Continuous-Variable Quantum Key Distribution

Master Thesis
submitted to the Faculty of the
Escola Tècnica d'Enginyeria de Telecomunicació de Barcelona
Universitat Politècnica de Catalunya
by

Walid Ben Saoud Benjerri

In partial fulfillment
of the requirements for the master in
Cybersecurity

Advisor: Rodriguez Fonollosa, Javier
Barcelona, September 2022

Contents

List of Figures

Revision history and approval record

Revision	Date	Purpose
0	01/06/2022	Document creation
1	01/09/2022	Document revision
2	04/10/2022	Document revision
3	18/10/2022	Document revision
2	22/10/2022	Document revision

DOCUMENT DISTRIBUTION LIST

Name	e-mail
[Walid Ben Saoud Benjerri]	walid.ben.saoud.benjerri@estudiantat.upc.edu
[Javier Rodriguez Fonollosa]	javier.fonollosa@upc.edu
[Project Supervisor 2]	

Written by:		Reviewed and approved by:	
Date	22/10/2022	Date	22/10/2022
Name	Walid Ben Saoud Benjerri	Name	Javier Rodriguez Fonollosa
Position	Project Author	Position	Project Supervisor

Abstract

Continuous-Variable Quantum Key Distribution is an alternative to the usual discrete variable quantum key distribution schemes, such as BB84, based on quantum states of light. Optical quantum states are defined in an infinite dimensional Hilbert space hence the ”continuous” prefix. Compared to DV-QKD, CVQKD has several advantages such as being compatible with existing telecommunications equipment as one only needs to modulate coherent states of light and perform homodyne/heterodyne detections, but suffers from a lower range, and security proofs are somewhat harder.

In the first chapter, we briefly remind the reader about the motivations for physical layer security(of which QKD and CVQKD are instances) against cryptography, the dominant approach for network security. Chapter 2 gives the necessary background for this topic, which requires both a mathematically rigorous theory of quantum mechanics in infinite dimensional Hilbert spaces, and a quantitative theory of information to be able to prove the security of protocols and measure the shared information between the legitimate and illegitimate parties. Chapter 3 studies in detail the protocol using coherent states with Gaussian modulation, the most common one.

1 Introduction

1.1 Cryptography VS Physical-Layer Security

One may legitimately enquire about the purpose of an exotic security scheme such as QKD, whereas cryptography is well-known, understood, and used in an ubiquitous fashion nowadays. For example, almost most Internet-based communications today rely on public and private key (asymmetric) cryptography to achieve both authentication and confidentiality without the need of a confidential or authenticated channel, and without the need of a preshared key (though the other party's public key has to be authenticated). Symmetric cryptography is also used as part of hybrid schemes once a session key has been agreed upon with these asymmetric protocols.

The answer to that question is that despite their ubiquity, public key cryptography protocols are only secure if two assumptions are true :

1. Some mathematical problem is hard (Integer factorization for RSA, discrete logarithm for Diffie-Hellman)
2. The attacker has limited computational power. This is commonly expressed as the attacker running a polynomial-on-n (the key size) computation, therefore unable to solve a hard problem in reasonable time for a large enough value of n.

If any of these two assumptions is false, security is compromised and the protocol is no longer secure.

Indeed, there is no known proof of the hardness of the mathematical problems underlying for example RSA or Diffie-Hellman, it is only presumed since efforts to find a polynomial algorithm didn't succeed.

Furthermore, even if these assumptions are true, computational power is ever increasing, which leads to key length being a determinant parameter in the security even against a polynomial adversary, so a key size that is secure today may cease to be in a few years. For example, RSA with a key length of 1024 is no longer considered secure for long-term.

Moreover, a mathematical problem that is hard for traditional computers can actually be easy for quantum computers : Shor famously devised an algorithm that can factorize integers in polynomial time, breaking the security of RSA if the attacker has access to a quantum computer with enough qubits. Today this is not possible (or not disclosed), but might be so in a few years, A solution to this last issue however is the design of algorithms based on mathematical problems that are resistants to attacks even by quantum computers. This area of research is called **post-quantum cryptography**, *but this is not the subject of this thesis.*

The protocols that we are interested in belongs to the wider class of **Information-theoretic security** also called **Unconditional Security** or **Physical layer security**. These designations are justified as we shall see next.

Unlike cryptography, security is guaranteed and proven in a mathematical sense against every possible attacker.

Roughly speaking, if X is the original message, Y is the message as observed by the legitimate recipient and Z is the message as observed by an eavesdropper, we want $H(X|Y)$ to be small (a small conditional entropy means there will be less uncertainty on the message for the legitimate recipient) but $H(X|Z)$ be high (a high entropy of X conditioned on Z means he will have a higher uncertainty, for this reason this quantity is also called **equivocation**). The ideal case would be that $H(X|Z) = H(X)$ so the attacker doesn't gain supplementary information (measured in bits) than what he had a priori before observing Z. This is called perfect secrecy, but it is impractical to achieve, so IT security focuses instead on the asymptotic analysis.

For the simplicity of analysis, we assume the source to encode a message M into a sequence of symbols X^n. Similarly Y^n and Z^n are the noisy observations of X^n by the intended recipient and the eavesdropper.

Then the goal is to compute the conditional entropy, that is compute $H(M|Z^n)$ as a function of n and study the *asymptotic* information leakage given by the limit as $n \to \infty$ of $I(M; Z^n)$. This limit should be 0 which implies that the eavesdropper's information regarding the message is arbitrary close to zero, so we can chose a high enough value of n to achieve the desired closeness. It can be shown that the optimal rate of encoding (from the message M to X^n) is $I(X, Y) - I(X, Z)$ which is intuitively true since it can be thought of as the difference of bits shared by Alice and Bob, and bits shared between Alice and Eve.

The model we just described is called Wyner's wiretap channel, since the attacker is able to get a noisy version of the message. Being a passive attacker we call it an eavesdropper. This paradigm received less attention than public key cryptography which was also a revolution in the same decade (1970s), perhaps due to the lack of a concrete channel. Information theoretical security therefore studies *fundamental bounds* on the possible information an attacker (or legitimate recipient) can get, regardless of the algorithm employed.

A key difference however from Public Key cryptography is the need for an authenticated channel, but it can be a public channel. Quantum Key Distribution in its discrete flavour is perhaps the best illustration of this paradigm.

1.2 Overview of discrete QKD

Quantum key distribution is a protocol relying on quantum mechanical properties to generate a shared key between two parties, here denoted Alice and Bob, knowledge of which is mathematically proven to be impossible for an outside eavesdropper (in an asymptotic sense that is similar to the previously described Wyner model, which is actually the secret-key agreement from a source model).

It is radically different in its principle than well known and used cryptography protocols such as Diffie-Hellman since it doesn't rely on a hard problem, but instead on the properties of quantum systems. QKD is an instance of the Wyner model, in which the channel is a quantum system, in the discrete case this channel is simply a qubit register. **Note that the qubits are generated independently i.e they are not entangled** Discrete QKD, here we illustrate the BB84 protocol, is based on the fact that measuring a quantum system will alter its state, but only if the system is not already in an eigenstate of the measurement operator. So even a passive eavesdropper will leave an ”evidence” of the measurement, unless she knows the encoding basis of each qubit. Based on this idea, the protocol is designed so that the possible prepared states are non-orthogonal making impossible for Eve the task of knowing precisely the state prepared by Alice, and there is a positive chance to modify the state when eavesdropping, thus introducing errors in Bob´s statistics.

More precisely there are 4 possible states prepared by Alice which are the result of the following encoding of her bit (classical) input : she encodes 0 as $|0\rangle$ or $|+\rangle$ at random, and 1 as $|+\rangle$ or $|-\rangle$ similarly. Introducing the Pauli X and Z basis, this means that 0 is encoded as the first eigenvector of either two basis, and 1 as the second vector. The choice of basis for a given pulse is made at random with 50% chance of choosing the Z basis, and 50% chance of choosing the X basis. Assume Bob measures in the same base Alice used, which happens with 50% probability. In this case the post-measurement state is necessarily the same as the original state prepared by Alice. So by reversing the encoding used by Alice he recovers the input bit for 50% of the pulses, in average. In other words, Bob identifies the original quantum state with certainty and hence the original bit.

