valid [3].

A linear map  $\mathbb{L}$  is *trace preserving* (TP) if

$$\text{Tr}(\rho_A) = \text{Tr}(\mathbb{L}(\rho_A)), \quad \forall \rho_A \in \mathcal{B}(\mathcal{H}_A). \quad (2.15)$$

Finally, we define a quantum channel  $\mathcal{E} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  as a linear, completely positive trace-preserving (CPTP) map. Alternative option to define quantum channel is the following equivalence: a map  $\mathcal{E} : \mathcal{H}_A \rightarrow \mathcal{H}_B$  is linear CPTP map  $\Leftrightarrow$  there exist such  $K_i : \mathcal{H}_A \rightarrow \mathcal{H}_B$  that

$$\mathcal{E}(\rho_A) = \sum_{i=1}^d K_i \rho_A K_i^\dagger, \quad (2.16)$$

where  $\rho_A \in \mathcal{B}(\mathcal{H}_A)$ ,  $d \leq \dim(\mathcal{H}_A)\dim(\mathcal{H}_B)$ , and  $\forall i \in \{1, \dots, d\}$ ,

$$\sum_{i=1}^d K_i^\dagger K_i = \mathbb{I}_A. \quad (2.17)$$

The equation (2.16) is called the Kraus decomposition and ensures that in order for a quantum operator to be a quantum channel acting on state  $\rho_A$ , it has to be possible to write the operator as (2.16) for some set of Kraus operators  $\{K_i\}$ . The Kraus representation is a powerful tool because it provides a flexible and comprehensive way to model the evolution of quantum systems under various influences, including interactions with the environment and noise, in a way that always respects the rules of quantum mechanics. This approach is especially useful because it does not require detailed knowledge of the environment or the specific interactions involved. Instead, we focus only on the outcomes (represented by the Kraus operators), such as "Will the qubit be flipped after traveling through the environment, or will it remain unchanged?" Using the Kraus representation, we can describe this process with two operations: the Kraus operators. One Kraus operator that leaves the state as it is, with a certain probability. Another that flips the state, also with some probability. Using the Kraus operators, we can calculate the final state of the qubit after it passed through the noisy environment, ensuring that the probabilities add up correctly and that the result is a valid quantum state.

## 2.3 Measurements

### 2.3.1 Projective measurement

In quantum mechanics, a measurement is described by operators  $M_X : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ , where  $x$  is the possible outcome from finite outcome set  $X$ . A set of measurement operators must satisfy the completeness relation

$$\sum_{x \in X} M_x^\dagger M_x = \mathbb{I}. \quad (2.18)$$

Additionally, if the measurement operators can be written as a spectral decomposition of orthogonal projectors, we call it projective measurement

$$M = \sum_x x P_x, \quad (2.19)$$

where  $x$  are the eigenvalues of  $M$ ,  $P_x P_{x'} = \delta_{xx'} P_x$ , and  $\sum_x P_x = \mathbb{I}$  are orthogonal projectors. Simple examples are measurement operators for  $Z$  basis

$$M_0 = |0\rangle \langle 0|, \quad (2.20a)$$

$$M_1 = |1\rangle \langle 1|, \quad (2.20b)$$

or measurement operators for  $X$  basis

$$M_+ = \frac{1}{2}(|0\rangle + |1\rangle)(\langle 0| + \langle 1|) = |+\rangle \langle +|, \quad (2.21a)$$

$$M_- = \frac{1}{2}(|0\rangle - |1\rangle)(\langle 0| - \langle 1|) = |-\rangle \langle -|, \quad (2.21b)$$

where the states  $|0\rangle, |1\rangle$  are defined as:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (2.22)$$

We distinguish between the state immediately before measurement happens, from which we get probability that result  $x$  occurs, i.e., the result is a number. And between the state after measurement, where the result is how the state looks like after the measurement, i.e., the result is a state. For a pure state, state before measurement is

$$P_\psi(x) = \langle \psi | M_x^\dagger M_x | \psi \rangle, \quad (2.23)$$

and the state after measurement is

$$|\psi'\rangle = \frac{M_x |\psi\rangle}{\sqrt{\langle \psi | M_x^\dagger M_x | \psi \rangle}}. \quad (2.24)$$

To explain this on an example, assume we have an arbitrary state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , where  $|\alpha|^2 + |\beta|^2 = 1$ . To find out the probability of measuring the state  $|0\rangle$ , we use equation (2.23), with measurement operator  $M_0$  from equation (2.20a), which yields

$$P_\psi(0) = \langle \psi | M_0^\dagger M_0 | \psi \rangle = \langle \psi | M_0 | \psi \rangle = |\alpha|^2, \quad (2.25)$$

where we used the fact that  $M_0^\dagger M_0 = M_0^2 = M_0$  and the same applies for  $M_1$ . Thus, the probability of measuring  $|0\rangle$  is  $|\alpha|^2$ . Analogously, we would get  $P_\psi(1) = |\beta|^2$ . When we actually measure the state  $|\psi\rangle$ , it collapses into one of the states from the superposition. This is denoted by the state of the system after the measurement

$$|\psi'_0\rangle = \frac{M_0 |\psi\rangle}{\sqrt{\langle \psi | M_0^\dagger M_0 | \psi \rangle}} = \frac{\alpha}{|\alpha|} |0\rangle, \quad (2.26a)$$

$$|\psi'_1\rangle = \frac{M_1 |\psi\rangle}{\sqrt{\langle \psi | M_1^\dagger M_1 | \psi \rangle}} = \frac{\beta}{|\beta|} |1\rangle. \quad (2.26b)$$

Hence, two states, which can be measured, are either  $|0\rangle$  or  $|1\rangle$ . The factors  $\frac{\alpha}{|\alpha|}$  and  $\frac{\beta}{|\beta|}$  are basically only a number  $e^{i\varphi}$ , where  $\varphi \in \mathbb{R}$ , called global phase. This global phase has modulus equal to one and hence it does not have any effect on the measurement probability.

### 2.3.2 POVM measurement

To achieve greater precision in distinguishing quantum states, especially when they are not orthogonal, we turn to the Positive Operator-Valued Measure (POVM)[3, 4]. We define the POVM operator as a set of operators  $E_x$ , known as POVM elements

$$E_x = M_x^\dagger M_x, \quad (2.27)$$

where the POVM elements fulfill the completeness condition

$$\sum_x E_x = \mathbb{I}, \quad E_x \geq 0 \quad \forall x \in X. \quad (2.28)$$

and the positivity condition, ensuring that each POVM element is positive semi-definite ( $E_x \geq 0$  for all  $x$ ), guaranteeing that the probability of any outcome is non-negative.

The probability of observing outcome  $x$  when measuring a state  $|\psi\rangle$  using a POVM is given by substituting the operator  $E_x$  into equation (2.23)

$$P_\psi(x) = \langle \psi | E_x | \psi \rangle. \quad (2.29)$$

This equation gives us the primary use of POVMs: determining the likelihood of various measurement outcomes. However, while POVMs excel in calculating the probabilities of outcomes, they do not directly provide the post-measurement state as projective measurements do. This is because POVMs involve a more abstract representation that may not correspond to a physical operation on the state alone.

In equation for the state after measurement (2.24), there is  $M_x$  alone, thus we cannot use the POVM for this type of measurement. We use the POVM only to find out the probability of the measurement. The advantage is the ability to better distinguish non-orthogonal states. With projective measurements, non-orthogonal states cannot be distinguished with certainty, we always end up with a probability to misidentify the state. However, a set of the POVM measurements can have larger dimensions, for example, consider three positive operators  $E_i$

$$E_0 = \frac{\sqrt{2}}{1 + \sqrt{2}} |1\rangle \langle 1|, \quad (2.30a)$$

$$E_1 = \frac{\sqrt{2}}{1 + \sqrt{2}} \frac{|-\rangle \langle -|}{2}, \quad (2.30b)$$

$$E_2 = \mathbb{I} - E_0 - E_1, \quad (2.30c)$$

and two non-orthogonal states  $|\psi_0\rangle = |0\rangle$  and  $|\psi_1\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle)$ . When we calculate the respective probabilities

$$\langle \psi_0 | E_0 | \psi_0 \rangle = 0,$$

$$\langle \psi_1 | E_1 | \psi_1 \rangle = 0,$$

$$\langle \psi_0 | E_2 | \psi_0 \rangle = \langle \psi_1 | E_2 | \psi_1 \rangle = \frac{1}{2},$$

we observe that if we measure 0, the measured state had to be  $|\psi_1\rangle$ , and when we measure 1, it had to be  $|\psi_0\rangle$ . However, if we measure outcome 2, we cannot deterministically determine which state was measured because both  $|\psi_0\rangle$  and  $|\psi_1\rangle$  yield a probability of  $1/2$ .

## 2.4 Communication channels

### 2.4.1 Classical channels

In information theory, channels serve as the medium for transmitting information. A channel is a probabilistic mapping from  $X$  to  $Y$  given by a conditional probability distribution  $p_{y|x}$ . If a channel is noiseless, transmitted information will remain the same. However, naturally occurring noise can lead to the partial or complete loss of transmitted information.

A fundamental model used to represent noise in classical communication is the *bit-flipping channel* [3]. When Alice is transmitting bit with value 0, Bob has a probability  $p$ , that the bit will remain 0 or probability  $1 - p$  that the bit will flip to 1. Mathematically speaking

$$0 \xrightarrow{\text{Channel}} p(0) + (1 - p)(1), \quad (2.31a)$$

$$1 \xrightarrow{\text{Channel}} p(0) + (1 - p)(1). \quad (2.31b)$$

Diagram of this channel is depicted in Figure 2.1. This channel, where the bit-flip probability for 0 and 1 is the same, is called *Binary symmetric channel*.

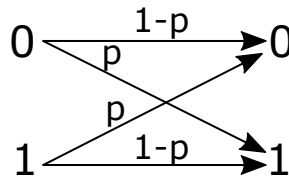


Figure 2.1: Diagram of a classical bit-flipping channel. Transmitted bit 0 has a probability of  $(1 - p)$  to be flipped to 1.

The simplest method to mitigate these errors involves sending three bits for each intended bit, where every bit has initially the same value. Assume Alice wants to send bit 1 to Bob. If she sends only one bit, Bob measures the probability  $p$  that the bit remained the same. However, if not, Bob has no way how to find out on his own that an error occurred. To solve this, we can use repetition code with majority voting. It means that we repeat the sent information multiple times and decide,

according to the majority of bits, what was the initial bit string. For example, if Alice sends three bits 111, Bob can receive string which is the same 111, with one bit flipped - 011, with two bits flipped - 001, or with all three bits flipped - 000. Bob will now choose the more probable option. In the first two cases, he can assume that the original bit was 1. In the latter two, he will think that the original bit was 0. Probability that all three qubits will flip is  $p^3$ . Probability that two or more qubits will flip is  $3p^2(1-p) + p^3 = 3p^2 - 2p^3$ . This method however requires as much as three as much bits to send. More robust technique for error correction will be described in section 3.3.2.

We will now present calculation of channel capacity for the binary symmetric channel, which can be applied also for a quantum key distribution protocol.

Assume we are randomly choosing between sending bit 0 or bit 1, thus  $p_X(0) = p_X(1) = 1/2$ . The probability of the bit-flips is given by equation (2.31). Hence the probability of receiving bit 0 is the sum of the case where we choose to send bit 0 times the bit is not flipped plus probability of sending bit 1 times the probability that the bit is flipped:  $p_X(0) \cdot (1-p) + p_X(1) \cdot p$ . The capacity (2.53) is then

$$\begin{aligned}
 C(\mathcal{N}) &= \max_{p_X(x)} H(Y) - H(Y|X) = \\
 &= \max_{p_X(x)} \left[ H\left(\frac{1}{2} \cdot (1-p) + \frac{1}{2} \cdot p, \frac{1}{2} \cdot p + \frac{1}{2}(1-p)\right) - h(p) \right] = \\
 &= H\left(\frac{1}{2}\right) - h(p) = 1 - h(p).
 \end{aligned} \tag{2.32}$$

where  $h(p)$  is the binary entropy from equation (2.44). Since both source probability distributions are the same, we did not have to maximize over them. This result tells us the theoretical limit of what bit error rate (BER) we are able to correct and how the error correction will reduce the channel's capacity. BER indicates the probability that a bit is flipped. From the graph in Figure 2.2, we can observe that the maximal channel capacity is with no BER and is minimal with BER of 0.5. When the BER exceeds 0.5, it implies that each bit is more likely to be flipped than not. Consequently, by simply inverting every received bit, the correct information can be retrieved, thereby effectively reducing the error rate to 1-BER.

Assuming that the BER is  $p = 10\%$ , this suggests that up to 53% of the channel's bandwidth can be used effectively for transmitting actual data reliably under ideal coding and decoding conditions. The remaining 47% would be effectively consumed by the need to introduce redundancy or error-correction codes that allow us to manage the errors introduced at a rate of 0.1 (10% of the bits being erroneous).

In a practical setting, such as when using QKD to generate and securely share a cryptographic key, a quantum bit error rate (QBER) of 0.1 necessitates using nearly half of the channel's capacity to handle errors. Effective error-correction mechanisms are designed to mitigate these errors. Some of them will be described in the Error Correction section 3.3.2 in the next chapter. However, they will require using a part

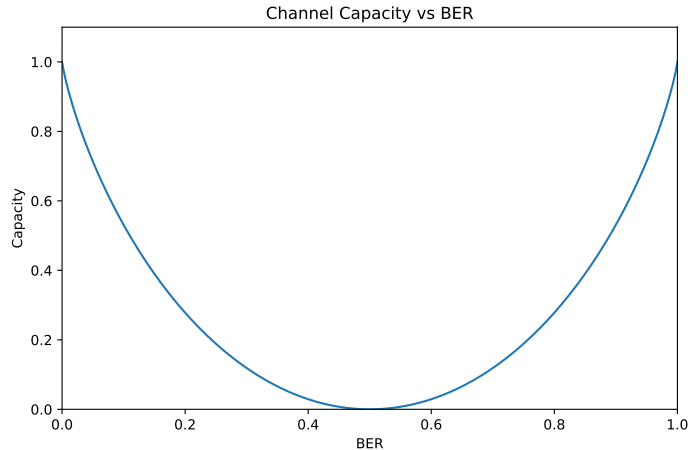


Figure 2.2: Capacity of a binary symmetric channel with respect to the binary error rate.

of the channel’s capacity to transmit the extra information needed to identify and correct errors.

In practice, classical channel can be, for example, an Ethernet cable, Wi-Fi network, or even the propagation of sound waves when two people speak.

### 2.4.2 Quantum channels

In quantum communication, the concept of quantum channels is fundamental for transmitting quantum information, represented by density matrices  $\rho$ . We defined quantum channel in the section 2.2 as a linear, completely positive trace-preserving (CPTP) map [3]. Actions of quantum channels on transmitted qubits are described through the action of quantum operators, which transform quantum states from one configuration to another. These channels can be for example optical fibers transmitting photons.

To model the general influence of both the environment and the noise altering the state, we use a noisy quantum channel, denoted as  $\mathcal{N}$ . The process begins with an input quantum state  $\rho_{in}$ , which represents the information for transmission. Alongside this, the environmental state, represented as  $\rho_E$ , models the surroundings assumed to interact with the quantum system. The input state together with the environment combined form a composite state  $\rho_{in} \otimes \rho_E$ .

As this composite system travels through the quantum channel, it undergoes a unitary transformation  $U$ , which represents the evolution of the system by including complex interactions between the quantum state and the environment.<sup>1</sup>

---

<sup>1</sup>Unitary transformation is reversible and conserves probability. It is defined by the condition  $U^\dagger U = I$ , where  $U^\dagger$  is the conjugate transpose of  $U$ .



After passing through the quantum channel, the output state of the system,  $\rho_{out}$ , is obtained by tracing out the environmental factors from the transformed composite state [5]:

$$\mathcal{N}(\rho_{in}) = \rho_{out} = \text{Tr}_E[U(\rho_{in} \otimes \rho_E)U^\dagger].$$

This operation effectively removes the environmental components, focusing on how the quantum system itself has changed. The environment's state post-transmission,  $\rho_E$ , is also of interest and is updated by tracing out the system's part from the transformed state:

$$\rho_E = \text{Tr}_B[U(\rho_{in} \otimes \rho_{in})U^\dagger],$$

indicating how the interaction has altered the environment.

In practical scenarios, the input to the channel often involves a probabilistic ensemble of states, not just a single state. Each quantum state  $\rho_i$  in the ensemble is prepared with a certain probability  $p_i$ , and the overall input to the channel is a mixture:

$$\sigma_{in} = \sum_i p_i \rho_i,$$

where  $\sigma_{in}$  is the average of all possible prepared states. The corresponding average output after the channel's influence is computed as:

$$\sigma_{out} = \mathcal{N}(\sigma_{in}) = \sum_i p_i \mathcal{N}(\rho_i).$$

We will now consider only simple noise models without complex interaction with the environment. The first model is again the bit-flipping quantum channel. It works on the same principle as the classical bit-flipping channel. But here, the classical binary states are replaced by quantum states  $|0\rangle$  and  $|1\rangle$  and the probability of the bit flip  $|0\rangle \rightarrow |1\rangle$  occurs with a probability of  $(1 - p)$ . The Kraus operators  $K_1, K_2$  are

$$K_1 = \sqrt{p}I = \sqrt{p}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right), \quad (2.33a)$$

$$K_2 = \sqrt{1-p}\left(|0\rangle\langle 1| + |1\rangle\langle 0|\right). \quad (2.33b)$$

Another type of quantum channel is the *Amplitude damping channel*, which models the energy dissipation from a quantum system to its environment. This channel is particularly relevant for understanding how excited states ( $|1\rangle$ ) decay to ground states ( $|0\rangle$ ) with a certain probability  $p$ , representing the loss of quantum information through interaction with the environment. Transmitted ground states remain

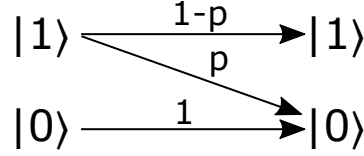


Figure 2.3: Diagram of amplitude damping channel. For state  $|1\rangle$ , we have probability of  $p$  to flip to bit 0 and  $1 - p$  to transmit correctly without altering its value. State  $|0\rangle$  remains in its state with probability of 1.

unchanged. This channel is mathematically represented by two Kraus operators  $K_1$  and  $K_2$

$$K_1 = \sqrt{p}|0\rangle\langle 1|, \quad (2.34a)$$

$$K_2 = |0\rangle\langle 0| + \sqrt{1-p}|1\rangle\langle 1|. \quad (2.34b)$$

Another quantum channel is the *phase damping channel*. This channel describes the loss of quantum information without energy dissipation. For example, this might occur when a photon is scattered randomly from particles as it travel through a medium like a waveguide. Unlike amplitude damping, the energy eigenstates of the quantum system does not change, but the relative phase between the states in the superposition is changed. A simple visualization can be imagined as a random rotation around z-axis by an angle  $\theta$ , where the magnitude of  $\theta$  follows a Gaussian distribution with a mean of 0 and a variance of  $2\lambda$ . Regarding the density matrix  $\rho$ , amplitude damping channel alters the diagonal elements, which reflect the superposition state's amplitude (in other words, the likelihood of measuring a specific state). In contrast, the phase damping channel reduces the off-diagonal elements, keeping the probability of measurement unchanged. Instead, it reduces the coherence of the phase relationship, making the off-diagonal elements of the density matrix, which represent the system's coherence, diminish over time.

For example a state  $|\psi\rangle = a|0\rangle + b|1\rangle$  with density matrix  $\rho$  will transform as

$$\rho = \frac{1}{2} \begin{bmatrix} \|a\|^2 & ab^* \\ a^*b & \|b\|^2 \end{bmatrix} \xrightarrow{\text{Phase damping}} \frac{1}{2} \begin{bmatrix} \|a\|^2 & ab^*e^{-\lambda} \\ a^*be^{-\lambda} & \|b\|^2 \end{bmatrix}, \quad (2.35)$$

where  $\lambda$  is the variance of a Gaussian distribution. We see that the diagonal elements remained unchanged and only the off-diagonal elements are being dumped to zero with time.

## 2.5 Entropy

Since entropy (in this case informational entropy) is usually misunderstood concept, we will begin with a more intuitive introduction in the terms of probability and

surprise, alternatively called information or information content. We define surprise  $s$  as the logarithm of the inverse of probability

$$s = \log_2(1/p). \quad (2.36)$$

From now on, all logarithms will be considered in the base 2. For  $p = 1$ , we get  $s = 0$ , which can be interpreted that there is no surprise when we know with a probability equal to 1 what will happen. On the other hand, for a limit scenario where  $p$  is approaching 0, we get  $s \xrightarrow{p \rightarrow 0} \infty$ , so now the surprise is maximal. In other words, the surprise for an event with 0% probability to occur is infinite if such event happens. For independent probabilities, the surprise is additive. For instance, with a biased coin, which has a probability of 70 % of getting heads and 30 % chance of getting tail, the two corresponding surprises will be equal to  $s_{head} = \log(1/0.7) \approx 0.51$  and  $s_{tail} = \log(1/0.3) \approx 1.74$ . For a sequence of head, head, tail, the probability for independent events is  $p = 0.7 \cdot 0.7 \cdot 0.3$ , and the surprise is

$$\begin{aligned} s &= \log(1/p) = \log\left(\frac{1}{0.7 \cdot 0.7 \cdot 0.3}\right) = \\ &= \log(1/0.7) + \log(1/0.7) + \log(1/0.3) = \\ &= s_{head} + s_{head} + s_{tail} \doteq 2.77, \end{aligned} \quad (2.37)$$

where we used the properties of logarithms. The average surprise per one coin flip will be the surprise of getting head times the probability of getting head + the same for tail. To generalize it for  $n$  possible outcomes

$$\langle s \rangle = \sum_{i=1}^n p_i s_i = \sum_{i=1}^n p_i \log(1/p_i) = - \sum_{i=1}^n p_i \log(p_i). \quad (2.38)$$

This is one of the ways to understand information entropy  $H$  - as an expected value of surprise  $H = \langle s \rangle$ . It's worth noting that for normalized probability  $p_i \in \langle 0, 1 \rangle$ ,  $\sum p_i = 1$ , the maximum entropy reaches maximal value  $\log(n)$  which corresponds to the least amount of knowledge if event 1 or 2 will occur. We can observe that this particular case happens when we are calculating entropy of uniform probability distribution. Because  $p_i$  is the same for all  $i$ , precisely  $p_i = 1/n \forall i$ , we can simplify the equation as

$$H = - \sum_{i=1}^n p_i \log(p_i) = -n \left( \frac{1}{n} \log\left(\frac{1}{n}\right) \right) = \log(n). \quad (2.39)$$

As we have discussed, the concept of surprise can be understood also as an information content. The degree of surprise provides a measure of the new information we gain from an observation. High surprise corresponds to a significant amount of new information, while low surprise indicates minimal new information learned. This concept is sometimes framed in terms of uncertainty: entropy quantifies the uncertainty inherent in a probability distribution. Understanding entropy as a measure of uncertainty sets the stage for introducing Shannon entropy, which further formalizes these ideas within the context of information theory.

### 2.5.1 Shannon entropy

As a natural extension of the discussion on informational entropy, we introduce Shannon's entropy. Shannon entropy is defined mathematically as in the equation (2.39)

$$H(X) = H(p_1, p_2, \dots, p_n) \equiv \sum_x p_x \log\left(\frac{1}{p_x}\right) = - \sum_x p_x \log(p_x), \quad (2.40)$$

where  $X$  is a random variable,  $x$ -s represent its possible values and  $p_x$  is corresponding probability distribution. For convenience, in quantum information, we define  $0 \log 0 = 0$  [3]. Note that the Shannon entropy does not depend on what phenomena  $X$  represents, but it only depends on the probability of occurring the random variable  $x$ , i.e., on the probability distribution of the outcomes.

It is practical to use log base 2 in defining entropy, particularly in the context of information theory, where information is typically encoded using bits. For example, a simple coin toss, which can result in one of two outcomes, uses one bit of information since  $\log(2^1) = 1$ . Similarly, for a dice throw with six uniformly distributed possible outcomes, the entropy is calculated as  $\lceil \log(6) \rceil \doteq \lceil 2.585 \rceil = 3$  bits. This rounding up ensures that enough bits are used to represent all possible outcomes (numbers 1 to 6). For  $n$  dice throws, the required number of bits scales accordingly and is given by  $\lceil \log(6^n) \rceil$  bits, reflecting the increase in possible outcome combinations.

To put this into perspective, we will now discuss the Shannon entropy's relation with data compression using so called *prefix-free variable length coding*. Prefix-free coding is a type of coding that ensures no code is a prefix of any other code, meaning that no encoded message can be mistaken for the beginning of another, which provides clear and unambiguous decoding. Variable-length encoding method assigns different lengths of bit strings to different symbols based on their probabilities or frequencies of occurrence, allowing for more frequent symbols to have shorter codes and less frequent symbols to have longer codes, thus optimizing the overall encoding efficiency.

As we will now see, the entropy is strongly related to data compression. When Shannon defined the entropy, he wanted to quantify what is the minimal amount of resources needed to store a given information. This describes *Shannon's noiseless coding theorem*, which states that the minimum resources are defined by Shannon's entropy  $H(X)$ . If we use the entropy as defined with  $\log_2$ , we will get the most optimal compression in the number of bits, in other words, the length of a bit string needed to describe all possible outcomes. For example, for a random variable  $X$ , suppose we have 5 possible outcomes  $\{x_1, x_2, x_3, x_4, x_5\}$  with respective probabilities

$\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}\}$ , the Shannon entropy will be equal to

$$\begin{aligned}
 H(X) &= \sum_x p_x \log\left(\frac{1}{p_x}\right) = \\
 &= -\frac{1}{2} \log\left(\frac{1}{2}\right) - \frac{1}{4} \log\left(\frac{1}{4}\right) - \frac{1}{8} \log\left(\frac{1}{8}\right) \\
 &\quad - \frac{1}{16} \log\left(\frac{1}{16}\right) - \frac{1}{16} \log\left(\frac{1}{16}\right) = \\
 &= 1.875.
 \end{aligned} \tag{2.41}$$

Therefore, the most optimal compression will be in average 1.875 bits. Nevertheless, this information does not tell us how this encoding should look like. Also note that important property of the encoding is the ability to decode it unambiguously. Let's take a closer look at which encoding we can actually use and if we can reach the optimal value predicted by Shannon entropy. First, encoding  $E_1$  assigns these bit values to the 5 outcomes:

$$x_1 \rightarrow 000 \quad x_2 \rightarrow 001 \quad x_3 \rightarrow 010 \quad x_4 \rightarrow 011 \quad x_5 \rightarrow 100$$

This scheme ensures each outcome is encoded with three bits, leading to an average length of 3 bits per outcome. However, according to Shannon's noiseless coding theorem, we can achieve better compression because this method uses more bits than the Shannon entropy suggests. Let's try to optimize further, we will assign shorter codes to more probable outcomes:

$$x_1 \rightarrow 0 \quad x_2 \rightarrow 10 \quad x_3 \rightarrow 110 \quad x_4 \rightarrow 1110 \quad x_5 \rightarrow 1111$$

then the average bit length of this encoding scheme  $E_2$  can be calculated as the number of bits of the outcome times the respective probability

$$\langle E_2 \rangle = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + 2 \times \frac{1}{16} \cdot 4 = 1.875 \text{ bits}, \tag{2.42}$$

which aligns exactly with Shannon's entropy, confirming its optimality. Moreover, each code in this scheme is uniquely decodable because no code is a prefix of any other (prefix-free property), ensuring that even concatenated sequences of these codes can be unambiguously decoded. So in a long string of bits, we will be always able to decode which bits are which outcome. For instance, string 0110101111 can be decoded in only one possible way as 0 110 10 1111.

We can try to compress further using encoding  $E_3$  that assigns the following bits to each variable:

$$x_1 \rightarrow 0 \quad x_2 \rightarrow 1 \quad x_3 \rightarrow 10 \quad x_4 \rightarrow 11 \quad x_5 \rightarrow 100$$

This scheme yields an average bit length of

$$\langle E_3 \rangle = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 1 + \frac{1}{8} \cdot 2 + \frac{1}{16} \cdot 2 + \frac{1}{16} \cdot 3 = 1.3125 \text{ bits}, \quad (2.43)$$

which seems more efficient than Shannon's predicted entropy. However, this encoding lacks the prefix-free property, leading to potential ambiguities during decoding. For example, the bit string '011010' could be interpreted in multiple ways, such as 0 1 10 10 or 01 10 10. Thus, despite its lower average bit length, is impractical due to its ambiguity and does not satisfy the requirements of a robust encoding scheme.

We can conclude that we have found the most optimal encoding  $E_2$ , since it aligns with prediction from Shannon's theorem and fulfils the prefix-free property. Thus, any further attempts in compression will result in a loss of information.

Special case of the Shannon entropy is the *Binary entropy* which is a Shannon entropy with only two possible outcomes, for example coin flipping. For two outcomes with the probability of occurring  $p$  and  $1 - p$ , we can write the binary entropy as

$$h(p) = -p \log(p) - (1 - p) \log(1 - p). \quad (2.44)$$

### 2.5.2 Entropies of two variables

For two variables  $X$  and  $Y$ , we define various types of entropies, similarly to probabilities, to express the relationships between  $H(X)$  and  $H(Y)$ . Joint entropy  $H(X,Y)$  represents the total uncertainty about the pair  $X, Y$  and is defined using the joint probability  $p(x,y)$

$$H(X,Y) = - \sum_{x,y} p(x,y) \log(p(x,y)). \quad (2.45)$$

This value essentially represents the union of the uncertainties in both entropies. Suppose Alice has a random variable  $X$  with entropy  $H(X)$ , but if Alice sends us  $Y$ , we then know the value of  $Y$ , in other words, we have learnt  $H(Y)$  bits of information about the pair  $(X,Y)$  and if  $X$  and  $Y$  are dependent on each other, we learn new information also about  $X$  from the value of  $Y$ . The conditional entropy  $H(X|Y)$  quantifies how much additional information about  $X$  we learn

$$H(X|Y) = H(X,Y) - H(Y), \quad (2.46)$$

or in terms of conditional probabilities

$$H(X|Y) = - \sum_{x,y} p(x|y) \log(p(x|y)). \quad (2.47)$$

From equation (2.46) we see that the relation between joint entropy and conditional entropies is given by:

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y), \quad (2.48)$$

which leads to the following inequalities:

$$H(X, Y) \geq H(X), \quad H(X, Y) \geq H(Y). \quad (2.49)$$

For conditional entropy, the following inequality holds

$$H(X) \geq H(X|Y). \quad (2.50)$$

Instinctively, this relation must be true because  $H(X)$  represents the uncertainty about  $X$  before any information about  $Y$  is known.  $H(X|Y)$  represents the remaining uncertainty about  $X$  after learning  $Y$ . Thus if  $X$  and  $Y$  are independent,  $H(X) = H(X|Y)$ . But if  $X$  and  $Y$  are dependent, learning  $Y$  reduces our uncertainty about  $X$ , hence  $H(X|Y) < H(X)$ . In other words, we can state that conditioning on another variable does not increase the entropy of a random variable [6, 3].

The mutual information  $I(X : Y)$  measures how much variables  $X$  and  $Y$  have in common, or in other words, how much we can learn about  $X$  from  $Y$ . To calculate mutual information, we simply add the entropy of  $X$  and  $Y$  together and subtract their joint entropy

$$I(X : Y) = H(X) + H(Y) - H(X, Y) = H(X) - H(X|Y). \quad (2.51)$$

The result is their intersection, as graphically depicted in Figure 2.4. In terms of probability distributions, we can write

$$I(X : Y) = \sum_{x,y} p(x,y) \log \left( \frac{p(x,y)}{p(x)p(y)} \right). \quad (2.52)$$

From this definition, one can observe that for independent events  $X$  and  $Y$ , the mutual information is equal to zero. We can easily verify if we substitute  $p(x,y) = p(x)p(y)$  to the logarithm. Thus if  $X$  and  $Y$  are independent, they have no mutual information. On the other hand, if  $X$  and  $Y$  are strongly dependent on each other, we can predict the outcome of  $X$  with high probability by knowing the outcome of  $Y$ .

In cryptography, particularly in systems like encryption, the security of a cryptographic algorithm is often assessed using the information theory. One of the critical measures used is the mutual information between the plaintext ( $X$ ) and the ciphertext ( $Y$ ). For a cryptographic system to be considered secure, the mutual information between the plaintext and the ciphertext should be as low as possible. Ideally,  $I(X : Y)$  should be zero, meaning that knowing the ciphertext  $Y$  gives no information about the plaintext  $X$ . On the other hand, if the mutual information is high, it indicates that a significant amount of information about the plaintext can be inferred from the ciphertext. To evaluate the security of an encryption system, one could analyze how each symbol of the plaintext affects the symbols of the ciphertext. If the transformation (encryption) carried out by the system results in a ciphertext

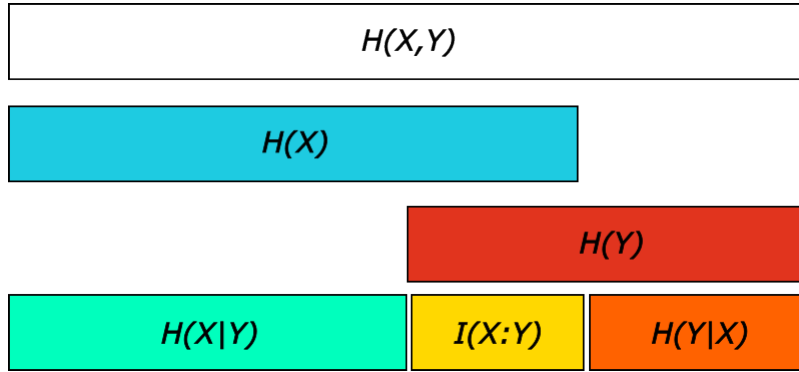


Figure 2.4: This diagram offers a visual representation of the relationships among various entropy concepts within the context of information theory. It demonstrates all the mentioned equations between each entropy. For example, the figure illustrates how the joint entropy  $H(X,Y)$ , which represents the total uncertainty of both random variables  $X$  and  $Y$  combined, can be decomposed into separate parts: the individual entropies  $H(X)$  and  $H(Y)$  with the conditional entropies  $H(X|Y)$  and  $H(Y|X)$ . Or that the sum of the mutual information and the conditional entropies equals the joint entropy, showcasing the balance between shared and exclusive information within the system. We also see that indeed  $H(X) \geq H(X|Y)$  with equality holding when  $X$  and  $Y$  are independent, in other words, they don't share any mutual information  $I(X : Y) = 0$ .

where the symbols are statistically independent of the symbols in the plaintext, the encryption is indeed secure.

$$C(\mathcal{N}) = \max_{p_X(x)} I(X : Y), \quad (2.53)$$

where  $p_X(x)$  represents the probability distribution from which Alice selects the realizations of  $X$ . This lower bound for channel capacity is sometimes called Holevo Information [7]. This equation underscores the maximum mutual information as the optimal measure of the channel's capacity to convey information reliably [6].

