NATIONAL UNIVERSITY OF SINGAPORE

Department of Physics

Optimization of Practical Covert Communication

Author:
Ignatius William
Primaatmaja

Supervisor:
Prof. Valerio Scarani
Mentor:
Dr. Juan Miguel Arrazola

A thesis submitted in partial fulfillment of the requirements for the degree of
Bachelor of Science (Hons)
April 2017

Abstract

Covert communication allows two parties to communicate with low probability of detection by the adversaries. As a trade-off, it usually takes a long time to send a message covertly. In this report, we consider the covert communication scheme proposed in Ref. [2] and perform an optimization over protocol parameters in order to significantly reduce the number of channel uses and the time required to transmit a classical message of fixed length. We propose a model in which the noise in one channel is due to the leakage from another neighboring channel, in which open communication is happening using bright signals. We also consider variations of the implementation and encoding of the protocol.

Acknowledgements

I would like to take this opportunity to express my gratitude to my project supervisor, Valerio, for his guidance and encouragement and also the invaluable advices. Working with him has not only nurtured me as a scientist, but also helped me to grow as a person. I could not express how fortunate I am to have the opportunity to work and form friendships with wonderful people in his group.

Next, I would also like to thank my mentor, Juan Miguel, who has patiently helped me throughout this journey, explaining many notions that I was not familiar with and also clearing any misconceptions that I might have. I cannot thank him enough for the many feedbacks that he has given me, be it in pointing out any calculation mistakes or any advices in communicating the results of this work.

I would also like to appreciate our collaborators from USTC, who have provided us with the details of the experimental parameters. I would also like to thank them for performing the practical demonstration of covert communication.

I am also grateful for the support from my friends and family, especially my parents. Their words of encouragement have helped me to persevere in completing this project. I would like to thank them for lifting me up when I felt depressed due to certain difficulties that I faced during this project.

Last, but not least, I would like to thank God for His Providence, without which, I couldn't finish this project.

Contents

Chapter 1

Introduction

Cryptography is the science of sending secure message in the presence of an adversary. This is usually done by encrypting the message such that the adversary cannot retrieve the message from the ciphertext in a short time. Using one-time pad, this can be done even if the adversary's power is only limited by the laws of physics. Furthermore, the discovery of quantum key distribution (QKD) enables exchange of secret key that is required for the information-theoretic secure encryption [7]. In most cases, encryption based cryptography is sufficient.

However, there are situations where the sole fact that two parties are communicating can be incriminating to them. For example, suppose that Alice is a secret agent who is working as an undercover spy. Even if the content of the communication is encrypted, if anyone detects that she is communicating with her spy agency can be problematic for her mission. In this case, she requires a method for covert communication, which is undetectable by any potential adversaries.

Covert communication has been practiced using different schemes in the past. In Ref. [6], it was shown that secure steganography could be performed by hiding information in the quantum shot-noise of digital photographs. It has also been shown that it is possible to communicate covertly via noisy optical channels [4], with a square-root law stating that $\mathcal{O}(\sqrt{N})$ bits of message can be transmitted given N channel uses. Such communication can even be kept covert against quantum adversaries [3]. More recently, covert communication has been extended into the quantum realm [2], where it was shown that qubits can also be transmitted covertly over noisy channels, while providing a protocol for covert QKD. In the light of these recent developments, it becomes increasingly important to optimize the performance of covert communication protocols in order to bring them closer to practical demonstrations.

In this report, we consider the simple protocols for covert communication of Ref. [2] and perform an optimization over protocol parameters in order to significantly reduce the number of channel uses and the time required to transmit a classical message of fixed length. We consider the situation in which the noise in the channel is due to the leakage from a neighboring channel, in which open communication is happening using bright signals. This permits a control of the noise level in the covert channel by adjusting the brightness of the signals, which leads to the possibility of attaining optimal values for the noise level.

The key parameter in the security proof of covert communication is the detection bias. It is challenging to bound the detection bias analytically, thus we performed a numerical optimization of protocol parameters. Furthermore, due to the nature of covert communication, a significant decoding error is expected. We consider a simple repetition code, which is straightforward to implement in the protocol.

In Chapter 2, we will introduce some notions that will be used extensively throughout this report. In Chapter 3, we will described the protocol that was proposed in Ref. [2]. In Chapter 4, we will optimize the protocol to minimize the number of channel uses and time taken to transmit classical messages. Finally, in Chapter 5, we will consider different variations in the encoding or implementations of the protocol.

Chapter 2

Preliminaries

In this chapter, we introduced some important notions which will be used extensively in the following chapters.

2.1 The density operator

In quantum mechanics, we can represent the state of a quantum system using the density operator ρ which has the following properties [8]:

1. It is positive semi-definite, that is if $\{p_1, p_2, ..., p_N\}$ are the eigenvalues of ρ, then $\forall i$, $p_i \geq 0$.

2. $\operatorname{Tr} \rho = 1$.

Given some pure states $|\psi_i\rangle$, and some real numbers $p_i \geq 0$, the density matrix ρ can be written as

$$\rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \tag{2.1}$$

The real number p_i gives the probability of finding the system in state $|\psi_i\rangle$. ρ is a pure state if and only if there exists i such that $p_i = 1$.

Now, suppose if the system that ρ describes have two subsystems, described by ρ_A and ρ_B, the state ρ can be written in terms of the tensor product of the states of its subsystems

$$\rho = \rho_A \otimes \rho_B \tag{2.2}$$

In particular, if we are interested in M modes of a system, we can decompose its state into

$$\rho = \bigotimes_{i=1}^{M} \rho_i \tag{2.3}$$

where ρ_i is the state of the i-th mode.

2.2 Functions of the density operator

Some quantities in quantum information theory are functions of the density operator. To evaluate this function, first, we have to find the eigenvalues and eigenvectors of the density operator. Suppose that we have done that, we can perform the spectral decomposition of the density operator ρ such that

$$\rho = \sum_i p_i \left|\varphi_i\right\rangle\!\left\langle\varphi_i\right| \tag{2.4}$$

where the eigenvectors are orthogonal, i.e. $\langle\varphi_i|\varphi_j\rangle = \delta_{ij}$. Then a function f of the density operator ρ is given by

$$f(\rho) = \sum_i f(p_i) \left|\varphi_i\right\rangle\!\left\langle\varphi_i\right| \tag{2.5}$$

In particular, when the density matrix is diagonal

$$\rho = \begin{pmatrix} p_1 & & & \\ & p_2 & & \\ & & ... & \\ & & & p_D \end{pmatrix} \tag{2.6}$$

then

$$f(\rho) = \begin{pmatrix} f(p_1) & & & \\ & f(p_2) & & \\ & & ... & \\ & & & f(p_D) \end{pmatrix} \tag{2.7}$$

Sometimes, we are also interested in a function of tensor products of states. For example, we consider the log function which is used frequently in quantum

information theory.

Theorem 2.2.1. *Let $\rho_A \in \mathcal{L}(\mathcal{H}_A)$ and $\rho_B \in \mathcal{L}(\mathcal{H}_B)$ be density operators, then*

$$\log(\rho_A \otimes \rho_B) = \mathbb{1}_A \otimes \log(\rho_B) + \log(\rho_A) \otimes \mathbb{1}_B \tag{2.8}$$

Proof. We express ρ_A and ρ_B in terms of their spectral decompositions

$$\rho_A = \sum_i p_i \left|\psi_i\right\rangle\!\left\langle\psi_i\right|$$
$$\rho_B = \sum_j q_j \left|\varphi_j\right\rangle\!\left\langle\varphi_j\right|$$

Then

$$\begin{aligned}
\log(\rho_A) \otimes \mathbb{1}_B + \mathbb{1}_A \otimes \log(\rho_B) &= \left(\sum_i \log p_i \left|\psi_i\right\rangle\!\left\langle\psi_i\right|\right) \otimes \left(\sum_j \left|\varphi_j\right\rangle\!\left\langle\varphi_j\right|\right) \\
&\quad + \left(\sum_i \left|\psi_i\right\rangle\!\left\langle\psi_i\right|\right) \otimes \left(\sum_j \log q_j \left|\varphi_j\right\rangle\!\left\langle\varphi_j\right|\right) \\
&= \sum_{i,j} \log p_i \left|\psi_i\varphi_j\right\rangle\!\left\langle\psi_i\varphi_j\right| + \sum_{i,j} \log q_j \left|\psi_i\varphi_j\right\rangle\!\left\langle\psi_i\varphi_j\right| \\
&= \sum_{i,j} (\log p_i + \log q_j) \left|\psi_i\varphi_j\right\rangle\!\left\langle\psi_i\varphi_j\right| \\
&= \sum_{i,j} \log(p_i q_j) \left|\psi_i\varphi_j\right\rangle\!\left\langle\psi_i\varphi_j\right| \\
&= \log\left(\sum_i p_i \left|\psi_i\right\rangle\!\left\langle\psi_i\right| \otimes \sum_j q_j \left|\varphi_j\right\rangle\!\left\langle\varphi_j\right|\right) \\
&= \log(\rho_A \otimes \rho_B)
\end{aligned}$$

This completes the proof of Theorem 2.2.1. □

2.3 The quantum relative entropy

The quantum relative entropy (also known as Kullback-Leibler divergence) is a function of two density operators that measures the difference between them. As we shall see later, the quantum relative entropy will determine the detection bias of our protocol, which is an important parameter in the security proof of covert communication. Given two density matrices ρ and σ, the relative entropy is defined as

$$D(\rho||\sigma) = \mathrm{Tr}\left[\rho(\log\rho - \log\sigma)\right] \tag{2.9}$$

Although the relative entropy quantifies how 'far' is one density operator from the other, it is not a distance measure in the mathematical sense since it does not satisfy the triangle inequality neither is it symmetric in its arguments, i.e. in general

$$D(\rho||\sigma) \neq D(\sigma||\rho) \tag{2.10}$$

The relative entropy is rather difficult to calculate since we have to diagonalize the state ρ and σ first. However, there is a special case in which the relative entropy can be calculated easily. This case corresponds to the situation when both density operators are diagonal in a given basis which we will use frequently in this report, as many states that we consider are diagonal in the Fock basis.

Theorem 2.3.1 (**Relative entropy of diagonal density matrices**). *When the density operators ρ and σ are both diagonal in the same basis, then the quantum relative entropy reduces to the classical relative entropy*

$$D(\rho||\sigma) = \sum_n \rho(n) \log\left(\frac{\rho(n)}{\sigma(n)}\right) \tag{2.11}$$

where $\rho(n)$ and $\sigma(n)$ are classical probability distributions.