For example, suppose Alice prepares in the X basis, and Bob measures in the same basis, let f the function that represent Alice’s encoding should she choose the X basis i.e $f(0) = |+\rangle$ and $f(1) = |-\rangle$. In this case Bob’s outcome is $f^{-1}(f(b)) = b$.

So assuming no attacker is present, there shouldn’t be more than 50% incorrect decodings, and they can check this without revealing their entire keys by ”sacrifying” half of their keys, a procedure that is called *sifting*. Note that sifting is rarely employed in CVQKD as we shall see later.

Having recast the protocol in Information Theory terms, the next steps are classical processing derived from the sequential key agreement model and results therein(such as using universal families to achieve privacy amplification), so the next steps are fully classical.

Several other protocols are possible expanding on this idea. For example, a six-state protocol is a natural generalization of this scheme, by using all three Pauli basis.

1.3 High level description of CVQKD

We provide a sketch of CVQKD with Gaussian modulation and using coherent states, which is arguably the most common flavour of CVQKD. In the following chapters, we analyze the protocol and its security in detail.

CVQKD is slightly different than the conventional discrete QKD protocol BB84 as its security principle is not based on a 50% chance (or more, for example in the six-state protocol) for the eavesdropper to measure in the wrong basis therefore introducing errors and leaving evidence. Instead, the encoding process (from the classical input to the quantum state) happens to be deterministic, but Alice´s classical input will be a pair of two reals numbers (or equivalently a complex number) encoded as the coherent state $|\alpha\rangle$. Coherent states are a family of states indexed by complex numbers which are defined in the next chapter. Despite this, the core principle BB84 is based on that is encoding information in a family of non orthogonal states is still the same. In fact, it can be shown that no two coherent states are orthogonal. This principle underpins its security as the no cloning theorem holds.

Instead of generating a sequence of i.i.d bits, Alice is generating pairs of real numbers (or equivalently, a sequence of complex numbers) that are still i.i.d. This is her (classical) input. We denote $(q, p) = (q_i, p_i), i \in \mathbb{N}$ her sequence. Both p_i and q_i are drawn from independent zero mean Gaussian distribution with a common variance $\tilde{V}_{mod}$.

She then encodes those numbers as quantum states where the ket labels are the generated numbers, in a certain family of quantum states called "coherent states". Alice generates the complex number and the corresponding coherent state so she encodes the complex number $\alpha = q+ip$ as $|\alpha\rangle$. Note that encoding is deterministic conditionally to the classical input, as opposed to the discrete QKD where she choses a random basis.

Next is Bob's measurement. Assume temporarily that $\hat{q}$ (resp $\hat{p}$) are physical observables. Bob's goal is to estimate either q or p. It turns out that when measuring using the operator $\hat{q}$, the resulting random variable is Gaussian variable with mean $2q$ and variance 1. The same is true for p and $\hat{p}$ So Bob observes a noisy version of Alice's data. Measuring both q and p is not possible but there is a third option, called heterodyne detection, which allows Bob to estimate a *linear combination* of q and p.

(these operators, called quadratures, are defined in the next chapter), it can be shown that their expectation correlates to Alice's input so that on the coherent state $\rho = |q+ip\rangle$

$$\mathbb{E}_\rho(\hat{p}) = p, \mathbb{E}_\rho(\hat{q}) = q$$

and

$$V_\rho(\hat{p}) = V_\rho(\hat{q}) = 1$$

Consequently, there is positive mutual information between Alice and Bob since these operators allow Bob to estimate Alice inputs. That is, given Alice's key (x_i) Bob can obtain a noisy version (y_i) by estimating either real or imaginary part, which by information-theoretical processing can be processed to derive a final key.

In fact, it can be shown that Bob‘s classical variables are related to Alice's in such a manner that they are an instance of the well-known **normal model** i.e

$$y_i = ax_i + z_i$$

with $z_i \sim \mathcal{N}(0, \sigma^2)$ that is independent of x_i Leverrier and Grangier [2010]

In reality, $\hat{p}$ and $\hat{q}$ don't correspond to physical devices, but can be emulated with the number operator using the principle of homodyne detection which will give us the same measurement statistics as if we were literally measuring using one of these quadrature operators.

There is also the fundamental question of calculating the mutual information, or an upper bound thereof, between Alice's input, and Bob's measurement (or an eavesdropper measurement), this is answered by Holevo's theorem as we shall see in the next chapter.

Among the information theoretical questions one could ask, the mutual information between Alice and Bob is crucial as it determines their key length. In fact, reverse reconciliation where Bob's data is regarded as the raw key and Alice corrects her observations, turns out to be more efficient than the conventional direct reconciliation.

2 Mathematical tools and formalisms

2.1 Basics of Information Theory

2.1.1 Information metrics

The (Shannon) entropy of a random variable X is the expected amount of information when observing the value of X. The information of an event A is defined as $i(A) := -\log p(A)$. So the entropy of X is the expected value of $i_X(X)$ that is:

$$\mathrm{E}\left[i_X(X)\right]$$

where

$$\mathrm{i}_X(x) := -\log\left[p_X(x)\right]$$

Or more explicitly;

$$\mathrm{H}(X) = -\sum_{i=1}^{n} \mathrm{P}(x_i)\log_b \mathrm{P}(x_i)$$

The entropy of X, of course, only depends on the distribution of X, but only through the set of probabilities $\{p_X\}$, not the individual values of the random variable . The choice of the basis of the logarithm is in principle arbitrary, but it is commonly taken as 2 in information theory, so the entropy is measured in *bits*.

The entropy can be viewed as the amount of bits required to encode X. This is not literally true considering only a single random variable whose entropy is n, because it can still take more than 2^n values each with positive probability (whereas the maximum distinct values one can encode with n bits is 2^n), but this interpretation is justified by the famous Shannon source coding theorem:

If we have n copies of X (i.e. n i.i.d random variables with the same distribution as X) which we denote X^n, then we can encode them using $nH(X)$ bits (which is the amount implied by the above interpretation) with a probability of error that tends to 0 as $n \to \infty$ so the probability of error becomes negligible.

Shannon defined the entropy for the purpose of coding, and also used it to justify some concepts of cryptography like the one-time pad, but it has found application to security as well, providing a useful mean to quantify the information an eavesdropper or attacker can gain from his observations.

The conditional entropy of X given that $Y = y$ is just the entropy of the marginal distribution of X given $Y = y$. It's noted $H(X|Y = y)$. The conditional entropy of X given Y is the expected value of $H(X|Y = y)$ where the expectation is taken over Y. In another words it is the average of $H(X|Y = y)$ weighted by the marginal distribution of Y and it can be conveniently written as (by multiplying these weights $p(x|y)$ by $p(y)$ to get the $p(x, y)$)

$$H\left(X|Y\right) = -\sum_{i,j} P(x_i, y_j) \log P\left(x_i|y_j\right)$$

In this form, it bears similarity (since it's a weighted sum with coefficients the joint probabilities of (X, Y) but is different from the joint entropy of X and Y, which is defined as just the entropy of the random vector (X, Y):

$$H\left(X, Y\right) = -\sum_{i,j} P(x_i, y_j) \log P\left(x_i, y_j\right)$$

The conditional entropy of $H(X|Y)$ can be thought of as the average amount of uncertainty left about X after observing Y. It is thus natural to expect that $H(X|Y) \leq H(X)$ and indeed this can be proven easily (**conditionning does not increase entropy**). However it turns out there are other entropies (than Shannon's) where conditionning *might* increase entropy. Another important formula is the chain rule, which is not much surprising with the above interpretations:

$$\mathrm{H}(X, Y) = H(X|Y) + H(Y)$$

Since the "worst case" (from the perspective of someone looking to estimate X, but is actually the best case against an eavesdropper) is $H(X|Y) = H(X)$ (from the conditionning does not increase entropy property), it makes sense to define the information Y gives us about X as the *difference* between $H(X)$ and $H(X|Y)$(such difference is non-negative), so that a value of 0 corresponds to a complete lack of information, and the highest value(which is $H(X)$) corresponds to observing Y reduces uncertainty about X completely:

$$\mathrm{I}(X, Y) := \mathrm{H}(X) - \mathrm{H}(X \mid Y)$$

The mutual information is symmetric and is in fact an ultimate bound about the length (in bits) of a common key Alice and Bob can extract, as precised by the communication over noisy or wiretapped channels theorems.