With the knowledge of mutual information, we can now answer how much information can be reliably transmitted through a classical noisy communication channel, denoted as  $\mathcal{N}$ . Through this channel, Alice sends the outcomes of a random variable  $X$ . The information Bob gets is affected by a noise. This output from the noisy channel  $\mathcal{N}$  is represented by a random variable  $Y$ . The maximum information rate (number of bits) that can be transmitted with negligible error across this noisy channel is determined by Shannon's noisy coding theorem. It states that the capacity of the channel  $C(\mathcal{N})$  is equal to

$$C(\mathcal{N}) = \max_{p_X(x)} I(X : Y), \quad (2.54)$$



where  $p_X(x)$  is the probability distribution from which Alice selects the realizations of  $X$ . [6].

Channel capacity quantifies the maximum mutual information, measured in bits, that can be reliably transmitted for each use of a communication channel. In this context, 'use' typically refers to 'bits per symbol'. For instance, in quantum cryptography, it is the amount of information that can be encoded and transmitted in a single photon. To determine the maximum data transmission rate, or bitrate, of a channel, the channel capacity is multiplied by the number of uses per second — that is, how frequently the channel can transmit a symbol.

The influence of noise on a communication channel significantly affects the reliability of transmitted data. It causes the output  $Y$  to deviate from the input  $X$ , and this deviation is quantitatively captured by mutual information  $I(X : Y)$ . This measure compares the actual output distribution with what would ideally occur in the absence of noise. In an ideal, noiseless channel,  $I(X : Y)$  would reach its maximum, that is equal to the entropy of  $X$  and indicating that no information is lost during transmission, meaning that  $X$  and  $Y$  would be identical.

The primary objective in optimizing communication strategies is to find a probability distribution  $p_X(x)$  that maximizes  $I(X : Y)$ . This optimization involves choosing among various distributions, each representing different strategies for using the channel's resources. For example, consider a scenario where we send one bit, or two bits, and three bits, where one out of the two, and two out of the three bits are redundant copies to enhance error correction. Here, a 'symbol' would represent a set of bits transmitted for each bit instance. So in triple repetition coding, the symbol would comprise of three bits. Determining the most effective distribution depends on the noise characteristics of the channel. In conditions of minimal noise, transmitting single bits per symbol may be the most efficient. Conversely, in highly noisy environments, a preferable strategy might be to use triple repetition per symbol/use to enhance error correction.

In the context of secure communication, private classical transmission is concerned with ensuring that any information sent over a communication channel does not get leaked to an unauthorized party (eavesdropper), for example via wiretapping. The private information rate, denoted as  $P(\mathcal{N})$ , for a channel  $\mathcal{N}$  quantifies the maximum rate at which information can be securely transmitted such that the information leakage to an eavesdropper is minimized. A key measure here is given by the expression:

$$P(\mathcal{N}) \geq \max_{p_X(x)} [I(X : Y) - I(X : E)], \quad (2.55)$$

where  $I(X : Y)$  represents the mutual information between the input  $X$  and the output at the legitimate receiver  $B$  and  $I(X : E)$  represents the mutual information between the input  $X$  and the output at the eavesdropper  $E$ . The difference  $I(X : B) - I(X : E)$  quantifies the net rate of information transfer that is se-

cure, in other words, the information gained by the legitimate receiver minus the information potentially intercepted by the eavesdropper. Our goal is to ensure safe communication prone to eavesdropping, which means to maximize this difference over all possible input distributions  $p(x)$ .

Relative entropy, also known as Kullback-Leibler divergence, serves as a measure of the discrepancy between two probability distributions  $p(x)$  and  $q(x)$ . It is defined as follows:

$$H(p(x) \parallel q(x)) = \sum_x p(x) \log \left( \frac{p(x)}{q(x)} \right) = -H(x) - \sum_x p(x) \log(q(x)), \quad (2.56)$$

where the term  $-p(x) \log(0) = +\infty$  for any  $p(x) > 0$  ensures that the divergence is unbounded when  $p(x)$  is positive and  $q(x)$  is zero, reflecting a fundamental incompatibility between the distributions in such cases.

This metric quantifies the difference between a real distribution  $q(x)$  and the original, error-less distribution is  $p(x)$ . The divergence is expressed in terms of the additional entropy introduced by using an incorrect assumption about the data's distribution. It provides yet another quantitative tool. It is used to calculate the loss of information when a model based on  $q(x)$  is used to approximate or represent a model based on  $p(x)$ . Relative entropy of 0 indicates that the two distributions contain identical quantities of information.

Consider a practical example where the true distribution of a random variable  $X$  is  $p(x)$ , and  $q(x)$  represents an estimated or assumed model. The relative entropy  $H(p(x) \parallel q(x))$  computes the expected logarithmic difference between the probabilities assigned by these two models, which can be interpreted as the expected cost in terms of additional information required to code samples from  $p(x)$  using the model  $q(x)$ . This cost becomes particularly critical in scenarios involving predictive modeling and decision-making processes, where choosing a model that minimally diverges from reality can significantly enhance performance and outcomes.

### 2.5.3 Von Neumann entropy

The quantum analog of Shannon's entropy is the von Neumann entropy  $S(\rho)$ , where instead of probability distribution, we define it using density matrix  $\rho$ . The von Neumann entropy is expressed as

$$S(\rho) = \text{Tr}(\rho \log(\rho)) = - \sum_i \lambda_i \log(\lambda_i), \quad (2.57)$$

where  $\lambda_i$  are eigenvalues of  $\rho$  derived from its spectral decomposition

$$\rho = \sum_i \lambda_i |\psi_i\rangle \langle \psi_i|, \quad (2.58)$$

where  $\{|\psi_i\rangle\}$  is an orthogonal basis of eigenvectors for  $\rho$ . For a  $d$ -dimensional space, the value of  $S(\rho)$  lies within interval  $S(\rho) \in \langle 0, \log(d) \rangle$ . This means, von Neumann entropy is non-negative,  $S(\rho) = 0$  if  $\rho$  is a pure state and  $S(\rho) = \log(d)$  when  $\rho$  is a maximally mixed state. We can interpret von Neumann entropy as the amount of uncertainty about a quantum state  $\rho$ .

As with Shannon entropy, we will now list entropies quantifying different relations between two systems, namely joint, conditional, mutual, and relative entropy. Additionally, we will discuss one more type of entropy called coherent entropy, which does not have classical analog.

Assume a bipartite state  $\rho_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$  of two quantum systems  $A$  and  $B$ ,  $\rho_{AB} = \rho_A \otimes \rho_B$ . Note that for better clarity, we will use notation  $S(A)_\rho = S(\rho_A)$ . Thanks to that, the connection between classical and quantum entropy will be more clear. The joint entropy is defined as

$$S(A,B)_\rho = -\text{Tr}(\rho_{AB} \log(\rho_{AB})), \quad (2.59)$$

and represents the total entropy of the combined system. While in the classical case the inequality (2.49) was true, in quantum world this is no longer true. For a pure bipartite state  $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ , the relation between marginal entropies is

$$S(A)_\psi = S(B)_\psi, \quad (2.60)$$

and for the joint entropy holds

$$S(A,B)_\psi = 0. \quad (2.61)$$

Therefore the inequalities  $S(A,B) \geq S(A)$  and  $S(A,B) \geq S(B)$  are not always true. On the contrary, inequality for conditional entropy still holds even in a quantum world

$$S(A)_\rho \geq S(A|B)_\rho, \quad (2.62)$$

where we define the conditional entropy as

$$S(A|B) = S(A,B) - S(B). \quad (2.63)$$

We notice that given the definition, it is possible for the conditional entropy to be negative. Consider again a pure bipartite state  $|\psi\rangle_{AB} = 1/\sqrt{2}(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B)$ . Since the state is a pure state  $|\psi\rangle_{AB}$ , its entropy is zero,  $S(A,B) = 0$ . This reflects our complete knowledge of the state as a whole. However, when considered independently, these qubit states are in a maximally mixed state, which means  $S(A) = S(B) = \log(d) = 1$  bit. From the definition (2.63), we get

$$S(A|B) = -S(B) = -1. \quad (2.64)$$

The negative conditional entropy of -1 bit means that the total uncertainty about both qubits (which is zero due to their entanglement) is less by one bit than the

uncertainty of knowing one qubit alone. This scenario illustrates that the knowledge of one qubit in an entangled pair provides a complete understanding of its partner, effectively reducing uncertainty beyond what exists in isolation—a phenomenon only possible due to quantum entanglement. In other words, knowing the state of system  $B$  not only provides information about system  $A$  but provides more information than is contained in the isolated system  $A$  alone.

Furthermore, since marginal entropies depend only on eigenvalues and from the Schmidt decomposition we know that these are the same. Hence, it also holds for pure bipartite states that

$$S(A) = S(B). \quad (2.65)$$

Because the composite system is in a pure state, we know exactly its state. But that does not mean we know everything about the individual states, which can be in maximally mixed state,  $\rho_A = \rho_B = \mathbb{I}/2$ , and thus we have less knowledge about them, which results in a negative conditional entropy. This is possible only in a quantum world. Actually, it can be proven that quantum conditional entropy is always negative for all pure entangled states [6].

The phenomena of negative conditional entropy is captured by coherent information  $I(A \rangle B)_\rho$ , and quantifies the rate of quantum correlation in a bipartite state

$$I(A \rangle B)_\rho = S(B)_\rho - S(A, B)_\rho = -S(A|B)_\rho. \quad (2.66)$$

It represents the amount of entropy (or uncertainty) reduced in subsystem  $B$  due to its correlations with subsystem  $A$ , relative to the total entropy of the combined system  $AB$ . Coherent information measures the "quantumness" of the information, specifically how much quantum information remains intact after being transmitted through a quantum channel. It quantifies the extent to which the entanglement and other quantum correlations between subsystems  $A$  and  $B$  are preserved despite the noise introduced by the environment.

Negative coherent information indicates that the quantum channel is dissipative or decohering, lowering quantum correlations and coherence. It can be proven that for a bipartite state transmitted through a quantum channel, the quantum correlations can only decrease [3].

Quantum mutual information, on the other hand, measures the total amount of correlations, both classical and quantum, between two subsystems of a quantum state. For a bipartite state  $\rho_{AB}$ , mutual information is defined as

$$I(A : B) = S(A) + S(B) - S(A, B) = S(A) - S(A|B) = S(B) - S(B|A). \quad (2.67)$$

Defining capacity of quantum channels is a bit more complicated than at the classical counterpart. There is no single numerical quantity to define capacity of transmitting information. Rather, quantum channels appear to have at least four different natural definitions of capacity, depending on the auxiliary resources allowed, the class of

protocols allowed, and whether the information to be transmitted is classical or quantum. Additionally, in quantum channels, single-use capacity is not equal to the asymptotic capacity. The asymptotic capacity describes the amount of information that can be reliably transmitted using the quantum channel infinitely many times. We will briefly show the four cases, as discussed by Peter Shor [8, 5].

The first case is the classical capacity of a quantum channel refers to the maximum rate at which classical information can be reliably transmitted over a quantum channel. In this case, the capacity is similar to the classical capacity

$$C(\mathcal{N}) = \max_{p_X(x)} [I(A : B)], \quad (2.68)$$

but with mutual information defined with von Neumann entropy (2.67). This capacity depends on the quantum channel being used and the encoding and decoding strategies. We also differentiate between channels with and without entanglement. If Alice and Bob share entanglement before communication begins, the capacity can potentially be higher.

The second case is private capacity, where we transmit classical information over quantum channel and we want to keep this information secret. It quantifies the rate at which information can be sent securely such that an eavesdropper (who might have access to the channel or its environment) gains almost no information about the transmitted message.

$$P(\mathcal{N}) = \max_{p_X(x)} [I(A : B) - I(A : E)] \quad (2.69)$$

The third case is a quantum capacity, which is the ability of a quantum channel to transmit quantum information. This capacity, denoted as  $Q$ , is defined based on the coherent information (2.66). The quantum capacity of a channel is defined as the maximum rate at which qubits can be transmitted such that the information can be recovered reliably at the output, despite the effects of noise and decoherence in quantum channels.

$$Q(\mathcal{N}) = \max I(A)B)_\rho \quad (2.70)$$

The final case is the entanglement-assisted classical capacity of a quantum channel. It is the rate at which classical information can be sent with the help of entanglement shared between the sender and receiver. This capacity typically exceeds the classical capacity because the pre-shared entanglement can be used to perform more efficient error correction and enhance the correlation between transmitted and received signals. This capacity reflects the maximal enhancement in communication rate and reliability achievable by utilizing the quantum resource of entanglement. It leverages the joint quantum state of the entangled pair to encode more information than could be done with classical states alone.

For any two density operators, we define relative entropy as

$$S(\rho||\sigma) = \text{Tr}(\rho \log(\rho)) - \text{Tr}(\rho \log(\sigma)). \quad (2.71)$$

As in the classical case, the relative entropy is non-negative and has no upper bound, i. e., can be infinite.

The upper limit on how much information can be obtained about the sequence of signals emitted by the source is quantified by accessible information. The amount of accessible information about a quantum state is given by the *Holevo bound* [3].

Suppose a random variable  $X$  that contains the information which state  $\rho_X$  from system  $\rho$  was emitted. This state is measured using fixed POVM elements  $\{E_0, \dots, E_m\}$  with measurement outcome  $Y$ . For such measurement it holds

$$I(X : Y) \leq S(X)_\rho - \sum_x p_x S(X)_{\rho_x}, \quad (2.72)$$

where  $\rho = \sum_x p_x \rho_x$ . This bound states the upper bound on the accessible information. It also says that the maximum information encoded in one qubit is one bit. This gives the capacity of a channel where at each time step the sender must choose one of the states  $\rho_X$  to send.

#### 2.5.4 One-Shot entropies

So far, all the definitions assumed that we can repeat the measurements independently and infinitely many times, which gives us the exact probabilities  $p(x)$ . However, in practical realization, we do not have such luxury, therefore we need to define min and max entropies, which take into account a finite number of repetitions of the measurements.

Min-entropy quantifies the uncertainty in a random experiment from the perspective of the "least uncertain" or most predictable outcome. Essentially, it measures the predictability of the most likely event. The definition of a classical min-entropy is hence

$$H_{min}(A) = -\log\left(\max_x(P(X = x))\right). \quad (2.73)$$

We can think of min-entropy as asking the question: "What is the highest probability with which I can predict an outcome?" For example when we return to our example with a die with 6 sides, but the die is weighted so that the side 6 comes up half the time, and the other five sides each have a 1/10 chance. The min-entropy here would focus solely on the probability of rolling a 6, as it's the most predictable outcome. Thus the min-entropy for this die would be  $H_{min} = -\log\left(\max_x(P(X = x))\right) = -\log(1/2) = 1$ .

Max-entropy, on the other hand, measures the maximum uncertainty in a single-shot experiment. It provides an upper bound on the entropy or uncertainty of a system. It's like saying, "Assuming the least we can know about a system, how uncertain can we be about its state?"

The real importance of one-shot entropies comes in to play in the quantum information theory, since it describes well the non-asymptotic scenarios. Von Neumann entropy serve well in idealized settings where states are assumed to be identically and independently distributed. However, in practical quantum mechanics, such scenarios are rare, and each quantum state often carries unique information that cannot be averaged across multiple systems and the information itself cannot be copied. This is where one-shot entropies, specifically min-entropy and max-entropy, become useful.

Remember that the conditional quantum entropy,  $S(A|B)$ , measures the uncertainty in a quantum system  $A$  when another quantum system  $B$  is known. For a joint state  $\rho_{AB}$ , and a reference state  $\sigma_B$  on  $B$ , the conditional entropy is expressed as:

$$S(\rho_{AB}|\sigma_B) = -\text{Tr} \left[ \rho_{AB} \left( \log(\rho_{AB}) - \log(\mathbb{I} \otimes \sigma_B) \right) \right], \quad (2.74)$$

where  $\sigma_B \in \mathcal{B}(\mathcal{H}_B)$  is a chosen state of the subsystem  $B$  used as a reference to calculate various entropy measures.

System  $B$  is assumed to be in this state  $\sigma_B$  and ideally,  $A$  would be in some state  $\rho_A$  such that  $\rho_{AB} = \rho_A \otimes \sigma_B$ . The conditional entropy  $S(\rho_{AB}|\sigma_B)$  measures how the joint state  $\rho_{AB}$  diverges from a simple, independent combination of  $A$  and  $B$  states  $\rho_A \otimes \sigma_B$ . In other words, this deviation means  $\rho_{AB}$  cannot be simply split into independent states of  $A$  and  $B$ , thus indicating entanglement between states  $A$  and  $B$ . The equation can be also written as

$$S(\rho_{AB}|\sigma_B) = S(\rho_{AB}) - S(\rho_B) - S(\rho_B||\sigma_B), \quad (2.75)$$

where  $\rho_B = \text{Tr}_A(\rho_{AB})$ . The reference state  $\sigma_B$  is used as a benchmark to quantify how much the actual state of  $B$ ,  $\rho_B$ , deviates from this reference state. This is maximal when  $\rho_B = \sigma_B$ , thus the relative entropy (measure of distinguishability between these two state)  $S(\rho_B||\sigma_B) = 0$ , and in this case  $S(A|B) = S(\rho_{AB}|\sigma_B)$ . As with the classical min-entropy, we are search for the most probable outcome. To write it down, we will make use of supremum

$$S(A|B) = \sup_{\sigma_B} H(\rho_{AB}|\sigma_B), \quad (2.76)$$

where the supremum iterates through all state  $\sigma_B \in \mathcal{B}(\mathcal{H})$ . We are looking for the highest conditional entropy of  $\rho_{AB}$  relative to all possible states  $\sigma_B$  on system  $B$ . By exploring all possible reference states  $\sigma_B$ , we capture the maximum uncertainty of system  $A$  given knowledge of system  $B$ . Practically, this can be used to assess the resilience of the system  $A$ 's information against any possible known configuration of  $B$ .

To adapt to one-shot scenarios in quantum settings, where the quantum state  $\rho_{AB}$  is available only once, we define the quantum min-entropy  $H_{min}(\rho_{AB}|\sigma_B)$  for a specific reference state  $\sigma_B$  as follows

$$H_{min}(\rho_{AB}|\sigma_B) = -\log \lambda, \quad (2.77)$$

where  $\lambda \in \mathbb{R}$  is defined as the smallest real number such that

$$\lambda \cdot \mathbb{I}_A \otimes \sigma_B - \rho_{AB} \geq 0. \quad (2.78)$$

The max-entropy of  $\rho_{AB}$  relative to  $\sigma_B$  is defined to capture the maximal uncertainty:

$$H_{max}(\rho_{AB}|\sigma_B) = \log[\text{Tr}((\mathbb{I} \otimes \sigma_B)\rho_{AB}^0)], \quad (2.79)$$

where  $\rho_{AB}^0$  is the projector onto the support of  $\rho_{AB}$  [9]. This measure looks at how large  $\rho_{AB}$  can extend when measured against  $\sigma_B$ , indicating the maximum spread or dispersion of the joint state with respect to  $B$ 's reference state  $\sigma_B$ .

To finish the definition of the min entropy, we define the conditional min-entropy of a state  $\rho_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$  conditioned on  $\mathcal{H}_B$  as

$$H_{min}(A|B) = \sup_{\sigma_B} S_{min}(\rho_{AB}|\sigma_B) = -\log\left(\min_{\sigma_B} \lambda\right), \quad (2.80)$$

in which the supremum and the minimum are taken over all states  $\sigma_B$  in  $\mathcal{B}(\mathcal{H}_B)$ . This represents the minimal scaling factor  $\lambda$  that, when applied to  $\sigma_B$  (extended to the entire system), can still "cover" the state  $\rho_{AB}$ . The min-entropy thus provides the amount of uniform randomness extractable from a classical random variable correlated with the quantum system  $B$ , ensuring independence from  $B$ 's specific quantum state [6]. The conditional min-entropy is directly related to the maximum achievable overlap with a maximally entangled state if only local actions on the B-part are allowed [10].

The conditional max-entropy is defined similarly but focuses on maximizing the overlap:

$$H_{max}(A|B) = \sup_{\sigma_B} H_{max}(\rho_{AB}|\sigma_B) = \max_{\sigma_B} \log \|\sqrt{\rho_{AB}}\sqrt{\mathbb{I}_A \otimes \sigma_B}\|^2. \quad (2.81)$$

This expression determines how large the quantum system  $A$  can be theoretically compressed while still ensuring that it can be completely reconstructed given access to quantum system  $B$  [6]. The max-entropy measures the highest fidelity of  $\rho_{AB}$  with any product state that remains entirely mixed over  $A$ [10].

The smooth min-entropy and smooth max-entropy extend the concepts of min-entropy and max-entropy by incorporating a tolerance for small deviations in the state, represented by the smoothness parameter  $\varepsilon$ . These entropies are defined within the  $\varepsilon$ -neighborhood of a given quantum state  $\rho$ , and measure the uncertainty in a quantum state  $\rho$  under this slight perturbations. This makes them particularly relevant for real-world applications where ideal conditions are not always met.

The smooth min-entropy,  $H_{min}^\varepsilon(A|B)$ , is defined as:

$$H_{min}^\varepsilon(A|B) = \sup_{\rho' \in \mathcal{B}^\varepsilon(\rho)} H_{min}(A|B)_{\rho'}, \quad (2.82)$$



where  $\mathcal{B}^\varepsilon(\rho)$  represents the set of all quantum states  $\rho'$  that are within an  $\varepsilon$ -proximity to  $\rho$ . This definition captures the least uncertainty (minimal entropy) about system  $A$  given system  $B$ , across all states close to  $\rho$  within the specified neighborhood.

Conversely, the smooth max-entropy,  $H_{max}^\varepsilon(A|B)$ , is defined as:

$$H_{max}^\varepsilon(A|B) = \inf_{\rho' \in \mathcal{B}^\varepsilon(\rho)} H_{max}(A|B)_{\rho'}, \quad (2.83)$$

This measures the maximal uncertainty (maximum entropy) under the same conditions, ensuring that the measurement of entropy considers the worst-case scenario within the  $\varepsilon$ -neighborhood.

The significance of smooth entropies extends to their asymptotic behavior, particularly how they converge to the von Neumann entropy in the limit of many copies of the state and for  $\varepsilon$  approaching 0. These limits are expressed as:

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} H_{min}^\varepsilon(A^n|B^n)_{\rho^{\otimes n}} = S(A|B)_\rho, \quad (2.84)$$

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} H_{max}^\varepsilon(A^n|B^n)_{\rho^{\otimes n}} = S(A|B)_\rho \quad (2.85)$$

where  $\rho^{\otimes n}$  indicates  $n$  independent copies of  $\rho$ . These expressions demonstrate that as the number of systems increases and the smoothness parameter decreases, the smooth entropies converge to the von Neumann entropy of the system, illustrating the foundational relation between practical, finite-resource measures of entropy and their idealized, infinite-resource counterparts [10].

### 2.5.5 Entropic uncertainty principle

The notorious Heisenberg's uncertainty principle  $\Delta X \Delta P \geq \hbar/2$  is a fundamental formula in quantum mechanics. However, with quantum entanglement and quantum memory, one is able to predict with certainty what he or she will measure. In other words, Heisenberg does not provide lower bound on uncertainties for entangled state, where the uncertainty depends on the rate of entanglement between particles and quantum memory [11]. Additionally, in quantum information theory, we would like to have the uncertainty in terms of entropies for better quantification of uncertainty and better description of transmitting information through a noisy quantum channel.

To demonstrate how entanglement can beat the uncertainty principle, we consider a scenario, where Bob has a source of qubits. He sends one qubit to Alice and he wants to predict her measurement. They agree on two possible measurements A, B. Bob wants to guess with lowest possible uncertainty. Alice measures using either A or B and announces her choice to Bob. Bob's uncertainty is described using a classical Shannon entropy because all the information he has is classical, e. g., he knows the

density matrix of the qubit. Then the uncertainty is described as

$$H(A) + H(B) \geq \log \frac{1}{c}, \quad (2.86)$$

where  $c$  is the complementarity of the observables

$$c = \max_{j,k} \|\langle \psi_j | \varphi_k \rangle\|^2, \quad (2.87)$$

where  $\psi_j$  are eigenstates of A and  $\varphi_k$  are eigenstates of B. However, with quantum memory, Bob can overcome this bound, as long as the qubit in his quantum memory is entangled with Alice's qubit. Suppose we have a maximally entangled state in possession of Bob

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B) = \frac{1}{\sqrt{2}}(|+\rangle_A |+\rangle_B + |-\rangle_A |-\rangle_B), \quad (2.88)$$

and he distribute qubit A to Alice and keeps qubit B in his quantum memory. Bob saves his qubit into a quantum memory while Alice measures her qubit. She chooses between two POVMs, in this case the POVM set for measuring in the  $Z$  basis and the POVM set for  $X$  on her system A

$$N_A^z = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}, \quad (2.89)$$

$$M_A^x = \{|+\rangle\langle +|, |-\rangle\langle -|\}, \quad (2.90)$$

where  $|\pm\rangle = 1/\sqrt{2}(|0\rangle \pm |1\rangle)$ . She announces her choice to Bob. If she chooses the  $X$  basis with POVM  $\{M_A^x\}$ , the bipartite state will have the following form

$$\sigma_{AB} = \sum_x |x\rangle_A \langle x| \otimes \text{Tr}((M_A^x \otimes \mathbb{I}_B)\rho_{AB}), \quad (2.91)$$

and the Bob's uncertainty will be the quantum conditional entropy  $S(X|B)_\sigma$ . Analogously when Alice chooses the  $Z$  basis, the uncertainty Bob encounters is equal to  $S(Z|B)$  and the current state is

$$\tau_{AB} = \sum_z |z\rangle_A \langle z| \otimes \text{Tr}((N_A^z \otimes \mathbb{I}_B)\rho_{AB}), \quad (2.92)$$

where  $\{|x\rangle_A\}$  and  $\{|z\rangle_A\}$  are orthonormal states of classical register  $X$  and  $Z$ , respectively. Bob's total uncertainty is the sum of  $S(X|B) + S(Z|B)$ . For POVMs, we can define the quantity of their incompatibility  $c$  as

$$c = \max_{j,k} \|\sqrt{M_A^x} \sqrt{N_A^z}\|_\infty^2, \quad (2.93)$$

where  $\|\cdot\|_\infty$  is the infinity norm of an operator, which is simply the largest eigenvalue in the finite-dimensional case [6]. The uncertainty with which Bob measures the bipartite state is generally

$$S(X|B)_\sigma + S(Z|B)_\tau \geq \log \frac{1}{c} + S(A|B), \quad (2.94)$$

where  $S(A|B)$  is negative if the state  $\rho_{AB}$  is entangled. Thus we have a lower bound as compared to the equation (2.86). Particularly, in our case  $c = 1/2$ , thus  $\log(1/c) = \log(2) = 1$  and  $S(A|B)_\Psi = -1$ , hence

$$H(X|B)_\sigma + H(Z|B)_\tau \geq 0, \quad (2.95)$$

which implies that Bob knows the measurement outcome in both cases with certainty. However,  $X$  and  $Z$  are incompatible and thus it should not be possible for Bob to measure the outcome precisely. This theorem can also be formulated using min, and max-entropy as [6]

$$H_{min}^\varepsilon(X|E) + H_{max}^\varepsilon(Z|B) \geq \log\left(\frac{1}{c}\right), \quad (2.96)$$

for  $\rho_{ABE} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B) \otimes \mathcal{H}_E$ . This formula is useful for quantum cryptography to determine the lower bounds of the amount of information an eavesdropper  $E$  has on the outcome of Alice's measurement  $X$ .

### 3 Quantum Key Distribution Protocols

Quantum key distribution (QKD) protocols consist of two main stages - quantum transmission and classical post-processing. In this quantum transmission stage, Alice and Bob communicate over a quantum channel using a QKD protocol to generate a raw bit key. This is followed by key sifting to produce a symmetric key for Alice and Bob. However, due to imperfections and potential eavesdroppers, they need to initiate a second part of the QKD protocol. In this purely classical post-processing stage, they first estimate the error in their sifted key (parameter estimation), and finally, they improve the security of the key through privacy amplification.

QKD refers to a symmetric key exchange protocol that utilizes quantum information carriers, primarily, but not limited to, photons. While any form of qubit or qudit [12] can be used to perform QKD, photons offer a significant advantage. They interact weakly with the environment, resulting in long coherence times, allowing them to be transmitted over tens of kilometers through optical fibers or even hundreds of kilometers through free space. Although QKD can be performed with other systems, such as superconducting qubits using quantum computers, these qubits have relatively short coherence times of around tens to hundreds of microseconds, making long-distance transmission infeasible.

Pioneered by Bennett and Brassard in 1984, quantum key distribution quickly emerged as a highly discussed topic. Since then, numerous publications have been written, focusing on new protocols, theoretical analysis of security, and even a development of a new field called Quantum Hacking arose [13, 14].

The quantum key exchange is proven to be theoretically secure because any eavesdropping on the quantum channel can be detected, resulting in the protocol to be aborted and restarted. In general, QKD can be divided into two categories. In the first category, called discrete variable QKD (DV-QKD), utilizes single photons as information carriers. The second category, known as continuous variable QKD (CV-QKD), employs quantum states of light such as squeezed states and Gaussian states to encode information. DV-QKD gave the origin of QKD when in the year 1984, C. H. Bennet and G. Brassard published their famous paper describing the DV-QKD protocol called BB84 [15]. While DV-QKD was proven to work and is already commercially available, it has several disadvantages, for instance, high losses and limited reach due to the transmission of single photons. Some of the problems associated with DV-QKD can be addressed by CV-QKD, which is currently under active development.

In this chapter, we will focus on the fundamentals and various protocols of Discrete Variable Quantum Key Distribution. We will explore key concepts such as the no-cloning theorem and the use of Mach-Zehnder interferometers in quantum cryptography, specifically examining protocols like BB84 and B92 with emphasis on time and phase encoding. The discussion will also cover crucial post-processing

steps including parameter estimation, error correction, and privacy amplification. This focus is chosen due to our experimental work with DV-QKD. The study of CV-QKD will not be covered in this thesis as per the thesis assignments.

### 3.1 Fundamental principles

In this chapter we introduce the *no-cloning* theorem which is a fundamental part in understanding the security of QKD. The next section deals with a quantum description of a Mach-Zehnder interferometer, which is substantial for the time and phase encoding protocol, discussed later.

#### 3.1.1 No-cloning theorem

Quantum Key Distribution achieves theoretical security through fundamental principles of quantum mechanics. Two main aspects that ensure this security are the quantum measurement process, which causes a collapse of superposition into one of the possible states, and the requirement that all quantum operations must be unitary. From these principles, the no-cloning theorem emerges as a cornerstone of quantum security.

The no-cloning theorem[3] states that it is impossible to create an exact copy of an arbitrary unknown quantum state. This concept can be understood through the following quantum mechanical analysis. Consider a general quantum state  $|\psi\rangle$ , which we aim to copy using a quantum operation  $U$  onto a state  $|s\rangle$ . The cloning process can be expressed as

$$|\psi\rangle \otimes |s\rangle \xrightarrow{U} U(|\psi\rangle \otimes |s\rangle) = |\psi\rangle \otimes |\psi\rangle. \quad (3.1)$$

Now consider another arbitrary state  $|\varphi\rangle$ , onto which we also apply the copy operator

$$U(|\psi\rangle \otimes |s\rangle) = |\psi\rangle \otimes |\psi\rangle, \quad (3.2a)$$

$$U(|\varphi\rangle \otimes |s\rangle) = |\varphi\rangle \otimes |\varphi\rangle. \quad (3.2b)$$

We now take the scalar product of these two equations

$$\langle\psi|\varphi\rangle = (\langle\psi|\varphi\rangle)^2, \quad (3.3)$$

which has only 2 solutions. Either the two states are the same  $|\psi\rangle = |\varphi\rangle$  or orthogonal  $\langle\psi|\varphi\rangle = 0$ . From this result, it follows that only orthogonal states can be cloned. Therefore together with the quantum nature of destructive measurements, any transmitted quantum information cannot be effectively eavesdropped without raising attention. Any eavesdropping attempt will be detectable as an increase of

the error in the sifted key between Alice and Bob. This theorem is fundamental to quantum cryptography, however, it is also very limiting in other fields of quantum computation. For instance, unlike classical optical networks, where signals can be amplified, quantum signals (single photons) cannot be amplified without altering their quantum state due to the no-cloning theorem. This limitation complicates the development of long-range quantum networks with quantum repeaters.

Another added challenge resulting from no-cloning theorem is the development of quantum memory. As we have learnt, we cannot make multiple copies to preserve the quantum information. Combined with the typically short coherence times of qubits, this makes the storage and preservation of quantum information difficult.

An example of copying quantum information with two orthogonal states is a CNOT gate. A CNOT gate is a quantum logic gate that flips the target qubit's state if the control qubit is in the state  $|1\rangle$ . Assume  $|s\rangle = |0\rangle$  and  $|\psi\rangle = |1\rangle$ . Then we can easily copy  $|\psi\rangle$  into  $|s\rangle$  using CNOT gate, where control qubit is  $|\psi\rangle$

$$|10\rangle \xrightarrow{\text{CNOT}} |11\rangle. \quad (3.4)$$

Here, both qubits retain their states, which shows that under specific conditions (like identical or orthogonal states), we can make copies of quantum states.

While quantum states cannot be cloned, quantum information can be transferred via a protocol called quantum teleportation, where the state of one qubit is transferred to another through entanglement and classical communication, albeit at the cost of destroying the original state. This phenomenon demonstrates the principle that quantum information is preservable, even if the physical qubits are not.

