PRACTICAL QUANTUM CRYPTOGRAPHY: DEMONSTRATION OF SECURE COMMITMENTS IN THE NOISY STORAGE MODEL

submitted by

NELLY NG HUEI YING

A final year project report

presented to

Nanyang Technological University

in partial fulfilment of the

requirements for the

Bachelor of Science (Hons) in Physics / Applied Physics

Nanyang Technological University

April 2012

Abstract

The establishment of quantum theory has sparked further advancements in computer science and information theory. In this work, we explore aspects of quantum cryptography, showing how physical assumptions combined with fundamental laws of nature allow for secure communication, without additional computational assumptions as frequently invoked in many present-day cryptographic protocols. These progresses have wide applications, which can be used to perform tasks which lie in the central interest of modern cryptography, such as secure identification of a customer to an ATM machine.

In particular, we have studied the *bit commitment* protocol within the scope of Noisy-Storage Model, and proved its robustness against experimental errors, demonstrating the feasibility of executing the protocol with real-world quantum devices. Having so, we conducted the implementation with actual optical devices. On the technical aspect, we developed a new set of uncertainty relations, which are extremely useful for actual implementations of protocols. These relations allow a significant reduction in the amount of classical information post-processing, directly shortening the computational time for secure two-party protocols in the model. Our approach involves the analytical optimization for a certain class of Rényi entropic measures conditioned on quantum measurements, which might be of independent interest.

Acknowledgements

First of all, I give thanks to the most high God, Whom has graciously granted me more than sufficient physical strength, mental resilience and intellectual knowledge to complete this work. I thank Him for creating joy and peace within me, enabling me to comprehend and appreciate the knowledge that I have acquired throughout the process.

This project would never have been successful without the supervision of my professors, Prof. Stephanie Wehner and Prof. Chew Lock Yue. Prof. Wehner guided me through each detail and discussed each argument with rigour. I thank her for always pointing out my faults, at the same time always encouraging and eager to discuss ideas. With a passionate but serious attitude towards research, she has inspired me tremendously. I also thank her for the giving me the chance to participate in the Quantum Cryptography conference held at ETH Zurich in September 2011.

I thank Prof. Chew for guiding me throughout the project, always reminding me to keep a clear understanding of concepts. His clear and simple logical flow of thoughts have taught me to look at the research topic from a broader perspective, capturing the essence of the project itself. I also greatly appreciate his patience and kindness throughout these years of undergraduate research.

I have been tremendously blessed by having the chance to work with my collaborators, Christian Kurtseifer, Siddarth Joshi and Mario Berta. I also appreciate the conducive environment in Centre for Quantum Technologies, and weekly discussions with fellow students and researchers. All these has greatly increased my knowledge and appreciation for the field of quantum information.

I would like to express my utmost gratitude to Professor Elbert Chia, who has been an incredible lecturer, educator and inspirer throughout my undergraduate studies in NTU. It is an honour to be his student. I thank my family and friends for constantly believing in me and giving me mental support. I would also like to thank two of my high school teachers, Mr. Tan Lak Tong and Mdm. Tan Sok Sin for giving me the very first inspiration and motivation to study physics and mathematics.

I dedicate this thesis to my younger brother Isaac, in hope that he will grow to be a creative, passionate, disciplined and responsible individual, and also learn to appreciate the beauty of science more.

Contents

Parts of this work are also found in the following articles:

1. First experimental implementation of bit commitment in the noisy-storage model,
Nelly Ng, Siddarth K. Joshi, Chia Chen Ming, Christian Kurtseifer, Stephanie Wehner
(in preparation)

2. A min-entropy uncertainty relation for finite size cryptography,
Nelly Ng, Mario Berta, Stephanie Wehner
(in preparation)

Chapter 1

Introduction

This chapter serves as a basic grounding, in order to explain how fundamental laws of quantum theory can be applied to cryptographic scenarios. An overview of ideas in quantum cryptography is presented. We discuss what it means for a cryptographic protocol to be correct and secure. Lastly, we introduce the Noisy Storage Model and discuss its significance in modern day quantum cryptography.

1.1 A brief history of quantum cryptography

Quantum information science has offered a whole new set of perspectives on computing and cryptography, by invoking fundamental laws of quantum physics to develop more sophisticated algorithms for problem solving. Such progress potentially provides solutions for many hard problems. For example, consider the problem of finding the prime factors of a large integer N. This mathematical problem is conjectured to be hard classically, i.e. having high computational complexity. However, the quantum Shor's algorithm has reduced its complexity to that of polynomial time, by invoking methods of quantum Fourier transform. This has been implemented by [24] and more recently, by [21] using photonic systems, showing that factorization is a relatively easy task for a quantum computer.

However, these new discoveries bring serious problems that impact real-life, i.e. they pose a threat to the security of a large class of cryptographic protocols, which rely crucially on the conjectured computational complexity of a mathematical problem. Taking the RSA public key algorithm for example, which is useful to generate a private key between two parties Alice and Bob. This protocol relies heavily on the assumption that large numbers are extremely difficult to factorize. Classically, indeed there is no known algorithm efficient

enough to do this for large numbers. However, given the advancements on implementing Shor's algorithm, a good enough quantum computer will potentially be able to break such cryptographic systems.

Fortunately at the same time, the same laws of nature remarkably offer an alternative to the task of generating a secret key between two parties, by communicating through a public quantum channel. Being provably secure, quantum key distribution (QKD) is probably the most successful protocol in quantum cryptography. We explain the famous BB84 protocol by [3] to obtain a rough idea of how this works.

In the BB84 QKD protocol, Alice chooses a random bit string P of length $4n$, with each string element defined over the binary field $\{0, 1\}$. She then chooses another random binary string Q of the same length. For each bit $P(i)$, if $Q(i) = 0$ she encodes $P(i)$ in the Z basis $\{|0\rangle, |1\rangle\}$. Otherwise if $Q(i) = 1$, $P(i)$ is encoded in the X basis $\{|+\rangle, |-\rangle\}$, where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. She then sends the quantum state over to Bob by a public quantum channel. For simplicity we first assume that the channel is noiseless.

When Bob receives the state, he randomly chooses to measure in the X or Z basis. The basis choice is recorded in a string R. Subsequently, Alice sends Q to Bob, and Bob sends R to Alice. They pick out the places where their measurement bases match, and obtain roughly $2n$ shared bits.

Subsequently, Alice chooses a random subset of n bits and announces it to Bob, while they compare the measurement outcomes. Note that if the channel is noiseless, Alice and Bob's bit values should match perfectly. In the case where the quantum channel is noisy, coding procedures such as error-correction and privacy amplification can always be performed such that the correctness and secrecy of the key is ensured.

If an eavesdropper Eve is tapping on the channel, she is able to intercept the quantum state sent from Alice to Bob. To gain information about the key, she can either choose to perform a measurement to guess Alice's basis, or she can manipulate the quantum state by acting it upon some unitary operation.

The success of quantum key distribution lies within the insight, that whenever an adversarial Eve attempts to obtain information transmitted from Alice to Bob, her act of observation inevitably disturbs the system in a way where Alice and Bob may detect, by examining the statistics of measurement outcomes. This can be successful due to various properties of quantum states, such as no-cloning (it is impossible to make an exact second copy of an arbitrary quantum state, given no prior information about the state itself) etc.

In other words, comparison of measurement statistics allows Alice and Bob to make a statement about how much information Eve, if exists, holds about the remaining n shared bits between them. If the amount of disagreements between Alice and Bob exceeds a certain limit, both parties abort the protocol. Otherwise, they can be convinced that Eve's information about the n shared bits is limited. By randomness distillation procedures, a secret key can then be established between Alice and Bob.

Since the pioneering work of [3], many variants of QKD have been proposed. Large number of researches are working towards realizing and commercializing QKD networks, justifying the importance and potential of quantum cryptography as a field of new science.

1.2 Information theoretic security

A cryptographic protocol usually contains two requirements: its correctness, and security. The correctness of the protocol means that whenever authorized parties are using the protocol, the outputs are the correct and desirable ones. Security means that whenever a dishonest user is involved, the information held by honest parties are protected, or sometimes in addition that the cheating is always detected by honest parties.

In other words, security is the protection of information from unauthorized users or malicious parties in the process of communication. It is worth noting that there exists many variants of security. *Information-theoretic security* for a protocol implies that there is no conjecture/assumption about the computational complexity of problem solving. In such cases, security holds even against a adversarial party with unlimited computational resources, because he/she simply does not gain enough information to break the protocol. The QKD protocol is an example where information-theoretic security is provably true.

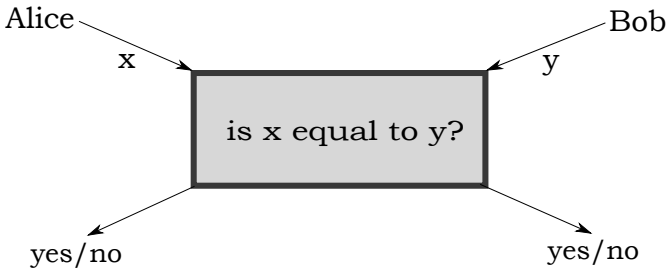
We have seen that computational assumptions are subject to large amount of changes, due to developments in quantum computing. Information-theoretic security is therefore interesting because of its robustness under the advancements of quantum algorithms.

1.3 Modern quantum cryptography

In general, it is desirable to design quantum cryptographic primitives so that information theoretic security holds, solely based on the validity of quantum physics. Besides key distribution, with advancements of cryptography, two-party computations have been of much interest. These protocols enable two-parties, Alice and Bob, to solve joint problems

even if they do not trust each other. Assuming Alice and Bob each have some information x and y. Two party computations then have a general form: the goal is to compute a joint function $f(x, y)$ such that Alice cannot know more about y other than provided by the output $f(x, y)$, and similarly for Bob.

Examples of such tasks include secure auctions, or the ever present problem of secure identification, e.g. of a customer to an ATM machine. For example, consider the problem of secure identification, whereby a user obtains access to a machine by entering a correct password. At the same time an honest user wants to protect himself from an adversarial machine, which might have been hacked to steal passwords. Essentially this implies a two-party computation that evaluates the equality function, returning a single answer "yes" if $x = y$, and "no" otherwise.



Security for two party protocols is generally hard to achieve, however most of them can be reduced to a combination of several fundamental protocols, such as commitments (in particular bit commitment BC, where the committed information is a single bit) and oblivious transfer (OT). The security of these protocols then becomes an interesting question, since one can hope to achieve many other complicated tasks by a composition of these building blocks.

It is long known that these fundamental schemes can never be perfectly secure given any classical scheme. Therefore, all present-day available schemes are based on computational security, which is highly undesirable. This poses a major challenge to modern quantum cryptography: if quantum communication is allowed (i.e. the information sent across from a party to another is allowed to be a quantum state, instead of a classical variable), can information-theoretic security be obtained for such protocols?

Many quantum-mechanical protocols have been suggested for implementing these tasks. Unfortunately, a large amount of results have disproved the security of these protocols, even if quantum communication is allowed [28, 7, 22, 23, 11]. It has subsequently been proven that given only quantum communication without any further assumptions, it is impossible for *any* bit commitment protocol to be secure. Only weak variants can be

obtained, in which the attacker can cheat with a very large probability, which renders them rather uninteresting for any practical application.

This might serve as a surprise for most people: how could this be when quantum communication offers such a great advantage in distributing encryption keys? However, note that in QKD Alice and Bob trust each other and want to defend themselves against an outside eavesdropper Eve. In particular, this allows Alice and Bob to perform checks on what Eve may have done, ruling out many forms of attacks. This is in sharp contrast to two-party cryptography where there is no Eve, and Alice and Bob do not trust each other. Intuitively, it is this lack of trust that makes the problem considerably harder, as each player now has to fend for himself.

1.4 Bounded and Noisy Quantum Storage Model

Yet, because two-party protocols form a central part of modern cryptography, one is willing to make *assumptions* on how powerful an attacker can be in order to implement them securely. Besides computational assumptions, *physical* assumptions are also used often in classical two-party schemes. For example, [27, 6] have shown that by bounding the amount of information stored by an adversary, security can often be achieved.

Physical assumptions are often reasonable, since if the amount of physical resources required to cheat is much more costly than the benefits of breaking the protocol, one can say that security is valid. However, bounding classical data storage has become a rather undesirable assumption, since classical data storage has become much more plentiful and cheap nowadays. One can easily store a few terabytes (8×10^{12} bits) of classical data, making these assumptions easily breakable. Quantum information, on the other hand, is much harder to store, since it is very difficult to preserve the superposition properties of a quantum state. Up to now, storing even a couple hundred of qubits (two-level quantum systems) is technologically hard, since these states are heavily subjected to decoherence, while interacting with the storage environment.

In the light of this, a natural strategy is to make physical assumptions of an adversarial party to store quantum information. The bounded/noisy quantum storage model works exactly only this physical assumption: quantum information is extremely hard and expensive to store. Therefore, an adversary has limited amount of quantum storage, and generally his quantum memory is subjected to quantum noise. In this model, any other assumptions on the abilities of cheating parties are discarded: they are allowed to avoid any

form of transmission noise, store unlimited amount of classical data, perform unlimited classical and simultaneous quantum computations upon the received state.

1.5 Overview of Work

Theoretical results in [19] have shown that schemes such as BC and OT are provably secure against such an all-powerful adversary in the Noisy Storage Model. This however, is not sufficient to prove that BC and OT are indeed implementable under experimental losses and errors. For oblivious transfer, [38, 35] has shown the *robustness* of OT under certain ranges of experimental errors, where honest parties are subjected to some unavoidable noise. This is an important problem, since any implementation of such protocols with real-world quantum devices are unavoidably subjected to noise, where a dishonest party might take further advantage of such imperfections to perform cheating.

Also, it is worth noting that only theoretical proofs of principle have been obtained: there has never been an actual demonstration or implementation of these protocols. The *practical feasibility* of these protocols are also of utmost importance, since it is desirable to perform secure implementations that can potentially replace the computationally-secure protocols widely used in the present.

In this work, our main goal is to prove the robustness and demonstrate the practical feasibility of *secure bit commitment* (BC) under the Noisy Storage assumption.

This thesis consists of 6 chapters: Chapter 1 gives an overview and brief introduction to the central ideas in quantum cryptography. In Chapter 2, we introduce a set of necessary tools required for our analysis, which includes basic knowledge about quantum measurements, and some classical information processing tools.

The subsequent three chapters mainly contain our main work and results. In Chapter 3, we derive uncertainty relations by developing bounds for a certain class of conditional Rényi entropic measures. In Chapter 4, we apply the results of Chapter 3 to study the commitment scheme under the Noisy Storage Model. After introducing the bit commitment protocol shown in [19], we sought to prove its robustness under experimental losses and errors. Chapter 5 introduces the experimental aspect of executing the protocol, and by executing the protocol for honest parties, we demonstrate the feasibility of commitments in this model.

Finally, Chapter 6 provides a brief summary of our work, including several prospects and open problems that might be of future interest.

Chapter 2

Tools and Preliminaries

An introduction of the tools used and some background information is presented to the reader. Firstly, the notation for describing classical and quantum information is introduced in Section 2.1. Subsequently, a special entropic measure called the smooth min-entropy is introduced in Section 2.2. This quantity will be used to measure the uncertainty of information with respect to quantum measurements. Furthermore, the concept of expressing quantum mechanical uncertainty by means of entropy is also explained. After that, the idea of privacy amplification by using cryptographic tools such as 2-universal hash functions is explained in Section 2.3. Lastly, a brief description of error-correcting codes and some asymptotic properties are given in Section 2.4.

2.1 Classical and quantum information

2.1.1 Density matrix formalism

Define a Hilbert Space $\mathcal{H}_n$ of dimension n. For any set of orthonormal basis vectors $\{|e_1\rangle, ..., |e_n\rangle\} \in \mathcal{H}_n$, any normalised pure quantum state can be expressed as a linear superposition of these bases

$$|\phi\rangle = \sum_{i=1}^{n} \nu_i |e_i\rangle \qquad \text{where for } \forall i,\ \nu_i \in \mathbb{C}, \text{ and } \quad \sum_{i=1}^{n} |\nu_i|^2 = 1. \tag{2.1}$$

Assuming a random variable X defined over the range $\{x_i\}$, with corresponding probabilities $\{p_i\}$. We denote such a distribution to be $\{x_i, p_i\}$. Mixed states are described by having a classical probability distribution over a set of pure quantum states $\{|\phi_i\rangle, p_i\}$. Note that such states cannot be described as a linear superposition of basis vectors, as

done for pure quantum states. This is the motivation of introducing the density matrix formalism, which incorporates the nature of statistical mixture into a quantum-mechanical framework.

The density matrix formalism maps each state vector from $\mathcal{H}_n$ to a matrix $\rho \in \mathbb{C}^{n\times n}$. For any pure quantum state $|\phi\rangle$, the corresponding density matrix is

$$\rho_{\text{pure}} = |\phi\rangle\langle\phi| = \sum_{i,j=1}^{n} \nu_i \nu_j^* \; |e_i\rangle\langle e_j|, \tag{2.2}$$

where the trace of density matrix is

$$\text{tr}(\rho) = \sum_k \langle e_k|\rho|e_k\rangle = \sum_{i,j,k=1}^{n} \nu_i \nu_j^* \delta_{ik}\delta_{jk} = \sum_{k=1}^{n} |\nu_k|^2 = 1. \tag{2.3}$$

On the other hand, a mixed state given by the statistical mixture of pure (not necessarily orthogonal) states $\{|\phi_i\rangle, p_i\}$, has a corresponding density matrix $\rho_{mixed} = \sum_i p_i |\phi_i\rangle\langle\phi_i|$. A mixed density matrix also has unity trace, since

$$\text{tr}(\rho_{\text{mixed}}) = \sum_i p_i \,\text{tr}(|\phi_i\rangle\langle\phi_i|) = \sum_i p_i = 1. \tag{2.4}$$

Although all states are characterized by the same form, $\text{tr}(\rho^2) = 1$ only if ρ is a pure state. Otherwise, $\text{tr}(\rho^2) < 1$. Furthermore, since the trace is invariant under basis transformations, the quantum-mechanical/statistical nature of the state is basis-independent.

A *classical* state simply denotes a probability distribution over a fixed basis (orthogonal) set $\{|x_i\rangle, p_i\}$, and can always be written as $\rho_X = \sum_i p_i |x_i\rangle\langle x_i|$.

Throughout this analysis, we describe both classical (statistical) and quantum information contained in a state by using the same notion of density matrices, hence it is useful to keep in mind the differences between them. Also, states can simultaneously contain both classical and quantum information. They are described as multipartite systems. For example, assuming a random variable over the distribution $\{x_i, p_i\}$, where conditioned on the value of $X = x_i$, a quantum state ρ_i is obtained. The state in the Hilbert space $\mathcal{H}_\mathcal{X} \otimes \mathcal{H}_\mathcal{Q}$ can then be described as

$$\rho_{XQ} = \sum_i p_i |x_i\rangle\langle x_i| \otimes \rho_i. \tag{2.5}$$

Such states are referred to as cq-states, or classical-quantum states.

2.1.2 Qubits

Consider the simplest type of quantum system, that lives within a two-dimensional Hilbert space. These systems are called qubits, which are analogues of classical bits in information theory. Any pure qubit state can be expressed in the simplest canonical basis

$$|\psi\rangle = a|0\rangle + b|1\rangle, \qquad \mathrm{a, b} \in \mathbb{C}, \quad \text{such that} \quad |a|^2 + |b|^2 = 1. \tag{2.6}$$

A few commonly used operators for qubit systems are introduced here. Firstly, we have the set of Pauli spin matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Denote the ordered eigenbases for each matrix σ_i as $\{|e_{i1}\rangle, |e_{i2}\rangle\}$. Then the Pauli matrices are traceless and have mutually unbiased eigenbases, i.e. $\forall i \neq j$,

$$|\langle e_{i1}|e_{j1}\rangle|^2 = |\langle e_{i1}|e_{j2}\rangle|^2 = |\langle e_{i2}|e_{j1}\rangle|^2 = |\langle e_{i2}|e_{j2}\rangle|^2 = \frac{1}{2}. \tag{2.7}$$

Mutually unbiased bases represent the eigenbasis of quantum-mechanical operators that exhibits the largest incompatibility. In other words the largest amount of quantum-mechanical uncertainty is obtained for the observables corresponding to such operators. Some general results on their existence and properties are shown in [1].

The Hadamard operator permutes the σ_x and σ_z bases, while the T operator cyclicly permutes the eigenbases of all three Pauli matrices. These operators take the form

$$\mathbf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \mathbf{T} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

For any qubit state ρ, the following relation is satisfied:

$$\sum_{i=x,y,z} \mathrm{tr}(\sigma_i \rho)^2 \leq 1 \tag{2.8}$$

where equality holds if ρ is pure. This is known as the Bloch sphere condition.

Any two-level system can be represented as a qubit. Examples of such systems are fermions with up/down spins, or photons with horizontal/vertical polarization. Measurements defined by the operators σ_x and σ_z are referred to as BB84 measurements [3], while measurements defined by the eigenbases of all three Pauli matrices are called the six-state

measurements. These measurements are experimentally implemented by Stern-Gerlach apparatus for spins, or using polarized beamsplitters for photons.

2.1.3 Measurement operators

A central task of importance in quantum information theory is to distinguish between different quantum states. To identify an arbitrary quantum state, the most general strategy is to perform unitary operations (in general, with the assistance of an ancillary system) and perform measurements on the state, observing its outcome. Concretely, let us assume that a state was chosen from a set of density matrices with the probability distribution $\{\rho_i, p_i\}$. The goal is to design a set of suitable measurements, such that after performing the measurement upon the state, it can be identified with optimal correctness based on the measurement outcome. In this section, we introduce the formalism of measurement operators and their functionalities.

Projective Measurements

Definition 2.1.1 (Projective measurements)**.** *A projective measurement is defined by an operator M, which satisfies*

1. *$M^\dagger = M$ is self-adjoint.*
2. *$M^\dagger M = M^2 = M$, implying that the measurement is idempotent.*

A complete set of projective measurements $\{M_1, ...M_k\} \in \mathbb{C}^{n\times n}$ defined in $\mathcal{H}_n$ satisfies the completeness relation: $\sum_{i=1}^{k} M_i = \mathbb{I}$. Projective measurements, when acted upon any density matrix ρ, can be viewed as a projection of ρ onto a subspace of $\mathcal{H}_n$, obtaining each measurement outcome m_i with probability

$$p(m_i) = \mathrm{tr}(M_i\rho). \tag{2.9}$$

The possible measurement outcomes are described by the distribution $\{m_i, \mathrm{tr}(M_i\rho)\}$.

To provide an example on how projective measurements can be used to distinguish states, consider an ensemble of pure normalized states $\{|\psi_i\rangle\}$, with corresponding density matrices $\{\rho_i\}$, for $1 \leq i \leq n$. If $\langle\psi_i|\psi_j\rangle = \delta_{ij}$ are orthonormal, distinguishing states becomes easy. In fact, since orthonormal states form a basis for the Hilbert space $\mathcal{H}_n$ of dimension n, by choosing the set of measurement operators $\{M_i\}$ such that $M_i = |\psi_i\rangle\langle\psi_i|$, the task of distinguishing states is complete. Let $\{m_1, ...m_n\}$ denote the corresponding measurement outcomes. Note that if the pure density matrix $\rho_i = |\psi_i\rangle\langle\psi_i|$ is subjected to these measurement operators, $p(m_j) = \mathrm{tr}(M_j\rho_i) = \delta_{ij}$.