Proof. We let ρ and σ be diagonal in certain basis

$$\rho = \sum_n \rho(n) \, |n\rangle\langle n|$$
$$\sigma = \sum_n \sigma(n) \, |n\rangle\langle n|$$

Then the relative entropy is given by

$$\begin{aligned} D(\rho||\sigma) &= \mathrm{Tr}\left[\sum_n \rho(n) \log\left(\frac{\rho(n)}{\sigma(n)}\right) |n\rangle\langle n|\right] \\ &= \sum_{n'} \langle n'| \left[\sum_n \rho(n) \log\left(\frac{\rho(n)}{\sigma(n)}\right) |n\rangle\langle n|\right] |n'\rangle \\ &= \sum_{n',n} \rho(n) \log\left(\frac{\rho(n)}{\sigma(n)}\right) \langle n|n'\rangle \, \delta_{nn'} \\ &= \sum_n \rho(n) \log\left(\frac{\rho(n)}{\sigma(n)}\right) \end{aligned}$$

□

Furthermore, the quantum relative entropy has the additivity property that will be used extensively in this report

Theorem 2.3.2 (**Additivity of quantum relative entropy**). *Given the states* $\rho_1, \rho_2, \sigma_1, \sigma_2$*, the quantum relative entropy is additive in this sense*

$$D(\rho_1 \otimes \rho_2 || \sigma_1 \otimes \sigma_2) = D(\rho_1 || \sigma_1) + D(\rho_2 || \sigma_2) \tag{2.12}$$

Proof.

$$D(\rho_1 \otimes \rho_2 || \sigma_1 \otimes \sigma_2) = \mathrm{Tr}\Big[(\rho_1 \otimes \rho_2)\Big(\log(\rho_1 \otimes \rho_2) - \log(\sigma_1 \otimes \sigma_2)\Big)\Big]$$

applying Theorem 2.2.1

$$\begin{aligned}
&= \mathrm{Tr}\Big[(\rho_1 \otimes \rho_2)\Big(\mathbb{1}_1 \otimes \log \rho_2 + \log \rho_1 \otimes \mathbb{1}_2 \\
&\qquad\qquad - \mathbb{1}_1 \otimes \log \sigma_2 - \log \sigma_1 \otimes \mathbb{1}_2\Big)\Big] \\
&= \mathrm{Tr}\Big[(\rho_1 \otimes \rho_2)\Big(\mathbb{1}_1 \otimes (\log \rho_2 - \log \sigma_2) + \log(\rho_1 - \log \sigma_1) \otimes \mathbb{1}_2\Big)\Big] \\
&= \mathrm{Tr}[\rho_1 \otimes \rho_2(\log \rho_2 - \log \sigma_2) + \rho_1(\log \rho_1 - \log \sigma_1) \otimes \rho_2] \\
&= \mathrm{Tr}[\rho_1(\log \rho_1 - \log \sigma_1)]\,\mathrm{Tr}\,\rho_2 + \mathrm{Tr}\,\rho_1\,\mathrm{Tr}[\rho_2(\log \rho_2 - \log \sigma_2)] \\
&= D(\rho_1 || \sigma_1) + D(\rho_2 || \sigma_2)
\end{aligned}$$

Since $\mathrm{Tr}\,\rho_1 = 1 = \mathrm{Tr}\,\rho_2$. □

We can deduce a simple corollary by repetitive applications of Theorem 2.3.2. Given the states ρ and σ and an integer N, we have

$$D(\rho^{\otimes N} || \sigma^{\otimes N}) = N D(\rho || \sigma) \tag{2.13}$$

This corollary is important since when both states are independent and identically distributed (i.i.d.), to calculate the relative entropy between the two states, it is sufficient to consider only two states which have much smaller dimension.

Next, we will give two theorems that we will also use in this report. But before that, we will give a definition of a convex function.

Definition (**Convex function**). *A function* $f(x)$ *is convex over* (a, b) *if for all* $x_1, x_2 \in (a, b)$ *and* $0 \leq \lambda \leq 1$*,*

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda) f(x_2) \tag{2.14}$$

Theorem 2.3.3 (**Joint convexity of quantum relative entropy**). *Let $p_X(x)$ be a probability distribution over finite alphabet $\mathcal{X}$. Let ρ_x and σ_x be density matrices for all $x \in \mathcal{X}$. Set $\bar{\rho} = \sum_x p_X(x)\rho_x$ and $\bar{\sigma} = \sum_x p_X(x)\sigma_x$. The quantum relative entropy is jointly convex in its arguments.*

$$\sum_x p_X(x) D(\rho_x||\sigma_x) \geq D(\bar{\rho}||\bar{\sigma}) \tag{2.15}$$

In particular, when $\mathcal{X} = \{1, 2\}$ and $p_X(x) = \{p, 1-p\}$, then

$$pD(\rho_1||\sigma_1) + (1-p)D(\rho_2||\sigma_2) \geq D(p\rho_1 + (1-p)\rho_2||p\sigma_1 + (1-p)\sigma_2) \tag{2.16}$$

The second theorem is the Pinsker's inequality which relate the relative entropy with a distance measure called the trace distance.

Definition (**Trace distance**). *Given two operators ρ, σ, the trace distance $T(\rho, \sigma)$ is as follows*

$$T(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\| = \frac{1}{2}\,\mathrm{Tr}\,\sqrt{(\rho-\sigma)^\dagger(\rho-\sigma)} \tag{2.17}$$

Unlike the relative entropy, the trace distance is symmetric in its arguments and it satisfies the triangle inequality. It is also non-negative. Thus, it has the metric property. The quantity $\|\rho - \sigma\|$ is sometimes referred as the trace norm. The trace distance is an important parameter in quantum state discrimination since the minimum probability of error in distinguishing state ρ from state σ is given by [5]

$$P_{\text{min error}} = \frac{1}{2}\Big[1 - T(\rho, \sigma)\Big] \tag{2.18}$$

Theorem 2.3.4 (**Pinsker's inequality**). *Let ρ and σ be density operators, then*

$$D(\rho||\sigma) \geq \frac{1}{2\ln 2}\|\rho - \sigma\|^2 \tag{2.19}$$

As such, the Pinsker's inequality provides the bound on the probability of error in distinguishing ρ from σ in terms of relative entropy. This method is also used when we bound the detection bias in covert communication.

The proofs for Theorem 2.3.3 and 2.3.4 are given in Section 11.9 of Ref. [10]

2.4 Some states of light

There will be some states of light which will be used in this report, they are Fock states, thermal states, coherent states.

2.4.1 The Fock state

The Fock state $|n\rangle$ is the state of well defined photon number n. It is the eigenstate of the number operator $\hat{N}$ with eigenvalue n

$$\hat{N} = \sum_i a_i^\dagger a_i \tag{2.20}$$

where a_i is the annihilation operator corresponding to the i-th mode. Since $\hat{N}$ is Hermitian, then for any n' and n, we have

$$\langle n'|n\rangle = \delta_{nn'} \tag{2.21}$$

The state $|0\rangle_i$ is the ground state the i-th mode. Any state $|n\rangle_i$ can be expressed in terms of the ground state

$$|n\rangle_i = \frac{1}{\sqrt{n!}} \left(a_i^\dagger\right)^n |0\rangle_i \tag{2.22}$$

A state that we use quite often in this report is the vacuum state $|vac\rangle$, which is what we have if every mode is in the ground state

$$|vac\rangle = ... \otimes |0\rangle \otimes |0\rangle \otimes |0\rangle \otimes ... \tag{2.23}$$

2.4.2 Thermal states

The thermal state can be expressed in terms of the Fock basis as

$$\rho_{\bar{n}} = \sum_{n=0}^{\infty} \frac{\bar{n}^n}{(1+\bar{n})^{n+1}} |n\rangle\langle n| \tag{2.24}$$

where $\bar{n}$ is the mean photon number of the state.

The thermal state is significant in the study of covert communication because it describes the photon number distribution of the blackbody radiation which is naturally present. In fact, the mean photon number $\bar{n}$ is related to the temperature T of the system by

$$\bar{n} = \frac{1}{\exp(\hbar\omega/k_B T) - 1} \tag{2.25}$$

where ω is the angular frequency of the mode, $\hbar$ is the Planck constant and k_B is the Boltzmann constant. Unfortunately, in room temperature, the mean photon number of the thermal state is too low for practical covert communication.

2.4.3 Coherent states

The coherent state $|\alpha\rangle$ is the eigenstate of the annihilation operator a with eigenvalue α

$$a\,|\alpha\rangle = \alpha\,|\alpha\rangle \tag{2.26}$$

where α is a complex number.

In the Fock basis, it can be written as

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{2.27}$$

and when the phase of the coherent state is randomized, we obtain a mixed state

$$\rho = \sum_{n=0}^{\infty} \frac{e^{-\mu}\mu^n}{n!} |n\rangle\langle n| \tag{2.28}$$

which corresponds to the Poisson distribution of Fock states. The mean of the Poisson distribution μ corresponds $|\alpha|^2$.

The coherent state is significant because many experiments use lasers which are approximately coherent state sources. In this work, we will use mainly coherent states with phase randomization for both signal and the noise.

Chapter 3

Quantum Covert Communication

In this chapter, we outline the covert communication scheme proposed by Arrazola and Scarani in Ref. [2]. This scheme allows covert transmission of qubits using single photon or coherent state encoding. We shall focus on the security proof of the protocol, however, the notions that are introduced in this chapter will be used extensively in the following chapters.

3.1 A model of noisy channel

Consider the setup illustrated in Fig. 3.1. In this model, Alice and Bob are connected by a noisy optical channel, which is modeled as a beam splitter of transmittivity η. The noise is originated from Alice's lab (represented by the red box). Eve has no control over the noise level coming into the channel.

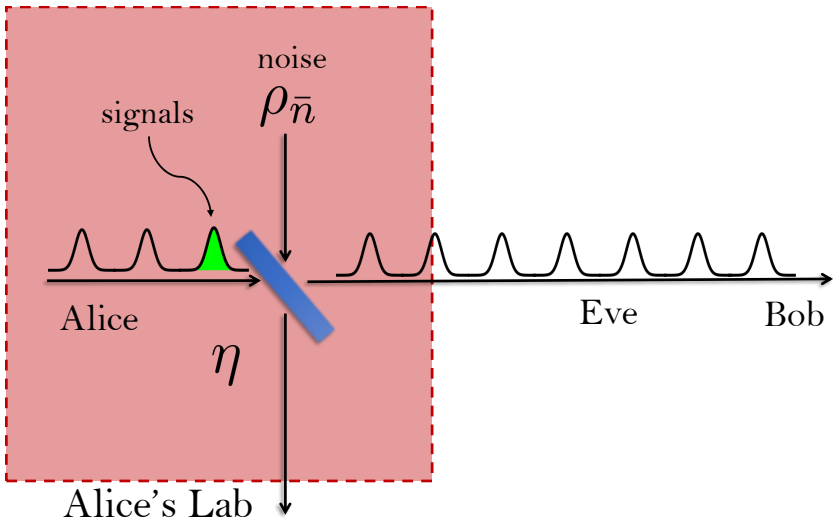


Figure 3.1: A model of the optical channel where the noise comes from Alice's lab

For concreteness, let Alice encodes her qubit states into the polarization degree of freedom of single photons. Suppose further that Alice and Bob have access to N time-bins, each of which may be used to send a qubit signal, and let the noise be thermal (the formalism does not rely on these assumptions and can be applied to other methods of implementation). The beam splitter reflects some of the noise into the channel such that Eve's two-mode state when Alice is not sending any signal to Bob is given by

$$\rho_E = \rho_{\bar{n}'} \otimes \rho_{\bar{n}'} \tag{3.1}$$

where $\bar{n}' = (1-\eta)\bar{n}$ is the mean photon number of the thermal state.

When a signal is sent, let the two-mode state after the beam splitter be ρ_S. When there is communication in the channel, for each time-bin, Alice will decide with constant probability $q \ll 1$, whether or not she will send a signal to Bob. Therefore, Eve's two-mode state when there is a communication in the channelis given by

$$\sigma_E = q\rho_S + (1-q)\rho_E \tag{3.2}$$

Now, since Alice and Bob have access to N time-bins, Eve's $2N$-mode states are simply the tensor products of the two-mode states

$$\rho = (\rho_E)^{\otimes N} \tag{3.3}$$

$$\sigma = (\sigma_E)^{\otimes N} \tag{3.4}$$

where ρ is the state when there is no communication and σ is the state when Alice and Bob are trying to communicate covertly.

3.2 Detection bias

Eve might still attempt to detect whether or not there is any communication between Alice and Bob. To prove the security of the protocol, we consider Eve's detection error, which is her probability of incorrectly guessing whether there is a communication in the channel. This probability is given by

$$P_e = \frac{1}{2} - \epsilon \tag{3.5}$$

where we want $\epsilon > 0$ to be small (we do not want Eve to perform much better than a random guess). We refer ϵ as the *detection bias*.

Now, we want to get a bound on the detection bias. Eve's attempt to determine whether or not there is a communication in the channel is equivalent to quantum state discrimination between ρ and σ. The probability of error in achieving this task is given by [5]

$$P_e \geq \frac{1}{2} - \frac{1}{4}\|\rho - \sigma\| \tag{3.6}$$

where $\|\rho - \sigma\|$ is the trace norm

Comparing (3.5) and (3.6), the bound on the detection bias is given by

$$\epsilon \leq \frac{1}{4}\|\rho - \sigma\| \tag{3.7}$$

and using Pinsker's inequality from Theorem 2.3.4

$$\|\rho - \sigma\| \leq \sqrt{2\ln 2}\sqrt{D(\rho||\sigma)} \tag{3.8}$$

and we can obtain the bound on the detection bias in terms of the relative entropy

$$\epsilon \leq \frac{1}{4}\|\rho - \sigma\|$$

using Pinsker's inequality

$$\leq \sqrt{\frac{\ln 2}{8}D(\rho||\sigma)}$$

since $\ln 2 < 1$

$$\leq \sqrt{\frac{1}{8}D(\rho||\sigma)}$$

and by additivity

$$= \sqrt{\frac{N}{8}D(\rho_E||\sigma_E)} \tag{3.9}$$

which, given ρ_E and σ_E, can be computed directly.

