Techniques for Improving Numerical Secret Key Rates of Quantum Key Distribution Protocols

by

Scott Johnstun

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics - Quantum Information

Waterloo, Ontario, Canada, 2023

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The block diagonal solver of Section 4.1 is fully my own work. The idea for and theory behind entrywise constraints, as explained in Section 4.2, was developed by Lars Kamin, with whom I worked to implement the constraints in the framework of numerical QKD. Similarly, the optimization of ε-security parameters in Section 4.3 was developed by Lars Kamin. We worked together to implement this optimization into our numerical QKD framework. Finally, the key throughput maximization techniques in Section 4.4 are fully my own work, though the examples presented utilize data produced by Jasminder Sidhu at the University of Strathclyde.

No parts of this manuscript were written for publication elsewhere.

Abstract

Quantum Key Distribution (QKD) is a method for producing secure keys between two parties whose security does not rely on computational assumptions potentially breakable by quantum computers. However, physical constraints, such as noise, imperfect devices, and the necessity of finite resources, limit the rate at which experimental implementations of QKD can produce key, and in some cases prevent the generation of secure key altogether. Determination of key generation rate is facilitated by a numerical framework for general QKD protocols [1], upon which we propose improvements. With protocols used in actual QKD experiments as examples, we present and demonstrate various methods for improving key rate calculations in the regime of a finite number of signals sent. Our methods include a block diagonal optimization for the state shared by the two parties, modifying constraints on acceptance of candidate states, optimizing security parameter distribution, and optimizing the grouping of data into blocks for time-binned data. Through these improvements, we are able to both reduce the computational cost of key rate calculations in our numerical framework and improve key rates in the case of a finite number of sent signals.

Acknowledgements

I am very grateful to my supervisor and advisor, Norbert Lütkenhaus, for supporting me throughout my time studying under him. I am thankful for the opportunity he provided to me to work in his group and for his wisdom, his levelheadedness, and his always-respectful demeanor.

I also want to express thanks to my close colleagues and collaborators in John Burniston and Lars Kamin, who have helped me to understand QKD and been willing to field even the stupidest of questions that I've come up with. Special thanks goes to John Burniston for his contributions and collaboration in developing our group's numerical key rate toolbox OpenQKDSecurity. This thanks also extends to other current and former members of the Optical Quantum Communication Theory group at University of Waterloo, including Florian Kanitschar, Devashish Tupkary, Shlok Nahar, Amir Arqand, Ernest Tan, Ireneo Membrere, Shihong Pan, and Wenyuan Wang, with or from whom I have learned much of the foundation of quantum information theory and semidefinite programming.

Lastly, and most importantly, I am infinitely grateful for my parents, Clay and Lara, who have loved and supported me in all my endeavors, and my wife, Hannah, who has always believed in me and made sacrifices to help me follow my dreams.

Dedication

This work is dedicated to the brightest lights I know:

Rachel, Ariel, Mara, and Diana.

(a stark contrast to the dim lasers of QKD)

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Modern computer security is a topic of ever-increasing importance in the information age. The field of cryptography focuses on techniques for encrypting and decrypting information such that two parties can communicate in secret. Cryptographic algorithms currently in use rely on computational security, which assumes that certain problems are inherently sufficiently difficult to solve that an attacker performing an optimal attack would require an infeasible amount of time to break the encryption. However, in 1994, it was shown that this assumption may not be true—at least for certain computationally difficult problems—when a quantum algorithm for solving the problems of discrete logarithms and integer factoring was discovered [2]. Although modern quantum computers have too few qubits and experience too much noise to run this algorithm and break cryptosystems of the size used in secure digital communication, a complete breach of RSA-2048, a standard cryptographic algorithm, within 10 years is likely [3].

Quantum key distribution (QKD) is the field of research dedicated to establishing secret communication between two parties through the use of quantum systems. QKD offers provable security that is not reliant on obfuscation or computational complexity, as it exploits properties of quantum mechanics that allow for the detection of an eavesdropper. It also only relies on the assumption that the eavesdropper must obey the laws of physics, which is an improvement over computational security as it allows for security even in the presence of an eavesdropper with arbitrary computational power. It requires the use of a quantum channel, over which quantum state are exchanged, and a classical channel that is authenticated, meaning the identity of the sender and their message may not be altered. Another advantage of QKD is that the keys generated enjoy time-independent security. Secure communications protected by a classical cryptosystem today can be stored by an adversary for years, during which the adversary can attempt various attacks. Given enough

time, the constantly increasing availability of computational power, algorithm development, and perhaps some luck, the adversary can break the encryption and decrypt the stored communication. On the other hand, once a key has been generated through a QKD protocol, no attack can be made on the key can be attempted—until it's used, the security of a key does not change.

QKD has been realized experimentally using a variety of protocols, at large distances, and in the presence of various imperfections and noise sources. Examples include free-space communication using the Ekert protocol over 144km between two islands [4]; communication over 148km of optical fiber with the BB84 protocol [5]; communication between a ground station and an aircraft using a six-four protocol [6]; and intercontinental communication facilitated by a satellite in orbit using a decoy-state BB84 protocol [7]. However, in these experiments, rates of bit transfer are low, and, for example, in [4], a provably secure framework for QKD was not used, so the generated key may be insecure.

It is generally true that as we increase the requirements on the security of the key (i.e., reduce the probability of the protocol being executed despite the key being insecure), imperfections from experimental challenges, such as imperfect devices, noisy quantum channels, and statistical fluctuations have the effect of reducing key rate, sometimes eliminating the possibility of any key generation at all. QKD theory, which develops protocol security, is simplest in the case of ideal devices, so extension to imperfect devices necessarily introduces complications. With imperfect devices, we face the problem of implementation security, where improperly characterized devices can spoil the security of a valid QKD protocol. Decoy state analysis [8, 9, 10] is one extension to QKD theory that has eliminated a need for single photon sources, as it makes secure communication in QKD using simple laser sources very viable, offering better loss scaling than a simple weak coherent pulse protocol. However, there are still strict requirements on devices used for QKD; in particular, we need reliable and fast single-photon detectors, which is an area of continuing research and development that needs more work before reliable quantum networks can become a reality [11].

In this work, we intend to describe challenges that arise in the interface between experimental results and computing key rates under the proper framework of a security proof. This is an important topic because loss and error rate allowances can vary greatly between QKD protocols. We will provide several methods of increasing key rates in the finite-size regime, which is the relevant regime in experimental QKD. We focus on two particular protocols, the four-six protocol and the three-state protocol, that have been used in QKD experiments and show how our improvements apply to key rate calculations for those protocols. In doing so, we hope to provide a general pathway for the key rate calculator to maximize the key rate obtained from the results of QKD experiment while staying within

a proper security framework.

In Chapter 2, we provide the theoretical background to build up quantum key distribution. Namely, we first build up from the foundation of linear algebra and use it as the backbone for the mathematics of quantum information. We then describe quantum key distribution, explaining the security proof and applying it to a simple BB84 protocol. Finally, we give a brief introduction to quantum optics and describe important states and devices for use in QKD.

Chapter 3 moves from generality to specificity, introducing decoy state analysis, a technique we rely on heavily in the protocols we study. We then give mathematical descriptions of the four-six and three-state protocols as well as possible experimental setups. We close the chapter by describing the numerical framework we use for key rate calculations in QKD, which forms the base for where our optimizations will come into play.

In Chapter 4, we delve into our techniques for improving key rate calculations. We first introduce the block diagonal solver and explain how it can reduce space and time complexity of the semidefinite program that facilitates key rate calculation. We then show how choosing entrywise constraints can improve key rate over 1-norm constraints that have been used in previous works, including examples demonstrating the improvement. Following that, we then present results relating to intelligently distributing sub-security parameters for the key rate calculation. We show that introducing a bias in the distribution can improve key rate and calculate an optimal distribution. Finally, we consider the case of time-binned data from simulated satellite data. We describe the tradeoff between time bin block length and channel similarity, then we present several different options for maximizing the total amount of key bit throughput given an example quantum channel that varies over time.

Finally, in Chapter 5, we give summarize the work, give concluding remarks, and suggest ideas for future research.

Chapter 2

Background

In this chapter, we introduce the concepts that are important for understanding quantum information. We build up quantum key distribution and the security framework behind it. Finally, we provide an introduction to the field of quantum optics and describe important states and devices that are used in optical quantum key distribution.

2.1 Quantum Information Theory

To begin, we will discuss the mathematics that build up to quantum mechanics, focusing on those topics most relevant to QKD. Much of this discussion follows definitions in common textbooks on quantum information [12, 13] and mathematical physics [14].

2.1.1 Complex Euclidean Spaces

The language of quantum information theory is linear algebra, so we first establish necessary definitions. A *vector* is a collection of complex numbers indexed by an *alphabet* Σ, which is a finite, nonempty set. For a member $j \in \Sigma$, we can denote the complex number at index j in the vector v by v_j.

Definition 2.1.1 (Vector Space)**.** A vector space $\mathcal{V}$ is a set of vectors with the following properties:

1. For any vectors $u, v \in \mathcal{V}$, the vector $u+v$, defined by $(u+v)_j = u_j + v_j$, is an element of $\mathcal{V}$.

2. For any scalar $\alpha \in \mathbb{C}$ and vector $u \in \mathcal{V}$, the vector αu defined by $(\alpha u)_j = \alpha u_j$, is an element of $\mathcal{V}$

We can denote the vector space of all vectors indexed by Σ with $\mathbb{C}^\Sigma$. This kind of vector space is known as a complex Euclidean space.

If a vector $u \in \mathbb{R}^\Sigma$ has entries that satisfy $0 \leq u_i \leq 1$ for all $i \in \Sigma$ and $\sum_i u_i = 1$, we call u a probability distribution. The set of probability distributions on Σ is denoted $\mathcal{P}(\Sigma)$.

We can augment this space with an inner product, which provides a measure of similarity between vectors in the space.

Definition 2.1.2 (Inner Product)**.** Let $\mathbb{C}^\Sigma$ be a complex Euclidean space. The inner product $\langle u, v \rangle$ of two vectors $u, v \in \mathbb{C}^\Sigma$ is defined as

$$\langle u, v \rangle = \sum_{j \in \Sigma} u_j^* v_j$$

An inner product satisfies the following properties:

1. Linearity in the second argument:

   For all $u, v, w \in \mathbb{C}^\Sigma$ and $\alpha, \beta \in \mathbb{C}$,

   $$\langle u, \alpha v + \beta w \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle .$$

2. Conjugate symmetry:

   For all $u, v \in \mathbb{C}^\Sigma$,

   $$\langle u, v \rangle = \langle v, u \rangle^* .$$

3. Positive definiteness:

   For all $u \in \mathbb{C}^\Sigma$,

   $$\langle u, u \rangle \geq 0,$$

   with $\langle u, u \rangle = 0$ if and only if $u = 0$.

We assign an inner product space the norm $||u|| = \sqrt{\langle u|u \rangle}$ and the Euclidean distance metric $d(u, v) = ||u - v||$.

Definition 2.1.3 (Inner Product Space)**.** An inner product space is a complex Euclidean space equipped with an inner product.

Complex Euclidean spaces can be combined using a tensor product. The *tensor product* of a collection of complex Euclidean spaces $\mathbb{C}^{\Sigma_1}, \cdots, \mathbb{C}^{\Sigma_n}$ is the complex Euclidean space

$$\mathbb{C}^{\Sigma_1} \otimes \cdots \otimes \mathbb{C}^{\Sigma_n} = \mathbb{C}^{\Sigma_1 \times \cdots \times \Sigma_n},$$

where $\times$ is the Cartesian product acting on sets. We can also consider the tensor product of vectors, defined for vectors $u^{(1)} \in \mathbb{C}^{\Sigma_1}, \cdots, u^{(n)} \in \mathbb{C}^{\Sigma_n}$ using

$$(u^{(1)} \otimes \cdots \otimes u^{(n)})_{j_1, \cdots, j_n} = u^{(1)}_{j_1} \cdots u^{(n)}_{j_n}$$

A related operation is the direct sum of two spaces. Given a collection of complex Euclidean spaces $\mathbb{C}^{\Sigma_1}, \cdots, \mathbb{C}^{\Sigma_n}$, their *direct sum* is the complex Euclidean space

$$\mathbb{C}^{\Sigma_1} \oplus \cdots \oplus \mathbb{C}^{\Sigma_n} = \mathbb{C}^{\Omega},$$

where $\Omega = \Sigma_1 \cup \cdots \cup \Sigma_n$.

2.1.2 Linear Operators

Given complex Euclidean spaces $\mathcal{U} = \mathbb{C}^{\Sigma_1}$ and $\mathcal{V} = \mathbb{C}^{\Sigma_2}$, we can consider mappings that transform vectors in $\mathcal{U}$ into vectors in $\mathcal{V}$. We denote the set of such mappings with $\mathcal{L}(\mathcal{U}, \mathcal{V})$. When the input and output spaces are both $\mathcal{U}$, we simply write $\mathcal{L}(\mathcal{U})$ to denote $\mathcal{L}(\mathcal{U},\mathcal{U})$. These mappings are referred to as *operators*. This set forms a vector space with the following operations:

1. For operators $X, Y \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, we define the operator $X + Y \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ by

   $$(X + Y)u = Xu + Yu$$

   for all $u \in \mathcal{U}$.

2. For an operator $X \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ and a scalar $\alpha \in \mathbb{C}$, we define the operator αX by

   $$(\alpha X)u = \alpha(Xu)$$

   for all $u \in \mathcal{U}$.

Each operator has a corresponding matrix, which is a mapping of the form

$$M : \Sigma_1 \times \Sigma_2 \to \mathbb{C}$$

for alphabets Σ_1 and Σ_2. For $a \in \Sigma_1$ and $b \in \Sigma_2$, we write $M_{a,b}$ to refer to the entry $M(a, b)$.

In the case where $\mathcal{V} = \mathbb{C}$, we have an operator known as a linear functional, which is a mapping from a higher-dimensional complex Euclidean space to a one-dimensional complex Euclidean space, or scalar. The set of linear functionals $\mathcal{L}(\mathcal{U}, \mathbb{C})$ is denoted $\mathcal{U}^*$ and is known as the *dual space* of $\mathcal{U}$. Elements in this space correspond to row vectors, which are matrices with a single row.

The *identity operator* I of a space $\mathcal{X}$ maps all vectors in $\mathcal{X}$ to themself, i.e. $Ix = x$ for all $x \in \mathcal{X}$. If the context does not make it clear which space the identity operator corresponds to, we write $I_{\mathcal{X}}$ with a subscript to denote the space.

The *standard basis* in a complex Euclidean space $\mathbb{C}^\Sigma$ is the orthonormal basis given by $\{e^{(a)} : a \in \Sigma\}$, where $e_b^{(a)} = 1$ if $a = b$ and $e_b^{(a)} = 0$ otherwise.

Definition 2.1.4 (Trace)**.** The *trace* of an operator $X \in \mathcal{L}(\mathbb{C}^\Sigma)$ is the sum of the diagonal entries of its corresponding matrix

$$\mathrm{Tr}\,(X) = \sum_{a \in \Sigma} X_{a,a} = \sum_{a \in \Sigma} \langle e^{(a)}, X e^{(a)} \rangle\,.$$

Definition 2.1.5 (Adjoint)**.** For each operator $X \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ we define an *adjoint* operator $X^\dagger \in \mathcal{L}(\mathcal{V}, \mathcal{U})$, which is uniquely defined by the equation

$$\langle v, Xu \rangle = \langle X^\dagger v, u \rangle$$

for all $u \in \mathcal{U}$ and $v \in \mathcal{V}$. Note that $(X^\dagger)^\dagger = X$.

Similar to vectors, it can be useful to define norms for operators. Though there are many ways to define norms for operators, we will only define the trace norm in this work, as we will not need any other norms.

Definition 2.1.6 (Trace Norm)**.** For an operator $X \in \mathcal{L}(\mathcal{U})$, the *trace norm* of X is defined as

$$\|X\|_1 = \mathrm{Tr}\,(|A|)\,.$$

It is useful to define several types of linear operators that are useful for the purposes of quantum information. All of the following definitions apply to a square operator $X \in \mathcal{L}(\mathcal{U})$.

Definition 2.1.7 (Classification of Linear Operators)**.**

Normal operators commute with their adjoint. That is, X is normal if $X^\dagger X = XX^\dagger$.

Hermitian operators are self-adjoint. That is, X is Hermitian if $X = X^\dagger$. All Hermitian operators are normal.

Positive semidefinite operators satisfy $\langle u, Xu\rangle \geq 0$ for all $u \in \mathcal{U}$. If $\langle u, Xu\rangle > 0$ for all $u \in \mathcal{U}$, we say X is **positive definite**. The set of positive semidefinite operators on a space $\mathcal{U}$ is denoted by $\mathrm{Pos}(\mathcal{U})$.

Unitary operators satisfy $X^{-1} = X^\dagger$, or, equivalently, $XX^\dagger = X^\dagger X = I$.

Density operators are positive semidefinite operators with unit trace. That is, X is a density operator if $\langle u, Xu\rangle \geq 0$ for all $u \in \mathcal{U}$ and $\mathrm{Tr}\,(X) = 1$. We denote the set of all such operators with $\mathcal{D}(\mathcal{U})$.

Projection operators project vectors onto smaller subspaces of a larger vector space. They are defined by the equation $X^2 = X$.

2.1.3 Quantum Mechanics

Quantum mechanics gives a framework in which we can describe the interaction and evolution of quantum states and operations.

Quantum States

Isolated physical systems have an associated complex Euclidean space known as the *state space* of the system. We will denote vectors in complex Euclidean spaces associated with quantum systems using a ket, such as $|\psi\rangle$. Functionals belonging to the corresponding dual space will be denoted with a bra: $\langle\psi|$. We can represent an inner product between two vectors with a braket $\langle\phi|\psi\rangle$.

Definition 2.1.8 (Quantum state)**.** A *quantum state* is a density operator $\rho \in \mathcal{D}(\mathcal{X})$, where $\mathcal{X}$ is the state space of the system.

States can be mixed or pure. Pure states satisfy $\mathrm{Tr}\,(\rho^2) = 1$ and are rank 1 operators. A state $\rho \in \mathcal{D}(\mathcal{X})$ is pure if and only if there is a vector $|u\rangle \in \mathcal{X}$ such that

$$\rho = |u\rangle\langle u|$$

Mixed states are probabilistic mixtures of pure states, and they take the form

$$\rho = \sum_{a \in \Gamma} p(a)\rho_a, \tag{2.1}$$

for an alphabet Γ, where $p(a) \in \mathcal{P}(\Gamma)$ is a probability vector and each ρ_a is a density matrix. Mixed states satisfy $\mathrm{Tr}\,(\rho^2) < 1$.

Definition 2.1.9 (Classical state)**.** A quantum state $\rho \in \mathcal{D}(\mathbb{C}^\Sigma)$ is *classical* in a given basis if its matrix representation is diagonal in that basis. In this case, the diagonal entries of ρ form a probability vector in $\mathbb{C}^\Sigma$.

It is also important to consider the composite state of multiple systems, known as multipartite systems. This can be established through the use of tensor products.

Definition 2.1.10. A state ρ in a space $\mathcal{D}(\mathcal{U})$, where $\mathcal{U} = \mathcal{U}_1 \otimes \cdots \otimes \mathcal{U}_n$, is a *product state* if it takes the form

$$\rho = \sigma_1 \otimes \cdots \otimes \sigma_n$$

for $\sigma_1 \in \mathcal{D}(\mathcal{U}_1), \cdots, \sigma_n \in \mathcal{D}(\mathcal{U}_1)$.

If ρ is a product state, then the subsystems that make up ρ are independent. Multipartite systems that are not product states are said to be entangled.

This leads us to the definition of partial trace, which allows for subsystems of a multipartite system to be isolated.

Definition 2.1.11 (Partial trace)**.** Given a bipartite density operator $\rho_{AB} \in \mathcal{D}(\mathcal{U})$, where $\mathcal{U} = A \otimes B$ for state spaces A and B, the *partial trace* of ρ_{AB} over system B is the density operator obtained by computing the trace of the portion of the state in the space B and considering the remainder of the state:

$$\mathrm{Tr}_B(\rho) = \sum_{b \in \Sigma_B} (I_A \otimes \langle b|)\rho_{AB}(I_A \otimes |b\rangle).$$

Here Σ_B denotes the alphabet of the state space B, and $|b\rangle$ is shorthand for the standard basis vector $|e^{(b)}\rangle$.

This definition extends simply to the partial trace of a multipartite space over any number of subsystems.

At times, it is useful to consider a state to be a subsystem of a larger system. This is important because a subsystem in a mixed state can be a component of a supersystem in an entangled state. Such a situation can be realized mathematically with a purification.

Definition 2.1.12 (Purification)**.** Let $\mathcal{U}$ and $\mathcal{V}$ be complex Euclidean spaces, and let $\rho \in \mathcal{D}(\mathcal{U})$ be a density operator. Let $|x\rangle \in \mathcal{U} \otimes \mathcal{V}$ be a vector. The vector $|x\rangle$ is a *purification* of ρ if

$$\mathrm{Tr}_{\mathcal{V}} \left(|x\rangle \langle x| \right) = \rho. \tag{2.2}$$

Quantum Channels

Quantum states can be transformed through the action of quantum channels, which are maps from operators to operators.

Definition 2.1.13 (Quantum channel)**.** Given complex Euclidean spaces $\mathcal{U}$ and $\mathcal{V}$, a *quantum channel* is a linear map $\Phi : \mathcal{L}(\mathcal{U}) \to \mathcal{L}(\mathcal{V})$ that satisfies the following properties:

1. Φ is *completely positive*, meaning the map $\Phi \otimes I_{\mathcal{W}}$ maps positive semidefinite operators in $\mathcal{U} \otimes \mathcal{W}$ to positive semidefinite operators in $\mathcal{V} \otimes \mathcal{W}$, for any complex Euclidean space $\mathcal{W}$.
2. Φ is trace preserving, meaning $\mathrm{Tr}\left(\Phi(X)\right) = \mathrm{Tr}\left(X\right)$ for all $X \in \mathcal{L}(\mathcal{U})$.

Each quantum channel $\Phi : \mathcal{L}(\mathcal{U}) \to \mathcal{L}(\mathcal{V})$ has an adjoint channel $\Phi^\dagger : \mathcal{L}(\mathcal{V}) \to \mathcal{L}(\mathcal{U})$ that satisfies

$$\langle \Phi^\dagger(Y), X \rangle = \langle Y, \Phi(X) \rangle$$

for all $X \in \mathcal{L}(\mathcal{U})$ and $Y \in \mathcal{L}(\mathcal{V})$.