In a scenario, where Alice prepares a state  $|\psi\rangle$  with the probability  $p$  and a state  $|\varphi\rangle$  with the probability  $1 - p$ , information needed to know about this state, in order to be able to copy it, is given by  $S(A)$ . The no-cloning theorem can be seen as a fact, that Bob's accessible information defined by Holevo's bound (2.72) is lower than the information needed to clone the state. In general, in quantum mechanics, the accessible information for two non-orthogonal states is less than the entropy of preparation [3].

To summarize, the no-cloning theorem is a fundamental result in quantum information theory that states the impossibility of creating an identical copy of an arbitrary unknown quantum state. This theorem has important implications for quantum computing and quantum communication, ensuring the security of quantum cryptography. While certain specific states, like orthogonal or identical states, can be copied under controlled conditions using mechanisms like the CNOT gate, the general prohibition against cloning arbitrary states preserves the integrity of quantum information.

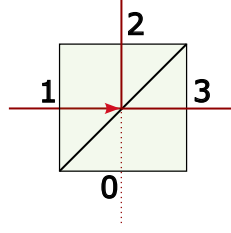


Figure 3.1: Beam splitter with input ports 0, 1 and output ports 2, 3.

### 3.1.2 Quantum description of Mach-Zehnder interferometer

To properly understand how phase encoding of quantum information into photons works, we need to describe a beam splitter and a Mach-Zehnder interferometer (MZI), as well as how two Mach-Zehnder interferometers can be connected using quantum mechanics. We will begin with a description of a beam splitter, as shown in Figure 3.1.

Consider an input state  $|\psi\rangle = |1_0 0_1\rangle$ , where the subscript denotes the port number of the beam splitter. The output from a lossless beam splitter with transmission  $t$  and reflection  $r$  coefficients can be described using annihilation and creation operators as follows [16]:

$$\hat{a}_3 = t\hat{a}_1 + r\hat{a}_0, \quad (3.5a)$$

$$\hat{a}_2 = r\hat{a}_1 + t\hat{a}_0, \quad (3.5b)$$

where  $\hat{a}_0, \hat{a}_1$  are the input field operators and  $\hat{a}_2, \hat{a}_3$  are the output field operators. These operators obey the commutation relations  $[\hat{a}_k, \hat{a}_l] = [\hat{a}_k^\dagger, \hat{a}_l^\dagger] = 0$  and  $[\hat{a}_k, \hat{a}_l^\dagger] = \delta_{kl}$ , for  $k, l \in \{0,1\}$ . The conditions on the coefficients  $r$  and  $t$  are given by

$$|t|^2 + |r|^2 = 1, \quad (3.6a)$$

$$rt^* + r^*t = 0. \quad (3.6b)$$

For a 50:50 beam splitter, these coefficients are  $t = 1/\sqrt{2}$  and  $r = i/\sqrt{2}$ , where  $i$  denotes the phase change caused by reflection. For a correct description that fulfills commutation relations, we also include vacuum states  $|0\rangle$  in input 0. Thus, with the input state  $|\psi\rangle = |0_0 1_1\rangle$ , we obtain

$$|0\rangle_0 |1\rangle_1 \xrightarrow{BS} \hat{a}_1^\dagger |0\rangle_2 |0\rangle_3 = \frac{1}{\sqrt{2}}(\hat{a}_3^\dagger + i\hat{a}_2^\dagger) |0\rangle_2 |0\rangle_3 = \frac{1}{\sqrt{2}}(|1\rangle_2 |0\rangle_3 + |0\rangle_2 |1\rangle_3). \quad (3.7)$$

This operation results in an entangled state. For a general BS, the output state would be

$$|\psi_{out}\rangle = r |1\rangle_2 |0\rangle_3 + t |0\rangle_2 |1\rangle_3. \quad (3.8)$$

The probability of coincident detection is given by

$$P_c = \langle \psi_{out} | \hat{a}_2^\dagger \hat{a}_3^\dagger \hat{a}_3 \hat{a}_2 | \psi_{out} \rangle = 0, \quad (3.9)$$

which confirms our expectations because only one photon entered the beam splitter. Applying a second beam splitter to  $|\psi_{out}\rangle$  results in

$$|\psi_{out}\rangle \xrightarrow{BS} 2rt |1\rangle_4 |0\rangle_5 + (r^2 + t^2) |0\rangle_4 |1\rangle_5. \quad (3.10)$$

For a 50:50 beam splitter, the output state simplifies to  $|1\rangle_2 |0\rangle_3$ , indicating a non-entangled state with a deterministic output port. This is typical for a symmetric Mach-Zehnder interferometer (MZI with both arms of the same length), depicted in Figure 3.2, without any delays or phase modifications. Introducing a delay  $\varphi$  modifies the operators  $\hat{a}_0^\dagger$  and  $\hat{a}_1^\dagger$  to  $\hat{a}_4^\dagger$  and  $\hat{a}_5^\dagger$ , using sine and cosine functions as follows

$$\hat{a}_0^\dagger = -\sin(\varphi/2)\hat{a}_4^\dagger + \cos(\varphi/2)\hat{a}_5^\dagger, \quad (3.11a)$$

$$\hat{a}_1^\dagger = \cos(\varphi/2)\hat{a}_4^\dagger + \sin(\varphi/2)\hat{a}_5^\dagger. \quad (3.11b)$$

Passing the state  $|\psi\rangle = |0_0 1_1\rangle$  through the Mach-Zehnder interferometer yields

$$|\psi_{out}\rangle = \cos(\varphi/2) |1\rangle_4 |0\rangle_5 + \sin(\varphi/2) |0\rangle_4 |1\rangle_5. \quad (3.12)$$

The detection probabilities are calculated as

$$P_1 = |\langle 1_4 0_5 | \psi_{out} \rangle|^2 = \cos^2(\varphi/2) = \frac{1}{2}(1 + \cos(\varphi)), \quad (3.13a)$$

$$P_2 = |\langle 0_4 1_5 | \psi_{out} \rangle|^2 = \sin^2(\varphi/2) = \frac{1}{2}(1 - \cos(\varphi)). \quad (3.13b)$$

The probability of detection at either detector 1 or detector 2 depends on the cosine of the delay  $\varphi$ . Moreover, the detection probabilities sum up to one:  $P_1 + P_2 = 1$ .

Now consider a double MZI setup with three 50:50 beam splitters, as depicted in Figure 3.3. In this configuration, the creation operators become more complex compared to the single MZI setup

$$\hat{a}_0^\dagger = \frac{-i \sin\left(\frac{\varphi_2}{2}\right) - e^{-i\varphi_1} \cos\left(\frac{\varphi_2}{2}\right)}{\sqrt{2}} \hat{a}_6^\dagger + \frac{i \cos\left(\frac{\varphi_2}{2}\right) - e^{-i\varphi_1} \sin\left(\frac{\varphi_2}{2}\right)}{\sqrt{2}} \hat{a}_7^\dagger, \quad (3.14a)$$

$$\hat{a}_1^\dagger = \frac{\sin\left(\frac{\varphi_2}{2}\right) + i e^{-i\varphi_1} \cos\left(\frac{\varphi_2}{2}\right)}{\sqrt{2}} \hat{a}_6^\dagger + \frac{-\cos\left(\frac{\varphi_2}{2}\right) + i e^{-i\varphi_1} \sin\left(\frac{\varphi_2}{2}\right)}{\sqrt{2}} \hat{a}_7^\dagger. \quad (3.14b)$$

The probability of detection at detectors  $D_3$  and  $D_4$  depends on the product of two sines, each with a respective delay in the argument:

$$P_3 = |\langle 1_6 0_7 | \psi_{out} \rangle|^2 = \frac{1}{2}(1 + \sin(\varphi_1) \sin(\varphi_2)), \quad (3.15a)$$



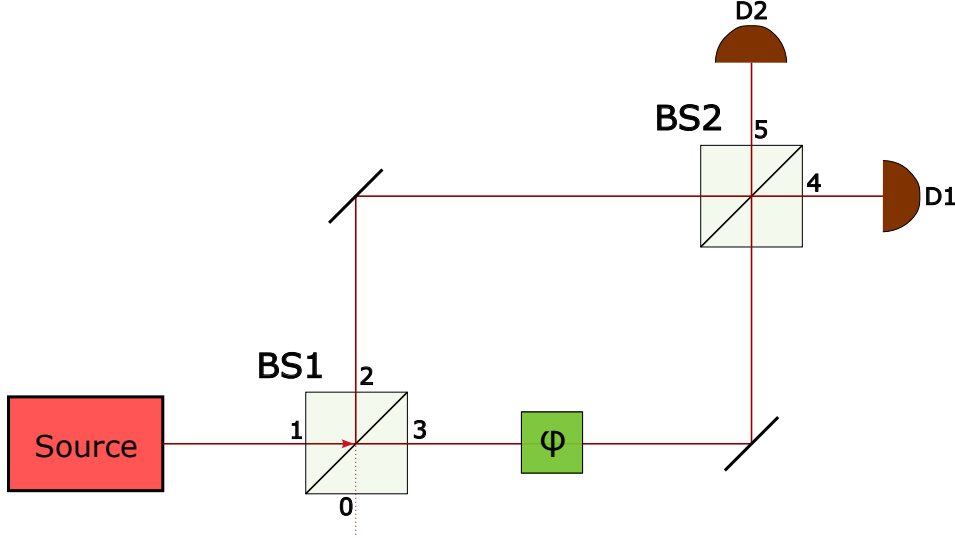


Figure 3.2: A Mach-Zehnder interferometer consisting of a laser source, two beam splitters (BS1 and BS2), a phase delay  $\varphi$ , and two detectors (D1 and D2) at the outputs from BS2. For a free-space MZI where the laser light is transmitted through air, there would be additionally two mirrors to direct the light from BS1 to BS2. Alternatively, optical fibers can be used without any mirrors.

$$P_4 = |\langle 0_6 1_7 | \psi_{out} \rangle|^2 = \frac{1}{2}(1 - \sin(\varphi_1) \sin(\varphi_2)). \quad (3.15b)$$

When the phase delay  $\varphi_1$  is set to 0 or a multiple of  $\pi$ , resulting in  $\sin(\varphi_1) = 0$ , the probabilities  $P_3$  and  $P_4$  both equal  $1/2$ , independent of the value of  $\varphi_2$ . Similarly, if  $\varphi_2$  is set such that  $\sin(\varphi_2) = 0$ , then  $P_3$  and  $P_4$  no longer depend on  $\varphi_1$ .

To understand why  $P_3$  and  $P_4$  do not depend on  $\varphi_2$  when  $\sin(\varphi_1) = 0$ , we need to look at the state evolution, especially at the second beam splitter BS2. Substituting  $\varphi = 0$  into the Equation (3.12) shows that the state after BS2 is  $|1\rangle_4 |0\rangle_5$ . This means the single photon always takes the same path (port 4) into the second MZI. Therefore, the phase delay  $\varphi_2$  in that interferometer becomes irrelevant - the photon never experiences the other path where  $\varphi_2$  is introduced. This is why  $P_3$  and  $P_4$  are both  $1/2$  regardless of  $\varphi_2$  value when  $\sin(\varphi_1) = 0$  [16].

In our experiment, we are utilizing two physically separate asymmetrical Mach-Zehnder interferometers, one at Alice's side and one at Bob's. At first glance, it may seem that this system will behave as two independent MZIs. However, due to their different arm length, the two asymmetric MZIs can be understood that they not act as isolated interferometers but as one symmetrical MZI in special case, when the photon travels once in the longer arm and once in the shorter arm. The phase delays  $\varphi_1$  and  $\varphi_2$  introduced by Alice and Bob in their respective MZIs are not independent variables in determining the overall output detection probabilities  $P_3$  and  $P_4$ . More details will be described in the 3.2.3 section.

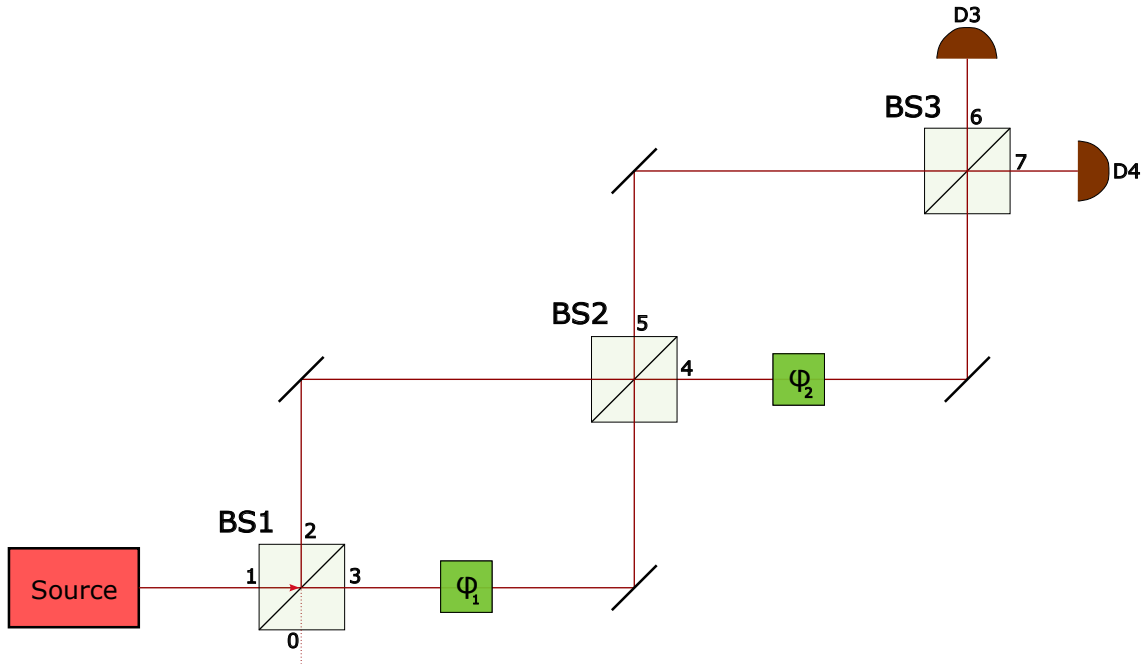


Figure 3.3: Diagram of two connected Mach-Zehnder interferometers, consisting of three beam splitters and two phase delays  $\varphi_1$  and  $\varphi_2$ .

### 3.2 Discrete variable QKD

In DV-QKD, we use single photons. The quantum information is encoded either in polarization or in phase and time together. Polarization encoding has a huge advantage in the simplicity of the setup. We need only a source of photons and polarization filters, but polarization is not generally preserved. Even with polarization maintaining fibers, polarization is not perfectly preserved for long distances. Also, the main goal of QKD is to use the already developed infrastructure of optical fibers, which are single-mode fibers and does not maintain polarization. The second approach is to use phase and time to encode information. This is done using phase modulators and asymmetric Mach-Zehnder interferometers. By using phase and time encoding, we can utilize the already deployed optical fibers for internet communication, but we need a more expensive and more complex setup than for polarization encoding.

In DV-QKD, single photons are utilized to encode quantum information, either through polarization or through a combination of phase and time. Polarization encoding offers significant advantages due to the simplicity of the setup, requiring only a photon source and polarization filters. However, polarization is not perfectly preserved over long distances, even with polarization-maintaining fibers. Furthermore, the primary goal of QKD is to utilize the already developed infrastructure of optical fibers, which are typically single-mode fibers that do not maintain polarization.

The alternative method involves using phase and time encoding, which is implemented using phase modulators and asymmetric Mach-Zehnder interferometers. Al-

though this approach allows for the use of existing optical fibers, it requires a more complex setup compared to polarization encoding.

To distribute the photons, Alice or Bob generates single photons and transmits them, as in the method originally designed by Bennett and Brassard. Alternatively, entangled photons can be generated, potentially even by a third party such as Eve. While the security of the BB84 protocol, introduced by Bennett and Brassard [15], relies on the no-cloning theorem and the uncertainty principle, entanglement-based protocols depend on maximally entangled states, which would be disrupted by any eavesdropping. Alice and Bob can verify the integrity of the entanglement using Bell's inequality.

There are two possible QKD setup layouts. In the *One way* setup, Alice owns a single photon generator and Bob possesses detectors. Alice generates a photon, encodes information, and sends it to Bob, who then processes and detects it. The photon travels a single path from Alice to Bob.

The second option is the *Plug and Play* setup, where Bob controls both the photon source and detectors. He generates a photon and sends it to Alice, who encodes the information and reflects the photon back to Bob using a Faraday mirror. This setup is advantageous for scalability and cost-efficiency in a star network topology. In this configuration, the central node, possessing the photon detectors and an asymmetric interferometer, can manage multiple nodes equipped only with phase modulators and Faraday mirrors. While this reduces costs as only the central node requires expensive detectors, it necessitates that all communications pass through the central node, requiring a high level of trust in the node operator.

### 3.2.1 BB84 protocol

The BB84 protocol, first proposed by C. H. Bennett and G. Brassard in 1984 [15], laid the groundwork for quantum key distribution (QKD) by demonstrating how unconditional security could be achieved during symmetric key exchanges. This security is grounded in the uncertainty principle of quantum mechanics, which ensures that any eavesdropping attempts can be detected and the reinitiate key transmission.

The core principle of BB84 is that two non-orthogonal quantum states cannot be precisely distinguished without knowledge of their preparation. Measuring a quantum state cause it to collapse into one of its superposition's basis states, thereby altering the original state. Furthermore, as detailed in section 3.1.1, non-orthogonal states cannot be cloned, ensuring that any intercepted state cannot be recovered by measurement and then re-transmitted or relayed without detection.

In the protocol, we have three parties: Alice and Bob, who are trying to establish a shared secret key, along with a potential eavesdropper, Eve. For the unconditional security, we assume that Eve is limited only by the laws of physics and has all

possible technology available. These three parties engage in the following process: Alice selects two non-orthogonal bases,  $X$  and  $Z$ , defined as:

$$|Z\rangle: \{|0\rangle, |1\rangle\}, \quad (3.16a)$$

$$|X\rangle: \{|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\}. \quad (3.16b)$$

Alice assigns one bit to two non-orthogonal state. She denotes the value 0 to states  $|0\rangle$  and  $|+\rangle$ , and bit value 1 to states  $|1\rangle$  and  $|-\rangle$ . She then randomly sends one of these four qubits to Bob via a quantum channel. Bob, upon receiving each qubit, randomly chooses to measure it in either the  $X$  or  $Z$  basis. Both parties record the basis used for each qubit. This record is called an alphabet.

After transmitting all qubits, Alice publicly shares her alphabet (the bases she used) through a classical channel. This information becomes available to any observer, including Eve. Bob then compares part of his alphabet with Alice's and discards any results where different bases were used, a process known as *key sifting*. In this process, statistically, Bob discards 50% of his measurements. The remaining bits forms the sifted key. All the possible combinations are summarized in Table 1.

However, in real-world scenarios, noise and measurement errors can alter some of the qubits. Consequently, Alice's and Bob's keys may differ by several bits. This discrepancy is quantified by the *Quantum Bit Error Rate* (QBER), defined as the probability that a bit in Bob's sifted string differs from the corresponding bit in Alice's [17]. To determine QBER, Alice and Bob publicly compare a subset of their sifted keys. If the QBER is higher than the acceptable threshold, it indicates either excessive noise in the quantum channel or potential eavesdropping by Eve.

Alice's Basis	Qubit Sent	Bob's Basis	Outcome	Keep/Discard
$Z$	$ 0\rangle$	$Z$	$ 0\rangle$	Keep
$Z$	$ 0\rangle$	$X$	$ +\rangle$ or $ -\rangle$	Discard
$Z$	$ 1\rangle$	$Z$	$ 1\rangle$	Keep
$Z$	$ 1\rangle$	$X$	$ +\rangle$ or $ -\rangle$	Discard
$X$	$ +\rangle$	$X$	$ +\rangle$	Keep
$X$	$ +\rangle$	$Z$	$ 0\rangle$ or $ 1\rangle$	Discard
$X$	$ -\rangle$	$X$	$ -\rangle$	Keep
$X$	$ -\rangle$	$Z$	$ 0\rangle$ or $ 1\rangle$	Discard

Table 1: BB84 Protocol: Combinations of Alice's and Bob's Bases and Qubits

The mutual information  $I_M$  between Alice and Bob, considering the sifting process,

can be calculated using the binary entropy function  $H$  introduced in (2.44) as follows:

$$I_M = 1 - H(\text{QBER}) = 1 - (-\text{QBER} \cdot \log_2(\text{QBER}) - (1 - \text{QBER}) \cdot \log_2(1 - \text{QBER})). \quad (3.17)$$

Here,  $I_M$  is maximal ( $I_M = 1$ ) when  $\text{QBER} = 0$ , indicating perfect agreement between Alice's and Bob's keys. Conversely,  $I_M = 0$  for  $\text{QBER} = 50\%$ , signifying completely random sifted keys with no mutual information. To include the sifting process, we need to divide  $I_M$  by  $1/2$ .

Eve's simplest eavesdropping strategy, known as the *Intercept and Resend* attack, is detailed in section 4.4.1. In this strategy, Eve intercepts the quantum channel and measures all qubits sent by Alice, attempting to guess the encoding basis. To correctly infer a qubit's state, she must choose the same basis that Alice used. However, since Eve has a 50% chance of choosing the wrong basis. In this case, the measurement is purely random and she will thus get guess the correct state in 50% of these case. Consequently, this introduces an average Quantum Bit Error Rate (QBER) of 25%. If the QBER calculated by Alice and Bob exceeds this threshold, they cannot trust their sifted key, as they cannot ascertain whether the high QBER is due to channel noise or Eve's eavesdropping attempts. Thus, they must suspect the interception by Eve.

Eve can enhance her strategy by storing the qubits in quantum memory until Alice publicly reveals her basis, at which point Eve measures all her intercepted qubits in the correct basis. While the quantum transmission between Alice and Bob is underway, Eve continues to resend qubits to Bob, but now she generates them entirely at random. This approach increases the QBER between Alice and Bob, yet it allows Eve to know correctly the value of all qubits transmitted by Alice. If Alice and Bob then correct their noisy key and decide to use it, Eve will have knowledge of the entire key. This attack scenario, which utilizes currently unavailable but theoretically possible technology, demonstrates that ensuring security requires considering all potential future developments in quantum technologies.

We will now demonstrate this protocol with a simple example using a few bits. Suppose Alice wants to send a bit value of 0; she thus sends the state  $|0\rangle$ , which belongs to the  $Z$  basis, and records this decision. Bob, also choosing the  $Z$  basis by chance, measures the state  $|0\rangle$ , correctly inferring the bit value as 0. Alice's bit string is now  $b_A = 0$ , and her chosen basis is  $\alpha_A = Z$ . Bob's bit string is  $b_B = 0$ , and his basis is  $\alpha_B = Z$ . This bit is later confirmed as valid because they both used the same basis and obtained the same outcome.

Next, Alice sends the bit 0 again but encoded in the  $X$  basis as the state  $|+\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle)$ . Bob, measuring in the  $Z$  basis, has a 50% chance of measuring either  $|0\rangle$  or  $|1\rangle$ . He measures  $|0\rangle$  and records a bit value of 0. Although they both ended with the same bit value, this bit is marked as invalid because it was measured in different bases. Alice's records now show  $b_A = 00$ ,  $\alpha_A = ZX$ , and Bob's are  $b_B = 00$ ,

$$\alpha_B = ZZ.$$

Finally, Alice opts to transmit the bit 1 in the  $X$  basis, sending the state  $|-\rangle$ . Bob measures in the  $X$  basis and incorrectly records  $|+\rangle$ , thus noting bit 0. Their raw keys are now  $b_A = 001$ ,  $b_B = 000$ , and their bases are  $\alpha_A = ZXX$ ,  $\alpha_B = ZZX$ .

After transmitting seven more qubits, their final keys and bases are:

$$b_A = 0011110011, \quad b_B = 0001111011, \quad (3.18a)$$

$$\alpha_A = ZXXZXXZXXZ, \quad \alpha_B = ZZXZXZZXZX. \quad (3.18b)$$

They stop the transmission and start the key sifting phase. Alice publicly shares her basis sequence  $\alpha_A$  and compares it with Bob's. For the first qubit, they used the same basis thus they keep this bit. Since the second bit was measured in different bases, it is discarded. The third bit, measured in matching bases, is kept even though they are unaware that their measurement results differed. Fast forward, their sifted alphabet after discarding mismatched measurements will be:

$$\alpha = Z_1X_3Z_4X_5Z_7X_8, \quad (3.19)$$

where the subscript denotes the sequence order of each qubit. For QBER estimation, they randomly select qubits 3, 4, and 7 and publicly compare their values. Alice's bits are 110, and Bob's are 010. They find that only the second and third bits match, leading to an estimated QBER of 33%. Given this high QBER, Alice and Bob would typically abort the protocol and restart, suspecting potential eavesdropping or unacceptable noise levels. However, assuming they continue, the final key, derived from the remaining bits 1, 5, and 8, would be 011, forming their shared secret for initiating secure classical communication. Naturally, for practical purposes, this key would need to be much longer than three bits.

In the original publication, Bennett and Brassard used photon polarization as a physical realization of qubits. Their set of non-orthogonal states consisted of  $|0^\circ\rangle = |0\rangle$ ,  $|90^\circ\rangle = |1\rangle$ , which is our  $Z$  basis, and  $|45^\circ\rangle = |+\rangle$  and  $|135^\circ\rangle = |-\rangle$ , which is our  $X$  basis [15]. The states are geometrically depicted in Figure 3.4.

There is also a modified BB84 protocol called the six-state protocol. This protocol is the same as BB84 but it uses 3 bases  $X$ ,  $Y$ , and  $Z$  instead of only two. Using more basis means that Eve will get less information from measure and resend attack, which make this protocol more tolerant to noise. On the flip side, Bob has now only 1/3 chance to choose the correct basis, which means that approximately 2/3 of measurements will be discarded [18].

The BB84 protocol is fundamental for most other protocols, which are based on BB84, but modified to withstand attacks on imperfections caused by non-ideal devices. This is discussed in more detail in subsection 4.6.

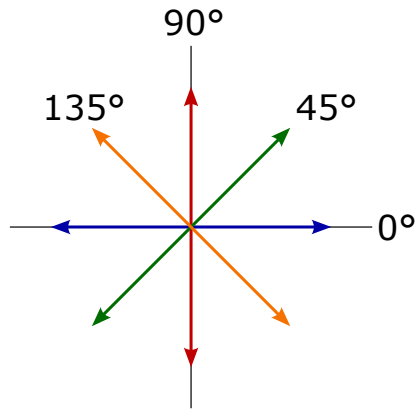


Figure 3.4: Set of polarization of photons representing quantum states.

### 3.2.2 B92 protocol

The B92 protocol, proposed by C. H. Bennett in 1992, simplifies the original BB84 by utilizing only two mutually non-orthogonal states, typically  $|0\rangle$  and  $|+\rangle$ . The critical aspect of this protocol is the non-orthogonality of these states, which ensures that the no-cloning theorem remains applicable, providing security against cloning attacks. Ideally, these states are chosen such that their scalar product equals  $1/\sqrt{2}$ , enhancing the protocol's effectiveness [19]. The security of B92 was formally proven in 2003 [20, 21].

In B92, Alice sends one of the two non-orthogonal states, and Bob measures the incoming qubit using either the  $X$  or  $Z$  basis. A key point in this protocol is that if Alice sends  $|0\rangle$ , Bob will never measure  $|1\rangle$ , regardless of his chosen basis. Similarly, if Alice sends  $|+\rangle$ , Bob will never measure  $|-\rangle$ . For example, if Alice sends  $|0\rangle$  and Bob measures in the  $Z$  basis, he will always detect  $|0\rangle$ . However, if he chooses the  $X$  basis, he has a 50% chance of measuring either  $|+\rangle$  or  $|-\rangle$ . Consequently, specific measurements by Bob allow him to infer which state Alice sent: if he measures  $|1\rangle$ , he knows Alice sent  $|+\rangle$ ; if he measures  $|-\rangle$ , he deduces that Alice sent  $|0\rangle$ .

During key sifting, Bob retains only the results where he measured  $|1\rangle$  or  $|-\rangle$ , and Alice keeps corresponding bits. Although B92 transmits only two states, resulting in a theoretical 75% discard rate of transmitted qubits (compared to 50% in BB84), it simplifies not needing to share basis choices. They assess the QBER by comparing parts of their keys. Due to its reliance on only two states, B92 demands a lower QBER for security—ideally below 12.5%, stricter than the 25% required in BB84. This is because during the intercepts-resend attack, Eve selects only between two possible states. Therefore QBER must be lower than  $1/4$  of the sifted key, which gives upper limit on QBER to 12.5%.

To summarize

$$P(0|-) = P(+|1) = 1, \quad (3.20)$$

where  $P(0|-)$  is a probability that Alice sent state  $|0\rangle$  given that Bob measured

state  $|-\rangle$  is equal to 1 and the same goes for the case when Alice sent state  $|+\rangle$  and Bob measured  $|1\rangle$ .

This protocol can also be visualized using photon polarization: Alice emits photons polarized at either  $0^\circ$  or  $45^\circ$ , while Bob sets his polarization filter to  $90^\circ$  or  $135^\circ$ . If Alice sends a photon polarized at  $0^\circ$  and Bob's filter is set to  $90^\circ$ , no photon passes through, confirming Alice's polarization. Conversely, if a photon does pass through, Bob concludes Alice's photon was polarized at  $45^\circ$ . This deterministic outcome based on filter settings underscores the protocol's reliance on the non-orthogonality of quantum states and Bob's measurements to deduce the correct bit values.

This protocol can also be visualized using photon polarization. Alice emits photons polarized at either  $|0^\circ\rangle$  or  $|45^\circ\rangle$ , while Bob receives the photons with his polarization filter set to  $90^\circ$  or  $135^\circ$ . Note that Bob chooses orthogonal polarizations compared to Alice. We now have 2 possible scenarios. If Alice sent a photon with polarization  $0^\circ$  and Bob set his axis in polarization filter to  $90^\circ$ , then he will not detect any photon. The same goes when Alice sends a photon with polarization  $45^\circ$  and Bob has a polarization filter with the axis rotated  $135^\circ$ . But when Bob chooses  $90^\circ$  and he detects a photon, then he knows that Alice must have sent a photon with polarization  $45^\circ$  and he writes this down as a bit with value 0. He knows for sure because as we have just said, in this case, polarization  $0^\circ$  never passes through the polarization filter with the axis rotated  $90^\circ$  and Alice sends only two states. Analogically, Bob notes bit with value 1, when a photon passes while he has its polarization filter set to  $135^\circ$ . In this case, he knows for sure that Alice sent a photon with polarization  $45^\circ$ .

### 3.2.3 BB84 and B92 protocols with time and phase encoding

Interferometric QKD encodes quantum information into the photon's phase rather than its polarization, differing from traditional polarization-based methods. This technique was initially explored by Charles Bennett in the same paper as the protocol B92 [19], as an extension to the foundational BB84 protocol. This protocol utilizes a Mach-Zehnder interferometer with 50:50 beam splitters as depicted in Figure 3.2, previously described in section 3.1.2. The primary difference involves utilizing two delay modules, one for Alice and one for Bob. Modifications in equations (3.12) and (3.13) include replacing  $\varphi$  with the phase difference induced by both phase modulators,  $\varphi_A - \varphi_B$ . It is crucial for interference that the coherence length of the light is larger than the path difference in the interferometer arms. The modified output intensity equation is

$$I_{out} = I_{source} \cdot \cos^2\left(\frac{\varphi_A - \varphi_B + k\Delta L}{2}\right), \quad (3.21)$$

where  $k$  is the wave number,  $\Delta L$  is the path difference between the interferometer arms,  $I_{source}$  is the input intensity, and  $I_{out}$  is the output intensity. This formula-



tion is valid for both classical light and single photons [22]. The resulting detection probabilities are

$$P_1 = |\langle 1_4 0_5 | \psi_{out} \rangle|^2 = \cos^2 \left( \frac{\varphi_A - \varphi_B}{2} \right), \quad (3.22a)$$

$$P_2 = |\langle 0_4 1_5 | \psi_{out} \rangle|^2 = \sin^2 \left( \frac{\varphi_A - \varphi_B}{2} \right), \quad (3.22b)$$

To implement the B92 protocol, Alice applies a phase shift of either 0, corresponding to bit 0, or  $\pi/2$  for bit 1. Bob applies phase shifts of  $3\pi/2$  for bit 0 and  $\pi$  for bit 1. Summary for this protocol is in Table 2. This yields deterministic outcomes under certain conditions. For instance, if Bob sets his modulator to  $3\pi/2$  and detector D1 clicks, he knows Alice sent a zero phase modulation, interpreting this as bit 0. The valid measurement for bit 1 occurs when detector D2 clicks while Bob's modulator is set to  $\pi$ .

Alice PM	Bob PM	$\Delta\varphi$	$P_1$	$P_2$
0	$\pi$	$-\pi$	1	0
0	$3\pi/2$	$-3\pi/2$	0	1
$\pi/2$	$\pi$	$-\pi/2$	0	1
$\pi/2$	$3\pi/2$	$-\pi$	1	0

Table 2: Table of all possible values of phase modulations (PM) for B92 protocol with corresponding probabilities  $P_1$ ,  $P_2$  of measurements either on detector D1 or D2.

This approach is straightforward, yet it faces a significant challenge if Alice and Bob need to communicate over large distances. They would require a single large Mach-Zehnder Interferometer (MZI), which becomes extremely unstable because the path length difference must remain within a fraction of the wavelength. Consequently, this configuration is impractical for transmission distances exceeding a few meters. A more robust solution involves each party owning an unbalanced or asymmetrical Mach-Zehnder interferometer (AMZI), depicted in Figure 3.5. Output from Alice's AMZI is multiplexed into a single fiber, serving as the quantum channel to Bob's AMZI. The path difference of the arms of the single AMZI,  $\Delta L$ , significantly exceeds the coherence time of the light source, preventing interference within individual AMZIs. Interference only occurs when two AMZIs are connected, and the photon passes once through the long and once through the short arm, occurring statistically in half of the cases. This setup yields four potential time-bins, with the remaining two being non-interfering paths (short-short and long-long), which detections are probabilistic and carry no useful information. For practical implementation, precise synchronization and detection capabilities are necessary to distinguish the three possible time-bins to discard all measurement for these non-interfering paths. For practical realization, we must ensure that we are able to differentiate the path that

a photon took. In other words, we must have synchronization and detection precise enough to distinguish the 3 time-bins / 3 paths.