3.3 Covert communication using coherent states

In the beginning of this chapter, we briefly mentioned that Alice can encode her qubit into the polarization of a single photon. However, the formalism does not rely on this assumption. It is still valid for other implementations of the protocol and there are cases where we want to encode our qubit differently. In particular, it is more practical to use coherent state signals rather than single photon. In this section, we are going introduce the *coherent state mapping* [1], a set of rules to map protocols that use single photon qubit into its coherent state encoding counterpart.

After that, we will consider the upper bound on detection bias for a particular case where we want to send covert signals using coherent state source in the presence of thermal noise.

3.3.1 Coherent state encoding

In general, the state $|\psi\rangle$ of a qubit, defined by the polarization a single photon, can be written as

$$|\psi\rangle = \lambda_1 |H\rangle + \lambda_2 |V\rangle \tag{3.10}$$

where $|H\rangle$ and $|V\rangle$ corresponds to a single photon in horizontal and vertical polarization modes respectively. We can map this single photon state into the coherent state

$$|\alpha, \psi\rangle = |\alpha\lambda_1\rangle_H \otimes |\alpha\lambda_2\rangle_V \tag{3.11}$$

where the value of α is a free parameter of the mapping and can be chosen freely to suit the requirements of the protocol. $|\alpha|^2 = \mu$ is the mean photon number of the coherent state and the states $|\alpha\lambda_1\rangle_H$ and $|\alpha\lambda_2\rangle_V$ are coherent states with parameters $\alpha\lambda_1$ and $\alpha\lambda_2$ in their own respective polarization modes. Intuitively, we map the original Hilbert space of the qubit protocol with canonical basis $\{|H\rangle, |V\rangle\}$ into orthogonal optical modes with corresponding annihilation operators $\{a_H, a_V\}$ that corresponds to each polarization mode.

3.3.2 Approximating the relative entropy

In this project, we will consider coherent states with phase-randomization which correspond to Poisson distribution of Fock states. For such coherent state with mean photon number μ, the density matrix is given by

$$\rho_\mu = \sum_{n=0}^{\infty} \frac{e^{-\mu}\mu^n}{n!} |n\rangle\langle n| \tag{3.12}$$

where $|n\rangle$ is the Fock state that corresponds to n photons.

The state for the signal can be calculated by considering the different numbers of photons that come from the 'Alice' and the 'noise' input modes and then taking partial trace. The state is given by

$$\rho_S = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} p(k,l)\sigma_{kl} \tag{3.13}$$

where σ_{kl} is the state when there is k photons coming from the 'noise' input mode and l photons coming from the 'Alice' input mode and $p(k,l)$ is the probability of that occurring. The calculation of σ_{kl} is given in Appendix A whereas $p(k,l)$ is given by

$$p(k,l) = \frac{\bar{n}^k}{(1+\bar{n})^{k+1}} \frac{e^{-\mu}\mu^l}{l!} \tag{3.14}$$

since we assume independence between the number of photons coming from the 'noise' input mode and the number of photons coming from the 'Alice' input mode.

Now, calculating ρ_S using (3.13) directly would be impossible to do because there are infinite terms to compute. However, we can assume that the protocol utilizes coherent state with low mean photon number $\mu \ll 1$ but is still much greater than the noise level $\mu \gg \bar{n}$ such that we can approximate the state for the signal as

$$\rho_S \approx \sum_{k=0}^{1} \sum_{l=0}^{2} \frac{\bar{n}^k}{(1+\bar{n})^{k+1}} \frac{e^{-\mu}\mu^l}{l!} \sigma_{kl} \tag{3.15}$$

because $p(k,l) \approx 0$ for higher order terms and therefore it suffices to consider just the first few terms.

The quantum relative entropy when both the noise and signal is diagonal in the Fock basis reduces to classical relative entropy (Theorem 2.3.1)

$$D(\rho_E||\sigma_E) = \sum_{n=0}^{\infty} \rho_E(n) \left[\log \rho_E(n) - \log \sigma_E(n)\right] \tag{3.16}$$

In the scenario that we considered (i.e. when $\bar{n} \ll \mu \ll 1$), for $n > 0$, we have $[\log \rho_E(n) - \log \sigma_E(n)] < 0$ and therefore, to consider the upper bound of $D(\rho_E||\sigma_E)$, it also suffices to only consider the first few terms.

3.4 The upper bound on detection bias

Due to the complicated expression of the relative entropy as well as σ_E, the analytical calculation of the detection bias is rather difficult to do. Nevertheless, we can give the numerical bound on the detection bias by considering a few photons (since the terms with higher photon numbers contribute negatively to the relative entropy, as discussed in the previous section).

We consider the case where Alice sends 20 photons covertly to Bob, using a

coherent state source with mean photon number of $\mu = 10^{-3}$ in the presence of thermal noise with mean photon number $\bar{n} = 10^{-5}$. The plot of the upper bound of detection bias against the number of time-bins is given in Fig. 3.2.

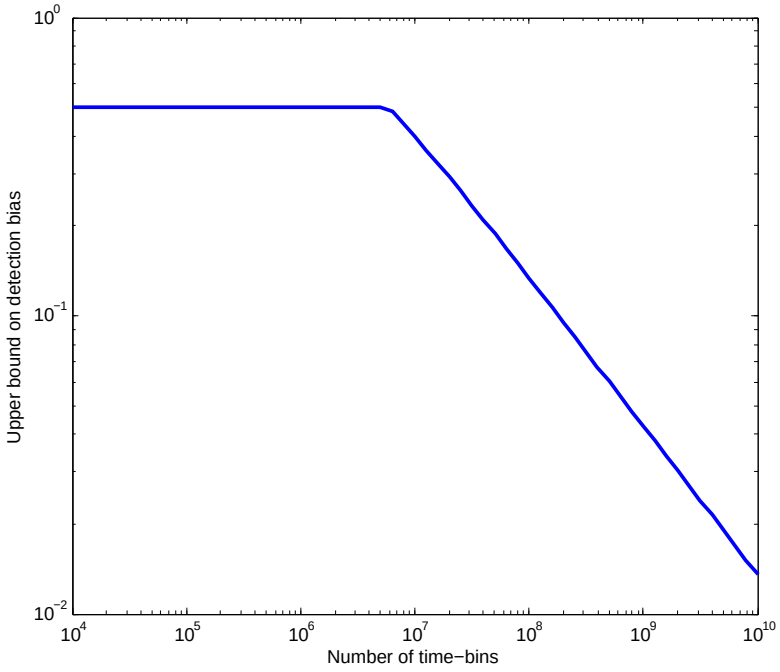


Figure 3.2: Log-log plot of the upper bound of detection bias ϵ as a function of the number of time-bins N. We set $\eta = 0.5$, $\mu = 10^{-3}$ and $\bar{n} = 10^{-5}$. This plot of mine reproduces the plot in Fig. 2 of Ref. [2]

From Fig. 3.2, we can see that the square root law from Ref. [4] also applies to coherent state encoding implementation of covert communication. To reduce the upper bound on detection by one order of magnitude, it takes an increase of approximately two order of magnitude in the number of time bins.

Chapter 4

Practical Covert Communication

Due to the enormous amount of classical post-processing that quantum key distribution (QKD) requires, demonstrating covert QKD with current level of technology is very challenging. In the light of this difficulty, we decided to do practical demonstration of classical covert communication instead. That being said, there are also some challenges in performing classical covert communication that we have to address. These challenges are:

1. The thermal noise in room temperature and in the frequency range that is feasible for practical communication is rather low. To communicate covertly within reasonable duration, we need higher noise level.

2. The low signal-to-noise ratio that is demanded by the nature of covert communication induces high probability of decoding error. As such, we need to do an error correction which in general is detrimental to the security of the protocol. Therefore, we need to find the optimal parameters that achieve both acceptable decoding reliability and detection bias.

To address these challenges, we will consider a protocol where Alice uses Poisson distribution of Fock states for both the signal and the noise, with mean photon number μ and $\bar{n}$ respectively. These states correspond to coherent states with phases-randomization. Alice has full control of both μ and $\bar{n}$ so she can tune them such that she can minimize the "cost" of covert communication. In this chapter, we will consider two different resources as the "cost": the number of time-bins (the number of channel uses) N and the actual time t that Alice and Bob need to spend in the experiment.

4.1 Another model of noisy optical channel

We consider the case where each coherent state source produces photons in two different modes. Suppose that the modes are centered around 1550 and 1560 nm and there are some overlap between the two modes as illustrated by the intensity profiles in Fig. 4.1. The noise in the channel is induced by the overlap.

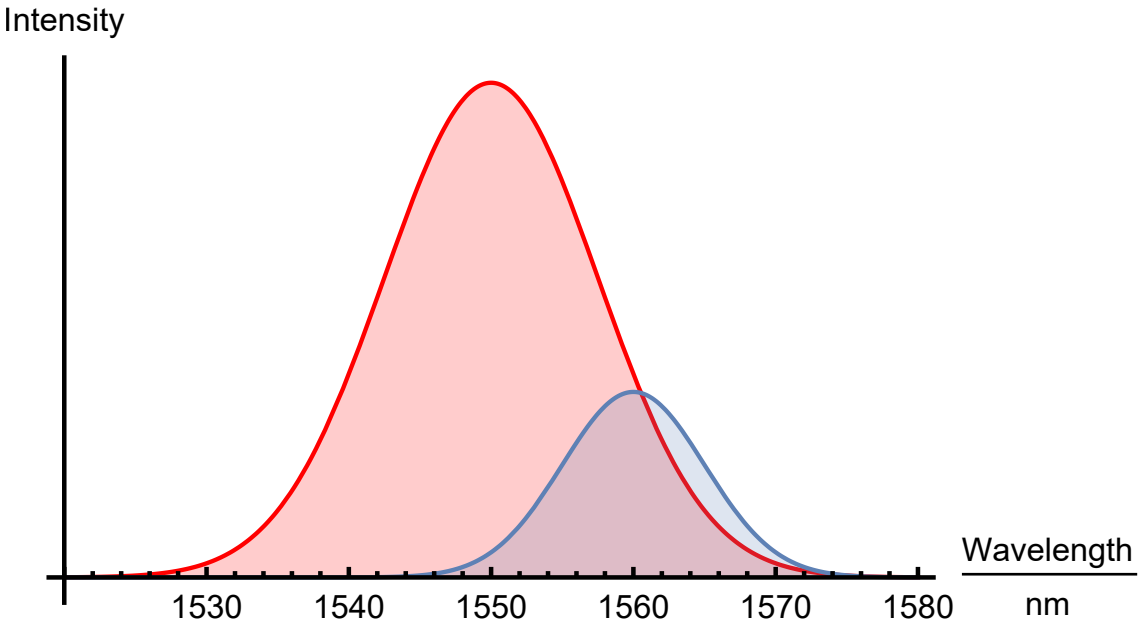


Figure 4.1: The plot of intensity against wavelength of two signals that have different frequency modes, the overlap between the two modes induces a noise in the channel between Alice and Bob which can be used to hide the signal for covert communication.

In Fig. 4.2, we make an adjustment to the model to suit our implementation that uses the induced noise to communicate covertly in the frequency range of the overlap region. In this model, the two modes in Fig. 4.1 correspond to two different communication channels. In the first channel, Alice is communicating non-covertly with Charlie by sending bright signals. Her communication with Charlie induces a background noise in the second channel which connects her to Bob. This background noise allows her to communicate covertly with Bob.

