Towards a Quantum Steganographic Capacity of Lossy Bosonic Channels

Vignesh Raman

Contents

1 Introduction

Steganography, the science of hiding information within an innocent looking message, has a long history that can traced back to Ancient Greece. The advent of the digital age has opened many new opportunities for hiding information and has led to the formalization of steganography using sound cryptographic principles. Briefly, the objective is for two legitimate parties, Alice and Bob, to use an innocent-looking covertext within which they hide a cypher-text, resulting in a stegotext. The stegotext is made available to an adversary, the warden Willie, from which Alice and Bob should hide the presence of the cyphertext, possibly using a shared secret key.

Quantum steganography is the extension of steganography to the quantum setting, wherein a quantum protocol (e.g.: a quantum error-correcting code) is used to hide classical or quantum information. Because of the unique nature of quantum states & channels [1], quantum steganography can be stronger than classical steganography.

A lot of effort has been devoted to characterizing how much information can be embedded into various quantum channels with or without noise. Back in 2016, Sanguinetti, et al., developed a steganographic protocol to transmit information privately, by hiding information in the quantum noise of a photograph [11]. More recently, Li & Liu [12] developed another novel quantum steganography scheme for color images that outperformed existing protocols in terms of embedding capacity. Bloch and Tahmasbi have recently developed and analyzed several quantum steganography protocols [10] that improve on earlier work, exploiting a concept [8] known in information theory as channel resolvability.

The central focus of the thesis is to provide a concise background survey of specific topics from relevant disciplines in classical and quantum information theory, and to put forth a formulation of the problem concerning the possible characterization of the steganographic capacity, i.e. the maximum rate of transmission of information under constraints of secrecy as dictated by steganography, for a specific type of quantum channel - channels that are linear, trace-preserving, positive maps - called the lossy bosonic channel (LBC).

An example to illustrate possible interest in the endeavor to formulate said problem, as well as to shed light upon why this is worth looking into for society, is to consider an instance where you have low-power fiber optics cables on which one party wants to discreetly send information to another party via a channel that is controlled by a third party.

The outline of the thesis is as follows:

- Section 2 is the literature review in conjunction with required background to yield a survey of topics required for Section 3.

- Section 3 presents the aforementioned problem formulation that concerns the possible characterization of the steganographic capacity.
- Section 4 is the conclusion of the presented background survey and problem formulation, which is ultimately closed off with a discussion on the future work that could build off this thesis.

2 Literature review and Background

The problem that is being attempted to be formulated is one that is likely to be worthy of consideration and important to society, because solving such a problem would be a significant step towards understanding the limitations concerning the rate of carrying/retrieval of information in lossy bosonic channels. The characterization of steganographic protocols for LBCs is of high interest to communication systems researchers in cyber-security. This is because LBCs are essentially (rather, from a slightly simplistic perspective) the canonical channel that models common communication channels supporting quantum information.

This section contains an overview of lossy bosonic channels, quantum steganography and channel resolvability - all of which are relevant to the problem formulation that follows.

2.1 Lossy Bosonic Channels

2.1.1 Introduction

An important, practically-relevant quantum channel in quantum communications, is the lossy bosonic channel (LBC). An LBC consists of a collection of bosonic modes that lose energy en route to the receiver from the transmitter. This channel can model the communication of photons over a fiber optic cable, or over free space, owing to the fact that the main source of noise in these situations is just the loss of photons.

A rigorous discussion of these channels requires the establishment of an applicable formalism of bosonic systems.

2.1.2 Bosonic Systems

It is of necessity to first consider the notion of a continuous variable system - this is a quantum system which has an infinite-dimensional Hilbert space described by observables with continuous eigenspectra[1]. Because these types of systems supply the quantum description of the propagating electromagnetic field, they are relevant for quantum communication.

[1] The set of possible eigenvalues of an observable is its eigenspectrum.

Examples of these systems are often represented by N bosonic modes[2], corresponding to N quantized radiation modes of the electromagnetic field. One can think of it corresponding to the situation dealing with N harmonic oscillators.

N bosonic modes are associated with a tensor product [3]: $\mathcal{H}^{\otimes N} = \bigotimes_{k=1}^{N} \mathcal{H}_k$ and corresponding pairs of bosonic field operators, i.e., annihilation and creation operators, $\{\hat{a}_k, \hat{a}_k^\dagger\}_{k=1}^N$. For convenience, represent these operators as a vector $\hat{b} := (\hat{a}_1, \hat{a}_1^\dagger, \cdots, \hat{a}_N, \hat{a}_N^\dagger)^T$, satisfying the commutation relations $[\hat{b}_i \hat{b}_j] = \Omega_{ij}$; $(i, j \in \{1, 2, \cdots, 2N\})$, where Ω_{ij} is the matrix element of the following symplectic matrix:

$$\Omega := \bigotimes_{k=1}^{N} \omega = \begin{bmatrix} \omega & & \\ & \ddots & \\ & & \omega \end{bmatrix}; \ \omega = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{2.1}$$

One thing to note about this system is that its Hilbert Space is separable and infinite dimensional. This is owing to the fact that the single-mode Hilbert space $\mathcal{H}$ is spanned by the Fock basis $\{|n\rangle_{n=0}^{\infty}\}$, with it being composed of eigenstates of the number operator $\hat{n} := \hat{a}^\dagger \hat{a}$. Over these states in particular, we have:

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1}\, |n+1\rangle, n \geq 0$$

$$\hat{a} |n\rangle = \sqrt{n}\, |n-1\rangle, n \geq 1; \ \ \hat{a} |0\rangle = 0, n = 0$$

These bosonic systems can also be described by another kind of field operators referred to as the quadrature field operators $\{\hat{q}_k, \hat{p}_k\}_{k=1}^N$, where $\hat{q}_k := \hat{a}_k + \hat{a}_k^\dagger$ and $\hat{p}_k := i\left(\hat{a}_k^\dagger - \hat{a}_k\right)$. We now note that quadrature operators are observables with continuous eigenspectra: $\hat{q} |q\rangle = q |q\rangle$, $\hat{p} |p\rangle = p |p\rangle$. The quadrature eigenvalues therefore can be used as continuous variables to describe the bosonic system in the phase-space representation.