Now, suppose we receive a quantum state selected randomly from the distribution $\{\rho_i, p_i\}$, with the ensemble of pure normalized density matrices $\{\rho_i\}$ as stated above. By performing the specified measurement upon the received state, we obtain some outcome m_j. Based on this outcome, we guess that the state received was ρ_j. The probability of guessing correctly is given by

$$p_{guess} = \sum_{i=1}^{n} p_i \operatorname{tr}(M_i \rho_i) = \sum_{i=1}^{n} p_i \delta_{ii} = 1. \tag{2.10}$$

Let us look at another example of projective measurements, where the task is to distinguish two arbitrary quantum states ρ_0 and ρ_1. Given a density matrix from the set $\{\rho_0, \rho_1\}$ where each state is chosen with equal probability. The guessing probability is

$$\begin{aligned} p_{guess} &= \frac{1}{2} \left[\operatorname{tr}(M_0\rho_0) + \operatorname{tr}(M_1\rho_1)\right] \\ &= \frac{1}{2} \left\{\operatorname{tr}(M_0\rho_0) + \operatorname{tr}[(\mathbb{I} - M_0)\rho_1]\right\} \\ &= \frac{1}{2} \left[1 + \operatorname{tr}(M_0 A)\right], \end{aligned} \tag{2.11}$$

where $A = \rho_0 - \rho_1$ is Hermitian, therefore it can be diagonalised as $A = \sum_{i=1}^{n} E_i |e_i\rangle\langle e_i|$. To maximize p_{guess}, define the operator M_0 to be

$$M_0 = \sum_{i, E_i \geq 0} |e_i\rangle\langle e_i|. \tag{2.12}$$

Indeed, it has been shown by [15] that for distinguishing any two states ρ and ρ', the optimal p_{guess} can be obtained by a set of projective measurements, often referred to as a von Neumann measurement.

Positive Operator-Valued Measurements (POVM)

POVMs are sometimes useful when an optimal measurement to identify a set of non-orthogonal quantum states is of interest. In general, these measurement operators do not have to be mutually orthogonal. Although they still need to be self-adjoint, they do not need to be idempotent either.

Once such conditions are relaxed, POVMs define a set of elements $\{A_i\}$,[1] with $\{a_i\}$ being the associated measurement outcomes. The only conditions required are that each

[1]The number of POVM elements can in general be larger than the dimension of Hilbert Space, in contrast to projective measurements, where for any complete set of projectors, the number of operators must be less or equal to the dimension of the Hilbert Space.

element $A_i \geq 0$ are positive matrices[2] , and the completeness relation is satisfied, namely

$$\sum_i A_i = \mathbb{I}. \tag{2.13}$$

Given density matrix ρ, the probability of obtaining outcome a_i is $p(a_i) = \mathrm{tr}(A_i\rho)$.

POVM elements are extremely useful in unambiguous state discrimination, meaning that given a measurement outcome, we either cannot determine the state at all, or we must identify it with full confidence. For example, assume that we are given either state $|\psi_1\rangle = |0\rangle$ or $|\psi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ with equal probability. The optimal unambiguous discrimination using only projective measurements is provided by $M_1 = |0\rangle\langle 0|$ and $M_2 = |1\rangle\langle 1|$, giving a 25% chance of identifying the state (obtaining the measurement value 1).

To see how POVMs provide a better strategy, we introduce $|\phi_1\rangle = |1\rangle$ and $|\phi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ such that $\langle\phi_i|\psi_i\rangle = 0$ for any i are orthogonal. A set of POVMs can be constructed by defining $A_1 = c|\phi_1\rangle\langle\phi_1|$ and $A_2 = c|\phi_2\rangle\langle\phi_2|$. Lastly, for completeness to hold, define an additional POVM element $A_3 = \mathbb{I} - A_1 - A_2$.

Suppose that the state $|\psi_1\rangle$ was given. The probability of obtaining measurement outcome a_1 is $\mathrm{tr}(A_1|\psi_1\rangle\langle\psi_1|) = \mathrm{tr}(c|\phi_1\rangle\langle\phi_1|\psi_1\rangle\langle\psi_1|) = 0$, implying that if the outcome is a_1, the state given must have been $|\psi_2\rangle$ instead. Also, if a_2 is obtained, the state can only be $|\psi_1\rangle$. The value of c can be chosen to be optimal, where the only constraint is such that $A_3 = \mathbb{I} - A_1 - A_2$ must be positive. Solving gives $c \approx 0.585$, which implies that no matter which state is given, there is 29% chance of identifying it without mistake.

Besides unambiguous state discrimination, POVMs can also provide a better strategy in maximizing the guessing probability, $p_{guess} = \sum_i p_i \,\mathrm{tr}(A_i\rho_i)$. For example, given a set of three symmetric pure density matrices

$$\begin{aligned} \rho_k &= |\psi_k\rangle\langle\psi_k| \\ |\psi_k\rangle &= \sin\theta_k|1\rangle + \cos\theta_k|0\rangle, \end{aligned} \tag{2.14}$$

where $\theta_k = \frac{2k\pi}{3}, k = 0, 1, 2$, and each state is obtained with equal probability. Since any complete set of projective measurements can only consist at most two elements $\{P_0, P_1\}$, and since all states are non-orthogonal to each other, $p_{guess} = \frac{1}{3}[\mathrm{tr}(P_0\rho_i) + \mathrm{tr}(P_1\rho_j)] < \frac{2}{3}$ for any distinct $i, j \in \{0, 1, 2\}$. However, it has been shown in [2] that the optimal guessing

[2] A positive matrix implies that all its eigenvalues are real and positive.

probability can be obtained by the square-root operators

$$A_k = \frac{1}{3}\ \rho^{-\frac{1}{2}}\rho_k\rho^{-\frac{1}{2}}, \tag{2.15}$$

with $\rho = \frac{1}{3}\sum_{i=0}^{2}\rho_i$ denoting the average state. For this example, $A_k = \frac{2}{3}|\psi_k\rangle\langle\psi_k|$ are *weighted* projective operators, while $p_{guess} = \frac{2}{3}$.

In summary, we see that POVM measurements can provide a much better strategy compared to projective measurements.

2.2 Entropic uncertainty relations

The term "entropy" is no stranger to any physics educated individual. We are familiar with the assumption of Equal A Priori Probability in statistical mechanics: for an isolated system in thermal equilibrium, each microstate appears to be equally accessible. The system reaches a state of full (scaled) entropy $S = kT\ln\Omega$, Ω being the multiplicity of the system, i.e. all possible ways for energy to be distributed amongst the particles.

In information theory, entropy is also commonly used to quantify the uncertainty of information. In fact, there exists various classes of such entropic measures. In this section, we introduce a few important quantities and discuss their physical significance.

2.2.1 Shannon and von Neumann entropy

For a system A described by the density matrix ρ_A, the von Neumann entropy is

$$S(A)_{\rho_A} = -\operatorname{tr}[\rho_A\log\rho_A]. \tag{2.16}$$

The term log here and throughout our description refers to the binary logarithm $\log_2$. Also, given a density matrix which can be diagonalized in the form $\rho = \sum_i E_i|e_i\rangle\langle e_i|$, with corresponding eigenvalues E_i, then $\log(\rho) = \sum_i \log(E_i)|e_i\rangle\langle e_i|$.

The von Neumann entropy is a generalisation of the Shannon entropy. If ρ is pure, then $S(\rho) = 0$. If ρ is a mixed state, by performing diagonalization we can always rewrite ρ as a statistical distribution over pure orthonormal quantum states $\{p_i, |\phi_i\rangle\}$. The von Neumann entropy is then equal to the Shannon entropy of a random variable X over the distribution $\{p_i, x_i\}$,

$$H(X)_\rho = -\sum_i p_i\log p_i. \tag{2.17}$$

Having correlated systems A and B, it is useful to describe the entropy of A conditioned

on B. Suppose we have a joint state ρ_{AB}. Then the conditional von Neumann entropy is

$$\mathrm{H}(A|B)_{\rho_{AB}} := \mathrm{H}(AB) - \mathrm{H}(B), \tag{2.18}$$

where $\rho_B = \mathrm{tr}_A(\rho_{AB})$ denotes the partial trace of ρ_{AB} over system A.

The Shannon entropy is used in many applications as a standard entropy measure. It considers the asymptotic fraction of randomness contained within infinite instances of independent and identically distributed (i.i.d.) events. For example, consider a source emitting piece of mutually independent data $\{X_1, X_2, ..., X_n\}$ according to probability distribution P_X. Let $X^n = X_1 X_2 \cdots X_n$. For some encoding of $\mathrm{Enc}(X^n)$ with length Rn, R is the data compression rate. Shannon's source coding theorem states that the theoretical lower limit of data compression rate such that given $\mathrm{Enc}(X^n)$, X^n can be recovered completely, is given by the Shannon entropy $\mathrm{H}(P_X)$ itself.

However, the usage of Shannon entropy has setbacks in cryptography, since it fails to capture extreme scenarios and correlations between events. To see why this is the case, consider the probability distribution of a random string $X^n \in \{0,1\}^n$. If the probability distribution is uniform, $\mathrm{H}(X^n) = n$. Now, let us consider a probability distribution such that

$$\mathrm{Prb}(X^n = Y) = \begin{cases} \frac{1}{2} & \text{if } Y = \mathbf{0^n} \text{ is the zero string of length } n, \\ \frac{1}{2} \cdot \frac{1}{2^n - 1} & \text{otherwise.} \end{cases} \tag{2.19}$$

The Shannon entropy $\mathrm{H}(X^n) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2^n-1} \approx \frac{1}{2}n$ still scales with the string length, and is large when n is large. Yet if we were to guess X^n, we have a staggering 50% chance of guessing it correctly! This is the motivation for introducing another entropy measure that gives a better measure of randomness.

2.2.2 Min entropy

Given a density matrix for the system A denoted as ρ_A, the min-entropy is defined as

$$\mathrm{H}_{\min}(A)_{\rho_A} := \sup \ \{\lambda \in \mathcal{R} : 2^{-\lambda}\ \mathbb{I}_A \geq \rho_A\} \tag{2.20}$$

where $\mathbb{I}_A$ is the identity matrix with dimension of system A. Essentially this is the negative logarithm of the largest eigenvalue of ρ_A. For a pure state, $\mathrm{H}_{\min}(A)_{\rho_A} = 0$, while for a mixed state diagonalized with respect to the eigenbasis $\{|\phi_i\rangle\}$ with respective eigenvalues $\{p_i\}$,

$$\mathrm{H}_{\min}(A)_{\rho_A} = -\log \ \max_i \ p_i. \tag{2.21}$$

Returning to the previous example of having the probability distribution in (2.19), we see that $\mathrm{H}_{\min}(X)_{\rho_X} = -\log\ \frac{1}{2} = 1 \ll n$. Instead of averaging over the whole distribution, the min-entropy captures the most probable event, making it useful while considering single-shot scenarios.

Given a joint state of ρ_{AB}, the conditional min-entropy is defined as

$$\mathrm{H}_{\min}(A|B)_{\rho_{AB}} := \max_{\sigma_B}\ \mathrm{H}_{\min}(A|B)_{\rho_{AB}|\sigma_B} \tag{2.22}$$

$$\mathrm{H}_{\min}(A|B)_{\rho_{AB}|\sigma_B} := \sup\ \{\lambda \in \mathcal{R} : 2^{-\lambda}\ \mathbb{I}_A \otimes \sigma_B \geq \rho_{AB}\}. \tag{2.23}$$

where the maximization is over all positive definite matrices $\sigma_B \in \mathcal{H}_B$.

Min-entropy conditioned on classical information

The min-entropy has a very useful physical significance, which links this quantity to the best strategy available for distinguishing quantum states. This can be seen in the theorem as stated below:

Theorem 2.2.1 (Min-entropy,[20]). *Let* $\rho_{XB} = \sum_i p_i |x_i\rangle\langle x_i| \otimes \rho_B^{x_i}$ *be a cq-state which is classical on X. Then*

$$\mathrm{H}_{\min}(X|B)_\rho = -\log p_{guess}(X|B)_\rho$$

$$p_{guess}(X|B)_\rho := \max_{\{E_B^{x_i}\}_i} \sum_i p_i\ \mathrm{tr}(E_B^{x_i}\rho_B^{x_i}). \tag{2.24}$$

where the maximization is over all possible sets of POVM elements $\{E_B^{x_i}\}_i$.

Theorem 2.2.1 is proven by using semi-definite programming methods, which we shall omit here. Essentially, if we acquire a state ρ from the ensemble $\{\rho_B^{x_i}\}$, by using the optimal strategy of generalised measurements to identify ρ we can attempt to guess X. $\mathrm{H}_{\min}(X|B)_\rho$ tells us the negative logarithm of this guessing probability.

We have seen that min-entropy considers the extreme case scenario, while Shannon entropy averages over the probability distribution. However, let us consider the probability distribution of random string $X^n \in \{0,1\}^n$ given as

$$\mathrm{Prb}(X_i = 1) = 0.2, \quad \mathrm{Prb}(X_i = 0) = 0.8 \qquad \text{for } \forall i \in [1, n]. \tag{2.25}$$

The min-entropy and Shannon entropy are evaluated as $\mathrm{H}_{\min}(X^n)_{\rho_X^{\otimes n}} = 0.322n$ and $\mathrm{H}(X^n)_{\rho_X^{\otimes n}} = 0.722n$ respectively, where their difference scales with n. This comes from the fact that min-entropy considers only the individually most probable event: X^n equals to the zero string of length n. If n were small, X^n is easy to guess, hence it is reasonable

that $\mathrm{H}_{\min}(X^n)$ is small. However as one might see, X^n becomes harder to guess as $n \to \infty$. The smooth min-entropy offers a bridge between the min and Shannon entropy, ignoring the most probable events up to some amount ϵ. It is defined as

$$\mathrm{H}^{\varepsilon}_{\min}(A)_{\rho_A} := \max_{\rho'} \ \mathrm{H}_{\min}(A)_{\rho'_A} \tag{2.26}$$

$$\rho'_A \in \{\rho \in \mathbb{C}^{n\times n} : C(\rho, \rho_A) = \sqrt{1 - F^2(\rho_A, \rho'_A)} \leq \epsilon\}. \tag{2.27}$$

where

$$F(\rho, \rho') = \mathrm{tr}\left(\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}\right), \tag{2.28}$$

maximizing over all density matrices ρ'_A which are ϵ-close to ρ_A.[3] Tabulating these quantities for different values of n w.r.t. the distribution specified in (2.25)

n	$\frac{1}{n}\,\mathrm{H}(X^n)$	$\frac{1}{n}\,\mathrm{H}^{0.01}_{\min}(X^n)_{\rho_X^{\otimes n}}$	$\frac{1}{n}\,\mathrm{H}_{\min}(X^n)_{\rho_X^{\otimes n}}$
10	0.722	0.336	0.322
150	0.722	0.542	0.322
500	0.722	0.594	0.322
1000	0.722	0.630	0.322

shows that for small n, the smooth min-entropy is close to the min-entropy, while as n grows large for i.i.d. events, it grows closer to the Shannon entropy. Given this, the smooth min-entropy manifests itself as the most suitable measure for cryptographic purposes.

The conditional smooth min-entropy is defined as

$$\mathrm{H}^{\varepsilon}_{\min}(A|B)_{\rho_{AB}} := \max_{\rho'_{AB}} \ \mathrm{H}_{\min}(A|B)_{\rho'_{AB}} \tag{2.29}$$

$$\rho'_{AB} \in \{\rho \in \mathbb{C}^{n\times n} : C(\rho, \rho_{AB}) = \sqrt{1 - F^2(\rho, \rho_{AB})} \leq \epsilon\}. \tag{2.30}$$

maximizing over all density matrices ρ'_{AB} which are ϵ-close to the state ρ_{AB}.

Another way to understand this quantity is to say that it guarantees at least an amount of min-entropy, *except for* a probabilistic error parameter of ϵ.

Smooth min-entropy conditioned on quantum information

Various general properties of the smooth min-entropy has been shown in [30]. Smooth min-entropy relations conditioned on classical information are relatively easy to obtain.

[3] The quantity $C(\rho, \rho')$, commonly referred to as the purified distance, is an important distance measure in quantum information theory. For commuting density matrices ρ, ρ' with ordered eigenvalues $\{p_i\}$ and $\{q_i\}$, it is equal to the trace distance $\frac{1}{2}\sum_i |p_i - q_i|$ between them.

Meanwhile, it is considerably harder to quantify smooth min-entropy conditioned on quantum side information. In the cases of bounded quantum storage model, one might consider bounding the smooth min-entropy by using a useful chain rule:

Theorem 2.2.2 (Chain rule of smooth min-entropy,[30]). *Given a joint system ρ_{ABQ}, and $\rho_{AB} = \mathrm{tr}_Q(\rho_{ABQ})$ denotes the reduced density matrix of ρ_{ABQ} taking partial trace over system Q. Then the following holds:*

$$\mathrm{H}_{\min}^{\varepsilon}(A|BQ)_{\rho_{ABQ}} \geq \mathrm{H}_{\min}^{\varepsilon}(AB|Q)_{\rho_{AQ}} - \log|\mathcal{B}|, \tag{2.31}$$

where $|\mathcal{B}|$ size of support of B[4]. Furthermore, if A and B are classical, (2.31) *reduces to*

$$\mathrm{H}_{\min}^{\varepsilon}(A|BQ)_{\rho_{ABQ}} \geq \mathrm{H}_{\min}^{\varepsilon}(A|Q)_{\rho_{AQ}} - \log|\mathcal{B}|. \tag{2.32}$$

In the more general case of noisy-storage model, define the set of valid density matrices in a Hilbert Space $\mathcal{H}$ as $\mathcal{B}(\mathcal{H})$. Then a quantum memory (or channel) can be modelled as a completely-positive trace-perserving map (CPTPM) $\mathcal{F} : \mathcal{B}(\mathcal{H}_{in}) \to \mathcal{B}(\mathcal{H}_{out})$. These maps represent the most general quantum operations that can be performed upon a quantum state. $P_{succ}^{\mathcal{F}^{\otimes n}}(Rn)$ is the success probability of recovering Rn classical bits which were chosen uniformly at random and sent through n uses of the channel, via any encoding/decoding scheme. Its formal definition is given by

$$P_{succ}^{\mathcal{F}^{\otimes n}}(Rn) := \max_{\{\rho_x, E_x\}} \frac{1}{2^{Rn}} \sum_x \mathrm{tr}(E_x \rho_x). \tag{2.33}$$

The quantity $P_{succ}^{\mathcal{F}^{\otimes n}}(Rn)$ of a quantum channel is determined by its *classical capacity*, $C_{\mathcal{F}}$. For some quantum channels, this can be shown to have a transition that sharpens in the large n limit: for $R < C_{\mathcal{F}}$, $P_{succ}^{\mathcal{F}^{\otimes n}}(Rn) < 1$ is a trivial bound, whereas for $R > C_{\mathcal{F}}$, $P_{succ}^{\mathcal{F}^{\otimes n}} < 2^{-n\gamma^{\mathcal{F}}(R)}$, where $\gamma^{\mathcal{F}}(R) > 0$, meaning that via any encoding scheme, it becomes exponentially improbable to recover the nR bits sent through the quantum channel.

This property of channels is known as the *strong converse*, which is not known to be true for quantum channels in general. This is because the classical capacity of quantum channels are not necessarily additive. In other words, entanglement between states which are sent down two copies of the channel might significantly assist the preservation of states. It has been shown in [18] that strong converse holds for any generalized depolarizing

[4]The support of a probability distribution $P(X)$ is the set of possible values X can take, i.e. the set of $x \in \mathcal{X}$ where $P(X = x) > 0$. It has a maximum value of d_X, the dimension of $\mathcal{X}$.

channel and quantum unital[5] channels. To see some examples, the classical capacity of a quantum identity channel for qubits, i.e. $\mathcal{F}(\rho) := \rho$ preserving the state completely has a classical capacity $C_{\mathcal{F}} = 1$, while for the depolarizing channel of dimension d defined as:

$$\mathcal{F}_r(\rho) := r\rho + (1-r)\frac{\mathbb{I}}{d} \qquad \text{for a fixed value of } 0 \leq r \leq 1, \tag{2.34}$$

has a classical capacity $C_{\mathcal{F}_r} = \log d + (r + \frac{1-r}{d}) \log(r + \frac{1-r}{d}) + (d-1)\frac{1-r}{d} \log \frac{1-r}{d}$. Both of these channels satisfy the strong converse property.

Having defined all these quantities, the theorem stated as below then holds:

Theorem 2.2.3 (Conditioning on quantum information, [19])**.** *Given a state ρ_{ABQ}, where A and B are classical, and Q denotes some quantum information. Let $\epsilon, \epsilon' \geq 0$ be arbitrary, and let $\mathcal{F} : \mathcal{B}(\mathcal{H}_{in}) \to \mathcal{B}(\mathcal{H}_{out})$ denote any CPTPM. Then*

$$\mathrm{H}_{\min}^{\epsilon+\epsilon'}(A|B\mathcal{F}(Q)) \geq -\log P_{succ}^{\mathcal{F}} \left(\lfloor \mathrm{H}_{\min}^{\epsilon}(A|B) - \log \frac{1}{\epsilon'} \rfloor \right). \tag{2.35}$$

By Theorem 2.2.3, if the channel $\mathcal{F}$ a satisfies strong converse, then in the asymptotic limit, once the number of bits nR sent through the quantum channel exceeds its classical capacity, it becomes exponentially improbable to recover the classical information.

In summary, we see that Theorem 2.2.2 and Theorem 2.2.3 provide lower bounds for the smooth min-entropy further conditioned on quantum information $\mathcal{F}(Q)$, for the case where the state is sent through a quantum channel $\mathcal{F}$. A quantum memory can be viewed as such a channel. With this, we have a sufficient description of the smooth min-entropy conditioned on both classical and quantum information.

2.2.3 Quantum mechanical uncertainty in entropic terms

The uncertainty in quantum measurements is governed by the generalised Heisenberg uncertainty principle, which states that two physical observables with non-commuting operators cannot be simultaneously determined for any quantum state. More precisely, let A and B represent two Hermitian operators in a n-dimensional Hilbert space $\mathcal{H}_n$, with corresponding eigenbases $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$ for $1 \leq i \leq n$. Given a quantum state ψ,

$$\Delta_\psi A \; \Delta_\psi B \geq \frac{1}{2} \left| \langle [\hat{A}, \hat{B}] \rangle_\psi \right| \tag{2.36}$$

[5] Unital channels map the completely mixed state from the input Hilbert space to the completely mixed state of its output Hilbert space.

where $\Delta_\psi A$ and $\Delta_\psi B$ denote the standard deviation of measurement outcomes represented by operators $\hat{A}$ and $\hat{B}$.

This conventional method of describing quantum-mechanical uncertainty has a setback: it is dependent on both measurement operators and the quantum state itself. Entropic uncertainty relations, in contrast, have been developed to describe the inherent uncertainty caused by measurement operators, and are independent of the quantum state.

Lemma 2.2.4. *Assuming that a state ρ is given, where L possible measurements $\theta \in \Theta$ can be performed upon the state, with outcomes labelled $x \in \mathcal{X}$. Each measurement is chosen with equal probability. Let $p_{x|\rho,\theta}$ denote the probability of observing outcome x when making the measurement labelled θ on the state ρ. Then*

$$\mathrm{H}(X|\Theta)_{\rho_{X\Theta}} = \frac{1}{L}\sum_{\theta} \mathrm{H}(X|\Theta=\theta). \tag{2.37}$$

where $\mathrm{H}(X|\Theta=\theta) = -\sum_x p_{x|\rho,\theta}\ \log p_{x|\rho,\theta}$.

Proof. Since X and Θ are classical, the state $\rho_{X\Theta}$ can be expressed as

$$\rho_{X\Theta} = \sum_x \sum_\theta \frac{1}{L}\ p_{x|\rho,\theta}\ |x\rangle\langle x| \otimes |\theta\rangle\langle\theta|. \tag{2.38}$$

By the definition of conditional von Neumann entropy in (2.18),

$$\begin{aligned}
\mathrm{H}(X|\Theta)_{\rho_{X\Theta}} &\equiv \mathrm{H}(X\Theta)_{\rho_{X\Theta}} - \mathrm{H}(\Theta)_{\rho_\Theta} \\
&= \sum_{x,\theta} -\frac{1}{L}\ p_{x|\rho,\theta}\ \log \frac{1}{L} p_{x|\rho,\theta} + \sum_\theta \frac{1}{L}\log\frac{1}{L} \\
&= \sum_{x,\theta} -\frac{1}{L}\ p_{x|\rho,\theta} \log\frac{1}{L} - \sum_{x,\theta}\frac{1}{L} p_{x|\rho,\theta}\ \log p_{x|\rho,\theta} + \sum_\theta \frac{1}{L}\log\frac{1}{L} \\
&= \sum_\theta \frac{1}{L}\ \mathrm{H}(X|\Theta=\theta).
\end{aligned} \tag{2.39}$$

□

This lemma shows that an uncertainty relation in terms of von Neumann entropy is given by the average over individual entropies conditioned upon a particular measurement θ. Such a relation states that for *all* states ρ ,

$$\mathrm{H}(X|\Theta)_\rho \geq c \tag{2.40}$$

where c depends solely on measurements. For example, consider a single qubit subjected to BB84 measurements. Since σ_x and σ_z are mutually unbiased, for any state where $H(X|\Theta=0)_\rho = 0$, we must required that $H(X|\Theta=1)_\rho > 0$ holds true. [25] has shown that $c = \frac{1}{2}$ regardless of the state. Besides the von Neumann entropy, such conditional

entropic uncertainty relations can be obtained for Rényi entropic measures, as we shall see in Chapter 3.