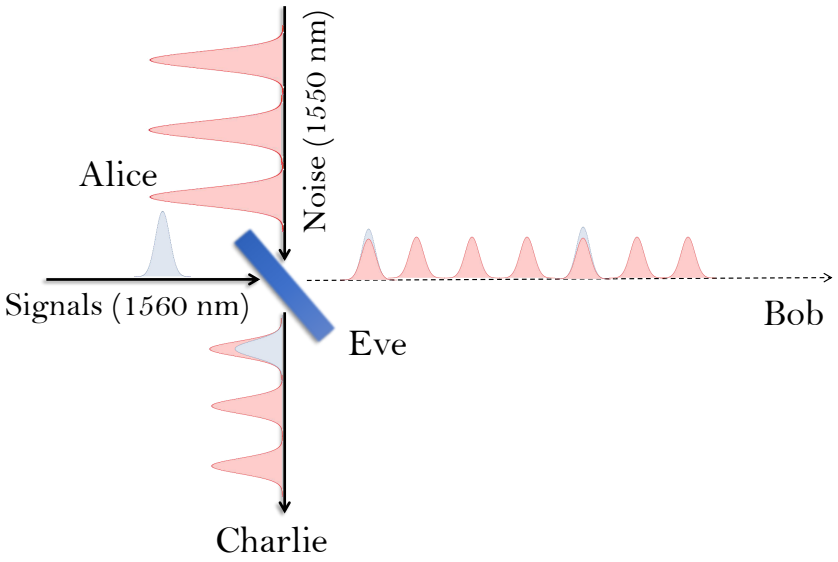


Figure 4.2: A model of noise for covert communication. Conducting open communication with Charlie induces a background noise in the covert channel (represented by the dashed arrow) connecting Alice and Bob. Note that since Alice has full control over both the signal and the noise level.

In this model, the states in the Fock basis are given by

$$\rho_{\bar{n}} = \frac{e^{-\bar{n}}\bar{n}^k}{k!}\,|k\rangle\langle k| \tag{4.1}$$

$$\rho_{(\mu+\bar{n})} = \frac{e^{-(\mu+\bar{n})}(\mu+\bar{n})^k}{k!}\,|k\rangle\langle k| \tag{4.2}$$

and therefore

$$\rho_E = \rho_{\bar{n}} \otimes \rho_{\bar{n}} \tag{4.3}$$

$$\rho_S = \rho_{(\mu+\bar{n})} \otimes \rho_{\bar{n}} \tag{4.4}$$

$$\sigma_E = q\rho_S + (1-q)\rho_E \tag{4.5}$$

where ρ_E is Eve's state when there is no communication and σ_E is the state when there is communication.

4.2 Error correction

In non-ideal situations, the channel connecting Alice and Bob will also be lossy. Due to losses and noise, Bob will observe significant errors in the signals sent by Alice.

To quantify Bob's ability to decode Alice's message correctly, we let Bob's decoding error probability be $\mathcal{E}$. We want this probability to be reasonably low. Suppose that Alice sends a message of length b bits, then Bob's bit error probability δ can be expressed in terms of $\mathcal{E}$ and b

$$\delta = 1 - (1 - \mathcal{E})^{1/b} \tag{4.6}$$

Since for each of the N time-bins, Alice decides with equal probability q whether or not she is going to send a signal, the expected number of signals d is given by

$$d = Nq \tag{4.7}$$

In addition to the loss in the fiber, Bob's detector will also have certain inefficiency. This will result in not all of the photons that Alice sends will be detected. To model this loss, we consider a lossless channel, a perfect detector and a beam splitter with transmitivity τ between them.

The probability p_C that Bob detects a click in the correct mode is

$$\begin{aligned} p_C &= 1 - \Pr[\text{no click in the correct mode}] \\ &= 1 - \exp[-\tau(\mu + \bar{n})] \end{aligned} \tag{4.8}$$

while the probability p_W that Bob detects a click in the wrong mode is

$$\begin{aligned} p_W &= 1 - \Pr[\text{no click in the wrong mode}] \\ &= 1 - \exp(-\tau\bar{n}) \end{aligned} \tag{4.9}$$

Ignoring the probability of having clicks in both modes, the probability of having a click is approximately $(p_C + p_W)$. Then, we get the probability that a click is in the correct mode, given that there is a click is given by

$$p_g = \frac{p_C}{p_C + p_W} \tag{4.10}$$

where p_C and p_W is given by (4.8) and (4.9) respectively.

Now, to correct these errors, we use the repetition code with majority vote decoding scheme. Repetition code R_k repeats every bit k times such that

$$d = bk \tag{4.11}$$

For example, suppose Alice wants to send the bit '0'. Instead of sending one '0' bit, she will send the string

$$\underbrace{'00...00...00'}_{\text{k times}}$$

Next, Bob will look at the string he receives. Bob will decode the bit as '0' if he receives more '0' bits than '1's, otherwise he will decode it as '1'. We then obtain an expression for δ

$$\begin{aligned}\delta &= \sum_{i=0}^{k} \Pr[i \text{ clicks}] \Pr[\text{correct clicks are not majority}] \\ &= \sum_{i=0}^{k} \Pr[i \text{ clicks}] \sum_{j=0}^{\lfloor i/2 \rfloor} \Pr[j \text{ correct cliks } | i \text{ clicks}] \end{aligned} \tag{4.12}$$

The number of clicks C is binomially distributed with probability $(p_C + p_W)$ and number of trials k. Now, since k is large and the probability of registering a click is small, C is also approximately Poisson distributed. We let μ_C and σ_C be the mean and standard deviation of the number of clikcs C respectively. It suffices to consider the range of 5 standard deviations σ_C from the mean μ_C in the first summation in equation (4.12). We get the equation

$$\delta \approx \sum_{i=\mu_C-5\sigma_C}^{\mu_C+5\sigma_C} \binom{k}{i} (p_C + p_W)^i (1 - p_C - p_W)^{k-i} \sum_{j=0}^{\lfloor i/2 \rfloor} \binom{i}{j} (p_g)^j (1 - p_g)^{i-j} \tag{4.13}$$

$$\approx \sum_{i=\mu_C-5\sigma_C}^{\mu_C+5\sigma_C} \underbrace{\frac{(\mu_C)^i e^{-\mu_C}}{i!}}_{\text{Poisson}} \sum_{j=0}^{\lfloor i/2 \rfloor} \underbrace{\binom{i}{j} (p_g)^j (1 - p_g)^{i-j}}_{\text{Binomial}} \tag{4.14}$$

where for Poisson distribution, μ_C and σ_C are given by

$$\begin{aligned}\mu_C &= k \times (p_C + p_W) \\ \sigma_C &= \sqrt{k \times (p_C + p_W)}\end{aligned}$$

Using the above equations and (4.14), we can find the codeword length k that will achieve the desired bit error probability δ.

4.3 Experimental implementation

In this section, we will describe some of the experimental considerations and constraints. To provide an overview of the experiment, we can refer to Fig. 4.3 for

the experiment. We use the 1550 nm laser as the source for the noise and the 1560 nm laser for the signal. Note that while the experiment uses time-bin encoding, our calculation assumes polarization encoding. However, we do not expect significant differences.

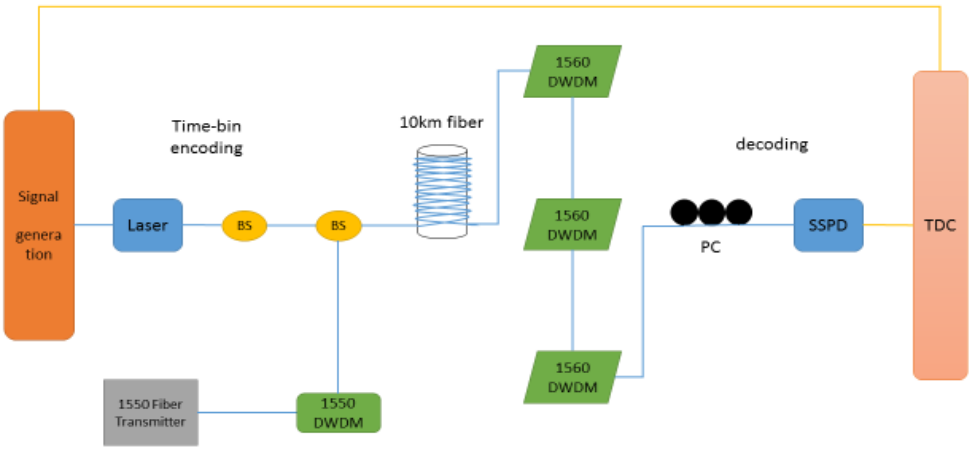


Figure 4.3: Illustration of the experimental setup provided by our collaborators. In this experiment, Alice uses time-bin encoding. We use a sync laser and time-to-digital converter (TDC) to synchronize Alice's and Bob's clock. Both the signal and the noise are passed through a beam splitter and go into the optical fiber. Bob uses the superconducting single-photon detector (SSPD) to detect the photon. Then he decodes the signal by measuring the time of arrival of the photon.

4.3.1 Loss and detector inefficiency

Due the losses in the fiber and detector inefficiency, some of the photons will not be detected. To model this, we consider that there is no loss in the channel and the detector is 100% efficient. However, we will add a beam splitter with transmittivity τ across the channel. Based on the loss and the detector inefficiency, we approximate the channel transmittivity to be $\tau \approx 0.1$.

4.3.2 Noise level

The noise level in the channel is fixed at 10^6 photons/s. However, we can adjust the pulse width such that for each time-bin, we get the noise level that we desire. Thus,

the mean photon number for the noise $\bar{n}$ is given by

$$\bar{n} = 10^6 w \tag{4.15}$$

where w is the pulse width. We also note that the smallest pulse width that we can achieve is $w_{\min} = 1$ ns.

4.3.3 Extinction ratio

The extinction ratio r_e is given by

$$r_e = \frac{\mu_{\text{on}}}{\mu_{\text{off}}} \tag{4.16}$$

where μ_{on} is the number of photons when the laser is on and μ_{off} is the number of photons when the laser is off.

For our experiment, the extinction ratio is $1000 : 1$. This will affect our expression for ρ_E. When we take into account the extinction ratio, it is given by

$$\rho_E = \sum_k \frac{e^{-\bar{n}'} \bar{n}'^k}{k!} |k\rangle\langle k| \tag{4.17}$$

where $\bar{n}' = \bar{n} + 0.001\mu$ and $|k\rangle$ is the Fock state with photon number k.

4.4 Optimization of parameters

Now, we have all the necessary tools to perform our optimization. Since the upper bound on the detection bias is a function of a number of parameters, namely the mean photon number of the signal (μ) and the noise ($\bar{n}$), as well as the probability of sending a signal q. The probability of sending a signal, in turn, depends on the number of of time-bins that are available (N), the length of the message b, and the length of the repetition code (k), which depends on our tolerance on the decoding error. Hence, in principle, optimizing the protocol is an optimization problem over several number of parameters. Since it is challenging to do analytical optimization, we will perform a numerical optimization over all the parameters $(\mu, \bar{n}, q)$.

4.4.1 Number of time-bins as objective function

To find the optimal performance of the protocol, we set a tolerance level for Eve's detection bias and Bob's decoding error. Next, we fix the number of bits that Alice is going to send and search for the values of μ and $\bar{n}$ that will achieve the desired performance using the least number of time-bins. To do so, we have created a numerical routine that outputs optimal values of μ and $\bar{n}$ for any protocol.

We aim to achieve the following performance:

- Number of bits in the message that Alice sends $b = 35$ bits (7 letters).
- Bob's decoding error probability $\mathcal{E} = 0.01$
- Eve's detection bias $\epsilon = 0.1$

Interestingly, when we have high signal-to-noise ratio, i.e. $\mu \gg \bar{n}$, Bob's decoding error is very small, hence Alice can send a smaller number of signals. However, at the same time, high signal-to-noise ratio is also detrimental to the security of the protocol because Eve can easily distinguish the bright signals from the noise. The optimal region actually, lies in the low signal-to-noise ratio, where Eve has difficulty distinguishing signals and noise, but Bob can still, through a large enough code length, retrieve the messages reliably.