We denote the set of quantum channels from $\mathcal{U}$ to $\mathcal{V}$ with $\mathcal{C}(\mathcal{U}, \mathcal{V})$. A convenient way to represent quantum channels mathematically is through the use of Kraus operators, which we define now.

Definition 2.1.14 (Kraus operators)**.** For every quantum channel $\Phi \in \mathcal{C}(\mathcal{U}, \mathcal{V})$, there exist an alphabet Σ and a set of operators $\{K_a : a \in \Sigma\}$, which is a subset of $\mathcal{L}(\mathcal{U}, \mathcal{V})$, such that

$$\Phi(X) = \sum_{a \in \Sigma} K_a X K_a^\dagger$$

for all $X \in \mathcal{L}(\mathcal{U})$. These operators are known as *Kraus operators* for Φ. Kraus operators are not unique.

A channel represented with Kraus operators has an adjoint that is simple to compute via

$$\Phi^\dagger(Y) = \sum_{a\in\Sigma} K_a^\dagger Y K_a,$$

which holds for all $Y \in \mathcal{L}(\mathcal{V})$.

Measurement

Measurements provide us with a means to obtain classical information about a quantum state. Though they are not the only form of measurement, we focus on Positive Operator-Valued Measures (POVMs), as they are the most general form of measurement and the only measurement we are concerned with in this work.

Definition 2.1.15 (Positive operator-valued measure (POVM))**.** Given a complex Euclidean space $\mathcal{U}$ and an alphabet Σ (note $\dim(\mathcal{U})$ is not necessarily equal to $|\Sigma|!$), a *POVM* is a set of matrices $\{\Gamma_i\}_{i\in\Sigma} \subset \mathrm{Pos}(\mathcal{U})$ such that

$$\sum_{i\in\Sigma} \Gamma_i = I.$$

The probability of obtaining outcome i when measuring a state $\rho \in \mathcal{D}(\mathcal{U})$ is given by

$$p(i) = \mathrm{Tr}\left(\Gamma_i \rho\right).$$

POVMs are useful because the requirements of positivity and summing to identity guarantee that the set $\{p(i)\}_{i\in\Sigma}$ is a probability distribution, as $\sum_{i\in\Sigma} p(i) = 1$ and, for all $i \in \Sigma$, $0 \leq p(i) \leq 1$.

A POVM $\{\Gamma_i\}_{i\in\Sigma}$ can be written in the form of a quantum-to-classical channel acting on $\rho \in \mathcal{D}(\mathcal{U})$ and producing a state $\sigma \in \mathcal{D}(\mathbb{C}^\Sigma)$. Since each $\Gamma_i \in \mathrm{Pos}(\mathcal{U})$, it has an operator square root $E_i = \sqrt{\Gamma_i}$ such that $E_i^2 = \Gamma_i$ and $E_i \in \mathrm{Pos}(\mathcal{U})$. Then, defining the Kraus operators

$$K_{i,j} = |i\rangle \langle j| E_i,$$

for $i \in \Sigma$ and $j \in \{1, \cdots, \dim(\mathcal{U})\}$, it holds that

$$\sum_{i\in\Sigma} \sum_{j=1}^{\dim(\mathcal{U})} K_{i,j} \rho K_{i,j}^\dagger = \sum_{i\in\Sigma} \sum_{j=1}^{\dim(\mathcal{U})} |i\rangle \langle j| E_i \rho E_i |j\rangle \langle i| = \sum_{i\in\Sigma} \mathrm{Tr}\left(E_i \rho E_i\right) |i\rangle \langle i| = \sum_{i\in\Sigma} \mathrm{Tr}\left(\Gamma_i \rho\right) |i\rangle \langle i|.$$

The output of this channel is a classical state; as a density matrix, it is a diagonal matrix with the value $p(i)$ on the ith diagonal.

Entropies

In information theory, an entropy is a quantity that gives a sense of the amount of randomness or uncertainty in a system. We will find entropies particularly useful for finding key rate in quantum key distribution protocols, as it is important to know how much information an adversary could glean from the execution of a protocol.

It is insightful to first define classical entropies and then connect those to the corresponding quantum entropies.

Definition 2.1.16 (Shannon Entropy)**.** Let Σ be an alphabet and let p be a probability distribution on Σ. The *Shannon entropy* of p is

$$H(p) = -\sum_{i \in \Sigma} p_i \log_2(p_i),$$

where, in the case that $p_i = 0$, we make use of the fact that $\lim_{x \to 0} x \log x = 0$.

The Shannon entropy $H(u)$ quantifies the amount of randomness, measured in bits, of a probability distribution. As an example, consider the probability vectors defined on $\Sigma = \{0, 1\}$

$$p_0 = 1, \;\; p_1 = 0$$

and

$$q_0 = \frac{1}{2}, \;\; q_1 = \frac{1}{2}.$$

Sampling from the probability distribution p gives us the result 0 with no uncertainty, so correspondingly, $H(p) = 0$. On the other hand, sampling from the probability distribution q gives an equal chance of obtaining the outcome 0 or 1, which is maximally uncertain. In this case, there is one bit of uncertainty, and thus $H(q) = 1$.

In quantum information theory, we define entropies on density operators instead of probability vectors. The von Neumann entropy, which extends the concept of classical Shannon entropy to semidefinite operators, is a quantitative measurement of the expected amount of randomness in a quantum state.

Definition 2.1.17 (von Neumann Entropy)**.** For a positive semidefinite operator σ in a complex Euclidean space $\mathcal{U}$, the *von Neumann entropy* of σ is defined as

$$\mathrm{H}(\sigma) = -\mathrm{Tr}\left(\sigma \log(\sigma)\right).$$

It is also useful to measure the dissimilarity of two probability distributions. For this purpose, we define the relative entropy, which measures the information content or "surprise" (defined for an event with probability p_i as $-\log_2(p_i)$) of using a model probability distribution q when the true distribution is p.

Definition 2.1.18 (Relative Entropy)**.** Let Σ be an alphabet and let p and q be probability distributions on Σ. The *relative entropy* of p with respect to the reference distribution q is defined as

$$\mathrm{D}\left(p||q\right) = \sum_{i \in \Sigma} p_i \log_2 \left(\frac{p_i}{q_i}\right).$$

In the case that $q_i = 0$ for some $i \in \Sigma$, we use the fact that $0 \log_2(0/0) \to 0$ if $p_i = 0$; otherwise, we define $\mathrm{D}\left(p||q\right) = \infty$, as there are variables in p that q does not predict.

Extending this to density operators gives us a quantitative measure of the dissimilarity of two quantum states.

Definition 2.1.19 (Quantum Relative Entropy)**.** For two positive semidefinite operators $\rho, \sigma \in \mathrm{Pos}(\mathcal{U})$, the *quantum relative entropy* of ρ with respect to σ is defined to be

$$\mathrm{D}\left(\rho||\sigma\right) = \mathrm{Tr}\left(\rho \log(\rho)\right) - \mathrm{Tr}\left(\rho \log(\sigma)\right).$$

The final entropic quantity we need is the conditional entropy. In classical information theory, this quantifies the amount of uncertainty about a random variable P given knowledge about the state of another random variable Q.

Definition 2.1.20 (Conditional Entropy)**.** For classical variables P and Q with corresponding probability distributions p and q, respectively, the *conditional entropy* is defined as

$$H(P|Q) = H(P, Q) - H(Q).$$

In this definition, we have made use of the convention that the Shannon entropy of a variable P is the Shannon entropy of the probability vector corresponding to its state: $H(P) = H(p)$. Thus, $H(P, Q)$ denotes the Shannon entropy of the joint probability distribution between variables P and Q.

The corresponding quantity in quantum information theory is the quantum conditional entropy, which gives a sense of the expected uncertainty in a bipartite state in the case that we have information about one part of the state. This can also be interpreted as a measure of how mixed the state is.

Definition 2.1.21 (Quantum Conditional Entropy)**.** For a bipartite quantum state $\rho_{AB} \in \mathrm{Pos}(\mathcal{U})$, which is made up of two subsystems labeled A and B, the *quantum conditional entropy* of ρ given σ is

$$\mathrm{H}(A|B) = \mathrm{H}(\rho_{AB}) - \mathrm{H}(\rho_B),$$

where $\rho_B = \mathrm{Tr}_A(\rho_{AB})$. Note that it is common to write this entropy in terms of the systems, in which case we omit the subscript on the state as follows:

$$\mathrm{H}(A|B)_\rho = \mathrm{H}(A,B)_\rho - \mathrm{H}(B)_\rho.$$

Note that the state is included as a subscript for clarity.

2.2 Quantum Key Distribution

Quantum key distribution (QKD) aims to allow for two parties to establish a shared secret key. These parties, typically called Alice and Bob, desire to generate a key that is fully unknown to an eavesdropper, Eve. The quantum nature of QKD allows for Alice and Bob to detect the interference of such an eavesdropper, as, unlike classical bits, quantum bits (and quantum systems in general) cannot be duplicated. In order to eavesdrop on the communication between Alice and Bob, Eve must interact with the quantum states exchanged between them; however, since interactions with quantum systems necessarily affect their state, Alice and Bob can use their own measurement results to detect Eve with high probability and abort communication if necessary.

QKD protocols can be classified as entanglement-based or prepare-and-measure. *Entanglement based* protocols have a third party prepare an entangled state and send half of the state each to Alice and Bob. *Prepare and measure* protocols involve Alice preparing a state at random, taking note of which state she prepared, and then sending the state to Bob. These classifications are ultimately equivalent by the so-called *source replacement scheme*, which will be described later. However, our focus in this work is on prepare-and-measure protocols.

It is important to consider the resource requirements for QKD. The simplest requirement is that Alice and Bob's systems exist in environments that are private to them; any violation of this is an obvious source of information leakage. Alice needs a quantum channel for sending states to Bob, who does not typically need to send quantum states back to Alice, meaning this channel may be one-way. Alice and Bob also need a shared, two-way, authenticated classical channel. This channel is allowed to be fully accessible to Eve, but if Eve is allowed to impersonate Alice or Bob, the protocol's security is compromised

(hence the authentication requirement). These first few requirements are often touted as the only requirements for QKD; however, any realistic implementation of QKD requires that Alice's signal sources and Bob's detectors work well. Specifically, we require that they are trusted, meaning they cannot be modified by an action of Eve, and they must be properly characterized, which means they behave according to their quantum mechanical description. Any incongruency between a device's quantum mechanical description and its operation is a source of insecurity for the QKD protocol, either in the form of allowing Eve to acquire more information or in the form of Alice and Bob simply being unable to determine a secure shared state.

We note here that device-independent QKD relaxes the assumption of trusted devices, allowing for provable security even in the case that their devices were created by Eve and do not behave as Alice and Bob expect [15, 16, 17]. This is accomplished by generating entangled pairs and using the devices to test Bell inequalities on them, which allows Alice and Bob to place a bound on the amount of information Eve can obtain from their communication. Device-independent QKD has been realized experimentally, but the requirements of generating high-fidelity entangled states and nearly perfect random quantum measurements provides a significant experimental challenge yet unrealized [18]. As a result, the protocols we are concerned with in this work are not device-independent.

A typical prepare-and-measure QKD protocol has the following structure:

1. **Signal Preparation** Alice sends a random sequence of signal states to Bob, taking note of what signal she sent for each state. These states are partitioned into sets typically defined by a basis choice. The states pass through a quantum channel and arrive at Bob's setup.

2. **Measurement** Bob performs a POVM measurement on Alice's signal in one of two bases, chosen randomly. He takes note of the basis used for each measurement as well as the measurement result, which is a specific signal state.

3. **Acceptance Test** Alice and Bob choose a subset of the exchanged signals and reveal all preparation and measurement information on this subset. From this information, which comes in the form of a joint probability distribution, they can identify if the signal states were altered during transmission.

4. **Public Discussion** For each signal that was sent, Alice announces functions of her local measurements corresponding to the aforementioned partition. Bob does the same based on his measurement result. Alice and Bob sift out signals based on their discussion.

5. **Key Map** Within each basis choice, Alice assigns a bit value to signals that survived sifting, determined as a function of her state sent and the public discussion made in the previous step. This function is known as the key map, and the result is a raw key.

6. **Error Correction** Alice sends error correction information over the public communication channel to Bob, who uses the information to correct errors in his bit string. This necessarily leaks some information to Eve about the raw key bit strings Alice and Bob hold.

7. **Privacy Amplification** Alice and Bob each apply a random two-universal hash function to their bit strings to obtain a shared secret key. Privacy amplification reduces the likelihood of Eve determining the secret key from information she has gleaned through the execution of the protocol.

The result of a QKD protocol is a secure string of l bits that is totally uncorrelated with Eve and where all possible keys are equally likely. The protocol is allowed to abort, in which case no secret key is produced. In addition, the key produced is random, not a particular key chosen by either party; this means that Alice and Bob may freely utilize randomness in their protocol.

In the ideal case, the state shared by Alice, Bob, and Eve has the form

$$\rho_{ABE}^{(\text{ideal})} = \frac{1}{|S|}\left(\sum_{s\in S} |s\rangle\langle s|_A \otimes |s\rangle\langle s|_B\right) \otimes \rho_E, \tag{2.3}$$

where S is the set of all possible bit strings of length l and $\rho_E = \text{Tr}_{AB}(\rho_{ABE})$ is the portion of the exchanged state that Eve holds, including all classical information. This case is ideal because of the following properties:

1. Randomness: all possible keys $s \in S$ are equally likely, with probability $\frac{1}{|S|}$.

2. Correctness: Alice and Bob have the same key s

3. Secrecy: Alice and Bob's state is completely uncorrelated with Eve's state (see the remark after Definition 2.1.10)

2.2.1 Security Proof for QKD

We utilize the numerical framework developed in [19, 20, 1, 21]. This framework will allow us to compute the amount of secret key bits Alice and Bob can generate per signal exchanged, which is a quantity known as *key rate*.

It is important to be specific about what security means in QKD; to this end, we provide a definition of security in QKD [22].

Definition 2.2.1 (ε-security)**.** Suppose we perform a QKD protocol that acts on the state $\rho_{AA'} \in \mathcal{D}(\mathcal{A} \otimes \mathcal{A}')$, where $\mathcal{A}$ and $\mathcal{A}'$, and $\mathcal{E}$ are complex Euclidean spaces corresponding to Alice's local and flying systems respectively. Denote by $\mathcal{F}$ a map that performs the steps of acceptance testing through privacy amplification, and let $\mathcal{F}^{\text{ideal}}$ be a map that outputs the ideal state of Eqn. 2.3. This protocol is ε-secure if the output state satisfies

$$\frac{1}{2} \left\| \left((\mathcal{F} - \mathcal{F}^{\text{ideal}}) \otimes I_E \right) \left[(I_A \otimes \Phi) (\rho_{AA'}) \right] \right\|_1 \leq \varepsilon$$

for all quantum channels Φ.

Intuitively, this definition says that the actual outcome of the protocol and the ideal outcome of the protocol are only ε-distinguishable. For sufficiently small ε, the output of the protocol is nearly indistinguishable from the ideal case. This security definition tells us that the probability that the protocol produces a key (i.e., it does not abort) *and* the produced key is not random, correct, and secure is no larger than ε.

Note that in this section and throughout this work, we assume that Eve is limited to collective attacks, which represents those attacks where Eve performs the same attack on each signal. This stands in contrast to coherent attacks, wherein Eve can perform an attack on the full collection of states Alice sends to Bob. Work has been done in QKD to allow for protection against coherent attacks, such as the postselection technique [23]. The postselection technique lifts a protocol that is ε'-secure against a collective attack to a protocol that is ε-secure against general attacks with $\varepsilon = \varepsilon'(N+1)^{(d^2-1)}$, where N is the number of signals sent by Alice and d is the dimension of the quantum systems used for communication (for qubits, $d = 2$). Thus, in the case of a moderate 10^9 signals sent by Alice, a protocol that is ε-secure against general attacks with $\varepsilon = 10^{-3}$ (that is, a protocol that produces an insecure key 1 in 1000 times) is equivalent to a protocol that is secure against collective attacks with $\varepsilon' = 10^{-33}$. This is an extremely tight tolerance, but due to the logarithmic scaling of security parameters in key rate calculations, this technique is viable.

We now briefly summarize the procedure for obtaining the secure key rate of a QKD protocol. Let A represent Alice's system (which may be classical) that she keeps and A' represent the system that she sends to Bob. Let B denote the system Bob receives, and let E represent Eve's system, which may be high-dimensional. We consider a prepare-and-measure protocol. Denote with X and Y the systems that hold the outcomes of the measurements that Alice and Bob perform on systems A and B, respectively.

Let the set of states Alice uses for the protocol, indexed by the alphabet Σ, be denoted by $\{|\phi_i : i \in \Sigma\rangle\}$, and let $\{p_i\} \in \mathbb{R}^{\Sigma}$ be a probability distribution. In the source replacement scheme, Alice will prepare state $|\phi_i\rangle$ with probability p_i and send that state to Bob, keeping track of the index i that was prepared. Alice's state at preparation is then

$$|\psi\rangle_{AA'} = \sum_{i\in\Sigma} \sqrt{p_i}\, |i\rangle_A\, |\phi_i\rangle_{A'}\,. \tag{2.4}$$

Alice sends the system A' to Bob via a quantum channel $\mathcal{E}_{A'\to B}$. Their shared state after transmission is

$$\rho_{AB} = (I_A \otimes \mathcal{E}_{A'\to B})(|\psi\rangle\langle\psi|_{AA'}) \tag{2.5}$$

Note the equivalence to an entanglement-based scheme in the above lines.

Alice and Bob then perform measurements on the systems A and B to obtain measurement results, stored in systems X and Y. After the process of key mapping, Alice obtains a raw key Z.

In the asymptotic limit, our key rate, in bits per signal sent, is obtained from the Devetak-Winter key rate formula [24]:

$$R_\infty = p_{\text{pass}} \left(\min_{\rho\in\mathbf{S}} \mathrm{H}(Z|E)_\rho - \mathrm{H}(Z|Y)_\rho \right), \tag{2.6}$$

where p_{pass} denotes the probability that a signal survives sifting in the protocol. The set $\mathbf{S}$ denotes a set of density matrices that match the probability distribution obtained during the acceptance testing step of the protocol.

The last term in Eqn. 2.6 is the minimum amount of information leaked during error correction. Since error correction at the minimum is not always posible in practice, we define a total error correction leakage $\delta_{\text{leak}} = f\mathrm{H}(Z|Y)$, using $f \geq 1$ to represent the efficiency of error correction.

Eqn. 2.6 can be reformulated to be agnostic of Eve's state if we first define two quantum channels to describe the measurements, announcements, and key mapping that Alice and Bob apply to their states.

1. The map $\mathcal{G}$ is a completely positve, trace nonincreasing quantum channel that performs Alice and Bob's joint measurements and ties those measurements to the corresponding announcements. It performs sifting (eliminating the need for p_{pass} in the key rate formula) based on these announcements, then connects the measurements to the appropriate key bits via the key map. Mathematically,

$$\mathcal{G}(\rho) = \sum_{\alpha,\beta\in\{z,x\}} K_{\alpha,\beta}\rho K_{\alpha,\beta}^{\dagger}, \tag{2.7}$$

where

$$K_{\alpha,\beta} = \sum_{x\in\{0,1\}} |g(x,\alpha,\beta)\rangle_R \otimes \sqrt{\Gamma_{\alpha,x}^{A}} \otimes \sqrt{\Gamma_{\beta,x}^{B}} \otimes |\alpha,\beta\rangle_C \tag{2.8}$$

are the Kraus operators of $\mathcal{G}$. In these equations, the subscripts R and C denote the key and announcement systems respectively, which are classical; α and β represent the announcement that Alice and Bob make; x labels the key bit assigned in Alice's POVM; and $g(x,\alpha,\beta)$ is a key map that takes in Alice's key bit and her and Bob's announcements and outputs the corresponding key bit in the final secret key.

2. The map $\mathcal{Z}$ is a completely positive, trace preserving quantum channel that projects onto subspaces of the image of $\mathcal{G}$ corresponding to each key bit. It has the form

$$\mathcal{Z}(\mathcal{G}(\rho)) = \sum_{x\in\{0,1\}} \left(|x\rangle\langle x|_R \otimes I_{ABC}\right)\mathcal{G}(\rho)\left(|x\rangle\langle x|_R \otimes I_{ABC}\right). \tag{2.9}$$

We call this a key projection map because each term in the application of the map extracts the projection of the input ρ onto the corresponding key bit. Note that the input $\mathcal{G}(\rho)$ includes key bit and announcement registers R and C, which are outputs of $\mathcal{G}$.

We can then rewrite Eqn. 2.6 as

$$R_\infty = \min_{\rho_{AB}\in\mathbf{S}} \mathrm{D}\left(\mathcal{G}(\rho_{AB})||\mathcal{Z}\circ\mathcal{G}(\rho_{AB})\right) - p_{\text{pass}}\delta_{\text{leak}}. \tag{2.10}$$

Throughout this work, we will define the shorthand function $f(\rho) = \mathrm{D}\left(\mathcal{G}(\rho)||\mathcal{Z}\circ\mathcal{G}(\rho)\right)$.

Eqn. 2.10 gives us the asymptotic key rate for a QKD protocol as a function of a minimization over the subset $\mathbf{S}$ of all shared density matrices between Alice and Bob. This set defines density matrices that satisfy the constraints imposed by the two parties' measurement results and thus depends on the specific choice of measurement POVM, the amount of loss in the protocol, and any sources of noise.

2.2.2 Qubit-Based BB84

The BB84 protocol is a QKD protocol proposed in 1984 [25] that is relatively simple while demonstrating all of the important components of a QKD protocol. We use this section to illustrate the application of each component of a protocol in our mathematical framework.