In the case with one large symmetric MZI, the photon is superposition of two paths (either traveling through one arm or the other), and in each path the photon can experience different environmental noise or optical length deviations. However, in AMZI setup the photon travels in both paths (short-long and long-short) the same fiber link between Alice and Bob. Therefore all the transmitted photons experience identical environmental fluctuations for most of the journey, except for the few meters within the AMZIs, which is almost negligible in comparison in the tens of kilometers that can separate Alice and Bob. This significantly simplifies maintaining the overall path length stable, as variations are slower than the temporal separation (typically around 5 ns) determined by the path imbalance in the AMZIs. Moreover, for good interference visibility and low error rates, the path imbalances in Alice and Bob's AMZIs only need slight adjustments, to be matched within a fraction of the photon's coherence time, which corresponds to a few millimeters. This makes the AMZI setup a feasible method for practical long-distance quantum key distribution.

Despite these advantages, the AMZI setup still presents practical challenges. The path imbalances in Alice's and Bob's interferometers must remain stable within a fraction of the photon's wavelength throughout the key exchange to preserve proper phase relations. This requires placing the interferometers in temperature-stabilized environments and, for extended operations, an active system to compensate for any residual drifts. Additionally, to ensure indistinguishability between the interfering paths (short-long and long-short), identical polarization transformations induced by the short and long paths in each interferometer must be achieved using polarization controllers. However, since the polarization in short, temperature-stabilized optical fibers without strain remains relatively stable, frequent adjustments for polarization matching are not necessary [22].

Note that for the two cases of short-long and long-short, the two interferometers act as a single MZI, but for the paths of long-long and short-short, it does not behave as interferometer at all. Thus, the probabilities (3.22) will be multiplied by an additional factor of 1/4, indicating that interference occurs only in half of the cases

$$P_1 = \frac{1}{4} \cos^2 \left( \frac{\Delta\varphi}{2} \right) = \frac{1}{8} (1 + \cos(\Delta\varphi)), \quad (3.23a)$$

$$P_2 = \frac{1}{4} \sin^2 \left( \frac{\Delta\varphi}{2} \right) = \frac{1}{8} (1 - \cos(\Delta\varphi)). \quad (3.23b)$$

For the BB84 protocol, Alice's two bases are  $\{0, \pi\}$  and  $\{\pi/2, 3\pi/2\}$ . Bob then applies only phase modulation of either 0 or  $\pi/2$  as a measurement in the first or second basis. Again, due to the single-photon interference, detectors can click

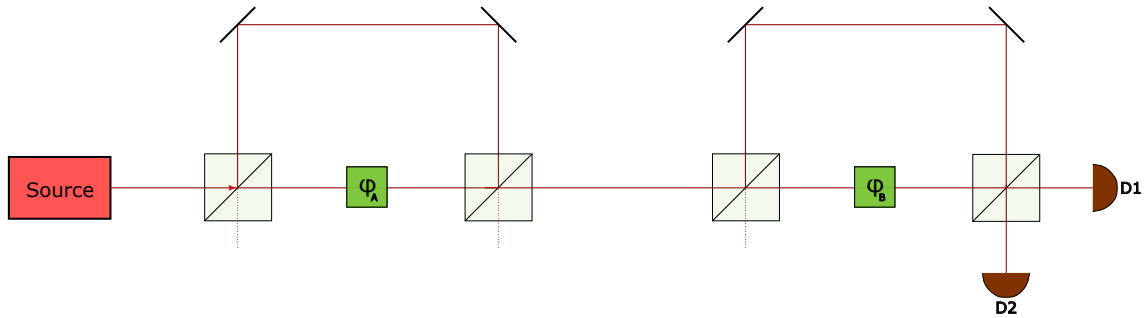


Figure 3.5: Illustration of two asymmetrical Mach-Zehnder interferometers used for time and phase encoded QKD. The left part consists of the source and the first AMZI with phase modulation  $\varphi_A$ , belonging to Alice. The right part includes Bob's AMZI with phase modulation  $\varphi_B$  and two single-photon detectors D1 and D2.

Alice PM	Bob PM	Bob's detector	Resulting bit
0	0	D1	0
$\pi/2$	$\pi/2$	D1	0
$\pi$	0	D2	1
$3\pi/2$	$\pi/2$	D2	1

Table 3: Table showing possible options in phase encoded BB84 protocol for cases when Alice and Bob choose the same basis and the photon travels short-long or long-short path. For paths short-short and long-long, the probabilities are 1/2 to detect a photon on the detector D1 or D2.

deterministically for specific cases. D1 clicks for  $(0,0)$  and  $(\pi/2, \pi/2)$ , and D2 clicks for  $(\pi,0)$  and  $(3\pi/2, \pi/2)$ . In all other cases, detectors click probabilistically. During the key sifting, Bob announces the time when he detected a photon and which phase modulation he applied. Alice tells Bob from which basis she chose the phase modulation. From this information, Alice and Bob can deduce the photon's path and whether the detection was deterministic or probabilistic. They retain only the deterministic measurements. For Alice, bit 0 is when she shifted phase by 0 or  $\pi/2$ , and bit 1 is when  $\varphi_A = \pi \vee 3\pi/2$ . Bob evaluates bit 0 as a click of detector D1 and bit 1 as a click of D2. They confirm both possess a symmetrical key and can test it for QBER [23, 24]. This protocol is summarized in Table 3.

Another setup, that reduces the stability requirements even more, is called the Plug and Play system. This setup is illustrated in Figure 3.6 and consists of only one AMZI, where a strong light signal travels from Bob to Alice, where the pulse is attenuated<sup>2</sup> and its phase modulated. This configuration benefits from self-stabilization

<sup>2</sup>Here we assume that Alice and Bob communicate using single photons, but they are actually using strongly attenuated pulses. More thorough discussion about this topic follows in section 4.2 and in section 4.4.2.

as the light passes twice through the same interferometer at Bob's end. At Alice's side, the light is reflected using a Faraday mirror, which compensates for polarization effects as well. The Faraday mirror combines a mirror with a  $\lambda/4$  Faraday rotator, shifting the photon's polarization by  $45^\circ$  upon passage and reflecting it back for another  $45^\circ$  shift. But the fundamental principle of phase-time BB84 still holds, Alice modulates phase by  $\varphi_A \in \{0, \pi/2, \pi, 3/2\pi\}$  and Bob by  $\{0, \pi/2\}$ , however Bob now owns almost all of the equipment [25].

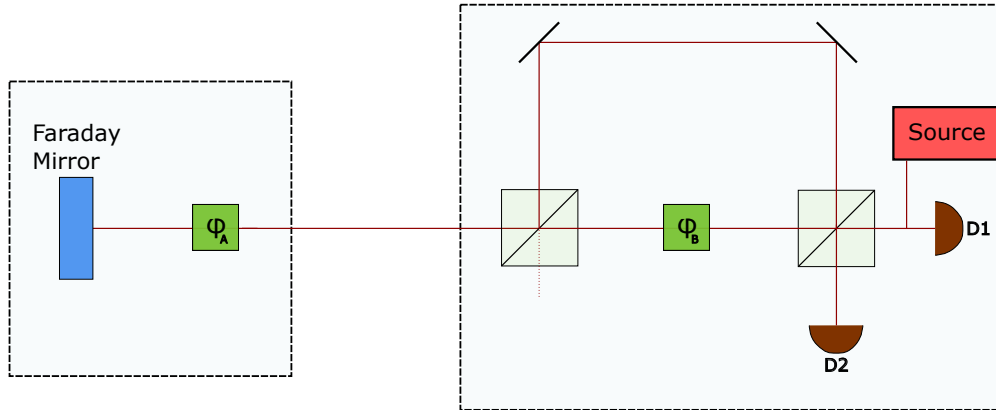


Figure 3.6: BB84 protocol in a Plug and play version. Bob owns the laser source, detectors, phase modulator  $\varphi_B$  and the AMZI, whereas Alice owns only phase modulator  $\varphi_A$  and a Faraday mirror.

The primary drawback of this configuration is its duty cycle, which limits the key rate. Bob must wait for his strong pulse to return from Alice; otherwise, if this pulse intersects with Alice's reflected weak coherent pulse en route, the back-scattered photons from the strong pulse would overwhelm the weak coherent pulse sent back by Alice. Thus, Bob has to send his pulses in batches (also called trains) to ensure all have passed the phase modulator before the first one returns [26].

This section covered interferometric QKD protocols that encode information in the phase of the photon rather than polarization. It described the basic Mach-Zehnder interferometer setup for the B92 protocol and its limitations for long distances. The asymmetric Mach-Zehnder interferometer setup was introduced as a practical solution for long-distance distribution by having interfering paths that experience the same environmental fluctuations over the fiber link. The BB84 protocol was adapted to the phase-time encoding scheme using AMZIs at both Alice's and Bob's ends. While providing stability advantages, the AMZI approach requires active phase tracking and polarization compensations. Finally, the "Plug and Play" system utilizing a single AMZI at Bob's side with a Faraday mirror reflector at Alice's end was presented as an even more stable configuration, though with a limitation on the duty cycle. Overall, these phase-encoding protocols offer an alternative to polarization techniques and enable long-range QKD with reasonably practical implementations.

### 3.2.4 E91 protocol

Developed by A. Ekert in 1991, the E91 protocol utilizes the entanglement of quantum states to enable secure quantum key distribution [27]. Unlike traditional QKD protocols that rely on the uncertainty principle, E91 leverages the quantum entanglement of the Bell states.

The protocol begins with a third party, Charlie, who generates a pair of entangled qubits in one of the Bell states, for example

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad (3.24)$$

and distributes these qubits to Alice and Bob. Charlie must announce the specific entangled state sent to ensure the protocol's integrity.

Alice and Bob randomly choose measurement bases from predefined sets. Alice's possible bases are  $A_1 = X$ ,  $A_2 = Z$ , and  $A_3 = \frac{1}{\sqrt{2}}(Z + X)$ , while Bob's bases are  $B_1 = Z$ ,  $B_2 = \frac{1}{\sqrt{2}}(Z - X)$ , and  $B_3 = \frac{1}{\sqrt{2}}(Z + X)$ . Key generation occurs when Alice and Bob measure in the  $(A_1, B_1)$  or  $(A_3, B_3)$  combinations, which are expected to produce correlated outcomes due to their entangled state.

The security of E91 relies on the quantum entanglement and on the fact that local manipulation with the bipartite entangled state can only introduce error to the state. For security verification, Alice and Bob use the remaining measurement combinations to test the entanglement's integrity by calculating the Clauser-Horne-Shimony-Holt (CHSH) inequality[3]

$$S = |\langle A_1 B_2 \rangle + \langle A_2 B_2 \rangle + \langle A_1 B_3 \rangle - \langle A_2 B_3 \rangle| \leq 2\sqrt{2}. \quad (3.25)$$

This inequality was presented as a proof of Bell's theorem which states that certain consequences of entanglement cannot be reproduced by local hidden-variable theories. In E91, this inequality helps determine if the qubits kept their entanglement during transmission. A value of  $S$  close to  $2\sqrt{2}$  indicates strong entanglement, while  $S \leq 2$  suggests either significant channel noise or potential eavesdropping.

E91's robustness against interception is underscored by the fact that even if Eve controls the qubit source, the protocol remains secure as long as Alice and Bob verify the entanglement through the CHSH inequality. Because if Eve tries to gain information about the key by entangling herself with the qubits sent to Alice and Bob, the quantum state between Alice's and Bob's qubits will be disturbed.

In summary, in the E91 protocol, Alice and Bob use pairs of entangled photons to establish a shared, secret key. When Alice and Bob measure the states of their respective qubits, the results are correlated due to entanglement. As in the previous protocols, the common information carrier is also a photon. The security check in

the protocol is not verified by calculating the QBER, but Alice and Bob test the violation of Bell's inequalities that distinguishes between correlations described by quantum physics (entanglement) and all classical forms of correlations (those that might be simulated by an eavesdropper). If the measurement results between Alice and Bob violate Bell's inequalities, it implies that the photons are entangled in a quantum manner, and no eavesdropper can replicate these results.

### 3.2.5 BBM92 protocol

The BBM92 protocol, developed by Charles Bennett, Gilles Brassard, and N. David Mermin in 1992[28], was designed to integrate the advantages of the E91 protocol with the simplicity of the BB84 framework. Like E91, BBM92 uses entangled photon pairs to enable quantum key distribution but simplifies the process by eliminating the need for testing the CHSH inequality.

In BBM92, a source generates pairs of entangled qubits (typically entangled in polarization) and distributes one qubit to Alice and the other to Bob. Both Alice and Bob independently choose their measurement basis from two options, typically denoted as  $X$  and  $Z$ . This is similar to the basis choices in BB84 but here used to utilize the entanglement. After measurements, Alice and Bob communicate via a public channel to announce the bases they used for each measurement. They disregard any results where they used different bases, as these would not be correlated.

For the cases where both used the same bases, the results should ideally be perfectly correlated due to the properties of entanglement. To verify this and compute the QBER, Alice and Bob compare a subset of their results. If the error rate is within acceptable limits, they confirm that no eavesdropping has occurred, and they can use the remaining correlated bits as their shared secret key. Unlike E91, BBM92 does not require the verification of the CHSH inequality to ensure security. The protocol assumes that any attempt by an eavesdropper (Eve) to measure or interact with the qubits would disrupt the entanglement, leading to detectable errors in the correlations of Alice's and Bob's measurements. Therefore, checking the QBER is sufficient to infer the presence of eavesdropping.

BBM92 thus combines the practicality of BB84's basis sifting with the robust security afforded by quantum entanglement, similar to E91 but without the complexity of testing quantum inequalities.

As in interesting note to conclude the list of different QKD protocols, a paper by V. Sharma et al. [29] concludes that there is no significant difference in effectiveness of single-qubit-based versus entangled-state-based quantum communication protocols. In ideal, noiseless conditions, both types of protocols perform equivalently; however, this balance shifts under realistic, noisy conditions. Single-qubit protocols are typically more effective in environments affected by amplitude damping and phase

damping noises, while entangled-state protocols are superior under collective noise conditions. Interestingly, the study suggests that while entanglement is a valuable resource for specific quantum tasks like teleportation and dense coding, its use in other cryptographic tasks might not always justify the higher physical costs compared to single qubits, especially in well-characterized channels where noise affects each qubit independently.

### 3.3 Post-processing

Post-processing is a purely classical part of a Quantum Key Distribution protocol, performed after the quantum phase, i.e., the transmission of qubits and their measurement. The goal of post-processing is to estimate and correct errors and to enhance security through several phases, specifically *parameter estimation*, *error correction*, and *privacy amplification*[6]. Initially, Alice and Bob randomly select a subset of their transmitted keys to perform QBER estimation. Given the inherent noise in the quantum channel, errors are inevitable. To address these errors, Alice and Bob employ specialized error correction algorithms designed for QKD systems. However, this process inevitably exposes some of their secret bits to potential public exposure. Consequently, they employ privacy amplification techniques to further secure the corrected key against potential eavesdroppers.

#### 3.3.1 Parameter estimation

Parameter estimation is the initial step in the post-quantum phase of the protocol where Alice and Bob assess errors in the key transmission. If the error rate exceeds the predefined threshold of the QKD protocol, the protocol is aborted since it cannot be confidently determined whether the communication was compromised or was only noisy. For error evaluation, Alice transmits a part of her sifted key to Bob, who then calculates the error rate. Although only a portion of the key is used for this estimation, statistical methods are used to confirm its sufficiency to correctly reflect the overall key's error rate. According to Chernoff–Hoeffding bounds, a large random subset is representative of the entire set's statistical properties. Therefore, if Alice and Bob detect a QBER of  $n\%$  in their sample, it is indicative of the entire key of the length  $N$ .

We introduce Hoeffding's inequality to illustrate this concept. Consider a finite set  $\mathcal{X} = (x_1, \dots, x_N)$  of  $N$  elements and a random sample  $X_1, \dots, X_n$  drawn from  $\mathcal{X}$  without replacement. Let  $a = \min_{1 \leq i \leq N} x_i$  and  $b = \max_{1 \leq i \leq N} x_i$ . Define the average of  $\mathcal{X}$  as

$$\bar{\mathcal{X}} = 1/N \sum_{i=1}^N x_i \text{ and the average of the sample } X \text{ as } \bar{X} = 1/n \sum_{i=1}^n X_i. \text{ For any } \varepsilon > 0,$$

Hoeffding's inequality is given by:

$$P(\bar{X} - \bar{\mathcal{X}} \geq \varepsilon) \leq \exp\left(-\frac{2n\varepsilon^2}{(b-a)^2}\right). \quad (3.26)$$

This inequality demonstrates that the probability of the sample mean  $X$  deviating from the set mean  $\mathcal{X}$  by more than  $\varepsilon$  decreases exponentially with the sample size  $n$  and the square of  $\varepsilon$ .

For binary data, where  $x_i \in \{0,1\}$ , Hoeffding's inequality simplifies to

$$P(\bar{X} - \bar{\mathcal{X}} \geq \varepsilon) \leq \exp(-2n\varepsilon^2), \quad (3.27)$$

where  $0 \leq \varepsilon \leq 1$ .

A refinement, Serfling's inequality, accounts for the total key length  $N$  minus the number of values drawn without replacement  $n$ :

$$P(\bar{X} - \bar{\mathcal{X}} \geq \varepsilon) \leq \exp\left(-\frac{2n\varepsilon^2}{(1 - \frac{n-1}{N})(b-a)^2}\right), \quad (3.28)$$

and for binary values, this simplifies to:

$$P(\bar{X} - \bar{\mathcal{X}} \geq \varepsilon) \leq \exp\left(-\frac{2nN\varepsilon^2}{N-n+1}\right). \quad (3.29)$$

These inequalities demonstrate the exponential convergence of the probabilities with the product of  $n\varepsilon^2$  as the argument. Serfling's inequality converges faster than Hoeffding's inequality, particularly reflecting dependence on the total key length  $N$  as illustrated in Figures 3.8 and 3.7.

To determine the optimal number of bits for parameter estimation  $k$ , Alice and Bob compute the probability that the error rate in the sampled bits  $\Lambda_k$  differs by a constant  $\delta$  from the error rate in the remaining bits  $\Lambda_n$ . This probability is conditioned on  $\Lambda_k$  being below a threshold  $\lambda_{max}$ ; otherwise, the protocol is aborted. The probability is denoted as:

$$P(\Lambda_n - \Lambda_k \geq \delta \mid \Lambda_k \leq \lambda_{max}), \quad (3.30)$$

which utilizes the statistical principles of Serfling's inequality to ensure robustness in error rate estimation across the key, allowing Alice and Bob to validate the integrity of the key transmission securely.

Suppose Alice and Bob's key strings are  $K_A$  and  $K_B$  respectively, with length  $N$ . We take  $k$  bits from the key, so  $n$  bits of the original key remains. We denote the  $k$  bits as  $K_A^k$  for Alice and  $K_B^k$  for Bob and the remaining keys as  $K_A^n$  and  $K_B^n$ . We write the original Alice and Bob's keys as  $K_A = K_A^k K_A^n$  and  $K_B = K_B^k K_B^n$  respectively. We also define Hamming weight

$$|K_A^k \oplus K_B^k| \quad (3.31)$$



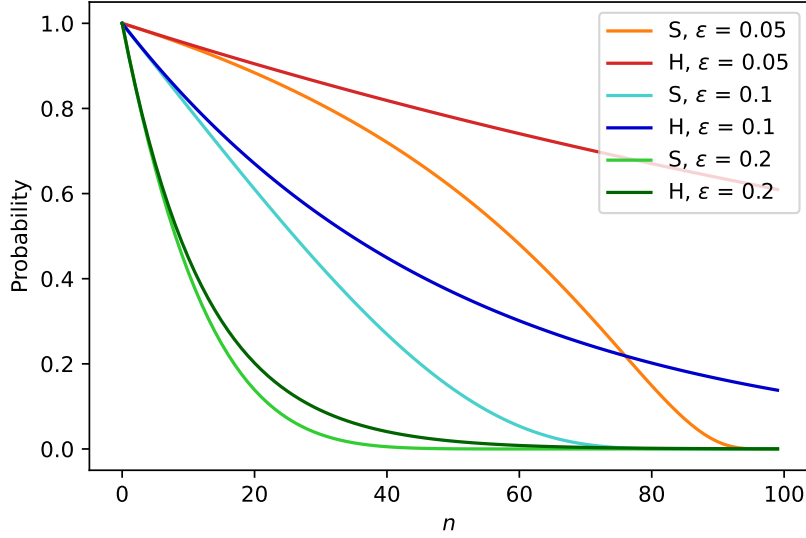


Figure 3.7: Graph depicting the exponentially decreasing probability  $P(\bar{X} - \bar{\mathcal{X}} \geq \varepsilon)$  for Serfling's inequality (S) and for Hoeffring's inequality (H) in relation with increasing length of the drawn key  $n$ . Each inequality is also shown for three values of  $\varepsilon = \{0,05, 0.1, 0.2\}$ . The total key length used was  $N = 100$ . Also we used simplified equations (3.27), (3.29) for binary values of  $x$ .

as the number of the error bits, i.e., how many bits in the sample  $K^k$  differs at Alice and Bob, where symbol  $\oplus$  denotes XOR operation. For the same bits, XOR results in bit 0, while if the bits are different, the result of XOR is bit 1. Basically, the Hamming weight tells us how many 1s are in the result of the XOR operation. Using the Hamming weight, we define error rates as

$$\Lambda_k = \frac{1}{k} |K_A^k \oplus K_B^k|, \quad (3.32a)$$

$$\Lambda_n = \frac{1}{n} |K_A^n \oplus K_B^n|, \quad (3.32b)$$

$$\Lambda_N = \frac{1}{N} |K_A \oplus K_B|. \quad (3.32c)$$

The total error rate can be rewritten as

$$\Lambda = \nu \Lambda_k + (1 - \nu) \Lambda_n, \quad (3.33)$$

where  $\nu = k/N$ . Using Bayes' theorem<sup>3</sup>, we can set a higher bound to the probability

---

<sup>3</sup>Bayes' theorem states that  $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

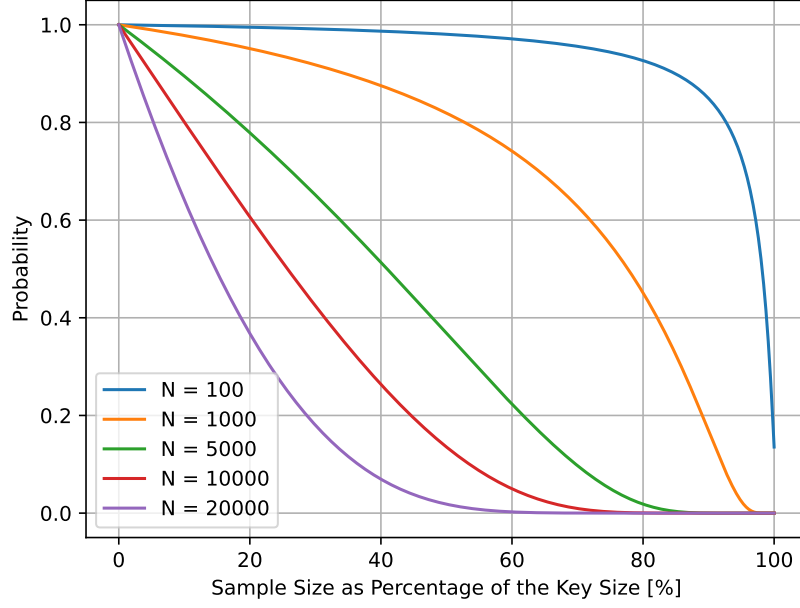


Figure 3.8: This figure illustrates the upper bounds on the probability that the sample of the key of the size  $n$  deviates from the total key size  $N$  by more than a specified epsilon  $\varepsilon = 0.01$ , calculated using Serfling's Inequality. The x-axis represents the sample size  $n$  as a percentage of the total key size  $N$ , demonstrating how the deviation probability depends on both the total key size and sample size. Different curves correspond to different population sizes  $N = \{100, 1000, 5000, 10000, 20000\}$ . The plot shows the importance of sample size selection in achieving desired precision levels in QBER estimation.

(3.30) as

$$P(\Lambda_n - \Lambda_k \geq \delta | \Lambda_k \leq \lambda_{max}) \leq \frac{P(\Lambda_n - \Lambda_k \geq \delta)}{P(\Lambda_k \leq \lambda_{max})}, \quad (3.34)$$

where  $P(\Lambda_k \leq \lambda_{max})$  is the probability that passes the check and  $P(\Lambda_n - \Lambda_k \geq \delta)$  can be estimated using Serfling's inequality (3.29) as

$$P(\Lambda_n - \Lambda_k \geq \delta) \leq \exp\left(-\frac{2k^2n\delta^2}{(k+1)N}\right). \quad (3.35)$$

Thus the whole probability has higher bound

$$P(\Lambda_n - \Lambda_k \geq \delta | \Lambda_k \leq \lambda_{max}) \leq \frac{\exp\left(-\frac{2k^2n\delta^2}{(k+1)N}\right)}{P(\Lambda_k \leq \lambda_{max})}. \quad (3.36)$$

This inequality says that the error rate on the remaining  $n$  bits and the  $k$  bits, which differs by arbitrarily small constant  $\delta$ , is exponentially decreasing with larger  $k$ , i.e., the more bits we use for parameter estimation, the higher chance that the rate of

the error in the sample will be that same as in the remaining  $n$  bits. Thus Alice and Bob can determine the errors in their samples using (3.32a) and they know it offers reliable information about the error in the whole key. Note that although the errors could result from channel noise or an eavesdropper, Alice and Bob must assume that the transmission was noise-free and that all errors contributing to the QBER were introduced by Eve. Consequently, they must consider that Eve might have acquired all the information that could potentially be inferred from the errors quantified by the QBER.

### 3.3.2 Error correction

The aim of the next step is to correct any error in the sifted key, also called *information reconciliation*. To do so, Alice and Bob need to locate the errors and determine what is the correct value. However, Alice and Bob share this information through the classical channel, which is prone to eavesdropping. Thus they have to share this information in a way, which ensures security for the derived key, but also it allows them to compare the keys if they are the same.

In QKD, we need to work with another variable: the number of exposed bits. These are the bits used for error correction that are revealed by Alice on the public channel, thus reducing key secrecy. We can use the *Reconciliation key rate*  $R_k$  to express

$$R_k = \frac{n_s - n_e}{n_s} \quad (3.37)$$

where  $n_s$  is the number of sifted key bits, and  $n_e$  is the number of bits exposed. The reconciliation key rate essentially measures the efficiency of the error correction process in a QKD setup, indicating the proportion of the original sifted key bits that remain secure and usable after the error correction process.

In the section 2.4.1, we introduced the binary symmetric channel along with the bit error rate. We also mentioned the use of error-correcting codes and the theoretical limit from the definition of channel capacity. In this section, we will describe the first widely used error-correcting code proposed by Brassard and Salvail [30] and discuss its application in QKD.

This protocol is called *Cascade* and is based on a parity check [31]. Parity check is a simple error correcting technique that appends a parity bit to a block of  $n$  bits. This parity bit is calculated as the modulo 2 sum of the  $n$  bits in the block. If the parity bit is 0, it means there is an even number of 1s in the block; otherwise, a 1 indicates an odd number of 1s. Alice and Bob compare their parities for each block to determine if any blocks contain errors.

Alice and Bob will agree on a block size beforehand and divide their sifted key into blocks of this length. Then they will add a parity bit and share the result. If the parities match, they assume the block is correct and proceed to the next block. If

not, they divide the block into two sub-blocks and iteratively search for the error bit. Note that this method is not perfect, as it can only detect some types of errors. For example, it cannot detect two bit-flip errors because the parity for a code with two bit-flip errors is the same as the parity for error-free bits. In general, we can detect if an odd number of bit-flips occurred, but we cannot determine how many. If an even number of bit-flips occurred (including zero), the parity check provides no information about the actual state of the bits.

The Cascade protocol is designed so that most of the computation is done on Bob's side [32]. This ensures potential scalability in cases where Alice acts like a server and has many clients communicating with her. The aim of the error correction is to create identical keys for Alice and Bob while keeping them secret. Thus, Alice and Bob can only share publicly so much information that anyone listening to the transmission channel will not be able to reconstruct the secret key. Also, note that this part is purely classical and uses only classical communication. Therefore, we must always assume that someone is eavesdropping on our communication.

The protocol is performed in iterations. The first iteration is the only one without shuffling. All consecutive iterations start with bit shuffling. The shuffling is not meant to add confusion to the key to protect it against eavesdroppers. Its purpose is to redistribute the errors from even parity blocks, because we have no information about errors in these blocks. Moreover, it helps to distribute errors uniformly, ideally to have one error per block maximum, because errors can often be grouped together. For example, if a disturbance in the transmission occurs, it could affect many consecutive bits, and parity check struggles with more than one error per block.

After the shuffling, bits are divided into blocks. In the first round, the block size is computed as a function of the QBER

$$k_1 \approx 0.73/QBER, \quad (3.38)$$

and for the next rounds,  $k_i$  are computed as the multiple of two of the previous block size

$$k_{i+1} = 2 \cdot k_i. \quad (3.39)$$

Note that block size of  $1/QBER$  means that there will be, on average, one error bit in the block. Therefore, by taking  $k_1$  smaller than  $1/QBER$ , we increase the probability that there will indeed be only one error in each block.

In each iteration, Bob computes the parity for each block and asks Alice for the correct parity of this block. If the parity checks are valid for a given block, Bob assumes that this block is correct. If the parities do not match and Bob's parity is 0, he knows there is an even number of errors, but for now, he will leave this block as it is. If the parities disagree and Bob's parity is 1, he knows he has an odd number of errors in this block and runs a binary algorithm (parity check with binary search) to repair these errors.

Bob's parity	Alice's parity	Error parity
0	0	0 (even)
0	1	1 (odd)
1	0	1 (odd)
1	1	0 (even)

Table 4: Table displaying different parities and the outcome. If the Bob's parity matches Alice's (1st and 4th row), Bob has no information if the block contains zero errors or an even number of errors. On the other hand, if the error parity is odd (2nd and 3rd rows), Bob knows that there is at least one or any odd number of errors in this block.

The binary algorithm repairs exactly one error in blocks that have an odd error parity. He divides this block in half, which ensures that one sub-block will now have even parity, and one will have odd parity (since odd parity means an odd number of 1s in the code block, both blocks will never have an equal amount of 1s after the division). Bob then asks Alice to provide the parity for the left sub-block, and he compares Alice's parity with his own parity of this sub-block. If the error parity is odd, Bob knows that the error is in this sub-block but still does not know exactly which bit is flipped. If the error parity is even, he knows that the error is in the second sub-block. He then recursively divides the identified errored sub-block until he reaches a block of one bit. When Alice sends him the parity of this one-bit sub-block, he finally knows which bit is flipped. Note that when Alice is sharing the parity of this single bit, she is basically sharing the actual value of this bit.

After Bob identifies and corrects all one-bit errors in all the odd error parity blocks, he will now have all of his blocks to have even error parity. However, he now does not know if the block contains 0, 2, 4, ... errors. Then the next iteration starts, where Bob again shuffles the bits and proceeds with the binary algorithm for odd error parity blocks. However, when performing the binary algorithm in the  $N + 1$  iteration, the parities will change also the parities in all the previous  $N$  iterations. Bob will go back to the  $N$  iteration and reapply the binary algorithm to correct newly emerged odd error parities. This will cause yet another cascade effect in the previous  $N - 1, N - 2, \dots$  iterations. Bob will backpropagate back to the first iteration while applying the binary algorithm along the way. This is called the cascade effect and thus the name Cascade protocol.

In conclusion, the Cascade protocol builds upon this simple parity check. Below is the summarized structure of the whole protocol:

1. Bob shuffles the bits of his noisy key.
2. Bob divides the key into smaller blocks.
3. Bob calculates the parity for each block.

4. Communication of parities:
  - a) Bob sends only the indices of the bits in his block to Alice.
  - b) Alice replies with the parity for the bits corresponding to these indices.
5. Error correction process for blocks with odd error parity:
  - a) Bob divides his block using binary search. This division continues iteratively, during which:
    - i. Bob sends indices of the subdivided blocks to Alice.
    - ii. Alice replies with the parity for the bits at these indices.
  - b) This process is repeated until the block size is reduced to one single bit.
  - c) Bob repairs this single bit based on Alice's parity result.
6. This process is repeated for all blocks that have an odd error parity.
7. Iterations and Cascade Effect:
  - a) The above steps are repeated for  $N$  iterations.
  - b) After completing  $N$  iterations, Bob recursively backpropagates from the  $N$ -th iteration back to the 1st iteration. This involves re-evaluating and potentially correcting earlier decisions based on the results of later iterations, thus causing the cascade effect.

A typical run of Cascade consists of 4 rounds with block sizes  $k_1, k_2, k_3, k_4$  iterations. After that all errors should be corrected. Then to ensure key privacy, Alice and Bob must delete one bit from the reconciled key for each parity bit exchanged over the classical channel, in order to reduce the amount of information gained by any potential eavesdroppers. These bits are called *exposed bits* as we introduced in the equation for the reconciliation key rate (3.37). Therefore, the final key size is the sifted key size minus the number of parity bits exchanged during the error reconciliation process.

After the error correction phase is completed, Alice and Bob can utilize hash functions to verify that they indeed have the same, error-less key. Hash functions are one-way functions that take as input of an arbitrary length and give out an output of a constant length. One-way means that it is easy to compute from the input the output, but it is computationally unfeasible to get the original input string from the output string. Additionally, an ideal hash function outputs a different string for every different input, i.e., there are no collisions. A collision is when two different inputs have the same output. In this case, we have two-universal hash functions, where we define that the probability of collision is lower than  $1/m$ .

Let  $U$  be a universe of inputs  $U = \{0, \dots, |U| - 1\}$ ,  $\mathbb{H}$  be a family of hash functions  $h_i$ ,  $\mathbb{H} = \{h_i\}_{i=1}^k$ , where  $h_i : U \rightarrow \{0, \dots, m - 1\}$ .  $\mathbb{H}$  is 2-universal if  $\forall x, y \in U$ , such

that  $x \neq y$  holds

$$P(h(x) = h(y)) \leq \frac{1}{m}, \quad (3.40)$$

where  $h$  is selected uniformly at random from  $\mathbb{H}$ .