A plot of the protocol performance with optimal parameters is shown in Fig. 4.4. As it can be clearly seen, compared to the values suggested in Ref. [2], which rely on thermal background noise at infrared frequencies, our protocol offers a reduction in the number of time-bins of several orders of magnitude.

4.4.2 Time as objective function

In most practical situations, the important resource is time. As such, it may be more important to minimize the duration of the protocol rather than minimizing the number of time-bins (number of channel uses). For a given pulse width w, the total time needed to execute the protocol t is related to the total number of time-bins N by

$$t = Nw \tag{4.18}$$

We also have to keep in mind that w is determined uniquely by equation (4.15). Therefore, in finding the optimal parameters that will minimize the running time of

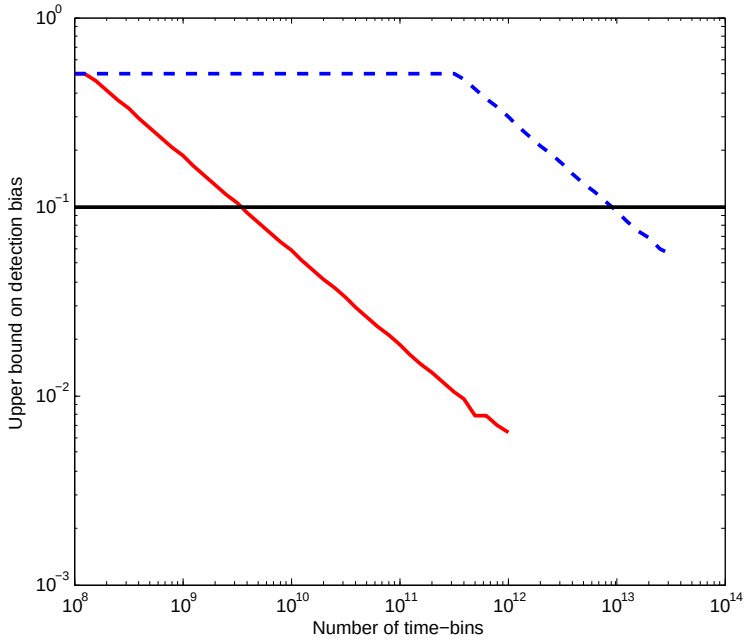


Figure 4.4: Plot of the upper bound on detection bias ϵ against the number of time-bins N. The target detection bias for Eve is $\epsilon = 0.1$, as shown by the black solid line. The red curve corresponds to the optimal parameters: $\bar{n}_{\text{opt}} = 0.640$, $\mu_{\text{opt}} = 0.855$, and $d = 13305$. The dashed blue curve corresponds to $\bar{n} = 10^{-5}$, $\mu = 10^{-3}$ and $q = 3.4023 \times 10^{-7}$, as proposed in Ref. [2].

the experiment, equation (4.15) puts an additional constraint on the noise level $\bar{n}$. After taking into account these changes, we can use a similar numerical routine to find the optimal parameters.

In this optimization, we consider two different cases. In both cases, we tolerate decoding error probability $\mathcal{E} = 0.01$.

1. Long message with higher detection bias

   - Alice sends $b = 35$ bits (7 letters)
   - Eve can only detect communication with detection bias not higher than $\epsilon = 0.1$
   - optimal parameters: $\mu = 0.0180$, $\bar{n} = 1.0 \times 10^{-3}$.

2. Short message with lower detection bias

   - Alice sends $b = 10$ bits (2 letters)

- Eve can only detect communication with detection bias not higher than $\epsilon = 0.01$

In both cases, we obtain the optimal noise level $\bar{n} = 0.001$ as this corresponds to the smallest pulse width that we can achieve $w = 1$ ns. In this optimization, the optimal signal-to-noise ratio is not very low. Due to the physical constraints that force us to operate at much lower $\bar{n}$ and μ (unlike in our first optimization where both $\bar{n}_{\text{opt}}$ and μ_{opt} are relatively high), means that we need higher mean photon numbers for the signal to accommodate the losses in the channel.

The plots for the upper bound on detection bias against the duration of the experiment are given in Fig. 4.5 (for the first case) and Fig. 4.6 (for the second case). In both cases, we can see that the optimal signal-to-noise ratio is much larger than the optimal signal-to-noise ratio when we consider time-bin as our objective function.

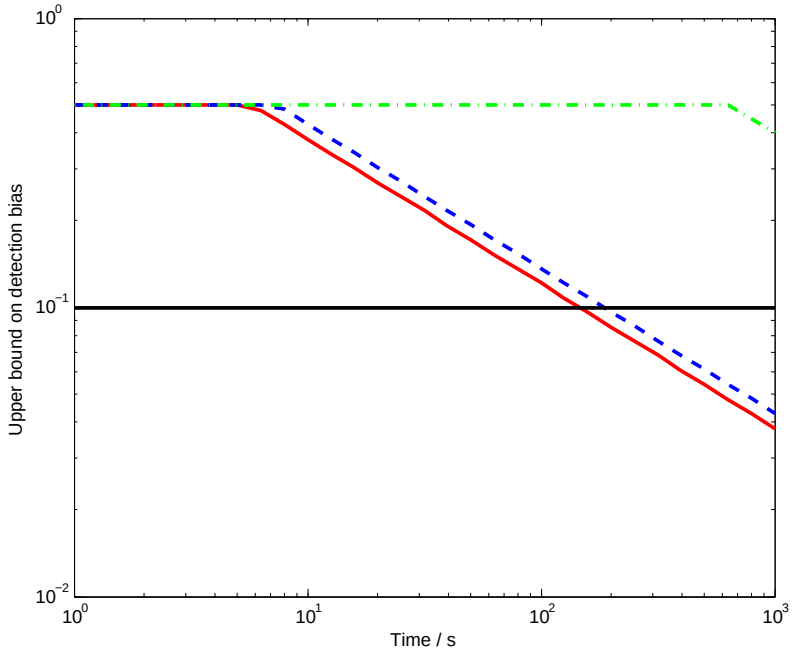


Figure 4.5: Plot of the upper bound on detection bias ϵ against total running time T to send $b = 35$ bits. The target detection bias for Eve is $\epsilon = 0.1$, as shown by the black solid line. The red curve corresponds to the optimal parameters: $\bar{n}_{\text{opt}} = 0.001$, $\mu_{\text{opt}} = 0.018$, and $d = 208357$. Total running time $t = 146$ s. The blue curve corresponds to $\bar{n}_{\text{blue}} = 0.001$ and $\mu_{\text{blue}} = 0.008$ and the green curve corresponds to $\bar{n}_{\text{green}} = 0.001$ and $\mu_{\text{green}} = 0.028$.

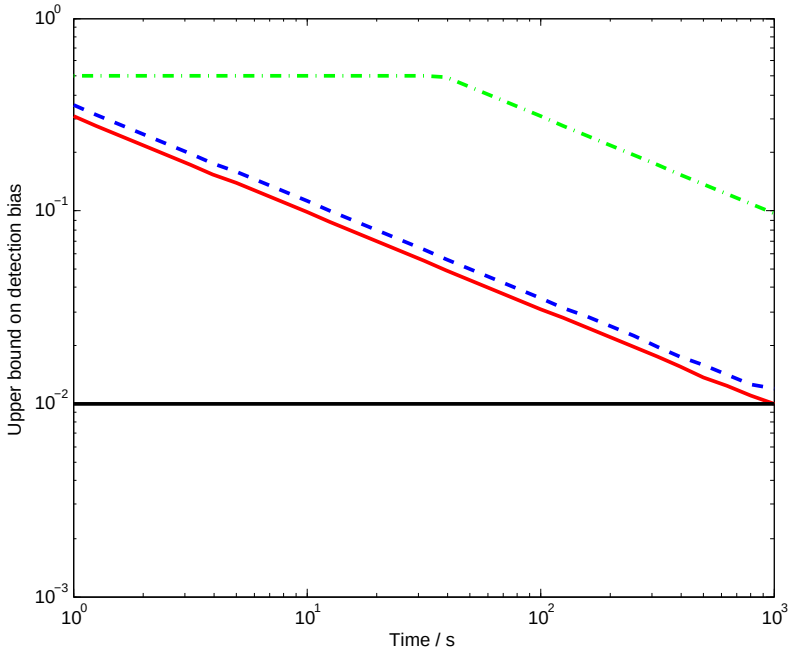


Figure 4.6: Plot of the upper bound on detection bias ϵ against total running time t to send $b = 10$ bits. The target detection bias for Eve is $\epsilon = 0.01$, as shown by the black solid line. The red curve corresponds to the optimal parameters: $\bar{n}_{\text{opt}} = 0.001$, $\mu_{\text{opt}} = 0.0095$, and $d = 107034$. Total running time $t = 875$ s. The blue curve corresponds to $\bar{n}_{\text{blue}} = 0.001$ and $\mu_{\text{blue}} = 0.006$ and the green curve corresponds to $\bar{n}_{\text{green}} = 0.001$ and $\mu_{\text{green}} = 0.012$.

Chapter 5

Other Attempts to Improve Covert Communication Protocols

In the previous chapter, we optimized the protocol by only considering the experimental parameters. In this chapter, we consider other strategies that Alice can use. In other words, we want to consider some modifications to our basic protocol. Here, we consider the case where the number of channel uses is the objective function.

5.1 Block implementation of covert communication

Now, consider the protocol where access to N channel uses corresponds to an access to N time-bins. Suppose Alice and Bob have K blocks, each containing N time-bins. For each block, Alice independently decides with probability p whether she will perform the basic protocol and with probability $(1-p)$, she will do nothing in that block of N time-bins. For each block, let

$$\rho = (\rho_E)^{\otimes N}$$
$$\sigma = (\sigma_E)^{\otimes N}$$

where ρ_E and σ_E is the same as before. Now let σ^* be

$$\sigma^* = p\sigma + (1-p)\rho \tag{5.1}$$

Thus, Eve's $2KN$-mode state is given by

$$\rho' = \rho^{\otimes K} \tag{5.2}$$
$$\sigma' = (\sigma^*)^{\otimes K} = (p\sigma + (1-p)\rho)^{\otimes K} \tag{5.3}$$

The detection bias ϵ for the block protocol is bounded by

$$\begin{aligned}\epsilon \leq \sqrt{\frac{1}{8}D(\rho'||\sigma')} &= \sqrt{\frac{K}{8}D(\rho||\sigma^*)} \\ &= \sqrt{\frac{K}{8}D(\rho||(p\sigma + (1-p)\rho))} \\ &\leq \sqrt{\frac{K}{8}pD(\rho||\sigma)} = \sqrt{\frac{NK}{8}pD(\rho_E||\sigma_E)}\end{aligned} \tag{5.4}$$

where we used the joint convexity of relative entropy, that is for states $\rho_1, \rho_2, \sigma_1, \sigma_2$ and for all $0 \leq p \leq 1$, we have from Theorem 2.3.3

$$D(p\rho_1 + (1-p)\rho_2||p\sigma_1 + (1-p)\sigma_2) \leq pD(\rho_1||\sigma_1) + (1-p)D(\rho_2||\sigma_2) \tag{5.5}$$

and by setting

$$\rho_1 = \rho_2 = \sigma_2 = \rho$$
$$\sigma_1 = \sigma$$

the second term in (5.5) vanishes and we get the inequality that we want.

If we want the expected number of signals d to be the same, we need $p = 1/K$ and thus we get back the bound that we had in the original protocol and there is no improvement.

5.2 Sending multiple bits per photon

Alice can also consider encoding multiple bits into a single photon. She can do this by exploring other degrees of freedom of the photon (for example we can consider both polarization and time of arrival of the photon). By encoding more bits into a photon, she hopes that it would be harder for Eve to detect the communication since Alice can send less number of photons to send the same number of bits.