2.1.3 Bosonic channels

We consider a multimode bosonic system (with N arbitrary modes) as described in the preceding discussion, whose quantum state is described by an arbitrary density operator $\hat{\rho} \in \mathcal{D}(\mathcal{H}^{\otimes N})$, where $\mathcal{D}(\mathcal{H}^{\otimes N})$ denotes the set of all density operators on $\mathcal{H}^{\otimes N}$.

Definition 2.1. *N-mode Bosonic Channel*: A completely positive, trace - preserving linear map $\mathcal{E} : \hat{\rho} \to \mathcal{E}(\hat{\rho}) \in \mathcal{D}(\mathcal{H}^{\otimes N})$

Because our focus is specifically on lossy bosonic channels (LBCs), we consider the model described by Figure 1.

[2]While one may note that the mode of a free quantum field is essentially the Fourier mode expansion of the field $\phi(\vec{x}) = \int \frac{d^3p}{(2\pi)^2\sqrt{2\omega_p}} \left(a(\vec{p})e^{i\vec{x}\cdot\vec{p}} + b(\vec{p})^\dagger e^{-i\vec{x}\cdot\vec{p}}\right)$, we more generally refer to any collection of creation/annihilation operators as modes, and classify them as bosonic (or fermionic) modes based on the obeyed commutation relations.

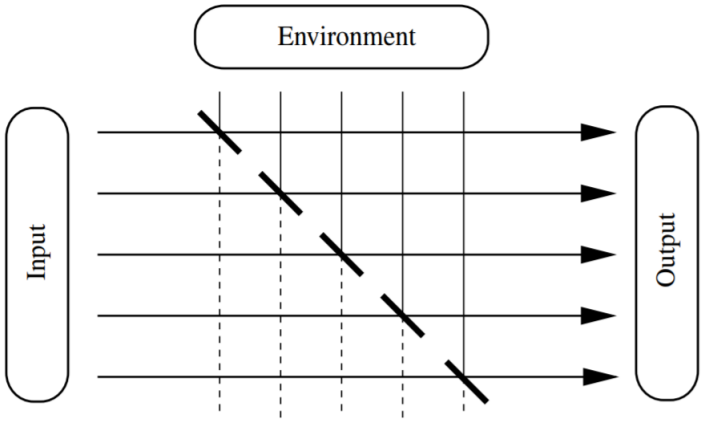

Figure 1: A schematic picture of the model of lossy bosonic channel [17]. Each input mode (left–right line), representing one use of the channel, interacts with the corresponding environment mode (top–bottom line) through a beam splitter.

2.1.4 Pure-Loss Bosonic Channels (PLBCs)

2.1.4.1 Introduction

We consider a scenario where the sender Alice attempts to transmit M bits to receiver Bob using n bosonic modes, whereas warden Willie attempts to detect her transmission attempt. We treat a single spatiotemporal polarization mode [4, Supplementary Note 1] of the electromagnetic field as the fundamental transmission unit over the channel.

Each of the 2^M possible M-bit messages maps to an n-mode codeword and their collection forms a codebook. Desirable codebooks ensure that the codewords, when corrupted by the channel, are distinguishable from one another. This provides reliability: a guarantee that the probability of Bob's error in decoding Alice's message $\mathbb{P}_e^{(b)} > \delta$ with arbitrarily small $\delta > 0$ over large n.

Willie's detector reduces to a binary hypothesis test of Alice's transmission state given his observations of the channel. Let $\mathbb{P}_{FA}$ denote the probability that Willie raises a false alarm when Alice does not transmit and $\mathbb{P}_{MD}$ denote the probability that Willie misses the detection of Alice's transmission. Under the assumption of equal prior probabilities on Alice's transmission state, Willie's detection error probability is $\mathbb{P}_e^w = \frac{\mathbb{P}_{FA}+\mathbb{P}_{MD}}{2}$.

Alice desires a reliable signalling scheme that is covert; in other words, ensures $\mathbb{P}_e^{(w)} > \frac{1}{2} - \epsilon$ for an arbitrarily small $\epsilon > 0$ regardless of Willie's quantum measurement choice (as $\mathbb{P}_e^{(w)} = \frac{1}{2}$ for a random guess).

Given the aforementioned scenario, consider a single-mode lossy bosonic channel $\varepsilon_{\eta_b}^{\bar{n}_T}$ of transmissivity $\eta_b \in (0, 1]$ and thermal noise mean photon number per mode $\bar{n}_T \geq 0$. Willie collects fraction $\eta_w = 1 - \eta_b$ of Alice's photons that do not reach Bob. For a pure loss bosonic channel ($\bar{n}_T = 0$), the environment

input is in the vacuum state $\hat{\rho}_0^E = |0\rangle\langle 0|^E$, corresponding to the minimum noise the channel must inject to preserve the Heisenberg inequality of quantum mechanics. The set-up is described in Fig. 2.

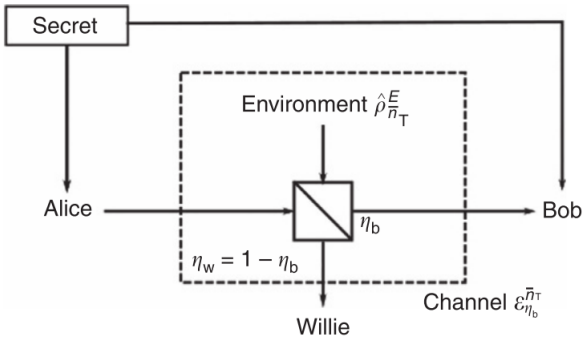

Figure 2: Willie collects fraction $\eta_w = 1 - \eta_b$ of Alice's photons that do not reach Bob. This lossy–noisy bosonic channel accurately models single-spatial-mode free space and single-mode fibre optical channels. Alice and Bob share a secret before the transmission [4].