Alice randomly chooses hash function  $h_A$ , she computes a hash of her key bit string and sends it together with the hash function she used to Bob. Bob then computes the hash with the same hash function  $h_A$  with his bits as input. If the two hashes match, then their key is without errors. A key distribution protocol is  $\varepsilon$ -correct if the probability that the two resulting keys of Alice and Bob differ at most by  $\varepsilon_{cor}$ . Probability that the keys are different even if the hashes are the same is [6]

$$P(K_A \neq K_B | h_A(K_A) = h_A(K_B)) \leq \varepsilon_{cor}. \quad (3.41)$$

There have been proposed many variants of the Cascade protocol, for example, Winnow protocol [33] or different error-correction techniques such as Low-Density Parity-Check (LDPC)[34]. And while Cascade protocols works quite well, they have some disadvantages. One of them is the need of Bob to communicate intensively with Alice. The other is the value of key reconciliation rate with respect to QBER. The Cascade protocol performed the best out of these 3 protocols for error rates less than 5% and its performance was very close to the theoretical limit. However, when increasing the error rate above 5%, Cascade diverges farther from the maximum achievable limit set by Shannon, and the LDPC protocol dominates. The performance of the Winnow is roughly the same distance from the theoretical limit regardless of the error rate, and is slightly better than LDPC for error rates less than 5% [35].

### 3.3.3 Privacy amplification

Privacy amplification is the concluding phase of classical post-processing in a QKD protocol. After error correction, Alice and Bob share a partially secure secret bit string that is  $\varepsilon$ -correct. However, this key might still be vulnerable as Eve could potentially have partial information about it. Either from the quantum transmission, of which our only evidence is the measured QBER in the parameter estimation phase, or during the error correction, where we assume that all shared parity information between Alice and Bob is known to Eve.

The aim of privacy amplification is to eliminate any potential advantage Eve might hold by reducing her knowledge about the final key to an insignificant level. This is achieved by shortening the key length while increasing its entropy, effectively rendering any partial information Eve holds useless.

Alice employs *randomness extractors* to accomplish this. A randomness extractor is a function that transforms a string with initially low entropy and a uniformly

random seed (possessing maximal entropy) into an output that is highly random and uniformly distributed. The output's security is guaranteed if its length is shorter than the min-entropy, which represents Eve's uncertainty about the input. The process is depicted mathematically as follows:

Suppose Alice and Bob's corrected key is represented by a classical random variable  $X$ , with realizations  $x$  of a set of values  $\mathcal{X}$  and a corresponding probability distribution  $p_X$ . Eve's partial information is represented by a quantum system  $E$ ,  $\rho_E$ , correlated with  $X$ . This relation is modeled by a classical-quantum state  $\rho_{XE} \in \mathcal{B}(\mathcal{H}_X \otimes \mathcal{H}_E)$ :

$$\rho_{XE} = \sum_{x \in \mathcal{X}} p_X(x) |x\rangle \langle x| \otimes \rho_E^x, \quad (3.42)$$

where  $\{|x\rangle\}$  forms an orthonormal basis and  $\rho_X = \sum_x p_X(x) |x\rangle \langle x|$  is a density operator of this classical ensemble  $\{p_X(x), |x\rangle \langle x|\}$ . Additionally, state  $\rho_Y \in \mathcal{B}(\mathcal{H}_Y)$  represents the seed. The correlation between  $X$  and  $E$ , i.e., how much information Eve has about their key  $X$ , is described by the conditional min-entropy from equation (2.80). We want to transform the key  $X$  to a new key  $Z$  with additional randomness, such that Eve's information about the key is below certain level. The definition of a set of secure functions, called *quantum-proof randomness extractors* that are  $(k, \varepsilon)$  strong is as follows:

$(k, \varepsilon)$ -strong quantum-proof randomness extractor is a function  $\{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^n$  that for all classical-quantum states  $\rho_{XE}$  with classical random variable  $X \in \{0, 1\}^n$  with conditional min-entropy greater than a certain threshold  $k$ ,  $H_{\min}(X|E) \geq k$ , and uniformly random seed  $Y \in \{0, 1\}^d$ , the following inequality holds

$$\frac{1}{2} \left\| \rho_{ext} - \frac{\mathbb{I}}{2^m} \otimes \rho_Y \otimes \rho_E \right\|_1 \leq \varepsilon, \quad (3.43)$$

where  $\|\rho\|_1 = \text{Tr}(\sqrt{\rho^\dagger \rho})$ . Density matrix  $\rho_{ext}$  represents the state, where key  $X$  was transformed using a randomness extractor with seed  $Y$  and with Eve's state intact. The second state  $\frac{\mathbb{I}}{2^m} \otimes \rho_Y \otimes \rho_E$  is an ideal state before the transformation, where  $\frac{\mathbb{I}}{2^m}$  represents the key  $X$ , which is in maximally mixed state, in other words, it is uniformly random, since it is the result of a 2-universal hash function from error correction phase. Furthermore, this state is independent of the seed  $Y$  and Eve's state  $E$ .

An example of quantum-proof randomness extractor function is the 2-universal hash function. The protocol of privacy amplification is as follows: Alice and Bob publicly generate a random seed  $Y$ , according to which they randomly choose a hash function  $h$  from the set of 2-universal hash functions  $\mathcal{H}$  and computes the hash of their raw key, i.e.,  $K_A \rightarrow h(K_A)$ . The result is a uniformly random bit string, uncorrelated with Eve's information about the original state  $X$ . Moreover, this has function cannot



worsen the  $\varepsilon$ -correct parameter  $\varepsilon_{cor}$  from error correction. In other words, for two keys  $K_A$  and  $K_B$ , applying the hash functions on both of them does not make more errors than we estimated in the error correction phase.

To verify the security of the generated key  $h(K_A)$ , we use so called *Quantum Leftover Hash Lemma*. This lemma states the maximum amount of uniform randomness that can be extracted from  $K_A$  while being independent of  $E$  with smoothing parameter  $\varepsilon'$  [6].

Let  $\rho_{h(K_A)YE}$  be the output state after applying 2-universal hash function to the Alice's key  $K_A$ . Then  $\forall \varepsilon' > 0$  it holds that

$$D(\rho_{h(K_A)YE}, \rho_U \otimes \rho_{YE}) \leq 2\varepsilon' + \frac{1}{2} \sqrt{2^{l-H_{min}^{\varepsilon'}(K_A|E)}}, \quad (3.44)$$

where  $D(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1$  is a trace distance<sup>4</sup>, and  $\rho_U = \frac{1}{|\mathcal{Z}|} \sum_{u \in \mathcal{Z}} |u\rangle \langle u|$  is the maximally mixed state over the space of possible keys  $\mathcal{Z}$ . Also note that in the lemma is smooth min-entropy  $H_{min}^{\varepsilon'}(K_A|E)$ . With the smoothing parameter  $\varepsilon'$ , we can define length of bits  $l$ , which are  $\varepsilon$ -secure, denoted as  $\varepsilon_{sec}$ , as

$$l \leq H_{min}^{\varepsilon'}(K_A|E) + 2 - \log\left(\frac{1}{\varepsilon_{sec} - 2\varepsilon'}\right), \quad (3.45)$$

where we used the Quantum Leftover Hash Lemma. The  $\varepsilon$ -security is a measure if the resulting key deviates at most by the value of  $\varepsilon$  from the perfect key. A typical value for practical realization is  $\varepsilon_{sec} = 10^{-10}$  [36].

To summarize, the post-processing stage in Quantum Key Distribution protocols involves three steps after the quantum phase to estimate and correct errors, as well as enhance security. It starts with parameter estimation, where Alice and Bob assess the error rate in a random subset of the transmitted key using statistical bounds like Serfling's inequality. Next, error correction techniques like the Cascade protocol are employed to locate and correct errors in the key while minimizing information exposure to potential eavesdroppers. Finally, privacy amplification using randomness extractors like universal hash functions is performed to eliminate any residual information an eavesdropper might have gained, resulting in a final secure key shared between Alice and Bob.

---

<sup>4</sup>Trace distances quantifies distinguishability between two states

## 4 Security of QKD

By the security of a quantum key distribution protocol, we understand if the protocol was secure even despite imperfections of the devices used and due to noise.

Firstly, we will discuss how a secure classical communication using *One-time pad* is established with the derived key from the QKD protocol. Alternatively, one may use AES symmetrical encryption, which requires only 128, 192, or 256 bits of the key. AES with a 256-bit key can be assumed to be secure even against quantum computers in near future. AES is based on Galois fields, however, we won't discuss its working principle since it is above the scope of this publication. More information can be found for example in [37].

Then we will discuss the definition of security and introduce a *photon number splitting attack*. Finally, we will describe three protocols resistant to that attack.

### 4.1 One-time pad

The One-time Pad (OTP) is the only known encryption that is unconditionally secure. However, it requires a pre-shared key, which can be used only once and has to be at least as long as the message. The key also has to be truly random, which severely limits its usage in classical communication, but not in quantum communication.

In QKD we usually use the Vernam cipher. It's based on the bitwise exclusive-or (XOR) operation, which takes two bits as an input and the result is a modulo 2 addition operation performed on these two bits. So in our case, Alice have a message  $m$ , converted into a bit string, that she wants to encrypt. She also has a bit key  $k$  she gained from a successfully performed QKD. To encrypt the message, she uses the bitwise XOR to add the message with the key. The result will be the encrypted message  $e$ . To summarize:

$$e = m \oplus k. \quad (4.1)$$

Bob then receives the encrypted message through classical channel. To decrypt it, he uses the shared key

$$e \oplus k = m \oplus k \oplus k = m. \quad (4.2)$$

As we see, adding the key to the encrypted message results in the original plaintext message  $m$  [38]. This simple operation provides absolutely secure encryption.

Since we stated that OTP is unconditionally secure, one may still ask if it is possible to break the cipher by frequency analysis or brute-force attack. And the answer is no. By applying Vernam cipher, we get uniformly distributed letter occurrence, therefore frequency analysis cannot be used. Moreover, the attacker cannot brute-force the encrypted messages because all possibilities are equally probable.

To ensure true randomness, Alice has to use a truly random generator to generate randomly the states she is sending to Bob. In the classical world, we use pseudo-random generators based on random processes. The simple ones use algorithms with an initial seed. Better generators use complex processes such as atmospheric noise or radioactive decay, but it is still not fully random, we only have not enough knowledge to predict it. On the other hand, quantum physics offers truly random generation. The simplest model is a source of single photons, which are sent to a 50:50 beam splitter, where we then measure if the photon went through or was reflected. In practice, it is a little bit more complicated because this experiment is device-dependent. For example, the beam splitter will never have a splitting ratio of exactly 50:50, and if we replace the beam splitter with another one, the generation will be slightly different. Therefore device-independent generators are being developed.

The main bottleneck of the OTP is the requirement that the message size cannot exceed the key size, and that each key can only be used once. Although it is theoretically possible to transmit keys in real-time over a quantum channel, thereby continuously generating new keys, this approach is constrained by the key bit rate of the channel. In practical applications, different symmetric encryption algorithms are commonly used to overcome these limitations. For instance, the widely adopted Advanced Encryption Standard (AES) is considered a robust choice [39]. This resilience is attributed to the fact that the only feasible quantum attack is a brute-force attack using Grover's algorithm, which merely halves the effective security level. Thus, AES with a key length of 256 bits still ensures an acceptable security level of 128 bits, which remains secure by contemporary standards.

## 4.2 Practical limitations

In practical realizations of QKD, several limiting factors exist, some of which have been exploited to perform attacks on QKD systems, as described in more detail in Section 4.4. These limiting factors include the generation of weak coherent pulses instead of single photons, fiber losses, quantum channel noise, detector efficiency, and dark counts on detectors [40].

The generation of weak coherent pulses, used in place of single photons, involves employing strongly attenuated laser pulses to approximate single-photon emissions. This approach and its vulnerabilities, particularly in the context of the Photon Number Splitting (PNS) attack, are further discussed in Section 4.4.2.

Fiber losses (transmittance) can be modeled as

$$\eta = 10^{(-\beta l + c)/10}, \quad (4.3)$$

where  $\beta$  represents the attenuation in decibels per kilometer,  $l$  is the transmission length, and  $c$  denotes a distance-independent constant loss in optical components.

Quantum channel noise, which may arise from a variety of sources (e.g., thermal fluctuations, stray light, or vibrational effects), can lead to phase-flip errors or phase deviations, as outlined in Section 2.4.2. Additionally, when utilizing entangled states, the phase-damping channel must be considered, as it reduces the entanglement between two photons [41].

Lastly, single-photon detectors face three primary constraints: detector efficiency, dark counts, and detector’s dead time. Detector efficiency is the probability that the detector will click if a photon reaches it, whereas dark counts refer to the unwanted detector clicks in the absence of photon incidence, caused by thermal fluctuations or stray counts. A third significant parameter is the detector’s dead time, the time required to reset the detector after a click. These error rates can be partially mitigated, for example, by cooling the detectors [26, 42].

In summary, the detection of a photon on Bob’s side relies upon a series of conditional probabilities: Will there be at least one photon in the weak coherent pulse? Will this photon travel from Alice to Bob without being lost due the fiber losses? Will the information encoded in the photon remain unchanged? And finally, will the detector correctly identify the presence or absence of a photon? All these scenarios are modeled in the simulation part of this thesis in the Chapter 5.

### 4.3 Security model of QKD

The security of Quantum Key Distribution is evaluated based on how closely the practical implementation of the protocol aligns with an ideal version of QKD. The protocol must demonstrate two key properties:  $\varepsilon$ -correctness and  $\varepsilon$ -secrecy.

The  $\varepsilon$ -correctness ensures that the keys generated by Alice ( $K_A$ ) and Bob ( $K_B$ ) are identical within a small error margin defined by  $\varepsilon$ . Specifically, the probability that their keys differ should be less than or equal to  $\varepsilon$ . If  $\varepsilon$  is small enough, it’s assumed that their keys are effectively the same.

The  $\varepsilon$ -secrecy measures how much information an eavesdropper (Eve) could potentially learn about the key. It compares the state of the system when Eve is eavesdropping to a scenario where the key is ideal—perfectly random and unknown to Eve. The metric used here is the trace distance, which quantifies the distinguishability of two states. The secrecy condition ensures that this distance is less than or equal to  $\varepsilon$ . The final measure of security ( $\varepsilon$ -security) combines both correctness and secrecy, asserting that both conditions are met to a degree determined by  $\varepsilon$ . This definition of security is called *universal security*.

We will start with  $\varepsilon$ -correctness. We say that a protocol is  $\varepsilon$ -correct if the probability that Alice’s and Bob’s key  $K_A, K_B$  differ less than by a constant  $\varepsilon \geq 0$ :

$$P(K_A \neq K_B) \leq \varepsilon. \quad (4.4)$$

Thus if the inequality hold for small enough  $\varepsilon$ , we can assume that Alice's and Bob's key are the same.

Definition of  $\varepsilon$ -secrecy is slightly more complex. Assume we have an ideal key, which is uniformly, randomly, distributed and independent of any eavesdropper. Suppose Alice's key is  $K_A$  with probability distribution  $p_{K_A}$ . Additionally, we have a space  $\mathcal{S}$  made of all possible sequences of the key, i.e, of the bit string. Then we denote Eve's state for particular key  $k_A \in \mathcal{S}$  as  $\rho_E^{k_A}$ . Alice's and Eve's classical-quantum joint state is described by

$$\rho_{K_A E} = \sum_{k_A \in \mathcal{S}} p_{K_A}(k_A) |k_A\rangle \langle k_A| \otimes \rho_E^{k_A}, \quad (4.5)$$

where  $\{|k_A\rangle\}_{k_A \in \mathcal{S}}$  is an orthonormal basis of a Hilbert space  $\mathcal{H}_{k_A}$ . We take into account two cases, if Alice and Bob abort the protocol, then Eve's state is the same for real and ideal protocol and it is pointless to continue. That happens with probability  $p^\perp$ . If the protocol is not aborted, the ideal (for Alice) Alice's and Eve's composite state is

$$\rho_{K_A E}^{\text{pass}} = \rho_U \otimes \rho_E, \quad (4.6)$$

where  $\rho_U = \frac{1}{|\mathcal{S}|} \sum_{u \in \mathcal{S}} |u\rangle \langle u|$  is a fully mixed state on  $\mathcal{H}_U$ , thus its entropy is the highest.

If the protocol was not aborted, we define the  $\varepsilon$ -secrecy as a case when the distinguishability between the perfect and the real state is lower than a non-negative constant  $\varepsilon$ , so

$$(1 - p^\perp) \frac{1}{2} \left\| \rho_{K_A E}^{\text{pass}} - \rho_U \otimes \rho_E \right\|_1 \leq \varepsilon. \quad (4.7)$$

Finally, we define the protocol as  $\varepsilon$ -secure if the protocol does not abort and the following inequality holds for all  $\varepsilon \geq 0$

$$(1 - p^\perp) \frac{1}{2} \left\| \rho_{K_A K_B E}^{\text{pass}} - \rho_{UU} \otimes \rho_E \right\|_1 \leq \varepsilon, \quad (4.8)$$

where state  $\rho_{K_A K_B E}^{\text{pass}}$  is the joint state of Alice's, Bob's and Eve's systems after a QKD protocol and  $\rho_{UU} = \frac{1}{|\mathcal{S}|} \sum_{u \in \mathcal{S}} |u\rangle \langle u| \otimes |u\rangle \langle u|$  is a fully mixed state on  $\mathcal{H}_U^2$ . Thus the key is universally secure if the real situation is  $\varepsilon$ -close to the ideal situation [6].

## 4.4 Attacks on QKD

In this section, we address practical vulnerabilities in Quantum Key Distribution systems by focusing on specific types of eavesdropping attacks. Real-world imple-

mentations of QKD often vary from theoretical models due to technological limitations and imperfections in quantum devices and channels. Notably, the generation of single photons, which is essential for QKD, presents significant challenges. Ideal single-photon sources are not yet feasible for widespread use due to their complexity and high cost. Consequently, QKD systems commonly employ heavily attenuated laser diodes to create weak coherent pulses, introducing inherent imperfections in the photon source that can be exploited by attackers.

We begin our discussion with the eavesdropping techniques that arise from these imperfections. The chapter details three principal attacks: the Intercept-Resend attack, Photon Number Splitting attack, and Detector Control attack. Each attack exploits different aspects of system vulnerabilities—ranging from the quantum state disturbance during interception, to leveraging multi-photon emissions in weak coherent pulses, and to manipulating detector responses under non-ideal conditions. By examining these attacks, we aim to highlight the challenges in preserving the integrity and security of QKD systems against sophisticated eavesdropping strategies that take advantage of the gap between theoretical security and practical implementation.

Certainly! Here's a revised and clarified version of your text on the intercept-resend attack in the context of the BB84 protocol, incorporating the corrections and improvements discussed:

#### **4.4.1 Eavesdropping**

The fundamental intercept-resend attack was briefly discussed in Section 3.2.1. Here, we delve deeper into how this attack influences the mutual information in the BB84 protocol and how it affects the quantum states. In the intercept-resend attack, Eve gains information about the key by measuring the qubits sent from Alice to Bob and then sending replacement qubits to Bob based on her measurement results.

Eve intercepts the qubit sent by Alice before it reaches Bob. She measures the intercepted qubit in one of the two bases ( $X$  or  $Z$ ), chosen at random. After measuring the qubit, Eve prepares a new qubit in the state corresponding to her measurement result and sends this qubit to Bob. Bob receives the qubit from Eve (thinking it is from Alice) and measures it using his chosen basis. When Eve measures the qubit, she introduces disturbance. In the BB84 protocol, if Eve chooses the correct basis to measure the qubit, she gets the correct bit value. If she chooses the incorrect basis, she has a 50% chance of obtaining the wrong result, thus increasing the quantum bit error rate to 25% on average, because she chooses the wrong basis half of the time and each wrong basis choice leads to a 50% error rate. The random variables  $A$ ,  $B$ , and  $E$  represent Alice's bit, Bob's received bit, and Eve's measured bit, respectively.

For perfect key distribution without eavesdropping, the mutual information of Alice's

and Bob's key bits is

$$I(A : B) = H(A) - H(A|B) = H(A) = H(B) = 1, \quad (4.9)$$

where  $H(A) = H(B) = 1$  denotes that Alice and Bob have the same bit, and  $H(A|B) = H(B|A) = 0$  implies that Alice and Bob's bits are perfectly correlated.

When Alice and Bob are eavesdropped by Eve, it introduces a disturbance in the QBER. If the error rate introduced by Eve is  $e'$ , where  $e' = 0.25$ , the mutual information between Alice and Bob decreases as:

$$I(A : B) = 1 - H(e'), \quad (4.10)$$

where  $H(e')$  is the binary entropy function. The mutual information between Alice and Eve,  $I(A : E)$ , is nonzero since Eve gains some information about the transmitted qubits. Specifically, Eve has complete information (1 bit) about the bits when she guesses the correct basis, which is 50% of the time. Otherwise, she gains no information about the bit because of quantum uncertainty. Hence, the average mutual information between Alice and Eve is:

$$I(A : E) = 0.5 \text{ bits.} \quad (4.11)$$

So, on average, Eve gains 0.5 bits of information for each intercepted bit. This does not account for any additional information Eve might have due to error correction procedures Alice and Bob might undertake, which could potentially leak further information if not done with privacy amplification in mind.

The second discussion will be in terms of single quantum states, to show the disturbance made by eavesdropping and highlight that it is not possible to eavesdrop without creating noise.

Suppose Alice sends two non-orthogonal states  $|\psi_0\rangle, |\psi_1\rangle$ . Eve uses an ancilla state  $|E\rangle$  with a unitary operation  $U$  to extract some information from Alice's states. The result is that Alice's states remain the same after the operation and Eve's state  $|E'\rangle$  will contain some info about Alice's state

$$U |\psi_0\rangle |E\rangle = |\psi_0\rangle |E'_0\rangle, \quad (4.12a)$$

$$U |\psi_1\rangle |E\rangle = |\psi_1\rangle |E'_1\rangle. \quad (4.12b)$$

If we take the scalar product of the left sides and the right sides, we get

$$\langle \psi_0 | \psi_1 \rangle \langle E | E \rangle = \langle \psi_0 | \psi_1 \rangle \langle E'_0 | E'_1 \rangle. \quad (4.13)$$

The scalar products  $\langle \psi_0 | \psi_1 \rangle$  cancel out and we are left with

$$\langle E | E \rangle = \langle E'_0 | E'_1 \rangle, \quad (4.14)$$

however,  $\langle E | E \rangle = 1$  and thus also the right hand side has to be equal to 1,  $\langle E'_0 | E'_1 \rangle = 1$ , which implies that  $|E'_0\rangle$  and  $|E'_1\rangle$  are the states. Because we applied a unitary operator  $U$  on two different states and got the same states, we conclude that these states do not carry any information about Alice's states.

The only case, when the states  $|E\rangle$  can contain part of the information about Alice's states is if her states are affected by the unitary operation  $U$ . In other words, Alice's states will be altered by Eve's attempt to gain some information about them. Suppose Alice's altered states are  $|\psi'_0\rangle$  and  $|\psi'_1\rangle$ , then

$$U |\psi_0\rangle |E\rangle = |\psi'_0\rangle |E'_0\rangle, \quad (4.15a)$$

$$U |\psi_1\rangle |E\rangle = |\psi'_1\rangle |E'_1\rangle. \quad (4.15b)$$

And we again take the scalar product of each side

$$\langle \psi_0 | \psi_1 \rangle \langle E | E \rangle = \langle \psi'_0 | \psi'_1 \rangle \langle E'_0 | E'_1 \rangle. \quad (4.16)$$

The more the states are orthogonal, the more are they distinguishable. Therefore Eve wants to maximize the scalar product  $\langle E'_0 | E'_1 \rangle$ . However, the bigger the scalar product of  $E$ , the smaller the scalar product of Alice's altered states, which means greater disturbances introduced by Eve. Hence, Eve can gain information about Alice's state without raising the error rate.

#### 4.4.2 Photon number splitting attack

Photon number splitting (PNS) aims at the incapability of generating always only single photons. While single-photon sources are available, they are quite costly and still not perfect. Thus in QKD, instead of single photon sources, a laser attenuated to a single photon level is utilized. Such photon generation is called *weak coherent laser pulses*.

Laser is source of coherent states  $|\alpha\rangle$ , hence one can find relation between coherent and Fock states  $|n\rangle$  as

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{+\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle, \quad (4.17)$$

where  $\alpha$  is a complex phase related to the mean photon number  $\mu$  by  $|\alpha|^2 = \mu$ . The number of photons in each pulse follows a Poisson distribution  $P_\mu(n)$ . The emitted



quantum state  $\rho_\mu$  is represented by [17]

$$\rho_\mu = \sum_{n=0}^{\infty} \frac{\mu^n e^{-\mu}}{n!} |n\rangle \langle n| = \sum_{n=0}^{\infty} P_\mu(n) |n\rangle \langle n|. \quad (4.18)$$

For a source attenuated to an average photon number  $\mu = 0.1$ , the typical photon distribution in the coherent state is:

$$|\alpha\rangle = 0.951 |0\rangle + 0.301 |1\rangle + 0.067 |2\rangle + 0.012 |3\rangle + \dots, \quad (4.19)$$

which yield measurement probabilities:

$$p_{|0\rangle} = 0.905, \quad p_{|1\rangle} = 0.090, \quad p_{|2\rangle} = 0.0045, \quad p_{|3\rangle} = 15 \times 10^{-5}, \dots \quad (4.20)$$

This distribution implies that about 90% of the pulses are empty (vacuum state), 9% contain exactly one photon, and approximately 1% contain two or more photons. Although the probability of generating two or more photons may seem minimal, it becomes significant when considered in the context of non-empty pulses. Specifically, out of the pulses that contain at least one photon (approximately 10% of all pulses), about 10% of these (or 1% of the total number of pulses) actually contain two or more photons. This means that approximately 10% of non-empty pulses pose a potential security risk, as they allow for the possibility of a PNS attack. This ratio highlights the importance of accurately assessing photon distributions in QKD systems to safeguard against vulnerabilities.

When we use a noisy channel, we define the probability that Bob receives a non-vacuum state as

$$P(|n\rangle \neq |0\rangle) = 1 - e^{-\mu\eta}, \quad (4.21)$$

where  $\eta$  is a *single-photon transmittance*, which is the probability that a single photon does not get lost during the transmission through a imperfect quantum channel [6]. The transmittance is defined by the equation (4.3).

Since Eve is limited only by the laws of physics, she can potentially build a photon number quantum non-demolition (QND) measurement device to count transmitted photons. When she sees a vacuum state, she let it pass to Bob. When Eve detects only one photon is being sent, she applies intercept-resend strategy. However, when she notices that two or more photons are being transmitted, she can split them and keep only one of them. Then she can either measure it without raising QBER or if she owns a quantum memory, she can store this qubit and wait until Alice publicly announces her basis. Eve can then measure correctly her stored qubits.

In a PNS attack, the eavesdropper could exploit multi-photon pulses. Using quantum non-demolition measurements, she could determine the number of photons in each pulse without disturbing them. If multiple photons are present, she can split off one

photon, storing it for later measurement once the basis is revealed in key sifting, while sending the remaining photons to the legitimate receiver (Bob). This strategy allows her to gain information without being detected, as it results in average a QBER below the theoretical limit of the intercept-resend attack.

The security of QKD against PNS attacks can be enhanced, for example, by using decoy states—pulses (introduced in section 4.6.2) with randomly varied intensities—to detect and quantify the presence of eavesdropping activities.

## 4.5 Detector control attack

This type of attack exploits vulnerabilities in single-photon avalanche detectors (SPADs), which facilitate the detection of single photons. They are introduced in more detail in the Experiment chapter 6.1.3. SPADs operate optimally in Geiger mode, ideal for detecting single photons. In this mode, a reverse-biased voltage is applied, and the absorption of a photon triggers an avalanche of ionization, resulting in a detectable increase in voltage, known as a "click." After detection, the voltage is lowered to quench the avalanche, preparing the detector for subsequent photons.

However, when exposed to light pulses exceeding a certain threshold intensity  $P$ , SPADs switch from Geiger mode to linear mode, where they lose sensitivity to single photons and become "blinded" by the intense light. This vulnerability can be exploited by an attacker, such as Eve, who can control the outcome of Bob's measurements by using bright pulses just above the threshold  $P$ .

In such a scenario, Eve conducts an intercept-resend attack using these bright pulses instead of single photons. Specifically, if Bob employs a phase-encoded BB84 protocol, and if he selects a different basis than Eve, his detectors will both click due to the non-interfering bright pulse being split by the beam splitter, leading him to discard these measurements as invalid. Conversely, if Bob selects the same basis as Eve, the bright pulse interferes constructively, ensuring that only one of his detectors clicks. This manipulation allows Eve to determine which Bob's detector will click, therefore perfectly aligning her measurements with Bob's without detection [43].

Moreover, intense light pulses can significantly degrade the detection efficiency of SPADs. Pulses with power ranging from 0.3 to 0.5 watts can permanently reduce the detection efficiency by 80% to 90%. This creates an imbalance where one detector is more likely to click than the other. Increasing the pulse power to between 1.2 and 1.7 watts can permanently force the SPAD into linear mode, where it remains fully controllable by Eve. Pulses exceeding 2 watts can irreversibly damage the detector [44].

## 4.6 Modified QKD protocols

These protocols are designed to cope with imperfect realization of the QKD systems. All these three protocols aim to be secure against the PNS attack. As we showed in the previous section, QKD protocols, while theoretically secure, can be vulnerable to attacks when implemented with imperfect devices. In particular, the photon number splitting attack exploits the fact that practical light sources sometimes emit pulses containing multiple photons.

To address this vulnerability, several modified QKD protocols have been developed. These protocols aim to maintain security even in the presence of imperfect devices and the threat of the PNS attack.

In this section, we will discuss three such protocols: the SARG04 protocol, the decoy state protocol, and the differential phase shift (DPS) protocol. Each takes a different approach to mitigating the risk of the PNS attack. SARG04 modifies the classical communication in the BB84 protocol to leak less information to an eavesdropper. The decoy state protocol uses additional "decoy" pulses with different photon number statistics to detect PNS attacks. And the DPS protocol encodes information in the relative phase between pulses, which is disrupted if an eavesdropper performs a PNS attack.

By cleverly modifying the original BB84 protocol, these schemes provide a practical way to perform secure QKD even with imperfect devices. In the following sections, we will examine each protocol in more detail to understand how they work and why they are effective against the PNS attack.

### 4.6.1 SARG04 protocol

This protocol was developed and named after its founders Scarani-Acin-Ribordy-Gisin in year 2004, it aims to solve the photon number splitting attack by Alice not publicly disclosing info about used basis [45]. While it has a different name, it is basically BB84 protocol with the difference in what Alice announces via public channel. In BB84, she announces all her measurement basis, which leaks a lot of information. Therefore, in SARG04 protocol, Alice shares with Bob only two possible states that Bob can measure, in such way that Bob will know to either remove the measurement or keep it. Assume Alice and Bob are again using  $X$  and  $Z$  basis. Alice sends qubit  $|0\rangle$  to Bob and on a classical channel, she announces two states  $|0\rangle$  and  $|+\rangle$ , which can Bob potentially measure. One of the states has to be the state she sent and the second state has to be the one of the two states from the second basis. Bob now measures the state received from Alice. If he chooses basis  $Z$ , he will always measure  $|0\rangle$ . On the other hand, if he chooses  $X$  basis, he has probability  $1/2$  to measure either  $|+\rangle$  and  $|-\rangle$ . Notice that in this case, Bob will never measure state  $|1\rangle$ . When he measures  $|0\rangle$  or  $|+\rangle$ , he does not know with certainty if he measured

correctly and he has to mark them as invalid. Only when Bob measures state  $|-\rangle$ , he can conclude that Alice had to send  $|0\rangle$  and Bob adds bit 0 to his key. All possible states during the SAGR04 protocol are denoted in Table 5.

Sent state	Announced states	Valid measurement
$ 0\rangle$	$ 0\rangle,  +\rangle$	$ -\rangle$
$ 1\rangle$	$ 1\rangle,  +\rangle$	$ -\rangle$
$ +\rangle$	$ 0\rangle,  +\rangle$	$ 1\rangle$
$ -\rangle$	$ 0\rangle,  -\rangle$	$ 1\rangle$

Table 5: Possible states during SAGR04 protocol. Note that there are always two possible combinations of announced states. For example, when Alice sends state  $|0\rangle$ , we can also announce  $|0\rangle$  and  $|-\rangle$ . Then the only one valid measurement will be state  $|+\rangle$ .

#### 4.6.2 Decoy states protocol

The Decoy states protocol [46, 47] is a variant of the BB84 protocol, designed to counter the Photon Number Splitting attack. While the security of BB84 relies only on encoding the information onto individual photons, decoy states introduces an additional layer of security by utilizing the photon number distribution of the signal and decoy pulses.

The standard BB84 protocol typically utilizes weak coherent pulses that might occasionally contain more than one photon. If multiple photons are present, an eavesdropper could potentially perform a PNS attack by splitting these photons and retaining one for measurement without disturbing the transmission, as described in detail in section 4.4.2.

In contrast, the Decoy states protocol uses the same basis systems as BB84 but transmit three different pulses: signal, decoy and vacuum. The pulses vary in intensities, defined by the mean photon number  $\mu$ . Typically, decoy states have highest photon number  $\mu_{decoy}$ , thus they contain more pulses in average. Signal pulses have lower  $\mu_{signal}$  and vacuum pulses ideally contain no photons  $\mu_{vacuum} \approx 0$ . Thus

$$\mu_{decoy} > \mu_{signal} \mu_{vacuum}. \quad (4.22)$$

Additionally, Alice has predefined probabilities with which she will randomly generate each state  $p_d$ ,  $p_s$ , and  $p_0$ .

The core concept of this protocol is that the eavesdropper cannot efficiently distinguish between signal and decoy pulses. Thus, when performing the PNS attack, Eve will more likely destructively measure the signal pulses than the decoy pulses, since decoy pulses contain, on average, two or more photons in one pulse. Thus, Eve will

more likely take one photon from these pulses, while with the signal pulses, she will, in more cases, measure only one photon, thus introducing an error.

For example, in the article [48], they performed the decoy states protocol with three states: vacuum states with  $\mu = 0$  and a probability of generation  $p_0 = 0.25$ , signal states  $\mu = 0.2$ ,  $p_s = 0.25$ , and decoy states  $\mu = 0.6$ ,  $p_d = 0.5$ . The resulting probabilities of the number of photons in one pulse for the signal state are

$$p_{|0\rangle} = 0.819, \quad p_{|1\rangle} = 0.164, \quad p_{|2\rangle} = 0.0045, \quad p_{|3\rangle} = 0.001, \dots, \quad (4.23)$$

and for the decoy state

$$p_{|0\rangle} = 0.549, \quad p_{|1\rangle} = 0.329, \quad p_{|2\rangle} = 0.099, \quad p_{|3\rangle} = 0.020, \dots \quad (4.24)$$

These equations show that in the pulses, there will still mostly be either one or two photons, thus Eve is unable to recognize decoy states from signal states.