5.2.1 Eve's states and detection bias

For fair comparison, we assume that Alice and Bob still have access to $2N$ modes. Suppose Alice encodes M bits into one photon. Now, for each channel use, the state when there is no communication

$$\rho'_E = \rho_{\bar{n}}^{\otimes 2^M} \tag{5.6}$$

and the signal state is given by

$$\rho'_S = \rho_{(\mu+\bar{n})} \otimes \rho_{\bar{n}}^{\otimes(2^M-1)} \tag{5.7}$$

where $\rho_{(\mu+\bar{n})}$ is the same as in (4.2). Eve's state when there is communication is therefore given by

$$\sigma'_E = q'\rho'_S + (1-q')\rho'_E \tag{5.8}$$

where q' is given by

$$q' = \frac{bk/M}{2N/2^M} = \frac{2^{M-1}}{M}q \tag{5.9}$$

Eve's $2N$-mode states are given by

$$\rho = (\rho'_E)^{\otimes N/2^{M-1}} \tag{5.10}$$

$$\sigma = (\sigma'_E)^{\otimes N/2^{M-1}} \tag{5.11}$$

Therefore the upper bound on detection bias is given by

$$\epsilon \leq \frac{1}{\sqrt{2^{M-1}}}\sqrt{\frac{N}{8}D(\rho'_E||\sigma'_E)} \tag{5.12}$$

We can check that we obtain the formulae in Chapter 4 when $M = 1$. The factor of $1/\sqrt{2^{M-1}}$ is favourable for covert communication. However, we can easily check that $q' > q$ for $M > 2$. This will increase the relative entropy and therefore the detection bias. Additionally, we have not taken into account any effect on Bob's reliability in decoding the message.

5.2.2 Decoding error probability

Since we operate at multiple modes, Bob's error probability will be different from our basic scheme's. The probability p_C that the detector detects a photon in the

correct mode is still the same

$$\begin{aligned} p_C &= 1 - \Pr[\text{no click in the correct mode}] \\ &= 1 - \exp[-\tau(\mu + \bar{n})] \end{aligned} \tag{5.13}$$

while the probability p_W that the detector detects a photon in one of the wrong modes is given by

$$\begin{aligned} p_W &= 1 - \Pr[\text{no click in any of the wrong modes}] \\ &= 1 - \exp(-\tau\bar{n})^{(2^M-1)} \end{aligned} \tag{5.14}$$

For any $M > 1$, p_W in (5.14) is larger than p_W in (4.9). It follows that the probability p_g that a click is in the right mode, given that there is a click, is smaller.

Meanwhile, the relationship between decoding error probability $\mathcal{E}$ and δ will be

$$\delta = 1 - (1 - \mathcal{E})^{M/b} \tag{5.15}$$

Note that δ no longer of the interpretation as bit error probability since each signal now contains M bits of information.

For error correction, if we use repetition code of length k, to guarantee correct decoding, we demand that there should be at least $(\lfloor k/2 \rfloor + 1)$ clicks in the correct mode. Typically, a smaller number is sufficient to achieve majority, however, in the worst case scenario (where all the errors are contributed by the same mode), we need at least that number of clicks. Therefore, the expression for δ in (4.14) is still valid in this protocol.

Since p_W is larger in this protocol, we need longer codeword to correct the errors. Consequently, we will need more time-bins if we want to keep the probability of sending a signal q unchanged.

5.2.3 Numerical results

Since it is inconclusive from our analysis whether there will be an improvement to the protocol, we will do numerical optimization of the new scheme. We will consider the case where $M = 2$ and $M = 3$ (as q' increases with M, we do not expect improvement in higher values of M) and compare it to the basic protocol ($M = 1$).

Unfortunately, it turns out that the modification does not improve the efficiency of the protocol even after the optimization of the parameters. We observes that increasing the number of bits encoded in one photon actually worsens the performance of the protocol as shown in Fig. 5.1. We learn that it is not enough to increase the length of the repetition code to compromise the increase in p_W. Alice needs to increase her signal-to-noise ratio, which leads to higher detection bias.

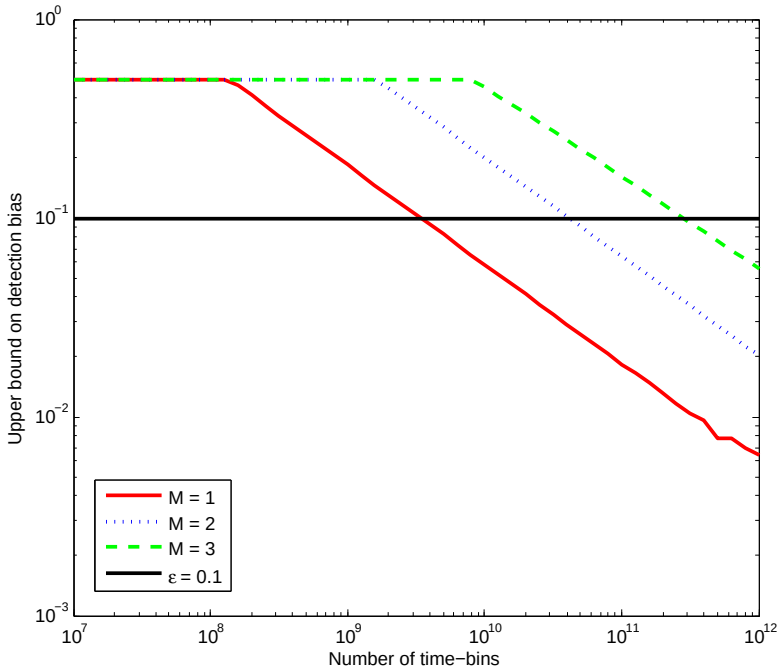


Figure 5.1: The plot of the upper bound on detection bias ϵ against the number of time-bins N to send a message of length $b = 35$ bits. We set the decoding error probability $\mathcal{E} = 0.01$. The red curve corresponds to our basic protocol where we send one bit per photon. The mean photon numbers are $\mu_{\text{red}} = 0.855$ and $\bar{n}_{\text{red}} = 0.640$ for the signal and the noise respectively. The blue dotted curve corresponds to the case when we encode two bits into per photon with $\mu_{\text{blue}} = 0.170$ and $\bar{n}_{\text{blue}} = 0.021$. The green dashed curve corresponds to the case when we encode three bits into per photon with $\mu_{\text{green}} = 0.160$ and $\bar{n}_{\text{green}} = 0.007$. The horizontal black line is the target detection bias $\epsilon = 0.1$.

5.3 Switching-off the noise

In situations when Alice has full control over both the signal and the noise level, like the one discussed in Chapter 4, one could think of turning-off the noise when a signal is sent. In implementation like ours (Fig. 4.1), this may be problematic since the modes are different: by sitting at 1550 nm, Eve could notice that there is no

light when a covert signal is sent. By assuming implementations where the modes are equal, like the one in Chapter 3, does this actually help? Here, we prove that it doesn't.

In this implementation of the protocol, when a signal is sent, Bob will no longer get a click from the noise. However, as the noise does not contribute to the error, now we need to take into account the error coming from the alignment of Bob's polarizing beam splitter (PBS). Suppose Alice sends a photon with horizontal polarization, when Bob's PBS is not perfectly aligned, the PBS sees some vertical component in the photon's polarization. There is a non-zero probability that the photon's polarization is measured as vertical. In this section, we will consider imperfect alignment as the main source of error, which is insignificant in our original scheme.

On the other hand, now Eve can also distinguish the noise level when there is a communication and when there is no communication. Intuitively, this will increase her detection bias. Thus, we need to perform optimization of this protocol and see whether the overall effect will lead to an improvement.

This modification can be applied to both the original protocol and the protocol where Alice encodes multiple bits into one photon. As the main drawback of the multiple bits per photon scheme is the increase in error, this modification can eliminate the problem.

5.3.1 Single bit per photon

When Alice encodes one bit per photon, the expression for signal state is replaced by

$$\rho_S = \rho_\mu \otimes |0\rangle\langle 0| \tag{5.16}$$

where ρ_μ is the coherent state with mean photon number μ and phase-randomization and $|0\rangle$ is the vacuum state.

The state when there is communication in the channel is given by

$$\sigma_E = q(\rho_\mu \otimes |0\rangle\langle 0|) + (1-q)(\rho_{\bar{n}} \otimes \rho_{\bar{n}}) \tag{5.17}$$

To calculate Bob's bit error probability, we model the imperfect alignment as a beam splitter that sometimes reflects a photon to the wrong mode. Let v be the transmittivity of Bob's beam splitter. Then the probability p_C that Bob detects a photon in the correct mode

$$p_C = 1 - \exp(-\tau\mu v) \tag{5.18}$$

while the probability p_W that Bob get a click in the wrong mode is

$$p_W = 1 - \exp(-\tau\mu(1-v)) \tag{5.19}$$

5.3.2 Multiple bits per photon

When Alice encodes M bits into one photon, the expression for the signal state is

$$\rho_S = \rho_\mu \otimes |0\rangle\langle 0|^{\otimes(2^M-1)} \tag{5.20}$$

The probability of sending a photon is given by (5.9).

To compute δ (here, we use the expression of δ in (5.15)), we use the same model for the alignment of Bob's PBS. The expression for the probability of having click in the correct mode (p_C) is the same as (5.18) while to calculate p_W we also consider the following sources of error:

1. Imperfect alignment of the PBS.
2. Dark count of Bob's detector. We let the detector to have a dark count at constant rate of $P_D = 10^{-6}$.

Therefore, the probability of Bob having a click in the wrong mode is given by

$$p_W = 1 - \underbrace{\exp(-\tau\mu(1-v))}_{\text{alignment}}\underbrace{(1-P_D)^{2^M-2}}_{\text{dark count}} \tag{5.21}$$

5.3.3 Evaluation of the protocol

Taking into account the extinction ratio, we change the vacuum state $|0\rangle\langle 0|$ to the coherent state $\rho_{\bar{n}'}$ with mean photon number $\bar{n}' = \bar{n}/r_e = 0.001\bar{n}$. However, after optimization, the overall performance of both protocols (single bits per photon and multiple bits per photon) is still significantly below the basic protocol. We conclude

that the effect of switching-off the noise on the detection bias is more significant than its effect on the decoding error. Thus, this modification fails to improve the efficiency of covert communication.

5.4 Multiplexing

Multiplexing is a method by which the multiple signals are combined into one signal over shared medium. A common technique of multiplexing is the frequency-division multiplexing (or sometimes called the wavelength-division multiplexing). Frequency-division multiplexing (FDM) allows a number of signals to be modulated onto different carrier frequencies and then carried simultaneously. It is used widely in our day-to-day communication technologies, such as radio and television set. To perform frequency-division multiplexing, the bandwidth of the medium must exceed the bandwidth of the signals. Secondly, the carrier frequencies must be well separated such that the bandwidth of the signals do not overlap [9]. In this section, we are looking of the possibility of implementing frequency-division multiplexing of the noise in covert communication to shorten the duration of the experiment.

The noise level in the channel connecting Alice and Bob depends on the profile of the intensity of the source as a function of wavelength by which they communicate. For concreteness, we assume that Alice and Bob communicates in two different carrier wavelengths λ_1 and λ_2 and the noise level in these two wavelength is given by

$$\bar{n}_1 \equiv \bar{n}(\lambda_1) \tag{5.22}$$

$$\bar{n}_2 \equiv \bar{n}(\lambda_2) \tag{5.23}$$

Similarly, for the mean photon number of the signal

$$\mu_1 \equiv \mu(\lambda_1) \tag{5.24}$$

$$\mu_2 \equiv \mu(\lambda_2) \tag{5.25}$$

We define the effective number of channel uses T

$$T \equiv \frac{t}{w} \tag{5.26}$$

where t is the duration of the experiment and w is the pulse width.

Now if T_0 is the effective number of channel uses when we do not use FDM, then

$$T_0 = N \tag{5.27}$$

where N is the total number of channel uses. Therefore, when we use FDM, we get

$$T = \frac{N}{2} \tag{5.28}$$

since for every pulse, we use the channel twice (we use both the λ_1 and λ_2 channel).