1. **Signal Preparation** Alice chooses the following qubit states to send to Bob (note $|D\rangle \equiv \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle)$ and $|A\rangle \equiv \frac{1}{\sqrt{2}}(|H\rangle - |V\rangle)$)

$$\{|H\rangle, |V\rangle, |D\rangle, |A\rangle\}.$$

   This set is partitioned into two: the z-basis contains the states $|H\rangle$ and $|V\rangle$, and the x-basis contains the states $|D\rangle$ and $|A\rangle$. With probability p_z, Alice prepares a random state from the z-basis, and with probability $p_x = 1 - p_z$, she prepares a random state from the x-basis. The prepared state is stored in system A'. Alice also takes a note of which state she sent in her system A, which we take to simply be the index of the prepared state in the above set. After signal preparation, Alice's systems hold the state

$$|\psi\rangle_{AA'} = \sqrt{\frac{p_z}{2}}|1\rangle_A|H\rangle_{A'} + \sqrt{\frac{p_z}{2}}|2\rangle_A|V\rangle_{A'} + \sqrt{\frac{p_x}{2}}|3\rangle_A|D\rangle_{A'} + \sqrt{\frac{p_x}{2}}|4\rangle_A|A\rangle_{A'}. \tag{2.11}$$

   For completeness, we note here that the POVM that Alice uses to measure her state is

$$\Gamma^A = \{|1\rangle\langle 1|, |2\rangle\langle 2|, |3\rangle\langle 3|, |4\rangle\langle 4|\}$$

2. **Measurement** Bob receives the state that he received from Alice into his system B and uses the POVM

$$\Gamma^B = \{p_z|H\rangle\langle H|, p_z|V\rangle\langle V|, p_x|D\rangle\langle D|, p_x|A\rangle\langle A|\}$$

   to measure it.

3. **Acceptance Test** Alice and Bob choose a small subset of their exchanged signals and reveal measurement information to obtain a joint probability distribution $p(a, b)$ for $a \in \{1, 2, 3, 4\}$ and $b \in \{H, V, D, A\}$. From this probability distribution, they decide whether to abort the protocol or continue to key generation.

4. **Public Discussion** For the remainder of the signals, Alice and Bob announce the partition their measurement belonged to, i.e. whether their measured state was in the x-basis or z-basis. They discard all events where they measured in different bases; events where they measured in the same basis contribute to key generation.

5. **Key Map** Alice assigns the key bit 0 to $|H\rangle$ and $|D\rangle$ signals and the key bit 1 to $|V\rangle$ and $|A\rangle$ signals.

6. **Error Correction** Alice sends error correction information to Bob as described previously.

7. **Privacy Amplification** Alice and Bob each implement privacy amplification as described previously.

In the asymptotic limit, Alice and Bob can test an arbitrarly large number of signals in the acceptance test step. This allows them to determine the expectation values of the operators $\Gamma_i^A \otimes \Gamma_j^B$ for $i \in \{1,2,3,4\}$ and $j \in \{H,V,D,A\}$, as

$$\gamma_{i,j} = \mathrm{Tr}\left((\Gamma_i^A \otimes \Gamma_j^B)\rho_{AB}\right). \tag{2.12}$$

From this information, they can define the set $\mathbf{S}$, which constrains the optimization in Eqn. 2.10 such that the search space includes only those density matrices that would produce the statistics observed in their protocol, so

$$\mathbf{S} = \left\{\rho_{AB} : \forall i,j, \mathrm{Tr}\left((\Gamma_i^A \otimes \Gamma_j^B)\rho_{AB}\right) = \gamma_{i,j}\right\}. \tag{2.13}$$

We use Eqns. 2.7, 2.8, and 2.9 to construct the $\mathcal{G}$ and $\mathcal{Z}$ maps. In the BB84 protocol, the announcements are simply the basis choices ("z-basis" or "x-basis"), and the key map function in $\mathcal{G}$ function has a simple form:

$$g(x,\alpha,\beta) = \begin{cases} x, & \alpha = \beta \\ \perp, & \alpha \neq \beta. \end{cases}$$

Note that $\perp$ represents a discarded signal.

Finally, Eqn. 2.10 can be used to determine the rate at which secret key can be generated per signal exchanged between Alice and Bob.

2.3 Quantum Optics

In order to extend the scope of QKD beyond ideal qubit protocols, we need to utilize results from quantum optics. In this section, we briefly introduce relevant details quantum optical theory, based primarily on [26]. For a summary of the foundations of the theory, see Appenix A.

2.3.1 Fock States

In Appendix A, we show how the electromagnetic field is quantized into fundamental excitations known as photons. A quantum state of light is composed of a number of photons grouped together. States with a definite number of photons are known as *Fock states*, and are the conceptually simplest states of light. A Fock state containing exactly n photons is written as $|n\rangle$.

Fock states can be transformed through the use of conjugate operators $\hat{a}$ and $\hat{a^\dagger}$, known as annihilation and creation operators respectively. These operators obey the commutation relation of Eqn. A.11 and can be used to add or remove one quanta of energy from a state:

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle \tag{2.14}$$

$$\hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\,. \tag{2.15}$$

As with all quantum states, we can also form superpositions and mixtures of Fock states to produce more complicated states.

Multimode Light

Assigned to each creation and annihilation operator pair is a mode l, which refers to an independent property of the light, such as its polarization, time of arrival, or position. In this work, we are interested in distinguishing light through its polarization. Light states in separate polarization modes do not interact with each other, which is apparent in the multimode commutation relations

$$[\hat{a}_l, \hat{a}_{l'}] = 0 = [\hat{a}_l^\dagger, \hat{a}_{l'}^\dagger] \tag{2.16}$$

$$[\hat{a}_l, \hat{a}_{l'}^\dagger] = i\hbar\delta_{l,l'} \tag{2.17}$$

(see Eqn. A.34).

For an example of a multimode state, consider the state

$$\frac{1}{\sqrt{2}}\left(|1\rangle_H\,|0\rangle_V + |0\rangle_H\,|1\rangle_V\right) \equiv \frac{1}{\sqrt{2}}\left(|1,0\rangle_{HV} + |0,1\rangle_{HV}\right) = \frac{1}{\sqrt{2}}\left(\hat{a}_H^\dagger + \hat{a}_V^\dagger\right)|0,0\rangle\,,$$

which is a superposition of the state of exactly one horizontally polarized photon and the state of exactly one vertically polarized photon. We can convert this state to be in the

diagonal polarization basis DA by converting the annihilation operators in the H and V modes to the D and A modes via the relations

$$\hat{a}_D = \frac{1}{\sqrt{2}} (\hat{a}_H + \hat{a}_V) \tag{2.18}$$

$$\hat{a}_A = \frac{1}{\sqrt{2}} (\hat{a}_H - \hat{a}_V) . \tag{2.19}$$

(the creation operators are obtained via a trivial conjugation). We then see that the above state is simply

$$\frac{1}{\sqrt{2}} \left(\hat{a}_H^\dagger + \hat{a}_V^\dagger \right) |0,0\rangle = \hat{a}_D^\dagger |0,0\rangle = |1,0\rangle_{DA} ,$$

exactly one photon in the diagonal polarization state.

2.3.2 Coherent States

Though useful and conceptually simple, Fock basis states are highly nonclassical and thus do not represent states that are easy to produce experimentally. In the discussion of quantization, we moved from the classical picture to the quantum picture by replacing the continuous variables p and q in Eqn. A.6 with operators $\hat{p}$ and $\hat{q}$ in Eqn. A.8. We now examine what happens when we reverse this process—instead of replacing a continuous variable with an operator, we will replace the operator $\hat{a}$ with a continuous variable α. One way to make this replacement is to look for eigenvalues of the annihilation operator:

$$\hat{a} |\alpha\rangle = \alpha |\alpha\rangle . \tag{2.20}$$

We can expand the state α in the Fock basis as

$$|\alpha\rangle = \sum_{n=0}^{\infty} C_n |n\rangle , \tag{2.21}$$

for $C_n \in \mathbb{C}$. Applying Eqn. 2.20 to Eqn. 2.21 yields a recurrence relation whose solution is (up to a phase)

$$C_n = \frac{\alpha^n}{\sqrt{n!}} C_0, \tag{2.22}$$

where C_0 can be determined by normalization to have the value $\exp\left(-\frac{1}{2}|\alpha|^2\right)$. This yields states of the form

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle . \tag{2.23}$$

These states are known as coherent states.

Coherent states are defined by the eigenvalue $\alpha \in \mathbb{C}$, which is a complex amplitude. This value is related to the mean number of photons in the pulse, which we denote μ, by

$$\mu = |\alpha|^2.$$

Given a field in a coherent state $|\alpha\rangle$, the probability of detecting n photons is

$$P(n) = |\langle n|\alpha\rangle|^2 = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!} = e^{-\mu} \frac{\mu^n}{n!} \equiv P_\mu(n), \tag{2.24}$$

which is a Poissonian distribution with mean μ.

Coherent states are especially useful because they describe the output of a laser source. Lasers produce coherent light of a particular frequency through spontaneous and stimulated electronic transitions [27]. As discussed in [28, 29], when a laser is rapidly turned on and off to send a large number of signals per second, the phase of the coherent light emitted is effectively random if the period of sending signals is large compared to the period of oscillation of the light emitted. We will assume that this is the case; see [28] for an analysis of phase-correlated light in QKD.

With this assumption of random phase, the state emitted from a laser has the form

$$|\psi\rangle = |\sqrt{\mu}e^{i\phi}\rangle$$

for a fixed $\mu \in \mathbb{R}^{\geq 0}$ and random $\phi \in [0, 2\pi)$. However, since we have no knowledge about the value of ϕ, we must consider the mixture of all possible values of ϕ:

$$\rho = \frac{1}{2\pi} \int_0^{2\pi} |\sqrt{\mu}e^{i\phi}\rangle \langle\sqrt{\mu}e^{i\phi}| \, d\phi \tag{2.25}$$

$$= e^{-\mu} \sum_n \frac{\mu^n}{n!} |n\rangle \langle n| . \tag{2.26}$$

By considering the quantity $\langle n|\rho|n\rangle$, we can see that this state follows the same Poissonian photon number detection distribution as Eqn. 2.24. States of the form in Eqn. 2.26 are known as *weak coherent pulses*.

2.3.3 Detectors

For the purposes of QKD, it is also useful to mention measurement of multiphoton states. For our purposes, we focus on threshold detectors, which are detectors that fire when 1

or more photons arrive but do not distinguish the number of photons. The measurement POVM of a perfect threshold detector for photons in a mode l looks like

$$\Gamma^l = \{\Gamma^l_{\text{vac}}, \Gamma^l_{\text{click}}\}$$

where

$$\Gamma_{\text{vac}} = |0\rangle\langle 0|_l \quad \text{and} \quad \Gamma_{\text{click}} = \sum_{n=1}^{\infty} |n\rangle\langle n|_l \,.$$

A detection event is referred to as a "click", which simply refers to the case that at least one photon reached the detector.

In the case of multiple modes, our setup will include a detector for each mode, so we need to consider the case where photons arrive at multiple detectors at once. For the H and V modes in the z-basis of the BB84 protocol, the multimode POVM has the form

$$\Gamma^z = \{\Gamma^z_{\text{vac}}, \Gamma^z_0, \Gamma^z_1, \Gamma^z_{\text{double}}\}$$

where

$$\begin{aligned}
\Gamma^z_{\text{vac}} &= |0,0\rangle\langle 0,0|_{HV} \\
\Gamma^z_0 &= \sum_{n=1}^{\infty} |n,0\rangle\langle n,0|_{HV} \\
\Gamma^z_1 &= \sum_{n=1}^{\infty} |0,n\rangle\langle 0,n|_{HV} \\
\Gamma^z_{\text{double}} &= \sum_{m,n=1}^{\infty} |n,m\rangle\langle n,m|_{HV} \,.
\end{aligned}$$

For the x-basis, the POVM Γ^x is identical if we replace the HV modes with DA modes.

Chapter 3

Interfacing with QKD Experiments

In this chapter, we delve into specific details of the framework we use to calculate key rates from the protocols we focus on. We delve first into decoy state analysis, showing how it allows us to enjoy the simplicity of qubit-based QKD while using easily implementable sources. We describe the considerations that must be taken into account in the case of a finite number of signals sent, which is an issue of fundamental importance when working with real-world protocols. We provide mathematical descriptions for the protocols we focus on in this work, the four-six protocol and the three-state protocol. Finally, we describe our numerical framework for computing tight, reliable lower bounds on key rates in quantum key distribution.

3.1 Decoy State Analysis

Decoy state analysis [8, 9, 10, 30, 31] is a tool of fundamental importance in QKD, as it allows for the theoretical simplicity of qubit-based QKD to be applied when using simple attenuated laser sources. Decoy state analysis is a postprocessing technique that provides protection against an attack known as the photon number splitting (PNS) attack by characterizing the action of Eve's attack on the quantum channel connecting Alice and Bob. Though it sacrifices key rate compared to qubit-based protocols, the tradeoff in ease of implementation far outweighs the reduced key rate, making decoy state analysis widely used throughout QKD.

3.1.1 QKD with WCP

The BB84 protocol can be formulated in terms of weak coherent pulses (WCP) [32, 33], which improves the ease of implementation [34]. Instead of requiring perfect single photon sources, these pulses can be simply implemented by attentuating a laser source low enough that the probability of sending more than one photon is low.

In the WCP regime, the majority of signals sent have either 0 photons or 1 photon, though the presence of signals with more than 1 photon is an unavoidable consequence of Poissonian distributions. In the case of 0 photons, Bob measures nothing, so he can map the corresponding event to the discard symbol $\perp$. The case of 1 photon is equivalent to the qubit BB84 protocol. If the pulse contains 2 or more photons, the signal is vulnerable to the PNS attack, wherein Eve splits off all but one photon from the multiphoton signal. With some careful signal discarding, Eve's interference can acquire a significant amount of information about the secret key while remaining undetectable to Alice and Bob based on their measurement statistics alone. Clearly, the PNS attack presents a significant security risk. Alice and Bob can utilize decoy state analysis to characterize Eve's attack and detect a potential PNS attack [9, 30].

Shield Systems

As weak coherent pulses are mixed states, they are inherently vulnerable to being a mixed subsystem of a larger pure state that Eve has control of. If Eve were to hold a purification of the state ρ_{AB} that Alice and Bob share, she could obtain full knowledge of the secret key. To preserve the purity of the shared state of Alice and Bob, we introduce a *shield system* [35].

A shield system is an extra quantum system, held by Alice and denoted A_s, that forms a purification of the state $\rho_{AA'}$. Following [36], we choose to let the shield system represent the photon number sent. It is not important that Alice has access to the system A_s after she sends the system A' to Bob, but it must remain outside of Eve's hands to preserve the purity of the state. The full state held by Alice after signal preparation, but before anything is sent to Bob, is

$$|\psi\rangle_{AA_sA'} = \sum_i \sqrt{p_i}\,|i\rangle_A \otimes \sum_n \sqrt{P_\mu(n)}\,|n\rangle_{A_s}\,|\rho_i\rangle_{A'}\,, \tag{3.1}$$

where we have represented with p_i the probability of Alice sending state i and where $|\rho_i\rangle$ represents the pure state that Alice sends to Bob when she sends state $|i\rangle$. The state $\rho_{AA'}$ is recovered by taking the partial trace over the shield system A_s.

The final state ρ_{AA_sB} is block diagonal in photon number of the shield system, so we can write

$$\rho_{AA_sB} = \sum_n \sqrt{P_\mu(n)} \left|n\right\rangle \left\langle n\right|_{A_s} \otimes \rho_{AB}^{(n)}, \tag{3.2}$$

where

$$\rho_{AB}^{(n)} = \left(\left\langle n\right|_{A_s} \otimes I_{AB}\right) \rho_{AA_sB} \left(\left|n\right\rangle_{A_s} \otimes I_{AB}\right)$$

is the state Alice and Bob share when n photons are sent. From Eqn. 3.2, we can write the acceptance set in terms of $\rho_{AB}^{(n)}$ as

$$\mathbf{S}_n = \{\rho_{AB}^{(n)} \ : \ \forall i,j, \mathrm{Tr}\left(\left(\Gamma_i^A \otimes \Gamma_j^B\right)\left(\rho_{AB}^{(n)}\right)\right) = \gamma_{i,j}^{(n)}\},$$

where the $\gamma_{i,j}^{(n)}$ are the measurement statistics for an n-photon state. We note that these values are not known to Alice and Bob directly but can be bounded based on their measurements. This will be the subject of decoy analysis in the following section.

As proven in [36], the block diagonal structure of ρ_{AA_sB} allows us to split the state in the key rate formula of Eqn. 2.10 on photon number and write it as a sum of minimizations:

$$R_\infty = \sum_{n=0}^{\infty} \min_{\rho_{AB}^{(n)} \in \mathbf{S}_n} P_\mu(n) f(\rho_{AB}^{(n)}) - p_{\mathrm{pass}} \delta_{\mathrm{leak}}. \tag{3.3}$$

It was previously mentioned that multiphoton pulses are vulnerable to being attacked by the PNS attack; therefore, the relative entropy $f(\rho_{AB}^{(n)}) = \mathrm{D}\left(\mathcal{G}(\rho_{AB}^{(n)}) || \mathcal{Z} \circ \mathcal{G}(\rho_{AB}^{(n)})\right)$ is zero for $n \geq 2$. Thus, our key rate is

$$R_\infty = \min_{\rho_{AB}^{(0)} \in \mathbf{S}_0} P_\mu(0) f(\rho_{AB}^{(0)}) + \min_{\rho_{AB}^{(1)} \in \mathbf{S}_1} P_\mu(1) f(\rho_{AB}^{(1)}) - p_{\mathrm{pass}} \delta_{\mathrm{leak}}.$$

The vacuum term corresponding to $n = 0$ contributes to key rate, but the effect is miniscule. Any vacuum signal that Alice generates is naturally unknown to Eve, as she sees no quantum state. The only way for Bob to measure a detection for such a signal is through dark counts in his detectors; this can contribute to key if Bob's detector that experiences a dark count corresponds to the signal state that Alice prepared. Thus, the vacuum contribution to key rate is on the order of the dark count rate, which must anyways be low to be useful for QKD.

In practice, it has been found that the zero-photon contribution is not worth the extra computational time [37, 31]. Decoy analysis is typically the most time-consuming part

of a key rate calculation, so doing additional analysis to find bounds on the zero-photon component for a miniscule improvement to key rate is not worthwhile. However, in the very high loss regime, where key rate is on the order of the dark count rate, including this additional computation is more worthwhile.

The key rate formula for only the single photon contribution is then

$$R_\infty \geq P_\mu(1) \min_{\rho_{AB}^{(1)} \in \mathbf{S}_1} f\left(\rho_{AB}^{(1)}\right) - p_{\text{pass}}\delta_{\text{leak}}. \tag{3.4}$$

In the following section, we will show how to constrain this set to simplify the optimization.

3.1.2 Decoy State Analysis in QKD

In decoy state analysis, Alice will choose to send her states with the mean photon number μ randomly selected from a predetermined set $\{\mu_i\}_{i=1}^{n_D}$, where n_D is the total number of decoy intensities to use—typically between two and four. It is important to note here that, while Eve may know the set $\{\mu_i\}$, she does not know which i was chosen and thus does not know the intensity of each sent signal individually. After communication has finished, Alice and Bob publicly discuss the intensity of each signal sent and sort their detection events, arriving at a three-dimensional probability distribution for their observed statistics $p(a, b, i)$. Based on these statistics, Alice and Bob can determine how Eve influenced their signal, if at all. Though Eve could potentially know the number of photons in the pulse Alice sends, she has no way to know which intensity was selected by Alice, so she has no choice but to make an attack independent of intensity.

Let Σ_A denote the set of states Alice sends in and Σ_B denote the set of Bob's possible measurement outcomes. The probability that a signal of intensity μ_i, sent in polarization state $a \in \Sigma_A$, is detected by Bob in mode $b \in \Sigma_B$ is an important detection probability $\gamma_{a,b}^{\mu_i}$. This value can be determined from the probability $P_{\mu_i}(n)$ that Alice sent that n photons for the intensity μ_i (see Eqn. 2.24) and the yield $Y_n(a, b)$, which is the probability of a detection given the signal had n photons

$$Y_n(a, b) = \Pr(\text{click in detector } b \mid n \text{ photons sent in polarization } a).$$

We denote the detection probability $\gamma_{a,b}^{\mu_i}$ as such because it corresponds to the joint POVM measurement $\Gamma_a^A \otimes \Gamma_b^B$, can be written as

$$\gamma_{a,b}^{\mu_i} = \sum_{n=0}^{\infty} P_{\mu_i}(n) Y_n(a, b). \tag{3.5}$$

The values Y_n characterize the attack Eve performs, and they are the only unknowns in the above equation; $\gamma_{a,b}^{\mu_i}$ is known from the measurement statistics, and $P_{\mu_i}(n)$ is known from Alice's source characterization. For signals with $n = 1$, the state that Bob receives from Alice is a single photon, which we can treat as a qubit. In this case, we note that the values

$$Y_1(a,b) = \Pr(\text{click in detector } b \mid 1 \text{ photon sent in polarization } a)$$

are particularly important, as they describe the subset of signals sent that behave like qubits for the purposes of QKD.

Eqn. 3.5 is a set of n_D equations with an infinite number of unknowns $\{Y_n\}_{n=0}^{\infty}$. We choose a truncation limit N_{ph} and then utilize the fact that $0 \leq Y_n \leq 1$ for all n to place bounds on $\gamma_{a,b}^{\mu_i}$:

$$\gamma_{a,b}^{\mu_i} = \sum_{n=0}^{\infty} P_{\mu_i}(n) Y_n(a,b) \geq \sum_{n=0}^{N_{ph}} P_{\mu_i}(n) Y_n(a,b) \tag{3.6}$$

and

$$\begin{aligned} \gamma_{a,b}^{\mu_i} = \sum_{n=0}^{\infty} P_{\mu_i}(n) Y_n(a,b) &\leq \sum_{n=0}^{N_{ph}} P_{\mu_i}(n) Y_n(a,b) + \sum_{n=N_{ph}+1}^{\infty} P_{\mu_i}(n) \cdot 1 \\ &= \sum_{n=0}^{N_{ph}} P_{\mu_i}(n) Y_n(a,b) + \left(1 - \sum_{n=0}^{N_{ph}} P_{\mu_i}(n)\right). \end{aligned} \tag{3.7}$$

Armed with the aforementioned n_D equations as well as bounds on $\gamma_{a,b}^{\mu_i}$, we can use linear programming to solve for bounds on the single photon tields $Y_1(a,b)$ for each a and b [38]. For each expectation value $\gamma_{a,b}$, we solve the following linear programs

$$\begin{aligned} \text{minimize} \quad & Y_1(a,b) \\ \text{subject to} \quad & 0 \leq Y_n \leq 1 \qquad\qquad n = 1, \ldots, N_{ph} \\ & \gamma_{a,b}^{\mu_i} \geq \sum_{n=0}^{N_{ph}} P_{\mu_i}(n) Y_n(a,b) \\ & \gamma_{a,b}^{\mu_i} \leq \sum_{n=0}^{N_{ph}} P_{\mu_i}(n) Y_n(a,b) + \left(1 - \sum_{n=0}^{N_{ph}} P_{\mu_i}(n)\right). \end{aligned} \tag{3.8}$$

$$
\begin{aligned}
\text{maximize} \quad & Y_1(a,b) \\
\text{subject to} \quad & 0 \leq Y_n \leq 1 \qquad\qquad n = 1, \ldots, N_{ph} \\
& \gamma_{a,b}^{\mu_i} \geq \sum_{n=0}^{N_{ph}} P_{\mu_i}(n) Y_n(a,b) \qquad (3.9) \\
& \gamma_{a,b}^{\mu_i} \leq \sum_{n=0}^{N_{ph}} P_{\mu_i}(n) Y_n(a,b) + \left(1 - \sum_{n=0}^{N_{ph}} P_{\mu_i}(n)\right).
\end{aligned}
$$

The result of these linear programs are lower bounds $Y_1^L(a,b)$ and upper bounds $Y_1^U(a,b)$, respectively, on the single photon conditional probability for Alice sending signal a and Bob measuring outcome b.