To determine if an eavesdropper has interfered with the transmission, Alice and Bob employ statistical analysis to compare the rates of detected decoy and signal pulses. They analyze the discrepancies in detection rates against expected outcomes calculated based on the quantum channel conditions, mainly by predicting the number of received photons by Bob based on the fiber length and attenuation, used mean photon numbers, and probabilities of generation of each pulse [49]. They also still calculate the QBER and verify if it is below the secure threshold. Any significant deviation from the expected photon statistics suggests potential eavesdropping, indicating that Alice and Bob should abort the protocol and start over.

In summary, the Decoy states protocol enhances the security of the BB84 protocol by introducing decoy pulses with a higher mean photon number than the signal pulses. This makes it challenging for an eavesdropper to distinguish between signal and decoy pulses, reducing the effectiveness of the Photon number splitting attack. By carefully analyzing the detection rates of decoy and signal pulses, Alice and Bob can detect potential eavesdropping and ensure the security of their quantum communication.

### 4.6.3 Differential phase shift

This *Differential phase shift* (DPS) protocol is also resistant to the PNS attack. We start again with Alice, who generates weak coherent photon pulses with  $\mu$  around 0.2, as in a classical QKD protocol. These pulses are generated in time intervals of length  $\tau$ . She then modifies the phase of the photon by either 0 or  $\pi/2$ , i.e., she generates two non-orthogonal states. Bob has a one-pulse delay Mach-Zehnder interferometer, where the arms of the interferometer are in such length that if the first pulse travels through the longer arm and the second pulse travels in the shorter arm, they will meet at the beam splitter and they can interfere with each other. I.e., The difference

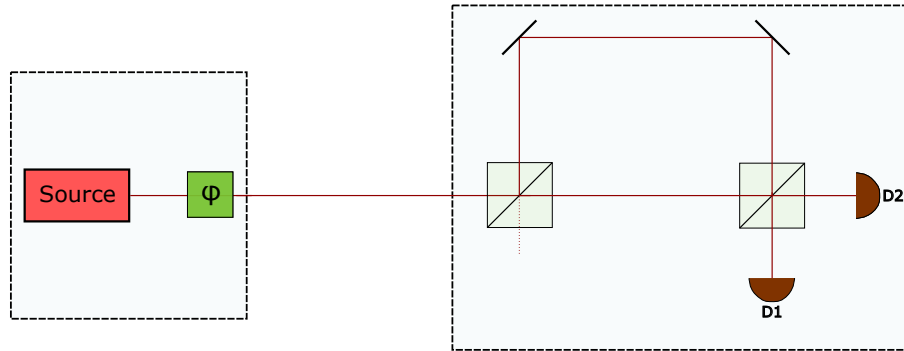


Figure 4.1: Setup for the Differential phase shift protocol. Alice (on the left) owns a laser source and a phase modulator with phase delay  $\varphi$ . She generates pulses in time intervals of length  $\tau$ . Bob (on the right) has a asymmetrical Mach-Zehnder interferometer with path difference equal to the time between two consecutive pulses  $\tau$ . If the two pulses interfere at the second beam splitter at Bob, a probabilistic click on the detectors will occur.

of the travel times in the arms is equal to  $\tau$ . The interference occurs when the two photons were modulated the same. If both photon were modulated with phase  $\varphi = 0$ , and the first one went through the longer arm, while the second one took the shorter arm at Bob's interferometer, they will constructively interfere and the detector 1 will click deterministically with probability 1. For the same scenario, but with phase modulation of  $\varphi = \pi/2$ , the detector 2 click with probability 1. After the transmission, Bob tells Alice via classical channel the result and the time of the detection. This information is enough for Alice to derive the same key as Bob [50]. The configuration of the protocol is depicted in Figure 4.1.

This protocol is secure against the PNS attack because the information is encoded to the phase difference between pulses. While Eve can use the non-destructive measurement to catch multi-photon pulses, it would disrupt the coherence at the beamsplitter, which results in measurement errors. Therefore the PNS attack can be detected [49].

#### 4.6.4 Coherent one-way protocol

The Coherent One-Way (COW) protocol [51, 52] is specifically designed to be simple and practical for implementation, especially with existing telecommunications infrastructure.

Alice uses a mode-locked laser and intensity modulator to generate a series of light pulses at a mean photon number  $\mu$  and a fixed time interval  $\tau$ . The protocol encoding uses pairs of two states: a pulse with non-zero photon number and a vacuum pulse with zero photon number. Each logical bit of information  $|0\rangle$ ,  $|1\rangle$ , is encoded into

pairs of pulses:

$$|0\rangle = |\sqrt{\mu}\rangle |0\rangle, \quad (4.25a)$$

$$|1\rangle = |0\rangle |\sqrt{\mu}\rangle. \quad (4.25b)$$

The pulses from Alice are not mutually orthogonal. These pairs travel through a quantum channel to Bob, whose setup incorporates a beam splitter with a specific transmission coefficient  $t_B$ . This beam splitter divides the incoming pulses into two paths: one leading to the data line for key generation and the other to the monitor line for eavesdropping detection.

In the data line, the pulses are detected by a single photon detector. Here, Bob distinguishes between the two bit values based on the time slot in which the photon is detected. Specifically, a photon detected in the first time slot of the pair indicates a bit value of  $|0\rangle$ , corresponding to the pulse configuration  $|\sqrt{\mu}\rangle |0\rangle$ . Conversely, a photon detected in the second time slot indicates a bit value of  $|1\rangle$ , corresponding to the configuration  $|0\rangle |\sqrt{\mu}\rangle$ .

The monitor line features an interferometer equipped with two detectors, designed to detect any deviations in the coherence of the incoming pulses. This setup is critical for identifying the presence of an eavesdropper. Under normal operation without eavesdropping, the interference pattern caused by consecutive coherent pulses should result in a predictable detection pattern. If two consecutive pulses are both non-empty (non-empty means that the light pulse from the pulse pair contains a photon) constructive interference should lead to a detection event at only one specific detector (usually  $D_{M1}$ ) due to the phase relationship of the pulses.

The protocol also employs decoy states, which consist of two consecutive coherent states  $|\sqrt{\mu}\rangle |\sqrt{\mu}\rangle$ . These states are used to statistically analyze discrepancies between the expected and actual pulse detection rates, providing a method to detect eavesdropping. If an eavesdropper (Eve) attempts to perform a quantum non-destructive measurement to count photons, it will disrupt the phase coherence between these pulses, altering the expected interference pattern. Since all the pulses are equally-spaced, the coherence of both decoy and bit sequences can be verified with a single interferometer.

Coherence is quantitatively assessed using the visibility metric, defined as:

$$V = \frac{p(D_{M1}) - p(D_{M2})}{p(D_{M1}) + p(D_{M2})}, \quad (4.26)$$

where  $p(D_{M1})$  and  $p(D_{M2})$  are the probabilities of detection at detectors  $D_{M1}$  and  $D_{M2}$ , respectively. High visibility indicates undisturbed coherence, while lower visibility suggests potential eavesdropping activities.

This allows Bob to differentiate between bit values based on detection location and detect eavesdropping by observing changes in interference patterns caused by decoy

states. This dual functionality enhances the security and reliability of the COW protocol in practical quantum key distribution systems.

In a paper practically implementing this protocol, they set mean photon number  $\mu = 0.5$  and on Bob's side, they used a 90/10 beam splitter that sends 90% of the photons directly to a single photon detector (bit or data channel). The remaining 10% of the photons went through a monitor channel - an unbalanced Michelson interferometer and were subsequently detected [53]. This protocol is used by the commercial device Cerberis XG of the company ID Quantique [54].



## 5 Simulation of phase-time BB84 protocol

In this chapter we present our simulation of the BB84 protocol with time and phase encoding with many customizable parameters, which are listed in the table 6. The purpose of this simulation is to emulate the operation of an advanced commercial QKD device and explore various stages and challenges involved in a practical QKD system such as tackling noise and eavesdropper detection, as our experiment was limited by time and resource constraints.

In a real-world QKD experiment, the implementation of a full-functioning system can be challenging due to the complex hardware and infrastructure requirements. This simulation aims to provide a comprehensive and flexible environment to study the behavior of a QKD system under various conditions, including noise, losses, and potential eavesdropping attacks.

Variable Name	Description	Default/Possible Values
num_of_bits_sent	Number of bits sent in the simulation	100 000
protocol_type	Type of QKD protocol used	BB84 or Decoy states BB84
signal_mu	Mean photon number for signal pulses	0.2
decoy_mu	Mean photon number for decoy pulses	0.6
decoy_vacuum_mu	Mean photon number for vacuum decoy pulses	0.01
signal_prob	Probability of choosing signal pulses	0.25
decoy_prob	Probability of choosing decoy pulses	0.5
decoy_vacuum_prob	Probability of choosing vacuum decoy pulses	0.25
fiber_length	Length of the optical fiber used	50 km
att_db_per_km	Attenuation of the fiber per kilometer	0.2 dB/km
det_efficiency	Efficiency of the photon detector	0.5
phase_noise_std_deviation	Standard deviation for phase noise	0.003
phase_flip_probability	Probability of a phase flip occurring	0.04
dark_count_probability	Probability of a dark count	0.005
noise	Noise settings in the simulation	0 = none 1 = noise+losses 2 = only transmission losses 3 = only phase noise
eavesdropping	Eavesdropping scenario in the simulation	0 = no eavesdropping 1 = Intercept-Resend attack 2 = PNS attack
error_correction_alg	Error correction algorithm used	Cascade or Winnow

Table 6: Customizable parameters of the QKD simulation

The simulation covers the entire QKD process, starting from the generation of quantum states (pulses) by Alice, followed by the transmission of these pulses through a simulated quantum channel, and their subsequent measurement by Bob. The code incorporates various features and techniques that are essential for a practical QKD system, such as decoy state protocols, error estimation, error correction, privacy amplification, and key authentication. Furthermore, the simulation allows for the investigation of two different attack scenarios: intercept-resend attack and photon-number splitting attack. By simulating these attacks, the code provides insights into their impact on the final QBER and the effectiveness of countermeasures.

Overall, this simulation can serve as a valuable tool for researchers and students working in the field of quantum communication, providing insights into the behavior of QKD systems under various conditions. We made the code for the simulation publicly available on the Github website [55]. It can also be used for further development and extension of this simulation to contain more advanced attacks, error-correction techniques or various QKD protocols.

## 5.1 Quantum transmission

We will now describe the quantum phase of the simulation. Since the simulation is running on a classical computer, no part of the simulation is quantum, but the probabilistic behaviour of quantum word is simulated using pseudo-random generators. The experiment architecture is the same as in a typical phase-time BB84 protocol, as depicted in Figure 3.5 or as in the Decoy states protocol, where a intensity modulator is added at Alice's side.

The simulation begins with the generation of quantum pulses by Alice. Each pulse represents a quantum state, and its properties are defined by the Pulse class. The pulses are generated with a specific mean photon number, determined by the Poissonian distribution, which is a common model for describing the statistics of attenuated laser pulses. The number of generated pulses depends on the parameter `num_of_bits_send`. There are two modes of generation, either generation of single photons, simulated with  $\mu = 0.9$  using the `__init__` method of Pulse class

```
1 pulse = Pulse(seq_num=i, mu_poisson=mu, pulse_type=pulse_type)
```

or generate  $n$  photons for decoy state protocol using the method

```
1 generate_pulses(n, decoy_states=True, signal_mu=0.2, decoy_mu=0.6,
   decoy_vacuum_mu=0.01)
```

with  $\mu$  values based on practical implementations [48].

The phase of the pulse is encoded according to the BB84 protocol, where Alice randomly chooses one of two non-orthogonal bases (X or Y) and prepares the pulse in one of the four possible states  $(0, \pi/2, \pi, 3\pi/2)$  corresponding to the chosen basis.

```
1 Alice_delays, Alice_basis = generate_bits_and_delays_for_Alice(n=
   num_of_bits_send)
2 simulate(pulse, phase=Alice_delays[i])
```

These pulses are then transmitted through a simulated quantum channel, where they are affected by noise (phase-flip channel and phase decoherence channel) and by fiber losses. The phase decoherence is model as a phase change from Gaussian distribution with mean 0 and standard deviation 0.05. This can be also viewed as a simulation of imperfect phase modulator. The phase-flip channel is modeled by a parameter `phase_flip_probability` and then the photon has 50% probability of

being flipped either by  $+\pi$  or  $-\pi$ . Finally, the fiber loss is simulated using the Beer-Lambert law, which models the attenuation of the optical signal as it propagates through the fiber with two parameters, fiber length and fiber attenuation, as

```

1 simulate_phase_noise(pulse, std_dev=phase_noise_std_deviation)
2 simulate_phase_flip(pulse, phase_flip_probability=0.01)
3
4 def attenuation(distance_km, attenuation_db_per_km=0.2):
5     return 10 ** (-(attenuation_db_per_km * distance_km) / 10)

```

where the attenuation of 0.2 dB/km is set as typical fiber attenuation for telecommunication wavelengths [56]. In the code, attenuation is implemented as a loop, where the probability that a multiphoton pulse loses a photon is conditioned on the attenuation experienced by the previous photons in the pulse. Each iteration represents a chance for a photon to be lost, continuing until the current photon in the pulse is not lost due to the losses, then no more photons are lost or until no more photons are left.

```

1 def simulate_fiber_losses(pulse, distance_km, attenuation_db_per_km
2     =0.2):
3     while True:
4         if np.random.rand() >= attenuation(distance_km,
5             attenuation_db_per_km) and pulse.photon_number > 0:
6             pulse.photon_number -= 1 # One photon is lost
7         else:
8             break

```

After simulating the quantum channel, the pulses are processed by Bob's measurement apparatus. The simulation models Bob's measurement process, which uses the asymmetrical Mach-Zehnder interferometer. The `interference_simulator` function simulates the interference pattern observed at the detectors, depending on the phase of the pulse and the relative delay between the two arms of the interferometer. It takes as an input the pulse of the Pulse class, which carries information about which path it took and about its current phase. It also contains the simulation of adding phase from Bob's phase modulator and then evaluating the interference based on these properties. These detections are then converted into bit values based on which detector clicked.

```

1 interference_simulator(pulse, bob_delay=Bob_delays[i],
2     detector_efficiency=det_efficiency, dark_counts=
3     dark_count_probability)

```

The simulation takes into account detector efficiency and dark counts. Detector efficiency models the probability of a photon being successfully detected, while dark counts simulate the scenario where the detector clicks even in the absence of an incoming photon, introducing errors in the measurement process. The detector efficiency was set to 0.5 for all simulations.

After the simulation of the quantum transmission phase, Alice and Bob now have

their noisy bit keys, which are going to be evaluated and corrected in the post-processing phase.

## 5.2 Post-processing

Following the quantum transmission and measurement phases, the simulation proceeds to the classical post-processing stage, which is crucial for establishing a secure and error-free symmetric key between Alice and Bob.

Key sifting is the first step in the classical post-processing phase, where Alice and Bob discard the bits corresponding to their measurement bases that did not match. Thus as explain in chapter, we will discard all measurements whose outcome was random, and hence useless for key generation. In the simulation, the `key_sifting` function compares Alice's and Bob's bases and identifies the shared bases and their corresponding sequence numbers. Only the pulses with matching bases and successful detection by Bob are included in the sifted key.

```
1 Bob_sifted_key, confirmed_seq_nums = key_sifting(pulses=all_pulses,
        Alice_bases=Alice_basis, Bob_delays=Bob_delays)
2 Alice_sifted_key = alice_sift_key(Alice_bits, confirmed_seq_nums)
```

Due to the presence of noise and to detect potential eavesdropping attacks, the sifted keys held by Alice and Bob may not be identical. To address this issue, the simulation employs error estimation and correction techniques.

The `compare_random_subset` function is used to estimate the quantum bit error rate (QBER) by comparing a random subset of the sifted keys held by Alice and Bob. This estimation is performed while taking into account the statistical bounds defined by Hoeffding's and Serfling's inequalities, to ensure the accuracy of the QBER estimation.

```
1 Alice_final_key, Bob_final_key, QBER = compare_random_subset(
        Alice_sifted_key, Bob_sifted_key, percentage=25, delete_bits=
        True)
```

Once the QBER is estimated, the simulation employs error correction algorithm to reconcile the differences between Alice's and Bob's keys. The code implements two different error correction protocols: Cascade and Winnow. The `cascade_protocol` function implements the Cascade algorithm, iteratively correcting errors until the keys are reconciled. Alternatively, the Winnow protocol can be used.

```
1 corrected_key = correct_key(Alice_final_key, Bob_final_key, QBER
        /100, version='cascade')
2 corrected_key, total_errors = winnow_protocol(Alice_final_key,
        Bob_final_key, QBER)
```

After error correction, Alice's and Bob's keys should be identical, assuming the QBER is within the error correction capability of the chosen protocol. However,

before using the reconciled key for secure communication, it is essential to authenticate the key to ensure its integrity that the key is the same for Alice and Bob. The simulation implements key authentication by comparing the hashes of Alice's and Bob's keys on a public channel using the `compare_hashes` function.

```
1 compare_hashes(Alice_final_key, corrected_key)
```

Finally, the simulation performs privacy amplification, which is a crucial step in QKD protocols to remove any potential information leaked to an eavesdropper during the quantum transmission and error correction phases. The `privacy_amplification` function applies a hash function (in this case, SHA-3) to the corrected key.

```
1 Alice_final_shared_key = privacy_amplification(Alice_final_key)
2 Bob_final_shared_key = privacy_amplification(corrected_key)
```

The resulting final shared keys, obtained by Alice and Bob after the privacy amplification step, can be used for secure communication, as demonstrated in the simulation by encrypting and decrypting a sample message using the AES algorithm.

```
1 send_encrypted_msg(Alice_final_shared_key, Bob_final_shared_key)
```

```
2
```

```
3 Output:
```

```
4 Alice message: "Hello, Bob! How are you?"
```

```
5 Encoded message: b'\xca\xa4\xaa\t\x1c...\xa3R\xef\xcc'
```

```
6 Bob decoded message: "Hello, Bob! How are you?"
```

Finally, the `QKDSimulationResults` class stores and records various statistics and results generated during the execution of the QKD simulation. It provides methods to capture relevant data at different stages of the simulation, such as pulse counts, sifted key length, quantum bit error rate estimation, error correction details, and final key length. This class facilitates the convenient retrieval and analysis of the simulation results, enabling the printing of summary information and the export of data for further processing or visualization purposes.

This is the end of the simulation. From now on, Alice and Bob can communicate securely even if someone with an access to a high-performing quantum computer is eavesdropping on their communication.

### 5.3 Simulation results

Throughout the entire simulation, we conducted 100 iterations, each with 100 000 sent qubits unless specified otherwise. The results are then the average over these iterations. For the estimation of QBER, 25% of the bits from the sifted key were utilized. In practical application, this proportion would be dynamically adjusted based on the length of the sifted key and the computed Serfling's bound. While

25% is necessary for shorter keys to ensure a reliable QBER estimate, it becomes inefficient for larger key lengths as illustrated in Figure 3.8.

### 5.3.1 QBER analysis for different scenarios

In each iteration, the perfect scenario without noise or eavesdropping resulted in an expected QBER of 0. We then introduced noisy channels, simulating phase-flip and phase deviation errors separately, with values ranging from 0 to 0.1. These simulations are depicted in Figures 5.1.

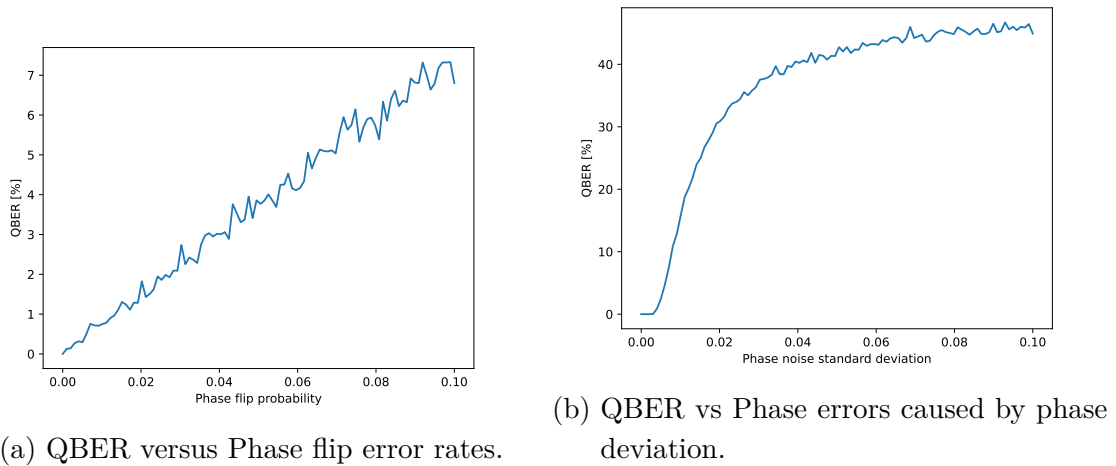


Figure 5.1: QBER variations due to phase errors under different noise conditions.

From these measurements, we set the parameter settings for phase noise standard deviation at 0.003, and for phase flip probability at 0.04:

```
1 phase_noise_std_deviation = 0.003
2 phase_flip_probability = 0.04
```

The resulting QBER was  $3.13 \pm 0.59\%$ , which is a value we tried to achieve to align with practical implementations where error rate is typically between 3-4% [47]. We set the dark count probability to 0.005, contributing a QBER of  $1.29 \pm 0.33\%$ . The total QBER was then  $4.30 \pm 0.69\%$ .

For noise-less channel with an eavesdropper, the average QBER measured was  $25.04 \pm 1.01\%$  with minimum and maximum values of 22.46% and 26.90%, respectively. These figures confirm the theoretical prediction that an eavesdropper guessing the correct base results in correct transmission approximately 75% of the time resulting in 25% QBER in average. The minimum value lower than 25% suggests that the eavesdropper could occasionally guess more accurately, though this deviation diminishes with larger data quantity.

When running the experiment with the simulation of a photon number splitting attack, the average QBER decreased to  $21.7 \pm 1.05\%$  with minimum and maximum

values of 19.29% and 24.1%, respectively. So with the PNS attack, the QBER never exceeded 25% and the average was around 3.3% lower than with Intercept-Resend attack. This showcases the danger of PNS attack, because it may be unnoticed by Alice and Bob.

Note that in real-life scenario, Alice and Bob has no way how to tell the origin of the noise and if it was caused by an eavesdropper or not. Thus, they cannot subtract the hypothetically known noise error from the total error. They classify only if the QBER is lower than their set limit that have to be lower than the theoretical limit of 25%. That is also the reason behind the pulse analysis in decoy states protocol. Alice and Bob gain new method how to measure if they were eavesdropped without relying solely on the QBER value.

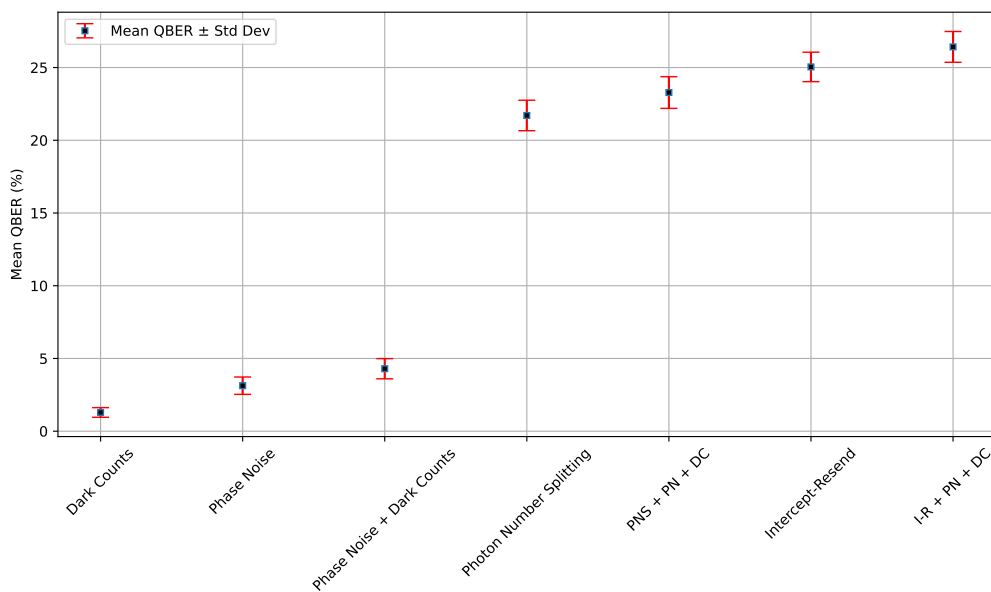


Figure 5.2: Mean QBER with Standard Deviation across various QKD Scenarios. PN denotes Phase Noise, and DC denotes Dark Counts.

We then performed the simulation again, now with the set noise (channel noise and dark counts) of  $4.30 \pm 0.69\%$  in average. The error for I-R attack was  $26.42 \pm 1.10$  with minimum QBER of 24.01 and maximum QBER of 28.26. For PNS attack with noise the QBER was  $23.28 \pm 1.09$  with minimum and maximum values of 21.10 and 26.00, respectively. All the simulation results are denoted in Figure 5.2 and in the Table 7.

In summary, we simulated and analyzed the impact of how different types of noise and eavesdropping affect the QBER in a QKD protocol. We started with a perfect scenario to establish a baseline QBER of zero and then introduced phase errors and dark counts to simulate real-world conditions. Our results matched expected QBER values revealed how different factors increase QBER and at which rate. We also examined how eavesdroppers affect QBER, especially under intercept-resend and

Scenario	Mean QBER (%)	Minimal value (%)	Maximal value (%)
Dark counts	1.29	0.56	2.20
Phase flip	3.13	1.78	4.45
Phase noise + Dark counts	4.30	2.67	6.25
Photon number splitting (PNS)	21.71	19.29	24.10
PNS + Phase + Dark	23.28	21.09	26.00
Intercept-resend (I-R)	25.04	22.46	26.90
I-R + Phase + Dark	26.42	24.01	28.26

Table 7: Simulation results showing the mean, minimum, and maximum QBER for various scenarios.

photon number splitting attacks, highlighting their potential to go undetected in the PNS attack.

### 5.3.2 Key length analysis

In the next part of the simulation, we focused on the core value of QKD protocol: the distilled key length. This value can be called as a *key bit rate* when we talk about the distilled key length per unit of time. Since our simulation is not performed in real-time, we examine the key length the key sifting and after the parameter estimation with respect to the number of sent photons. The parameters for these simulations were the same as for the previous section so the number of sent photons was 100 000 and we iterated 6 times for fiber length step size of 2 km.

We started by simulating the probability that a photon sent by Alice will reach Bob for the single photon scenario and for the decoy states protocol. We performed measurement for fiber lengths from 0 to 200 km for values of two fiber attenuation: 0.2 dB/km and 0.3 dB/km. The results are denoted in Figure 5.3. We continue the measurements with attenuation set to 0.2 dB/km.

in Figure 5.4, we examine the impact of dark counts on the QBER as fiber length increases. With longer fiber length, photon losses intensify, leading to a higher QBER and reduced final key size. This increase in QBER is due to a diminishing number of received photons compared to the constant rate of dark counts, effectively increasing the relative proportion of dark counts. Despite the increase in QBER, which approaches 50% at around 150 km, there is still non-zero final key length. This is thanks to the error correction algorithms that corrects even the noisy key. However, this level of error correction potentially compromises the secrecy of the key by leaking significant amount of information.

The Figure 5.5 further proves that the error caused by dark counts increases with longer fiber length. In this particular Figure, we tried to examine the influence of dark counts and dark counts with the noise. Since the noise is set independently on the fiber length, the difference is visible for the small fiber lengths and then the QBER caused by dark counts takes over.



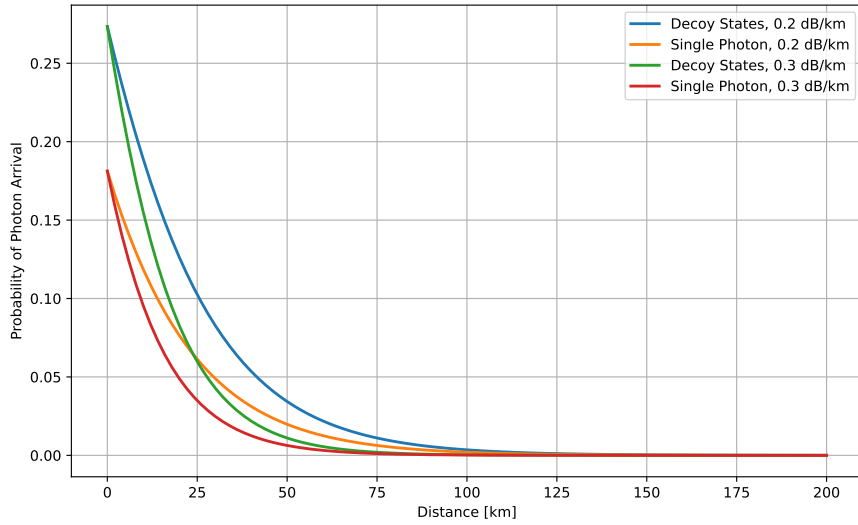


Figure 5.3: Comparison of photon transmission probabilities over distance for classical BB84 and Decoy states BB84 protocol. This graph displays two and two curves: two for decoy states (blue with fiber attenuation of 0.2 dB/km and green line with attenuation 0.3 dB/km) and two for the single photon protocol (red (0.2 dB/km) and orange line (0.3 dB/km)). The x-axis represents the distance in kilometers (from 0 to 200 km), and the y-axis represents the probability of at least one photon successfully arriving after transmission. The graph demonstrates that the decoy state protocol maintains higher probabilities of photon arrival and allows to transmit over longer distances compared to the single photon protocol.

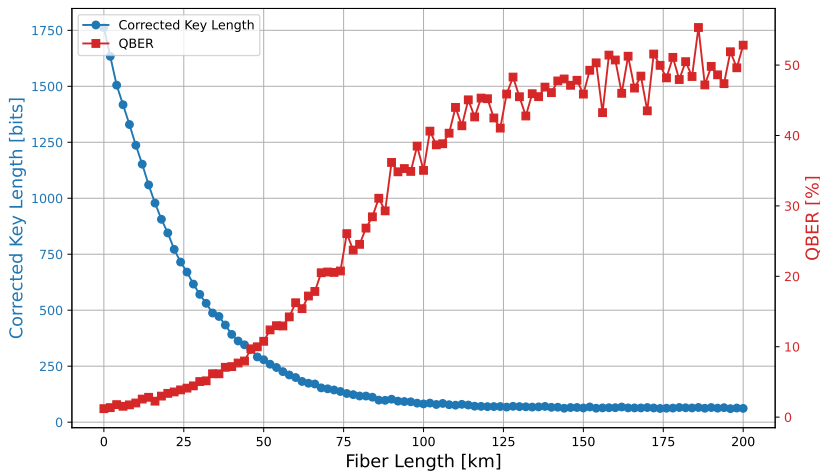


Figure 5.4: Relationship between the final key size and QBER in the sifted key versus fiber length, considering only fiber losses and dark counts.

In Figure 5.6 we present the simulation for three scenarios and compare the final key length after error correction. The first scenario is the blue line, which represents fiber losses together with dark counts. The orange line denotes the fiber losses without dark counts. As can be observed from the graph, for small fiber length (small losses), these two curves are very similar. However as the loss rate increases so does the corrected key length. We see that for length around 100 km, the corrected key length of the orange line is almost zero, showcasing the limited reach of QKD. For single photon BB84 this limit however comes around the fiber length of 50 km. Moreover, the final key length is many times greater for fiber lengths below 50 km. After 100 km, we observe that the simulation with dark counts still produce some key, which is however solely noise on the detectors caused by dark counts and does not present nor viable nor secure key. Note that these values are for 100 000 photon pulses, from which most of them do not contain any photon. Thus the huge difference between the scenario for single photon BB84 and decoy states.

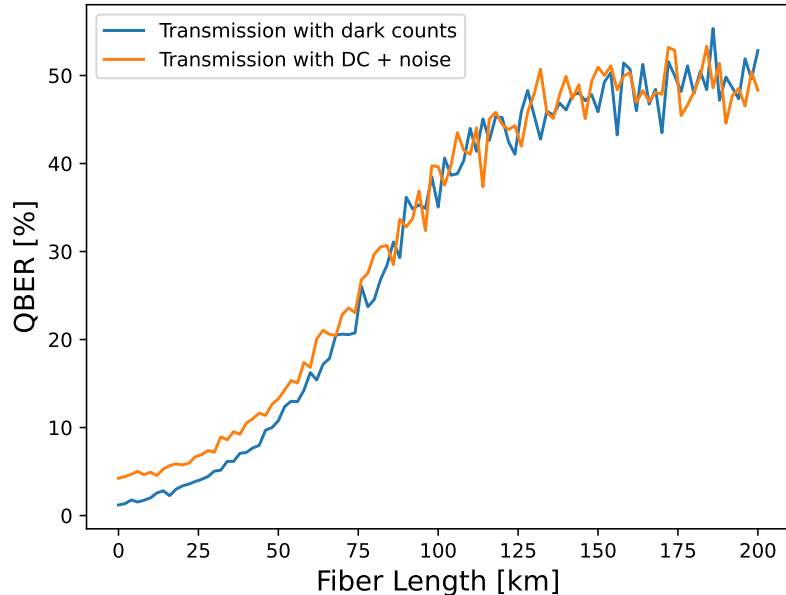


Figure 5.5: QBER in relation with the increasing fiber length for two scenarios: transmission losses with dark counts and transmission losses with dark counts and noise.

In Figure 5.6, we present simulations for three different QKD scenarios to compare the final key lengths after error correction. The blue line, representing fiber losses combined with dark counts, and the orange line, representing only fiber losses, perform similarly at shorter fiber lengths (smaller loss rate). However, as fiber length increases, the key length is exponentially decreasing for both scenarios. The corrected key length for the orange line approaches zero near 100 km, illustrating the practical limitations of QKD reach under these specific conditions. In contrast, the green line representing single photon BB84 protocol reaches its limit around 50 km.