In general, Alice does not need to send equal number of signals in the two frequencies. In other words, suppose Alice is sending b bits to Bob, she can send ηb bits in λ_1 and $(1-\eta)b$ bits in λ_2 where $0 \leq \eta \leq 1$. Furthermore, for error correction, the length of the codewords sent in λ_1 would be different from the length of the codewords sent in λ_2. Thus, the number of signals sent in each wavelength is given by

$$d_1 = \eta b k_1 \tag{5.29}$$

$$d_2 = (1-\eta) b k_2 \tag{5.30}$$

where d_1 and d_2 are the number of signals sent in λ_1 and λ_2 with k_1 and k_2 as the length of the codewords in the respective wavelength. Then, the probability of sending a signal in each wavelength is given by

$$q_1 = \frac{d_1}{N/2} \tag{5.31}$$

$$q_2 = \frac{d_2}{N/2} \tag{5.32}$$

Now, let

$$\rho_1 = \rho_{\bar{n}_1} \otimes \rho_{\bar{n}_1} \tag{5.33}$$

$$\rho_{S1} = \rho_{(\mu_1+\bar{n}_1)} \otimes \rho_{\bar{n}_1} \tag{5.34}$$

$$\rho_2 = \rho_{\bar{n}_2} \otimes \rho_{\bar{n}_2} \tag{5.35}$$

$$\rho_{S2} = \rho_{(\mu_2+\bar{n}_2)} \otimes \rho_{\bar{n}_2} \tag{5.36}$$

where $\rho_{\bar{n}_i}$ and $\rho_{(\mu_i+\bar{n}_i)}$ are coherent states with mean photon number $\bar{n}_i$ and $(\mu_i+\bar{n}_i)$

respectively and $i = 1, 2$. We also let

$$\sigma_1 = q_1\rho_{S1} + (1 - q_1)\rho_1 \tag{5.37}$$

$$\sigma_2 = q_2\rho_{S2} + (1 - q_2)\rho_2 \tag{5.38}$$

Then, we can compute Eve's states

$$\rho = \rho_1^{\otimes N/2} \otimes \rho_2^{\otimes N/2} \tag{5.39}$$

$$\sigma = \sigma_1^{\otimes N/2} \otimes \sigma_2^{\otimes N/2} \tag{5.40}$$

Thus, the detection bias is bounded by

$$\epsilon \leq \sqrt{\frac{1}{8}\frac{N}{2}\Big(D(\rho_1||\sigma_1) + D(\rho_2||\sigma_2)\Big)} \tag{5.41}$$

By the same token, we can further generalize the protocol to allow communication in X wavelengths by setting

$$d_i = \eta_i b k_i \tag{5.42}$$

$$q_i = \frac{d_i}{N/X} \tag{5.43}$$

$$\epsilon \leq \sqrt{\frac{1}{8}\frac{N}{X}\sum_{i=1}^{X} D(\rho_i||\sigma_i)} \tag{5.44}$$

$$T = \frac{N}{X} \tag{5.45}$$

where $i = 1, 2, ..., X$ and $\sum_i \eta_i = 1$. However, considering the number of parameters that are involved, we will not optimize the protocol for the most general case. Instead, we will optimize the protocol in which Alice and Bob communicate in two wavelengths simultaneously. In this case, the parameters that we need to consider are $\eta, \bar{n}_1, \bar{n}_2, \mu_1, \mu_2$.

For optimal performance of the protocol, we demand that

$$\bar{n}_1 = \bar{n}_{\text{opt}} \tag{5.46}$$

$$\mu_1 = \mu_{\text{opt}} \tag{5.47}$$

where $\bar{n}_{\text{opt}}$ and μ_{opt} are the optimal mean photon number of the noise and the signal in the basic protocol.

Furthermore, since λ_1 is the more efficient wavelength for the communication, we want to send more bits in λ_1 than in λ_2. We can set

$$\begin{aligned} \eta &\geq 1 - \eta \\ \Rightarrow \eta &\geq \frac{1}{2} \end{aligned} \tag{5.48}$$

In general, whether multiplexing leads to an improvement in the protocol depends on how the noise level depends on the wavelengths in which the communication take place. In this report, we do not consider a particular function which governs the relationship of the noise level with the wavelength associated to the channel. Instead, we consider how different noise level and evaluate what is the range of noise level, in which multiplexing is suitable.

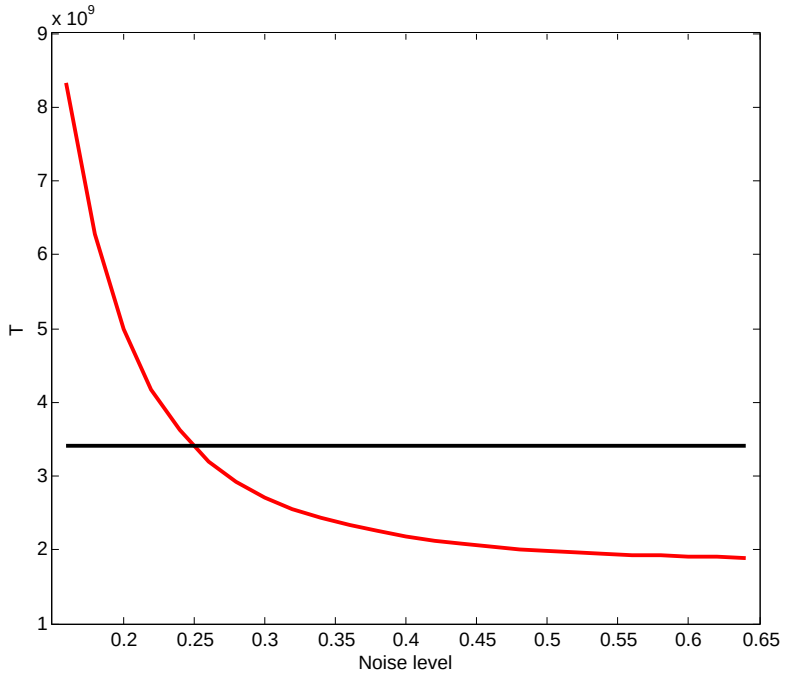
Again, we consider the case where Alice sends $b = 35$ bits covertly, with Eve's detection bias of $\epsilon = 0.1$ and Bob's decoding error probability $\mathcal{E} = 0.01$. As some may have expected, the optimal case is when $\bar{n}_2 = \bar{n}_{\text{opt}}$ and $\mu_2 = \mu_{\text{opt}}$ as well with $\eta = 0.5$. In this case, we have

$$T = 1.8951 \times 10^9 \approx \frac{N_{\text{opt}}}{2} \tag{5.49}$$

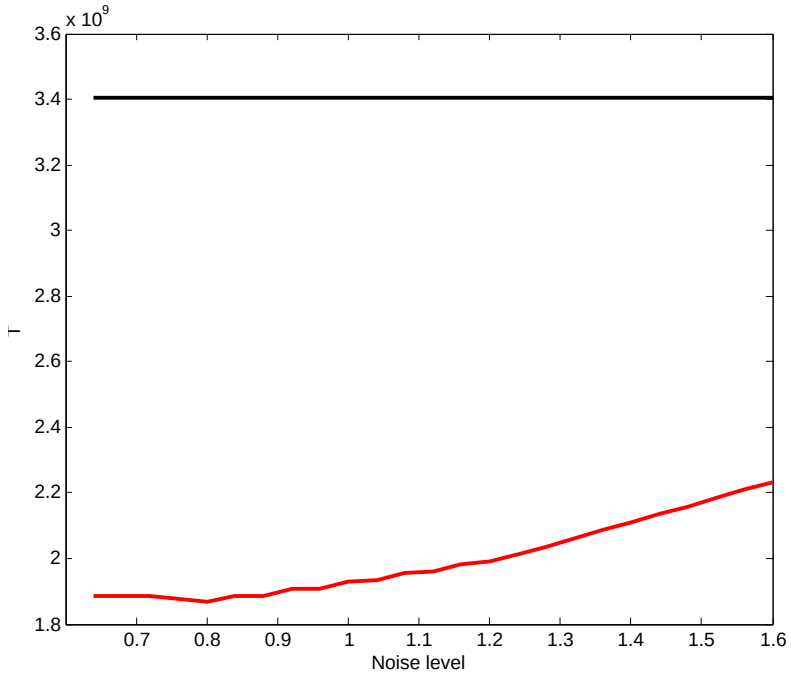
where N_{opt} is the optimal number of channel uses in the original protocol.

We may also be interested to know whether we still get an improvement when the $\bar{n}_2$ differs from the optimal noise level. Suppose we can achieve $\mu_2 = \mu_{\text{opt}}$ and setting $\eta = 0.5$, we want to know how much deviation from optimal noise level can we tolerate while still having an improvement from multiplexing. In Fig. 5.2, we plot the graph of the effective number of channel uses (T) against the noise level in λ_2 channel ($\bar{n}_2$).

We can see that multiplexing can lead to improvement in relatively large range of noise level (from less than half of the optimal noise level to more than twice of the optimal noise level). Although this result may sound promising, whether multiplexing is practically feasible also depends on the relationship of the noise level and the wavelength in which the communication take place.



(a) $\bar{n}_2 < \bar{n}_{\text{opt}}$



(b) $\bar{n}_2 > \bar{n}_{\text{opt}}$

Figure 5.2: The plot of effective number of channel uses (T) against the noise level in the λ_2 channel ($\bar{n}_2$) to send 35 bits covertly with detection bias $\epsilon = 0.1$ and decoding error probability $\mathcal{E} = 0.01$. The black horizontal line represents the optimal number of channel uses $N_{\text{opt}} \approx 3.4 \times 10^9$ in the basic protocol.

Chapter 6

Conclusion

6.1 Summary

The recent developments in covert communication has made it very important to optimize the performance of covert communication protocols in order to bring them closer to practical demonstrations. To do that, we propose a model where the noise in Alice-Bob channel arises from the leakage in Alice-Charlie channel where an open communication is taking place. Using this model, we minimize the number of channel uses and the time taken to send a classical message covertly.

Our method of optimization is as follows. First, we set a target for detection bias and tolerance on Bob's decoding error probability. Next, we perform a numerical routine to find the experimental parameters that will minimize the number of time-bins or the time taken to perform the experiment. We discovered that the optimal signal-to-noise ratio that will minimize the number of channel uses is relatively low. This causes difficulty for Eve to distinguish the signal-to-noise ratio, but Bob can still decode the message reliably through sufficiently large code length. On the other hand, the optimal mean photon number of both the signal and the noise is relatively small when we want to minimize the duration of the protocol. This is because the lower noise level corresponds to smaller pulse width, which is an important factor which affects the duration of the experiment. To compromise the loss in the fiber, we need a sufficiently high mean photon number for the signal. Thus, we have a slightly higher signal-to-noise ratio compared to the case where we want to minimize the number of channel uses.

Finally, we considered variations on the encoding or implementations of the protocol but, other than multiplexing, none of those provide any advantage over the

basic scheme.

6.2 Future directions

In this project, we consider the implementation of covert communication in blocks where Alice decides in i.i.d. manner whether she will perform the protocol in that block. This scheme is easy to analyze, however Alice can, in principle, use a different strategy in which she randomly choose only one block where she will perform the covert communication protocol. Although it would be more difficult to analyze the performance of such protocol, since the state will not be of tensor products structure, it is interesting to know whether this modification will lead to an improvement.

In our model, we consider noise that is induced by communication using bright signals in the neighbouring channel. Although this gives Alice the ability of controlling the noise level in the channel, such source of noise is rather artificial. The fact that Alice and Bob is connected by an optical fiber may also be incriminating to them. We can consider a covert communication protocol that is performed in free space subjected to the noise naturally present in daylight. The difficulty in this protocol is in modeling the state for daylight as well as the usual challenges in free space communication.

So far, we only considered the covert transmission of classical information. It would also be interesting to perform an optimization of quantum covert communication protocol. For example, one could look at the signals as encoding a qubit and perform state tomography to reconstruct its state.

Appendix A

Calculation of the Signal State

In this appendix, we present the explicit calculation of the signal state ρ_S. The signal state ρ_S can be computed directly by considering Eve's state when there is k photons coming from the noise and l photons coming from the signal. We denote that state by σ_{kl}. As discussed in section 3.3.2, it is sufficient to consider only few number of photons, so we will only consider $k \leq 1$ and $l \leq 2$.