Fano's inequality lies at the heart of information theory as it links entropy to a concrete and operational quantity : the probability of error when estimating X

$$H(X|Y) \leq H_b(e) + P(e) \log(|\mathcal{X}| - 1)$$

where $\mathcal{X}$ is the set of values X can take, and H_b is the binary entropy function.

A generalization of Shannon entropy is possible, even though some properties may be lost (conditionning may increase entropy). This is given by Rényi family of entropies:

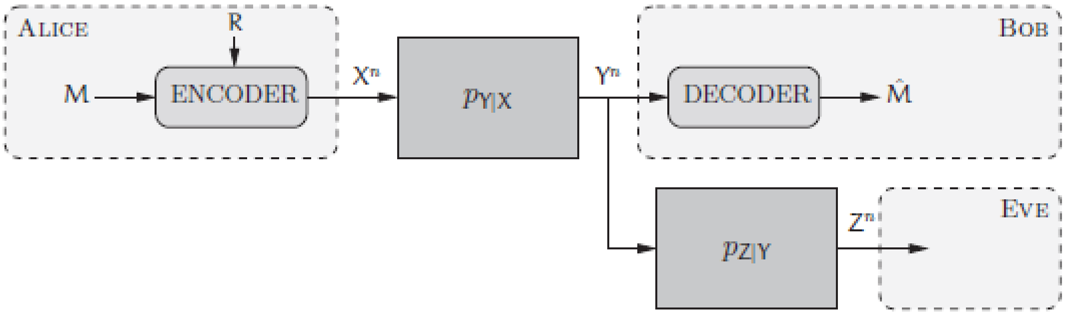
$$\mathrm{H}_\alpha(X) = \frac{1}{1-\alpha} \log \left(\sum_{i=1}^{n} p_i^\alpha \right)$$

The special case $\alpha = 2$ is called collision entropy and is particularly relevant to QKD as it serves to analyze the last step, privacy amplification.

2.1.2 Wiretap models

In his seminal paper Wyner applied entropy to a different problem than coding.

Imagine two parties Alice and Bob wishing to communicate while ensuring confidentiality, over a channel subject to eavesdropping by Eve (passive attacker). After emitting X^n (i.e after Alice encodes her message M into X^n), Bob and Eve observations are correlated to X^n through some specified pdf $P(y, z|x)$. This pdf is specified for a single symbol, but is extended to X^n, Y^n and Z^n by considering all the symbols independent. So after emitting the value x the probability that Bob observes y and Eve z is $P(y, z|x)$. The challenge is therefore defining suitable encoding schemes that maximize the MI between Alice and Bob while minimizing the one between Eve and Alice. The pdf depends on the channel but is not considered secret, i.e the statistics of the channel are known by the parties. In the simplest case, called the **degraded wiretap channel**, Eve is at explicit disadvantage because X^n gets corrupted to become Bob's observation Y and is this Bob's observation that gets corrupted again(specified by some pdf $p(z|y)$ to give Z^n. This means that the pdf $p(y, z|x)$ factors as $p(y|x)p(z|y)$ in another words $X^n \to Y^n \to Z^n$ is a Markov chain. **This places Eve at an explicit disadvantage** because of the data processing inequality that states that her mutual information with Alice is less than Bob's mutual information with Alice. Illustrated below is the DWTC:

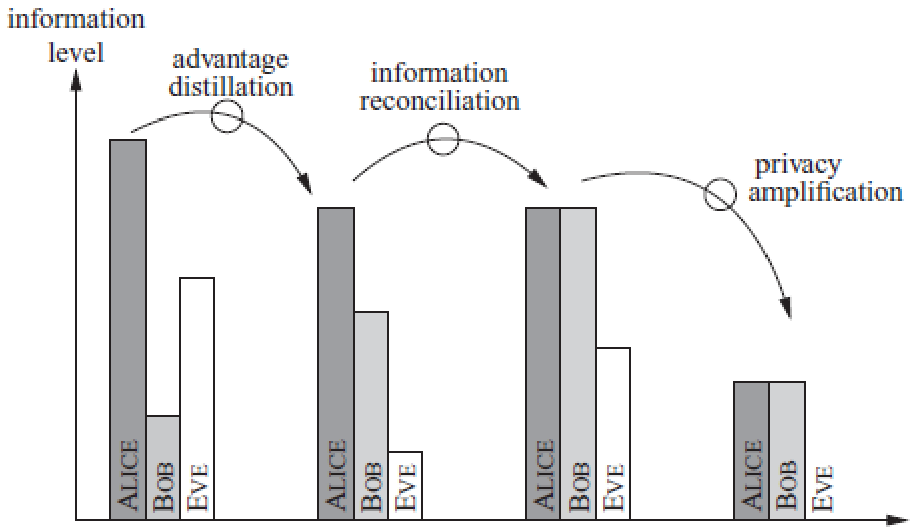


2.1.3 Key agreement from a source, and sequential key distillation

More relevant to QKD is the Key agreement from a source. This model looks similar to Wyner's wiretap model, but the difference is that Alice is not interested in sending a *speficic* message M. Instead, the goal for Alice and Bob is to end up with a common sequence S("agreement"), Eve has no information about. Unlike a message M, S is essentially random and unknown to Alice ahead of time and thus carries no information by itself, so we call it a key. There is still a secrecy requirement from Eve, so $I(S; Z^n)$ should be as small as possible. Furthermore, we assume the existence of a public authenticated channel that enables one-way communication. This shared key can be used as input to a classical cryptography algorithm, such as AES.

The process of extracting such key from the observations of the realizations of three

Figure 1: Evolution of information through key distillation steps



correlated random variables (X^n, Y^n, Z^n) is called **Key Distillation**, and its analysis is not trivial. However the analysis can be greatly simplified by assuming that Alice and Bob communicate through one-way messages through a public authenticated channel, and by assuming that key distillation is done in *stages*, each stage being concerned with a specific goal.

More precisely, we could define a new triplet of random variables (X', Y', Z') from (X, Y, Z) by transforming it (i.e. applying functions, which may take an additional random parameter). For example in a hypothetical scenario we can define $Y' = Y + sha256(X)$ and Bob would be still able to calculate it (it is as if he was observing a **new random variable**) as the value of $sha256(X)$ can be disclosed through the public channel without fully disclosing Alice's observation trivializing the problem. We can thus study the mutual information between these new random variables.

Sequential Key Distillation consists precisely of a succession of such well-chosen transformations. After each stage, the mutual information evolves to closely resemble our ultimate goal, that is $I(X^n; S) = I(Y^n; S) = H(S)$ (agreement on a common sequence) and $I(S; E) = 0$ (secrecy from Eve) where E is the random variable representing Eve's knowledge up to that stage. This knowledge includes the information transfered through the public channel.

Figure 1 from Bloch and Barros [2011] summarizes those steps in high level manner.

The first stage called advantage distillation seeks at correcting the potential imbalance between the Mutual Information between Alice and Bob and between Alice and Eve. The former should be greater than the latter, or else the bound for secrecy capacity is useless and so are the algorithms that achieve it.

The second stage is concerned with reconciliation: The objective is that with probability almost 1, Bob can reconstruct X^n knowing Y^n and the information communicated in the public channel. When this is achieved, then the common sequence becomes $S = X^n$ (it is a random variable). There is no secrecy concern at this point so no requirements on $I(S;E)$, all we need is Alice and Bob observing the same random variables.