Moreover, the difference between final key length between decoy states protocols and the single photon BB84 is notable, showcasing the benefits of decoy states protocol. The blue line shows some residual key length even at distances over 100 km. This is attributed to noise from dark counts and this derived key is neither viable nor secure. These results highlight the critical differences between using single photon BB84 and decoy state protocols. Note that these values are simulated for batches of only 100 000 photon pulses, from which most of them will not contain any photon. This further enlarges the gap between decoy states and single photon protocols.

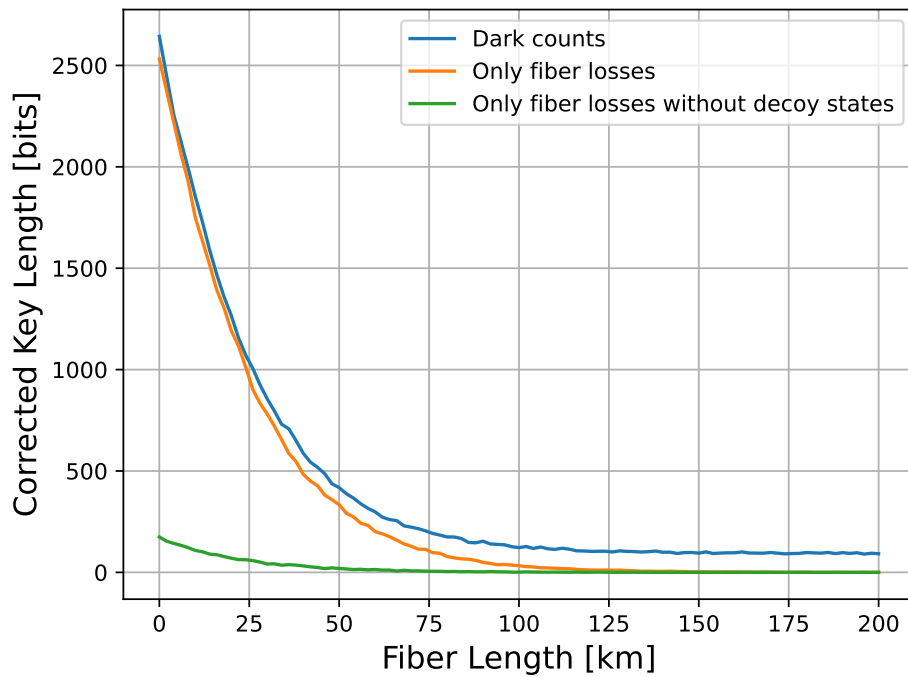


Figure 5.6: Simulation of fiber losses and the resulting key length for three different scenarios. The blue curve represents fiber losses with dark counts, the orange curve simulates only fiber losses without additional noise, and the green curve shows results from fiber losses without using decoy states, transmitting only single photons. The sample size used for QBER estimation was 25% of the sifted key.

## 6 Experimental part

Following the comprehensive simulation of the quantum key distribution protocols discussed in the previous chapter, this section transitions from theoretical models to the practical implementation of the BB84 protocol using phase and time encoding. The simulations provided a theoretical foundation and predicted outcomes under both idealized and real-world conditions. The purpose of the experimental setup described here is to showcase the principles on which the QKD protocols are based on and construct a ready-to-use optical setup for BB84 protocol with time-phase encoding. It should serve as the quantum part of the simulation and further assess the practical challenges of implementing QKD. The experiment also aimed to demonstrate the feasibility of the BB84 protocol using low resources. Utilizing a off the shelf microcomputer Red Pitaya STEMLab 125-14, the setup involved connecting a series of optical components to construct a quantum cryptography system divided into Alice and Bob and using the microcomputer to control parts of the experiment. This experiment was performed in laboratory at Palacký University in Olomouc.

This chapter begins by describing the experimental apparatus and setup of each component's role and the configuration of the overall setup. This is followed by a step-by-step description of the experimental procedure, providing a method on how to construct a QKD system. Furthermore, the results obtained from these experiments are analyzed, providing insights into the encountered challenges.

### 6.1 Experimental setup description

The experimental configuration implemented the BB84 protocol with time and phase encoding in a unidirectional mode as outlined in the section 3.2.3. The setup was split into two sections: Alice and Bob, linked by an optical fiber. Components of the experiment were controlled using the microcomputer Red Pitaya STEMLab 125-14. The role of this microcomputer was to activate a laser diode, modulate the phase on a phase modulator, and detect photons using single-photon detectors.

The layout of the setup is illustrated in Figure 6.1. In the diagram, Alice and Bob are connected through two channels: a classical channel and a quantum channel. The quantum channel is an optical fiber connecting the optical components of the experiment. The classical channel could also utilize an optical fiber or an Ethernet cable. However, in our experiment, we used only one Control unit and thus eliminating the need for a classical channel.

This modified setup is in Figure 6.2. Alice owns the source, a pulsed laser diode Sharp LT015MD0, which operates at a wavelength of 830 nm and has an average output power of 40 mW. This diode is powered by an Avtech AVO-9 pulsed laser diode driver, which can either independently power the diode or it can be triggered

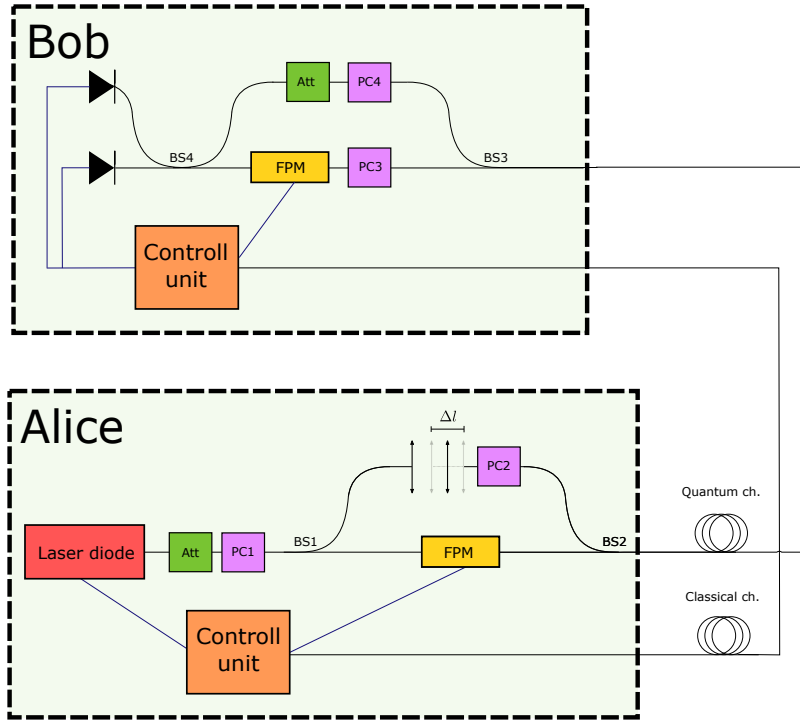


Figure 6.1: The setup designed for this work involved a laser diode, followed by an attenuator (Att) to diminish the laser’s output power to single photons. Prior to reaching the first beam splitter (BS1), a polarization controller (PC) was used to ensure correct input polarization into the fiber phase modulator (FPM), located in the shorter arm of the interferometer. In the longer arm, an air gap ensured that the lengths of Alice’s and Bob’s interferometers were equal, followed by PC2, which is unnecessary when Alice and Bob’s systems are connected together because this connection makes the photons which go through different arms of the Alice’s interferometer indistinguishable in polarization. At Bob’s end, two polarization controllers, PC3 and PC4, are installed - one in each arm of the interferometer. PC3 sets the appropriate polarization for the FPM while PC4 adjusts the polarization to maximize interference visibility. The attenuator is used to enhance indistinguishability by compensating for losses in the phase modulator

via TTL pulses produced by the Red Pitaya STEMLab 125-14. The maximum pulse frequency of the driver is 1 MHz. Photons emitted from the laser are channeled directly into an optical fiber. Initially, photons pass through an attenuator (Att), reducing their intensity to create weak coherent pulses, although this attenuator was omitted during initial experiments with strong pulse signal light. Subsequently, a Thorlabs’ fiber paddle polarization controller (PC) adjusts the polarization to align with a polarization filter at the input of the phase modulator.

At the first 50:50 beam splitter (BS1), the photon beam splits; one path leads to the

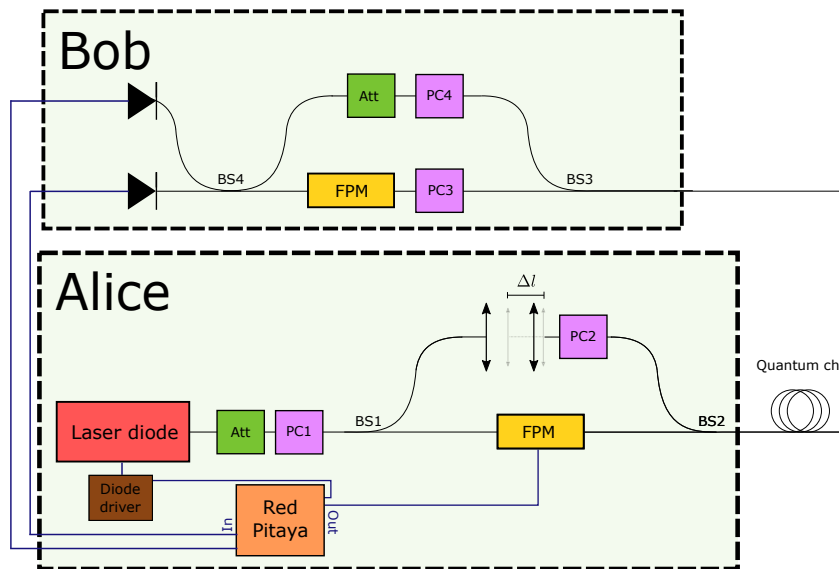


Figure 6.2: The setup actually used in this experiment. Instead of employing two control units, there is a single Red Pitaya at Alice's end that connects to the single-photon detectors and manages both the diode driver and Alice's phase modulator.

shorter arm of the interferometer, passing through a fiber phase modulator (FPM), also referred to simply as a phase modulator (PM). As photons traverse the PM, their phase is altered to a specified value by an applied electric field, controlled by the Red Pitaya. In the longer arm of the interferometer, photons travel through an air gap and PC2. The air gap consists of two lenses separated by air, where the second lens is adjustable in three dimensions. Adjusting the  $x$  and  $y$  coordinates ensures the photon from the first lens reaches the second lens. Variations in the  $z$  direction compensates for the minor differences in the lengths of optical fibers in Alice's and Bob's interferometers, with each fiber varying by millimeters. The  $z$  coordinate ranges from 0 to 16 mm. Adjustments in the  $x$  and  $y$  dimensions are manually made via screws, while the  $z$  coordinate is controlled by an Oriel Encoder Mike Controller, a motor control system utilizing an encoder that applies back electromotive force (back EMF) for smooth modifications. PC2 adjusted the output polarization for a phase where we evaluated Alice's symmetrical Mach-Zehnder interferometer without connection to Bob, as will be described later. However, for the fully integrated setup with Bob, this PC becomes redundant. Ultimately, both arms merge at BS2, and the photon proceeds through the quantum channel to Bob.

Bob's configuration also employs the asymmetrical Mach-Zehnder interferometer. In the shorter arm, Bob utilizes a PC3 and FPM, with PC3 ensuring the correct input polarization for the FPM. The longer arm includes an Att to maintain equal intensity from both arms, thereby ensuring indistinguishability. Additionally, a fourth PC aligns the polarization from the opposite arm to maximize interference visibility. There is a single-photon detector at each exit of BS4. However, for preliminary

measurements without Alice’s attenuator, a single PIN detector replaces them to measure signal power. In the next part, each component is presented and described in more detail.

### 6.1.1 Polarization controller

In our experiment, we utilized Thorlabs’ Fiber Polarization Controllers FPC030, which consists of three paddles. Each paddle winds a specific number of loops that induce stress-related birefringence, forming two or three independent fractional wave plates to modify the polarization in a single-mode fiber wrapped around two or three separate spools. Thus creating the independent fractional wave plates (fiber retarders) [57]. The degree of retardation is determined by the fiber cladding diameter, the diameter of the spool, the number of loops per spool, and the wavelength of light used. By rotating the paddles, we twist the fiber, thereby adjusting its fast axis. We can calculate the retardation as follows:

$$\varphi = \frac{2\pi^2 a N d^2}{\lambda D} \text{ rad}, \quad (6.1a)$$

$$\varphi = \frac{\pi a N d^2}{\lambda D}, \quad (6.1b)$$

where  $a$  represents the material constant,  $N$  is the number of loops,  $d$  is the diameter of the fiber cladding,  $\lambda$  is the wavelength used, and  $D$  is the loop diameter. For silica fiber,  $a = 0.133$  and our polarization controller has a loop diameter of  $D = 27$  mm. The first equation’s units (6.1a) are radians, and the second equation (6.1b) denotes the number of waves, e.g.,  $1/2$  wave or  $1/4$  wave. By setting the number of loops as  $\lambda/4$ ,  $\lambda/2$ , and  $\lambda/4$  wave plates, we can transform any input polarization into any desired output polarization by turning the paddles.

### 6.1.2 Phase modulator

Phase modulators in optical fibers utilize the electro-optic effect in  $\text{LiNbO}_3$  crystals to alter the phase of light passing through the fiber. This effect involves shifts in the electron positions or crystal lattice structures within the material due to an applied electric field, leading to changes in the refractive index of the modulator material.  $\text{LiNbO}_3$  crystals are particularly favored for their low dielectric constant, minimal power consumption, and rapid response capabilities, making them ideal for high-speed optical communications.

The relationship between the refractive index  $n$  of the  $\text{LiNbO}_3$  crystal and the applied electric field  $E$  is given by the expanded equation:

$$\frac{1}{n^2} = \frac{1}{n_0^2} + rE + hE^2 + \dots, \quad (6.2)$$

where  $n_0$  is the original refractive index in the absence of an electric field,  $r$  represents the linear electro-optic coefficient, and  $h$  is a higher-order electro-optic coefficient, accounting for more complex interactions between the field and the crystal structure [58].

During operation, light, typically a coherent laser beam, enters the phase modulator and travels along the optical fiber. As an electric field is applied across the modulator using electrodes, the refractive index changes, affecting the optical path length and causing a phase shift in the light wave. This shift with respect to the changed optical path can be expressed as:

$$\Delta\varphi = \frac{2\pi}{\lambda}\Delta L, \quad (6.3)$$

where  $\Delta\varphi$  is the phase shift,  $\lambda$  is the wavelength of the light, and  $\Delta L$  represents the change in optical path length resulting from the altered refractive index .

The output light from the modulator has a phase altered in accordance with the characteristics of the input electric signal. The typical characteristic of the phase modulator is the half-wave voltage, which is the voltage required to induce a phase shift of  $\pi$ . The phase modulators employed in our experiment were UTP APE PM-0.8-0.5, featuring a half-wave voltage of 1 V at a wavelength of 830 nm.

### 6.1.3 Single-photon avalanche diode

Single-photon avalanche diodes (SPADs) are highly sensitive devices that detect and amplify signals from single photons using the avalanche effect in a semiconductor material. These devices operate by biasing a diode above its breakdown voltage in a state known as Geiger mode, where no current flows until a photon strikes the diode. Upon the arrival of a photon, it interacts with the p-n junction and induces inner photoelectric effect, triggering an avalanche of secondary photoelectric effects through the multiplication of carriers, greatly amplified by the high reverse bias. This process generates a detectable short electric pulse, signaling the detection of the photon and allowing for precise tracking of the time of the incident.

When a photon detection occurs, the detector sends a roughly 2 V pulse through a coaxial cable which is then detected by a electric pulse counter device. This device counts the number of received pulses per unit of time. In our case, this device was either Red Pitaya or Tausand Abacus AB1504 coincidence counter.

Particularly in QKD, the detection efficiency of SPADs over telecommunications wavelengths, ranging from 1350 nm to 1550 nm, is crucial as QKD aims to perform transmissions on these wavelengths. SPADs work relatively well at wavelengths around 400-1000 nm, however, for the longer telecommunications wavelengths, the detectors exhibit lower detector efficiency. For these wavelengths, InGaAs/InP diodes are typically used [26]. For example, in the first paper that used InGaAs/InP for QKD, the InGaAs/InP detectors were cooled to 210 K and achieved



only about 10% efficiency [42]. Although advancements by 2016 improved this to 26% at the same temperature [59]. A more recent development in 2021 saw these detectors operate at a room temperature of +20 °C achieving an efficiency of 20.9% [60]

Since our laser diode operates on wavelength of 830 nm, we could utilize the common single photon detectors. In our experiment, we specifically used EG&G SPCM-AQR-14-FC SPADs. The measured dark count was around 4700 detections per second. The operating wavelength ranges from 400 nm to 1060 nm and the reported detector efficiency was 56.8%.

#### 6.1.4 Microcomputer and programming environment

As previously mentioned, one objective of the experiment was to utilize the Red Pitaya STEMLab 125-14 as the main control unit. This device is captured in Figure 6.3. The programs were developed in the Python programming language, leveraging the Red Pitaya's library `redpitaya.overlay.mercury`. Specifically, the Python code enabled the use of the pre-programmed Field-Programmable Gate Array (FPGA) module to transmit and receive signals through analog inputs and outputs called Radio Frequency (RF) inputs and outputs.

The RF inputs feature two channels with a sampling rate of 125 Msps, where 'sps' denotes samples per second. The analog-to-digital conversion (ADC) resolution is 14 bits, indicating the number of bits used to digitize the input samples. For a 14-bit ADC, this equates to  $2^{14}$ , or 16384 discrete digital levels, with a minimum voltage increment of 61  $\mu$ V. The input impedance is 1 M $\Omega$ , and the buffer size (the total amount of data the FPGA can hold at one time before processing is required) of the FPGA is 16384 bits.

Similarly, the RF outputs include two channels, also with a sampling rate of up to 125 Msps. The digital-to-analog conversion (DAC) resolution is 14 bits, covering a voltage range from -1 V to +1 V. The typical rise and fall time for a 2 V signal is 10 ns.

Parallely with assembling the optical setup, we developed software for the Red Pitaya to generate RF pulses to trigger the laser diode, establish a phase delay at the phase modulator, and count detections from the SPADs. The Red Pitaya is capable of performing these tasks simultaneously. Further discussion on its application in the experiment and any associated limitations will follow later in this chapter.

## 6.2 Stages of the experiment

Our experiment can be divided into 4 parts. The first part was to test the equipment, for example, verify the splitting ratio of each beam splitter and measure the length of each optical fiber to get an estimate on the optimal size of the air gap.

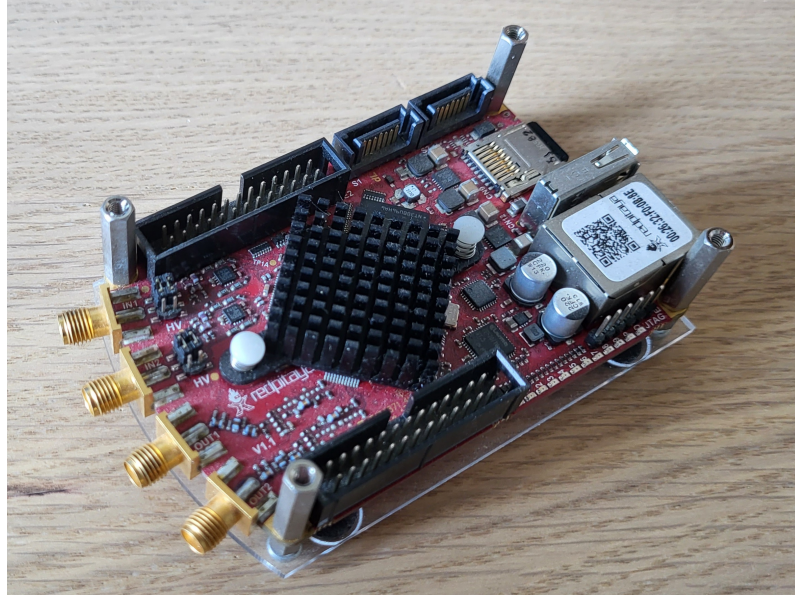


Figure 6.3: Photograph of the Red Pitaya STEMLab 125-14 utilized in our experiment. The golden pins represent the RF inputs and outputs for generating RF pulses.

The second part was to construct a symmetric Mach-Zehnder interferometer, representing Alice without any connection to Bob and verify the interference.

In the third part, we constructed the setup for Bob, connected him with Alice via an optical fiber and measured classical visibility with the help of conventional controlling device. Then we programmed the Red Pitaya to control the experiment on its own and measured the classical visibility of the two asymmetrical MZIs.

In the fourth part, we transitioned to a single-photon level by generating weak coherent pulses and measured visibility of a quantum signal by the detected photon counts.

### 6.2.1 Equipment testing

In the first phase, we remeasured the length of all optical fibers, which will be used in the experiment, and designed how they can be connected such that there is the smallest possible difference in the length of the same parts in Alice's and Bob's interferometer and thus to achieve highest possible indistinguishability and maximal interference visibility by only adjusting the  $z$  coordinate at the air gap. Using a PIN detector, we measured the raw output power of the laser diode without any additional devices. Then we measured the splitting ratio (SR) of each beam splitter to verify if the value given by the manufacturer represents the true value. Our measurements are denoted in Table 8.

We also conducted tests on the digital variable attenuator from OZ Optics, which

	SR by manufacturer	Measured SR
BS1	50:50	55.1:44.9
BS2	50.1:49.9	50.1:49.9
BS3	50:50	50.4 : 49.6
BS4	49:51	48.3 : 51.7

Table 8: Comparison of splitting ratio (SR) for each beam splitter given by the manufacturer and remeasured by us.

was placed immediately after the output from the laser diode at Alice’s part. The attenuator was capable of maximum attenuation of 60.25 dB. We evaluated the relationship between attenuation and power and illustrated these data alongside an exponential fit in Figure 6.4. The graph confirms that the attenuation follows an exponential decay.

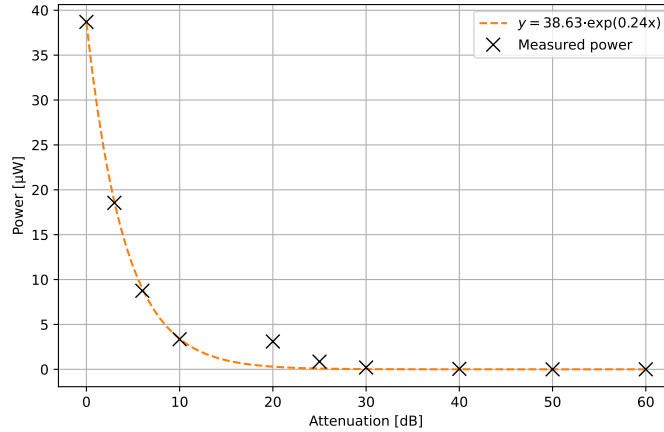


Figure 6.4: Measured power levels at varying attenuation set by the digital attenuator. The data fit well to the exponential decay function  $P = Ae^{-kx}$ , where  $A \approx 38.63$  and  $k \approx 0.243$ . This illustrates that the power output decreases exponentially as attenuation increases.

Furthermore, we calculated the overall attenuation of the system (Alice + Bob) with the digital attenuator set to 0 dB. Initial power measurements at the laser diode (Alice) showed  $P_{in} = 53 \mu\text{W}$ , and the power measured after the last beam splitter (BS4 at Bob) was  $P_{out} = 1.8 \mu\text{W}$ . Given that the beam splitter divides the power equally in two, we doubled the output measurement to estimate the total system attenuation:

$$Att = 10 \cdot \log_{10} \left( \frac{P_{init}}{2P_{out}} \right). \quad (6.4)$$

The calculated attenuation was  $Att = 11.67$  dB. This attenuation can be attributed to losses in the optical fibers and at the connectors of the fibers, imperfect polarization alignment to the polarization modulators, and losses in the air gap.

## 6.2.2 Constructing Alice

We continued to construct the setup of Alice. Specifically, we constructed a symmetrical Mach-Zehnder interferometer, with the two arms having equal optical path lengths of two meters. To have the same length of the two arms is crucial for achieving interference at the output. This whole setup is depicted in Figure 6.5 and a photography of the interferometer is captured in Figure 6.6.

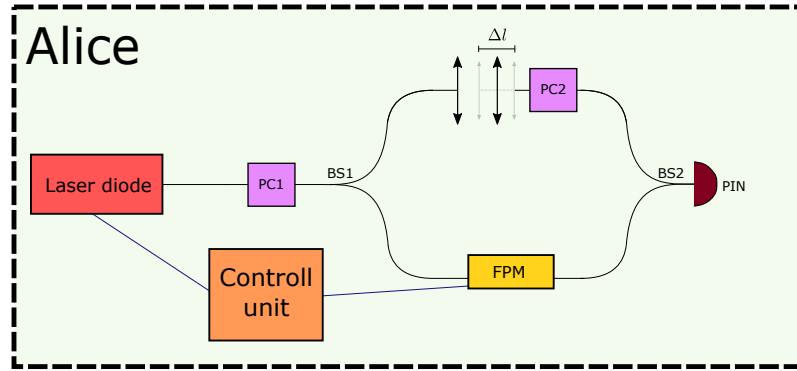


Figure 6.5: Setup used to measure interference at Alice’s symmetrical MZI. It consisted of a laser diode in a pulse mode of operation generating strong signal light, two polarization controllers PC1 and PC2, one fiber phase modulator FPM and one PIN diode as a detector.

Following the laser source, a polarization controller was installed to align the light’s polarization with the principal axis of the phase modulator and thus minimizing the polarization-dependent losses. Then the light beam encounters the BS1. In one arm of the interferometer lies the PM, and in the other arm, an air gap composed of two lenses—the first decouples photons from the fiber into air, while the second focuses the transmitted photons back to the fiber. The precise arrangement of both lenses is critical for minimizing losses for the light beam traveling through the air gap. The  $x$  and  $y$  positions of the lenses are finely tuned for this purpose, and the Oriel motor adjusts the  $z$  coordinate to precisely control the air gap length. By altering the gap length, we dynamically modify the length of one arm of Alice’s interferometer. By doing so, we can equal the interferometer arms to micrometers to optimize the interference at the output.

The second polarization controller PC2, placed after the second arm, is optimized to match the output polarization from the second arm, optimizing the interference at the output. Although PC2 becomes redundant later in the experiment, it was kept in the setup.

To measure the interference, we applied a voltage  $V_A$ , modeled as a sine wave of amplitude  $\pm 2$  V at a frequency of 1 Hz,

$$V_A = 2 \sin(2\pi t). \quad (6.5)$$

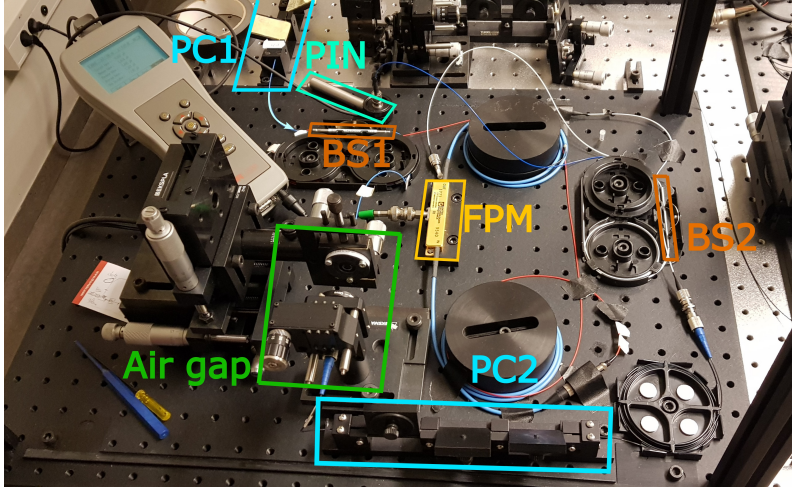


Figure 6.6: Photograph of the Alice's MZI with each component highlighted.

A Python script was used to control the counter-EMF motor to automatize the measurement for the entire range of  $z$  coordinate of the air gap. The visibility of the interference pattern was calculated using the formula:

$$Vis = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}, \quad (6.6)$$

where  $I_{\max}$  and  $I_{\min}$  denote the maximum and minimum intensities recorded by the PIN diode, respectively. The results are presented in Figure 6.7. The maximum visibility achieved was slightly above 25%. The objective of this measurement was not to maximize the visibility but to verify the occurrence of interference over the range of  $z \in \langle 0, 25 \rangle$  mm, which was successful.

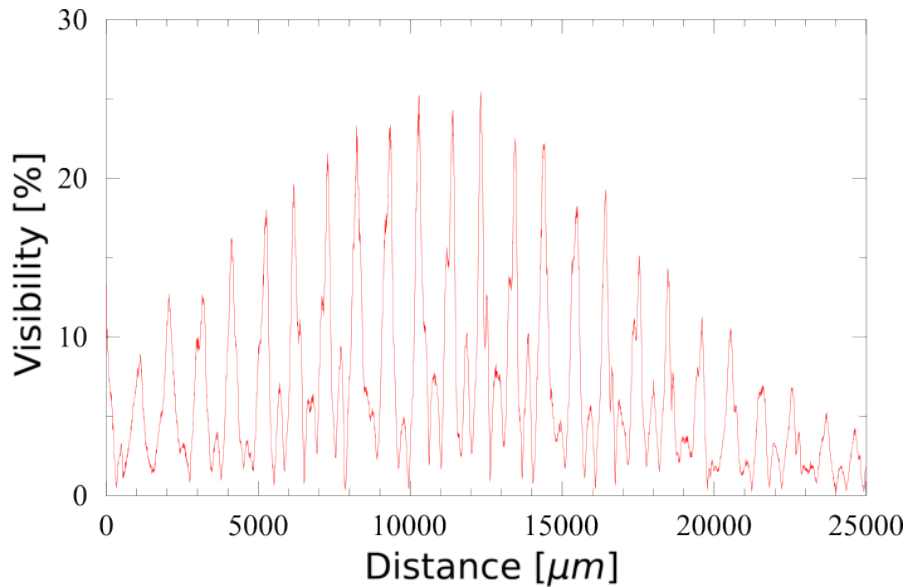


Figure 6.7: Result of visibility measurement for Alice's symmetrical MZI for the whole range of  $z \in \langle 0, 25 \rangle$  mm.

### 6.2.3 Connecting Alice and Bob

In the next stage of our experiment, we reassembled Alice's setup to replace the MZI with an asymmetrical Mach-Zehnder interferometer (AMZI). We proceeded to construct and integrate Bob's setup with Alice's, as depicted in Figure 6.2. Since the stage of the experiment still uses classical signal, the attenuator was omitted at Alice's end. Alice's setup was enclosed within a protective box, with Bob's components assembled atop this box, as depicted in Figure 6.8.

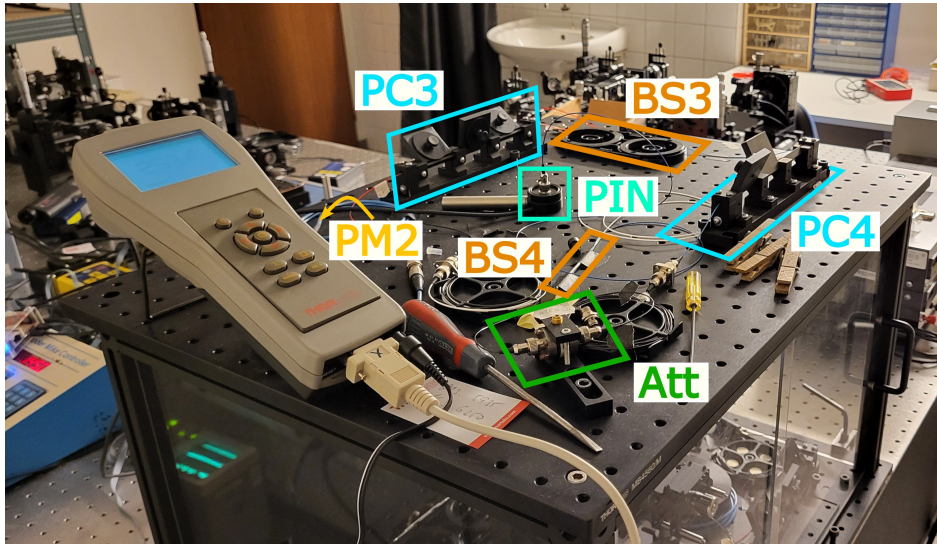


Figure 6.8: Picture of the whole experiment. Alice's equipment is covered in the box and Bob's components mounted on top.

Our goal was to verify the coherence and integration of the two AMZI segments to form a single interferometer. Visibility measurements were taken along the  $z$ -axis (0 to 25 mm), and performed three times. The measurement with the best visibility is depicted in graph 6.9. Each measurement featured main peak and secondary peaks at roughly 2 mm intervals, with the highest visibility recorded at 39.41%. We adjusted the  $z$  coordinate of the air gap to align with the peak of highest visibility and fine-tuned the polarization using PC4 to enhance visibility. The highest visibility reached with the fine-tuning was around 40%.

We noted minor shifts in the interference peak positions. The observed slight shifts could be partially caused by different temperatures in the room, and the instability of the laser diode.

To achieve higher visibility, we had to stabilize the interferometer. Even a small mechanical vibration of the optical fiber can have impact on the visibility. We also re-adjusted the air gap to reduce the signal losses while coupling and decoupling the laser pulse from the fiber. We taped carefully the optical fibers coming from the laser diode to PC1, since even a small shift or twist of the fiber was immediately noticeable, inducing a change of the polarization entering the Alice's FPM, thus

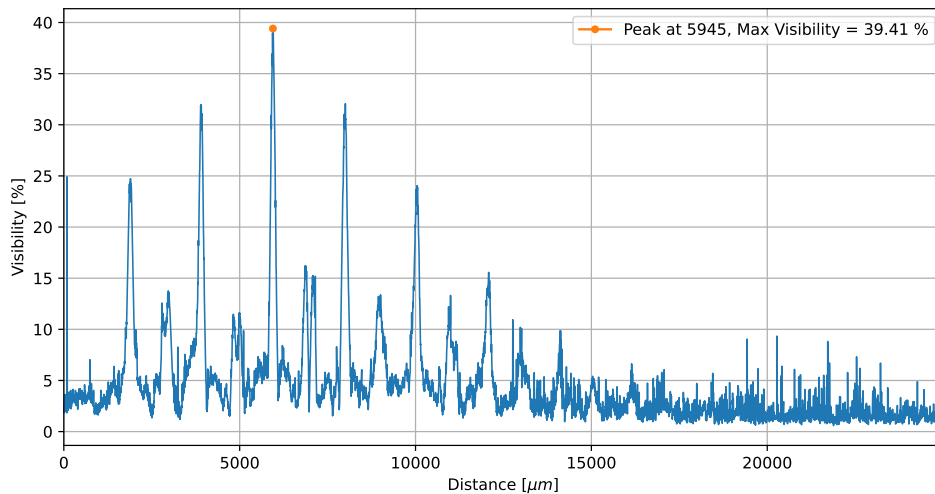


Figure 6.9: Measurement before the stabilization. The peak at  $z = 5945$ , the visibility were later enhanced to value slightly above 40% by configuring polarization at Bob's AMZI using PC4.

reducing the power in this arm. We also covered Bob's setup within a box, similar to Alice.