With these statistics, we redefine $\mathbf{S}$ from Eqn. 2.13 for the case of a single-photon subspace of a WCP:

$$
\mathbf{S}_1 = \left\{ \rho_{AB} : \forall a \in \Sigma_A, b \in \Sigma_B, \gamma_{a,b}^L \leq \mathrm{Tr}\left((\Gamma_a^A \otimes \Gamma_b^B) \rho_{AB} \right) \leq \gamma_{a,b}^U \right\}. \qquad (3.10)
$$

Note that the outcome of the linear programs are conditional probabilities, whereas the bounds $\gamma_{a,b}^L$ and $\gamma_{a,b}^U$ must be joint probabilities. To perform this conversion, we must simply multiply the linear program results by the probability $\Pr(a)$ of Alice sending signal a:

$$
\gamma_{a,b}^L = \Pr(a) Y_1^L(a,b) \qquad \gamma_{a,b}^U = \Pr(a) Y_1^U(a,b).
$$

With this updated constraining set, we can include it in the key rate formula in Eqn. 3.4:

$$
R_\infty = P_\mu(1) \min_{\rho \in \mathbf{S}_1} f(\rho) - p_{\text{pass}} \delta_{\text{leak}}. \qquad (3.11)
$$

3.2 Squashing Maps

As a starting point in theoretical analysis, QKD protocols are typically defined with respect to a finite-dimensional quantum system, such as qubits. However, when dealing with the implementation of optical QKD protocols, our signals reside in infinite-dimensional Fock space. This is problematic because, in a key rate calculation, we need to find Eve's optimal attack by optimizing over density matrices shared by Alice and Bob, and numerical optimization over an infinite dimensional space is an intractible problem. This discrepancy cannot be resolved by naïvely truncating photon number statistics, as this leaves a protocol

vulnerable to eavesdropping unless a bound can be placed on the amount of information an eavesdropper could obtain. Through the use of a so-called "squashing" map [39, 40], we can rigorously convert the infinite-dimensional protocol to a virtual finite-dimensional protocol whose key rate lower bounds the key rate of the original protocol.

A squashing map is implemented by the composition of a squashing channel Λ to the quantum channel Φ connecting Alice and Bob; without loss of generality, we assume that Eve has full control over it. This channel reduces the dimension of the sent state to finite dimensions. Bob measures this reduced-dimension state and obtains statistics that reproduce the original statistics of the infinite dimensional state.

Definition 3.2.1 (Squashing Model)**.** Let Γ^U be a finite POVM describing the measurement outcomes of a physical device acting on states in the infinite dimensional complex Euclidean space $\mathcal{U}$, and let Γ^V be the corresponing outcomes for a measurement on states in the finite dimensional complex Eucliean space $\mathcal{V}$. If there exists a completely positive map $\Lambda : \mathcal{D}(\mathcal{U}) \to \mathcal{D}(\mathcal{V})$ such that, for any state $\rho \in \mathcal{D}(\mathcal{U})$,

$$\mathrm{Tr}\left(\Gamma_k^U \rho\right) = \mathrm{Tr}\left(\Gamma_k^V \Lambda[\rho]\right)$$

for all k, then Λ is a *squashing map*.

The squashing maps defined in [40] accomplish this task through random reassignment of multi-detector click events to single click events, which introduces an artificial qubit error rate. In addition, these squashing maps do not exist for the case of asymmetric basis choice probabilities, which can be a difficult requirement for experimental implementations of QKD. Instead, we implement the flag state squasher of [41, 42], which does not suffer from either of these drawbacks. However, the flag state squasher does have the drawback of requiring a bound on the population of a subspace of the total Fock space.

The flag-state squasher works by splitting the infinite-dimensional Fock basis into lower and upper subspaces based on a photon number cutoff. Any signal in the upper subspace is assumed to be fully known by Eve. Though this cutoff number may be chosen freely, in this work we focus on the flag-state squasher in the case of reducing down to the single photon subspace.

Definition 3.2.2 (Flag-state Squasher)**.** Let the set $\{\Gamma_k^U\}$ be a POVM on a space $\mathcal{U}$ whose K elements commute with the projection operators $\Pi_{\leq 1}, \Pi_{>1}$, which project onto the orthogonal $n \leq 1$ photon and $n > 1$ photon subspaces of $\mathcal{U}$, respectively. Suppose also that $\rho \in \mathcal{D}(\mathcal{U})$ is block diagonal in total photon number, so we can write $\rho = \bigoplus_{n=0}^{\infty} p_n \rho_n$.

For each k, define $\Gamma_k^{\leq 1}$ to be the restriction of Γ_k^U to the $n \leq 1$ subspace. There exists a completely positive, trace preserving *flag-state squashing map*

$$\Lambda : \mathcal{U}_{\leq 1} \oplus \mathcal{U}_{>1} \mapsto \mathcal{U}_{\leq 1} \oplus \mathcal{U}_f,$$

where $\mathcal{U}_f$ is a classical system of so-called "flags", such that

$$\mathrm{Tr}\left(\Gamma_k \rho\right) = \mathrm{Tr}\left(\Gamma_k^{\leq 1} \Lambda(\rho)\right)$$

for all $\rho \in \mathcal{D}(\mathcal{U})$. This is accomplished by defining a new POVM

$$\tilde{\Gamma}_k = \Gamma_k^{\leq 1} \oplus |k\rangle\langle k| = \left[\begin{array}{c|c} \Gamma_k^{\leq 1} & 0 \\ \hline 0 & |k\rangle\langle k| \end{array}\right], \tag{3.12}$$

where $|k\rangle \in \mathbb{C}^K$. The resulting POVM $\{\tilde{\Gamma}_k\}$ has three dimensions corresponding to the $n \leq 1$ photon subspace, inherited from $\Gamma_k^{\leq 1}$, as well as an extra dimension for each outcome k labeling the POVM $\{\Gamma_k\}$.

To avoid the issue of Eve flooding the channel with signals outside of the $n \leq 1$ photon subspace, we must obtain an estimate of the population of the $n > 1$ subspace. The following reasoning is based on work presented in [42, 43]. Given an observable A which is block diagonal in the total photon number, suppose its expectation value $\langle A \rangle = \mathrm{Tr}\left(\rho A\right)$ is known. Then, defining $a_n = \mathrm{Tr}\left(\rho_n A_n\right)$, we can write

$$\langle A \rangle = \sum_{n=0}^{\infty} p_n \mathrm{Tr}\left(\rho_n A_n\right) = \sum_{n=0}^{\infty} p_n a_n.$$

In our case, we choose the observable

$$A = 1 - \Gamma_{vac} - \sum_{i=0}^{1} \sum_{\alpha} \Gamma_i^{\alpha}, \tag{3.13}$$

where we have used the notation of Section 2.3.3 and used α to index measurement basis choice. This observable represents the situation in which multiple detectors click simultaneously, which can only happen when Bob receives multiple photons (Note, however, that this does not cover *all* multiphoton events; receiving multiple photons in one detector is still a possibility).

If there exist bounds $\bar{a}_{\leq 1}$ and $\bar{a}_{>1}$ such that $\bar{a}_{>1} > \bar{a}_{\leq 1}$,

$$a_n > \bar{a}_{>1} \quad \text{for} \quad n > 1$$

and

$$a_n > \bar{a}_{\leq 1} \quad \text{for} \quad n \leq 1,$$

then the population in the $n \leq 1$ subspace is lower bounded by

$$P_{\leq 1} \geq \frac{\bar{a}_{>1} - \langle A \rangle}{\bar{a}_{>1} - \bar{a}_{\leq 1}}. \tag{3.14}$$

In our case, we obtain $\langle A \rangle$ from the statistics corresponding to any combination of multiple detectors clicking; trivially, we know $\bar{a}_{\leq 1} = 0$, and

$$\bar{a}_{>1} = 1 - \sum_{\alpha} p_{\alpha}^2, \tag{3.15}$$

where p_a is the probability that Bob measures in basis α. This lower bound, which is a result from Theorem 1 of [44], is valid for the $n \leq 1$ subspace of the flag state squasher in the case of perfect 4 or 6 state receivers with threshold detectors.

Inserting these bounds into 3.14 yields the lower bound

$$P_{\leq 1} = 1 - \frac{\langle A \rangle}{1 - \sum_{\alpha} p_{\alpha}^2}, \tag{3.16}$$

where $\langle A \rangle$ is determined from the observations $\gamma_{a,b}$ corresponding to the POVM elements in Eqn. 3.13.

3.3 Finite Size Effects

Thus far, we have only discussed key rates in the asymptotic limit of infinite signals sent. This is overly optimistic, as unlimited signals means the testing rounds can exactly reproduce expectation values of POVM measurements while leaving an unlimited amount of signals for key generation. However, in real QKD experiments, we do not have access to an unlimited number of signals.

To consider the finite number of signals sent, we must make a slight modification to the signal preparation step of our QKD protocol. When the signal is prepared, Alice also randomly decides to denote it as a *test round* or a *generation round*. Test rounds, which

occur with probability $p(test)$, are used for the acceptance test step, whereas generation rounds, which occur with probability $p(g) = 1 - p(test)$, are used for key generation.[1]

In our security framework, the intermediate steps of acceptance testing, error correction, and privacy amplification all contribute to the security of the protocol [45]. We denote these contributions with the corresponding symbols ε_{AT}, ε_{EC}, and ε_{PA}. These values represent the probability that the respective step in the protocol proceeds without aborting *and* the state shared by Alice and Bob is not secure from Eve. We also include a "smoothing" parameter $\bar{\varepsilon}$, which describes the extent to which we include similar states in an intermediate entropy calculation leading up to the Devetak-Winter formula in Eqn. 2.6. In total, a QKD protocol is ε-sound if it is $\varepsilon_{cor} = \varepsilon_{EC}$ correct and $\varepsilon_{sec} = \max\{\varepsilon_{AT}, \varepsilon_{PA} + 2\bar{\varepsilon}\}$ secure. This maximization results from the fact that we have two cases for achieving ε_{sec}-security: either Alice and Bob communicate and the statistics are unacceptable, in which case there is a probability ε_{AT} that the acceptance test step fails to cause an abortion; or the statistics were acceptable and Alice and Bob can generate a secure key with length determined by their smoothed entropy calculation and privacy amplification, which process fails with probability $\varepsilon_{PA} + 2\bar{\varepsilon}$ [46].

In the regime of finite signals sent, sampling error results in an inevitable deviation between the statistics observed $\gamma_{i,j}$ and the expectation values $\mathrm{Tr}\left((\Gamma_i^A \otimes \Gamma_j^B)\rho_{AB}\right)$ on the shared state between Alice and Bob. Because of this, we need to define a reference frequency distribution $\bar{F}$ to which the observed statistics $\gamma_{i,j}$ must be close in order for the protocol to proceed at all. We use $\mathbf{Q}$ to denote the set of collections of observed statistics that are sufficiently close to this reference distribution

$$\mathbf{Q} = \left\{\{\gamma_{i,j}\} \in \mathcal{P}(\Sigma) \ : \ \forall i,j, |\bar{F}_{i,j} - \gamma_{i,j}| \leq t\right\}, \tag{3.17}$$

where t is the finite acceptance parameter. A larger finite acceptance parameter results in an increased likelihood of acceptance but reduces key rate. This is because a larger acceptance set means a larger tolerance for Eve's interference in the protocol. Ideally, t should be chosen to be as small as possible while still allowing the protocol to accept.

With $\mathbf{Q}$ defined, we can update the constraint set in Eqn. 2.13 to include only those density matrices whose calculated expectation values are sufficiently close to the observed statistics:

$$\mathbf{S}_{\mu_e} = \left\{\rho_{AB} \ : \ \exists\{\gamma_{i,j}\} \in \mathbf{Q} \ s.t. \ \forall i,j, |\gamma_{i,j} - \mathrm{Tr}\left((\Gamma_i^A \otimes \Gamma_j^B)\rho_{AB}\right)| \leq \mu_e\right\}, \tag{3.18}$$

[1] In principle, for any protocol implementing decoy state analysis, Alice should only randomly choose the intensity in test rounds, whereas in generation rounds, she should always choose the primary signal intensity. However, this principle is not necessary for security; failure to follow it does not compromise security, though it does reduces key rate.

where μ_e is an entrywise bound on variation for finite size effects.

Combining this new optimization set with the contributions from our security parameters, we arrive at the finite key rate formula

$$R_N = \frac{\Pr(s) - t}{\Pr(s|g)} \min_{\rho \in \mathbf{S}_{\mu_e}} f(\rho) - \Pr(s,g)\delta_{\text{leak}} - \frac{1}{N}\log\frac{2}{\varepsilon_{EC}} - \sqrt{\frac{\Pr(s)-t}{N}}\Delta(\bar{\varepsilon}) - \frac{2}{N}\log\frac{1}{2\varepsilon_{PA}}, \tag{3.19}$$

a result proven in [46]. In this formula, we used the text s to denote the event wherein a detected signal survives sifting and g to label generation rounds. We also define

$$\Delta(\bar{\varepsilon}) = 2\log(1+\dim(X))\sqrt{\log\frac{2}{\bar{\varepsilon}^2}},$$

which contributes to lost key rate due to aforementioned smoothing.

3.3.1 Finite Effects in Decoy Analysis

Analysis of the application of finite-size QKD to decoy analysis has been done in the past [47, 48]. These analyses have relied on techniques that give tighter bounds than we will find in our work; however, they rely on assumptions that are difficult to overcome in the case of imperfect sources or detectors. This is particularly problematic in the case of experimental QKD, as even well-calibrated instruments perform worse than ideal devices in theory. We instead use the framework developed and proved in [46] to address the issue of decoy analysis with finite size effects. In this subsection, we merely present the main results of the framework.

Following the discussion in Section 3.1.2, we define the yields to be

$$Y_n(a,b) = \Pr(b|a,g,n), \tag{3.20}$$

which is the probability Bob measures a click in detector b given that n photons were sent by Alice in state a in round denoted by $g \in \{test, gen\}$, referring to testing and generation rounds, respectively. Since we only know acceptance testing results based on test signals, these are the signals we use in the decoy analysis. To account for looseness in expectation values due to the finite number of test signals, the bounds on $\gamma_{a,b}^{\mu_i}$ in Eqs. 3.6 and 3.7 become

$$\gamma_{a,b}^{\mu_i}(test) \geq \Pr(a,\mu_i,test)\left(\sum_{n=0}^{N_{ph}} P_{\mu_i}(n)Y_n(a,b)\right) - t - \mu_e \tag{3.21}$$

and

$$\gamma_{a,b}^{\mu_i}(test) \leq \Pr(a, \mu_i, test) \left(\sum_{n=0}^{N_{ph}} P_{\mu_i}(n) Y_n(a,b) + 1 - \sum_{n=0}^{N_{ph}} P_{\mu_i}(n) \right) + t + \mu_e. \tag{3.22}$$

Here $\Pr(a, \mu_i, test)$ is the joint probability that Alice sends signal a with intensity μ_i in a test round.

We use the results of decoy analysis to define the quantities

$$\gamma_{a,b}^{L} = \Pr(a|gen) Y_1^L(a,b) \quad \text{and} \quad \gamma_{a,b}^{U} = \Pr(a|gen) Y_1^L(a,b),$$

with which we define the set

$$S_{\mu_e}^{(1)} = \left\{ \rho_{AB} \ : \ \gamma_{a,b}^{L}(test) \leq \mathrm{Tr}\left((\Gamma_a^A \otimes \Gamma_b^B) \rho_{AB} \right) \leq \gamma_{a,b}^{U}(test) \right\} \tag{3.23}$$

Our final key rate formula, accounting for finite signals sent and the usage of decoy state analysis, is

$$R_N = \frac{\Pr(s) - t}{\Pr(s|g)} \min_{\rho \in S_{\mu_e}^{(1)}} f(\rho) - \Pr(s,g)\delta_{\text{leak}} - \frac{1}{N} \log \frac{2}{\varepsilon_{EC}} - \sqrt{\frac{\Pr(s) - t}{N}} \Delta(\bar{\varepsilon}) - \frac{2}{N} \log \frac{1}{2\varepsilon_{PA}}. \tag{3.24}$$

3.4 Important Protocols in this Work

In this section, we will briefly describe the QKD protocols that will be focused on in this work.

3.4.1 Four-Six Protocol

The four-six protocol, drawing inspiration from [49], utilizes a four-state sender and a six-state receiver.[2] We present an diagram of an experimental setup that could be used for this protocol in Figure 3.1. This diagram is not intended to describe all components that are actually present in a QKD experiment, but serves as a reference point to indicate how the protocol could be implemented in reality.

[2]Note that the work cited here uses a four-state receiver and a six-state sender. As noted in the text, a six-state receiver allows Alice and Bob to counteract misalignment in their quantum channel; by contrast, the advantage of a six-state sender is that Alice and Bob can generate a secure key regardless of the phase ϕ in their shared entangled state $|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}} \left(|H\rangle_A |V\rangle_B + e^{i\phi} |V\rangle_A |H\rangle_A \right)$.

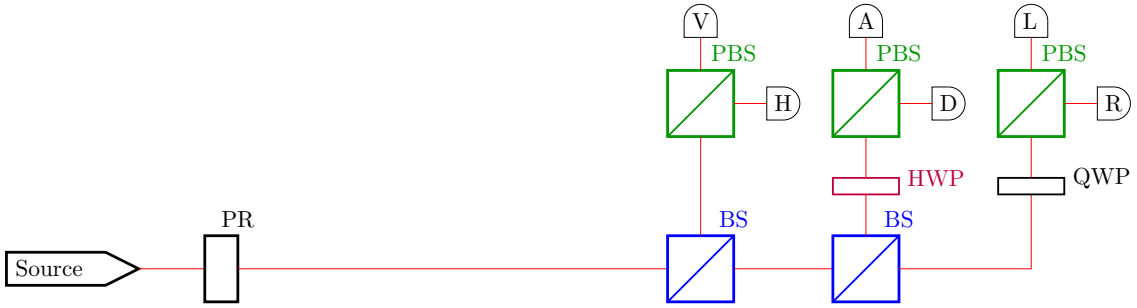


Figure 3.1: Conceptual diagram of the source and detector setup for the 4-6 protocol. The "Source" component denotes Alice's source for weak coherent pulses, and "PR" indicates a polarization rotator, with which Alice selects the polarization state sent. The labels "BS" and "PBS" refer to beamsplitters and polarizing beamsplitters, respectively. Beamsplitters split the input beam into two output beams whose relative intensities are determined by the reflectance of the beamsplitter; polarizing beamsplitters instead split the horizontal and vertical polarizations of the incoming light into different paths. The labels "HWP" and "QWP" indicate half wave plates and quarter wave plates, respectively, which perform passive polarization rotations; specifically, a half wave plate rotates the D and A polarizations to H and V respectively, and a quarter wave plate rotates R and L polarizations into H and V respectively. Finally, the half-rounded square shapes indicate detectors dedicated to measuring incoming light in the basis denoted on the detector.

On Alice's side, the protocol is identical to BB84: she chooses to send states in the set $\{|H\rangle, |V\rangle\}$ with probability p_z^A and states in the set $\{|D\rangle, |A\rangle\}$ with p_x^A, where $p_x^A + p_z^A = 1$. Bob has a setup that allows him to measure in all three polarization bases, meaning he measures the states $\{|H\rangle, |V\rangle\}$ with probability p_z^B, the states $\{|D\rangle, |A\rangle\}$ with probability p_x^B, and the states $\{|R\rangle, |L\rangle\}$ with probability p_y^B, where $p_x^B + p_y^B + p_z^B = 1$.

The choice to measure in 3 bases addresses the problem of misalignment in the source. Misalignment by real angle θ occurs when Alice sends the state $\cos\theta\,|H\rangle + \sin\theta\,|V\rangle$, for some small θ, instead of the state $|H\rangle$ due to faulty polarization alignment between the source and detectors. Measuring in all three polarization bases allows the parties to partially counteract misalignment. This is because a three-basis measurement is tomographically complete on the polarization qubit space, so the measurement statistics can be used to find a unitary $R_{\hat{n}}(\theta)$ representing misalignment along axis $\hat{n}$ by angle θ.

The states that Alice sends lie in an infinite dimensional Fock space, but Bob's mea-

surement results need to eventually reduce to the observations $\gamma_{a,b}^{\mu_i}$ from Eqn. 3.5, which correspond to the measurement results of a POVM with finite dimensions. Therefore, we implement the flag-state squashing map described in Section 3.2.

In this framework, part of the work of squashing is done by our threshold detectors, which map infinite-dimensional fock states to a binary outcome $b \in \{click, noclick\}$. For the four-six protocol, we have six detectors, giving a total of $2^6 = 64$ outcomes. We label these outcomes with the six-digit binary number with a 1 for each detector that clicked and a 0 for each detector that didn't click. The order of detectors in this binary string is ADLRVH. For example, if a multi-photon signal caused the D and H detectors to click simultaneously, the detection is assigned the label `010001`. These detection patterns can then be binned into four classifications:

- **Vacuum**: Events where no detector fires, which is only the `000000` event.
- **Single clicks**: Events where exactly one detector fires. These are the six events `000001`, `000010`, `000100`, `001000`, `010000`, and `100000`.
- **Double clicks**: Events where both detecors in a basis click simultaneously (and no other detectors click). These are the events `000011`, `001100`, and `110000`.
- **Cross clicks**: Every other event, which involves a click in multiple measurement bases, is classified as a cross click.