The third stage is called Privacy Amplification where the final sequence S' is derived from S by appropriately transforming it. It turns out this can be done using families of Hash function known as universal families. A family of hash functions from $\mathcal{A}$ to $\mathcal{B}$ is called universal if it satisfies the following property for any x_1 and x_2 in its domain : By choosing a hash function at random, the probability that it takes the same value on x_1 and x_2 is less than $\frac{1}{|\mathcal{B}|}$ where |B | denotes the cardinality of the set, . We denote G the random variable (function valued) arising from such draw.

Since the chosen hash function G is disclosed via the public channel, further analysis of Eve's information about the sequence must take into account this fact, so her conditional entropies are conditioned on G as well.

Let E be the random variable representing Eve's knowledge before this step, The objective is thus to prove that $H(K|E=e,G)$ is close to k since K is a string of k bits, this implies K is almost uniformly random for Eve. The main tool for this step is Bennet's theorem, as it gives a lower bound on $H(K|E,G)$. However, it involves another entropy than Shannon's, called the collision entropy. The collision entropy is defined as

$$\mathrm{H}_2(X) = -\log \sum_{i=1}^{n} p_i^2 \leq \mathrm{H}(X)$$

where Y is a copy of X, justifying the name "collision". Bennet's theorem gives a lower bound on $H(K|E=e,G)$ in terms of the same main variable K and conditionning variables/values $E=e$ and G but using the collision entropy

$$H(K|E=e,G) \geq H_c(K|E=e,G) \geq k - \frac{2^{k-c}}{\ln 2}$$

So by chosing k much smaller than c, Eve's conditional entropy is close to k, so K seems a uniform random variable.

2.2 Basics of Quantum Mechanics

2.2.1 The four postulates

The four postulates of Quantum Mechanics define a mathematical framework, but the physicist still has to identify the appropriate states, operators, Hamiltonian etc.

1. State space postulate The first postulate posits that the state of any system can be fully specified as an element of a well-chosen complex Hilbert space, of norm 1. Furthermore, the global phase is irrelevant, that is $x \in H$ and $e^{i\theta}x$ both denote the same physical state.

Another way to formulate this is that a physical state is an element of a *projective* Hilbert space, since by definition equivalence classes, called rays, are in 1-to-1 correspondence with norm 1 states when ignoring the global phase. The individual components of these vectors(complex numbers) are called **probability amplitudes**, for a reason made clear by the measurement postulate.

It is common to write such a state with the ket notation $|\psi\rangle$ The ψ is just a label whose meaning depends on the context. A basic fact of Hilbertian analysis is that $\mathcal{H}$ is its own dual, that is each element $|x\rangle$ of $\mathcal{H}$ induces a continuous linear form given by the inner product with x: $\psi \rightarrow \langle x|\psi\rangle$.

In the finite case, after fixing a basis, we can view $|\psi\rangle$ as a column vector of complex numbers, $\langle\psi|$ can therefore be seen as the row vector containing the same values but conjugated, that is, the Hermitian transpose of $|\psi\rangle$.

The simplest non trivial Hilbert space is the one of dimension 2. All Hilbert spaces of the same finite dimension are isomorphic, that is essentially the same for the purpose of linear algebra, so we can consider this space to be $\mathbb{C}^2$. Such an element is called a **Qubit**. Qubits are the bread-and-butter of discrete-variable quantum computing and quantun information theory, arguably the most common and studied paradigm. The other one being continuous variable quantum computing and information theory, which is more challenging mathematically, but sometimes easier to implement. Since a qubit register of n qubits can be seen as qudit with $d = 2^k$, so registers of qubits can simulate any qudit (where we call qudit the elements of the d-dimensional complex Hilbert space). The formula $d = 2^k$ might seem surprising as classically one might except $d = 2k$, but the reason for why such exponentiation is given by the 4th axiom. Intuitively, we see that qubits can "hold" much more information due to superposition, although retrieving it is not trivial as we show later in the measurement postulate and Holevo theorem.

It is also important to consider infinite dimensional Hilbert spaces as they arise in some situations : for example a particle's state can be modeled as an element of L^2 (informally speaking a function representing the probability to find the particle at a given location when modulus squared). More relevant to this thesis are quantum states of light, which also live in infinite-dimensional Hilbert spaces and can in fact be represented by tensor product of (separable) Hilbert spaces. They are used in quantum computing as well.

As a concrete example of the realization of a qubit, we can consider the spin of an electron, or the polarization of a photon.

2. Evolution postulate The time evolution of a quantum system, *in the absence of measurement and interaction with environment*, is fully deterministic and even reversible, which does not display quantum "weirdness". It can be described by a unitary operator U, so that starting from a state $|\alpha\rangle$ at time t_1, the state of the system at time t_2 is $U|\alpha\rangle$ where U only depends on the time.

   The postulate does not specify the form of the unitary U which is delegated to the physicist. A more down to earth formulation is the Schrodinger equation which explicitly introduces the Hamiltonian:

$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle$$

Both formulation are, in fact, equivalent.

3. Composite systems postulate The state space of a joint system of two systems each of which would be described individually by $\mathcal{H}_A$ and $\mathcal{H}_B$, shall be described by *tensor product* of Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$: $\mathcal{H}_A \otimes \mathcal{H}_B$ Moreover if the two systems have never interacted then the state of the system is $|\psi\rangle_A|\phi\rangle_B$ The tensor product can be defined rigorously via its Universal Property, however for most purposes it is enough to see it as all the expressions of the form, where (e_i) and (f_i) are basis of $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively.

   $$\sum_{i\in I, j\in J} c_{ij}|e_i\rangle \otimes |f_j\rangle$$

   subject to the normalization condition

   $$\sum_{i\in I, j\in J} |c_{ij}|^2 = 1$$

   By manipulating such expressions as if the symbol $\otimes$ was the usual multiplication, we get alternative expressions of a given state, for example if it turns out that $\forall i, j\ c_{ij} = a_i b_j$ then the usual factorization of a sum of products gives us:

   $$|\psi\rangle = \sum a_i e_i \otimes \sum b_i e_i$$

   In such case, the two systems can be studied independently and there is no further need for this postulate. Such states are called separables.

   Even if elements of $\mathcal{H}_A$ and $\mathcal{H}_B$ are not in general separable, the shares can be at two different physical locations(say Alice's and Bob's laboratories), which brings the question: What is the result of applying a measurement represented by an operator M on $\mathcal{H}_A$ on the first share? This question can be answered by the tensor product of operators. The product operator of M and N is defined as the unique operator such that

   $$\forall u, v\ M \otimes N|uv\rangle = M|u\rangle \otimes N|v\rangle$$

   The uniqueness is trivial because such condition constrains the values it takes on every$e_i \otimes e_j$, and the existence follows from developing the expressions of

   Note that the order of measurement doesn't matter.

   When dealing with composite systems, the Kronecker product of matrices turns out to be handy, as the density matrices can be obtained by taking the Kronecker product of subsystems. Density matrices are defined in the next chapter.

4. Measurement postulate Measurement is perhaps the most surprising postulate along with the composite system postulate as it introduces inherent randomness to the theory, contrasting with classical physics where evolution is always deterministic, in principle.

   Several interpretations exist to explain this randomness, such as the Many Worlds interpretation which has the advantage of eliminating randomness by simply speculating all outcomes are in fact realized, but in different universes. But one needs not to care about the interpretations since we can still calculate probabilities which is arguably all we need to make predictions. This is the point of view of the Copenhagen interpretation jokingly referred to as ”Shut up and calculate”.

   In fact, this inherent randomness is the main resource exploited by quantum security protocols. The BB84 protocol, for example, does not even make use of the evolution or the composite system postulates.

   This states that the value of an observable quantity (like the position, or momentum) is random, even if the state is purely determined physically according to the first postulate (by a state vector or wavefunction). However the probabilities are well defined (given the underlying state) by the theory and therefore we can still make ”predictions”. There are two kinds of measurements considered: projective measurements are the most common, and POVM are an abstract generalization.