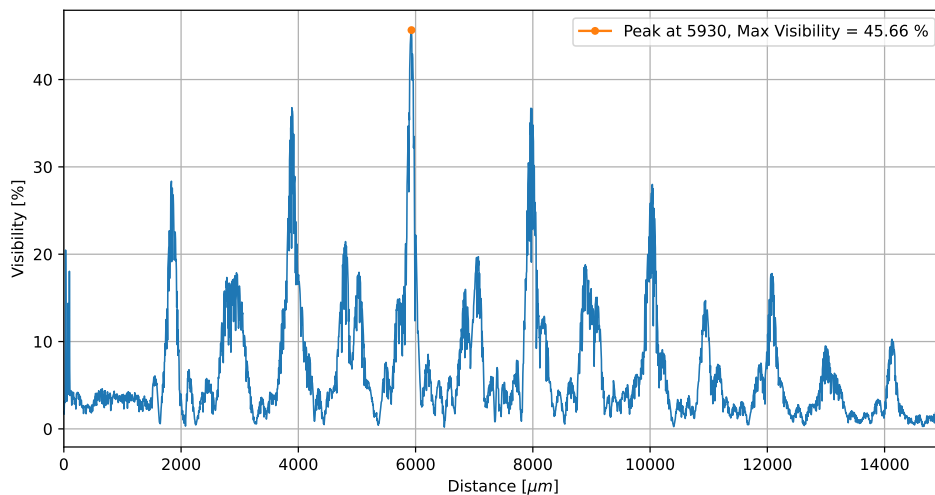


Figure 6.10: Post-stabilization visibility measurement achieved a maximum of 45.66%.

After stabilizing the entire setup, we performed new measurement, and the maximal visibility improved to 45.66%. This visibility pattern is depicted in Figure 6.10. We observe two types of peaks: the sharp, high peaks correspond to the longitudinal modes of the laser diode, while the smaller, wider peaks represent the transverse modes. These patterns demonstrate the interference effects associated with different

modes of the laser's output. This structure with one central peak and with each peak slightly lower shows the interference relation with respect to the path length. The measured classical visibility of 45.66% establishes a robust basis for further single-photon measurements. Notably, this result is very close to the theoretical maximum visibility of 50%, which is limited by the fact that only half of the photons contribute to interference, specifically those that traverse either a short-long or a long-short path through the interferometer.

## 6.2.4 Quantum measurements

Before integrating the single photon detectors into our experimental setup, we measured the dark counts of each detectors using the Tausand Abacus coincidence counter. The first detector, showing dark counts exceeding 1 million counts per second, was labeled as malfunctioning and used only for testing purposes. The second tested detector showed an acceptable dark count frequency of approximately 4.5 kHz.

We also used the dark count detections to try to measure photon counts using Red Pitaya. The Red Pitaya's FPGA module is controlled by code written in C programming language. There is also a Python library built upon the C code to control the FPGA. The detection module has a buffer with size set to  $2^{14} = 16384$  bits.

To initiate the FPGA module of Red Pitaya, we set several parameters. The data rate decimation was set to the smoothest. We configured the system to capture  $\frac{1}{4}$  of the buffer size before the trigger event and  $\frac{3}{4}$  after the trigger. The trigger itself was initialized on a positive edge greater than 0.4 V. The initialization for the Red Pitaya's input '0' in the code was as follows:

```

1 osc = [fpga.osc(ch, 1.0) for ch in range(2)]
2 for ch in osc:
3     # data rate decimation
4     ch.decimation = 1
5     # trigger timing [sample periods]
6     N = ch.buffer_size
7     ch.trigger_pre = N//4 * 1
8     ch.trigger_post = N//4 * 3
9
10 # trigger source is the level trigger from the same input
11 osc[0].sync_src = fpga.sync_src["osc0"]
12 # trigger level hysteresis [V] and edge ['neg', 'pos']
13 osc[0].level = [0.4, 0.5]
14 osc[0].edge = 'pos'
15
16 # start the detector
17 osc[0].reset()

```



During the tests, a sample pulse from a single photon detector was captured by the Red Pitaya, showing a maximal amplitude of 2.40 V (Figure 6.11).

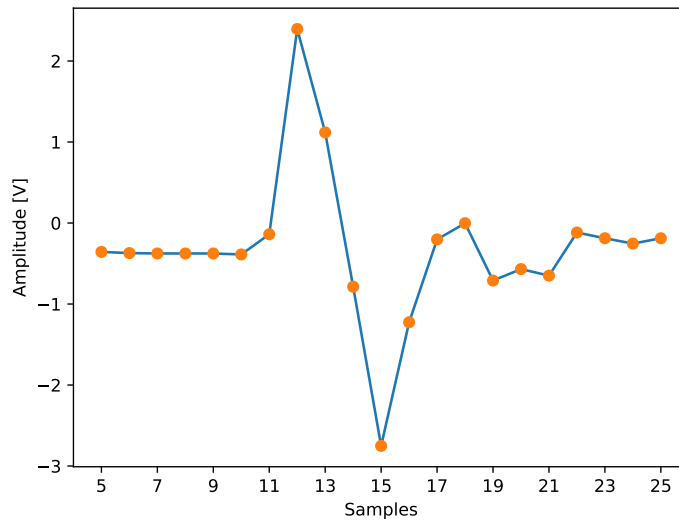


Figure 6.11: Single pulse from the single photon detector captured by Red Pitaya

We developed three different methods for counting photon detections. The simplest method involved waiting for a trigger event and then incrementing a counter. However, this approach proved inefficient for detection rates exceeding a few thousand photons per second due to the need for frequent restarts of the trigger after each detection event. For example, compared to a commercial coincidence counter, this method achieved about 80% of the actual detection rate for dark counts ranging from 4400 to 5000 counts per second. In contrast, for the malfunctioning detector with over 1 million detections per second, the method only detected about 12,000 counts, which is approximately 1% of the expected value.

The initial idea was to implement real-time data acquisition directly from the FPGA to facilitate immediate processing of detections. However, this was not fully realized within the work's timeframe due to several challenging factors. The real-time data reading from the FPGA required a sophisticated data handling and processing framework and the FPGA used in our experiments had a limited on-chip memory, which complicates the continuous streaming of data without intermediate buffering. This approach allowed us to focus on other critical aspects of the work while laying the groundwork for future work aimed at enabling real-time data acquisition from the FPGA.

To overcome these challenges, we delved deeper into the source code to develop a more efficient data retrieval technique using the existing FPGA modules. We

proposed an approach where the entire buffer is read and saved into a variable, repeating this process  $n$  times. We then used the number of detected samples to calculate the total photon count as follows:

$$N_{total} = \frac{\text{sample\_rate}}{n \cdot \text{buffer\_size}} \cdot N_{detected} \quad (6.7)$$

where  $N_{detected}$  is the number of counted pulses in the data from  $n$  buffers, and  $N_{total}$  is the calculated total number of detected photons per second to aggregate the measurements for one second but the overall time to fill the data array was longer time than 1 second. This approach was based on the assumption that the average photon count would remain consistent over multiple samples. Given that the sampling rate of the Red Pitaya is 125 Msps, and based on the observation that one pulse in Figure 6.11 took approximately 10 samples, we estimated the upper limit of detecting up to 12.5 million photon detections per second. So Red Pitaya should not have a problem with 1 million detections per second. We experimented with two versions of the code that employed slightly different data-saving methods, but both yielded similar results.

From the analysis of the malfunctioning detector, we obtained an estimate of roughly 1,213,000 detections per second, which aligns well with the actual count from the commercial counting device. However, this method proved ineffective for lower count rates, such as the dark count of the SPAD, which registered 4,500 dark detections. In such cases, only one or two pulses were present in the entire buffer. Since the detector only saves the buffer upon detecting a pulse, it resulted in a significant overcount, approximately recording twice as many detections ( $\approx 10,000$ ) as were actually present (4,500). Therefore, while this method functions well for hundreds of thousands of detections, it suffers from a lower limit determined by the ratio  $\text{sampling\_rate}/\text{buffer\_size}$  which is approx 7630 photon detections per second. Below this threshold, the method will overestimate the actual number of detections.

Further analysis and code optimization were undertaken by delving into the C source files, which revealed that the Python library we utilized was directly using this source files. So if we rewrote our code into C language, we would be still using the same technique of retrieving the data: *trigger*  $\rightarrow$  *save data*  $\rightarrow$  *iterate* until one second of buffer size data was accumulated. The primary bottleneck remains our inability to read and transmit data from the buffer instantaneously.

Therefore to conclude, Red Pitaya can serve as a cheap variant for a photon counting device but has its limitations. Moreover, for BB84 we would need to perform continuous detection, now the detections were gathered inconsistently from many samples.

After we tested the Red Pitaya on dark detections, we replaced the PIN detector in the experiment setup for the single photon detectors and measured visibility at the

output of BS4 as with the classical signal. Since the detections of Red Pitaya were not completely reliable, we chose to deploy back the Tausand Abacus photon counter, although, Red Pitaya still controlled the phase modulation and laser triggering. We attenuated the signal using the digital attenuator to 60 dB. Firstly, we tested again the number of dark counts on the detector without laser signal to see, if the dark counts frequency differ from the case where the detector was closed. After that we turned on the laser diode attenuated by digital attenuator and counted the photon detections on the detector. Without interference, there was around 450 000 detections per second. We then set the detection window to 100 ms and by using modified script from the previous experiment measurement, we denoted the output from the Abacus coincidence counter and periodically moved the air gap by  $2 \mu\text{m}$ , we repeated the same visibility measurement as for the classical signal. At each position, we measured 100 times the detections per 100 ms. The results are plotted in Figure 6.12.

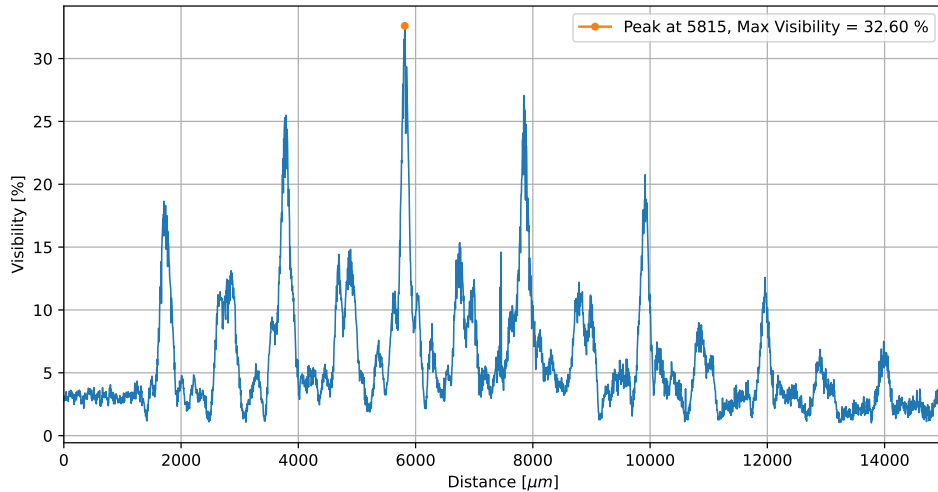


Figure 6.12: Quantum visibility measured by counting photon on the detector.

We achieved an interference of 32.60% with roughly the same maximal peak position. All together, both graphs for classical and quantum visibility were very similar, proving that the experimental setup is indeed working reliably even on a single photon level. The overlap of the classical visibility with quantum visibility can be seen in Figure 6.13.

In our latest experiment, the Red Pitaya was configured to control the entire setup. We adjusted the phase on Alice’s phase modulator to vary consistently from  $-\pi$  to  $\pi$ . The Bob’s phase modulation was fixed with value of 0. For each phase setting, we collected data from 220 buffers and counted the detected photons using the method previously described. The total photon count per second was verified using an Abacus coincidence counter. The results are depicted in Figure 6.14, where we

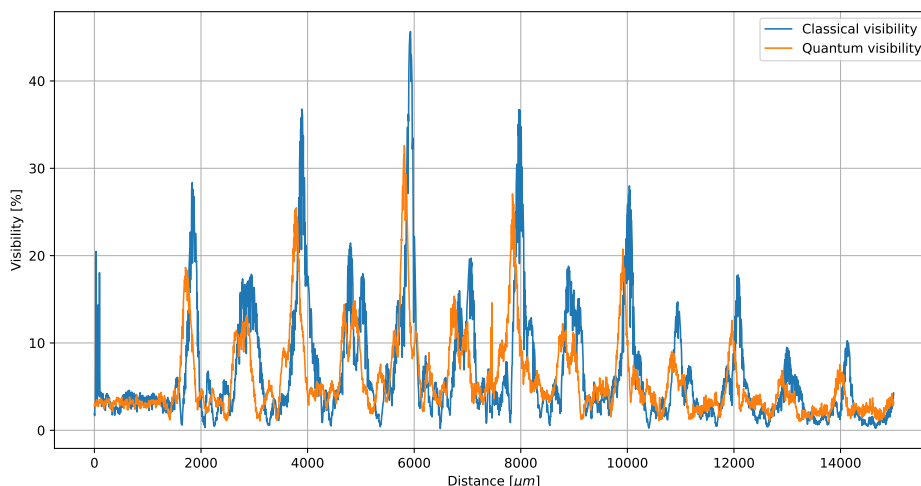


Figure 6.13: Comparison of classical and quantum visibility. The classical visibility was measured using PIN diode and calculated based on the power difference and the quantum visibility was computed using the number of photons. We see that both visibility curves aligns with small shift in the quantum visibility.

achieved a visibility of 30.67%. To reach a visibility above 50%, it would be necessary to discard counts from photons that traveled exclusively through either the short or long arms of the interferometers, as these photons do not contribute to interference. This adjustment would require a synchronized setup capable of precisely tracking the emission, travel, and detection times of each photon to accurately identify the path taken.

To proceed with transmitting quantum information using the BB84 protocol, a more complex classical architecture is necessary. This would involve ensuring precise synchronization between Alice's trigger pulse to the laser diode and the voltage setting on the phase modulator. Additionally, synchronization between Alice and Bob is crucial for them to accurately identify each pulse during the key sifting phase. Typically, this synchronization is achieved by Alice sending a high-intensity laser pulse to Bob to synchronize their clocks.

Furthermore, an advanced device capable of detecting single photons within narrow time windows would be essential, ideally equipped with a gating pulse for the Single Photon Avalanche Diodes (SPADs) to reduce dark counts. While FPGAs are commonly utilized in quantum key distribution as noted in several research papers [61, 62, 63] and it is also a part of the Red Pitaya, a custom-designed FPGA module specifically tailored for high-rate QKD experiments would be required. Our attempts to record the timing of each trigger on the Red Pitaya allowed us to estimate the arrival times of pulses based on the sampling rate of the device. However, the limited

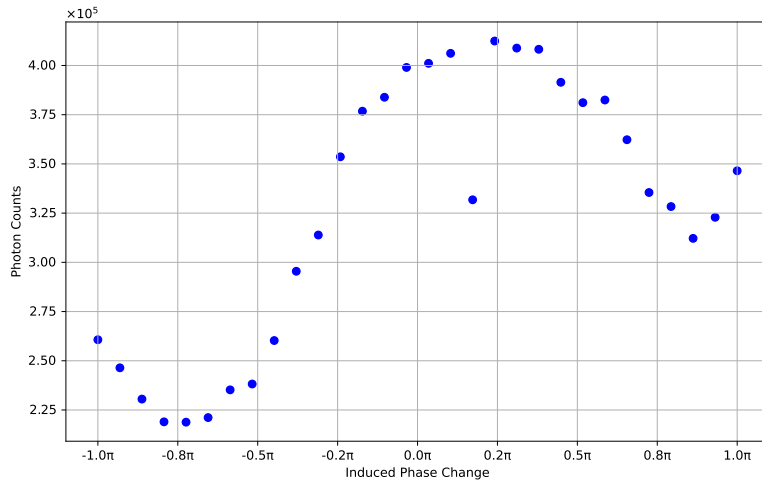


Figure 6.14: Photon Counts vs. Induced Phase Change. Photon counts are calculated from detected pulses across 220 buffers and scaled to detections per second. The phase change is calculated based on the applied voltage and the half-wave voltage of the phase modulator. This measurement yielded a visibility of 30.67% and was conducted at a path difference of 5940  $\mu\text{m}$ .

buffer size and the time precision of the Red Pitaya’s internal FPGA pose significant constraints. For practical implementation, an external clock would be necessary to enhance timing accuracy and overall system performance.

Once these technical challenges are addressed, Alice and Bob could implement the code from our simulation, which handles everything from key sifting onwards. The only modification required would be to establish a classical communication channel between them, potentially using the TCP/IP protocol, to transmit post-processing information.

In summary, we tested and prepared all the optical parts for BB84 protocol and verified its functioning with both classical and quantum signal. We also successfully tested Red Pitaya as a controlling device for laser triggering and phase modulation. We were also able to perform single photon counting in some sense. For performing the whole BB84 protocol, the crucial thing missing is an advanced computer control capable of synchronization between Alice and Bob. We would also need to develop custom FPGA module capable of faster and more reliable photon detection.

## 7 Results and conclusions

In this work, we studied both the theoretical part of QKD and we also proposed an experiment. We prepared theoretical fundamentals for our future work in this field. Particularly we introduced formalism of quantum operators and discussed Shannon, and Von Neuman entropies, together with finite-case of One-shot entropies. While these one-shot entropies were not abundantly used in the later concepts, it is important to cover this topic for the ability to construct proofs for real-life implementations of QKD protocols with a finite number of sent qubits. Thus even though we did not provide any rigorous security proof for the protocol, which would be above the scope of this work, we established the theoretical fundamentals for it.

In chapter number 3, we showed the whole process of Quantum key distribution with its post-processing phase. Besides presenting traditional QKD protocols, we also discussed entanglement-based QKD protocols E91 and BBM92. We also introduced the practical realization of the BB84 protocol with time and phase encoding.

We then analyzed the basics of QKD security. For instance, a one-time pad for absolutely secure communication with a shared key. Then we examined the photon number splitting attack, which is the most important real-life scenario attack, and enhanced protocols resistant to this attack.

The obtained simulation results provide valuable insights into the practical implementation of the BB84 quantum key distribution protocol with phase and time encoding. The simulations cover various scenarios, including the presence of noise, transmission losses, and eavesdropping attacks, allowing for a comprehensive analysis of the protocol's behavior under different conditions.

One of the critical aspects evaluated in the simulations is the quantum bit error rate. The results demonstrate that in the absence of noise and eavesdropping, the QBER is zero, as expected. However, when phase-flip errors and phase deviations are introduced, the QBER increases proportionally to the magnitude of the noise. Specifically, with a phase noise standard deviation of 0.003 and a phase-flip probability of 0.04, the simulated QBER reaches  $3.13 \pm 0.59\%$ . Additionally, incorporating dark counts with a probability of 0.005 contributes a QBER of  $1.29 \pm 0.33\%$ . The combined effect of phase noise, phase-flip errors, and dark counts results in a total QBER of  $4.30 \pm 0.69\%$ , aligning with practical implementations where the error rate typically falls between 3-4%.

The simulations also investigate the impact of eavesdropping attacks on the QBER. In the absence of noise, the average QBER for an intercept-resend attack is  $25.04 \pm 1.01\%$ , consistent with the theoretical prediction that an eavesdropper guessing the correct basis will result in approximately 25% QBER on average. Interestingly, the photon number splitting (PNS) attack exhibits a lower average QBER of  $21.7 \pm 1.05\%$ , highlighting the potential danger of this attack going unnoticed by Alice and

Bob. When noise is introduced, the QBER for the intercept-resend attack increases to  $26.42 \pm 1.10\%$ , while the PNS attack with noise results in a QBER of  $23.28 \pm 1.09\%$ .

The simulations further analyze the relation between the final key length and the QBER in the sifted key as a function of fiber length. As the fiber length increases, photon losses intensify, leading to a higher QBER and a reduced final key size. The simulations demonstrate that the error caused by dark counts increases with longer fiber lengths, and the QBER approaches 50% at around 150 km with no emitted photon reaching Bob.

The key length analysis compares the performance of the classical BB84 protocol, the decoy states BB84 protocol, and the single-photon BB84 protocol. The results show that the decoy states protocol maintains higher probabilities of photon arrival and enables transmission over longer distances compared to the single-photon protocol. Additionally, the final key length for the decoy states protocol is significantly greater than the single-photon protocol for fiber lengths below 50 km, highlighting the advantages of the decoy states approach.

Experimentally, the thesis presents the successful implementation of a quantum key distribution system utilizing the BB84 protocol with time and phase encoding. The experimental setup consists of Alice and Bob's components connected via an optical fiber, with a microcomputer (Red Pitaya STEMLab 125-14) serving as the control unit for various components, including laser diode triggering, phase modulation, and single-photon detection.

Throughout the experiment, visibility measurements were conducted to evaluate the interference pattern at different stages. Initially, a symmetrical Mach-Zehnder interferometer (MZI) was constructed for Alice, achieving a maximum visibility of slightly above 25%. Subsequently, Alice's setup was transitioned to an asymmetrical MZI, and Bob's components were integrated, resulting in a maximum classical visibility of 45.66%. This visibility pattern exhibited two types of peaks, reflecting the interference effects associated with different modes of the laser's output.

Transitioning to the quantum regime, the experiment generated weak coherent pulses and measured the quantum visibility by counting photons. A quantum visibility of 32.60% was achieved, aligning closely with the classical visibility measurements and demonstrating the successful operation of the experimental setup at the single-photon level.

In the final stage, the Red Pitaya microcomputer was programmed to control the entire experiment, including phase modulation and photon counting. A visibility of 30.67% was obtained, showcasing the capability of the low-cost Red Pitaya platform to function as a control unit for quantum key distribution experiments, albeit with limitations in continuous detection and synchronization between Alice and Bob.

Overall, the simulation results provide a comprehensive understanding of the practical challenges and considerations involved in implementing quantum key distribution protocols, while the experimental work successfully demonstrates the principles of the BB84 protocol with time and phase encoding using a low-resource setup. The thesis highlights the importance of addressing noise, losses, and potential eavesdropping attacks, as well as the advantages of decoy state protocols over single-photon approaches. Additionally, the experimental results showcase the feasibility of constructing a quantum cryptography system using off-the-shelf components and a low-cost microcomputer for control purposes, although further advancements in synchronization and real-time data acquisition are necessary for practical applications.

Therefore we conclude that all the goals of the thesis were successfully achieved.



## Bibliography

- [1] G. Alagic et al. *Status Report on the Third Round of the NIST Post-Quantum Cryptography Standardization Process*. 2022.
- [2] National Institute of Standards and Technology. *Post-Quantum Cryptography Standardization*. Accessed: 2024-03-28. 2024. URL: <https://csrc.nist.gov/Projects/post-quantum-cryptography/post-quantum-cryptography-standardization>.
- [3] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.
- [4] M. M. Wilde. “From Classical to Quantum Shannon Theory”. In: *Quantum Information Theory*. Cambridge University Press, 2019, pp. xi–xii.
- [5] L. Gyongyosi, S. Imre, and H. V. Nguyen. “A Survey on Quantum Channel Capacities”. In: *IEEE Communications Surveys & Tutorials* 20.2 (2018), pp. 1149–1205.
- [6] R. Wolf. *Quantum Key Distribution : An Introduction with Exercises*. Springer, 2021.
- [7] G. Smith. “Quantum Channel Capacities”. In: *CoRR* 1007.2855 (2010).
- [8] P. W. Shor. “Capacities of quantum channels and how to find them”. In: *Mathematical Programming* 97.1 (2003), pp. 311–335.
- [9] R. Renner. “Security of quantum key distribution.” PhD thesis. ETH Zurich, 2005.
- [10] R. König, R. Renner, and C. Schaffner. “The Operational Meaning of Min- and Max-Entropy”. In: *IEEE Transactions on Information Theory* 55.9 (2009), pp. 4337–4347.
- [11] M. Berta, M. Christandl, and R. Colbeck. “The uncertainty principle in the presence of quantum memory”. In: *Nature Phys* 6 (2010), pp. 659–662.
- [12] Y. Ding et al. “High-dimensional quantum key distribution based on multicore fiber using silicon photonic integrated circuits”. In: *npj Quantum Information* 3.1 (2017), p. 25.
- [13] Y. Zhao C. H. et al. “Quantum hacking: Experimental demonstration of time-shift attack against practical quantum-key-distribution systems”. In: *Phys. Rev. A* 78 (4 2008), p. 042333.
- [14] J. Z. Huang et al. “Quantum hacking of a continuous-variable quantum-key-distribution system using a wavelength attack”. In: *Phys. Rev. A* 87 (6 2013), p. 062329.
- [15] C. H. Bennett and G. Brassard. “Quantum cryptography: Public key distribution and coin tossing”. In: *Theoretical Computer Science* 560 (2014), pp. 7–11.

- [16] S. Ataman. *The quantum optical description of a double Mach-Zehnder interferometer*. 2014.
- [17] S. Pirandola et al. “Advances in quantum cryptography”. In: *Advances in Optics and Photonics* 12.4 (2020), p. 1012.
- [18] D. Bruß. “Optimal Eavesdropping in Quantum Cryptography with Six States”. In: *Phys. Rev. Lett.* 81 (14 1998), pp. 3018–3021.
- [19] C. H. Bennett. “Quantum cryptography using any two nonorthogonal states”. In: *Phys. Rev. Lett.* 68 (21 1992), pp. 3121–3124.
- [20] K. Tamaki, M. Koashi, and N. Imoto. “Unconditionally Secure Key Distribution Based on Two Nonorthogonal States”. In: *Phys. Rev. Lett.* 90 (16 2003), p. 167904.
- [21] K. Tamaki and N. Lütkenhaus. “Unconditional security of the Bennett 1992 quantum key-distribution protocol over a lossy and noisy channel”. In: *Phys. Rev. A* 69 (3 2004), p. 032316.
- [22] N. Gisin et al. “Quantum cryptography”. In: *Reviews of Modern Physics* 74.1 (), pp. 145–195.
- [23] K. Inoue. “Quantum key distribution technologies”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 12.4 (2006), pp. 888–896.
- [24] P. D. Townsend. “Quantum Cryptography on Optical Fiber Networks”. In: *Optical Fiber Technology* 4.4 (1998), pp. 345–370.
- [25] A. Muller et al. ““Plug and play” systems for quantum cryptography”. In: *Applied Physics Letters* 70.7 (1997), pp. 793–795.
- [26] V. Scarani et al. “The security of practical quantum key distribution”. In: *Rev. Mod. Phys.* 81 (3 2009), pp. 1301–1350.
- [27] A. K. Ekert. “Quantum cryptography based on Bell’s theorem”. In: *Phys. Rev. Lett.* 67 (6 1991), pp. 661–663.
- [28] C. H. Bennett, G. Brassard, and N.D. Mermin. “Quantum cryptography without Bell’s theorem”. In: *Phys. Rev. Lett.* 68 (5 1992), pp. 557–559.
- [29] V. Sharma et al. “A comparative study of protocols for secure quantum communication under noisy environment: single-qubit-based protocols versus entangled-state-based protocols”. In: *Quantum Information Processing* 15.11 (2016), pp. 4681–4710.
- [30] G. Brassard and L. Salvail. “Secret-Key Reconciliation by Public Discussion”. In: *Advances in Cryptology — EUROCRYPT ’93*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994, pp. 410–423.
- [31] Jesus M. M. et al. “Demystifying the information reconciliation protocol cascade”. In: *Quantum Info. Comput.* 15.5–6 (2015), pp. 453–477.

- [32] B. Rijsman. *Quantum Key Distribution (QKD) Protocols*. Accessed: 2024-04-24. URL: <https://cascade-python.readthedocs.io/en/latest/protocol.html#quantum-key-distribution-qkd-protocols>.
- [33] W. T. Buttler et al. “Fast, efficient error reconciliation for quantum cryptography”. In: *Physical Review A* 67.5 (2003).
- [34] D. Elkouss et al. “Efficient reconciliation protocol for discrete-variable quantum key distribution”. In: *2009 IEEE International Symposium on Information Theory*. IEEE, 2009.
- [35] J. S. Johnson et al. “An analysis of error reconciliation protocols used in Quantum Key Distribution systems”. In: *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* 12 (2015), pp. 217–227.
- [36] F. Grasselli. *Quantum Cryptography: From Key Distribution to Conference Key Agreement*. Quantum Science and Technology. Springer International Publishing, 2021. ISBN: 9783030643591.
- [37] C. Paar and J. Pelzl. *Understanding Cryptography: A Textbook for Students and Practitioners*. 1st. Springer Publishing Company, Incorporated, 2009. ISBN: 3642041000.
- [38] F. Xu et al. “Secure quantum key distribution with realistic devices”. In: *Rev. Mod. Phys.* 92.2 (2020), p. 25002.
- [39] M. Mehic et al. “Quantum Key Distribution: A Networking Perspective”. In: *ACM Comput. Surv.* 53.5 (2020).
- [40] G. Brassard et al. “Limitations on Practical Quantum Cryptography”. In: *Phys. Rev. Lett.* 85 (6 2000), pp. 1330–1333.
- [41] X. H. Li, F. G. Deng, and H.Y Zhou. “Efficient quantum key distribution over a collective noise channel”. In: *Phys. Rev. A* 78 (2 2008), p. 022321.
- [42] D. Stucki et al. “Photon counting for quantum key distribution with peltier cooled InGaAs/InP APDs”. In: *Journal of Modern Optics* 48.13 (2001), pp. 1967–1981.
- [43] L. Lydersen et al. “Hacking commercial quantum cryptography systems by tailored bright illumination”. In: *Nature Photonics* 4.10 (2010), pp. 686–689.
- [44] S. Sun and A. Huang. “A Review of Security Evaluation of Practical Quantum Key Distribution System”. In: *Entropy* 24 (2022), p. 260.
- [45] V. Scarani et al. “Quantum Cryptography Protocols Robust against Photon Number Splitting Attacks for Weak Laser Pulse Implementations”. In: *Phys. Rev. Lett.* 92 (5 2004), p. 057901.
- [46] W. Y. Hwang. “Quantum Key Distribution with High Loss: Toward Global Secure Communication”. In: *Phys. Rev. Lett.* 91 (5 2003), p. 057901.

- [47] D. Rosenberg et al. “Long-Distance Decoy-State Quantum Key Distribution in Optical Fiber”. In: *Phys. Rev. Lett.* 98 (1 2007), p. 010503.
- [48] Y. Liu et al. “Decoy-state quantum key distribution with polarized photons over 200 km”. In: *Opt. Express* 18.8 (2010), pp. 8587–8594.
- [49] H. Takesue et al. “Differential phase shift-quantum key distribution”. In: *IEEE Communications Magazine* 47.5 (2009), pp. 102–106.
- [50] K. Inoue. “Differential Phase-Shift Quantum Key Distribution Systems”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 21 (2015), pp. 109–115.
- [51] S. Chianga et al. “Towards practical quantum cryptography”. In: *Applied Physics B* 69.5 (1999), pp. 389–393.
- [52] D. Stucki et al. “Fast and simple one-way quantum key distribution”. In: *Applied Physics Letters* 87.19 (2005).
- [53] D. Stucki et al. “Continuous high speed coherent one-way quantum key distribution”. In: *Optics Express* 17.16 (2009), p. 13326.
- [54] Id Quantique. *Brochure for Cerberis XG QKD System*. Accessed: 2024-05-05. URL: [https://marketing.idquantique.com/acton/attachment/11868/f-2e621d25-e414-4772-a482-b1b272c24c11/1/-/-/-/-/Cerberis%20XG%20QKD%20System\\_Brochure.pdf](https://marketing.idquantique.com/acton/attachment/11868/f-2e621d25-e414-4772-a482-b1b272c24c11/1/-/-/-/-/Cerberis%20XG%20QKD%20System_Brochure.pdf).
- [55] O. Cermak. *Simulation of BB84 Phase-Time Quantum Key Distribution protocol*. Accessed: 2024-05-9. 2024. URL: [https://github.com/stopzer0/QKD\\_simulator](https://github.com/stopzer0/QKD_simulator).
- [56] R. Paschotta. *Telecom Fibers*. RP Photonics Encyclopedia. Accessed: 2024-02-18. URL: [https://www.rp-photonics.com/telecom\\_fibers.html](https://www.rp-photonics.com/telecom_fibers.html).
- [57] *Manual Fiber Polarization Controllers*. Accessed: 2023-11-09. URL: [https://www.thorlabs.com/newgrouppage9.cfm?objectgroup\\_id=343](https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=343).
- [58] H. Okayama. “Lithium Niobate Electro-Optic Switching”. In: (2006), pp. 39–81.
- [59] X. Meng et al. “InGaAs/InAlAs single photon avalanche diode for 1550 nm photons”. In: *Royal Society Open Science* 3.3 (2016), p. 150584.
- [60] S. H. Baek et al. “Room temperature quantum key distribution characteristics of low-noise InGaAs/InP single-photon avalanche diode”. In: *Journal of the Korean Physical Society* 78.7 (2021), pp. 634–641.
- [61] H. F. Zhang et al. “A Real-Time QKD System Based on FPGA”. In: *Journal of Lightwave Technology* 30.20 (2012), pp. 3226–3234.
- [62] X. Lu, L. Zhang, Y. Wang, et al. “FPGA based digital phase-coding quantum key distribution system”. In: *Science China Physics, Mechanics & Astronomy* 58 (2015), p. 120301.

- [63] Damien Stucki et al. “Continuous high speed coherent one-way quantum key distribution”. In: *Opt. Express* 17.16 (2009), pp. 13326–13334.