2.1.4.2 Insufficiency of PLBCs for covert communication

One interesting thing to now note, is the inability to instantiate covert communication over a channel that lacks excess noise. That is, regardless of Alice's strategy, reliable and covert communication over a pure-loss bosonic channel ($\bar{n}_T = 0$) to Bob is impossible [4, Theorem 1].

Theorem 2.2. *Suppose Willie has a pure-loss channel from Alice and is limited only by the laws of physics in his receiver measurement choice. Then Alice cannot communicate to Bob reliably and covertly even if Alice and Bob have access to a pre-shared secret of unbounded size, an unattenuated observation of the transmission, and a quantum-optimal receiver.*

Intuitive explanation

When the noise mean photon number per mode $\bar{n}_T = 0$, i.e. we use a PLBC, we note that no photons impinge on Willie's single photon detector[3] (SPD) when Alice is silent. But even on the detection of a single photon or more, it is apparent that Alice is attempting to transmit information. This means that Alice is restricted to usage of codewords that are nearly indistinguishable from vacuum - this coupled with the fact that Willie can detect an attempt

[3]The detector is utilized by Willie to collect the fraction of Alice's photons that do not reach Bob

to communicate covertly with high probability for large n, simply means that Alice cannot reliably communicate[4] with Bob covertly!

Formal proof

Having gone over an intuitive explanation as to why a PLBC cannot enable reliable, covert communication between Bob and Alice, we now go over the formal proof of Theorem 2.2, using the approach and methods detailed in the supplementary information to [4], which demonstrates that Willie can use an ideal SPD on each mode of the pure-loss bosonic channel (PLBC) to effectively discriminate between any non-vacuum state in Alice's codebook and an arbitrary n-mode vacuum state.

Proof. Let $|\mathbf{k}\rangle \equiv \bigotimes_{i=1}^{n} |k_i\rangle$, where $\mathbf{k} \in \mathbb{N}_0^n$, with $\mathbb{N}_0$ being the set of non-negative integers. Alice sends one of 2^M (equally likely) M-bit messages by choosing an element from an arbitrary codebook $\{\hat{\rho}_x^{A^n}, x = 1, \cdots, 2^M\}$, where a state $\hat{\rho}_x^{A^n} = |\psi_x\rangle^{A^n A^n} \langle\psi_x|$ encodes an M-bit message W_x. Note that $|\psi_x\rangle^{A^n} = \sum_{\mathbf{k}\in\mathbb{N}_0^n} a_{\mathbf{k}}(x) |\mathbf{k}\rangle^{A^n}$ is a general n-mode pure state.[5]

Now, let Willie use an ideal SPD (Single Photon Detector) on all the n-modes. This is represented by a POVM (positive operator-valued measure): $\left\{|0\rangle\langle 0|, \sum_{j=1}^{\infty} |j\rangle\langle j|\right\}^{\otimes n}$.

Additionally, suppose W_u is transmitted - then, the task becomes to distinguish between $\hat{\rho}_0^{W^n} = |\mathbf{0}\rangle^{W^n W^n} |\mathbf{0}\rangle$ and $\hat{\rho}_1^{W^n} = \hat{\rho}_u^{W^n}$, where $\hat{\rho}_u^{W^n}$ is the output state of a PLBC with transmissivity η_w corresponding to an input state $\hat{\rho}_u^{A^n}$. With the messages being sent equiprobabilistically, we have Willie's average error probability to be:

$$\boxed{\mathbb{P}_e^{(w)} = \frac{1}{2^{M+1}} \sum_{u=1}^{2^M} {}^{W^n}\langle\mathbf{0}| \hat{\rho}_u^{W^n} |\mathbf{0}\rangle^{W^n}} \tag{2.2}$$

Using a lemma proved in [4] [Supplementary Note 3, Lemma 3], we have:

$$\begin{aligned} {}^{W^n}\langle\mathbf{0}| \hat{\rho}_u^{W^n} |\mathbf{0}\rangle^{W^n} &= \sum_{\mathbf{k}\in\mathbb{N}_0^n} |a_{\mathbf{k}}(u)|^2 (1-\eta_w)^{\sum_{i=1}^n k_i} \\ &\leq |a_{\mathbf{0}}(u)|^2 + (1 - |a_{\mathbf{0}}(u)|^2)(1-\eta_w) \\ &= 1 - \eta_w(1 - |a_{\mathbf{0}}(u)|^2) \end{aligned}$$

[4] Here, Willie is assumed to have access/authority to control the environment, and sets it to a vacuum

[5] We limit our analysis to pure input states since, by convexity, using mixed states as inputs can only degrade the performance (since that is equivalent to transmitting a randomly chosen pure state from an ensemble and discarding the knowledge of that choice).

Substituting this into equation (2.2) for Willie's average error probability, we get:

$$\boxed{\mathbb{P}_e^{(w)} \leq \frac{1}{2} - \frac{\eta_w}{2}\left(1 - \frac{1}{2^M}\sum_{u=1}^{2^M} |a_{\mathbf{0}}(u)|^2\right)} \tag{2.3}$$

From the above inequality, we see that Alice must use a code book with the following upper bound on the probability of transmitting one or more photons, such that $\mathbb{P}_e^{(w)} \geq \frac{1}{2} - \epsilon$ is ensured:

$$\boxed{\frac{1}{2^M}\sum_{u=1}^{2^M}\left(1 - |a_{\mathbf{0}}(u)|^2\right) \leq \frac{2\epsilon}{\eta_w}} \tag{2.4}$$

With this result in hand, we move on to show the existence of an interval $(0, \epsilon_0\ \epsilon_0 > 0$, such that if some $\epsilon \in (0, \epsilon_0]$, then Bob's average decoding error probability $\mathbb{P}_e^{(b)} \geq \delta_0; \delta_0 > 0$ (hence making covert communication unreliable over the PLBC).