2.3 2-Universal Hash Functions

Sources that produce partial randomness are abundant, however pure randomness is hard to obtain. The basic principle of randomness extractors is to distill close-to-perfect randomness from a partially random source. Hash functions are a special type of randomness extractors, which have a wide usage in cryptography.

To understand on how this is done, consider a binary string X^n of length n, where a system B holds partial information about X^n. The conditional smooth min-entropy is lower bounded by some $\mathrm{H}_{\min}^{\varepsilon}(X^n|B) > k \cdot n > 0$. [6] A two-universal hash function $\mathrm{Ext} : \{0,1\}^n \otimes \mathcal{R} \to \{0,1\}^l$ picks a random seed $r \in \mathcal{R}$, and maps X^n to a shorter string $S = Ext(X^n, R)$ of length l, such that S is almost completely randomized. For example, if $l = 1$, randomness extraction can be achieved by generating a random seed that determines a substring of X^n, and computing the parity of bits for the substring. Intuitively, the parity of this substring will be completely random, if there is at least one bit in the substring which is unknown. The maximum length l such that this is possible is determined by the amount of smooth min-entropy available.

A formal description of privacy amplification by using 2-universal hash functions is given as below:

Theorem 2.3.1 (Privacy Amplification,[31]). *Consider a set of 2-universal hash functions* $\mathrm{Ext} : \{0,1\}^n \otimes \mathcal{R} \to \{0,1\}^l$, *where R is a random variable uniformly distributed on $\mathcal{R}$, and independent of X^n. Any joint state of X^n, the system B and the random seed R is described by $\rho_{X^nBR} = \rho_{X^nB} \otimes \tau_{\mathcal{R}}$, where τ denotes a uniform probability distribution. Then*

$$\rho_{Ext(X^n,R)RB} \approx_{\epsilon'} \tau_{\{0,1\}^l} \otimes \rho_{RB}, \text{[7]} \quad \textit{for } \epsilon' = 2\epsilon + 2^{-\frac{1}{2}(H_{min}^{\epsilon}(X^n|B)-l)-1}, \epsilon \geq 0. \tag{2.41}$$

where $\mathrm{H}_{\min}^{\varepsilon}(X|B)$ *is the smooth min-entropy of Bob's information about X^n.*

By this theorem, we see that the distribution of $\mathrm{Ext}(X^n, R)$ is close to uniform over the range $\{0,1\}^l$, and independent of the joint system of B and the random seed R. This

[6] The information B holds is allowed to be quantum or classical in general.

[7] The notation of state $\rho' \approx_\epsilon \rho$ means that the trace distance between the two matrices $\|\rho - \rho'\| = \frac{1}{2}\,\mathrm{tr}(\sqrt{A^\dagger A}) \leq \epsilon$, where $A = \rho - \rho'$.

implies that even if the random seed is known, $\text{Ext}(X^n, R)$ remains close to uniform. These functions can be efficiently constructed by using simple randomized linear functions, and we will use them subsequently.

2.4 Error-correcting codes

In a setting where information transfer is subjected to noise, encoding/decoding schemes provide ways to correct errors that occur in transmission. Essentially, error-correcting codes achieve this by sending extra information across the noisy channel.

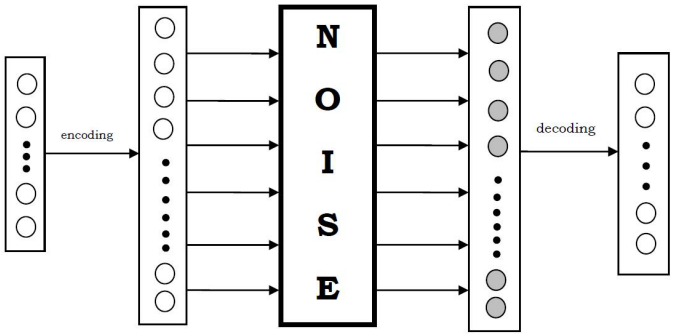


Figure 2.1: A diagram showing the encoding and decoding of message, such that k elements were encoded into n elements and sent through the noisy channel, then recovered completely after undergoing the transmission process.

The information sent through is usually defined over a finite field. Thus, the error-correcting code used specifies the alphabet (number of distinct elements) of the finite field, which information is defined in. For our purposes, we place the main focus on introducing *linear* codes defined over the *binary* field.

2.4.1 General properties of binary linear codes

An error-correcting code C can be completely specified by either of the two: its generator matrix usually denoted as G, or parity check matrix, usually denoted as H. G is the kernel of H transposed, meaning that $G \cdot H^T = \mathbf{0}$ yields the zero matrix.

For each error-correcting code specified, there are three important parameters: the **block length** n, **code rate** R and **minimum distance** d. A binary linear code has a generating matrix with dimensions $Rn \times n$. It can always be written in the standard row-echelon form $G = [\, \mathbb{I}_{Rn} \mid P \,]$, where P is a binary matrix of dimensions $Rn \times (1-R)n$. All the possible linear combinations of rows in the generating matrix form a set of **codewords**. Hence, given a code with rate R, the total number of codewords is 2^{Rn}. The parity

check matrix H, on the other hand, has dimensions $(1-R)n \times n$, and is given by $H = [\ P^T \mid \mathbb{I}_{(1-R)n}\ ]$.

We work out a small example to see how error-correcting codes provide a way to transmit information reliably. Consider the generating matrix

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix},$$

with its corresponding parity check matrix is given by

$$H = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

One can verify straightforwardly by matrix multiplication that $G \cdot H^T = \mathbf{0}$. The code rate is given as $R = \frac{4}{7}$.

Now, suppose we have a message comprising of 4 bits, $M \in \{0,1\}^4$. The encoding scheme specified by this code is to compute $w = M \cdot G$. For example, given $M = [1,0,0,1]$, the encoded message reads $c = [1,0,0,1,0,0,1]$, which is transmitted down the noisy channel. Next, we define the function that outputs the parity check syndrome to be $\text{Syn} : \{0,1\}^n \to \{0,1\}^{n-k}$, such that $\text{Syn}(c) = c \cdot H^T$. Note that for any message M, the corresponding encoded message w is essentially a linear combination of rows of G, and since the code is linear, we conclude that $\text{Syn}(w) = 0$.

After sending w through a noisy channel, let the corrupted message be $c' = [1,0,0,1,0,0,0]$, where one bit was flipped. We find that $\text{Syn}(c') = [0,0,1]$, concluding that errors have indeed occurred during transmission. Various methods can be used to correct the errors. For example, assuming that the number of errors do not exceed a certain limit, a list of all possible errors that give rise to the corresponding syndrome can be constructed. This method is called syndrome decoding. Also, algorithms can be designed to output the nearest codeword possible.

The maximum number of errors correctable depend on the construction of the code. This is governed by the **minimum distance** d, which is defined as the minimum Hamming

weight[8] over all codewords specified by the error-correcting code.

Theorem 2.4.1 (Error-correcting capabilities of codes, [17]). *Given an error-correcting code $\mathcal{C}$ that has a minimum distance d. Then the code $\mathcal{C}$ can always detect up to d-1 number of errors, while always being able to correct up to $t = \lfloor \frac{d-1}{2} \rfloor$ errors.*

This theorem states that the minimum distance directly bounds the ability of the code to detect errors. To see why this is the case, recall that detecting errors come from the fact that syndromes before and after transmission do not match. Now, consider sending a string b over the channel, and let e be a codeword which Hamming weight is exactly d. Thus, $\text{Syn}(e) = \mathbf{0}$. If the codeword sent through the channel gets corrupted such that $b' = b \oplus e$, calculation of the syndrome $\text{Syn}(b') = \text{Syn}(b) + \text{Syn}(e) = \text{Syn}(b)$ reveals that the syndrome of b' matches that of b, hence the code fails to detect errors. On the other hand, since there is no codeword with Hamming weight smaller than d, any number of errors less than d can be detected by showing that the syndrome is non-zero.

The error-correcting capability of codes can also be seen, by noting that error-correcting codes work by picking out the nearest string that has the correct syndrome. Returning to the above example, we have two strings b and $b \oplus e$ that have identical syndromes. Let us assume any two strings x, y with Hamming weight $\Delta(x) = t + 1 = \lfloor \frac{d-1}{2} \rfloor + 1$ and $\Delta(y) = d - \Delta(x)$ respectively, such that $x \oplus y = e$. Note that $\Delta(y) \leq \Delta(x)$. Now suppose that b is transmitted through the channel such that the output is $b'' = b \oplus x$. Then b'' is closer to $b \oplus e$ compared to b, and the code fails to recover b.

A good error-correcting code should have both high code rate and minimum distance. Furthermore, for codes of asymptotic block length n, it is favourable to have a minimum distance which scales with n, hence the relative minimum distance $\delta = \frac{d}{n}$ is introduced. The higher both R and δ are, the more efficient is the code performance in information transfer and error-correcting respectively. However, there is a trade-off between these parameters, and much work has been done in attempt to develop explicit constructions of codes that produce optimal parameters of R and δ.

2.4.2 Explicit constructions

We first have a quick look at the theory of binary concatenated codes. These codes, introduced by Forney in [13] are constructed by combining two codes: an outer code

[8] The Hamming weight of a binary string X denotes the non-zero elements in the string, often denoted as $\Delta(X)$.

defined over a finite field F_{out} of larger alphabet, with an inner binary code with relatively smaller block length. The concatenation works as follows: if F_{out} has an alphabet $q = 2^m$, each element in F_{out} can be mapped by an isomorphism to m binary bits. The inner code then performs an encoding on the m bits for each element in F_{out}.

These codes have the advantage that they can be constructed to encode and decode efficiently, also giving a definite statement about the minimum distance, which is generally hard to evaluate for large block lengths.[9] Furthermore, outer codes over large finite fields generally give a good trade-off between rate and minimum distance. For example, the Reed-Solomon (RS) code provides the optimal trade-off, such that given an RS code with block length n and code rate R, the minimum distance is given by $\delta = (1-R)+\frac{1}{n}$. On the other hand, a brute force search can be done for binary codes over a relatively small inner block length. By concatenation of such two codes, one can hope to achieve an overall code with good performance.

Theorem 2.4.2 (Concatenated codes, [17])**.** *Given an outer code C_{out} with parameters $[n_1, R_1, \delta_1]$, and inner code C_{in} with parameters $[n_2, R_2, \delta_2]$, the concatenated code $C = C_{out} \circ C_{in}$ has parameters n, R, δ where*

$$n = n_1 \cdot n_2, \qquad R = R_1 \cdot R_2, \qquad \text{and} \quad \delta \geq \delta_1 \cdot \delta_2. \tag{2.42}$$

Note that instead of a precise value, a lower bound is obtained for δ. This bound is highly non-optimal, since the explicit value of δ depends on the individual outer and inner codes themselves.

2.4.3 Randomized constructions

As seen previously, explicit constructions give definite lower bounds on the minimum distance, but these bounds are generally non-tight. One of the more useful bounds is given by Gallager for random constructed binary linear codes, which means that each element of the matrix P is chosen with uniform probability over the binary field.

Theorem 2.4.3 (Minimum distance of a linear binary code, [14])**.** *Given a linear binary code defined by its parity check matrix $H = [\ P \mid \mathbb{I}_{(1-R)n}\ ]$, where P is a randomly constructed binary matrix of dimensions $(1-R)n \times Rn$. Then the probability that the code $\mathcal{C}$*

[9] While the code rate is easy to obtain by observing the dimensions of either G or H, the exact value of minimum distance is known to be a non-polynomial hard problem which is practically impossible to evaluate for asymptotically large block lengths.

specified by H has minimum distance $d \leq \delta n$ *is lower bounded by*

$$Prb(d \leq \delta n) \leq 2^{(R-C_\delta)n} \tag{2.43}$$

where $C_\delta = 1 - h(\delta)$, $h(x) = x \log x + (1-x)\log(1-x)$ *being the binary entropy defined over* $x \in [0, 1]$.

A detailed proof of this theorem is shown in the appendix. This theorem tells us that for $R < C_\delta$, the probability that $d \leq \delta n$ is exponentially small w.r.t. the block length n. This theorem will be used later in our security proof in Chapter 4. Although random codes are seldom used because they are generally hard to decode, for our use we only require the properties of minimum distance, while decoding is not necessary. A similar bound is also obtained for Gallager's class of low-density parity check (LDPC) codes, which will also be discussed later.

Chapter 3

Smooth Min-entropy relations

Measurements in BB84 and six state bases[1] are very common in quantum cryptographic protocols. Being mutually unbiased, they provide entropic uncertainty relations with a large lower bound, which renders them useful in cryptographic settings, particularly in randomness distribution. Besides in QKD, they are also used in all two-party cryptographic protocols in the bounded [10, 9] and noisy-storage model [38, 35, 19].

The security of all protocols in this model crucially rests on the existence of uncertainty relations in terms of smooth min-entropy. For a state of n qubits subjected to BB84 measurements, a binary string X^n of measurement outcomes can be obtained. The central question of interest is to find an uncertainty relation for smooth min-entropy $\mathrm{H}_{\min}^{\varepsilon}(X^n|\Theta^n)_\rho$, that holds for any quantum state $\rho \in \mathcal{H}_2^{\otimes n}$. In particular, any form of entanglement is allowed between the qubits, and the state can be either pure or mixed. In a more general case, it is desirable to have $\mathrm{H}_{\min}^{\varepsilon}(X^n|\Theta^n K)_\rho$, where K is some additional classical information about the quantum state ρ.

To see why this is important, let us assume the following setting: Alice selects a string X^n where each bit value is chosen at random. For all n bits, she encodes each bit in a identically and independently chosen BB84 basis, and sends the qubits to Bob. In practice, she does so by creating entangled photon pairs and keeping half on her side as ρ_A, while sending the other half as an n qubit state ρ_B to Bob. An honest Alice measures her state with i.i.d. BB84 bases. As for any general strategy of cheating Bob, he is capable of performing instantaneous quantum operations (for example, entanglement swapping or arbitrary unitary operations) and performing any type of classical measurement on

[1]Recall that BB84 measurements are defined by the Pauli spin operators σ_x and σ_z, while six state measurements are defined by the full set of Pauli spin operators $\sigma_x, \sigma_y, \sigma_z$.

his quantum system (for example an optimal POVM measurement), thus obtaining some classical information K.

After a certain time t, Alice sends Bob the basis information Θ^n used to encode these qubits. Based on K and Θ^n, Bob's uncertainty about the string X^n is quantified by the smooth min-entropy $\mathrm{H}_{\min}^{\varepsilon}(X^n|\Theta^n K)_{\rho_{AB}}$.

The same scenario can be described in the opposite direction: Bob prepares a state ρ_k with probability p_k for Alice, which she then performs i.i.d. BB84 measurements upon, obtaining X^n as a result. Due to non-signalling properties, the exact scenario is equivalent to such a description, i.e. the order in which Alice and Bob performs classical measurements do not affect the measurement statistics. It is then clear that the distribution of Alice's basis Θ^n and Bob's classical information K are independent of each other.

In this chapter, our goal is to obtain a bound for the smooth min-entropy of Bob's information about Alice's string X^n, such that

$$\frac{1}{n}\,\mathrm{H}_{\min}^{\varepsilon}(X^n|\Theta^n, K) \geq c'. \tag{3.1}$$

We first study previous developed smooth min-entropy uncertainty relations for BB84 measurements. Subsequently, we present a derivation for new uncertainty relations, and show the advantage of our findings compared to previous work.

3.1 Previous smooth min-entropy relations

It has been shown in previous works that the smooth conditional min-entropy satisfies an uncertainty relation, as presented in the theorem below:

Theorem 3.1.1 (Uncertainty relations for BB84 measurements, [9])**.** *Given any n-qubit state ρ subjected to BB84 measurements, K denoting some classical information about ρ, then for any $\delta \in (0, \frac{1}{2}]$,*

$$\mathrm{H}_{\min}^{\varepsilon}(X^n|\Theta^n K)_{\rho} \geq \left(\frac{1}{2} - \delta\right) n \;\; = c'\; n, \tag{3.2}$$

where

$$\varepsilon = \exp\left[-\frac{\delta^2 n}{512(2+\log\frac{2}{\delta})^2}\right]. \tag{3.3}$$

Technical details of this theorem are omitted here, but essentially this relation is obtained by lower bounding the smooth min-entropy by the *conditional von Neumann entropy*, minus a small fraction δ. This can be justified mathematically by constructing a

martingale difference sequence by the von Neumann entropy conditioned on all past measurement history, and applying the result from [25] that the conditional von Neumann entropy w.r.t. BB84 measurements is lower bounded by $h = \frac{1}{2}$. Applying the Azuma's inequality provides this bound. A similar relation for six-states hold, for $h = \frac{2}{3}$.

Using this relation in a cryptographic protocol only yields an error ε which is exponentially decreasing in n. With this, it has been concluded that this uncertainty relation is sufficient for a theoretical proof of principle for a lower bound on smooth min-entropy. Therefore, Theorem 3.1.1 is widely employed in the protocols of [9, 8, 37, 33].

3.2 Rényi entropic bounds for smooth min-entropy

In this section, we develop an alternate lower bound on the conditional smooth min-entropy for BB84 measurements. We complete this analysis in four steps: firstly, we evaluate a tight uncertainty relation in terms of the Rényi entropy $\mathrm{H}_\alpha(X|\Theta)_\rho$, when ρ is just an $n = 1$ qubit state. Secondly, we show by mathematical induction that this result can be extended to an uncertainty relation for $n > 1$ qubits. The third step is to reintroduce the conditioning on classical information K. Finally, we relate the Rényi entropies of order $\alpha \in (1, 2]$ to the smooth min-entropy.

3.2.1 A single qubit uncertainty relation

The conditional α-Rényi entropies are defined as

$$\mathrm{H}_\alpha(A|B)_{\rho|\sigma} := \frac{1}{1-\alpha} \log \mathrm{tr}[\ \rho_{AB}^\alpha (\mathbb{I}_A \otimes \sigma_B)^{1-\alpha}\ ]. \tag{3.4}$$

It is worth noting that $\mathrm{H}_\alpha(A|B)$ has various different definitions. However, this particular quantity is used because a specific bound on the smooth min-entropy with respect to this conditional quantity has been proven recently in [36]. Also, the quantity $\mathrm{H}_\alpha(A|B)_{\rho_{AB}|\rho_B}$ for $\rho_B = \mathrm{tr}_A(\rho_{AB})$ recovers the conditional von Neumann entropy $\mathrm{H}(A|B)_\rho$ defined in (2.18), as $\alpha \to 1$.

Denote the ordered eigenbases of σ_z and σ_x as $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$ respectively. The value $\theta = 0$ represents measurements in the σ_z basis, while $\theta = 1$ represents measurements in the σ_x basis. Then the measurement operators associated with these bases are given by $M_{x|\theta} = \mathbf{H}^\theta |x\rangle\langle x| \mathbf{H}^\theta$, with $\mathbf{H}$ being the Hadamard matrix. For strings $\boldsymbol{x}, \boldsymbol{\theta} \in \{0,1\}^l$, the measurement operator $M_{\boldsymbol{x}|\boldsymbol{\theta}} = M_{x_1|\theta_1} \otimes M_{x_2|\theta_2} \otimes \cdots \otimes M_{x_l|\theta_l}$ is just the tensor product

of individual operators for each bit and basis.

For a single qubit labelled with classical information $K = k$, and subjected to BB84 measurements,

$$\mathrm{H}_\alpha(X|\Theta)_{\rho_{X\Theta K=k}|\rho_{\Theta K=k}} = \frac{1}{1-\alpha}\log\ P_\alpha(X|\Theta), \tag{3.5}$$

where

$$\begin{aligned} P_\alpha(X|\Theta) &= \mathrm{tr}[\ \rho^\alpha_{X\Theta K=k}(\mathbb{I}_X \otimes \rho_{\Theta K=k})^{1-\alpha}\ ] \\ \rho_{X\Theta,K=k} &= \sum_{\theta,x} p_{x|\theta,K=k}\ |x\rangle\langle x| \otimes p_{\theta|K=k}\ |\theta\rangle\langle\theta| \\ &= \sum_{\theta,x} p_{x|\theta,K=k}\ |x\rangle\langle x| \otimes p_\theta\ |\theta\rangle\langle\theta| \end{aligned} \tag{3.6}$$

since both X and Θ are classical variables, and Alice chooses Θ independent of the quantum state ρ_k. The quantity $P_\alpha(X|\Theta)$ can then be evaluated as follows:

$$\begin{aligned} P_\alpha(X|\Theta) &= \mathrm{tr}\left[\left(\sum_{\theta,x} p^\alpha_{x|\theta,K=k}\ |x\rangle\langle x| \otimes p^\alpha_\theta|\theta\rangle\langle\theta|\right)\left(\mathbb{I}_X \otimes \sum_{\theta'} p^{1-\alpha}_{\theta'}|\theta'\rangle\langle\theta'|\right)\right] \\ &= \mathrm{tr}\left[\sum_{\theta,x} p^\alpha_{x|\theta K=k}|x\rangle\langle x| \otimes p_\theta|\theta\rangle\langle\theta|\right] \\ &= \sum_{\theta,x} p^\alpha_{x|\theta K=k}\cdot p_\theta. \end{aligned} \tag{3.7}$$

Note that $p_{x|\theta,K=k} := \mathrm{tr}(M_{x|\theta}\ \rho_k)$. To minimize the α-Rényi entropy for values of $\alpha \in (1,2]$, it it sufficient to maximize $P_\alpha(X|\Theta)$. To do so, we are first required to establish the following lemma, which will be used subsequently in the proof:

Lemma 3.2.1. *Given a function $g : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, where g is defined as*

$$g(a,s) := s\cdot[(1+a)^{s-1} + (1-a)^{s-1}] - \frac{1}{a}\cdot[(1+a)^s - (1-a)^s]. \tag{3.8}$$

Then $g(a,s) \geq 0$ over the range $a \in [0,1)$ and $s \in (0,1]$.

The proof of this lemma is done in the appendix. Essentially, this is achieved by expanding the function in Taylor's series and showing that each term in the series is positive. Based on Lemma 3.2.1, we are now ready to establish a theorem that gives a tight uncertainty relation for $\mathrm{H}_\alpha(X|\Theta)_{\rho|\rho}$.