We first consider the input and output modes of the beam splitter in the model (Fig 3.1). For simplicity, we associate those modes with their corresponding creation operators $\{a_A^\dagger, a_N^\dagger, a_E^\dagger, a_L^\dagger\}$ and we let the transmittivity of the beam splitter be η, as illustrated in Fig A.1.

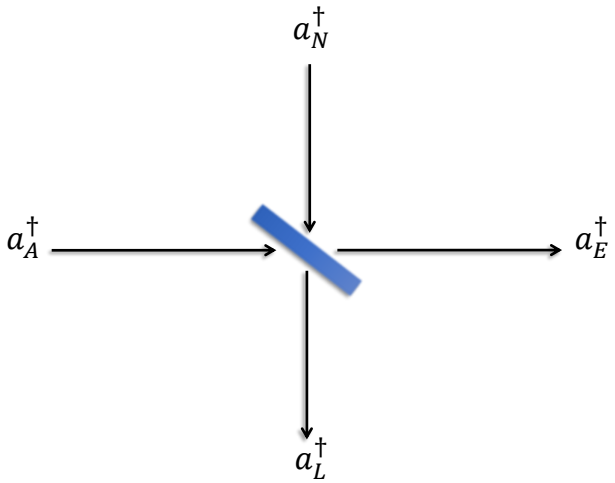


Figure A.1: The beam splitter (with transmittivity η) in Fig 3.1 with its input and output modes

The transformation of the beam splitter can be written as

$$a_A^\dagger \to \sqrt{\eta} a_E^\dagger + i\sqrt{1-\eta} a_L^\dagger$$
$$a_N^\dagger \to \sqrt{\eta} a_L^\dagger + i\sqrt{1-\eta} a_E^\dagger$$

Using this transformation rule, we will compute the σ_{kl}'s. In this appendix, we will adopt the convention of $|n_E n_L\rangle$ for the Fock state describing n_E photons in the E mode and n_L photons in the L mode.

Computing σ_{00} is trivial. Since we have vacuum in both input modes A and N, we will also have vacuum in both output modes E and L. Thus

$$\sigma_{00} = |0\rangle\langle 0|$$

For σ_{01}, we have

$$\begin{aligned} a_A^\dagger |vac\rangle\langle vac| a_A &\to (\sqrt{\eta} a_E^\dagger + i\sqrt{1-\eta} a_L^\dagger) |vac\rangle\langle vac| (\sqrt{\eta} a_E - i\sqrt{1-\eta} a_L) \\ &= \eta |10\rangle\langle 10| + (1-\eta) |01\rangle\langle 01| + i\sqrt{\eta(1-\eta)} |10\rangle\langle 01| \\ &\quad - i\sqrt{\eta(1-\eta)} |01\rangle\langle 10| \\ &\equiv \sigma'_{01} \end{aligned}$$

To obtain σ_{01}, we take the partial trace of σ'_{01} with respect to L.

$$\begin{aligned} \sigma_{01} &= \mathrm{Tr}_L(\sigma'_{01}) \\ &= (\mathbb{1}_E \otimes \langle 0|_L)\sigma'_{01}(\mathbb{1}_E \otimes |0\rangle_L) + (\mathbb{1}_E \otimes \langle 1|_L)\sigma'_{01}(\mathbb{1}_E \otimes |1\rangle_L) \\ &= (1-\eta) |0\rangle\langle 0| + \eta |1\rangle\langle 1| \end{aligned}$$

Therefore, we can see that the effect of taking partial trace is removing the off-diagonal terms of the E-L system, and then only keeping n_E in the remaining $|n_E n_L\rangle\langle n_E n_L|$'s. As such, we will skip the details of the partial trace step for rest of the σ_{kl}'s.

For σ_{02}:

$$\begin{aligned} \sigma_{02} &= \mathrm{Tr}_L \left[\frac{1}{\sqrt{2}} (\sqrt{\eta} a_E^\dagger + i\sqrt{1-\eta} a_L^\dagger)^2 |vac\rangle\langle vac| (\sqrt{\eta} a_E - i\sqrt{1-\eta} a_L) \frac{1}{\sqrt{2}} \right] \\ &= \frac{1}{2} \mathrm{Tr}_L \Big[\left(\eta (a_E^\dagger)^2 + 2i\sqrt{\eta(1-\eta)} a_E^\dagger a_L^\dagger - (1-\eta)^2 (a_L^\dagger)^2 \right) |vac\rangle \\ &\qquad \langle vac| \left(\eta a_E^2 - 2i\sqrt{\eta(1-\eta)} a_E a_L - (1-\eta)^2 a_L^2 \right) \Big] \end{aligned}$$

$$\sigma_{02} = (1-\eta)^2 \left|0\right\rangle\!\left\langle 0\right| + 2\eta(1-\eta) \left|1\right\rangle\!\left\langle 1\right| + \eta^2 \left|2\right\rangle\!\left\langle 2\right|$$

For σ_{10}:

$$\begin{aligned}\sigma_{10} &= \mathrm{Tr}_L \left[(\sqrt{\eta} a_L^\dagger + i\sqrt{1-\eta} a_E^\dagger) \left|vac\right\rangle\!\left\langle vac\right| (\sqrt{\eta} a_L - i\sqrt{1-\eta} a_E) \right] \\ &= \eta \left|0\right\rangle\!\left\langle 0\right| + (1-\eta) \left|1\right\rangle\!\left\langle 1\right|\end{aligned}$$

For σ_{11}:

$$\begin{aligned}\sigma_{11} = \mathrm{Tr}_L \Bigg[& (\sqrt{\eta} a_E^\dagger + i\sqrt{1-\eta} a_L^\dagger)(\sqrt{\eta} a_L^\dagger + i\sqrt{1-\eta} a_E^\dagger) \left|vac\right\rangle \\ & \left\langle vac\right| (\sqrt{\eta} a_E - i\sqrt{1-\eta} a_L)(\sqrt{\eta} a_L - i\sqrt{1-\eta} a_E) \Bigg] \\ = \mathrm{Tr}_L \Bigg[& \Big((2\eta-1) a_E^\dagger a_L^\dagger + i\sqrt{\eta(1-\eta)} (a_L^\dagger)^2 + i\sqrt{\eta(1-\eta)} (a_E^\dagger)^2 \Big) \left|vac\right\rangle \\ & \left\langle vac\right| \Big((2\eta-1) a_E a_L - i\sqrt{\eta(1-\eta)} a_L^2 - i\sqrt{\eta(1-\eta)} a_E^2 \Big) \Bigg] \\ = {} & 2\eta(1-\eta) \left|0\right\rangle\!\left\langle 0\right| + (1-2\eta)^2 \left|1\right\rangle\!\left\langle 1\right| + 2\eta(1-\eta) \left|2\right\rangle\!\left\langle 2\right|\end{aligned}$$

For σ_{12}:

$$\begin{aligned}\sigma_{12} = \mathrm{Tr}_L \Bigg[& \frac{1}{\sqrt{2}} (\sqrt{\eta} a_E^\dagger + i\sqrt{1-\eta} a_L^\dagger)^2 (\sqrt{\eta} a_L^\dagger + i\sqrt{1-\eta} a_E^\dagger) \left|vac\right\rangle \\ & \left\langle vac\right| (\sqrt{\eta} a_L - i\sqrt{1-\eta} a_E)(\sqrt{\eta} a_E - i\sqrt{1-\eta} a_L)^2 \frac{1}{\sqrt{2}} \Bigg] \\ = \frac{1}{2} \mathrm{Tr}_L \Bigg[& \Big(i\eta\sqrt{(1-\eta)} (a_E^\dagger)^3 + (3\eta-2)\sqrt{\eta} (a_E^\dagger)^2 a_L^\dagger + i(3\eta-1)\sqrt{1-\eta} a_E^\dagger (a_L^\dagger)^2 \\ & - (1-\eta)\sqrt{\eta} (a_L^\dagger)^3 \Big) \left|vac\right\rangle\!\left\langle vac\right| \Big(- i\eta\sqrt{(1-\eta)} a_E^3 + (3\eta-2)\sqrt{\eta} a_E^2 a_L \\ & - i(3\eta-1)\sqrt{1-\eta} a_E a_L^2 - (1-\eta)\sqrt{\eta} a_L^3 \Big) \Bigg] \\ = {} & 3\eta(1-\eta)^2 \left|0\right\rangle\!\left\langle 0\right| + (1-\eta)(1-3\eta)^2 \left|1\right\rangle\!\left\langle 1\right| + \eta(2-3\eta)^2 \left|2\right\rangle\!\left\langle 2\right| + 3\eta^2(1-\eta) \left|3\right\rangle\!\left\langle 3\right|\end{aligned}$$

In summary, these are the six σ_{kl}'s that we compute

$$\sigma_{00} = \left|0\right\rangle\!\left\langle 0\right| \tag{A.1}$$

$$\sigma_{01} = (1-\eta) \left|0\right\rangle\!\left\langle 0\right| + \eta \left|1\right\rangle\!\left\langle 1\right| \tag{A.2}$$

$$\sigma_{02} = (1-\eta)^2 \left|0\right\rangle\!\left\langle 0\right| + 2\eta(1-\eta) \left|1\right\rangle\!\left\langle 1\right| + \eta^2 \left|2\right\rangle\!\left\langle 2\right| \tag{A.3}$$

$$\sigma_{10} = \eta \left|0\right\rangle\!\left\langle 0\right| + (1-\eta) \left|1\right\rangle\!\left\langle 1\right| \tag{A.4}$$

$$\sigma_{11} = 2\eta(1-\eta) \left|0\right\rangle\!\left\langle 0\right| + (1-2\eta)^2 \left|1\right\rangle\!\left\langle 1\right| + 2\eta(1-\eta) \left|2\right\rangle\!\left\langle 2\right| \tag{A.5}$$

$$\sigma_{12} = 3\eta(1-\eta)^2 \left|0\right\rangle\!\left\langle 0\right| + (1-\eta)(1-3\eta)^2 \left|1\right\rangle\!\left\langle 1\right| + \eta(2-3\eta)^2 \left|2\right\rangle\!\left\langle 2\right| + 3\eta^2(1-\eta) \left|3\right\rangle\!\left\langle 3\right| \tag{A.6}$$

Using those six states, we can calculate the signal state ρ_S using

$$\rho_S \approx \sum_{k=0}^{1} \sum_{l=0}^{2} \frac{\bar{n}^k}{(1+\bar{n})^{k+1}} \frac{e^{-\mu}\mu^l}{l!} \sigma_{kl}$$

Bibliography

[1] Arrazola, J. M., and Lütkenhaus, N. Quantum communication with coherent states and linear optics. *Phys. Rev. A 90* (Oct 2014), 042335.

[2] Arrazola, J. M., and Scarani, V. Covert quantum communication. *Phys. Rev. Lett. 117* (Dec 2016), 250503.

[3] Bash, B. A., Gheorghe, A. H., Patel, M., Habif, J. L., Goeckel, D., Towsley, D., and Guha, S. Quantum-secure covert communication on bosonic channels. *Nature communications 6* (2015).

[4] Bash, B. A., Goeckel, D., and Towsley, D. Limits of reliable communication with low probability of detection on awgn channels. *IEEE Journal on Selected Areas in Communications 31*, 9 (2013), 1921–1930.

[5] Helstrom, C. W. *Quantum Detection and Estimation Theory.* Academic Press New York, 1976.

[6] Sanguinetti, B., Traverso, G., Lavoie, J., Martin, A., and Zbinden, H. Perfectly secure steganography: Hiding information in the quantum noise of a photograph. *Phys. Rev. A 93* (Jan 2016), 012336.

[7] Scarani, V., Bechmann-Pasquinucci, H., Cerf, N. J., Dušek, M., Lütkenhaus, N., and Peev, M. The security of practical quantum key distribution. *Rev. Mod. Phys. 81* (Sep 2009), 1301–1350.

[8] Shankar, R. *Principles of Quantum Mechanics.* Springer Science & Business Media, 2012.

[9] Stallings, W. *Data and Computer Communications*, 4 ed. Maxwell Macmillan International, 1994.

[10] Wilde, M. M. *Quantum Information Theory.* Cambridge University Press, 2013.