Denote the event that the transmitted message W_u is decoded by Bob as $W_v \neq W_u$, by $E_{u\to v}$. Given that W_u is transmitted, the decoding error probability is the probability of the union of events $\bigcup_{v\neq u, v=0}^{2^M} E_{u\to v}$. Let Bob then, choose a POVM $\{\Lambda_j^*\}$ that minimizes the average probability of error over n modes:

$$\mathbb{P}_e^{(b)} = \inf_{\{\Lambda_j\}} \frac{1}{2^M}\sum_{u=1}^{2^M} \mathbb{P}\left(\bigcup_{v\neq u, v=0}^{2^M} E_{u\to v}\right) \tag{2.5}$$

We will now work with a codebook that meets the necessary condition for covert communication given by the upper bound on number of photons as previously derived (2.4). Define the subset of this codebook $\{\hat{\rho}_u^{A^n}, u \in \mathcal{A}\}$, where $\mathcal{A} = \left\{u : 1 - |a_{\mathbf{0}}(u)|^2 \leq \frac{4\epsilon}{\eta_w}\right\}$. Let us then lower bound (2.5) as below[6]:

$$\begin{aligned}\mathbb{P}_e^{(b)} &= \frac{1}{2^M}\sum_{u\in\bar{\mathcal{A}}} \mathbb{P}\left(\bigcup_{v\neq u, v=0}^{2^M} E_{u\to v}\right) + \frac{1}{2^M}\sum_{u\in\mathcal{A}} \mathbb{P}\left(\bigcup_{v\neq u, v=0}^{2^M} E_{u\to v}\right) \\ &\geq \frac{1}{2^M}\sum_{u\in\mathcal{A}} \mathbb{P}\left(\bigcup_{v\neq u, v=0}^{2^M} E_{u\to v}\right)\end{aligned} \tag{2.6}$$

Assume $|\mathcal{A}|$ to be even WLOG, and split[7] $\mathcal{A}$ into $\mathcal{A}^{(\text{left})}$ and $\mathcal{A}^{(\text{right})}$.

[6] the probabilities in (2.6) are with respect to the POVM $\{\Lambda_j^*\}$ that minimizes (2.5) over the entire codebook

[7] these are two equal-sized non-overlapping subsets

Let $g : \mathcal{A}^{(\text{left})} \to \mathcal{A}^{(\text{right})}$ be a bijection, so that we can write (2.6) as follows:

$$\begin{aligned}\mathbb{P}_e^{(b)} &\geq \frac{1}{2^M} \sum_{u \in \mathcal{A}^{(\text{left})}} 2 \left(\frac{\mathbb{P}\left(\bigcup_{v \neq u, v=0}^{2^M} E_{u \to v}\right)}{2} + \frac{\mathbb{P}\left(\bigcup_{v \neq g(u), v=0}^{2^M} E_{g(u) \to v}\right)}{2} \right) \\ &\geq \frac{1}{2^M} \sum_{u \in \mathcal{A}^{(\text{left})}} 2 \left(\frac{\mathbb{P}(E_{u \to g(u)})}{2} + \frac{\mathbb{P}(E_{g(u) \to u})}{2} \right)\end{aligned} \tag{2.7}$$

The second lower bound follows from the fact that the events in the latter are contained in the unions. For convenience, we define the summation term to be Bob's average probability of error when Alice only sends messages W_u and $W_{g(u)}$, i.e.:

$$\mathbb{P}_e(u) \equiv \frac{\mathbb{P}(E_{u \to g(u)})}{2} + \frac{\mathbb{P}(E_{g(u) \to u})}{2} \tag{2.8}$$

Now, the lower bound on the probability of error in discriminating two received codewords, can be obtained by lower-bounding the probability of error in discriminating these two codewords before they are sent[8]. From [4] [Chapter IV.2 (c), Equation (2.34)], we have the the lower bound on the probability of error in discriminating between $\left|\psi_u^{A^n}\right\rangle$ and $\left|\psi_{g(u)}^{A^n}\right\rangle$ to be:

$$\boxed{\mathbb{P}_e(u) \geq \frac{1}{2}\left[1 - \sqrt{1 - F\left(|\psi_u\rangle^{A^n}, \left|\psi_{g(u)}\right\rangle^{A^n}\right)}\right]} \tag{2.9}$$

Here, $F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2$ is the fidelity between the pure states. To lower bound the RHS of equation (2.9), one can lower bound $F\left(|\psi_u\rangle^{A^n}, \left|\psi_{g(u)}\right\rangle^{A^n}\right)$. Now, an equivalent way of writing the fidelity quantity [19, equation (9.134)], in conjunction with the triangle inequality (for trace distance), yields:

$$\begin{aligned}F\left(|\psi_u\rangle^{A^n}, \left|\psi_{g(u)}\right\rangle^{A^n}\right) &= 1 - \left(\frac{1}{2}\left\|\hat{\rho}_u^{A^n} - \hat{\rho}_{g(u)}^{A^n}\right\|_1\right)^2 \\ &\geq 1 - \left(\frac{\left\|\hat{\rho}_u^{A^n} - |\mathbf{0}\rangle^{A^n A^n}\langle\mathbf{0}|\right\|_1}{2} + \frac{\left\|\hat{\rho}_{g(u)}^{A^n} - |\mathbf{0}\rangle^{A^n A^n}\langle\mathbf{0}|\right\|_1}{2}\right) \\ &= 1 - \left(\sqrt{1 - |^{A^n}\langle\mathbf{0}|\psi_u\rangle^{A^n}|^2} + \sqrt{1 - |^{A^n}\left\langle\mathbf{0}\middle|\psi_{g(u)}\right\rangle^{A^n}|^2}\right)\end{aligned}$$

$$\implies \mathbb{P}_e(u) \geq \frac{1}{2}\left[1 - \left(\sqrt{1 - |^{A^n}\langle\mathbf{0}|\psi_u\rangle^{A^n}|^2} + \sqrt{1 - |^{A^n}\left\langle\mathbf{0}\middle|\psi_{g(u)}\right\rangle^{A^n}|^2}\right)\right] \tag{2.10}$$

[8] this is equivalent to Bob having an unattenuated unity-transmissivity channel from Alice