Theorem 3.2.2. *Given any one-qubit density matrix ρ, and denote $\alpha \in (1,2]$. Then for*

BB84 measurements,

$$\mathrm{H}_\alpha(X|\Theta)_{\rho|\rho} \geq \frac{1}{\alpha-1}\,[\alpha - \log(2^{\alpha-1}+1)]. \tag{3.9}$$

Proof. Let $\alpha = 1+s$, where $s \in (0,1]$. Firstly, we evaluate the term

$$\begin{aligned} P_{1+s}(X|\Theta) &= \frac{1}{2}\sum_{\theta\in\{0,1\}}\sum_{x\in\{0,1\}} p_{x|\theta}^{1+s} \\ &= \frac{1}{2}\left[\mathrm{tr}(\rho|0\rangle\langle 0|)^{1+s} + \mathrm{tr}(\rho|1\rangle\langle 1|)^{1+s} + \mathrm{tr}(\rho|+\rangle\langle +|)^{1+s} + \mathrm{tr}(\rho|-\rangle\langle -|)^{1+s}\right] \\ &= \frac{1}{2^{2+s}}\,[(1+z)^{1+s} + (1-z)^{1+s} + (1+x)^{1+s} + (1-x)^{1+s}], \end{aligned} \tag{3.10}$$

where we denote $x := \mathrm{tr}(\sigma_x\rho)$ and $z := \mathrm{tr}(\sigma_z\rho)$. The third equality holds because of the following relations:

$$\begin{aligned} |0\rangle\langle 0| &= \frac{1}{2}\,(\mathbb{I}+\sigma_z)\,, \qquad |1\rangle\langle 1| = \frac{1}{2}\,(\mathbb{I}-\sigma_z), \\ |+\rangle\langle +| &= \frac{1}{2}\,(\mathbb{I}+\sigma_x)\,, \qquad |-\rangle\langle -| = \frac{1}{2}\,(\mathbb{I}-\sigma_x). \end{aligned} \tag{3.11}$$

For any density matrix ρ, the following Bloch sphere condition as stated in (2.8), we can therefore parametrize x and z by polar coordinates:

$$x = r\sin\phi, z = r\cos\phi \tag{3.12}$$

where $r \in [0,1]$, and $\phi \in [0, \frac{\pi}{2}]$. $P_\alpha(X|\Theta)$ can then be rewritten as a function depending on the variables s,r,and ϕ:

$$Q(s,r,\phi) = \frac{1}{2^{2+s}}[(1+r\cos\phi)^{1+s}+(1-r\cos\phi)^{1+s}+(1+r\sin\phi)^{1+s}+(1-r\sin\phi)^{1+s}]. \tag{3.13}$$

Evaluating the partial differential of $Q(s,r,\phi)$ with respect to r,

$$\frac{\partial Q(s,r,\phi)}{\partial r} = \frac{1+s}{2^{2+s}}[\cos\phi(1+r\cos\phi)^s - \cos\phi(1-r\cos\phi)^s + \sin\phi(1+r\sin\phi)^s - \sin\phi(1-r\sin\phi)^s]. \tag{3.14}$$

Since in the range of ϕ, $\sin\phi$ and $\cos\phi$ are positive, we obtain $\frac{\partial Q(s,r,\phi)}{\partial r} \geq 0$, which implies the maximum is attained at r=1.

Evaluating the partial differential of Q(s,r,ϕ) with respect to ϕ at r=1,

$$\begin{aligned} \frac{\partial Q(s,1,\phi)}{\partial \phi} &= \frac{1+s}{2^{2+s}}[-\sin\phi(1+\cos\phi)^s + \sin\phi(1-\cos\phi)^s + \cos\phi(1+\sin\phi)^s - \cos\phi(1-\sin\phi)^s] \\ &= \frac{1+s}{2^{2+s}}\{\sin\phi[(1-\cos\phi)^s - (1+\cos\phi)^s] + \cos\phi[(1+\sin\phi)^s - (1-\sin\phi)^s]\} \end{aligned} \tag{3.15}$$

For a stationary point of $Q(s,1,\phi)$, (3.15) is zero and the solution is obtained at $\phi =$

$\{0, \frac{\pi}{4}, \frac{\pi}{2}\}$. The characteristics of the endpoints $\phi = 0, \frac{\pi}{2}$ are the same, hence it suffices to analyse either. It remains to analyse the characteristic of these stationary points. To do so, we evaluate the second partial derivative at these points as a function of s:

$$f_1(s) = \left.\frac{\partial^2 Q(s,1,\phi)}{\partial\phi^2}\right|_{\phi=0} = \frac{1+s}{2^{1+s}}\left(s - 2^{s-1}\right), s \geq 0 \tag{3.16}$$

$$\begin{aligned} f_2(s) &= \left.\frac{\partial^2 Q(s,1,\phi)}{\partial\phi^2}\right|_{\phi=\frac{\pi}{4}} \\ &= \frac{1+s}{2^{2+s}}\left\{s\cdot\left[(1-a)^{s-1} + (1+a)^{s-1}\right] - \sqrt{2}\left[(1+a)^s - (1-a)^s\right]\right\}. \end{aligned} \tag{3.17}$$

where $a = \frac{1}{\sqrt{2}}$. To determine if the stationary point is a local minima or maxima, we show the positivity/negativity of these functions over the interval $s \in (0, 1]$. Note that $f_1(0) = -1$ and $f_1(1) = 0$, while $f_1(s)$ is always increasing since $\frac{\partial f_1(s)}{\partial s} = 1+(1-s)2^{s-2} \geq 0$. Hence $f_1(s)$ is negative, implying the endpoints correspond to a local maxima.

On the other hand, note that the second term in (3.17) is exactly of the form $g(a, s)$ as stated in Lemma 3.2.1. With this, we conclude that the point $\phi = \frac{\pi}{4}$ is a local minimum. This leaves the endpoints as the only candidates for optimal parameters that achieve the maxima of $Q(s, 1, \phi)$. Evaluating $Q(s, 1, 0)$ then provides us a bound on

$$P_{1+s}(X|\Theta) \leq Q(s,1,0) = \frac{1}{2^{1+s}}(2^s + 1), \tag{3.18}$$

and plugging this back into (3.5) gives (3.9). □

3.2.2 A relation for n-qubits

To extend the one-qubit uncertainty relation for multiple qubits, the central problem is to prove that the lower bound on conditional entropy scales linearly with block length n. Essentially this means we need to prove that for a system of n qubits, entanglement across qubits *do not* give rise to a lower minimal α-Rényi entropy for BB84 measurements.

We show that this is true by examining the last qubit measured, conditioned on all previous $n-1$ measurement history of basis and outcomes. Given a normalized density operator ρ_{ABk} for n qubits, denote B as the last qubit, while A represents the remaining $n-1$ qubits. Then

$$P(X_B|\Theta)_{\rho_{ABk}} = \frac{1}{2}\sum_{\theta_B, x_B \in \{0,1\}} p^{\alpha}_{x_B|\theta_B x_A \theta_A k}, \tag{3.19}$$

where $p_{x_B|\theta_B x_A \theta_A k} = \mathrm{tr}(M_{x_B|\theta_B}\ \sigma_B)$,

$$\sigma_B = \mathrm{tr}_A \left[\frac{M_{x_A|\theta_A}\ \rho_{ABk}\ M^\dagger_{x_A|\theta_A}}{\mathrm{tr}(M_{x_A|\theta_A}\ \rho_{ABk}\ M^\dagger_{x_A|\theta_A})} \right] \tag{3.20}$$

is the corresponding normalized density operator. Since the uncertainty relation for 1 qubit holds for any density operator, it holds in particular for σ_B. By induction, it is then easily shown that the minimal entropy is additive.

Having proven the uncertainty relation for general α-Rényi entropies for a single qubit state, we now prove that for any multiple-qubit state ρ_k, under i.i.d. BB84 measurements, the minimal output α-Rényi entropy is additive.

Lemma 3.2.3. *For any state ρ_k of n qubits, the minimal conditional α-Rényi entropy of X^n with respect to Θ^n is additive.*

Proof. Consider

$$\begin{aligned} P_\alpha(X^n|\Theta^n)_{\rho_k|\rho_k} &= \sum_{\theta^n \in \{0,1\}^n} p_{\theta^n} \sum_{x^n \in \{0,1\}^n} p^\alpha_{x^n|\theta^n, K=k} \\ &= \frac{1}{2^n} \sum_{\theta^n \in \{0,1\}^n} \sum_{x^n \in \{0,1\}^n} \left(\prod_{i=1}^n p_{i|x^{i-1},\theta^{i-1}} \right)^\alpha \end{aligned} \tag{3.21}$$

where $p_{i|x^{i-1},\theta^{i-1}} = p_{x_i|X^{i-1}=x^{i-1},\Theta^{i-1}=\theta^{i-1},K=k}$ for $i \geq 2$ and $p_1 = p_{x_1|\theta_1,K=k}$. Then

$$\begin{aligned} P_\alpha(X^n|\Theta^n)_{\rho_k|\rho_k} &= \frac{1}{2^{n-1}} \sum_{\theta^n \in \{0,1\}^n} \sum_{x^n \in \{0,1\}^n} \left(\prod_{i=1}^{n-1} p_{i|x^{i-1},\theta^{i-1}} \right)^\alpha \cdot \frac{1}{2} p^\alpha_{n|x^{n-1},\theta^{n-1}} \\ &\leq c \cdot \frac{1}{2^{n-1}} \sum_{\Theta^{n-1}, X^{n-1} \in \{0,1\}^{n-1}} \left(\Pi_{i=1}^{n-1} p_i \right)^\alpha \\ &\leq c^n. \end{aligned} \tag{3.22}$$

Hence we have

$$\mathrm{H}_\alpha(X^n|\Theta^n)_{\rho_k|\rho_k} \geq \frac{1}{1-\alpha} n \cdot \log c. \tag{3.23}$$

□

Combining this with the one-qubit uncertainty relation derived in Section 1, we obtain the following corollary:

Corollary 3.2.4. *Given any state ρ_k of n-qubits, for i.i.d. BB84 measurements,*

$$\mathrm{H}_\alpha(X^n|\Theta^n)_{\rho_k|\rho_k} \geq \frac{1}{\alpha - 1} \left[\alpha - \log(1 + 2^{\alpha-1}) \right] \cdot n. \tag{3.24}$$

3.2.3 Further conditioning on classical side information K

We have obtained an uncertainty relation $\mathrm{H}_\alpha(X^n|\Theta^n)_{\rho_k|\rho_k}$ in Corollary 3.2.4 for a particular ρ_k, where conditioning is on the basis string Θ^n. Generally an adversary has the state ρ_k labelled with classical information K, hence we need to relate this quantity to $\mathrm{H}_\alpha(X^n|\Theta^n, K)_\rho$, for the state $\rho = \sum_k p_k \rho_k$, such that the α-Rényi entropy is also conditioned on the classical information K. This quantity is evaluated as

$$\begin{aligned}\mathrm{H}_\alpha(X^n|\Theta^n K)_{\rho|\rho} &= \frac{1}{1-\alpha}\log\sum_k \sum_{\theta^n\in\{0,1\}^n} p_{k,\theta^n} \sum_{x^n\in\{0,1\}^n} p^\alpha_{x^n|\theta^n,k} \\ &= \frac{1}{1-\alpha}\log\sum_k p_k \sum_{\theta^n\in\{0,1\}^n} p_{\theta^n|k} \sum_{x^n\in\{0,1\}^n} p^\alpha_{x^n|\theta^n,k}, \end{aligned} \tag{3.25}$$

where the subtle difference is that $p(\Theta|k)$ is conditioned on classical information $K = k$. In general this quantity is **not** equal to the uniform distribution, meaning that $p(\Theta^n|K = k) \neq p(\Theta^n) = \frac{1}{2^n}$, which never decreases the value of the above summation, therefore implying that conditioning does not increase entropy.

However, note that in our case Θ^n is chosen randomly, regardless of the state the adversary prepares. In other words, no matter what state ρ_k is prepared, it remains impossible for the adversary to gain additional information about Θ^n at all. Then it is clear that $p(\Theta^n|k) = p(\Theta^n) = \frac{1}{2^n}$, and further conditioning on K yields the same lower bound

$$\begin{aligned}\mathrm{H}_\alpha(X^n|\Theta^n, K)_{\rho|\rho} &= \frac{1}{1-\alpha}\log\sum_k p_k \sum_{\theta^n\in\{0,1\}^n} p_{\theta^n} \sum_{x^n\in\{0,1\}^n} p^\alpha_{x^n|\theta^n,k} \\ &\geq \frac{1}{\alpha-1}\,[\alpha - \log(1+2^{\alpha-1})]n. \end{aligned} \tag{3.26}$$

3.2.4 Relation to the min-entropy

After obtaining $\mathrm{H}_\alpha(X^n|\Theta^n)_{\rho|\rho}$, we finally link this quantity to the final desired measure: $\mathrm{H}^\varepsilon_{\min}(X^n|\Theta^n)_{\rho|\rho}$. This is provided by applying some known results in about bounding the smooth min-entropy with this class of conditional Rényi entropies.

Theorem 3.2.5 (Lower bound on smooth min-entropy, [36])**.** *Given any quantum state $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB}), \sigma_B \in \mathcal{S}(\mathcal{H}_{AB})$, $\epsilon \geq 0$ and $\alpha \in (1, 2]$,*

$$\mathrm{H}^\varepsilon_{\min}(A|B)_\rho \geq \mathrm{H}^\varepsilon_{\min}(A|B)_{\rho|\sigma} \geq \mathrm{H}_\alpha(A|B)_{\rho|\sigma} - \frac{1}{\alpha-1}\log\frac{2}{\epsilon^2}. \tag{3.27}$$

This implies that the smooth min-entropy is lower bounded by a general α-Rényi entropy, with a correction term growing logarithmically with ϵ^2. For the Shannon entropy ($\alpha = 0$) this term diverges, hence the bound is not useful. However, considering the vicinity of $\alpha \in (1, 2]$, a good convergence can be obtained.

Lastly, the min-entropy of X^n can then be bounded by using [36]

$$\frac{1}{n}\,\mathrm{H}^{\varepsilon}_{\min}(X^n|\Theta^n K)_\rho \geq \frac{1}{n}\mathrm{H}^{\varepsilon}_{\min}(X^n|\Theta^n K)_{\rho|\rho} \geq \max_{s\in(0,1]} \frac{1}{s}[1 + s - \log(1 + 2^s)] - \frac{1}{sn}\log\frac{2}{\epsilon^2} = c'. \tag{3.28}$$

The maximum value of (3.28) is obtained for different values of s as n and ϵ varies. For finite size cryptography, such a bound might be useful in the following manner: by fixing a small value of error ϵ, maximizing the RHS expression for (3.28) given a desirable value of n will provide the maximum min-entropy rate achievable with this result.

3.3 Advantages in finite-size cryptography

At the first glance, it may be hard to see that this new result (3.28) provides a better bound on the smooth min-entropy than previously derived relations as stated in (3.2). Indeed, in the limit of large n, both error parameters vanish exponentially. By fixing the last correction term in (3.28) to be $\frac{1}{sn}\log\frac{2}{\epsilon^2} = \delta$, we obtain that $\epsilon = 2^{\frac{-\delta sn+1}{2}}$. The relation of (3.2) is therefore still preferable, since $\mathrm{H}(X^n|\Theta^n K)_\rho > H_\alpha(X^n|\Theta^n K)_\rho$ for $\alpha > 1$, it will yield a larger amount of smooth min-entropy for an exponentially vanishing error.

Yet, when it comes to applying into a practical experiment with finite block lengths, (3.2) has a small caveat: whereas ε decreases exponentially in the number of qubits n, for a large amount of uncertainty i.e. $c' = 1/2 - \delta \approx 1/2$ the convergence is extremely slow. For example, in the implementation of secure oblivious transfer protocol for $\delta = 0.0106$ [33] corresponding to $c' = 0.4788$ in (3.2), we need $n \geq 2 \times 10^8$ to even have $\varepsilon = 0.1$!

In an experiment using weak coherent pulses, with frequency of 1GHz and Poisson parameter $\mu = 1$ it takes approximately 2.5 seconds to generate such an amount of n [33] photons, if there are absolutely no losses of any kind. However, compared to the generation time, a more significant inconvenience is that the classical post-processing of such large block lengths is extremely time-consuming. To implement such protocols, it would thus be desirable to have a relation that is useful for significantly smaller values of n.

Applying our new uncertainty relation to the same example, by plugging in $s = 0.1$ clearly demonstrates that for the same $\varepsilon = 0.1$, $c_{BB84} \geq 0.4837$ with $n = 1 \times 10^4$.

Comparing this with calculations in the previous section, the required block length n is approximately 10^{-4} times smaller, implying that this result guarantees the same amount of smooth min-entropy with a much smaller block length. Figure 3.1 provides a comparison of these two bounds.

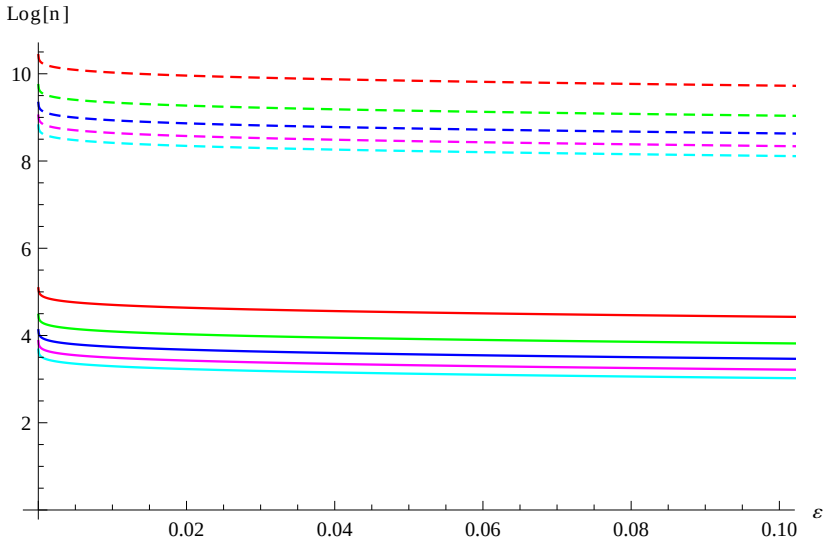


Figure 3.1: This plot shows the minimal required block length n on a logarithmic scale of base 10, in order to achieve error parameter ϵ. The dashed curves are plotted for the bounds stated in Section 3.1 while the solid lines are obtained from our new analysis in Section 3.2. The different colors represent the fixed values of the lower bound c', with values 0.45, 0.46, 0.47, 0.48, and 0.49 respectively. The plotted bounds are relatively higher on this logarithmic scale as c' increases.

This newly developed relation has far fetching implications on the practical feasibility of quantum-mechanical cryptographic protocols. It can readily be applied to any BB84 based two-party protocols in the bounded(or noisy)-storage model, and enables experiments for significantly smaller values of n. In the case study described in Chapter 4 and 5, it enables the experimental implementation [29] by offering security for $n = 2.5 \times 10^5$ instead of $n > 10^9$ for the same error parameter ε. This is particularly useful, since it enables the computation of syndrome for a relatively much smaller error-correcting code, which we are able to generate, store and use efficiently.

Lastly, it is worth noting that our approach to derive this uncertainty relation can be extended to the case of six-state measurements, i.e. measurements in σ_x, σ_z and σ_y eigenbases, which are useful for other classes of protocols. The derivations are shown in

the appendix for completeness. For this case we obtain

$$\mathrm{H}_{\min}^{\varepsilon}(X|\Theta K) \geq c_6 \text{ with} \tag{3.29}$$

$$c_6 := \max_{s\in(0,1]} \frac{-1}{s} \log \left[\frac{1}{3}\left(1+2^{1-s}\right)\right] - \frac{1}{sn}\log\frac{2}{\epsilon^2}. \tag{3.30}$$

This yields a similar improvement over the relation analogous to (3.2) proven in [9].

Chapter 4

The security of Commitments

In this chapter, we apply our results for smooth min-entropy uncertainty relations in Chapter 3 to study the commitment protocol within the Noisy Storage Model. We show conditions under which information theoretic security holds, paving the way towards a bit commitment protocol robust against experimental errors.

Sections 4.1 and 4.2 define two cryptographic protocols: the Commitment Scheme and Weak String Erasure. Their functionalities and security conditions are explicitly stated. Section 4.3 shows how security can be achieved for these protocols when considering an ideal experimental setup.

In Section 4.4 we explain how experimental errors affect the security analysis of these protocols. By showing the robustness of security under such errors, we prove that these protocols can indeed be implemented with practical devices.

4.1 Commitment Scheme

Assuming two parties Alice and Bob. Alice has a piece of information C, which she wants to commit to Bob. This scheme consists of two phases: in the *commit phase*, Alice provides Bob with some form of evidence, that she has chosen a particular C. Later on in the *open phase* Alice reveals C to Bob.

A straightforward way to think about this is to imagine that at $t = 0$, Alice sends a locked safe box to Bob containing a slip of paper, with C written on it. When $t = t_1$, she hands Bob the key to the safe. A secure commitment should have the following properties:

- Binding: After the commit phase, Alice is not allowed to change C. In other words, she cannot convince Bob to accept an opening of any $\hat{C} \neq C$.

- Hiding: Before the open phase, Bob is not allowed to know C.

Let us consider an example where this protocol is useful. Assume Alice and Bob are undergoing a divorce, and wants to decide who gets the family car. To do so, they agree to play a fair game: coin toss. Alice flips a coin, while Bob guesses the outcome. If Bob predicts correctly, he wins the car; otherwise Alice wins the car. This can be easily achieved if the two parties are face-to-face, but let us suppose that Alice and Bob do not want to meet in person. By the following steps, the commitment protocol can then be applied to execute this coin-toss game fairly:

Step 1. Alice flips a fair coin. If the outcome is heads, $C = 0$; and if the outcome is tails, $C = 1$. Alice then commits $C \in \{0, 1\}$ to Bob.

Step 2. Bob makes a random guess D, and sends his guess $D \in \{0, 1\}$ to Alice.

Step 3. After Alice receives Bob's guess, she opens C to B. Both parties compute $E = C \otimes D$. If $E = 0$, Bob wins; otherwise Alice wins.

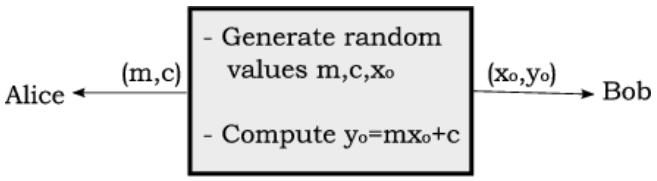
In general, one concludes that a commitment is useful to ensure that a piece of information is completely pre-determined, but shielded. Summarizing this, the formal definition of string commitment (the commitment of a binary string) is given as below:

Definition 4.1.1. *A string commitment protocol consists of two parties Alice and Bob. At $t = t_0$, Alice commits $C \in \{0, 1\}^l$ to Bob. At $t = t_1$, Alice opens the commitment and both parties output C.*

- *Correctness of the protocol: Whenever both parties are honest, Bob always accepts C as an honest commitment from Alice.*
- *Security against Alice (binding): If Bob is honest, whenever Alice cheats by attempting to change C after the commit phase, Bob will always detect the cheating and reject C as a valid commitment.*
- *Security against Bob (hiding): If Alice is honest, a dishonest Bob cannot learn C before the open phase. In other words, the probability distribution of C conditioned on Bob's information is uniform over the set $\{0, 1\}^l$.*

This protocol is the core of many seemingly impossible cryptographic tasks, however it is not easy at all to execute. For example, if Alice were to commit C by sending it in a locked safe, Bob could easily cheat by breaking the safe without Alice's knowledge. In classical cryptography, it has been proven that no matter what scheme is used, the protocol cannot be perfectly secure against at least one of the cheating parties. All practical commitments today are only computationally secure. However, this implies that

the commitment cannot last for longer times, since given enough computational resources and time, eventually one party will be able to break the protocol.

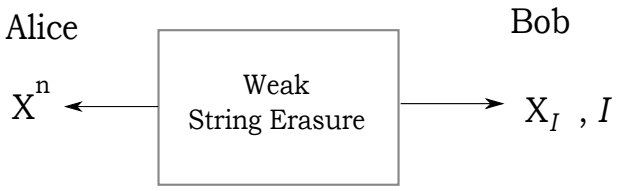


Theoretically, the key to secure execution of commitments arises if *correlated randomness* distributed between Alice and Bob as a physical resource can be obtained. To see why, let us consider a perfect, impenetrable simulator as in [32]: to Alice it outputs two random values (m,c), specifying the function of a straight line $y(x) := mx + c$. To Bob, it outputs a (x_0, y_0) such that the point (x_0, y_0) lies on $y(x)$. Alice commits to y_1 by sending to Bob x_1, such that $f(x_1) = y_1$. At the open phase, Alice sends m and c to Bob, who checks if (x_0, y_0) indeed lies on $f(x)$, and if it does, he accepts the commitment.

Note that Bob cannot learn y_1 beforehand since he does not know the values m, c which specify the function. On the other hand, any attempt of Alice to cheat implies that she has to provide Bob with m', c', where Bob will easily detect that (x_0, y_0) does not lie on $y(x) = m'x + c'$. Hence, it is clear that distributed randomness amongst both parties enables commitment schemes to work.