   A protective measurement is defined by a set of real numbers m_i representing the possible measurement outcomes, and a collection of corresponding states (kets) $|i\rangle$ According to this postulate, the probability of observing outcome m_i given that the underlying state is ρ is the **modulus squared** of the coordinate of ψ along the basis. Since this is an orthogonal basis, these coordinates are simply inner products. Moreover the postulate states that after measurement, the state will *collapse* into $|i\rangle$. This is called the Born rule. An obvious consequence of is that measurement is repeatable : $\langle \imath|\imath\rangle = 1$ so we will keep observing the value m_i if we repeat the measurement on the post-measurement states.

   Conversely, any Hermitian operator can be written in such fashion and thus seen as a measurement, thanks to the spectral theorem.

   An important consequence of this axiom is that the order in which we perform measurement matters, but it can be shown that as long that the observables commute, the order does not.

   The expectation of this measurement (on a given state $|\psi\rangle$) is defined as the expectation of the induced random variable. So it is the average result we would observe if we repeated the measurement on multiples copies of the state, it is denoted $\langle A\rangle$. We can write it $\langle\psi|\hat{A}|\psi\rangle$ and this can be seen by grouping the basis elements into columns of a matrix N, and by defining the diagonal matrix D whose diagonal elements are the measurement outcome.

   An alternative representation of measurement that is widely used is through Hermitian operators. For each basis elements we can define the projector Π_i. Define the Hermitian operator $\sum \Pi_i$. This Hermitian operator contains all the information about the measurement outcomes (the probability distribution). Conversely, any Hermitian operator can be written in such fashion and thus seen as a measurement, thanks to the spectral theorem. In this context, we often call them **Observables**. The observable representation is handy as for example the expectation of an operator M on a state $|\psi\rangle$ is just $\langle\psi|M|\psi\rangle$

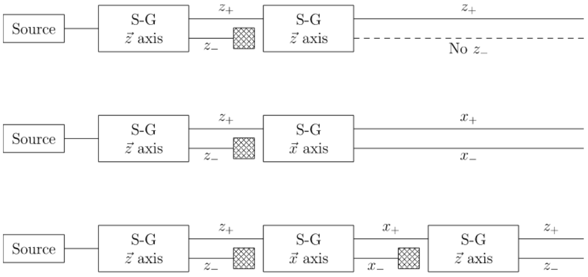
2.2.2 Motivation for the measurement postulate: The Stern-Gerlach experiment

When trying to measure the deflection of electrons along the Z axis, physicists observed that although this deflection should have been 0 as they have 0 angular momentum, particles still deflected as if they had an angular momentum, furthermore these values were quantized i.e taking only a discrete set of values, in fact just 2 : $1/2$ and $-1/2$. This suggests that particles posses an intrinsic property (like the mass or charge) known as Spin(a priori one for each axis), but at this stage one can still assume the randomness might simply be due to the preparation of the particles as an *ensemble* with half the particles having spin $-1/2$ and the other $1/2$ However, when cascading the experiments, results that are counter-intuitive showed up. Let's consider an apparatus where we first measure the spin along the Z axis. According to our classical intuition, no matter which measurement we perform next, this should not affect the value already measured(”realism”).

It turns out that even while filtering say particles with Z spin $1/2$, these particles reappaear with probability $1/2$ after measuring them again, with a measurement of the X spin in between. So we have the apparent contradiction of electrons having an intrinsic property, but whose value can change.

The measurement postulate provides a clear explanation of these observations: If we assume that the eigenbasis of X expressed in the Z basis are the sum and differences of them, then the Born rule correctly predicts that the outcomes of the last Z measurement should be a $Bern(1/2)$ In fact we say that the Z and X are mutually unbiased bases, and it is precisely on this fact that BB84 is based.

Figure 2: Stern-Gerlach apparatus



2.2.3 QKD as an instance of the secret key agreement from a source model

QKD can effectively be seen as an instance of secret key agreement from a source model, as we can clearly model X, Y and Z as random variables and find their joint distribution once Eve's intervention has been modelled. Let's say that Eve measures in the Breidbart basis, which eigenstates are the qubits obtained by rotating the computational basis by $\pi/8$:

$$|\phi_0\rangle = \cos(\pi/8)|0\rangle + \sin(\pi/8)|1\rangle, |\phi_1\rangle = -\sin(\pi/8)|0\rangle + \cos(\pi/8)|1\rangle$$

We could assume that the components of X^n are i.i.d Bern(1/2), it remains to explicit the conditional distribution of (Y^n, Z^n). This can be done by separating cases, that is further conditioning on the choices of basis. We only examine the case of one symbol since they are i.i.d.

For example if we wish to calculate $p(Y = 0, Z = 0|X = 0)$, we can further condition on the basis chosen, there are 4 cases. Given that Alice has chosen the Z basis, Bob the Z basis, and Eve the Breidbart basis, we find by applying the Born rule (inner product of the first vector of the Z basis, first vector Breidbart basis)

$$p(Z = 0|X = 0, bobbasis = Z, alicebasis = Z) = |\cos\frac{\pi}{8}|^2$$

Then the probability of Bob still observing 0 after the intervention is again with the Born rule

$$p(Y = 0|Z = 0, X = 0, bobbasis = Z, alicebasis = Z) = |\cos\frac{\pi}{8}|^2$$

So we have calculated $p(y, z|x, C)$ given some condition C, then the usual law of total probabilities(conditional case) can be used to collect them to get $p(y, z|x)$. Here the set of conditions is all the possible choice of basis, which happens with probabilities 1/4

So by a weighted sum of such conditional probabilities(here the weights would be 1/4 since each of the 4 choices of basis are equiprobable), we can explicit the distribution(X, Y, Z) and the theoretical results of the secret key agreement from a source, hold.

2.2.4 Density matrix : A useful formalism for mixed states

Density matrices provide a powerful alternative to state vectors for dealing with mixed states.

A mixed state is just a superposition, in the classical sense, of pure states(i.e. a probability distribution over pure states, capturing the fact that the pure state is unknown). So it is described with a set of probabilities p_j, and the corresponding pure states ψ_j. So with probability p_j, the system is in the state ψ_j.

Mixed states allow for some combinations of observables and probabilities distributions that wouldn't occur using pure states alone. For example, consider the mixed states that is in either $|0\rangle$ or $|1\rangle$ with probability 1/2. When measured in the Z basis, the probabilities of obtaining 0 and 1 are obviously 1/2 respectively. This is similar to the pure state $|+\rangle$, but when measured in the X basis the situation is different : $|+\rangle$ will obviously always yield the same outcome since it is an eigenstate of X, while the mixed state has $1/2$ chance to be in either the Z basis states, and these states projected on the X basis have the same probabilities. This behaviour isn't possible using a pure state since such pure state will be similar to $|+\rangle$ (with a relative phase difference).

Another motivation for density matrices is that a mixed state can be represented in a compact manner. Indeed, consider a mixed state, it could be fully specified by the set of pure states it can be in and their respective probabilities. Based on this knowledge, we can calculate the probabilities for an arbitrary observable by simply applying the law of total probabilities as in classical probabilities

$$p(m) = \sum_j p_j \langle \psi_j | \Pi_m | \psi_j \rangle$$

However the above sum involves an arbitrary large number of terms.

Even by considering a single qubit, the specification of single mixed state could be arbitrary long. Taking advantage of the well-known identity $tr(AB) = \langle x, Ax \rangle$ where B is the projector associated to x i.e $B = x^t x$ and by taking out the trace from the sum, which is justified since the trace is linear, the formula becomes

$$p(m) = \sum_j p_j \langle \psi_j | \Pi_m | \psi_j \rangle = \mathrm{tr}\left[\Pi_m \left(\sum_j p_j |\psi_j\rangle\langle\psi_j| \right) \right]$$

which only has d^2 *complex degrees of freedom for a single qudit that are the entries of the matrix* $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ *(in fact less since the matrix is hermitian)* For example for a qubit is is fully determined by two complex numbers (since it is Hermitian) in this formalism instead of two arbitrary large sets as the naive approach. The expectation of an operator M in this formalism can be shown to be $tr(M \cdot \rho)$ The density matrix (or operator, since it generalizes to arbitrary Hilbert spaces) also have some additional properties which are trivial from its definition : it is positive semidefinite and has a trace 1.