Since $|^{A^n}\langle \mathbf{0}|\psi_u\rangle^{A^n}|^2 = |a_{\mathbf{0}}(u)|^2$, $1-|a_{\mathbf{0}}(u)|^2 \leq \frac{4\epsilon}{\eta_w}$ and $1-|a_{\mathbf{0}}(g(u))|^2 \leq \frac{4\epsilon}{\eta_w}$, we have:

$$\boxed{\mathbb{P}_e(u) \geq \frac{1}{2} - 2\sqrt{\frac{\epsilon}{\eta_w}} \implies \mathbb{P}_e^{(b)} \geq \frac{|\mathcal{A}|}{2^M}\left(\frac{1}{2} - 2\sqrt{\frac{\epsilon}{\eta_w}}\right)} \tag{2.11}$$

Now, restating (2.4),

$$\frac{2\epsilon}{\eta_w} \geq \frac{1}{2^M}\sum_{u\in\bar{\mathcal{A}}}^{2^M}\left(1-|a_{\mathbf{0}}(u)|^2\right) \implies \frac{4\epsilon}{\eta_w}\frac{2^M-|\mathcal{A}|}{2^M} \implies \frac{|\mathcal{A}|}{2^M} \geq \frac{1}{2} \tag{2.12}$$

From (2.11) and (2.12), we have the positive lower bound on Bob's probability of decoding error $\boxed{\mathbb{P}_e^{(b)} \geq \frac{1}{4} = \sqrt{\frac{\epsilon}{\eta_w}}}$ for any n and $\epsilon \in (0, \frac{\eta_w}{16}]$, thus demonstrating that reliable covert communication over a pure-loss channel is impossible. □

2.2 Quantum Steganography

2.2.1 Brief overview of Steganography

The aim in steganography is for two parties to successfully embed information within an innocent looking message (cover text), without being detected by an undesired party.

We consider a general model of steganography systems [1]. Alice wishes to send some data (embedded data E) secretly, to another party Bob, without being detected by Eve. The message to be sent then, i.e. the stego-data, is computed using E, a key K, and an innocent looking cover data C. Note that C itself is computed from environmental data V using an algorithm referred to as a cover generating algorithm $\mathcal{G}$.

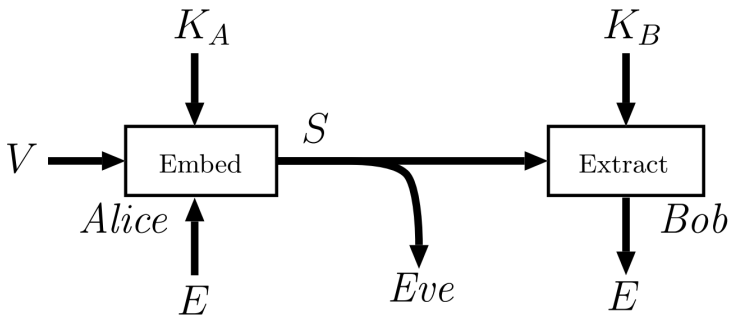

Figure 3: Communication with Classical Steganography [1]

What makes a communication system (such as that in Figure 3, reproduced from [1]) steganographic, is a pair of an embedder $\mathcal{E}$ and an extractor/decoder $\mathcal{D}$, such that the following hold:

$$\mathcal{E}(V, E, K) = S; \;\; \mathcal{D}(S, K) = E$$

Finally then, note that Eve's task is to detect usage of steganography by detecting C or S. The system is *perfectly secure* if the probability distribution of S is equal to that of C, i.e. C and S are indistinguishable to Eve.

2.2.2 Comparing Quantum Steganography to Classical Steganography

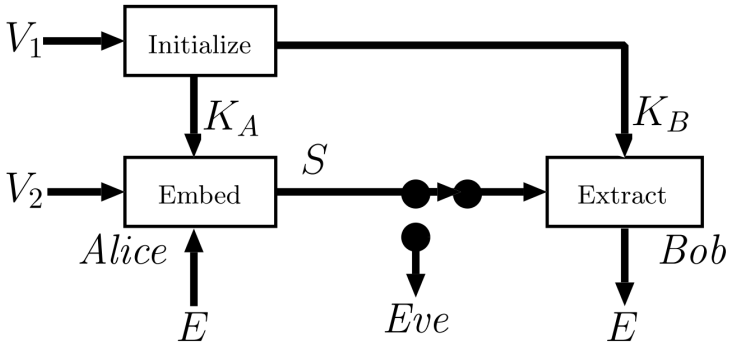

Figure 4: Communication with Quantum Steganography [1]

In the quantum steganography model [1] (Figure 4, reproduced from [1]), all random variables are replaced by quantum registers, allow Eve to destructively measure C or S, and add an initialization step for the keys, since K cannot be cloned owing to the no-cloning theorem.

In this model, the perfect security condition is $\rho_c = \rho_s$, where ρ_c and ρ_s are the density matrices of C and S respectively.

It has been shown that quantum steganography can be strictly more secure than the classical steganography system [1], since no classical steganography system can be perfectly secure if its cover data (whose distribution is unknown) is the result of a measurement[2].

2.2.3 Recent work

Many different models of steganography have been put forth that contain efforts on characterizing how much information can be embedded into various quantum channels with/without noise, and to assess how much key is required to achieve the task. We are interested in the model that [5],[14] considers, which assumes

that the warden has inaccurate knowledge of what the channel is. Specifically, we assume that the knowledge of the channel that the warden possesses, is a degraded version of the real channel - this can be achieved by intentionally cascading another channel at the transmitter.

Unlike earlier results [14],[5], Tahmasbi and Bloch [10] show that no shared key is required to run the stego protocol when the channel is noiseless. This is achieved through the use of a random encoder obtained from privacy amplification and source coding with side information techniques. Furthermore, they relax the assumption on the cover code in [14] that "on a valid codeword in the QECC, the typical errors all have distinct error syndromes, and act as unitaries that move the state to a distinct, orthogonal subspace," by relying on one-shot coding results.