4.2 Weak String Erasure

The functionality of Weak String Erasure (WSE) is as follows: it provides to Alice a random binary (classical) string X^n of length n, and Bob a randomly selected substring $X_{\mathcal{I}}$, together with a set $\mathcal{I}$, indicating locations of elements in the substring Bob possesses.



A formal definition of the protocol and its security conditions can be found in [19, Section 3.1]. We state a simpler version here for convenience:

Definition 4.2.1. *An (n, λ, ϵ)- Weak String Erasure (WSE) scheme is a protocol between two parties Alice and Bob, satisfying the following conditions:*

- *Correctness: Whenever both parties are honest, Alice obtains a random binary string X^n of length n, while Bob receives a random substring of X^n, with the corresponding set of location indices $\mathcal{I}$. The state created for Alice and Bob is*

$$\rho_{AB} = \tau_{\{0,1\}^n,A} \otimes \tau_{2^{[n]},B} \tag{4.1}$$

where τ_H denotes the uniform distribution over the range of possible values H.

- *Security against Alice: If Bob is honest, Alice does not learn $\mathcal{I}$ at all.*
- *Security against Bob: If Alice is honest, quantifying Bob's information about X^n by the smooth min-entropy, it is required that*

$$\frac{1}{n} \mathrm{H}_{\min}^{\varepsilon}(X^n|B) \geq \lambda > 0. \tag{4.2}$$

Essentially, security for Alice implies that Bob has limited amount of information about the string X^n, while security for Bob means that Alice does not know anything about the set of indices $\mathcal{I}$ Bob obtains from WSE.

It is easy to see that the WSE protocol distributes correlated randomness between Alice and Bob. The security of the WSE then becomes crucial: once established, one can make use of WSE as a subprotocol to implement the commitment scheme by simply adding some classical information post-processing.

4.3 Ideal Setting

In this section, we first consider an ideal setting of protocols for WSE and commitments, whereby security holds under the absence of any experimental losses or errors.

Protocol 1: Weak String Erasure

Outputs: $X^n \in \{0,1\}^n$ to Alice, $(\mathcal{I}, z_{|\mathcal{I}|}) \in 2^{[n]} \times \{0,1\}^{|\mathcal{I}|}$ to Bob.

1. **Alice:** Chooses string $X^n \in \{0,1\}^n$ and basis $\theta_n \in \{0,1\}^n$ uniformly at random. Encodes bit X_i in the basis θ_i (as $H^{\theta_i}|x_i\rangle$), and sends the state to Bob.
2. **Bob:** Chooses basis string $\tilde{\theta}_n \in \{0,1\}^n$ uniformly at random. Upon receiving the state from Alice, measures in basis $\tilde{\theta}_i$ to obtain outcome $\tilde{X}_i$.

Both parties wait time Δt.

3. **Alice:** Sends the basis information θ_n to Bob, and outputs X^n.
4. **Bob:** Computes $\mathcal{I} = \{i \in [1,n] \mid \theta_i = \tilde{\theta}_i\}$, and outputs $(\mathcal{I}, z_{|\mathcal{I}|}) := (\mathcal{I}, \tilde{X}_{\mathcal{I}})$.

WSE can be achieved by using any QKD device: an honest Alice has a source that emits EPR pairs. She measures one of the qubits in a randomly chosen BB84 basis, records the

measurement outcome in a string X^n, and sends the other half to Bob. If Bob is honest, he measures the state he receives in a random BB84 basis. After some time Δt, Alice sends her basis to Bob, while Bob picks out the locations where both bases match, and obtains a substring and a corresponding set of location indices.

This protocol satisfies correctness and security against Alice trivially. However, security against Bob relies on the physical assumption of bounded/noisy quantum storage. The goal of a dishonest Bob is to gain full information about Alice's string X^n. Upon receiving ρ_B from Alice, since the basis information is not readily available, and since BB84 bases are non-orthogonal, Bob cannot identify the quantum state with full certainty.

We consider a general strategy of attack dishonest Bob can make. First of all, he is free to perform any type of instantaneous unitary operations on the system, and perform any measurement upon the qubits, obtaining some classical information. For example, instead of BB84 measurements, he can perform an optimal set of POVMs such that the probability of guessing is maximized. Lastly, he can store a limited amount of quantum information. The waiting time Δt correspond to the amount of time Bob's quantum memory is subjected to noise, before Alice announces her basis.

4.3.1 Conditioning on classical information

In Chapter 3 we developed uncertainty relations with respect to i.i.d. BB84 measurements. However, these relations cannot be applied to Bob, since he **does not** necessarily measure in BB84 bases. To make proper use of them, note that the WSE scenario is equivalent to the following description: Bob performs a measurement on the state ρ_B, and obtains some classical information $K = k$ with probability p_k. This effectively prepares a state ρ_k on Alice's side, while an honest Alice measures ρ_k in i.i.d. BB84 measurements, obtaining X^n. Bob's information about X^n is then quantified by $\mathrm{H}^{\varepsilon}_{\min}(X^n|\Theta^n K)_\rho$, as we have derived previously in (3.28).

Theorem 4.3.1 (Smooth min-entropy conditioned on classical information)**.** *Consider a set of n binary variables $X_1, X_2, ..., X_n$ (not necessarily independent), which are measurement outcomes obtained by an honest Alice performing i.i.d. BB84 measurements on her state of n qubits. Given any smoothing parameter ϵ, the smooth min entropy satisfies*

$$H^{\epsilon}_{min}(X_1, ...X_n|\Theta, \mathcal{K}) \max_{s\in(0,1]} \geq g(s)n + \frac{2\log\epsilon - 1}{s}, \tag{4.3}$$

maximizing over the range $s \in (0, 1]$ where

$$g(s) = \frac{-1}{s}\left[\log(1+2^s) - (1+s)\right]. \tag{4.4}$$

4.3.2 Conditioning on quantum information

Theorem 4.3.1 gives us a bound on the min-entropy rate of Bob's classical information of X^n, whenever Alice is honest. We then use Theorem 2.2.3 to bound the min-entropy rate of X^n when he has quantum information about X^n, by evaluating the smooth conditional min-entropy for the joint ccq-state $\rho_{X^n K \Theta \mathcal{F}(\mathcal{Q})}$.

Lemma 4.3.2 (Smooth min-entropy conditioned on quantum information)**.** *For any attack of dishonest Bob, let his storage be given by $\mathcal{F} : \mathcal{B}(\mathcal{H}_{in}) \to \mathcal{B}(\mathcal{H}_{out})$, where $\mathcal{F} = \mathcal{N}^{\otimes \nu n}$, $\nu > 0$ and $\mathcal{N}$ being a quantum channel that satisfies the strong converse relation, with classical capacity $C_{\mathcal{N}}$. Let $\mathcal{Q}$ be the quantum state sent by an honest Alice. Then Bob's smooth min-entropy about Alice's string X^n is lower bounded by*

$$\mathrm{H}^{\varepsilon}_{\min}(X^n|\Theta^n K \mathcal{F}(\mathcal{Q})) > -\log P^{\mathcal{F}}_{succ}(\hat{R}n) = \lambda n, \tag{4.5}$$

where

$$\begin{aligned}\hat{R} &= \max_{s\in(0,1]} \; g(s) - \frac{2\log\epsilon - 3}{sn} - \frac{1}{n}\log\frac{2}{\epsilon} \\ &= \max_{s\in(0,1]} \; g(s) - \frac{(2+s)\log\frac{1}{\epsilon} + 3 + s}{sn}\end{aligned}$$

$$P^{\mathcal{F}}_{succ}(\hat{R}n) = 2^{-(\hat{R} - \nu\cdot C_{\mathcal{N}})n}. \tag{4.6}$$

In the large n limit, $\hat{R}$ approaches $\frac{1}{2}$, hence $P^{\mathcal{F}}_{succ}(\hat{R}n)$ will be exponentially small in n once $\nu \cdot C_{\mathcal{N}} < \frac{1}{2}$, meaning that the security condition against Bob for WSE is satisfied.

Note that this bound is dependent on the quantum storage device used by Bob. For example, if the memory were to be noiseless and of arbitrarily large size, P_{guess} is unity and the min-entropy of Bob's information about X^n would be simply zero. However, the assumption of noisy and bounded storage comes in here to give a sufficiently high min-entropy which is crucial for the security proof. For simplicity in further proofs, we also introduce a simpler version, considering only bounded storage:

Corollary 4.3.3 (Smooth min-entropy for bounded quantum storage)**.** *Assuming the smooth min-entropy of Bob's information of X^n be $H^{\epsilon}_{\infty}(X^n|T)$, and Bob has a quantum memory $\mathcal{Q}$ that has a storage size of νn qubits. Then*

$$H^{\epsilon}_{min}(X^n|T\mathcal{Q}) \geq H^{\epsilon}_{min}(X^n|T) - \nu n. \tag{4.7}$$

Having established security for WSE, we introduce the protocol for commitments.

Protocol 2: Non-Randomized String Commitment

By using an binary linear error-correcting code C, let Syn:$\{0,1\}^n \to \{0,1\}^{n-k}$ be the function that outputs the parity check syndrome for C. Also, let Ext : $\{0,1\}^n \times \mathcal{R} \to \{0,1\}^l$ be a 2-universal hash function.

(A) **Commit Phase**

1. **Alice and Bob :** execute WSE. Alice obtains X^n while Bob obtains $X_{\mathcal{I}}$, $\mathcal{I}$.
2. **Alice:**
   a) computes $w = \text{Syn}(X^n)$ and sends it to Bob.
   b) picks a 2-universal hash function $r \in_R \mathcal{R}$ and sends it to Bob.
3. **Alice:** Commits to $D \in \{0,1\}^l$ by computing $C^l = \text{Ext}(X^n, r)$ and sends $E = C \oplus D$ to Bob. Note that since C is unknown to Bob, D is also unknown.

(B) **Open Phase**

1. **Alice:** reveals the complete string X^n to Bob.
2. **Bob:** checks commitment by:
   a) computing the syndrome and check that it agrees with w sent by Alice.
   b) checking X^n against $X_{\mathcal{I}}$, ensuring that for all $i \in \mathcal{I}$, $X_{\mathcal{I}_i} = X_i$.
3. **Bob:** If conditions are satisfied, he accepts commitment and calculates $C^l = \text{Ext}(X^n)$. Both of them output D.

Correctness of the protocol is easy to verify, since the syndrome Bob computes with X^n will match the syndrome provided by an honest Alice in the commit phase. Also, for $\forall i \in \mathcal{I}$, $X_{\mathcal{I}_i} = X_i$ will hold true in the absence of errors. Once Bob obtains the full string of X^n in the open phase, he can immediately compute C^l to verify Alice's commitment.

4.3.3 Security against Alice

Whenever Alice attempts to cheat by changing C, we require that an honest Bob will detect the cheating. This is achieved by the backchecking procedures done by Bob. We establish the security against Alice in the following lemma:

Lemma 4.3.4 (Security against Alice, [19])**.** *If Bob is honest, and the distance satisfies $d > 2\log\frac{1}{\epsilon}$ for some $\epsilon > 0$, then the pair of protocols (Commit, Open) are ϵ-binding.*

Proof. For any attempt of Alice to cheat in such a protocol, she is able to choose to open a string $\tilde{X}^n$ to Bob such that $\tilde{X}^n \neq X^n$ and $\text{Syn}(\tilde{X}^n) = \text{Syn}(X^n)$. Hence, $\Delta(\tilde{X}^n, X^n) \geq d$. Let $\hat{X}^n$ be the string Alice picks to execute WSE with Bob. Then $\Delta(\tilde{X}^n, \hat{X}^n) \geq \frac{d}{2}$,

meaning that $\tilde{X}^n$ and $\hat{X}^n$ must differ at least $\frac{d}{2}$ locations, where d is the minimum distance of the error correcting code used.

However, since an honest Bob obtains each bit X_i with probability $\frac{1}{2}$, the probability of him not finding at least one bit in $X_{\mathcal{I}}$ which disagrees with $\tilde{X}^n$ is

$$\text{Prb}(X_{\mathcal{I}_i} = \tilde{X}_i,\ \forall i \in \mathcal{I}) = 2^{-\frac{d}{2}} \leq \epsilon. \tag{4.8}$$

□

4.3.4 Security against Bob

When the syndrome $\text{Syn}(X^n)$ is provided to Bob, the conditional min-entropy of Bob's information about X^n decreases, and can be bounded by the chain rule in Theorem 2.2.2. Subsequently, by using privacy amplification, we show that the cq-state of C^l and Bob's information is ϵ'-close to a product state, with C^l having uniform distibution over $\{0,1\}^l$.

Lemma 4.3.5 (Security against Bob)**.** *If Alice is honest, and the code rate satisfies* $R > 1 - \lambda + \frac{1+2\log\frac{1}{\epsilon'}}{n}$*, where* $\lambda = H_{min}^{\epsilon'/4}(X^n|\mathcal{B}')$ *is the* $\frac{\epsilon'}{4}$*-smooth min-entropy rate of Bob's information* [1]*, then the pair of protocols (Commit, Open) are* ϵ'*-hiding.*

Proof. From Lemma 4.3.2, weak string erasure between Alice and Bob produces smooth min-entropy

$$\begin{aligned} H_{min}^{\epsilon'/4}(X^n|\mathcal{B}') &\geq \max_{s\in(0,1]} -\log P_{guess}^{\mathcal{F}} \left[\left(g(s) - \frac{(2+s)\log\frac{1}{\epsilon'} + 7 + 3s}{sn}\right) n\right] \\ &= \lambda n. \end{aligned} \tag{4.9}$$

where we have denoted the smooth min-entropy rate for Bob's quantum information about X^n as λ. After obtaining the syndrome, the smooth min-entropy of Bob's total information about X^n is

$$\begin{aligned} H_{min}^{\epsilon'/4}(X^n|\mathcal{B}', \text{Syn}(X^n)) &\geq H_{min}^{\epsilon'/4}(X^n, \text{Syn}(X^n)|\mathcal{B}') - L \\ &\geq (\lambda - 1 + R)n. \end{aligned} \tag{4.10}$$

where L is the length of the syndrome, given by $L = n - k = (1 - R)n$, R being the code rate of the specified error-correcting code.

Next, we show that by using privacy amplification, we can distill the committed information C^l from X^n such that Bob is ignorant about C^l. Denoting the committed string

[1] The information holds by a dishonest Bob denoted by $\mathcal{B}'$ consists of: the basis information Θ^n provided by Alice, the classical information K obtained by Bob's measurement, and the quantum information after undergoing Bob's quantum memory $\mathcal{F}(\mathcal{Q})$.

as $C^l = Ext(X^n, r) \in \{0,1\}^l$ having length l,

$$\rho_{C^l,\mathcal{B}',\mathrm{Syn}(X^n)} \approx_{\epsilon'} \tau_{\{0,1\}^l} \otimes \rho_{\mathcal{B}',\mathrm{Syn}(X^n)} \tag{4.11}$$

where

$$\epsilon' = 2 \cdot \frac{\epsilon'}{4} + 2^{-\frac{1}{2}[H_{min}^{\epsilon_o}(X^n|\mathcal{B}',\mathrm{Syn}(X^n))-l]-1} \tag{4.12}$$

and $\tau_{\mathcal{A}}$ is the uniform distribution over the entire set $\mathcal{A}$. Setting the second term to be smaller than $\frac{\epsilon'}{2}$ gives

$$\begin{aligned} \frac{-1}{2}[(\lambda - 1 + R)n - l] &\leq \log \epsilon' \\ \lambda - 1 + R &\geq \frac{l + 2\log\frac{1}{\epsilon'}}{n}. \end{aligned} \tag{4.13}$$

For l at least 1, this implies we require $R > 1 - \lambda$ in the large n limit. □

It is worth noting that conventional usage of error-correcting codes often rely on the transmission and correction of codewords. However, here we focus on using the parity check syndrome as a back-checking tool. Honest parties are even not required to decode: they only need to perform a straightforward computation of the syndrome. We summarize some crucial differences from conventional error-correcting code research and the usage for this commitment protocol as below:

Properties	Conventional error-correcting	Usage in commitment protocol
Transmitted message	A codeword	Any selected randomly binary string
Syndrome	Zero	Non-zero
Detecting errors	Check if syndrome is zero	Check if syndrome fits original string
Decoding	A code which is easy to decode is always of better interest	There is no need to decode
Code size	Explicit constructions and analysis usually involve block lengths ranging from 10 to 10^5	Codes of sizes at least 2.5×10^5 are required

In an ideal setting, the conditions on the code are easily achieved. By setting small error parameters ϵ and ϵ', a typical concatenated code can be used, by using a Reed-Solomon error-correcting code defined over the field $GF(q)$ where q is of the form $q = 2^m - 1$ for some integer m. Each element of the code is then directly mapped to m binary bits, using a trivial code. Security for the protocol is then readily obtained.

4.4 How errors affect security

So far, performing commitments via WSE seems highly feasible. However, experimental errors can cause security to fail. While honest parties use standard devices subjected to noise, dishonest parties can be all powerful: being able to avoid *all* losses and errors. By tricking an honest party to believe that he is subjected to noise, an adversary can demand more information, or justify his cheating as an effect of noise. Also, an adversary gains complete knowledge about an honest party's setup. Such additional freedom makes it particularly challenging to prove the robustness of security under such scenarios.

In this section, we look at how these imperfections affect the security proof and sought to prove the robustness of this protocol against these errors.

4.4.1 Erasures

During transmission of qubits from Alice to Bob, significant amount of losses happen, especially for optical setups where photon detectors generally have efficiency around 40%. This by itself would not have affected the security. However, considering that Alice's photon source might not be a perfect single-photon source, this potentially compromises security. If for each qubit Alice detects, more than one photon was sent to Bob, he can perform both BB84 measurements, increasing his advantage of guessing the bit correctly.

Protocol 3: Weak String Erasure with Errors (WSEE)

Outputs: $X_n \in \{0,1\}^n$ to Alice, $(\mathcal{I}, z_{|\mathcal{I}|}) \in 2^{[m]} \times \{0,1\}^{|\mathcal{I}|}$ to Bob.

1. **Alice:** Chooses bit string $X_M \in_R \{0,1\}^M$ and basis string $\theta_M \in_R \{0,1\}^M$ uniformly at random. She encodes bit X_i in the basis given by θ_i (i.e., as $H^{\theta_i}|x_i\rangle$), and sends the resulting state to Bob.
2. **Bob:** Chooses basis string $\tilde{\theta}_M \in_R \{0,1\}^M$ uniformly at random. Measures in basis $\tilde{\theta}_i$ to obtain outcome $\tilde{X}_i$. If Bob obtains no click, he records round i as missing.

Both parties wait time Δt.

3. ***Both:*** *Reports the missing rounds to the other party.*
4. ***Alice:*** *If Bob reports q missing rounds, where q does not lie within $[(\epsilon-\beta)M, (\epsilon+\beta)M]$, ϵ being the erasure probability, Alice aborts protocol. Otherwise, she discards bits that Bob reported missing. Let $X_n \in \{0,1\}^n$ denote the remaining string, and θ_n the basis. Let $\tilde{X}_n$, and $\tilde{\theta}_n$ be the corresponding strings for Bob.*
5. **Alice:** Sends the basis information θ_n to Bob, and outputs x_n.
6. **Bob:** Computes $\mathcal{I} = \{i \in [m] \mid \theta_i = \tilde{\theta}_i\}$, and outputs $(\mathcal{I}, z_{|\mathcal{I}|}) := (\mathcal{I}, \tilde{X}_{\mathcal{I}})$.

In [37], WSE has been modified to Weak String Erasure with Errors (WSEE), where each party is allowed to report a fraction of missing rounds during the transmission. A dishonest Bob can then discard a certain fraction of rounds where he obtained single-photons, by reporting them as missing to Alice, while validating rounds where multi-photons were obtained. This further lowers Bob's smooth min-entropy about X^n. On the other hand, losses do not affect the security against Alice, since she remains ignorant of Bob's basis choice for all of the valid rounds. Therefore, we are allowed to condition all events based on the fact that Alice registers a valid round. Detailed methods to quantify the correction on smooth min-entropy are given in [37]. By simplifying the analysis, we narrow down the range of necessary parameters required and summarize them as below:

Probabilities	Description
p^1_{sent}	Probability that a single photon was sent to Bob.
$p^h_{B,noclick}$	Probability that honest Bob observes no click.
$p^d_{B,noclick}$	Probability that dishonest Bob observes no click.
p_{err}	Probability that the measurement outcome for honest Alice and honest Bob is different, when the same basis is used for both parties.

Table 4.1: Probabilistic quantities required for security proof of commitments. **All above parameters are conditioned on the event that Alice registered a single click**.

These quantities are dependent on various parameters of the setup, such as the mean photon number of photon source, efficiencies of detectors for Alice and Bob, and dark count rates for each detectors. A high p^1_{sent} indicates that whenever Alice detects a single photon, the more likely it is that a single photon was sent to Bob. This is crucially important since if Bob receives multiple photons, the probability of him guessing the bit value correctly increases significantly. On the other hand, a low value of $p^h_{B,noclick}$ indicates a high detection efficiency for Bob's apparatus, which restricts the additional amount of single-photons that cheating Bob reports missing. Lastly, $p^d_{B,noclick}$ is caused by the non-zero probability p^0_{sent} that no photons were sent to Bob at all. Essentially this models a dishonest Bob using a perfect photon detector, but tricking Alice into accepting a certain fraction of erasures.

We denote the number of signals sent as M (events where Alice registers a single click on her detector). The number of valid rounds which will be used subsequently for the commitment will be $n \approx (1 - p^h_{B,noclick})M$. With the parameters in Table 4.4.1, under a simpler assumption of bounded storage, the min-entropy can be evaluated. The following theorem is obtained by considering min-entropy only for an amount of rounds where Bob

obtains a single photon, then averaging over all the rounds that he considered valid.

Theorem 4.4.1. *Given parameters p^1_{sent}, $p^h_{B,noclick}$, and $p^d_{B,noclick}$, and denoting M as the number of signals sent. Assuming Bob has a storage of size N. Then except for an error ϵ, the lowest fraction of valid single-photon rounds left m^1_{left}, and the total fraction of valid rounds left m_{frac} is given by*

$$m^1_{left} = p^1_{sent} - p^h_{B,noclick} + p^d_{B,noclick} - 3\zeta$$
$$m_{frac} = 1 - p^h_{B,noclick} - 3\zeta,$$

where $\zeta = \sqrt{\frac{\ln \frac{2}{\epsilon}}{2M}}$. The smooth min-entropy of Bob's information is bounded by

$$\lambda = \frac{m^1_{left} \cdot L - \frac{N}{M}}{m_{frac}}, \tag{4.14}$$

where

$$L = \max_{s\in(0,1]} \frac{-1}{s}\left[\log(1+2^s) - 1 - s - \frac{3\log\frac{1}{\epsilon}}{m^1_{left}\cdot M}\right]. \tag{4.15}$$

4.4.2 Bit Flips

The main cause of bit flip errors are due to the polarization changes of the photons, while routed across the experimental setup through optical fibres. This is especially significant if small devices are used, since alignment of the lasers become a much harder task. In a typical setup, an error within the range 2-5% is expected.

The impact of bit flip errors on the security proof is even larger: If a finite fraction of errors were allowed to persist throughout the experimental data, Alice obtains more freedom to cheat, as a malicious Alice might be able to avoid errors and choose to corrupt the string X^n herself. In other words, the actual bit flip error in such a scenario equals zero, but Bob can be tricked into accepting an amount of $\approx p_{err}n$ wrong bits, where $p_{err} > 0$. It then becomes harder for Bob to identify cheating Alice, since whenever he finds a discrepancy between $X_{\mathcal{I}}$ and X_n, he cannot be sure if it was due to an error or a malicious Alice scheme. Note that however, security against Bob is not compromised since errors can only serve as a disadvantage to him, increasing his smooth min-entropy.