We note that this binning of detection events is a specific choice of coarse graining detection statistics, and that other choices could be made. For example, double clicks and cross clicks could be binned together as general multiphoton detection events, which would reduce information about the basis choice. If we were to use the squashing map of [40], this would be a loss of important information for mapping multiphoton detection events to single clicks.

Using the flag-state squasher, Bob's POVM post-squashing contains 11 POVM elements, labeled according to the vacuum event $\perp$; single clicks D, A, R, L, H, and V; the double clicks DA, RL, and HV; and cross clicks CC. With one flag for each POVM element as well as the $n \leq 1$ qubit subspace $\Gamma_k^{\leq 1}$ (note we have combined the zero-photon subspace with its flag $\perp$), these POVM elements are 13×13.

We use the observable A in Eqn. 3.13 for subspace estimation; its expectation value is

computed as the sum of the statistics corresponding to DA, RL, HV, and CC events.

$$\tilde{\Gamma}_k^B = \left[\begin{array}{c|c} \Gamma_{\tilde{k}}^{\leq 1} & 0 \\ \hline 0 & \begin{array}{ccccccccccc} \perp & & & & & & & & & & \\ & \mathrm{D} & & & & & & & & & \\ & & \mathrm{A} & & & & & & & & \\ & & & \mathrm{R} & & & & & & & \\ & & & & \mathrm{L} & & & & & & \\ & & & & & \mathrm{H} & & & & & \\ & & & & & & \mathrm{V} & & & & \\ & & & & & & & \mathrm{DA} & & & \\ & & & & & & & & \mathrm{RL} & & \\ & & & & & & & & & \mathrm{HV} & \\ & & & & & & & & & & \mathrm{CC} \end{array} \end{array}\right]$$

Implicitly, each symbol in the lower-right block of $\tilde{\Gamma}_{\tilde{k}}^{\leq 1}$ is the quantity $\delta_{s,k}$, where s is the specific POVM label, i.e. $\mathrm{A} = \delta_{\mathrm{A},k}$. The $n \leq 1$ photon subspace POVM takes the form

$$\Gamma_{\tilde{k}}^{\leq 1} = \{p_x^B \left|D\right\rangle \left\langle D\right|, p_x^B \left|A\right\rangle \left\langle A\right|, p_y^B \left|R\right\rangle \left\langle R\right|, p_y^B \left|L\right\rangle \left\langle L\right|, p_z^B \left|H\right\rangle \left\langle H\right|, p_z^B \left|V\right\rangle \left\langle V\right|\}.$$

On the other hand, Alice's POVM is simply

$$\Gamma^A = \{p_x^A \left|D\right\rangle \left\langle D\right|, p_x^A \left|A\right\rangle \left\langle A\right|, p_z^A \left|H\right\rangle \left\langle H\right|, p_z^A \left|V\right\rangle \left\langle V\right|\}.$$

Alice's POVM elements have been reduced to two dimensions via a Schmidt decomposition.

With these POVMs defined, we can use Eqn. 2.7 to construct $\mathcal{G}$ and Eqn. 2.9 to construct $\mathcal{Z}$, the two maps necessary for key rate calculation.

3.4.2 3-state Protocol

The 3-state protocol we use in this work is based on [50]. A diagram of an experimental setup for this protocol is presented in Figure 3.2. In this protocol, Alice only chooses from three signal states to communicate with Bob, two in the same basis and one in a different basis. Bob measures in all four modes corresponding to the two bases Alice sends in. In our case, Alice chooses states from the set $\{\left|R\right\rangle, \left|L\right\rangle\}$ with probability p_y^A and sends $\left|H\right\rangle$ with probability $p_z^A = 1 - p_y^A$. Bob measures the y-basis states $\left|R\right\rangle, \left|L\right\rangle$ with probability p_y^B and the z-basis states $\left|H\right\rangle, \left|V\right\rangle$ with probability $p_z^B = 1 - p_y^B$.

Figure 3.2: Conceptual diagram of the source and detector setup of the 3-state protocol. See the caption of Fig. 3.1 for component descriptions.

Similar to the 4-6 protocol, we must account for and squash the detector measurements; however, in the case of only four detectors, we have $2^4 = 16$ outcomes, and our detector order is LRHV. In this case, using the flag state squasher gives Bob POVM elements of the form

$$\tilde{\Gamma}_k^B = \left[\begin{array}{c|cccccccccccccc} \Gamma_k^{\leq 1} & & & & & & & 0 & & & & & & & \\ \hline & \perp & & & & & & & & & & & & & \\ & & \mathrm{R} & & & & & & & & & & & & \\ & & & \mathrm{L} & & & & & & & & & & & \\ 0 & & & & \mathrm{H} & & & & & & & & & & \\ & & & & & \mathrm{V} & & & & & & & & & \\ & & & & & & \mathrm{RL} & & & & & & & & \\ & & & & & & & & \mathrm{HV} & & & & & & \\ & & & & & & & & & & \mathrm{CC} & & & & \end{array}\right],$$

where

$$\Gamma_k^{\leq 1} = \{p_y^B \left|R\right\rangle \left\langle R\right|, p_y^B \left|L\right\rangle \left\langle L\right|, p_z^B \left|H\right\rangle \left\langle H\right|, p_z^B \left|V\right\rangle \left\langle V\right|\}.$$

Alice's POVM is

$$\Gamma^A = \{\left|R\right\rangle \left\langle R\right|, \left|L\right\rangle \left\langle L\right|, \left|H\right\rangle \left\langle H\right|\},$$

and thus we are able to construct the maps $\mathcal{G}$ and $\mathcal{Z}$.

We note here that, in the experiment for which we developed our key rate analysis, the source used produced thermal states, not weak coherent pulses. The only effect this has on

our framework is to replace the Poisson distribution $P_{\mu_i}(n)$ in Eqn. 3.5 with the probability distribution corresponding to a thermal state with temperature T (or inverse temperature β), as in Eqn. B.4. See Appendix B for a summary of thermal states in quantum optics.

3.5 Numerical Framework for QKD

We use the numerical framework derived in [1] for calculation of key rates of general QKD protocols. This technique takes in a numerical description of a discrete variable protocol as input and calculates a lower bound on the key rate. We also allow for the inclusion of parameters, such as loss, misalignment, and depolarization, to alter the effect of quantum channel between Alice and Bob. In this section, we briefly explain how a protocol is defined as well as the process for computing key rate.

3.5.1 Input Specification

To facilitate a QKD calculation, a protocol must be specified. Protocols are specified by describing the POVMs Γ^A and Γ^B that Alice and Bob have, respectively, with their corresponding announcements. These specify the signal states Alice uses and the measurement device Bob uses. The user must also define the key map g as well as the $\mathcal{G}$ and $\mathcal{Z}$ maps, as these specify the implementation of announcements and key bit generation. This information is also used for error correction purposes.

However, defining these quantities alone does not suffice for key rate calculation. It is also necessary to know what the resulting channel statistics are, given Alice's input state. To this end, we use a set of statistics $\{\gamma_{ij}\} \in \mathcal{P}(\Sigma)$ corresponding to the joint observables $\Gamma_i^A \otimes \Gamma_j^B$, where Σ is the alphabet formed by the Cartesian product of Alice and Bob's alphabets. These statistics and observables form constraints on Alice and Bob's shared state ρ_{AB}:

$$\mathrm{Tr}\left(\left(\Gamma_i^A \otimes \Gamma_j^B\right)\rho_{AB}\right) = \gamma_{ij}. \tag{3.25}$$

These are known as *equality constraints*. In some cases, such as when finite effects or decoy state analysis are considered, we instead have a set of bounds, $\{\gamma_{ij}^U\}$ and $\{\gamma_{ij}^L\}$, on the statistics such that

$$\gamma_{ij}^L \leq \mathrm{Tr}\left(\left(\Gamma_i^A \otimes \Gamma_j^B\right)\rho_{AB}\right) \leq \gamma_{ij}^U, \tag{3.26}$$

which we refer to as *inequality constraints*. In a typical prepare-and-measure protocol with decoy analysis or finite effects, there will be some subset $\Omega^A \subseteq \Gamma^A$ of k observables

with corresponding equality constraints on Alice's reduced system, since Eve cannot attack it, with the remaining observables corresponding to inequality constraints. Due to their separate treatment, however, we typically define Ω to be the set of observables with equality constraints ω and use Γ to refer to only those observables cooresponding to inequality constraints γ.

Simulated quantum channels can include parameters such as loss, misalignment, and depolarization, and postprocessing maps such as dark counts can be applied to the statistics, allowing us to analyze the performance of QKD protocols in a multitude of noise scenarios. We are also able to import frequencies obtained from experimental observations in QKD experiments, allowing for determination of key rate in QKD experiments.

3.5.2 Computing a Reliable Lower Bound

Given the inputs that allow us to compute $\mathcal{G}$ and $\mathcal{Z}$, our goal is to find the solution to the convex optimization problem

$$\alpha = \min_{\rho \in \mathbf{S}} f(\rho), \tag{3.27}$$

where $\mathbf{S}$ is the appropriate constraint set (as defined, for example, in Eqs. 2.13, 3.10, and 3.18). This is a convex optimization problem for which we need not an approximate solution, but a reliable and, ideally, tight lower bound on the solution.

The key rate calculation happens in two steps.

In the first step, we construct an approximate solution to Eqn. 3.27. The SDP in this case is

$$\begin{aligned} \text{minimize} \quad & f(\rho) \\ \text{subject to} \quad & \mathrm{Tr}\left((\Omega_k^A \otimes I_B)\rho_{AB}\right) = \omega_k \qquad \forall k \\ & \gamma_{ij}^L \leq \mathrm{Tr}\left((\Gamma_i^A \otimes \Gamma_j^B)\rho_{AB}\right) \leq \gamma_{ij}^U \quad \forall i,j \\ & \rho \succeq 0. \end{aligned} \tag{3.28}$$

Conceptually, we are searching for an eavesdropping attack that is close to optimal. This solution ρ can be found in a number of ways, but our approach uses the Frank-Wolfe algorithm [51]. We then compute the gradient of the objective function, $\nabla f(\rho)$, at this point, and use it to linearize f.

In the second step, we first construct the linearized SDP, which effectively bounds f

with a hyperplane, and then compute its dual, which we solve. The dual SDP is

$$\begin{aligned}
\text{maximize} \quad & \sum_k \omega_k u_k^{(eq)} + \sum_{i,j} \frac{\gamma_{i,j}^{(U)} + \gamma_{i,j}^{(L)}}{2} u_{i,j}^{(neq)} - \frac{\gamma_{i,j}^{(U)} - \gamma_{i,j}^{(L)}}{2} z_{i,j}^{(neq)} \\
\text{subject to} \quad & -z_k^{(eq)} \leq u_k^{(eq)} \leq z_k^{(eq)} \qquad \forall k \qquad (3.29) \\
& -z_{i,j}^{(neq)} \leq u_{i,j}^{(neq)} \leq z_{i,j}^{(neq)} \qquad \forall i,j \\
& \nabla f(\rho)^T \geq \sum_k u_k^{(eq)} (\Omega_k \otimes I_B) + \sum_{i,j} u_{i,j}^{(neq)} (\Gamma_i^A \otimes \Gamma_j^B), \qquad (3.30)
\end{aligned}$$

where $u^{(eq)}$ and $z^{(eq)}$ are vectors corresponding to equality constraints and $u^{(neq)}$ and $z^{(neq)}$ correspond to inequality constraints, with each vector being identically indexed to its corresponding set of observations $\{\omega_k\}$ and $\{\gamma_{i,j}\}$.

This dual SDP is a maximization, meaning that even a suboptimal solution to the SDP will be a valid lower bound on key rate. However, this method fails whenever ∇f does not exist, which can happen when our channel statistics are "too ideal": if we have specific channel statistics that are exactly zero (such as $\gamma_{H,V}$ in the case of BB84 with no error), we are effectively at an edge of $\mathbf{S}$, meaning ∇f may not exist. To ameliorate this, $\mathcal{G}$ is modified to include a slight perturbation that maps points ρ on the boundary of $\mathbf{S}$ to interior points. The work in Appendix C of [52] provides a small improvement to this process through the use of tighter solver tolerances.

Chapter 4

Methods for Improving Key Rate Calculations

4.1 Block Diagonal Solver

Our first technique for optimizing key rate calculations is the block diagonal solver. Using our framework, it is common for Alice and Bob to have auxiliary dimensions in their POVMs to accommodate extra information about the signal exchange.

For example, to guarantee that the shared state Alice and Bob is normalized, any protocol involving transmission loss will add an extra dimension corresponding to events where a signal was lost. Thus, a qubit protocol with loss represent Bob's state as a 3x3 block diagonal matrix with a 2x2 block corresponding to the qubit measurement and a 1x1 block corresponding to the extra loss dimension, denoted by $\perp$:

$$\rho_B = \left[\begin{array}{cc|c} & \sigma_B & \begin{matrix}0\\0\end{matrix} \\ \hline 0 & 0 & \perp \end{array}\right]$$

As another example, when using the flag state squasher of Definition 3.2.2, we introduce several additional dimensions to Bob's system. In the case of the four-six protocol, using the flag state squasher increases the size of Bob's system from 3 to 13, as we include flags for the single-photon polarization events D, A, R, L, H, and V as well as for the double click events DA, RL, HV and cross clicks. Without any optimization, even in the case that Alice's system is only two-dimensional, our numerical calculations are optimizing

over 26×26 density matrices. This is not an intractable problem, but it can cause the Frank-Wolfe iteration to experience instability or take longer to converge.

We can improve this situation significantly by informing the underlying SDP solver about the block diagonal structure of Alice and Bob's shared state. When dealing with numerical SDP solvers, when we declare a matrix variable of size $m \times n$, exactly mn variables are defined. We can place constraints on the matrix to reduce the number of degrees of freedom in these variables. However, when we know that certain entries of a matrix variable must be zero, it is inefficient to declare these entries as variables and then constrain them to be zero; instead, where possible, it is preferable to not declare these variables at all. This can be realized when the matrix variable is known to have block diagonal structure.

For a block diagonal matrix A formed by k square blocks of size $a_1 \times a_1$, $a_2 \times a_2$, $\cdots$, $a_k \times a_k$, we say that A has block dimensions $\vec{a} = (a_1, a_2, \cdots, a_k)$. It is often the case that we find block dimension vectors with a large number of repeated elements. In this case, we use a superscript to denote the quantity of repeated elements. For example, we could denote the block dimension vector $(3, 2, 2, 1, 1, 1)$ as $(3^1, 2^2, 1^3)$. If block diagonal matrices A and B have block dimensions $\vec{a}$ and $\vec{b}$ respectively, then their tensor product $A \otimes B$ may not, in general, be a block diagonal matrix, and even in cases where it is block diagonal, the block diagonal structure may not be optimal. This means that there may be an appropriate permutation of the rows and columns of $A \otimes B$ to produce an equivalent block diagonal matrix with block dimensions $\vec{a} \otimes \vec{b}$, which guarantees that no entries that must be zero are contained within these blocks.

As an example, let A and B be matrices with block dimensions $\vec{a} = (2, 1) = \vec{b}$. In the representations that follow, a block labeled with the number n is an $n \times n$ block of possibly nonzero elements, and a block with no numbers is a block of only zeros. We can symbolically describe the tensor product $A \otimes B$ with the following:

$$A \otimes B = \left[\begin{array}{c|c} 2 & \\ \hline & 1 \end{array}\right] \otimes \left[\begin{array}{c|c} 2 & \\ \hline & 1 \end{array}\right] = \left[\begin{array}{cc|cc|cc} 2 & & 2 & & & \\ & 1 & & 1 & & \\ \hline 2 & & 2 & & & \\ & 1 & & 1 & & \\ \hline & & & & 2 & \\ & & & & & 1 \end{array}\right] \tag{4.1}$$

We note that the resulting matrix has block dimensions $(6, 2, 1)$. In this case, we could declare three matrix variables with sizes 6×6, 2×2, and 1×1. This would result in

the declaration of $6^2 + 2^2 + 1^2 = 41$ variables. However, within the 6×6 block, there are sixteen entries that must be zero, so declaring these entries as variables to optimize over is inefficient.

With careful rearranging of rows and columns, we can instead convert this matrix into an equivalent block diagonal matrix with block sizes $(4^1, 2^2, 1^1) = (4, 2, 2, 1) = (2, 1) \otimes (2, 1) = \vec{a} \otimes \vec{b}$. That is, for some choice of a permutation matrix P, we have

$$P(A \otimes B)P^{\mathrm{T}} = \begin{bmatrix} 2 & 2 & & & \\ 2 & 2 & & & \\ & & 2 & & \\ & & & 2 & \\ & & & & 1 \end{bmatrix}, \tag{4.2}$$

which clearly has block dimensions $(4, 2, 2, 1)$. All potentially nonzero entries are contained exactly within these blocks, and no entries that are guaranteed to be zero are inside of any of the blocks, so by declaring four matrix variables one with size 4×4, two with size 2×2, and one with size 1×1, this arrangement minimizes the number of variables $(4^2 + 2^2 + 2^2 + 1 = 25)$ we optimize. For this reason, we refer to a matrix of the form in Eqn. 4.2 as an optimally block diagonal matrix.[1]

This permutation can be accomplished through the use of commutation matrices, which commute subsystems in a tensor product.

Definition 4.1.1 (Commutation matrix)**.** Given an $m \times m$ matrix A and an $n \times n$ matrix B, the *commutation matrix* $K_{(m,n)}$ is the matrix that commutes the tensor product $A \otimes B$. That is,

$$K_{(m,n)}(A \otimes B)K_{(m,n)}^T = B \otimes A.$$

We note that this definition depends only on the dimensions of A and B, not on the values of their entries.

Commutation matrices are commonly applied to quantum information as SWAP operators; however, our usage of these matrices has no physical effect and is purely a numerical convenience, so we do not use the SWAP terminology.

[1] Although our example of an optimally block diagonal matrix has the resulting block dimensions in descending order, this is not necessary for the purposes of improving numerical calculations.

To perform a rearrangement such as that taking Eqn. 4.1 to Eqn. 4.2, we must commute tensor products between each block of A and the entirety of B. This works because, for matrices C and D, if C is an optimally block diagonal matrix and D is not block diagonal, then $C \otimes D$ is optimally block diagonal. In the following equation, we present an example of this fact in the case that C has block dimensions $\vec{c} = (2, 1)$:

$$C \otimes D = \left[\begin{array}{cc:c} 2 & & \\ & & \\ \hdashline & & 1 \end{array}\right] \otimes D = \left[\begin{array}{cc:c} C_{1,1}D & C_{1,2}D & \\ C_{2,1}D & C_{2,2}D & \\ \hdashline & & C_{3,3}D \end{array}\right]. \tag{4.3}$$

Commuting individual tensor products between blocks in A and all of B guarantees that the result is optimally block diagonal.

We can construct P through the direct sum of the commutation matrices between each block of A and B:

$$P = \bigoplus_{i=1}^{|\vec{a}|} K_{(a_i, d_B)}, \tag{4.4}$$

where we note that $K_{(1,n)} = I_n$ for any $n \in \mathbb{N}$. This construction commutates individual tensor products in each sub-block of $A \otimes B$ based on the block dimensions of A.

To apply this to key rate calculation, we first construct P given the block diagonal structures $\vec{a}$ and $\vec{b}$ of Alice and Bob. We also compute the optimal block diagonal structure of the joint state ρ_{AB} via $\vec{c} = \vec{a} \otimes \vec{b}$. When we define ρ in the SDP for solving Eqn. 3.27, we construct a $c_i \times c_i$ matrix σ_i for each i indexing $\vec{c}$, then compute the direct sum $\rho_{AB} = \bigoplus_i \sigma_i$. At this point, ρ_{AB} is optimally block diagonal, but the observables $\Gamma_i^A \otimes \Gamma_j^B$ are not. It is more computationally efficient to invert the block diagonal structure of ρ_{AB} than to apply the transformation $\Phi(X) = PXP^T$ to each POVM element. This does not undo our original computational shortcut, as the reduced number of variables have already been declared and entries that are known to be zero will not be changed.

In Table 4.1, we present the resulting computation time improvements for the three protocols that we have discussed in this work: the BB84, Four-six, and 3-state protocols. We consider both asymptotic and finite size regimes for each of these protocols. The key rate calculation time is averaged over 30 key rate calculations with different channel parameters for each calculation. As is expected, a greater reduction in the number of optimization variables results in a greater speedup in key rate calculation.

We note that the block diagonal solver did not improve key rate in any protocol we tested. Given a sufficiently large problem, however, the reduction in search space offered

Protocol	No Rearrangement		Block Diagonal		Speedup	$\vec{a}$	$\vec{b}$
	Avg. time	Variables	Avg. time	Variables			
Qubit BB84 Asymptotic	0.846 s	36	0.842 s	20	0.5%	(2^1)	$(2^1, 1^1)$
Qubit BB84 Finite	1.28 s	36	1.26 s	20	2%	(2^1)	$(2^1, 1^1)$
Decoy Four-six Asymptotic	26.4 s	156	24.7 s	60	6%	(2^1)	$(2^1, 1^{11})$
Decoy Four-six Finite	19.6 s	156	18.9 s	60	4%	(2^1)	$(2^1, 1^{11})$
Decoy 3-state Asymptotic	11.7 s	900	10.3 s	108	12%	(3^1)	$(2^1, 1^8)$
Decoy 3-state Finite	13.8 s	900	10.9 s	108	21%	(3^1)	$(2^1, 1^8)$

Table 4.1: Runtime comparison for the block diagonal solver. The average times are given as an average over 31 different key rate calculations using different channel parameters for each calculation. The variables column refers to the number of optimization variables declared to the SDP solver. The $\vec{a}$ and $\vec{b}$ columns denote the block dimensons of Alice and Bob's respective systems; our block dimension convention is described in the paragraph preceding Eqn. 4.1. The block diagonal structure of the POVMs for Bob in the four-six and 3-state protocols can be found in Sections 3.4.1 and 3.4.2 respectively.

by the block diagonal solver can, in practice, improve stability and thereby improve the numerically computed lower bound on key rate. However, this is purely a numerical result; with perfect numerical representation and unlimited time, the block diagonal solver does not improve key rate. This is a result of the fact that the block diagonal solver is merely a rearrangement of the optimization problem, so it cannot affect the true lower bound on key rate.