2.3 Channel Resolvability

2.3.1 Classical Channel resolvability

We first define a few things.

Definition 2.3. *Variational Distance*: Letting P_Z and $P_{\tilde{Z}}$ be probability distributions on a countably infinite set $\mathcal{Z}$, the variational distance between them is:

$$d(P_Z, P_{\tilde{Z}}) := \frac{1}{2} \sum_{z \in \mathcal{Z}} |P_Z(z) - P_{\tilde{Z}}(z)|$$

Definition 2.4. *$\varepsilon-$Limit Superior (in probability)*: For $\varepsilon \in [0, 1]$,

$$\varepsilon p - \lim_{n\to\infty} \sup Z_n := \inf\{\alpha : \lim_{n\to\infty} \inf \Pr\{Z_n > \alpha\}\}$$

$$\varepsilon p^* - \lim_{n\to\infty} \sup Z_n := \inf\{\alpha : \lim_{n\to\infty} \sup \Pr\{Z_n > \alpha\}\}$$

Definition 2.5. *$\delta-$ Channel Resolvability*: Let $\delta \in [0, 1)$ be fixed arbitrarily. A resolvability rate $R \geq 0$ is said to be $\delta-$achievable at $\boldsymbol{X}$ if there exists a deterministic mapping $\varphi_n : \{1, \cdots, M_n\} \to \chi^n$ satisfying:

$$\lim_{n\to\infty} \sup \frac{1}{n} \log M_n \leq R$$
$$\lim_{n\to\infty} \sup d(P_{Y^n}, P_{\tilde{Y}^n}) \leq \delta$$

Where $\tilde{Y}^n$ denotes the output via W^n due to the input $\tilde{X}^n = \varphi_n(U_{M_n})$. Note that U_{M_n} denotes a uniform random number of the size M_n, which is a random variable uniformly distributed over $\{1, \cdots, M_n\}$. The $\delta-$channel resolvability then [18], at $\boldsymbol{X}$, is:

$$S(\delta|\boldsymbol{X}, \boldsymbol{W}) := \inf\{R : R \text{ is } \delta\text{- achievable at } \boldsymbol{X}\}$$

Finding the asymptotically minimum rate of the size of the uniform random numbers which can approximate a given target output distribution via a channel is called the problem of channel resolvability. Another formulation of the channel resolvability result, in the case of a Discrete Memoryless Channel (DMC), is presented as Lemma 5 in [20]. However, we will be interested in channel resolvability results for classical-quantum (cq) channels.

2.3.2 Channel Resolvability result for cq-channels

We now present a lemma [20, Lemma 6] for a cq-channel $x \mapsto \rho_W^x$, whose usage we shall note in the problem formulation section as being important in the proof of a related problem described in [10].

Lemma 2.6. *Let W be a message uniformly distributed over $\mathcal{M}$, P_X be a probability distribution over χ^n and $F : \mathcal{M} \to \chi^n$ be a random encoder whose codewords are iid according to a distribution P_X. Let $\rho_{\boldsymbol{W}}^{\boldsymbol{x}} := \rho_W^{x_1} \otimes \cdots \rho_W^{x_n}$ and $\phi(s) := \log\left(\sum_{\boldsymbol{x}} P_{\boldsymbol{X}}(\boldsymbol{x}) \operatorname{Tr}\left\{(\rho_{\boldsymbol{W}^{\boldsymbol{x}}}^{1-s}\rho_{\boldsymbol{W}}^{s})\right\}\right)$. We then have:*

$$\mathbb{E}_F\left(\left\|\frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}\rho_{\boldsymbol{W}}^{F(m)} - \rho_{\boldsymbol{W}}\right\|_1\right) \leq 2\sqrt{2^{\gamma s+\phi(s)}} + \sqrt{\frac{2^{\gamma}\nu(\rho_{\boldsymbol{W}})}{|\mathcal{M}|}} \tag{2.13}$$

3 Problem Formulation and discussion

Let W be the classical message with uniform distribution over $[\![1, M]\!]$, which is required to be reliably transmitted. To use the channel (trace-preserving completely positive map) $\mathcal{N}_{A\to B}$ n times, for running the classical communication protocol, we look at the classical communication code, $(M, \epsilon)^{\mathrm{CC}}$, where:

$$\frac{1}{M}\sum_{w=1}^{M} \operatorname{tr}\left(\Lambda^n \mathcal{N}_{A\to B}^{\otimes n}(f(w))\right) \geq 1 - \epsilon$$

Note that the code consists of:

- Function: $f : [\![1, M]\!] \to \mathcal{D}(\mathcal{H}_A^{\otimes n})$.[9] This is to encode message w into an input state $\rho_{A^n}^w \triangleq f(w)$.
- POVM: $\Lambda = \{\Lambda^w\}_{w\in[\![1,M]\!]}$. This is to decode W.

We note that the cover protocol $\mathcal{P}^c$ (inducing the quantum state $\rho_{B_n}^c$) is known to the warden, and that the perceived/assumed channel is in fact $\mathcal{N}_{A\to B}^{\otimes n} \circ \mathcal{M}_{A\to B}^{\otimes n}(f(w))$, which is a degraded version compared to just $\mathcal{N}_{A\to B}^{\otimes n}(f(w))$.

Our overarching objective then is to exploit this gap in the warden's knowledge of the channel characteristics, and run a stego protocol $\mathcal{P}^c$ to induce the state $\rho_{B_n}^s$ at the output of $\mathcal{N}_{A\to B}^{\otimes n}$, such that $\left\|\rho_{B_n}^c - \rho_{B_n}^s\right\|_1$ is small.

[9] Here, $\mathcal{D}(\mathcal{H}_A^{\otimes n})$ denotes the set of density operators on $\mathcal{H}_A$.