In this section, we derive new conditions on the error correcting code used such that the protocol can remain robust under a fraction of errors $p_{err} > 0$. Throughout this proof, we need to bound the occurrence probability of bad events. To do so, we make use of the Hoeffding inequality as stated below:

Theorem 4.4.2 (Hoeffding's inequality,[16])**.** *Given a set of identically and indepen-*

dently distributed random variables $\{X_j\}$, $1 \leq i \leq n$ *with a Bernoulli distribution, where* $Prb(X_i = 0) = 1 - p$ *and* $Prb(X_i = 1) = p$. *Define* $Y = \sum_{i=1}^{N} X_i$. *Then*

$$Prb[Y \leq (p - \alpha)N] = Prb\ [Y \geq (p + \alpha)N] = e^{-2\alpha^2 N}. \tag{4.16}$$

First of all, we introduce a modification to the commitment protocol in Section 4.3.2. For Bob to allow a certain number of errors, step 2b) of the protocol in the Open phase should be modified: instead of Bob rejecting the commitment whenever a disagreement betweeen $X_{\mathcal{I}}$ and X^n is obtained, he tolerates a certain number of disagreements. Also, an error-correcting code with different properties should be used. The fully modified protocol can be found as below:

Protocol 4: Modified Non-Randomized String Commitment

By using an binary linear error-correcting code C, let Syn:$\{0,1\}^n \rightarrow \{0,1\}^{n-k}$ be the function that outputs the parity check syndrome for C. Also, let Ext: $\{0,1\}^n \times R \rightarrow \{0,1\}^l$ be a 2-universal hash function.

(A) **Commit Phase**

1. **Alice and Bob :** execute WSEE. Alice obtains X_n while Bob obtains $X_{\mathcal{I}}$ and $\mathcal{I}$.
2. **Alice:**

   a) computes w=Syn(X_n) and sends it to Bob.

   b) picks a 2-universal hash function $r \in_R \mathcal{R}$ and sends it to Bob.
3. **Alice:** Commits $D \in \{0,1\}^l$ by computing $C^l = \text{Ext}(X^n, r)$ and sending $E = C \oplus D$ to Bob.

(B) **Open Phase**

1. **Alice:** reveals the complete string X_n to Bob.
2. **Bob:** checks commitment by:

   a) computing the syndrome and check that it agrees with w sent by Alice.

   b) checking X_n *against* $X_{\mathcal{I}}$, *ensuring that the number of disagreements lie between interval* $[(p_{err} - \alpha)m, (p_{err} + \alpha)m]$ *where* $m = |\mathcal{I}|$.
3. **Bob:** If conditions are satisfied, he accepts commitment and calculates $C^l = \text{Ext}(X_n)$. Both of them output D.

We shall first prove correctness of this modified protocol: whenever Alice and Bob are both honest, Bob always accepts the protocol except with some minimal probability ϵ.

Lemma 4.4.3 (Correctness of the protocol)**.** *If Alice and Bob are both honest, then for any* $\epsilon_1, \epsilon_2 > 0$, *there exists* α_1 *and* α_2 *such that the protocol is* ϵ*-correct, where* $\epsilon \leq \epsilon_1 + \epsilon_2$.

Proof. There are two steps in this proof. Firstly, we show that Bob receives at least m bits from WSEE except for probability ϵ_1. Secondly, we prove that the number of erroneous bits Bob picks up is close to the expected value p_{err}m, except for probability ϵ_2. The total probability of error for either events occuring is then $\epsilon = \epsilon_1 + \epsilon_2 - \epsilon_1\epsilon_2 \leq \epsilon_1 + \epsilon_2$.

Since each bit X_i from Alice is obtained by Bob with probability $\frac{1}{2}$, by applying the Hoeffding's inequality to the random variable $Y = |\mathcal{I}|$, i.e. the length of substring Bob obtains from WSEE, we see that

$$Prb\left[Y \leq \left(\frac{1}{2} - \alpha_1\right)n\right] \leq e^{-2\alpha_1^2 n} = \epsilon_1, \tag{4.17}$$

where n is the length of string X_n Alice has. Hence for any desired ϵ_1, we can set $\alpha_1 = \sqrt{\frac{\ln\frac{1}{\epsilon_1}}{2n}}$ so that Bob gets at least $m = (\frac{1}{2} - \alpha_1)n$ bits except for probability ϵ_1.

We proceed to show that the number of erroneous bits Bob obtains lie within the interval $[(p_{err} - \alpha_2)m, (p_{err} + \alpha_2)m]$ except for some probability ϵ_2. Note that previously we have stated that Bob obtains at least m bits except for a minimal probability, we can safely fix m in our second proof, by setting that if Bob gets additional bits, he always randomly truncates his substring to length m. Again by applying the Hoeffding's inequality, with the random variable of interest Z to be the number of erroneous bits Bob obtains,

$$Prb[|Z - p_{err}m| \geq \alpha_2 m] \leq 2e^{-2\alpha_2^2 m} = \epsilon_2. \tag{4.18}$$

holds when $\alpha_2 = \sqrt{\frac{\ln\frac{2}{\epsilon_2}}{2m}}$.

Hence correctness of the protocol is guaranteed except for probability $\epsilon \leq \epsilon_1 + \epsilon_2$. □

We now proceed to prove security against Alice. Recall that a malicious Alice can avoid bit flip errors and tamper with the bit string directly. To prove security, we need to show that no matter how Alice tampers with the string, Bob detects her cheating with probability close to 1. Previously where bit flip errors were not accounted for, whenever Bob checks $X_{\mathcal{I}}$ against X^n and finds one faulty bit, he aborts the protocol directly. However, in a realistic setup Bob accepts a number of roughly p_{err} m bits.

We know from the analysis of [19] that for any attack of Alice, she has to change at least $\frac{d}{2}$ bits such that Bob will accept the syndrome to be consistent. This can be seen through the fact that the syndrome fails to detect more than $\lfloor\frac{d-1}{2}\rfloor$ errors, as explained in Section 2.4. By this, we are able to set a constraint on the code distance used such that whenever Alice attempts to cheat, Bob picks up enough faulty bits to detect the cheating except for some minimal probability.

Lemma 4.4.4 (Security against Alice)**.** *If Bob is honest, given that Alice and Bob use an error-correcting code with minimum distance* $d > \frac{2(p_{err}+\alpha_2)(\frac{1}{2}-\alpha_1)n}{\frac{1}{2}-\alpha_3}$*, the pair of protocols(Commit, Open) is* $\epsilon_1 + \epsilon_3$*-binding.*

Proof. In our proof we assume that a malicious Alice can avoid all bit flip errors in WSEE, so that the scenario reduces to WSE where $p_{err} = 0$. From the proof of correctness, we know that Bob obtains enough bits from WSEE except with probability ϵ_1.

For any general attack Alice can attempt, to satisfy Bob's check on the syndrome she has to change at least $\frac{d}{2}$ bits in the original string X_n, where d is the code distance. Since Bob picks up each faulty bit with probability $\frac{1}{2}$, by defining W to be the number of faulty bits where Bob obtains in his substring, and applying Hoeffding's inequality,

$$Prb\left[W \leq \left(\frac{1}{2} - \alpha_3\right)\frac{d}{2}\right] \leq e^{-\alpha_3^2 d} = \epsilon_3. \tag{4.19}$$

Bob obtains at least $(\frac{1}{2} - \alpha_3)\frac{d}{2}$ flipped bits except with probability ϵ_3, for $\alpha_3 = \sqrt{\frac{\ln\frac{1}{\epsilon_3}}{d}}$. Combining with the fact that he accepts at most $(p_{err} + \alpha_2)m$ faulty bits, we require

$$(\frac{1}{2} - \alpha_3)\ \frac{d}{2} > (p_{err} + \alpha_2)m. \tag{4.20}$$

The requirement for code distance is then given by

$$d > \frac{2(p_{err} + \alpha_2)(\frac{1}{2} - \alpha_1)n}{\frac{1}{2} - \alpha_3}, \tag{4.21}$$

where in the large n limit, $\delta = \frac{d}{n} > 2\ p_{err}$ is required. Hence generally when Alice and Bob use a code with minimum distance that satisfies the above requirement, whenever Alice attempts to cheat, the pair of protocols is proven to be $\epsilon_1 + \epsilon_3$-binding. □

With this we end the security proof against Alice. The security proof for Bob is same as established in Lemma 4.3.5. We end the analysis by summarizing conditions on the error-correcting code, for a secure commitment:

Theorem 4.4.5 (Conditions for successful execution of the string commitment protocol)**.** *For fixed parameters of n and* ϵ*, if the error correcting code used satisfies:*

Relative minimum distance: $\delta > \frac{2(p_{err}+\alpha_2)(\frac{1}{2}-\alpha_1)}{\frac{1}{2}-\alpha_3}$.

Code rate: $R > 1 - \lambda + \frac{l+2\log\frac{1}{\epsilon}}{n}$.

where

$$\begin{aligned} \alpha_1 &= \sqrt{\frac{\ln\frac{2}{\epsilon}}{2n}}, \quad \alpha_2 = \sqrt{\frac{\ln\frac{4}{\epsilon}}{n}}, \quad \alpha_3 = \sqrt{\frac{\ln\frac{2}{\epsilon}}{d}}, \\ \lambda &= \frac{H_\infty^{\epsilon/4}(X^n|\mathcal{B}')}{n} \end{aligned} \tag{4.22}$$

then the modified Commitment protocol in Section 4.4.2 is ϵ-correct, ϵ-binding and ϵ-hiding for a string commitment of length l.

Proof. Using Lemma4.4.3 for correctness, while using Lemma 4.4.4 for security against Alice, set $\epsilon_1 = \epsilon_2 = \epsilon_3 = \frac{\epsilon}{2}$. On the other hand, use Lemma 4.3.5 for security against Bob and set $\epsilon' = \epsilon$ completes the proof.

□

4.4.3 Error-correcting codes for security

Given these new conditions on the error correcting code, we need to ask the question: for what amount of bit flip errors p_{err} can error-correcting codes satisfying conditions on the relative minimum distance δ and code rate R simultaneously as stated in Theorem 4.4.5 be obtained? After a survey of different codes, we present three different classes of codes that are able to achieve security for different ranges of p_{err}. We discuss their advantages and disadvantages.

Random codes

This class of codes are generated by specifying a binary matrix P with dimensions $(1-R)n \times Rn$, such that the parity check matrix H is defined as $H = [P|\mathbb{I}_{(1-R)n}]$. Each element in P is chosen randomly. By Theorem 2.4.3, for large block lengths, we can see that this bound approaches a step function where for rates $R < C_\delta$, minimum distance is expected to be larger than δn except with extremely small probability, which is later added into the ϵ-error of the protocol.

This offers a lower bound for randomly constructed codes in the asymptotic limit. We plot this bound below with respect to the parameter $\delta = \frac{d}{n}$, which we refer to as the relative distance. Also after a survey, we plot the MRRW2 bound [17], which is the smallest asymptotic upper bound on minimum distance derived up to present.

It is not known if the MMRW2 bound is tight, i.e. it might not be achievable. It is also not known if there exists any linear binary codes in the region between these bounds. This implies that for asymptotic block lengths, a fully randomized construction of the error-

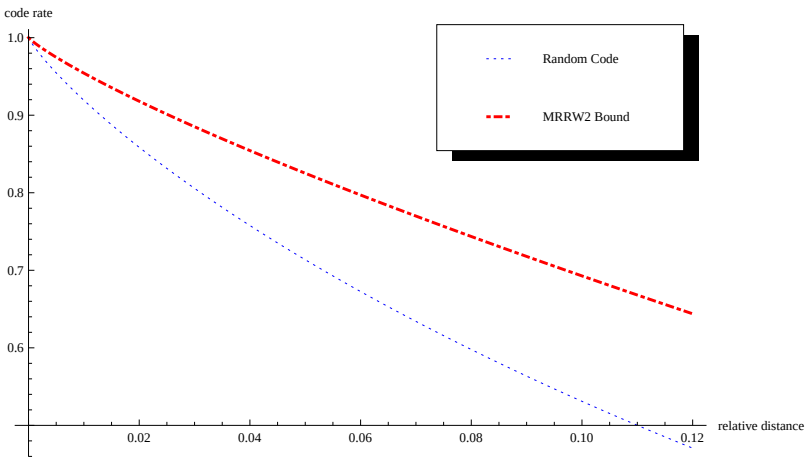


Figure 4.1: The random code bound versus the MMRW2 bound, giving a lower and upper bound respectively on the relative distance of random code, versus code rate.

correcting code provides the best estimate of $\delta - R$ tradeoff known up to the present. The existence of better constructions is a long standing open question in the problem of classical cryptography.

The advantage of using such codes is that they provide the highest tolerable p_{err} according to our analysis, which is approximately 5% in the large n limit. This guarantees that it is sufficient to use a randomly generated code for practical purposes, which will provide us both a good enough distance and code rate, only at the expense of contributing a small error probability.

These codes are easy to construct. Although these codes are usually not used in error-correcting since they are difficult to decode, but since honest parties are never required to decode in the protocol, the usage of such codes are straightforward. For the practical implementation in Chapter 5, we will use a random code to execute commitments.

Low density codes

The fully random binary codes have a relatively high density [2], which in large block length limit is computationally time-consuming and difficult. For efficiency purposes, it is of interest whether we can construct codes with lower density.

Initially, we obtained some low-density parity check (LDPC) codes from [12] that have been proven to work sufficiently for certain classes of QKD protocols. Implementing the calculation of syndrome in Matlab for block lengths of $n \approx 2\times10^5$, these codes offer a much shorter computational time, approximately a few minutes, which is a contrast compared

[2]The density of a code is the fraction of non-zero elements in the parity check matrix.

to the random code we constructed, consuming a computational time as approximately an hour. However, it was realised subsequently that their construction cannot provide a statement about the minimum distance of the code. The assessment of codes done was based on their error-correcting capabilities instead, since in QKD, error-correcting codes are used for the purpose of information reconciliation. However, since error-correcting codes serve as a back-checking tool in our protocol, we need a statement about the minimum distance.

It might not be clear why these two statements can be very different. While the minimum distance bounds the number of errors correctable as explained in Section 2.4, a statement in the reverse direction does not apply. Therefore, we render these constructions unsufficient for our purposes.

However, Gallager has shown in [14] that a specific ensemble of low density codes do attain the same limit given for the random codes as above, when considering large enough block lengths. These codes involve using random permutations of a submatrix, and the construction is straightforward. This class of codes might be of future interest, because their usage will significantly shorten the computation time used in the protocol. However, this is achieved at the expense of introducing an larger error probability ϵ of constructing a bad code (one with unsatisfactory minimum distance), where only a qualitative expression can be obtained for ϵ. Again, there is no way of verifying the distance for such codes with large block length.

Concatenated codes

Having understood two different types of randomized constructions, one might not be convinced that a randomized argument is sufficient to prove the robustness of the protocol against errors in a practical implementation, since the minimum distance cannot be verified via any methods. Given so, one naturally asks if there is a way to guarantee a certain amount of minimum distance without fail, as previously in the ideal setting. Indeed, this can be realised by using explicit constructions of concatenated codes.

By using a Reed-Solomon code as an outer code, while using a smaller binary linear code as an inner code, one can use Theorem 2.4.2 to show a lower bound for the minimum distance, since the minimum distance for inner binary linear codes of small sizes can be obtained by brute force methods.

In the case of low bit-flip errors, classes of explicit concatenated codes might be generated such that the minimum distance is guaranteed without introducing any probabilistic

errors from a randomized construction. For example, by exploiting this construction, by performing a search over small linear binary codes of size $n_{in} \approx 20$, a linear binary concatenated code with rate R=0.53 and relative minimum distance $\delta \geq 0.052$ can be constructed, where the code length $n = 311296$. This value of δ has a large discrepancy compared to the probabilistic argument for a random code, $\delta \geq 0.098$. A general bound for such concatenated codes in the asymptotic limit can be given with the Blokh-Zyablov radius [17], which is much lower than the random code bound.

From here it is clearly shown that, if a definite statement regarding the minimum distance of such large error-correcting codes is wanted without introducing any further probabilistic errors, security can still be obtained for smaller ranges of experimental parameters. For the given example of concatenated RS code, this corresponds to security for bit flip error rates $p_{err} \leq 2\%$.

We acknowledge that there are other classes of codes such that security can be obtained for different ranges of p_{err}, such as the randomized Wozencraft ensemble for concatenated codes. The main purpose of identifying these three classes of examples is to give an overview of choices for codes, highlighting their advantages and disadvantages. Randomized constructions (either fully randomized or partially randomized) always attain the highest bound, at an expense of introducing a small probability error of producing a bad code. Explicit constructions that give definite statements on the other hand, provide a much lower bound on the amount of bit flip error tolerable.

This proof has shown that by using a suitable error-correcting code, the commitment protocol is robust against bit flip errors up to approximately 5%, which is well achieved within experimental range.

Chapter 5

First Secure Experimental Implementation

In this chapter, we provide full statements about the security of commitments, by combining the analysis accounting for both erasures and errors. In Section 5.1, we work towards a simplified expression for the rate of commitment, i.e. to commit one bit securely, what is the required number of signals to send. We present region plots showing where security holds for the protocol.

In Section 5.2, we introduce the setup consisting a parametric down-conversion source. The relevant parameters were estimated by our collaborators in [29], while here we state the assumptions made. Subsequently, by using data obtained from this apparatus, we write the program for modified protocols WSEE and Commitment using Matlab. By implementing the protocol, we execute 500 secure bit commitments in Section 5.2.4, therefore demonstrating the practical feasibility of two-party cryptographic protocols in general under the Noisy Storage Model.

5.1 Security of Commitments for real-world devices

By combining the various parts of analysis in Chapter 4, we evaluate a summarizing result for the security of executing commitments with real world devices, w.r.t. to a range of experimental parameters. In this section, the goal is to derive an expression describing the **number of signals required** for security to hold except for an error parameter ϵ.

We accomplish this by three steps. First, we start by assuming a certain amount of errors p_{err}. With this, we show the condition on block length n, such that a random code

with a specified rate R provides enough minimum distance. Secondly, by applying the value of random code R, we obtain the minimal required smooth min-entropy $\hat{\lambda}$. Thirdly, by fixing the remaining parameters in Table 4.4.1, we summarize all conditions on the number of signals sent by Alice, M such that $\lambda > \hat{\lambda}$.

Firstly, by fixing p_{err} and error ϵ, we can state the condition on block length n such that the conditions on minimum distance is satisfied.

Theorem 5.1.1 (Condition on minimum distance)**.** *Given fixed parameters of ϵ and p_{err},*

$$\delta = \frac{2(p_{err} + 0.7\beta)}{1 - 4\sqrt{5}\beta}, \tag{5.1}$$

for any small value of $\beta \in [0, 0.01]$, where $\delta \in [0.05, 0.11]$. Then for values of smooth min-entropy rate higher than 0.3, the condition on minimum distance will be satisfied except for error probability ϵ, if a random error-correcting code is generated with code rate

$$R = h(\delta) - \frac{1}{n} \log \frac{1}{\epsilon}. \tag{5.2}$$

where h is the binary entropy function $h(x) = -x\log(x) - (1-x)\log(1-x)$, whenever

$$n \geq \frac{1}{\beta^2} \log \frac{2}{\epsilon}. \tag{5.3}$$

Proof. Fix ϵ and assume $\sqrt{\frac{\ln \frac{2}{\epsilon}}{n}} \leq \sqrt{\frac{\log \frac{2}{\epsilon}}{n}} \leq \beta$.[1] Furthermore, assume $\delta \geq 0.05$. This assumption leads to the condition for code rate in Theorem 4.4.5 such that smooth min-entropy $\lambda \geq 0.3$, which is known to be achieved in most cases. Hence from Theorem 4.4.5,

$$\alpha_1 \leq \sqrt{\frac{1}{2}}\,\beta, \qquad \alpha_2 \geq \beta, \qquad \text{and } \alpha_3 \leq \sqrt{20}\beta. \tag{5.4}$$

Comparing this to the condition on minimum distance, we obtain

$$\delta > \frac{2(p_{err} + 0.7\beta)(1 - 2\beta)}{1 - 4\sqrt{5}\beta} > \frac{2(p_{err} + \alpha_2)(\frac{1}{2} - \alpha_1)}{\frac{1}{2} - \alpha_3}. \tag{5.5}$$

□

Meanwhile, the condition on code rate can be translated as a lower bound on the smooth min-entropy rate λ. Combining with Theorem 5.1.1 gives the following lemma:

[1] It might seem unclear at this point why β is defined as an upper bound for the expression containing log instead of ln. The reason for this is to combine β with other terms containing binary logarithm functions, which occur in subsequent proofs.

Lemma 5.1.2 (Conditions on smooth min-entropy for secure bit commitment)**.** *By fixing p_{err} and ϵ, given δ as defined in Theorem 5.1.1. Then the protocol is 2ϵ-secure whenever the smooth min-entropy is*

$$\lambda \geq \hat{\lambda} = h(\delta) + \frac{1 + 3\log\frac{1}{\epsilon}}{n}. \tag{5.6}$$

Proof. From Theorem 2.4.3, for the minimum distance to be $d \geq \delta n$ except for probability less or equal to ϵ, we require

$$\begin{aligned} 2^{(R-C_\delta)n} &\leq \epsilon \\ R &\leq C_\delta - \frac{\log\frac{1}{\epsilon}}{n}. \end{aligned} \tag{5.7}$$

Substituting this into the condition for code rate yields (5.6). □

Now, we need to evaluate the min-entropy rate λ created according to parameters p^1_{sent}, $p^h_{B,noclick}$ and $p^d_{B,noclick}$. By combining all these statements, we construct a theorem describing the number of signals M needed for a secure commitment, given the relevant parameters and assuming the case of bounded storage. Recall that the **block length used in the commitment** n, relates to the **number of signals sent** M, by $n = m_{frac}M$.

Theorem 5.1.3. *Let a dishonest Bob's storage size be bounded by N qubits. For a fixed $\epsilon \leq 0.1$, and given the experimental probabilities listed in Table 4.4.1, let*

$$\delta = \frac{2(p_{err} + \beta)}{1 - 4\sqrt{5}\beta} \tag{5.8}$$

for some $\beta \in (0, 0.01]$ be the required relative minimum distance. Then

$$\hat{\lambda} = h(\delta) + 3\beta^2, \tag{5.9}$$

Subsequently, defining the terms

$$\begin{aligned} m_2 &= p^1_{sent} - p^h_{B,noclick} + p^d_{B,noclick} \\ m_3 &= 1 - p^h_{B,noclick} \\ L' &= \max_{s\in(0,1]} \frac{-1}{s}\left[\log(1 + 2^s) - 1 - s\right] - \frac{3\epsilon}{s}. \end{aligned} \tag{5.10}$$

For security to hold at all, the following condition is required:

$$m_2 L' - m_3\hat{\lambda} > 0. \tag{5.11}$$

If (5.11) *is true, then bit commitment can be implemented 2ϵ-securely by using a randomly constructed error-correcting code, whenever*

$$M > \max\ \{M_1, M_2, M_3, M_4\}, \tag{5.12}$$

where

$$M_1 = 4.5 \times 10^4 \ln \frac{2}{\epsilon}, \quad M_2 = \frac{\log \frac{1}{\epsilon}}{\epsilon \cdot (m_2 - 0.01)},$$
$$M_3 = \frac{\log \frac{2}{\epsilon}}{\epsilon \cdot m_3}, \qquad M_4 = \frac{N}{(m_2 - 0.01)L' - m_3\hat{\lambda}}. \tag{5.13}$$

Proof. By Lemma 4.4.1, the min-entropy rate has the form

$$\lambda = \frac{m^1_{left} \cdot L - \frac{N}{M}}{m_{frac}} > \frac{(m_2 - 3\zeta) \cdot L - \frac{N}{M}}{m_3}, \tag{5.14}$$

where

$$L = \max_{s \in (0,1]} \frac{-1}{s} [\log(1 + 2^s) - 1 - s] - \frac{2 \log \frac{1}{\epsilon} + 1}{m^1_{left} \cdot M}. \tag{5.15}$$

Since $s \leq 1$, assuming $\epsilon \leq 0.1$ gives

$$L \geq \max_{s \in (0,1]} \frac{-1}{s} [\log(1 + 2^s) - 1 - s] - \frac{3 \log \frac{1}{\epsilon}}{m^1_{left} \cdot M}, \tag{5.16}$$

while setting $\frac{\log \frac{1}{\epsilon}}{m^1_{left} \cdot M} \leq \epsilon$ provides M_2, simplifying the lower bound to L'. The extreme condition (5.11) is given by setting $N = 0$ and taking limit as $n \to \infty$, $\zeta \to 0$.