Since every ensemble correspond to a PSD operator of trace 1 and vice-versa, we can take density matrices/operators as the fundamental definition of "state" in quantum mechanics, and the others postulates can be reframed with density matrices. For example, the measurement postulate in this form tells us that the expectation of an observable A is $tr(M\rho A)$.

It is easy to verify that the density matrix of a product state in $\mathcal{H}_A \otimes \mathcal{H}_B$ is obtained by the Kronecker product of the two states.

2.2.5 Partial trace and purification

Motivation : Reduced states

Consider a general state ρ on $\mathcal{H}_A \otimes \mathcal{H}_B$. Physically, we can apply operators defined on $\mathcal{H}_A$ to the first share even though the two subsystems may be entangled and impossible to describe by describing each one individually (as respectively a mixed state of $\mathcal{H}_A$ and a mixed state of $\mathcal{H}_B$. So there is no a priory defined state ρ^A on $\mathcal{H}_A$ which gets us the measurement statistics of applying an operator M of $\mathcal{H}_A$ (through $\mathrm{Tr}(M \cdot \rho^A)$) but we still have a mathematically well defined operator on the product space given by $M \times I$ which is the one matching the experience.

It is therefore natural to ask if there is a state ρ^A which when measured by M would give us statistics *as if* the first subsystem was in ρ^A. In other words, ρ^A must fulfill the following condition, for any operator M:

$$\mathrm{Tr}(M \cdot \rho^A) = \mathrm{Tr}(M \otimes I \cdot \rho)$$

Such state would allow us to view the system as if it wasn´t entangled. The answer is positive (and unique) and is given by the **partial trace**. **Partial trace** The partial trace over A is defined as a liner application that maps operators of $\mathcal{H}_A \otimes \mathcal{H}_B$ to operators on $\mathcal{H}_A$ in the following manner

Given such application T, we can fix i and j and consider the linear map F_{ij} that maps u to a vector v by applying T to $e_i \otimes u$ and having decomposed this image as $T(e_i \otimes u) = \sum_q e_k \otimes v_k$ we define $v = v_j$. The trace of this application is denoted f_{ij}.

The expression of f_{ij} is trivial to calculate from the coefficients of T since the coefficients of the tensor T with i fixed reappear in $T(e_i \otimes u)$ and by keeping only the term $e_j \otimes v_j$, we are fixing two coefficients of the tensor so we are left with a sum along one indice:

$$F_{i,j} = \sum_{k=1}^{n} T_{i,k}^{k,j}.$$

The partial trace is the linear application of $\mathcal{H}_A$ defined by such coefficients f_{ij}. It is then easy to see that this definition is equivalent to this coordinate-free definition : $\mathrm{Tr}_W(R \otimes S) = \mathrm{Tr}(S)\, R \quad \forall R \in \mathrm{L}(V) \quad \forall S \in \mathrm{L}(W)$.

Now back to the original motivation, we can interpret ρ as an operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ (which has additional properties : positive semidefinite with trace 1). If we define ρ^A as its partial trace, then the sought-after identity will hold, for any operator M :

$$\mathrm{Tr}(M \cdot \rho^A) = \mathrm{Tr}(M \otimes I \cdot \rho)$$

This is because

$$\begin{aligned} \mathrm{Tr}(M \cdot \rho^A) &= \mathrm{Tr}(M)\,\mathrm{Tr}(\rho^A) \\ &= \frac{1}{dim\mathcal{H}_B}\,\mathrm{Tr}(M \otimes I) dim\mathcal{H}_B\,\mathrm{Tr}(\rho) \end{aligned} \tag{1}$$

where $dim\mathcal{H}$ denotes the dimension of Hilbert space $\mathcal{H}$.

2.2.6 Purification

An interesting feature of the partial trace is that even a pure state on $\mathcal{H}_A \otimes \mathcal{H}_B$ can become a mixed state on $\mathcal{H}_A$ when tracing out $\mathcal{H}_B$. A natural question is thus if any mixed state on $\mathcal{H}_A$ can be written as the partial trace of some pure state on $\mathcal{H}_A \otimes \mathcal{H}_B$ for some Hilbert space $\mathcal{H}_B$. The answer is yes, although such purification is not necessarily unique. Analysis of the possible purification show that all purifying state differ only by an unitary operator **which is crucial to the security analysis of CVQKD as we will see**

Obtaining a purification of a mixed state is in principle easy because we can just tensor each element of its ensemble with the respective state of an orthonormal basis of the ancilla space, and the partial trace will ”remove” this basis because $\mathrm{Tr}_K(M) = M$.

2.3 Quantum Information theory

2.3.1 Encoding classical variables

Classical states can simply be encoded as elements of a chosen family of orthogonal quantum states. Random variables are convex combinations(i.e linear combinations with positive coefficients summing to 1). So their density matrix is diagonal.

2.3.2 Quantum channels

A quantum channel is a formalization of the evolution of a quantum state(represented by a density matrix) in the most general manner. Such evolution could include but is not limited to, unitary evolution and measurements. There are two usual ways to define it, both equivalent. The constructive approach is based on the observation that one kind of transformation a system ρ can undergo is through interaction with the outside environment, represented by an ancilla a, then the global system will undergo some unitary transformation. At this stage if we measure our system, the answer will be given by the partial trace as explained in the Partial trace section before.

So from this observation, one kind of quantum channel must be those of the form:

$$\mathcal{E}(\rho) = \mathrm{Tr}_K U(\rho \otimes a)U^\dagger$$

In fact, every quantum channel can be written in this form.

An alternative but equivalent definition of quantum channels is possible. It is more axiomatic as it a set of conditions a map $\mathcal{E}$ must verify instead of an explicit construction. First, since density operators are positive, $\mathcal{E}$ must be positive, and it must be trace preserving since density operators have trace 1. Finally, these properties must hold for the induced application $\mathcal{E} \otimes Id_n$.

Such maps are called completely positive trace preserving (CPTP), these two definitions of quantum channels are equivalent by Stinespring’s dilation theorem.

2.3.3 Holevo Theorem

Consider the situation where we have a mixed state, but we don't know which one. That is, the mixed state is drawn at random from a collection $\{\rho_1, \rho_2, ...\rho_n\}$. The total state is therefore $\sum p_i\rho_i$. The challenge is thus identifying i, we can therefore ask the following question: What POVM would maximize $I(X;Y)$ (Y being the r.v representing the outcome of the measurement, and X the r.v whose outcomes are the indexes i)? An upper bound is known and is called Holevo theorem. This bound involves the von Neumann entropy:

$$I(X;Y) \leq H(\rho) - \sum_i p_i H(\rho_i)$$

2.4 Quantum optics for CVQKD

2.4.1 Quantum states of light

The main difference between discrete and continuous quantum key distribution is that the underlying Hilbert space is infinite dimensional. The analogue of a qubit is one mode of light, sometimes called qumode. Just as qubits can be represented as elements of two-dimensional complex Hilbert space(and multiple qubits by elements of the tensor product of such spaces) with a chosen orthogonal basis, we can represent states of light with an Infinite dimensional separable Hilbert space, with a chosen Hilbert basis. A rigorous derivation of this space requires the quantization of the electromagnetic field, however we assume the space is given and only present the main operators and types of states. The next section sheds some light as why operators are defined as such. Although it is an infinite dimensional space, is has the topological property of being separable, which according to basic Hilbert theory implies that a Hilbert basis can be found, and in this basis, called Fock basis, any state can be expressed as an infinite linear combination of the Fock states.

It should be noted that this is not a basis in the sense of linear algebra since an algebra basis always has a *finite* cardinality. Here we deal with infinite sums as well. That is, the meaning is that of a limit, and not a usual finite sum. It is a well-known fact that $\sum_{n=0}^{\infty} c_n|n\rangle$ converges if and only if $\sum_{n=0}^{\infty} c_n^2 < \infty$ so even though the Hilbert space is an arbitrary infinite dimensional separable vector space with inner product, we can still see its elements as sequences of complex numbers(albeit infinite) who are square summable just like an arbitrary vector space of dimension n can be seen as the set of sequences of length n, and in particular just as a qubit can be seen as a couple of complex numbers(up to normalization and global phase).