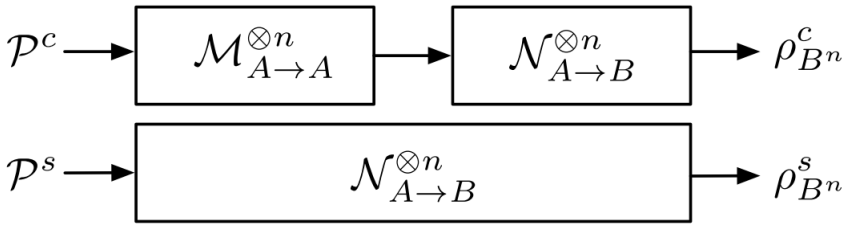

Figure 5: Top: Channel structure that exists according to Warden. Bottom: Actual channel structure [20].

More specifically, let A and B be two multimode bosonic systems, and $\mathcal{N}_{A\to B}^{\otimes n}$ be a lossy bosonic channel (LBC) [note that it is noisy, and that it is used n times]. $\mathcal{H}_A$ and $\mathcal{H}_B$ here are infinite dimensional Hilbert spaces (i.e. the Hilbert spaces are associated with the bosonic modes used in communication). Our aim then, would be to show the existence of a stego protocol that produces $\rho_{B_n}^s$ satisfying $\left\|\rho_{B_n}^c - \rho_{B_n}^s\right\|_1 \leq \zeta$ for any $\zeta > 0$ under certain constraints on ζ.

Now, there exists prior work [10] for a related problem with the same set-up of the model, except for the fact that systems A and B involved in the quantum channel $\mathcal{N}_{A\to B}$ are described by finite-dimensional Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. In this case, the existence of a stego-protocol - which produces $\rho_{B_n}^s$ satisfying $\left\|\rho_{B_n}^c - \rho_{B_n}^s\right\|_1 \leq \zeta$ for any $\zeta > 0$, provided certain constraints on ζ - is proven in part by utilization of specialized one-shot coding results, as alluded to in the end of Section 2.2.3. One of the one-shot coding results, which is an achievability result that states that there exists a classical communication code for a cq-channel inducing a pre-specified state at the output, utilizes both quantum channel coding as well as channel resolvability results. Specifically, Lemma 2.6. from Section 2.3.2. is used in proving this achievability result, and so we see that channel resolvability is inherently useful in solving such kind of problems. Therefore, this concept needs to be kept in mind when proceeding with future work that solves the problem formulated in Section 3 of this paper.

In order to make more progress, it might be worth it to slightly change the problem. Framing the proposed problem with the addition of an assumption that the warden does not have access to a shared secret between Alice and Bob [10], and using the model of a single-mode lossy thermal (bosonic) noise channel, can aid us in getting some parametric constraints.

[10] While a seemingly restrictive assumption at first glance, it may be so that cost of transmission being detected (or the cost of having a high quality/detection-resistant channel) might be significantly higher than sharing a secret, so in such situations it may be worth noting this assumption.

Let us therefore be more specific with the considered channel model. Consider the channel model for a single mode lossy thermal noise channel (which happens to be the quantum mechanical description of the transmission of a single spatiotemporal-polarization mode of the electromagnetic field at a given transmission wavelength over linear loss and additive Gaussian noise).

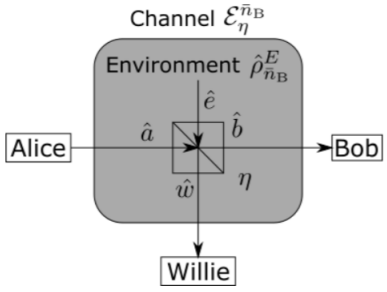

Figure 6: Single-mode bosonic channel $\varepsilon_\eta^{\bar{n}_B}$ modeled by a beam-splitter with transmissivity η and an environment injecting a thermal state $\hat{\rho}_{\bar{n}_B}$ with mean photon number $\bar{n}_B$. $\hat{a}, \hat{b}, \hat{w}, \hat{e}$ label the input/output modal annihilation operators [19].

To model loss, we take a beam-splitter with transmissivity η. As deducible from the LBC section (2.1), the relationship between bosonic mode operators (input/output modal annihilation operators) $\hat{b} = \sqrt{\eta}\hat{a} + \sqrt{1-\eta}\hat{e}$ needs $\hat{e}$ to ensure that $[\hat{b}, \hat{b}^\dagger]$. Now, Bob captures a fraction η of Alice's transmitted photons, while the warden Willie have $1-\eta$. Noise is modeled here (note the deviation from the PLBC figure 3 in section 2.1.4.1.) by mode $\hat{e}$ being in a zero-mean thermal state $\hat{\rho}_{\bar{n}_B}$, with $\bar{n}_B$ being the mean photon number per mode injected by the environment.

If Alice desires to transmit a message x, she modulates an n-mode state $\hat{\rho}_x^{A^n}$ [11] using the shared secret we mentioned in the assumption. Willie then basically performs a hypothesis test to determine whether Alice transmits anything or not. When Alice does not transmit any message, Willie observes a product thermal state $\hat{\rho}_0^{W^n} = \hat{\rho}_{\eta\bar{n}_B}^{\otimes n}$, and when Alice does transmit a message, let us say Willie observes $\hat{\rho}_1^{W^n}$ state. H_0 hypothesis corresponds to no transmission, and H_1 hypothesis corresponds to transmission. Then the probability of Willie's detection error [12] is,

$$P_e^{(w)} = \frac{P_{FA} + P_{MD}}{2}$$

where P_{FA} is probability of false alarm and P_{MD} is probability of missed detection.

[11] May be entangled across n modes.
[12] We assume equally likely hypotheses.