4.2 Constraint Types

In order to calculate key rates of QKD protocols, we construct sets of density operators on Alice and Bob's systems that must fall within constraints set by the channel statistics,

as explained in Section 3.5.1. The constraints explained there, which are used to define the sets $\mathbf{Q}$ and $\mathbf{S}_{\mu_e}$ in Section 3.3, are known as *entrywise* constraints, as they are a set of inequality constraints for each expectation value on an entrywise basis. However, the original generalized framework for finite size analysis of arbitrary QKD protocols [45] defines the set $\mathbf{S}_{\mu_1}$ as

$$\mathbf{S}_{\mu_1} = \left\{\rho_{AB} \ : \ \exists F \in \mathcal{P}(\Sigma) \ s.t. \left\|\Phi(\rho_{AB}) - F\right\|_1 \leq \mu_1 \ \& \ \left\|\bar{F} - F\right\|_1 \leq t\right\},$$

where F denotes the set of observed statistics $\{\gamma_{i,j}\}$ and $\Phi(X) = \sum_{i,j} \mathrm{Tr}\left(X\Gamma_{i,j}\right)|i,j\rangle\langle i,j|$ is a diagonal matrix holding the computed expectation value of the observable $\Gamma_{i,j}$ for each candidate ρ_{AB}. These constraints are referred to as *1-norm* constraints, as they make use of the operator 1-norm for distinguishing probability distributions.

The finite variation bounds μ_e and μ_1 for entrywise and 1-norm constraints, respectively, determine the allowable distance between candidate density matrices found during optimization and the observed statistics. They are computed as functions of the acceptance test security parameter ε_{AT}, the number of signals sent N, the expected statistics $\bar{F}$ and the finite variation threshold t, and the testing probability p_{test}. We present here the formulae for computing these bounds, leaving the derivations to their source works ([46] for μ_e and [45] for μ_1): for μ_e, we have

$$\mu_e = \min_{k,\mu'} \left\{\mu' \in [0,1] \ : \ \max\left\{C_k^{\downarrow}(\mu'), C_k^{\uparrow}(\mu')\right\} = \varepsilon_{AT}\right\}, \tag{4.5}$$

where

$$C_k^{\downarrow}(\mu) = 1 - I_{1-\left(\bar{F}_k - t - \mu\right)}\left(N - \lfloor N\left(\bar{F}_k - t\right)\rfloor - 1, \lfloor N\left(\bar{F}_k - t\right)\rfloor\right), \tag{4.6}$$

$$C_k^{\uparrow}(\mu) = I_{1-\left(\bar{F}_k + t + \mu\right)}\left(N - \lfloor N\left(\bar{F}_k + t\right)\rfloor, \lfloor N\left(\bar{F}_k + t\right)\rfloor + 1\right) \tag{4.7}$$

and $I_x(a,b)$ is the incomplete beta function; whereas for μ_1, we have

$$\mu_1 = \sqrt{2}\sqrt{\frac{\ln(1/\varepsilon_{AT}) + |\Sigma|\ln(p_{test}N + 1)}{p_{test}N}}. \tag{4.8}$$

In general, a smaller μ value gives higher key rate, the space Eve has for attacks is reduced. However, since μ_1 bounds a norm on the full statistics whereas μ_e bounds each entry individually, their values alone are not comparable.

Instead, we present a plot comparing the loss scaling of key rates for the qubit-based BB84 protocol using both types of constraints in Figure 4.1. We note that the key rate for

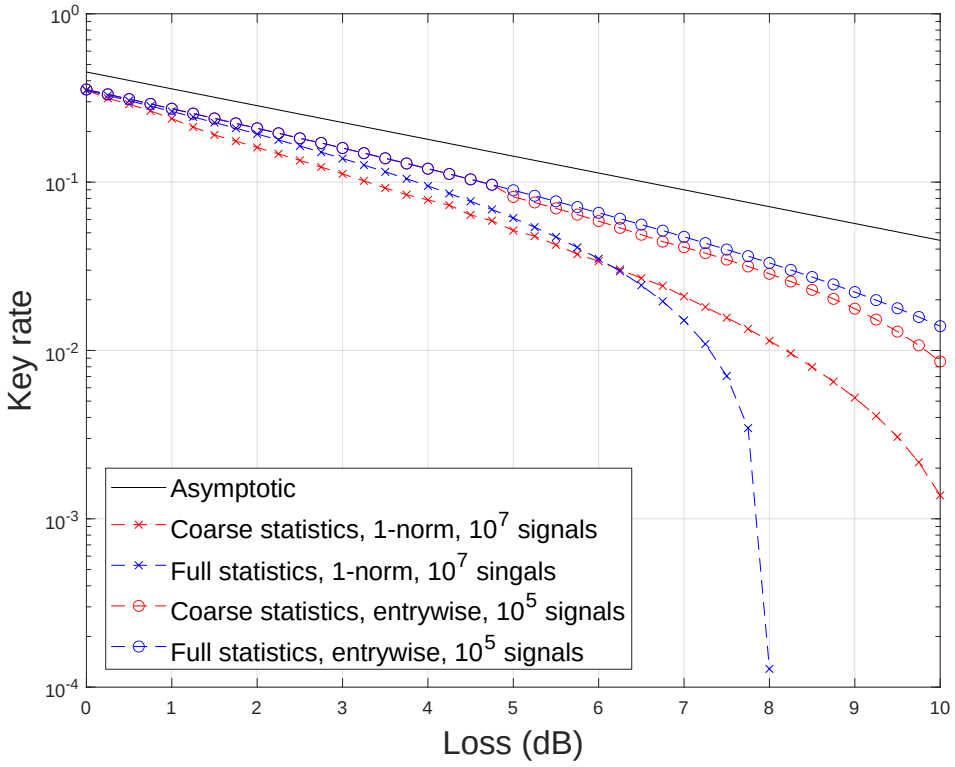


Figure 4.1: Comparison of calculated key rate for the qubit-based lossy BB84 protocol using entrywise constraints versus 1-norm constraints from previous works. We also demonstrate key rate improvement from coarse graining statistics when using 1-norm constraints, which is an undesirable effect. The set of statistics chosen for coarse graining is given in Eqn. 4.9; the fine statistics category uses the full bipartite statistics from Alice and Bob's observations. Physical parameters: depolarization = 0.01, misalignment = 0, $p_z = 0.5 = p_x$, testing probability = 0.0288

entrywise constraints approaches the asymptotic limit at smaller N compared to 1-norm constraints.

In this plot we also demonstrate another advantage of entrywise constraints over 1-norm constraints, which is in the usage of coarse-grained statistics. *Coarse-grained* statistics are statistics that are the result of a postprocessing map, known as a coarse graining map, on the original statistics. When we do not use a coarse graining map to postprocess statistics, we say we have *fine-grained* statistics. Coarse graining maps typically discard parts of

the data that are less essential to characterizing Eve's attack. As described in Section II.B of [45], 1-norm constraints lead to improved key rate in the case of coarse grained data, which is an undesirable effect. In Figure 4.1, we also demonstrate this effect by comparing key rates in the situations of full statistics and coarse statistics for entrywise and 1-norm constraints. In this case, our coarse graining map turns the full bipartite statistics $F = \{\gamma_{i,j}\}$, where $i \in \{H, V, D, A, \}$ and $j \in \{H, V, D, A, \perp\}$, into a new set of statistics

$$\tilde{F} = \{\gamma_{H,H}, \gamma_{H,V}, \gamma_{V,H}, \gamma_{V,V}, \gamma_{D,D}, \gamma_{D,A}, \gamma_{A,D}, \gamma_{A,A}\}, \tag{4.9}$$

which are the statistics of the signals where Alice and Bob measure in the same basis. This coarse graining map discards data, as do all coarse graining maps, so the key rate improvement indicates a situation wherein less-complete data gives better constraints on Alice and Bob's state than complete data would, which is a significant drawback. With 1-norm constraints, optimizing the numerical key rate for a given protocol requires a complicated optimization of coarse graining maps to balance the information loss from coarse graining with the key rate improvement. As is evident in Figure 4.1, entrywise constraints lead to an improvement in key rate when full statistics are used, which is the expectation in the case of more complete information. This also simplifies the task of maximizing key rate for a given protocol, as no optimal coarse graining map needs to be found.

Overall, we see that entrywise constraints lead to better key rates both by better scaling as the number of signals sent N increases and by providing tighter bounds as the amount of data Alice and Bob use increases.

4.3 Intelligent ε-Security Parameters

In determining the security of a QKD protocol, we must choose a small probability ε that the protocol generates an insecure key. As described in Section 3.3, this parameter is a combination of security parameters relating to the acceptance test ε_{AT}, error correction ε_{EC}, privacy amplification ε_{PA}, and entropy smoothing $\bar{\varepsilon}$. The soundness of the protocol is given by the total security parameter $\varepsilon = \varepsilon_{EC} + \max\{\varepsilon_{AT}, \varepsilon_{PA} + 2\bar{\varepsilon}\}$. For a given total security parameter ε, the individual security parameters may be chosen freely.

A naïve starting point is to choose these parameters uniformly, i.e. $\varepsilon_{AT} = \varepsilon_{EC} = \varepsilon_{PA} = \bar{\varepsilon} = \varepsilon/4$. However, since ε_{AT} directly affects the finite variational bounds μ_e in Eqn. 4.5 and μ_1 in Eqn. 4.8, it is advantageous to give it a larger slice of the ε-pie. Increasing ε_{AT} reduces these bounds, which reduces the space to optimize over in a key rate calculation and limits Eve's potential attacks.

However, arbitrarily weighting the security parameters is not as effective as performing an optimization, which was done in [46]. We reproduce the derivation here. We first choose $\varepsilon_{AT} = \varepsilon_{PA} + 2\bar{\varepsilon}$, which is its the maximum allowable value with a fixed total security parameter. Then $\varepsilon = \varepsilon_{EC} + \varepsilon_{PA} + 2\bar{\varepsilon}$. If we then set the key length penalties from error correction and privacy amplification to be equal, we get (see Eqn. 3.19)

$$\log \frac{2}{\varepsilon_{EC}} = 2 \log \frac{1}{2\varepsilon_{PA}} \tag{4.10}$$

$$\iff \varepsilon_{EC} = 8\varepsilon_{PA}^2. \tag{4.11}$$

(Note that this choice is not known to be optimal, but was chosen for a simple relationship between ε_{EC} and ε_{PA}) Since ε is fixed, it follows that for $\bar{\varepsilon}$,

$$\bar{\varepsilon} = \frac{1}{2}(\varepsilon - \varepsilon_{EC} - \varepsilon_{PA}) \tag{4.12}$$

$$= \frac{1}{2}(\varepsilon - 8\varepsilon_{PA}^2 - \varepsilon_{PA}). \tag{4.13}$$

We have now written every security parameter in terms of ε_{PA}, which allows us to compute the derivative of Eqn. 3.19 with respect to ε_{PA} to find the value that maximizes finite key rate:

$$\frac{dR_N}{d\varepsilon_{PA}} = \frac{4}{\varepsilon_{PA} N \ln 2} - \frac{(16\varepsilon_{PA} + 1)\log(1 + \dim(X))\sqrt{p_{\text{sift}} - t}}{\bar{\varepsilon}\sqrt{N \log\left(\frac{2}{\bar{\varepsilon}^2}\right)}}. \tag{4.14}$$

Note that the optimization term in Eqn. 3.19 is independent of ε_{PA}, which significantly simplifies this calculation. Let ε_{PA}^* be the value of ε_{PA} satisfying

$$\left.\frac{dR_N}{d\varepsilon_{PA}}\right|_{\varepsilon_{PA}=\varepsilon_{PA}^*} = 0, \tag{4.15}$$

which we calculate numerically. Finally, we choose the security parameters

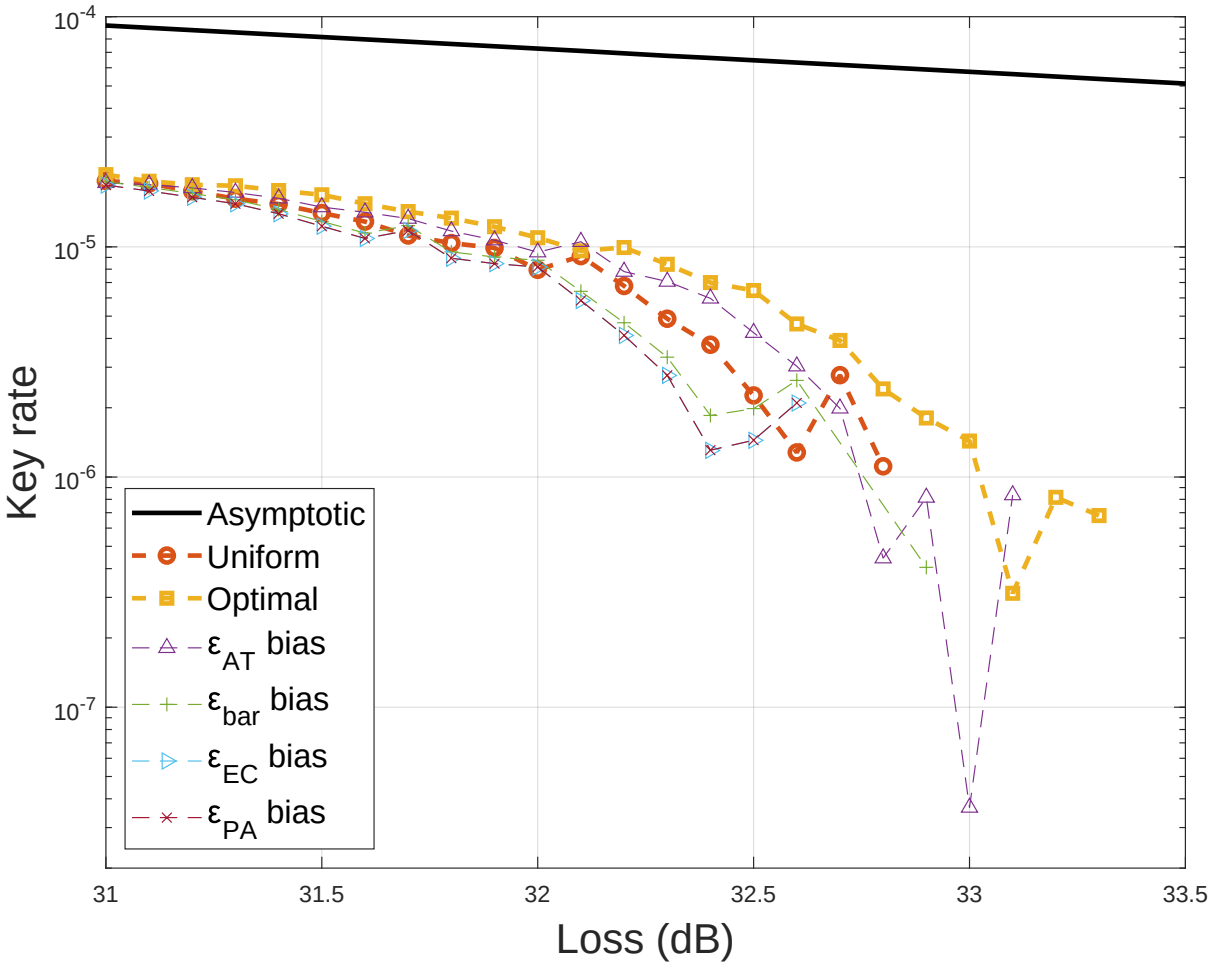
$$\varepsilon_{PA} = \varepsilon_{PA}^* \tag{4.16}$$

$$\varepsilon_{EC} = 8(\varepsilon_{PA}^*)^2 \tag{4.17}$$

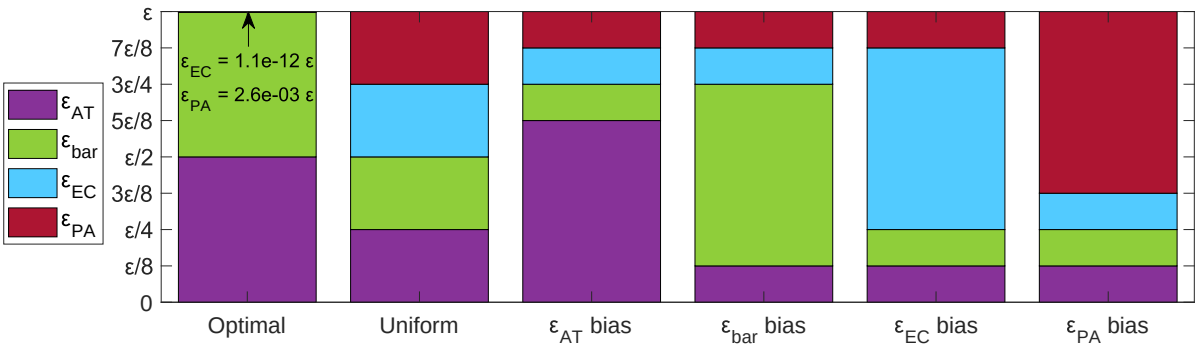
$$\bar{\varepsilon} = \frac{1}{2}\left(\varepsilon - \varepsilon_{PA}^*(8\varepsilon_{PA}^* + 1)\right) \tag{4.18}$$

$$\varepsilon_{AT} = \varepsilon_{PA}^* + 2\bar{\varepsilon}, \tag{4.19}$$

which can be tuned by the total security parameter ε alone. This distribution of security parameters is empirically observed to be a good choice given our assumption about ε_{AT}, an assumption we give veracity to below.



(a) Key rate



(b) ε-security parameters

Figure 4.2: (a) Demonstration of the effect on key rate of security parameter choice in the 3-state protocol with $N = 10^9$ signals sent and using weak coherent pulses and decoy state analysis. The six dotted lines represent different distributions of the security parameters, all with the same total security parameter $\varepsilon = 10^{-8}$. We also provide the asymptotic key rate for this parameter regime, though our intention in this regime is not to approach the asymptotic key rate. (b) Distribution of the ε-security parameters for the plot in (a).

In Figure 4.2a we demonstrate the effect of security parameter choice on key rate for the three-state protocol using weak coherent pulses and decoy state analysis. We have chosen the number of signals sent $N = 10^9$ and focus on the region of greatest tolerable loss, a regime in which we are far from the asymptotic limit. We note the noisiness of the plot indicates these key rate calculations experience numerical instability, which is typically inevitable in any protocol pushed to its limits, but also inherent to the 3-state protocol due to Alice's sending of only three states. The security parameters for each line in Figure 4.2a are represented in Figure 4.2b. We see that the total security epsilon is almost entirely dominated by the contributions from ε_{AT} and $\bar{\varepsilon}$.

When the security parameters are chosen optimally, as described above, we see an improvement in key rate at high loss, and, of all scenarios we tested, the largest maximum tolerable loss of 33.3 dB. In addition, Figure 4.2a supports our assumption that allowing ε_{AT} to be large has the biggest key rate improvement, as the key rates using a security parameter distribution biased towards ε_{AT} performed better than all other distributions except for the optimal distribution.

It is interesting to note that at a loss value of 33.1 dB, the line corresponding to the ε_{AT} bias rises above the optimal distribution due to the computed lower bound dropping at that point for the optimal distribution. We suspect that this is merely a numerical artifact, and we note that, in our experience, this is a common problem in numerical QKD when protocols are pushed to their limits. Such a numerical artifact does not indicate a point where the key rate is unreliable, as our key rate calculation method is guaranteed to find a lower bound on the key rate. Thus, the erratic behavior is not indicative of the true key rates, but indicates that the loss tolerance of the protocol is being pushed to its limit. Evidently, there is gap between the true key rate and our lower bound. This gap could be tightened through a higher accuracy solver, more run time, and optimizing the problem input to avoid badly conditioned problems.

Finally, we also see how a suboptimal distribution of security parameters can reduce key rate compared to a uniform distribution, as biasing towards any security parameter other than ε_{AT} generally reduces key rate compared to the uniform security parameter distribution.

4.4 Maximizing Key Throughput

In many QKD experiments, including ours utilizing the four-six protocol, Alice and Bob communicate for a duration of time over which the quantum channel is not constant.

For example, during a satellite overpass, the channel loss will vary. This is because the channel loss is dependent on the amount of atmosphere the signal transmits through, which is largest at the beginning and end of the overpass and smallest in the middle. If the properties of the channel change during a QKD experiment, Eve's attack is more difficult to characterize, and we have to assume the worst case attack for the full block length. In the case of high loss at the beginning and end of the overpass, this can result in zero secure key.

If, instead, the data is collected and binned into time steps, then we can partition the data such that the underlying channel is similar across all time steps within each partition. We use the statistics obtained within each partition to set up a separate key rate calculation for that partition. The total key length is then the sum of the key rate in each partition multiplied by the total number of signals in this partition.

Maximizing the total number of secret key bits requires balancing the effects of partition size. With a constant quantum channel, key rate is improved for large partitions due to a smaller finite size effect cost, as we can place a tighter bound on the allowed statistical fluctuation μ_e of the measured frequencies. However, the channel characterization can change more over the course of a larger partition, which can negatively impact key rate because a varying channel requires a larger acceptance set threshold t. As more time steps are included in a partition, the key rate changes, increasing with partition size at small sizes until reaching a maximum, after which point key rate decreases as partition size increases. However, when we consider the number of key bits generated in a block of data, the optimal width for bit throughput is much larger.

4.4.1 Channel Description

To demonstrate the benefits of our key throughput maximization approach, we present results of maximized key for predicted from REFQ collaborations [53], which is building a QKD source expected to fly on QEYSSAT [54]. We use data produced by researchers from the University of Strathclyde from a Monte Carlo simulation of a quantum channel through Earth's atmosphere linking a satellite to a ground station. This data is binned into 1-second time bins, which form the basis of our data blocks. A *time bin block* is a collection of time bins whose data is averaged over, representing the communication between the satellite and ground station during that time.