In this basis, we can runiquely define two important operators called the **ladder operators**, or creation and annihilation operators :

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$$
$$\widehat{a^\dagger}|n\rangle = \sqrt{n+1}|n+1\rangle$$

As the notation hints, it is easy to check that they are adjoints of each other. The name ladder or creation/annhilation is justified as they add or remove an element of the Fock basis. In fact, the Fock basis correspond to number of photons, so they can be interpreted as literally adding or removing a photon.

The photon number operator is defined as

$$\hat{n}|n\rangle = n|n\rangle$$

From the definition, it is clearly an observable. It can be used to measure(count) the number of photons as the Fock states are its eigenstates, and the corresponding eigenvalue of $|n\rangle$ is precisely n.

This observable is actually implemented as photon detectors, but is experimentally challenging, so what we usually have is the simpler POVM $\{|1\rangle, \mathbb{1} - |1\rangle\}$

and it is trivial to see that $\hat{n} = \widehat{a^\dagger}\hat{a}$ which is an important factorization of the number operator.

Coherent states, which form the backbone of most CVQKD protocols are defined as the eigenvectors of annihilation operators, it it possible to find their expression in the Fock basis as :

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

This expansion is easy enough to derive, as the coefficient c_n must divide by $\sqrt{n}$ the previous coefficient(to cancel the $\sqrt{n}$ introduced by the definition of $\hat{a}$ and multiply by $|\alpha$(since α is the desired eigenvalue. The outer coefficient follows from normalization. The above expansion shows (Born rule) that the number of a photons follows a Poisson distribution.

The quadrature operators are defined as

$$\hat{p} = \hat{a} + \hat{a}^\dagger$$
$$\hat{q} = -i(\hat{a} - \hat{a}^\dagger)$$

The name quadrature comes from the fact that the expectation value of these operators applied to coherent states, viewed as a function of time, is offset in phase by a quarter. Some other definitions exist for the quadrature operators, which only differ by a scale parameter (i.e. different units). The convention used here is called **Shot-Noise** because the uncertainty on coherent states is 1 (Heisenberg uncertainty principle).

$$\delta\hat{q}\delta\hat{p} \geq 1$$

The quadratures operators are also called position and momentum-like operators(as the notation hints), because they satisfy the same commutation relation as the position and momentum operators in quantum mechanics, the so called Canonical commutation relation (CCR).

$$[\hat{q}, \hat{p}] = 2i$$

As such, their eigenstates form a continuous basis for the Hilbert space.

$$|\psi\rangle = \int dx\psi(x)|x\rangle$$

The quadratures operators can be seen as the real and imaginary parts of a coherent state (of its complex ket label) because of the fundamental relation giving the expectation of these quadrature operators on coherent states$|q + ip\rangle$:

$$\langle\hat{q}\rangle = q, \langle\hat{p}\rangle = p$$

They are defined as such because we can prove that they minimize the uncertainty product on the quadratures $\delta\hat{p}\delta\hat{q} = 1$. In fact the uncertainty is always at least 1 on any state.

2.4.2 Physical origin: Quantization of the electromagnetic field

The above mathematical framework is sufficient for information theory, but we could also have a look at its origin. Following Leonhardt et al. [1997] we can postulate that the electromagnetic field is given by:

$$\hat{E} = u^*(x,t)a + u(x,t)a^*$$

We then replace the amplitudes by *operators* $\hat{a}$ and $\widehat{a^\dagger}$ subject only to a certain commutation relation: We assume the operators a and $a^\dagger$ to be *bosonic*:

$$[a, a^\dagger] = 1$$

From the commutator alone, we find that $\hat{a}$ and $\widehat{a^\dagger}$ are actually creation and annihilation operators in some Hilbert basis, so the previous mathematical section was backwards as we introduced the Fock basis first, then the ladder operators. In fact, denoting $\hat{n} = a^\dagger a$(number operator), and let $|n\rangle$ the eigenstate with eigenvalue n, we find that indeed $\hat{a}$ lowers its eigenvalue:

$$\hat{n}\hat{a}|n\rangle = \hat{a}^\dagger\hat{a}^2|n\rangle = (\hat{a}\hat{a}^\dagger\hat{a} - \hat{a})|n\rangle = (n-1)\hat{a}|n\rangle$$

So $\hat{a}|n\rangle$ is an eigenstate with eigenvalue $n - 1$ so it is a multiple of $|n - 1\rangle$. Since $|n - 1\rangle$ has norm 1, this scalar must be the square root of the norm of $\hat{a}|n\rangle$, which is $\sqrt{n}$.

One also needs to prove that the annhilation of the vacuum(i.e the eigenstate of the number operator with eigenvalue 0 is the null vector, this is easy as the norm of $a|0\rangle$ is obtained by multypling it with its conjugate which gives us $n|0\rangle = 0$ so it must be the null vector as its norm is 0.

Another reason we are interested in the number operator besides that its eigenstates give us a Hilbert basis, is that the Hamiltonian of the mode is just $\hat{n} + 1/2$

2.4.3 Generating coherent states: q/p modulation

Coherent states can be generated starting from an initial state α by repeated use of a beam splitter. A beam splitter has two inputs and two outputs, it can be shown that when the inputs are two coherent states, the outputs will be two coherent states as well whose eigenvalues are related to the input eigenvalues through the following matrix:

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

So in particular when the second input is a vacuum, we end up with the following coherent states:

$$\frac{1}{\sqrt{2}}\begin{pmatrix}\alpha\\ \alpha\end{pmatrix}$$

Repeating the operation by pairing these coherent states with two vacuaas, i.e by applyting two beamsplitters to the following vectors of coherent states:

$$\frac{1}{\sqrt{2}}\begin{pmatrix}\alpha\\ 0\end{pmatrix}$$

We get

$$\frac{1}{2}\begin{pmatrix}\alpha\\ \alpha\\ \alpha\\ \alpha\end{pmatrix}$$

Now by the opto-electric effect, we can shift the phase in each arm, we apply a phase of $\pm\phi_1$ to the first two arms and $\pm\phi_2$, so we end up with

$$\frac{1}{2}\begin{pmatrix}\alpha e^{i\phi_1}\\ \alpha e^{-i\phi_1}\\ \alpha e^{i\phi_2}\\ \alpha e^{-i\phi_2}\end{pmatrix}$$

The first two outputs and the last two outputs are fed to two beamsplitters yielding:

$$\frac{1}{\sqrt{8}}\begin{pmatrix}\alpha e\left(^{i\phi_1}+e^{-i\phi_1}\right)\\ \alpha e\left(^{i\phi_1}-e^{-i\phi_1}\right)\\ \alpha e\left(^{i\phi_2}+e^{-i\phi_2}\right)\\ \alpha e\left(^{i\phi_2}-e^{-i\phi_2}\right)\end{pmatrix}$$

We only keep the first and third states, and add a phase of $\pi/2$, so we have:

$$\frac{1}{\sqrt{2}}\begin{pmatrix}\alpha\cos(\phi_1)\\ \alpha i\cos(\phi_2)\end{pmatrix}$$

2.4.4 Measurement : Homodyne and Heterodyne detection

Measurement is a key step in QKD protocols, as we want Bob's values to be correlated to Alice's. If we see $\hat{p}$ and $\hat{q}$ as measurement operators(observables) then these can be used to estimate Alice's input as seen in the introduction.

The principle of homodyne detection is to "emulate" these quadratures, using the physically observables photon number opertors $\hat{n}$.