The system then, will be said to be covert [15] if for any $\delta_P > 0, P_e^{(w)} \geq \frac{1}{2} - \delta_P$ for large enough n. This is equivalent to (deducible from the quantum optimal receiver's $\min P_e^{(w)}$ yield), i.e. this criterion is satisfied, if

$$\frac{1}{4}\left\|\hat{\rho}_0^{W^n} - \hat{\rho}_1^{W^n}\right\|_1 \leq \delta_P$$

In order obtain a constraint on existing physical parameter(s) due to the covertness criteria, let us represent this criteria for covertness with a more convenient measure such as QRE (quantum relative entropy):

$$D(\hat{\rho}||\hat{\sigma}) = \mathrm{Tr}[\hat{\rho}\log\hat{\rho} - \hat{\rho}\log\hat{\sigma}]$$

It is considered to be more convenient because this measure happens to be additive over product states. Using Chernoff's lemma and Pinsker's inequality, we can relate it to performance of optimal hypothesis test, so to maintain slightly higher level of covertness, we set $\delta_Q RE = 2\delta_P^2$ (this is derivable from the fact that $\|\hat{\rho} - \hat{\sigma}\|_1 \leq \sqrt{D(\hat{\rho}||\hat{\sigma})}$. So, recasting the aim with this measure in mind, we would require our aim to prove the system covertness, given that the criterion for covertness is if, for any $\delta_{QRE} > 0$, $D\left(\hat{\rho}_1^{W^n}||\hat{\rho}_0^{W^n}\right) \leq \delta_{QRE}$ for large enough n.

Explicitly, [15] shows that the constraint (which manifests onto the transmitted mean photon number per mode $\bar{n}_S$ from Alice's end), owing to this criterion for this channel model, is:

$$D\left(\hat{\rho}_1^{W^n}||\hat{\rho}_0^{W^n}\right) \leq \delta_{QRE} \implies \bar{n}_S \leq \frac{\sqrt{2\eta\bar{n}_B(1+\eta\bar{n}_B)}}{1-\eta}\sqrt{\frac{\delta_{QRE}}{n}}$$

So we have obtained a constraint that happens to be the converse theorem on the transmitted mean photon number per mode, for a more restrictive version [13] of the the problem we set out to formulate.

4 Conclusion and Future Work

In this work, a concise survey of the topics of lossy bosonic channels, channel resolvability, and quantum steganography has been carried out. This has been followed by a problem formulation with a formally described channel model that makes progress towards the possibility of establishment of the existence of a stego protocol in the presented lossy bosonic channel model. With some additional assumptions to the proposed problem, an approach that is heavily based on [15] to obtaining the converse theorem for a similar problem is also presented.

[13] This is owing to the added assumptions of (1) shared secret between Alice and Bob and (2) equally likely hypotheses.

Future work would aim to successfully develop and frame the mathematical constraints alluded to in the problem formulation without the assumption of a shared secret in the first place. Additionally, it would frame a tangible theorem based on the above problem formulation for the described channel model, and provide the proof for the stego-protocol's existence given constraints on ζ, that would proceed similarly in the vein of Theorem 1's proof in [10].

5 Acknowledgement

The author would like to thank Dr. Bloch for introducing them to quantum information theory, providing guidance with advice and feedback, and for being an encouraging, patient and accommodating thesis advisor. The author also thanks Dr. Barry for being the second reader for the thesis, Dr. Madden for suggestions concerning the aspect of thesis writing, and their family for their love and support.

References

[1] S. Natori, Why Quantum Steganography Can Be Stronger Than Classical Steganography

[2] S. Natori: One-time hash steganography, in LNCS (1999) pp. 17–28

[3] Gaussian quantum information: Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd. Rev. Mod. Phys. 84, 621

[4] Boulat A. Bash, et al. : Quantum-secure covert communication on bosonic channels, Nature

[5] Chris Sutherland and Todd A. Brun Phys. Rev. A 100, 052312: Quantum steganography over noisy channels: Achievability and bounds

[6] H. Yagi, "Channel resolvability theorems for general sources and channels," 2017 IEEE International Symposium on Information Theory (ISIT), Aachen, 2017, pp. 2741-2745

[7] Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, Phys. Rev. Lett. 92, 027902 – Published 15 January 2004: Classical Capacity of the Lossy Bosonic Channel: The Exact Solution

[8] M. R. Bloch and J. N. Laneman, "Strong Secrecy From Channel Resolvability," in IEEE Transactions on Information Theory, vol. 59, no. 12, pp. 8077-8098, Dec. 2013.

[9] M. Bloch, "Channel intrinsic randomness," 2010 IEEE International Symposium on Information Theory, Austin, TX, 2010, pp. 2607-2611.

[10] M. Tahmasbi and M. Bloch, "Steganography Protocols for Quantum Channels," in Proc. of IEEE International Symposium on Information Theory, Paris, France, Jul. 2019, pp. 2179–2183.

[11] Sanguinetti, B.; Traverso, G.; Lavoie, J.; Martin, A. & Zbinden, H. Perfectly secure steganography: Hiding information in the quantum noise of a photograph Phys. Rev. A, American Physical Society, 2016 , 93 , 012336

[12] Li, P. & Liu, X. A novel quantum steganography scheme for color images, International Journal of Quantum Information, 2018 , 16 , 1850020

[13] Hayashi, M. Quantum Information Theory - Mathematical Foundation; Graduate Texts in Physics; Springer: Berlin, Germany, 2017.

[14] Chris Sutherland and Todd A. Brun Phys. Rev. A 101, 052319: Quantum steganography over noiseless channels: Achievability and bounds

[15] M. S. Bullock, C. N. Gagatsos, S. Guha and B. A. Bash, "Fundamental Limits of Quantum-Secure Covert Communication over Bosonic Channels," 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, IL, USA, 2019, pp. 56-63, doi: 10.1109/ALLERTON.2019.8919793.

[16] Ludovico Lami et al 2018 New J. Phys. 20 113012: All phase-space linear bosonic channels are approximately Gaussian dilatable

[17] Cosmo Lupo et al 2009 New J. Phys. 11 063023: Capacities of lossy bosonic channel with correlated noise

[18] H. Yagi, "Channel resolvability theorems for general sources and channels," 2017 IEEE International Symposium on Information Theory (ISIT), Aachen, 2017, pp. 2741-2745, doi: 10.1109/ISIT.2017.8007028.

[19] M.M. Wilde, Quantum Information Theory (Cambridge University Press, 2013).

[20] Tahmasbi, Mehrdad. Covert Communication: from classical channels to quantum channels. Diss. Georgia Institute of Technology, 2020.