In Figure 4.3, we characterize the quantum channel that signals are sent over during the satellite overpass by plotting the channel loss and the qubit error rates (QBER) in the X and Z bases. Defining $\gamma_{a,b}$ as the observed frequency for the event where Alice sends

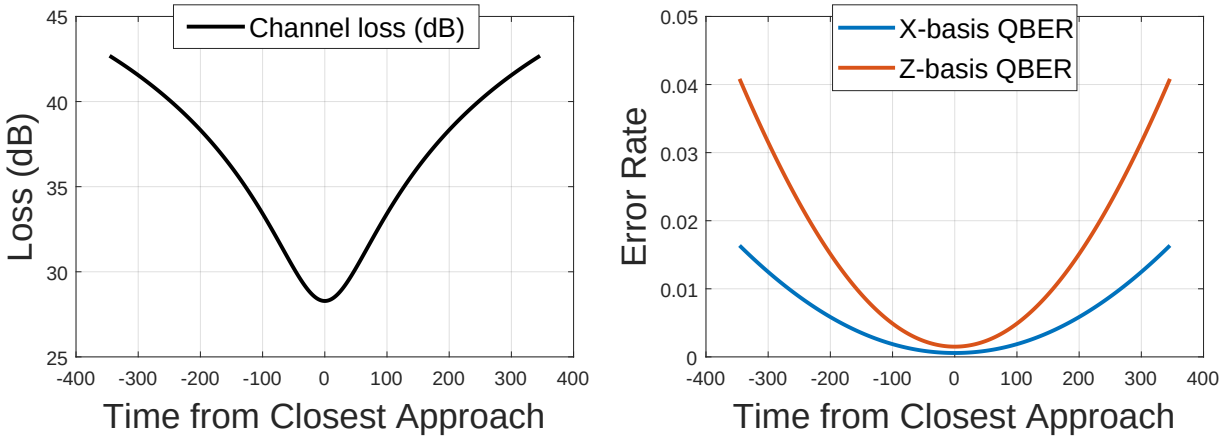


Figure 4.3: Characterization of the quantum channel for simulated satellite data.

state a and Bob measures state b, where $a, b \in \{H, V, D, A\}$, we have

$$QBER_X = \frac{\gamma_{D,A} + \gamma_{A,D}}{\gamma_{D,D} + \gamma_{D,A} + \gamma_{A,D} + \gamma_{A,A}}$$

and

$$QBER_Z = \frac{\gamma_{V,H} + \gamma_{H,V}}{\gamma_{H,H} + \gamma_{H,V} + \gamma_{V,H} + \gamma_{V,V}}.$$

Though these quantities do not completely characterize all that is happening to the quantum state, they give a sense of the noisiness of the channel.

4.4.2 Algorithms for Blocking Time Bin Data

We present the results of different optimizations of the time bin block length and key rate in Figure 4.4. In this discussion, block lengths, which refer to the number of time bins in a key rate caclulation, are denoted by ℓ, and the key rate computed using statistics from a block of size ℓ is denoted $K(\ell)$. For discussion of runtime, we use n to represent the number of time bin blocks and $O(K)$ to represent the cost of one key rate calculation (i.e., the cost of computing the result of Eqn. 3.27). The optimization methods are summarized below.

A. In Method A, we first choose a block size ℓ that is large enough that key rate can be generated in the noisier section of the time bin data. Once ℓ is decided, we group every

ℓ time bins in the overpass into a block for key rate calculation. This method requires only one optimization to choose ℓ, but produces the fewest number of key bits of all methods. An example of the results of this process is depicted in Figure 4.4a, which uses $\ell = 5$, giving us time bin blocks of 5 seconds. This method has a runtime of $nO(K)$.

B. In Method B, we choose a block length multiple m and, starting from the center of the data, optimize the block size as a multiple of m (up to a limit m_{max}), i.e. we compute the key rate of blocks centered on the middle of the data with sizes $\ell = m$, $2m$, $3m$, $m_{max}m$ and choose the multiple of m which gives the highest key rate. This process is repeated for blocks to the left and right of the center. This method can significantly increase the number of key bits generated compared to Method A but is more computationally expensive due to the optimization at each step. Figure 4.4b demonstrates the results of this method on our data with $m = 5$, $m_{max} = 8$. This method has a runtime of $nm_{max}O(K)$.

C. In Method C, we set $\ell = 1$ and search for the choice of block size (between 1 and ℓ_{max}) that maximizes key rate for the given starting point of the data, starting with the center point of the data. This method is slower than Method B, as it searches a larger optimization space, but it can also improve key bit throughput significantly. See 4.4c for the usage of this method on our data with $\ell_{max} = 50$. This method has a runtime of $n\ell_{max}O(K)$.

D. In Method D, we change the objective function we maximize. Instead of searching for an ℓ that maximizes key rate, we choose an ℓ that maximizes the number of key bits contributed by the block, computed as the product $\ell K(\ell)$. Compared to Method C, this produces a larger central block size and similarly sized outer blocks. This choice of optimization further increases the number of key bits generated compared to Method C. It also is not computationally more complex than Method C, as the same optimization routine is used with a different objective function. We present the results of Method D on our data in Figure 4.4d, where ℓ_{max} has been increased to 120. This method also has a runtime of $n\ell_{max}O(K)$.

E. Finally, in method E, we begin at the center and choose the block length ℓ that maximizes $\ell K(\ell)$, as in Method D. We then proceed outwards from the center by choosing blocks of equal length on each side of the previously chosen block, combining the statistics. For example, if the central block had a width $\ell_0 = 7$, combining time steps -3 to 3, the next block would include time steps -4 and 4, -5 and 5, $\cdots$, continuing on through the data. This method attempts to exploit the fact that the loss and the

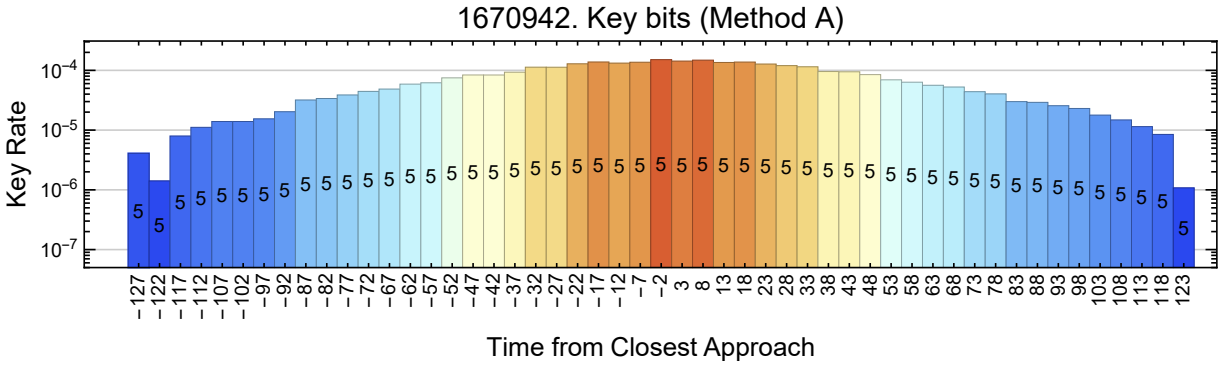
Method	n	Runtime	Limit Parameter	Worst Case (s)	Throughput (bits)
A	51	$nO(K)$		2.04×10^3	1.67×10^6
B	30	$nm_{max}O(K)$	$m_{max} = 8$	9.60×10^3	2.07×10^6
C	28	$n\ell_{max}O(K)$	$\ell_{max} = 50$	5.60×10^4	2.33×10^6
D	25	$n\ell_{max}O(K)$	$\ell_{max} = 120$	1.20×10^5	2.58×10^6
E	20	$n\ell_{max}O(K)$	$\ell_{max} = 120$	9.60×10^4	2.47×10^6

Table 4.2: Runtime results of our key throughput optimization methods. In the second column, n refers to the number of time bin blocks. The values m_{max} and ℓ_{max} represent the number of problems solved to optimize each block; they are differentiated by their relationships to the actual block size as described in the text. In the Worst Case column, we have calculated the maximum amount of time to compute each method based on the empirical estimate $O(K) \approx 40$ s. In all cases, we used the block diagonal solver, entrywise constraints, and an optimal ε-security parameter distribution, as described in previous sections.

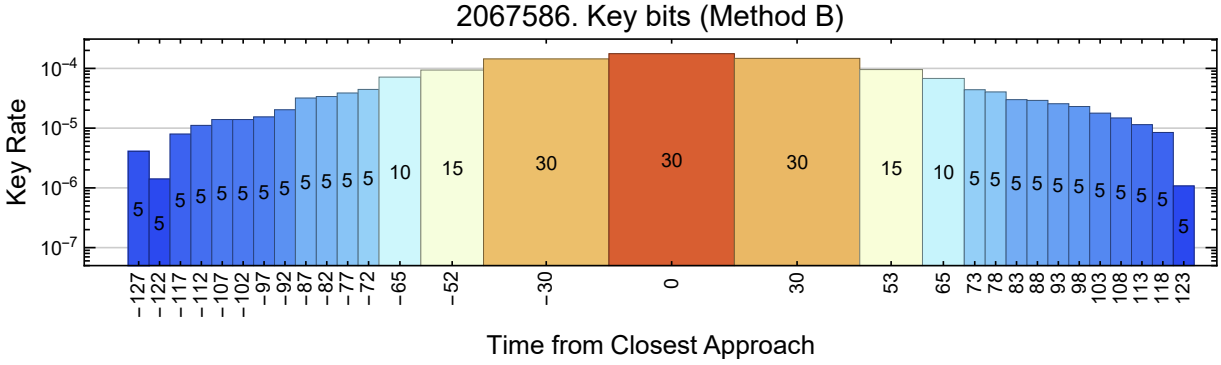
QBER in each basis are all symmetric around the central time, meaning time step t and time step $-t$ have similar noise characterizations. However, as seen in Figure 4.4e, we found that this method produced fewer key bits compared to Method D despite having the same runtime.

We found that, when computation time is not a concern, Method D produces the highest key bit throughput compared to other methods. Interestingly, Method E did not produce better key bit throughput despite being designed as an improvement to Method D.

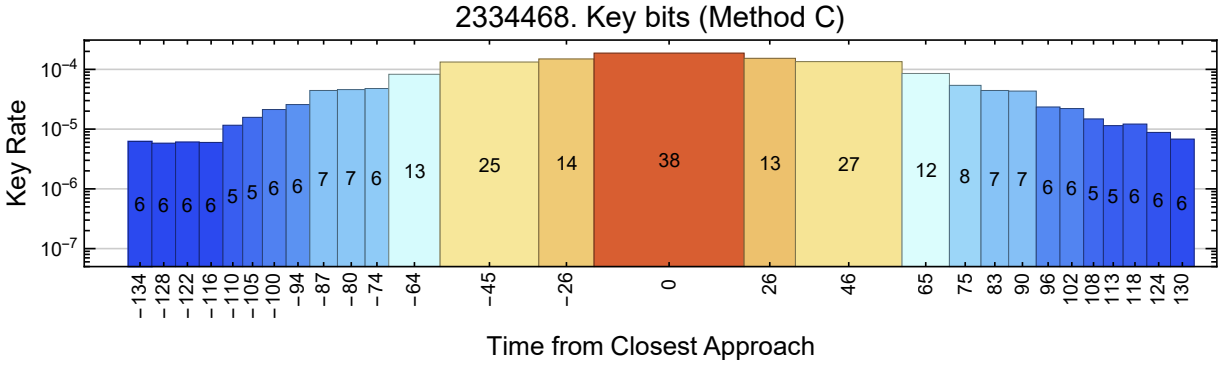
The absolute computation time of these methods depends highly on the protocol, the implementation of decoy state analysis, and the method used to compute the key rate. In our case, which we present in Table 4.2, the time to complete one key rate calculation was found to have an average close to 40 seconds. We see that methods that optimize time bin block size increase key throughput by up to a factor of about 0.5 at the cost of an increase in runtime of up to two orders of magnitude.


1670942. Key bits (Method A)
Key Rate
Time from Closest Approach

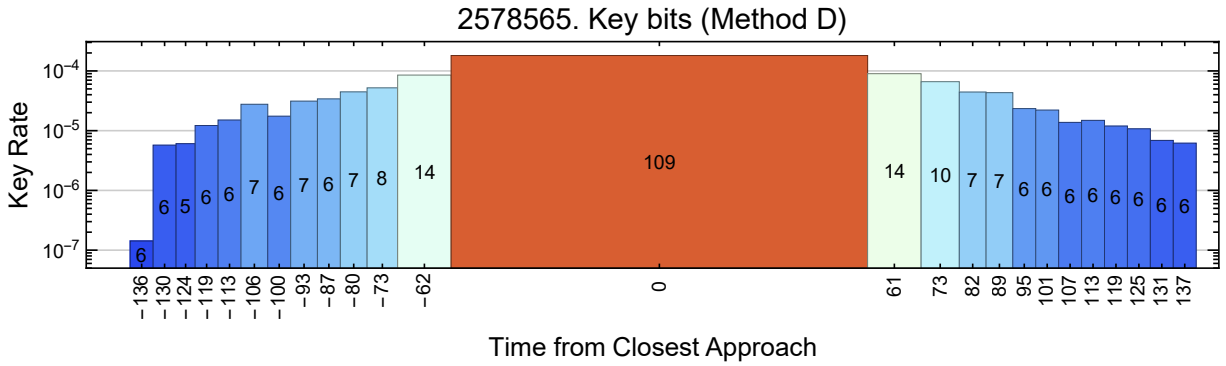
(a)


2067586. Key bits (Method B)
Key Rate
Time from Closest Approach

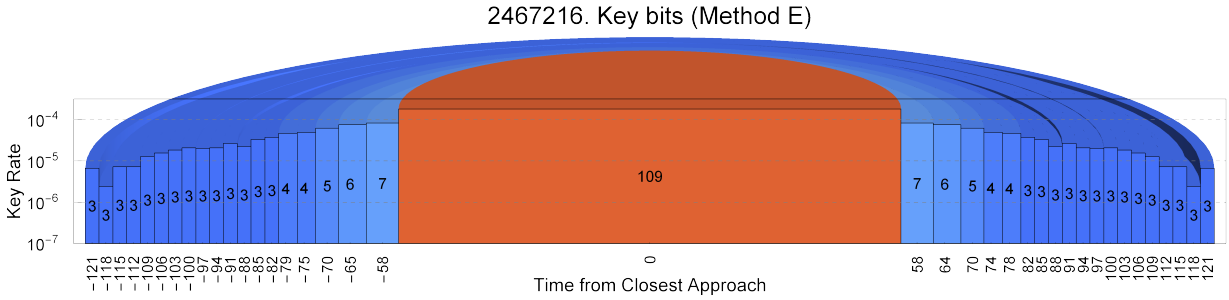
(b)


2334468. Key bits (Method C)
Key Rate
Time from Closest Approach

(c)



(d)



(e)

Figure 4.4: Results of different optimization strategies for maximizing key bit throughput on time binned, simulated satellite data using a base block size of 1 second. Each bar in these figures represents a block and is labeled with a number, which represents the number of time bins in that block. The axis label for each block is the time step on which the block is centered. The height of the block represents the key rate calculated on that block, (a) Method A performs no optimization and simply uses a block of $\ell = 5$ time steps across the time-binned data. (b) Method B begins at the center time step and searches through blocks of sizes that are multiples of $m = 5$, choosing the block length giving the highest key rate and then proceeding similarly forwards and backwards through the data. (c) Method C is the same as method B, but removes the limitation to block sizes that are multiples of $m = 5$. This increases the search time, but also increases the resulting key rate. (d) Method D chooses the block size that maximizes the bit throughput, calculated as the product of key rate and block size. (e) Method E combines data into blocks that are symmetric across time 0. In this figure, all non-central blocks have half of their data at a positive time and half at a negative time; for example, the two blocks of length 7 on either side of the central block are grouped together in a key rate calculation with a time bin block size of 14.

Chapter 5

Conclusion

In this work, we have presented the necessary background to introduce and justify quantum key distribution, from the background of quantum mechanics and information theory to a security proof and example protocols. We also covered the essentials of quantum optics as is needed for the experimental implementation of QKD; explained important techniques that are useful for interfacing with QKD experiments, namely decoy analysis and squashing maps, which allow for simpler physical devices to be used in experiments at the cost of only a minor increase in theory overhead; and discussed finite size effects in QKD, which must be accounted for in any physical implementation. Two specific protocols were presented, with which work has been done to interface between numerical key rate calculations and experimental data.

We demonstrated an improvement in runtime previous numerical key rate calculation methods through the form of our block diagonal solver, which reduces computational overhead. The choice to use entrywise constraints was shown to improve key rate and reduce undesirable scaling effects, such as improved key rate under coarse-grained statistics. Optimizing the distribution of security parameters for a given key rate calculation was also explored, and it was seen that biasing towards a larger acceptance test security parameter gives better key rate for a fixed total security parameter. Finally, we presented a number of methods for maximizing key throughput in the case of data split into time bins, which is a common experimental result. We saw that maximizing the key rate may be less beneficial to key throughput compared to a lower key rate with a larger bin of data. However, we found that the optimization algorithm yielding the largest key throughput also had a much larger runtime, indicating the presence of a tradeoff based on resource requirements.

There is a great amount of work that remains to done to improve and optimize numerical

key rate calculations.

Work has been done on adaptive key rate calculations, which allows for secure key rate generation when the quantum channel Alice sends her state over has time variations, which is the problem we addressed when attempting to maximize key throughput. Applying adaptive key rate methods to this problem may allow for larger time bin block widths, which could further increase key throughput.

Our block diagonal solver improvement is only one step out of a multitude of potential improvements. Our solver applies block diagonal structure to only ρ_{AB}, which only significantly affects the complexity of applying constraints including the quantity $\mathrm{Tr}\left(\Gamma\rho_{AB}\right)$. Further work can be done to apply block diagonal structure to the $\mathcal{G}$ map based on Alice and Bob's announcements, which could significantly reduce the cost of computing $\mathrm{D}\left(\mathcal{G}(\rho)||\mathcal{Z}(\mathcal{G}(\rho))\right)$. This would likely have a larger effect than our improvement, as the map $\mathcal{G}$ increases the dimensions of ρ, so a block diagonal reduction will naturally be more efficient. Vectorization of the construction of the block diagonal ρ_{AB} could also improve the efficiency, as could defining Alice and Bob's joint POVM in optimal block diagonal structure to eliminate the need for rearrangement.

Another direction for future optimization is improvements on decoy analysis. This direction is twofold: first, decoy analysis is oftentimes the primary computational sink for any protocol that uses it, so a faster way of performing decoy analysis could significantly reduce time to calculate key rates; and second, more precise decoy methods, such as those that optimize over channels using Choi matrices, can provide tighter bounds that increase key rate.

References

[1] A. Winick, N. Lütkenhaus, and P. J. Coles, "Reliable numerical key rates for quantum key distribution," *Quantum*, vol. 2, p. 77, July 2018. arXiv:1710.05511 [quant-ph].

[2] P. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," in *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pp. 124–134, 1994.

[3] M. Mosca, "Cybersecurity in an era with quantum computers: will we be ready?." Cryptology ePrint Archive, Paper 2015/1075, 2015. https://eprint.iacr.org/2015/1075.

[4] R. Ursin, F. Tiefenbacher, T. Schmitt-Manderbach, H. Weier, T. Scheidl, M. Lindenthal, B. Blauensteiner, T. Jennewein, J. Perdigues, P. Trojek, B. Ömer, M. Fürst, M. Meyenburg, J. Rarity, Z. Sodnik, C. Barbieri, H. Weinfurter, and A. Zeilinger, "Entanglement-based quantum communication over 144 km," *Nature Physics*, vol. 3, pp. 481–486, July 2007. Number: 7 Publisher: Nature Publishing Group.

[5] P. A. Hiskett, D. Rosenberg, C. G. Peterson, R. J. Hughes, S. Nam, A. E. Lita, A. J. Miller, and J. E. Nordholt, "Long-distance quantum key distribution in optical fibre," *New Journal of Physics*, vol. 8, p. 193, Sept. 2006.

[6] C. J. Pugh, S. Kaiser, J.-P. Bourgoin, J. Jin, N. Sultana, S. Agne, E. Anisimova, V. Makarov, E. Choi, B. L. Higgins, and T. Jennewein, "Airborne demonstration of a quantum key distribution receiver payload," *Quantum Science and Technology*, vol. 2, p. 024009, June 2017. Publisher: IOP Publishing.

[7] S.-K. Liao, W.-Q. Cai, J. Handsteiner, B. Liu, J. Yin, L. Zhang, D. Rauch, M. Fink, J.-G. Ren, W.-Y. Liu, Y. Li, Q. Shen, Y. Cao, F.-Z. Li, J.-F. Wang, Y.-M. Huang, L. Deng, T. Xi, L. Ma, T. Hu, L. Li, N.-L. Liu, F. Koidl, P. Wang, Y.-A. Chen, X.-B. Wang, M. Steindorfer, G. Kirchner, C.-Y. Lu, R. Shu, R. Ursin, T. Scheidl, C.-Z. Peng,

J.-Y. Wang, A. Zeilinger, and J.-W. Pan, "Satellite-relayed intercontinental quantum network," *Physical Review Letters*, vol. 120, p. 030501, Jan. 2018. arXiv:1801.04418 [quant-ph].

[8] W.-Y. Hwang, "Quantum Key Distribution with High Loss: Toward Global Secure Communication," *Physical Review Letters*, vol. 91, p. 057901, Aug. 2003. arXiv:quant-ph/0211153.

[9] X.-B. Wang, "Beating the Photon-Number-Splitting Attack in Practical Quantum Cryptography," *Physical Review Letters*, vol. 94, p. 230503, June 2005. Publisher: American Physical Society.

[10] J. W. Harrington, J. M. Ettinger, R. J. Hughes, and J. E. Nordholt, "Enhancing practical security of quantum key distribution with a few decoy states," Mar. 2005. arXiv:quant-ph/0503002.

[11] R. H. Hadfield, "Single-photon detectors for optical quantum information applications," *Nature Photonics*, vol. 3, pp. 696–705, Dec. 2009. Number: 12 Publisher: Nature Publishing Group.

[12] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge ; New York: Cambridge University Press, 10th anniversary ed ed., 2010.

[13] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 1 ed., Apr. 2018.

[14] S. Hassani, "Vectors and Linear Maps," in *Mathematical Physics: A Modern Introduction to Its Foundations* (S. Hassani, ed.), pp. 19–61, Cham: Springer International Publishing, 2013.

[15] U. Vazirani and T. Vidick, "Fully device-independent quantum key distribution," *Phys. Rev. Lett.*, vol. 113, p. 140501, Sep 2014.

[16] D. P. Nadlinger, P. Drmota, B. C. Nichol, G. Araneda, D. Main, R. Srinivas, D. M. Lucas, C. J. Ballance, K. Ivanov, E. Y.-Z. Tan, P. Sekatski, R. L. Urbanke, R. Renner, N. Sangouard, and J.-D. Bancal, "Experimental quantum key distribution certified by Bell's theorem," *Nature*, vol. 607, pp. 682–686, July 2022. Number: 7920 Publisher: Nature Publishing Group.