A bound on numbers of signals can be obtained by setting $3\zeta \leq 0.01$ in Theorem 4.4.1, which gives the value for M_1. M_3 comes from the condition given for n at (5.3), while $M = \frac{n}{m_{frac}}$. Lastly,

$$\lambda \geq \frac{(m_2 - 0.01) \cdot L' - \frac{N}{M}}{m_3} \geq \hat{\lambda} \tag{5.17}$$

while rearranging gives the value for M_4. □

The calculations for a storage with general depolarizing noise is even more complicated, but it can be done with the aid of Mathematica. Based on the calculations, we include some security region plots for the relevant parameters as follows. These plots are done for fixed values of $n = 250000$, $\epsilon = 10^{-4}$. The number of signals sent is given by $M = \frac{n}{1 - p^h_{B,noclick}}$. Also, the storage noise parameter is set for a low noise of $r = 0.9$. Finally, Bob's storage size is quantified by $N = \nu m_{store}$, where $m_{store} = p^1_{sent} \cdot (1 - p^h_{B,noclick})$ defines an upper bound for Bob's storage.

Figures 5.1-4 show the security region for $p^h_{B,noclick}$ and p_{err}, quantifying the amount of erasures and errors in the protocol. The only difference between Figure 5.1-2 and Figure 5.3-4 is the sum $p^d_{B,noclick} + p^1_{sent}$, which are 0.99 and 0.999 respectively. The higher this

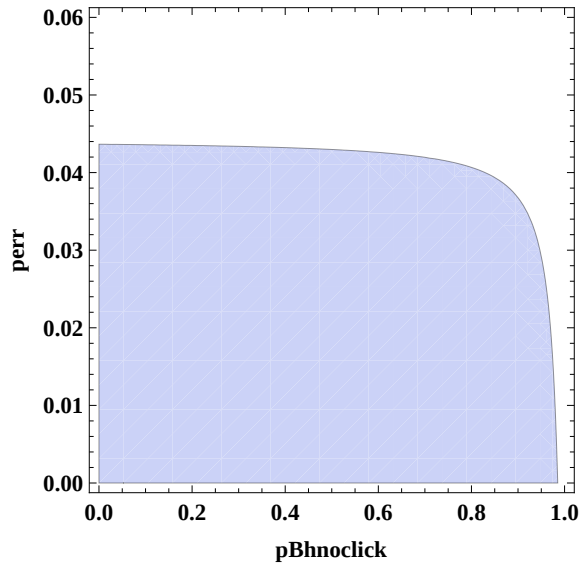


Figure 5.1: Security region plot for $p^h_{B,noclick}$ versus p_{err}, with other parameters $\nu = 0.01$, $p^1_{sent} = 0.765$ and $p^d_{B,noclick} = 0.225$.

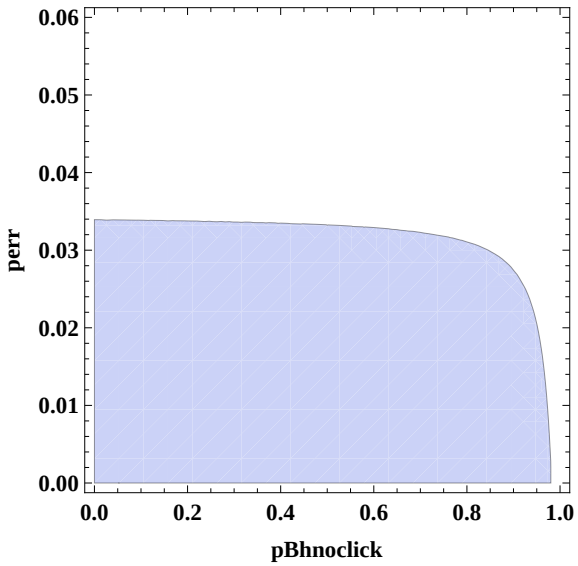


Figure 5.2: Security region plot for $p^h_{B,noclick}$ versus p_{err},with a larger storage rate of $\nu = 0.1$, $p^1_{sent} = 0.765$ and $p^d_{B,noclick} = 0.225$.

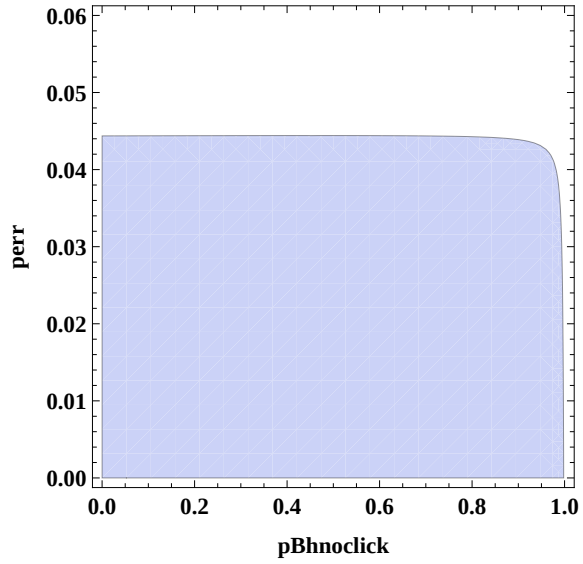


Figure 5.3: Security region plot for $p^h_{B,noclick}$ versus p_{err}, with other parameters $\nu = 0.01$, $p^1_{sent} = 0.765$ and $p^d_{B,noclick} = 0.234$.

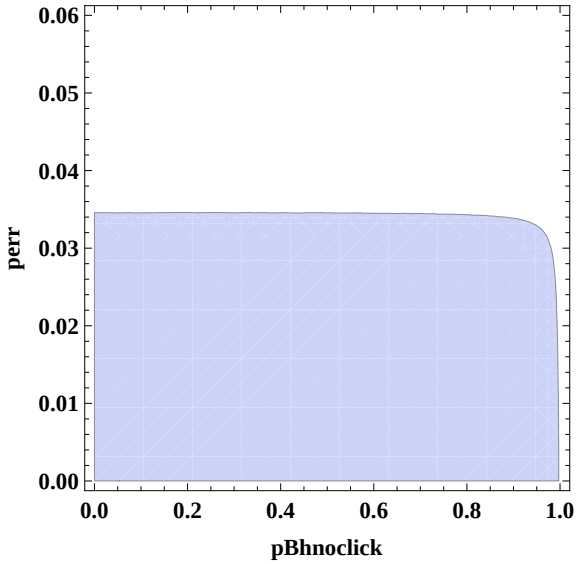


Figure 5.4: Security region plot for $p^h_{B,noclick}$ versus p_{err},with a larger storage rate of $\nu = 0.1$, $p^1_{sent} = 0.765$ and $p^d_{B,noclick} = 0.234$.

summation value, the less multiphotons Bob gets, hence confirming our reasoning that if the source is extremely close to ideal, erasures alone do not impact security. Dependences in the security region between erasures and errors become more obvious when $p^d_{B,noclick} + p^1_{sent}$ is low. Also, larger assumptions on the storage rate directly decreases the amount of min-entropy guaranteed, causing the amount of tolerable p_{err} to drop consistently for

all amounts of erasures.

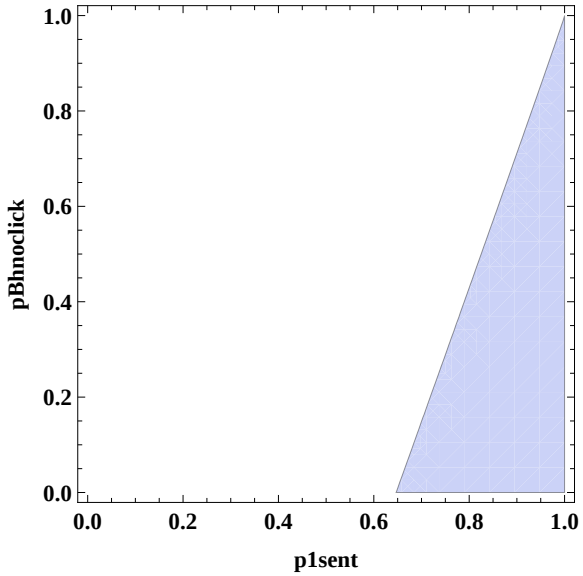


Figure 5.5: Security region plot for p^1_{sent} versus $p^h_{B,noclick}$, with $\nu = 0.02$, $p_{err} = 0.02$.

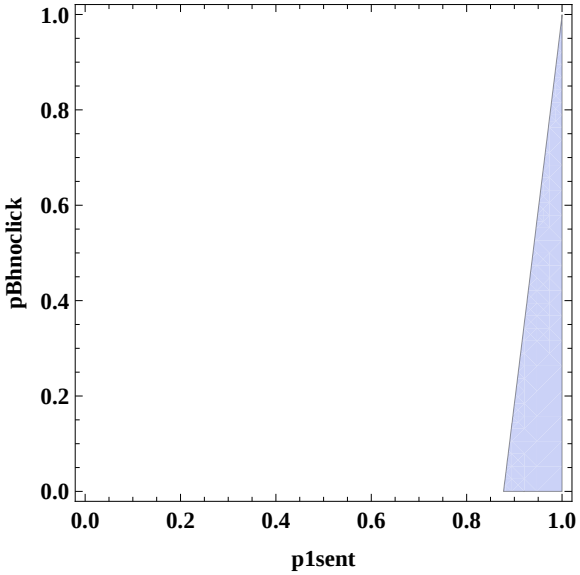


Figure 5.6: Security region plot for p^1_{sent} versus $p^h_{B,noclick}$, with $\nu = 0.02$, for a larger $p_{err} = 0.035$.

Figures 5.5-6 show the allowed values of p^1_{sent} for two different values of p_{err}, assuming that $p^d_{B,noclick} = 0$. For extremely small values of $p^h_{B,noclick}$ (indicating large amounts of losses), there exists a threshold on p^1_{sent} such the the protocol holds secure. This threshold increases with p_{err}, and for extremely small storage rates, it gives a maximum tolerable $p_{err} \approx 0.046$.

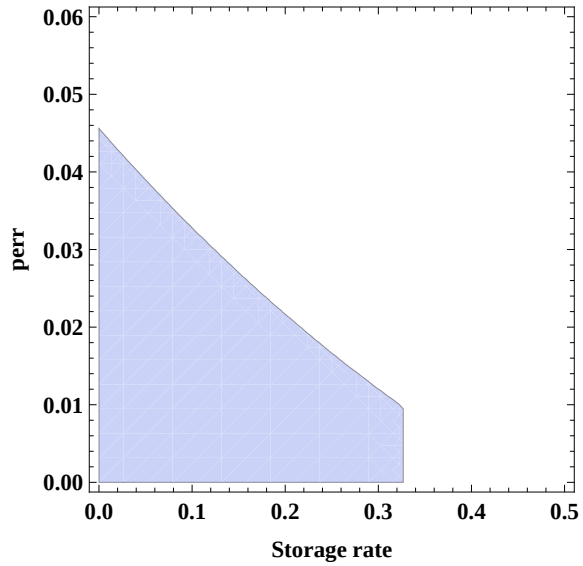


Figure 5.7: Security region plot for storage rate ν versus p_{err}, with $p^1_{sent} = 0.765$, $p^d_{B,noclick} = 0.234$ and $p^h_{B,noclick} = 0.1$.

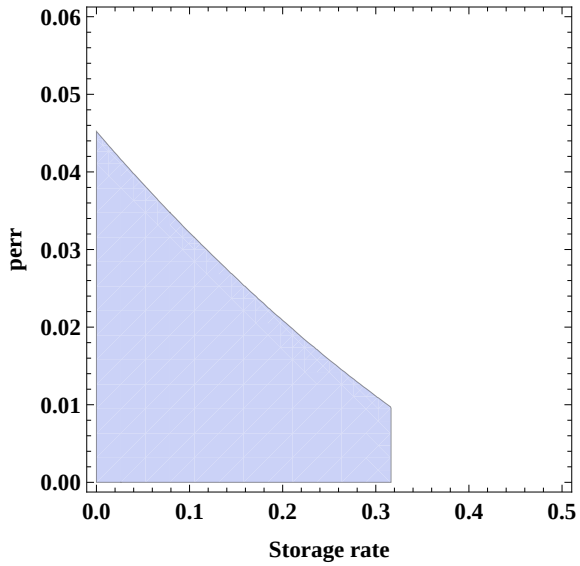


Figure 5.8: Security region plot for storage rate ν versus p_{err}, with $p^1_{sent} = 0.765$, $p^d_{B,noclick} = 0.234$ and $p^h_{B,noclick} = 0.88$.

Figures 5.7-8 show a monotonic decreasing trend for tolerable p_{err} w.r.t storage rate ν, as ν directly affects the min-entropy rate generated. The sharp cut-off of security for storage rate varies with $p^h_{B,noclick}$, since the higher $p^h_{B,noclick}$, the lower detection efficiency and the more dishonest Bob is allowed to cheat reporting rounds as missing, hence the lower his storage rate is needed for a high min-entropy. Also, this shows security for depolarizing noise storage for mostly low values of storage rate. This is non-optimal, since it has been shown in [4] that security can be achieved with arbitrarily larger storage rates, if the depolarizing noise parameter $r \lesssim 0.7$. This is due to the fact that we have bounded the smooth min-entropy of an adversarial Bob by the *classical capacity* of a quantum memory, while [4] shows security in terms of *entanglement cost*. Since the entanglement cost of a quantum channel is generally smaller than classical capacity, this poses an even better advantage for security, which is not shown in our part of the analysis.

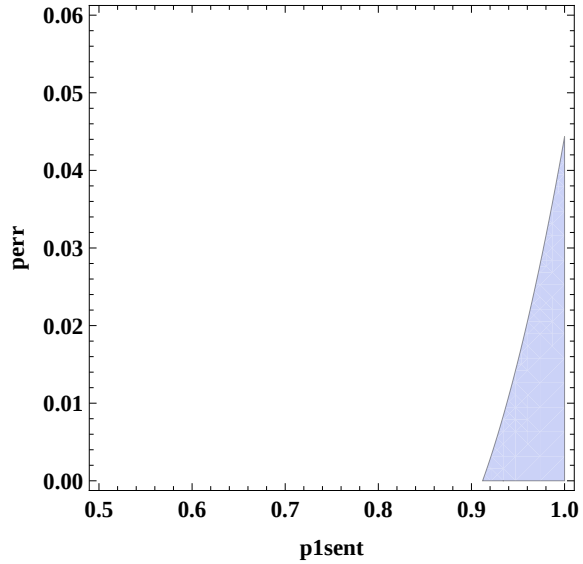


Figure 5.9: Security region plot for p^1_{sent} versus p_{err}, with $p^d_{B,noclick} = 0$, $p^h_{B,noclick} = 0.88$ and $\nu = 0.01$.

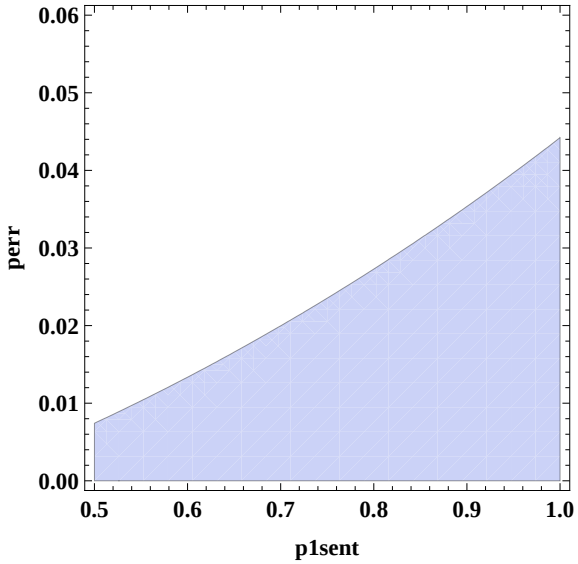


Figure 5.10: Security region plot for p^1_{sent} versus p_{err}, with $p^d_{B,noclick} = 0$, $p^h_{B,noclick} = 0.15$ and $\nu = 0.01$.

Comparing Figures 5.9-10 directly, it is clearly shown again that to compensate for low detection efficiency, i.e. the source should approximate a single-photon source.

5.2 Execution of protocol

Figure 5.9 shows the experimental conducted in [29]. Polarization-entangled photon pairs are generated via spontaneous parametric down conversion of blue light from a laser diode in Barium-betaborate crystal (BBO), and distributed to polarization analyzers

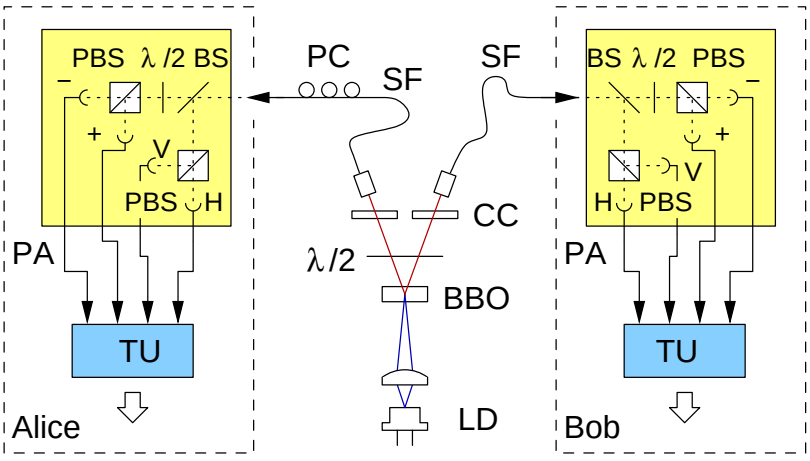


Figure 5.11: Experimental setup.

(PA) at Alice and Bob via single mode optical fibres(SF). A non-polarizing beam splitter (BS) decides a random measurement base choice for an honest party, with a half wave plate ($\lambda/2$) at one of the outputs. Polarizing beam splitters (PBS) are placed in front of single-photon counting photodiodes. Detection events on both sides are timestamped and recorded for processing. A polarization controller (PC) ensures that polarization anti-correlations are observed in all measurement bases.

Experimental data have been obtained for honest parties. In this section, we briefly discuss the estimation of parameters, and the processing of data. We use the data to perform honest and secure commitments.

5.2.1 Estimation of parameters

The parameters for honest parties given this setup, as previously stated in Section 4.4.1 are evaluated by our collaborators. Recall that each parameter has to be *conditioned* on the event that Alice registers a single click on her detector.

The main observables in this experiment are the rates of detection events for both Alice and Bob individually, and the pair rate where both parties obtain a detection event. The parameter $p^h_{B,noclick}$ can be directly obtained from these measured quantities.

For the parameter estimation on p^1_{sent} and $p^d_{B,noclick}$ [2], we need some assumptions about the source itself. The source in reality is a thermal light source. However, under short durations of time, the photon counting statistics can be approximated by Poissonian distribution, which is the only assumption in parameter estimation. Furthermore, only

[2]Recall that the value $p^d_{B,noclick}$ is equal to p^0_{sent}, the probability that no photons were *sent* to Bob, when Alice registers a single detection.

appropriate upper/lower bounds can be obtained for all of these values. However, a lower bound on p^1_{sent} and $p^d_{B,noclick}$ is sufficient for the security analysis.

After a combined effort, the probabilistic parameters required are concluded to be:

$$\begin{aligned} p^1_{sent} &\geq 0.7370 \ \pm 0.1258 \\ p^d_{B,noclick} &\leq 0.2630 \ \pm 0.1263 \\ p^1_{sent} + p^d_{B,noclick} &\geq 0.9991 \ \pm 0.0014 \\ p^h_{B,noclick} &= 0.8738 \ \pm 0.0036 \\ p_{err} &= 0.0429 \ \pm 0.0006 \end{aligned} \tag{5.18}$$

It is verified that these parameters satisfy the security condition from our analysis of the commitment protocol, for a small storage rate assumption of $\nu = 0.01$, corresponding to an actual storage size of approximately 2000 qubits.

5.2.2 Data symmetrization

In this protocol, it is crucial that both the fraction of rounds basis used, and the bit value should be uniform,

$$p(\theta = 0) = p(\theta = 1) = \frac{1}{2}, \qquad \text{and} \qquad p(X = 0) = p(X = 1) = \frac{1}{2}. \tag{5.19}$$

Otherwise, Bob can lower his min-entropy by always measuring the more likely basis. Also, whenever bases do not match, he can guess a more likely value of X.

After obtaining the data from experiment, a calculation was performed to show that a small but finite bias in the choice of basis and bit value. This is because the setup comprises of 4 photon detectors for both Alice and Bob. Each detector has a slightly different detection efficiency, which directly causes this asymmetry. In fact, this error persists in any typical BB84 experiment! However, this does not affect the security proof for an honest Alice. Note that since erasures do not affect Alice, she is free to discard any rounds without affecting the security proof. However, by throwing away a certain number of rounds, she can effectively change the detection efficiency of her own detectors.

We calculate the fraction of rounds $p_{X\theta}$, for $X = \{0, 1\}$ and $\theta = \{0, 1\}$. These values were obtained for 579 sets of data, each set having a length of $n = 250,000$. The values obtained are listed below:

$$\begin{aligned} p_{00} &= 0.2323 \pm 0.0010 & \qquad p_{10} &= 0.2395 \pm 0.0011 \\ p_{01} &= 0.2622 \pm 0.0009 & \qquad p_{11} &= 0.2660 \pm 0.0009 \end{aligned} \tag{5.20}$$

Subsequently, defining $p_{\mathrm{k}|X,\theta}$ to be the probability of keeping a round conditioned on the values of X and Θ. Note that our goal is to achieve $p_{x,\theta|\mathrm{k}} = \frac{1}{4}$ for any combinations of X and θ. Evaluting by Bayes' rule

$$p_{x\theta|\mathrm{k}} = \frac{p_{x\theta\mathrm{k}}}{p_{\mathrm{k}}} = \frac{p_{\mathrm{k}|x\theta} \cdot p_{x\theta}}{\sum_{x,\theta} p_{x,\theta} p_{\mathrm{k}|x,\theta}}, \tag{5.21}$$

it is implied that $p_{\mathrm{k}|x\theta} \cdot p_{x\theta}$ should be equal for all x and θ. Taking the lowest value $p_{00} = 0.231515$ and $p_{\mathrm{k}|00} = 1$, we can calculate the set of values $\{p_{\mathrm{k}|00}, p_{\mathrm{k}|01}, p_{\mathrm{k}|10}, p_{\mathrm{k}|11}\} = \{1, 0.963707, 0.882338, 0.871353\}$. $p_{\mathrm{k}} = 0.9292$ meaning that 92% of the rounds are kept. By processing the data and discarding certain fraction of rounds using these probabilities, we obtain a set of symmetrized data. The new values are listed below:

$$\begin{aligned} p_{00} &= 0.2508 \pm 0.0011 & p_{10} &= 0.2498 \pm 0.0009 \\ p_{01} &= 0.2491 \pm 0.0011 & p_{11} &= 0.2502 \pm 0.0009 \end{aligned} \tag{5.22}$$

Bias affects security for honest Bob in a similar fashion. Symmetrization can also be done on Bob's side by exactly the same method. Since bit values are already symmetrized on Alice's side. This effectively changes the value of $p^h_{B,noclick}$, and the increase can be calculated by

$$p^{h,new}_{B,noclick} = 1 - p_{B,keep} \cdot (1 - p^h_{B,noclick}) \tag{5.23}$$

where $p_{B,keep}$ is the fraction of rounds Bob keep during the symmetrization process.

In our set of data, the bias on Bob's side is however considerably smaller, therefore we omit the process. We read a total amount of 1.25×10^8 rounds from symmetrized experimental data, and use it to perform 500 commitments of a single bit.