[17] W. Zhang, T. van Leent, K. Redeker, R. Garthoff, R. Schwonnek, F. Fertig, S. Eppelt, W. Rosenfeld, V. Scarani, C. C.-W. Lim, and H. Weinfurter, "A device-independent

quantum key distribution system for distant users," *Nature*, vol. 607, pp. 687–691, July 2022. Number: 7920 Publisher: Nature Publishing Group.

[18] V. Zapatero, T. van Leent, R. Arnon-Friedman, W.-Z. Liu, Q. Zhang, H. Weinfurter, and M. Curty, "Advances in device-independent quantum key distribution," *npj Quantum Information*, vol. 9, pp. 1–11, Feb. 2023. Number: 1 Publisher: Nature Publishing Group.

[19] R. Renner, "Security of Quantum Key Distribution," Jan. 2006. arXiv:quant-ph/0512258.

[20] P. J. Coles, E. M. Metodiev, and N. Lütkenhaus, "Numerical approach for unstructured quantum key distribution," *Nature Communications*, vol. 7, p. 11712, May 2016. Number: 1 Publisher: Nature Publishing Group.

[21] J. Lin, T. Upadhyaya, and N. Lütkenhaus, "Asymptotic Security Analysis of Discrete-Modulated Continuous-Variable Quantum Key Distribution," *Physical Review X*, vol. 9, p. 041064, Dec. 2019. Publisher: American Physical Society.

[22] J. Müller-Quade and R. Renner, "Composability in quantum cryptography," *New Journal of Physics*, vol. 11, p. 085006, Aug. 2009.

[23] M. Christandl, R. König, and R. Renner, "Postselection Technique for Quantum Channels with Applications to Quantum Cryptography," *Physical Review Letters*, vol. 102, p. 020504, Jan. 2009. Publisher: American Physical Society.

[24] I. Devetak and A. Winter, "Distillation of secret key and entanglement from quantum states," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 461, pp. 207–235, Jan. 2005. Publisher: Royal Society.

[25] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theoretical Computer Science*, vol. 560, pp. 7–11, Dec. 2014.

[26] C. Gerry and P. Knight, *Introductory Quantum Optics*. Cambridge: Cambridge University Press, 2004.

[27] J. Peatross and M. Ware, *Physics of Light and Optics*. Available at optics.byu.edu, 2015 ed.

[28] S. Nahar, "Decoy-State Quantum Key Distribution with Arbitrary Phase Mixtures and Phase Correlations," 2022.

[29] F. Grünenfelder, A. Boaron, D. Rusca, A. Martin, and H. Zbinden, "Performance and security of 5 GHz repetition rate polarization-based Quantum Key Distribution," *Applied Physics Letters*, vol. 117, p. 144003, Oct. 2020. arXiv:2007.15447 [quant-ph].

[30] H.-K. Lo, X. Ma, and K. Chen, "Decoy State Quantum Key Distribution," *Physical Review Letters*, vol. 94, p. 230504, June 2005. Publisher: American Physical Society.

[31] W. Wang and N. Lütkenhaus, "Numerical security proof for the decoy-state BB84 protocol and measurement-device-independent quantum key distribution resistant against large basis misalignment," *Physical Review Research*, vol. 4, p. 043097, Nov. 2022. Publisher: American Physical Society.

[32] N. Lütkenhaus, "Security against individual attacks for realistic quantum key distribution," *Physical Review A*, vol. 61, p. 052304, Apr. 2000. Publisher: American Physical Society.

[33] G. Brassard, N. Lütkenhaus, T. Mor, and B. C. Sanders, "Security Aspects of Practical Quantum Cryptography," *Physical Review Letters*, vol. 85, pp. 1330–1333, Aug. 2000. arXiv:quant-ph/9911054.

[34] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, "Experimental quantum cryptography," *Journal of Cryptology*, vol. 5, pp. 3–28, Jan. 1992.

[35] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, "General paradigm for distilling classical key from quantum states," *IEEE Transactions on Information Theory*, vol. 55, pp. 1898–1929, Apr. 2009. arXiv:quant-ph/0506189.

[36] N. K. H. Li and N. Lütkenhaus, "Improving key rates of the unbalanced phase-encoded BB84 protocol using the flag-state squashing model," *Physical Review Research*, vol. 2, p. 043172, Nov. 2020. Publisher: American Physical Society.

[37] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo, "Practical decoy state for quantum key distribution," *Physical Review A*, vol. 72, p. 012326, July 2005. Publisher: American Physical Society.

[38] P. Rice and J. Harrington, "Numerical analysis of decoy state quantum key distribution protocols," Jan. 2009. arXiv:0901.0013 [quant-ph].

[39] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, "Security of quantum key distribution with imperfect devices," Sept. 2004. arXiv:quant-ph/0212066.

[40] O. Gittsovich, N. J. Beaudry, V. Narasimhachar, R. R. Alvarez, T. Moroder, and N. Lütkenhaus, "Squashing model for detectors and applications to quantum key distribution protocols," *Physical Review A*, vol. 89, p. 012325, Jan. 2014. arXiv:1310.5059 [quant-ph].

[41] N. K. H. Li, "Application of the Flag-State Squashing Model to Numerical Quantum Key Distribution Security Analysis," Master's thesis, University of Waterloo, Sept. 2020. Accepted: 2020-09-18T13:35:53Z.

[42] Y. Zhang, P. J. Coles, A. Winick, J. Lin, and N. Lutkenhaus, "Security proof of practical quantum key distribution with detection-efficiency mismatch," *Physical Review Research*, vol. 3, p. 013076, Jan. 2021. arXiv:2004.04383 [quant-ph].

[43] N. Lütkenhaus, "Applied Quantum Key Distribution." Course lecture notes, 2020.

[44] L. Kamin and N. Lütkenhaus, "Improved decoy methods." unpublished, N.D.

[45] I. George, J. Lin, and N. Lütkenhaus, "Phys. Rev. Research 3, 013274 (2021) - Numerical calculations of the finite key rate for general quantum key distribution protocols."

[46] L. Kamin, S. Johnstun, and N. Lütkenhaus, "Finite-size effects of decoy state methods." unpublished, N.D.

[47] C. C. W. Lim, M. Curty, N. Walenta, F. Xu, and H. Zbinden, "Concise security bounds for practical decoy-state quantum key distribution," *Physical Review A*, vol. 89, p. 022307, Feb. 2014. Publisher: American Physical Society.

[48] D. Rusca, A. Boaron, F. Grünenfelder, A. Martin, and H. Zbinden, "Finite-key analysis for the 1-decoy state QKD protocol," *Applied Physics Letters*, vol. 112, p. 171104, Apr. 2018.

[49] R. Tannous, Z. Ye, J. Jin, K. B. Kuntz, N. Lütkenhaus, and T. Jennewein, "Demonstration of a 6 State-4 State Reference Frame Independent channel for Quantum Key Distribution," *Applied Physics Letters*, vol. 115, p. 211103, Nov. 2019. arXiv:1905.09197 [quant-ph].

[50] F. Grünenfelder, A. Boaron, D. Rusca, A. Martin, and H. Zbinden, "Simple and high-speed polarization-based QKD," *Applied Physics Letters*, vol. 112, p. 051108, Jan. 2018.

[51] M. Frank and P. Wolfe, "An algorithm for quadratic programming," *Naval Research Logistics Quarterly*, vol. 3, no. 1-2, pp. 95–110, 1956. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nav.3800030109.

[52] T. Upadhyaya, T. van Himbeeck, J. Lin, and N. Lütkenhaus, "Dimension Reduction in Quantum Key Distribution for Continuous- and Discrete-Variable Protocols," *PRX Quantum*, vol. 2, p. 020325, May 2021. arXiv:2101.05799 [quant-ph].

[53] "Reference-Frame Independent Quantum Communication for Satellite-Based Networks (ReFQ)." NSERC UK-Canada Alliance Grant, 2020-2023.

[54] T. Jennewein, J. P. Bourgoin, B. Higgins, C. Holloway, E. Meyer-Scott, C. Erven, B. Heim, Z. Yan, H. Hübel, G. Weihs, E. Choi, I. D'Souza, D. Hudson, and R. Laflamme, "QEYSSAT: a mission proposal for a quantum receiver in space," in *Advances in Photonics of Quantum Computing, Memory, and Communication VII*, vol. 8997, pp. 21–27, SPIE, Feb. 2014.

APPENDICES

Appendix A

Field Quantization in Quantum Optical Theory

In this appendix, we present the foundations of quantum optical theory from the starting point of classical electromagnetism. Our presentation follows the presentation in [26].

A.1 Field Quantization

From classical electromagnetic theory we get Maxwell's equations, which describe the dynamics of electric and magnetic fields produced by currents and charges.

Consider an electromagnetic excitation inside a one-dimensional cavity with effective volume V. The cavity has perfectly conducting walls and no currents or charges within. The resulting electric and magnetic fields must form standing waves in the cavity. In this case, Maxwell's equations for $\mathbf{E}$ and $\mathbf{B}$, the electric and magnetic fields, respectively, can be written as

$$\begin{aligned} \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{B} &= \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \cdot \mathbf{E} &= 0. \end{aligned}$$

In these equations, μ_0 and ϵ_0 are the permittivity and permeability of free space, respectively.

A.1.1 Single-mode Fields

If we assume that the cavity extends in the z-direction, then a single-mode solution to the given equations and boundary conditions has the following form

$$E_x(z,t) = c_\omega q(t) \sin(kz) \tag{A.1}$$

$$B_y(z,t) = \frac{\mu_0 \epsilon_0}{k} c_\omega \dot{q}(t) \cos(kz), \tag{A.2}$$

where $c_\omega = \sqrt{\frac{2\omega^2}{V\epsilon_0}}$ is a constant determined by the frequency ω, k is the wave number $k = \omega/c$, and $q(t)$ is a factor that holds all time-dependence and has dimensions of length.

The classical Hamiltonian can be obtained from the total electromagnetic energy of this field

$$H = \frac{1}{2} \int \left(\epsilon_0 \mathbf{E}^2(\mathbf{r},t) + \frac{1}{\mu_0} \mathbf{B}^2(\mathbf{r},t) \right) dV \tag{A.3}$$

$$= \frac{1}{2} \int \left(\epsilon_0 E_x^2(z,t) + \frac{1}{\mu_0} B_y^2(z,t) \right) dz \tag{A.4}$$

$$= \frac{1}{2} \left(\dot{q}^2 + \omega^2 q^2 \right), \tag{A.5}$$

which we recognize as the Hamiltonian for a harmonic oscillator if we define $\dot{q} \equiv p$:

$$\frac{1}{2} \left(p^2 + \omega^2 q^2 \right). \tag{A.6}$$

With appropriate scaling factors, the electric and magnetic fields play the roles of canonical position and momentum.

Now, to move to the quantum mechanical picture, we appeal to the correspondence principle to replace the canonical variables p and q with operators $\hat{p}$ and $\hat{q}$. By nature of their conjugate relationship, these operators must satisfy the canonical commutation relation

$$[\hat{q}, \hat{p}] = i\hbar \hat{I}. \tag{A.7}$$

The Hamiltonian operator is then

$$\hat{H} = \frac{1}{2} \left(\hat{p}^2 + \omega^2 \hat{q}^2 \right). \tag{A.8}$$

The operators $\hat{q}$ and $\hat{p}$ are directly related to the electric and magnetic fields and thus are observable quantities, meaning they are Hermitian operators. We can introduce non-Hermitian annihilation and creation operators $\hat{a}$ and $\hat{a}^\dagger$, defined as

$$\hat{a} = \frac{1}{\sqrt{2\hbar\omega}}(\omega\hat{q} + i\hat{p}) \tag{A.9}$$

$$\hat{a}^\dagger = \frac{1}{\sqrt{2\hbar\omega}}(\omega\hat{q} - i\hat{p}). \tag{A.10}$$

These operators reveal the discrete, or quantum, nature of this system. Loosely speaking, applying the creation operator $\hat{a}^\dagger$ to a state corresponds to adding a photon to it; Conversely, applying the annihilation operator $\hat{a}$ corresponds to removing a photon from it.

The creation and annihilation operators obey the commutation relation

$$[\hat{a}, \hat{a}^\dagger] = \hat{I}. \tag{A.11}$$

We can re-write Eqn. A.8 in terms of these operators as

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \tag{A.12}$$

$$= \hbar\omega\left(\hat{n} + \frac{1}{2}\right), \tag{A.13}$$

where, in the second equation, we have defined the number operator $\hat{n} \equiv \hat{a}^\dagger\hat{a}$. This is because, for each integer $n \geq 0$, there is an eigenstate $|n\rangle$ of $\hat{H}$ which satisfies the equation $\hat{H}|n\rangle = E_n|n\rangle$ (for eigenenergy E_n) that is also an eigenstate of $\hat{n}$ with eigenvalue n.

$$\hat{n}|n\rangle = n|n\rangle \tag{A.14}$$

Combining Eqs. A.13 and A.14 and the eigenvalue equation gives

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right).$$

This reveals that a system in the state $|n\rangle$ has a collection of n quanta of energy, each with energy $\hbar\omega$, with an additional $\hbar\omega/2$ of energy regardless of the number of energy quanta

present. The operators $\hat{a}$ and $\hat{a}^\dagger$ respectively remove or add one quanta of energy, as seen in their behavior when acting on energy eigenstates:

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle \tag{A.15}$$

$$\hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle \tag{A.16}$$

In an electromagnetic field, these quanta are referred to as photons. This basis is known as the *Fock basis* and uses optical states of definite photon number. As a basis, it is a complete set, meaning any photon state can be expressed as a superposition of number states

$$|\psi\rangle = \sum_n C_n\,|n\rangle$$

for some set of weights $\{C_i\}_{i=1}^{\infty} \subset \mathbb{C}$.

Note that we can recover operators for the electric and magnetic fields in terms of $\hat{a}$ and $\hat{a}^\dagger$:

$$\hat{E}_x(z,t) = \sqrt{\frac{\hbar\omega}{\epsilon_0 V}}\left(\hat{a} + \hat{a}^\dagger\right)\sin(kz) \tag{A.17}$$

$$\hat{B}_y(z,t) = -i\sqrt{\frac{\hbar\omega}{\epsilon_0 V}}\left(\hat{a} - \hat{a}^\dagger\right)\cos(kz). \tag{A.18}$$

A.1.2 Multi-mode Fields

So far our analysis has only concerned single-mode fields, but extension to multimode fields is not prohibitively difficult. We can consolidate the electric and magnetic fields through use of the vector potential $\mathbf{A}(\mathbf{r},t)$, as

$$\mathbf{E}(\mathbf{r},t) = -\frac{\partial \mathbf{A}(\mathbf{r},t)}{\partial t} \tag{A.19}$$

and

$$\mathbf{B}(\mathbf{r},t) = \nabla \times \mathbf{A}(\mathbf{r},t). \tag{A.20}$$

Working in the Coulomb gauge, where

$$\nabla \cdot \mathbf{A}(\mathbf{r},t) = 0, \tag{A.21}$$

$\mathbf{A}(\mathbf{r},t)$ satisfies the wave equation

$$\nabla^2\mathbf{A} = \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2}. \tag{A.22}$$

The wave equation for $\mathbf{A}$ is solved generally by a superposition of plane waves

$$\mathbf{A}(\mathbf{r},t) = \sum_l \mathbf{e}_l \left(c_l e^{i(\mathbf{k}_l\cdot\mathbf{r}-\omega_l t)} + c_l^* e^{-i(\mathbf{k}_l\cdot\mathbf{r}+\omega_l t)}\right), \tag{A.23}$$

where l indexes independent modes, $\mathbf{e}_l$ is a real polarization vector, and c_l is the complex amplitude of the field. From this equation for $\mathbf{A}$, we can find $\mathbf{E}$ and $\mathbf{B}$ using Eqs. A.19 and A.20:

$$\mathbf{E}(\mathbf{r},t) = i\sum_l \omega_l \mathbf{e}_l \left(c_l e^{i(\mathbf{k}_l\cdot\mathbf{r}-\omega_l t)} - c_l^* e^{-i(\mathbf{k}_l\cdot\mathbf{r}-\omega_l t)}\right) \tag{A.24}$$

$$\mathbf{B}(\mathbf{r},t) = \frac{i}{c}\sum_l \omega_l(\hat{\mathbf{k}}_l\times\mathbf{e}_l) \left(c_l e^{i(\mathbf{k}_l\cdot\mathbf{r}-\omega_l t)} - c_l^* e^{-i(\mathbf{k}_l\cdot\mathbf{r}-\omega_l t)}\right) \tag{A.25}$$

It can then be shown that the Hamiltonian of this system, using Eqn. A.3, can be expressed in terms of the amplitudes c_l as

$$H = 2\epsilon_0 V\sum_l \omega_l^2|c_l|^2. \tag{A.26}$$

In order to quantize this system, we need to introduce canonical variables p_l and q_l for each mode, defined by the equations

$$q_l = \sqrt{V\epsilon_0}(c_l + c_l^*) \tag{A.27}$$

$$p_l = -i\omega_l\sqrt{V\epsilon_0}(c_l - c_l^*), \tag{A.28}$$

for which the Hamiltonian takes the form

$$H = \frac{1}{2}\sum_l (p_l^2 + \omega_l q_l^2). \tag{A.29}$$

From this point, we proceed analogously to before, applying the correspondence principle to replace classical variables p_l and q_l with operators $\hat{p}_l$ and $\hat{q}_l$, which obey the commutation relations

$$[\hat{q}_l,\hat{q}_{l'}] = 0 = [\hat{p}_l,\hat{p}_{l'}] \tag{A.30}$$

$$[\hat{q}_l,\hat{p}_{l'}] = i\hbar\delta_{l,l'}. \tag{A.31}$$

Note $\delta_{a,b}$ is the Kronecker delta, defined to be 1 if $a = b$ and 0 otherwise.

We proceed to define creation and annihilation operators for each mode,

$$\hat{a}_l = \frac{1}{\sqrt{2\hbar\omega_l}}(\omega_l \hat{q}_l + i\hat{p}_l) \tag{A.32}$$

$$\hat{a}_l^\dagger = \frac{1}{\sqrt{2\hbar\omega_l}}(\omega_l \hat{q}_l - i\hat{p}_l), \tag{A.33}$$

which satisfy the commutation relations

$$[\hat{a}_l, \hat{a}_{l'}] = 0 = [\hat{a}_l^\dagger, \hat{a}_{l'}^\dagger] \tag{A.34}$$

$$[\hat{a}_l, \hat{a}_{l'}^\dagger] = i\hbar\delta_{l,l'}. \tag{A.35}$$

The quantized Hamiltonian can be written in terms of $\hat{a}_l$ and $\hat{a}_l^\dagger$ as

$$\hat{H} = \sum_l \hbar\omega_l \left(\hat{a}_l^\dagger \hat{a}_l + \frac{1}{2}\right) \tag{A.36}$$

$$= \sum_l \hbar\omega_l \left(\hat{n}_l + \frac{1}{2}\right), \tag{A.37}$$

where the number of excitations for mode l is given, analogously to before, by the operator $\hat{n}_l = \hat{a}_l^\dagger \hat{a}_l$.

A.1.3 Coherent States

Coherent states are sufficiently useful to QKD that we provide the primary introduction to them in Section 2.3.2. In this section of the appendix, we simply show that taking the expectation value of the electric field operator on a coherent state yields

$$\langle\alpha|\hat{E}_x(\mathbf{r}, t)|\alpha\rangle = i\sqrt{\frac{\hbar\omega}{2\epsilon_0 V}}\left(\hat{a}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} - \alpha^* e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\right) \tag{A.38}$$

$$= 2|\alpha|\sqrt{\frac{\hbar\omega}{2\epsilon_0 V}}\sin(\omega t - \mathbf{k}\cdot\mathbf{r} - \arg(\alpha)), \tag{A.39}$$

which aligns with what we would expect for the electromagnetic field of classical light in a cavity (see Eqn. A.1).

The expectation value of energy for a coherent state is given by

$$\langle\alpha|\hat{H}|\alpha\rangle = \hbar\omega\,\langle\alpha|\hat{a}^\dagger\hat{a}|\alpha\rangle = \hbar\omega|\alpha|^2, \tag{A.40}$$

where we have discarded the constant vacuum contribution $\frac{\hbar\omega}{2}$. From this, we define the *mean photon number* of a coherent state to be $\mu = |\alpha|^2$.

Appendix B

Thermal States

It is also useful to consider the representation of radiation emitted by a blackbody in quantum optics. We can model a black body using a cavity at temperature T containing radiation at thermal equilibrium with the walls, with a small hole in one of the walls. From statistical mechanics, the probability that the thermal mode is excited to the nth energy level is

$$p(n) = \frac{\exp(-E_n/kT)}{\sum_n \exp(-E_n/kT)},$$

where k is the Boltzmann constant. This corresponds to the density operator,

$$\rho_{th} = \frac{\exp(-\hat{H}/kT)}{\mathrm{Tr}\left(\exp(-\hat{H}/kT)\right)}, \tag{B.1}$$

where $\hat{H}$ is the Hamiltonian of Eqn. A.13.

Using the fact that $\sum_n |n\rangle \langle n| = \hat{I}$, an expression for ρ_{th} in terms of Fock states is

$$\rho_{th} = \frac{1 - \exp(\hbar\omega/kT)}{\exp(-\hbar\omega/2kT)} \sum_{n=0}^{\infty} \exp(-\hbar\omega n/kT) \, |n\rangle \langle n| \,, \tag{B.2}$$

which tells us that the probability of finding n photons in a thermal state ρ with termperature T is

$$P_T(n) = (1 - \exp(-\hbar\omega/kT)) \exp(-\hbar\omega n/kT). \tag{B.3}$$

With the substitution $\beta = \frac{\hbar\omega}{kT}$, where β functions as an inverse temperature, this simplifies to

$$P_\beta(n) = \left(1 - e^{-\beta}\right) e^{-\beta n}. \tag{B.4}$$

This distribution has a larger width than the Poisson distribution, meaning that for a Poisson distribution and a thermal distribution with small mean and similar values of $P(1)$, the thermal distribution will have a larger number of $n \geq 2$ photon number states.

We can still perform decoy analysis and thereby prove security in QKD, however. This is accomplished by applying the results of Section 3.1.2, replacing instances of the Poisson distribution $P_{\mu_i}(n)$ with corresponding instances of the thermal state photon number distribution $P_\beta(n)$ from Eqn. B.4.