5.2.3 Generating Error correcting code and hash functions

By creating a random seed, a random binary matrix of dimensions 132500×117500 was generated in Matlab and used to define the the submatrix P of the parity check matrix $H = [\ P \mid \mathbb{I}_{(1-R)n}\ ]$. The code rate and block length are $R = 0.53$ and $n = 250000$ respectively. Meanwhile, 2-universal hash functions r are generated similarly, by creating 500 random binary strings of length n. The weight of hash functions generated is $(0.5000 \pm 0.0010)n$. The function $\mathrm{Ext}(X^n, r)$ is defined by the modulo two of the summation $\sum_{i=1}^{n} X(i) \cdot r(i)$.

5.2.4 Discussion

After executing the protocol, we find that correctness holds for all the 500 honest bit commitments executed. We tabulate the results of various parameters as below:

Quantity	Mean value	Standard Deviation	Min. value	Max. value	Skewness
$\|X_{\mathcal{I}}\|$	125142	254	124383	125899	0.1242
p_{err}	0.0426	0.0006	0.0402	0.4421	-0.2248
$\Delta(r)$	125003	246	124109	125711	0.0294

According to the symmetrized data, Bob obtains a substring of length $|X_{\mathcal{I}}| = (0.5005 \pm 0.0010)n$, while the fraction of errors were calculated to be $\bar{p}_{err} = 0.0426 \pm 0.0006$. The mean value $\bar{p}_{err}$ is slightly less compared to the value of $\bar{p}_{err} = 0.0428 \pm 0.0007$ before symmetrization, indicating that there might be correlations between the occurrence of erasures and errors in the apparatus.

For a security parameter $\epsilon = 10^{-4}$, Bob accepts an amount of errors within the interval $[(\bar{p}_{err} - 0.004, \bar{p}_{err} + 0.004)]$. Even for a security parameter of $\epsilon = 0.01$ in our proof, Bob accepts an interval $[(\bar{p}_{err} - 0.003, \bar{p}_{err} + 0.003)]$, suggesting that the data collected from the apparatus has a much narrower distribution compared to that bounded of the Hoeffding's inequality in our security analysis.

With this we end the discussion on the protocol execution, concluding that we have successfully demonstrated the feasibility of bit commitment in the Noisy Storage Model. The computational time however, is long due to the large density of the random code. Rewriting in different computing languages such as C, or the use of a different class of error-correcting code would increase the efficiency of this software. In other words, further refinements and improvements are necessary for widely implementing these protocols.

Chapter 6

Conclusions and Future Work

The Noisy Storage Model is a present-day scenario where useful two-party cryptographic protocols can achieve information-theoretic security. It works on the physical assumption that resources for quantum memory is scarce. Within the scope of this model, we have studied the implementation of two-party cryptography protocols and the necessary conditions for security to hold.

Results

Protocol and analysis: We adapt the commitment protocol of [19] and provide a full analysis of its security. In essence, we have modified the back-checking procedures for honest parties, and introduce the usage of new error-correcting codes, such that the protocol becomes robust against experimental losses and errors. By such methods, we obtain security for a wide range of experimental parameters. Any QKD apparatus with parameters satisfying our security condition can then be used to implement bit commitment securely. Working with a group of experimentalists, we perform the first ever implementation of bit commitment in the noisy-storage model, demonstrating the feasibility of two-party protocols in such models.

New uncertainty relation: We developed new uncertainty relations for BB84 measurements, which have fundamental implications for the practical feasibility of two-party protocols in present day quantum cryptography. Previously, the security of two party protocols was based on the uncertainty relation developed in [9], which yields an exponentially decreasing error in the limit of large block length. While this is sufficient for a proof of principle, an implementation based on this relation is extremely time-consuming due to the size of block length required for a small error parameter. We have proven

entropic uncertainty relations that pave the way for practical implementations of both BB84 and six-state protocols [9, 38, 34] at small block length. As part of our proof, we have shown tight uncertainty relations for a family of Rényi entropies that may be of independent interest. For our experimental implementation of commitments, this result has significantly reduced the amount of classical information post-processing required.

Further Work

Optimality and efficiency of the protocol: The preliminary implementation done in this work has not been optimized. The optimality of our analysis is unconfirmed: for example, we do not know if providing the syndrome to Bob actually enables him to gain knowledge about $(1 - R)n$ bits in the string X^n. A short survey into list decoding has been inconclusive, hence we omit this from our discussion. It would be interesting to find ways to construct an error-correcting code with small leakage, i.e. Bob does not learn much about the string X^n from its syndrome.

Also, to systematically use the protocol for applications, it would be desirable to refine and optimize the execution process, in order to achieve an optimal commitment rate (the number of commitments per number of block length).

Variants of the protocol: The WSEE protocol is useful, since most of two-party cryptographic protocols can achieve security under the assumption that both parties share correlated randomness. Oblivious transfer and secret sharing are examples of such protocols. In [26], variants of the WSE protocol have been investigated, by using the same working principle to distribute a string with elements defined over a larger alphabet. An implementation of this potentially involves the transmission and measurement of many-level quantum systems (or qudits), where the analysis of security might be complicated.

For the commitment protocol, there are a few advantages and disadvantages that can be forseen. Firstly, such quantum systems are even harder to store, implying that the Noisy Storage Model assumption might be even stronger. Also, in our protocol the error-correcting codes are required to be binary. Higher $\delta - R$ tradeoffs are however known for codes defined over larger alphabets. For example, algebraic geometric codes offer such constructions. Such variants of WSE and Commitment schemes might potentially provide such freedom. However, the transmission of such quantum systems will also be subjected to a larger amount of noise. Also, the security proof will be subjected to many changes.

Further investigation of quantum side information: Although methods have been

developed to quantify smooth min-entropy conditioned on classical side information, the known results of further conditioning entropy on quantum side information is very limited. Some approaches to this problem by using quantum-to-classical randomness extractors have been developed in [5] very recently. These fundamental problems offer a better security condition for the Noisy-Storage Model, by linking the ability of a quantum memory to preserve information by quantities such as its *quantum capacity* and *entanglement cost*, which are in general much lower than the classical capacity used in our analysis.

In general, the various capacities of quantum channels are very interesting to explore, and serves as a useful fundamental tool in understanding the preservation of quantum information. It would be desirable to prove security for larger classes of realistic quantum channels, by investigating these quantities.

Bibliography

[1] S. Bandyopadhyay, P. Oscar Boykin, Vwani Roychowdhury, and Farrokh Vatan. A new proof for the existence of mutually unbiased bases, 2001.

[2] Stephen M Barnett and Sarah Croke. Quantum state discrimination. *Advances in Optics and Photonics*, 1(2):43, 2008.

[3] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, 1984.

[4] M. Berta, F. Brandao, M. Christandl, and S. Wehner. Entanglement cost of quantum channels. arXiv:1108.5357, 2011.

[5] M. Berta, O. Fawzi, and S. Wehner. Quantum to classical randomness extractors. page 33, 2011. arXiv:1111.2026v2.

[6] C. Cachin and U. M. Maurer. Unconditional security against memory-bounded adversaries. In *Proceedings of CRYPTO 1997*, Lecture Notes in Computer Science, pages 292–306, 1997.

[7] H.F. Chau and H-K. Lo. Making an empty promise with a quantum computer. *Fortschritte der Physik*, 46:507–520, 1998. Republished in 'Quantum Computing, where do we want to go tomorrow?' edited by S. Braunstein, arXiv:quant-ph/9709053v2.

[8] I. Damgaard, Serge Fehr, Louis Salvail, and Christian Schaffner. *LNCS*, 4622:22, 2007.

[9] I. B. Damgård, S. Fehr, R. Renner, L. Salvail, and C. Schaffner. A tight high-order entropic quantum uncertainty relation with applications. In *Proceedings of CRYPTO 2007*, Springer LNCS, pages 360–378, 2007.

[10] I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Cryptography in the Bounded-Quantum-Storage Model. In *Proceedings of 46th IEEE Symposium on Foundations of Computer Science*, pages 449–458, 2005.

[11] G. D'Ariano, D. Kretschmann, D. Schlingemann, and R.F. Werner. Quantum bit commitment revisited: the possible and the impossible. arXiv:quant-ph/0605224v2, 2007.

[12] D. Elkouss, A. Leverrier, R. Alléaume, and J. J. Boutros. Efficient reconciliation protocol for discrete-variable quantum key distribution. In *Proceedings of the 2009 IEEE international conference on Symposium on Information Theory - Volume 3*, ISIT'09, pages 1879–1883, Piscataway, NJ, USA, 2009. IEEE Press.

[13] G. D. Froney, Jr. Concatenated codes. 1965.

[14] Robert G. Gallager. *Low-Density Parity-Check Codes*. PhD thesis, 1963.

[15] C. W. Helstrom. Detection theory and quantum mechanics. *Information and Control*, 10:254–291, 1967.

[16] W. Hoeffding. Probability Inequalities for Sums of Bounded Random Variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[17] W. C. Huffman and Vera Pless. *Fundamentals of error-correcting codes*. Cambridge University Press, 2003.

[18] R. König and S. Wehner. A strong converse for classical channel coding using entangled inputs. *Physical Review Letters*, 103:070504, 2009. arXiv.org:quant-ph/0903.2838v1.

[19] R. König, S. Wehner, and J. Wullschleger. Unconditional security from noisy quantum storage. *IEEE Transactions on Information Theory - To appear*, 2009.

[20] Robert König, Renato Renner, and Christian Schaffner. The operational meaning of min- and max-entropy. *IEEE Trans. Inf. Theor.*, 55:4337–4347, September 2009.

[21] B. P. Lanyon, T. J. Weinhold, N. K. Langford, M. Barbieri, D. F. V. James, A. Gilchrist, and A. G. White. Experimental Demonstration of a Compiled Version of Shor's Algorithm with Quantum Entanglement. *Physical Review Letters*, 99(25):250505, December 2007.

[22] H-K. Lo. Insecurity of quantum secure computations. *Physical Review A*, 56:1154, 1997.

[23] H-K. Lo and H. F. Chau. Is quantum bit commitment really possible? *Physical Review Letters*, 78:3410, 1997.

[24] C.-Y. Lu, D. E. Browne, T. Yang, and J.-W. Pan. Demonstration of a Compiled Version of Shor's Quantum Factoring Algorithm Using Photonic Qubits. *Physical Review Letters*, 99(25):250504, December 2007.

[25] H. Maassen and J. Uffink. Generalised entropic uncertainty relations. *Physical Review Letters*, 60:1103–1106, 1988.

[26] P. Mandayam and S. Wehner. Achieving the physical limits of the bounded-storage model. *Physical Review A*, 83:022329, 2011. arXiv:1009.1596v2.

[27] U. Maurer. Conditionally-perfect secrecy and a provably-secure randomized cipher. *Journal of Cryptology*, 5:53–66, 1992.

[28] D. Mayers. Unconditionally secure quantum bit commitment is impossible. *Physical Review Letters*, 78:3414–3417, 1997.

[29] N. Ng, S. K.Joshi, C. M. Chia, C. Kurtsiefer, and S. Wehner. First experimental implementation of bit commitment in the noisy-storage model. arXiv, 2012.

[30] R. Renner. Security of quantum key distribution. *International Journal of Quantum Information*, 6:1, 2008. arXiv:quant-ph/0512258v2.

[31] R. Renner and R. König. Universally composable privacy amplification against quantum adversaries. *Theory of Cryptography*, pages 407–425, 2005. arXiv:quant-ph/0403133v2.

[32] L. Rivest. Unconditionally secure commitment and oblivious transfer schemes using private channels and a trusted initializer, 1999.

[33] C. Schaffner. Simple protocols for oblivious transfer and secure identification in the noisy-quantum-storage model. *PHYS.REV.A*, 82:032308, 2010.

[34] C. Schaffner. Simple protocols for oblivious transfer and secure identification in the noisy-quantum-storage model. *Physical Review A*, 82:032308, 2010. arXiv:1002.1495v2.

[35] C. Schaffner, B. Terhal, and S. Wehner. Robust cryptography in the noisy-quantum-storage model. *Quantum Information & Computation*, 9:11, 2008.

[36] M. Tomamichel, R. Colbeck, and R. Renner. A fully quantum asymptotic equipartition property. *IEEE Transactions on Information Theory*, 55:5840, 2009.

[37] S. Wehner, M. Curty, C. Schaffner, and H.-K. Lo. Implementation of two-party protocols in the noisy-storage model. *Physical Review A*, 81:052336, 2010.

[38] S. Wehner, C. Schaffner, and B. Terhal. Cryptography from noisy storage. *Physical Review Letters*, 100:220502, 2008. arXiv:0711.2895v3.

Appendix A

Proof: minimum distance of a random linear binary code

Consider a parity check matrix H defined as $H = [\ P \mid \mathbb{I}_{(1-R)n}\ ]$, where P is a randomly generated binary linear matrix with dimensions $(1-R)n \times Rn$. Any codeword x defined by this error-correcting code has to satisfy the $Hx = \mathbf{0}$, i.e. the inner product of each row vector in H with x must be zero. This can be seen as each rows of H defining a subset, where the parity of bits for x defined by this subset is even. Then regardless of the Hamming weight of x, the probability of x satisfying all the parity checks is equal to $2^{-(1-R)n}$.

Now, we calculate the expected number of codewords with Hamming weight l. This is given by

$$N(l) = \binom{n}{k} 2^{-(1-R)n}. \tag{A.1}$$

This expression can be bounded by the Stirling's approximation. We state this as a theorem below:

Theorem A.0.1. *Let $N(l)$ be the expected number of codewords of weight l, satisfying parity checks defined by a parity check matrix chosen from an equiprobable ensemble of codes with block length n and rate R. Then*

$$N(l) = \binom{n}{k} 2^{-(1-R)n} \leq [2\pi n\lambda(1-\lambda)]^{-\frac{1}{2}} 2^{[h(\lambda)-(1-R)]n} \tag{A.2}$$

where $\lambda = \frac{l}{n}$.

Proof. For a binomial coefficient $\binom{n}{l}$, writing $l = \lambda n$, the Stirling's approximation can be

used to bound it by

$$\binom{n}{\lambda n} \leq \frac{1}{\sqrt{2\pi n\lambda(1-\lambda)}} e^{nh(\lambda)}. \tag{A.3}$$

where $h\lambda = -\lambda \log \lambda - (1-\lambda)\log(\lambda)$ is the bianry entropy. □

The probability of having a codeword less than weight δn, is then less than the sum of probabilities that individual sequences are codewords. Hence,

$$\begin{aligned}
\text{Prb}(d \leq \delta n) &\leq \sum_{l=1}^{\delta n} \binom{n}{l} 2^{-n(1-R)} \\
&\leq 2^{-n(1-R)} \binom{n}{\delta n} \left[1 + \frac{n\delta}{n - n\delta + 1} + \frac{n\delta(n\delta - 1)}{(n - n\delta + 1)(n - n\delta + 2)} + ...\right] \\
&\leq 2^{-n(1-R)} \binom{n}{\delta n} \frac{1-\delta}{1-2\delta}.
\end{aligned}$$

where the third inequality is obtained by bounding the summation by an infinite geometric series. Hence we obtain the expression

$$\text{Prb}(d \leq \delta n) \leq \frac{1}{1-2\delta}\sqrt{\frac{1-\delta}{2\pi n\delta}}\ 2^{-n(C_\delta - R)} \tag{A.4}$$

where $C_\delta = 1 - h(\delta)$. Since the polynomial term scales with $\sqrt{\frac{1}{n}}$, this expression can be simplified to

$$\text{Prb}(d \leq \delta n) \leq 2^{-n(C_\delta - R)}. \tag{A.5}$$

Appendix B

Proof: Lemma 3.2.1

Since a lies within the convergence radius of the function $(1 \pm a)^s$, we expand the function in Taylor's series

$$
\begin{aligned}
& s \cdot [(1+a)^{s-1} + (1-a)^{s-1}] - \frac{1}{a}(1+a)^s - (1-a)^s] \\
&= 2s \cdot \left[1 + \sum_{n=2,4\ldots} \frac{(s-1)(s-2)\ldots(s-n)}{n!} a^n \right] - \frac{1}{a} \left[2as + 2 \sum_{n=3,5\ldots} \frac{s(s-1)\ldots(s-n+1)}{n!} a^n \right] \\
&= 2s \cdot \left[\sum_{n=2,4\ldots} \frac{(s-1)(s-2)\ldots(s-n)}{n!} a^n - \sum_{n=3,5\ldots} \frac{(s-1)(s-2)\ldots(s-n+1)}{n!} a^{n-1} \right] \\
&= 2s \cdot \left[\sum_{n=2,4\ldots} \frac{(s-1)(s-2)\ldots(s-n)}{n!} a^n - \sum_{j=2,4\ldots} \frac{(s-1)(s-2)\ldots(s-j)}{(j+1)!} a^j \right] \\
&= 2s \cdot \sum_{n=2,4\ldots} (s-1)(s-2)\ldots(s-n) \; \frac{n}{(n+1)!} \; a^n \\
&\geq 0. \qquad \text{(B.1)}
\end{aligned}
$$

The first equality holds by a straightforward expansion of Taylor's series, the second equality by extracting $2s$ and absorbing $\frac{1}{a}$ into the second summation term, the third equality follows from redefining the summation variable $j = n-1$, and the last inequality follows because $(s-1)\ldots(s-n) \geq 0$ for $s \in (0,1]$ and n being an even integer.

Appendix C

Uncertainty relation for six states

In this section, we make use of similar methods in Chapter 3.2 to derive an uncertainty relation of an n-qubit state based on the measurement bases σ_x, σ_y and σ_z. The proof is essentially the same: except for Step 1 where an additional free parameter is introduced, and Step 2 where the measurement operator of interest is given by the operator defined on the eigenbases of σ_x, σ_y and σ_z.

Theorem C.0.2. *Given any one-qubit density matrix ρ, and let $\alpha = 1+s$, where $s \in (0, 1]$. Then for measurements of six-states,*

$$\mathrm{H}_\alpha(X|\Theta)_{\rho|\rho} \geq \frac{1}{\alpha - 1}\left[\log \frac{1 + 2^{2-\alpha}}{3}\right]. \tag{C.1}$$

Proof. We evaluate the term

$$\begin{aligned} P_{1+s}(X|\Theta) &= \frac{1}{3}\sum_{x\in\{0,1\}}\sum_{\theta\in\{0,1,2\}} p_{x|\theta}^{1+s} \\ &= \frac{1}{3}\cdot\frac{1}{2^{1+s}}\sum_{i=0}^{2}[(1+x_i)^{1+s} + (1-x_i)^{1+s}] \end{aligned} \tag{C.2}$$

where $\{x_0, x_1, x_2\} := \{x, y, z\}$ and $x_i := \mathrm{tr}(\sigma_{x_i}\rho)$. Parametrizing this in terms of spherical coordinates, we write

$$x_0 = r\sin\phi\sin\theta, \qquad x_1 = r\cos\phi\sin\theta, \qquad x_2 = r\cos\theta. \tag{C.3}$$

where $0 \leq r \leq 1$, $0 \leq \phi, \theta \leq \frac{\pi}{2}$. The expression (C.2) can be rewritten in terms of these

new coordinates as

$$M(s,r,\phi,\theta) := \frac{1}{3}\cdot\frac{1}{2^{1+s}}\sum_{p=0,1}[1+(-1)^p r\sin\phi\sin\theta]^{1+s}+\frac{1}{3}\cdot\frac{1}{2^{1+s}}\sum_{p=0,1}[1+(-1)^p r\cos\phi\sin\theta]^{1+s} + \frac{1}{3}\cdot\frac{1}{2^{1+s}}\sum_{p=0,1}[1+(-1)^p\cos\theta]^{1+s} \tag{C.4}$$

Evaluating the partial differential of Q(s,r,ϕ) with respect to r,

$$\frac{\partial M(s,r,\phi,\theta)}{\partial r} = \frac{1+s}{3}\cdot\frac{1}{2^{1+s}}\sin\theta\cdot\{\sin\phi[(1+r\sin\phi\sin\theta)^s-(1-r\sin\phi\sin\theta)^s] + \cos\phi[(1+r\cos\phi\sin\theta)^s-(1-r\cos\phi\sin\theta)^s]\}. \tag{C.5}$$

Again we see that since in the range of ϕ,θ, all values of sines and cosines are positive, we obtain $\frac{\partial M(s,r,\phi,\theta)}{\partial r}\geq 0$, which implies the maximum is attained at r=1. Subsequently, evaluating the partial derivative

$$\frac{\partial M(s,1,\phi,\theta)}{\partial \phi} = \frac{1+s}{3}\cdot\frac{1}{2^{1+s}}\sin\theta\cdot\{\cos\phi[(1+\sin\phi\sin\theta)^s-(1-\sin\phi\sin\theta)^s] - \sin\phi[(1+r\cos\phi\sin\theta)^s-(1-r\cos\phi\sin\theta)^s]\}. \tag{C.6}$$

which solution gives the points $\phi=0,\ \frac{\pi}{4},\ \frac{\pi}{2}$. We continue by evaluating the second partial derivative at:

$$\left.\frac{\partial^2 M(s,1,\phi,\theta)}{\partial\phi^2}\right|_{\phi=0} = \frac{1+s}{3}\cdot\frac{1}{2^{1+s}}\cdot\sin\theta\ [2s\sin\theta-[(1+\sin\theta)^s-(1-\sin\theta)^s]]$$

$$\left.\frac{\partial^2 M(s,1,\phi,\theta)}{\partial\phi^2}\right|_{\phi=\frac{\pi}{4}} = \frac{1+s}{3\cdot 2^s}\cdot c^2\left\{s\left[(1+c)^{s-1}+(1-c)^{s-1}\right]-\frac{1}{c}\left[(1+c)^s-(1-c)^s\right]\right\} \tag{C.7}$$

where $c=\frac{\sin\theta}{\sqrt{2}}$. By expanding in Taylor's series, the first equation is clearly negative for $s\in(0,1]$, whereas the second equation is positive by 3.2.1, hence the maximum is obtained at $\phi=0$. The last step is to evaluate

$$\frac{\partial M(s,1,0,\theta)}{\partial\theta} = \frac{1+s}{3}\cdot\frac{1}{2^{1+s}}\sin\theta\cdot\{\cos\phi[(1+\sin\phi\sin\theta)^s-(1-\sin\phi\sin\theta)^s] - \sin\phi[(1+r\cos\phi\sin\theta)^s-(1-r\cos\phi\sin\theta)^s]\}. \tag{C.8}$$

But then this is of similar form as (3.15), hence the maxima is obtained at $\theta=0$. Evaluating M(s,1,0,0) then gives us

$$P_{1+s}(X|\theta)\leq M(s,1,0,0)=\frac{1}{3}\cdot(1+2^{1-s}). \tag{C.9}$$

$$\mathrm{H}_\alpha(X|\Theta)\leq\frac{-1}{s}P_\alpha(X|\Theta) \tag{C.10}$$

taking the limit $\alpha\to 1$ gives us $\frac{2}{3}$, which is exactly the Shannon entropy.

□

The additivity of minimal entropy holds, by using the same argument in Step 2 of Appendix A. Given a string divided into parts A and B, where B is a single qubit, note that the uncertainty relation for B holds for the state

$$\sigma_B = \mathrm{tr}_A \left[\frac{N_{x_A|\theta_A} \ \rho_{AB} \ N^\dagger_{x_A|\theta_A}}{\mathrm{tr}(N_{x_A|\theta_A} \ \rho_{AB} \ N^\dagger_{x_A|\theta_A})} \right] \tag{C.11}$$

, for $N_{x_A|\theta_A} = \mathbf{T}^{\theta_A}|x_A\rangle\langle x_A|\mathbf{T}^{\theta_A} \otimes \mathsf{id}_B$, where

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

is the operator that cyclically permutes the eigenbases of σ_x, σ_y and σ_z.

By exactly the same arguments in Appendix A, the min-entropy of a string $X^n \in \{0,1\}^n$ conditioned on the basis $\theta^n \in \{0,2\}^n$can then be bounded by using [36]

$$\frac{1}{n} \ \mathrm{H}^\varepsilon_{\min}(X^n|\Theta^n K)_\rho \geq \ \frac{1}{n}\mathrm{H}^\varepsilon_{\min}(X^n|\Theta^n K)_{\rho|\rho} \geq \max_{s\in(0,1]} \ \frac{-1}{s} \log \left[\frac{1}{3}(1 + 2^{1-s}) \right]. \tag{C.12}$$