This is possible using a beamsplitter. A beamsplitter has two input ports and two output ports. When the beamsplitter is fed a coherent state as the first input, along with a classical optical field, represented simply a complex number α_{LO}, then its action can be represented in the Heisenberg picture in the following way. The Heisenberg picture means that formally, the output state is identical to the input state(our coherent state), so we have formally:

$$|\Psi_{\text{out}}\rangle = |\Psi_{\text{in}}\rangle$$

However, the annihilation operators (and thus creation operators, since they are adjoint of each others) of the resulting quantum optical fields are different. This means the number operators are different as well, so the photon statistics are different than those we would get without a beamsplitter. But we can still calculate their expressions. We use the subscripts 1 and 2 to identify which outputs of the beamsplitter the operators are acting on:

$$\hat{n}_1 = \hat{a}_1^\dagger \hat{a}_1 = (\hat{a}^\dagger + \alpha_{LO})(\hat{a} + \alpha_{LO})$$

and:

$$\hat{n}_2 = \hat{a}_2^\dagger \hat{a}_2 = (\hat{a}^\dagger - \alpha_{LO})(\hat{a} - \alpha_{LO})$$

The number-operators difference is obtained by subtracting the above operators, and by replacing α_{LO} by its polar expression(θ is the phase) and $\hat{a}$ by its expression as $\frac{1}{2}\hat{q} + i\hat{p}$:

$$\Delta\hat{n} = |\alpha_{\text{LO}}|\hat{q}\cos\theta + \hat{p}\sin\theta$$

So we can use the number operators $\hat{n}_1$ and $\hat{n}_2$ as measurement devices, and substract the outcomes, the expectation of this difference would be the same as if we used $\hat{p}$ (case $\theta = \pi/2$) or $\hat{q}$ (case $\theta = 0$) as measurement devices, Indeed, if X_1 ans X_2 are the r.v corresponding to these measurement then

$$\begin{aligned} \mathbf{E}(X_1 - X_2) &= \mathbf{E}(X_1) - \mathbf{E}(X_2) \\ &= \langle\psi|\hat{n}_1|\psi\rangle - \langle\psi|\hat{n}_2|\psi\rangle \\ &= \langle\psi|\hat{n}_1 - \hat{n}_2|\psi\rangle \\ &= |\alpha_{\text{LO}}|\langle q\rangle\cos\theta + \langle p\rangle\sin\theta \end{aligned} \tag{2}$$

So it is a linear combination of the expectations of the quadrature operators, and a scaling factor.

2.4.5 Multiples modes

To fully exploit quantum mechanical effects such as entanglement, one needs to consider composite states as well, which in quantum optics are usually called multi modes states. Studying these states through density matrices is not always the easiest approach, as these matrices are effectively infinite even for a single mode, however another representation is possible, through the Wigner function of a state. The Wigner function of an N mode state is a real function of $2N$ variables that contains all the information of the density matrix, hence it can be seen as the fundamental representation of a state without reference to the density matrix. A detailed review can be found inLeonhardt et al. [1997] For the purpose CVQKD, we are only interested in a family of states known as **Gaussian states**. A Gaussian state is defined as a state whose Wigner function is Gaussian(i.e. it is the probability density function of a Gaussian variable). This has many important implications.

First of all, Gaussian states are entirely determined by their mean vector, and their covariance matrix. The mean vector of a state is just the vector of expectations of each quadrature operator, there are two quadrature operators hence its size is twice the number of modes: $2N$. The covariance of any two operators A and B is generally defined as:

$$Cov(A, B) = \frac{1}{2}([A, B] + [B, A]) - \langle A\rangle\langle B\rangle$$

This formula is analog to the formula defining the covariance for classical random variables, but since operator composition is not commutative in general, there is no preferred order hence the average over the two possible choices. The covariance of operators can be hard to interpret for non commuting operators, however and we will make use of this property, the covariance of two commuting operators on ρ is simply the covariance of the classical random variables one gets by applying A on ρ then B on the post-measurement state. Using the *anti*commutator, the covariance of can be expressed in a more compact way:

$$Cov(A, B) = \frac{1}{2}\{A, B\} - \langle A\rangle\langle B\rangle$$

We are interested in the $2N$ quadrature operators of N mode states, their commutation relations can be nicely summarized using the symplectic form, where we use subscripts to indicate on which mode an operator is acting:

$$\Omega = \bigoplus_{1}^{N} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

This is because the quadrature operators of two different modes are commuting, so for $i \neq j$, we have $[p_i, q_j] = 0$. On the other hand, the quadrature operators within the same mode are identical to the single mode case so for all k, we have $[p_k, q_k] = 2i$.

Gaussian states are easy to manipulate from a theoretical perspective, for example taking the partial trace is just keeping the relevant submatrix, that is we have an $M + N$ modes state whose covariance matrix is $\begin{pmatrix} A & C \\ C^T & B \end{pmatrix}$ The resulting covariance matrix is simply A

(and the new mean vector is just the old mean vector, but truncated to only keep the values corresponding to the modes not being traced out) The quadrature operators for any mode, when evaluated on a Gaussian state, have a Gaussian distribution. This follows from the properties of the Wigner function as the probability density of a quadrature is obtained by integrating over the rest of modes.

3 Protocol with Gaussian modulation using coherent states (GG02)

3.1 Prepare and measure scenario

Any QKD protocol can be realized in two different ways, namely the Prepare-and-measure (PM) procedure and the Entanglement Based (EB). The Prepare-and-measure (PM) scenario is experimentally more realistic, but the security analysis is difficult. On the contrary, although practical implementation of the Entanglement based protocol is difficult its security analysis is simpler.

In the following we describe the PM scenario, which consists of Alice preparing coherent states, then sending them through a quantum channel to Bob, who will then measure them using Homodyne or Heterodyne detection.

The protocol starts with Alice generating a sequence of pairs of real numbers, or equivalently a sequence of complex numbers, according to a Gaussian distribution with mean 0 and fixed variance $\tilde{V}_{mod}$. Denote by (Q_i) and (P_i) her two sequences of (real) random variables. All these real variables are pairwise independent. The list of outcomes $(q_i, p_i), i \leq N$ makes her initial key i.e before the classical processing for error correction and privacy amplification.

To encode a given outcome (q_i, p_i), Alice will modulate a coherent state whose quadrature components are q_i and p_i. That is, her ith signal pulse is the coherent state $|q_i + p_i\rangle$. Note that unlike the discrete-variable QKD protocol BB84, the encoding of her classical key into a quantum state is deterministic : it does not involve choosing a basis at random. She then sends this coherent state to Bob through a quantum channel. Bob, upon receiving the coherent states, proceeds to estimate either q_i or p_i, using the quadrature operators which are observable (although they are unphyiscals, the exact mechanism is homodyne detection which gives the same statistics). This is possible because these operators when applied to a coherent state $|q + p\rangle$ yield a Gaussian distribution whose mean is $2q$ (for $\hat{q}$) and $2p$ (for $\hat{p}$) and variance is 1. So Bob recovers the chosen quadrature component, affected with variance 1 Gaussian noise, which is called shot-noise unit. This is called Homodyne detection, but we also consider the case that Bob performs Heterodyne detection, that is measuring a weighted sum of the quadratures. The implications of such heterodyne measurement on the SNR and hence the mutual information are examined in the next section.

We can now derive some basic but important probabilities relations that are helpful to later show that the PM and EB scenario are equivalent:

First of all, we calculate the variance of Bob's outcome. Suppose he performs a homodyne measurement with $\widehat{q}$, the $\widehat{p}$ case in analogous. Our argument is based on the fact that Bob's outcome is essentially the sum of $2q_i$ and a zero mean unit variance Gaussian variable. Therefore the variance at Bob (for the ensemble) is just $4\tilde{V}_{mod} + 1$

In Laudenbach et al. [2018] the approach for calculating Bob´s variance is a little different. We see that for a coherent state $|q + ip\rangle$, the conditional expectation of $\widehat{q}^2$ is $4q^2 + 1$. So since Alice generates q as a zero mean unit variance Gaussian, the variance at Bob of $\widehat{q}$ that is the expectation of $\widehat{q}^2$ is $4\tilde{V}_{mod} + 1$

One might also be interested in the photon statistics (number operators). We can calculate the mean number of photons given a a determined coherent state $|\alpha\rangle$ and use the law of total expectation to get the mean number of photons of the ensemble.

$$\langle n_i \rangle = \langle \alpha | n_i | \alpha \rangle = |\alpha|^2 = q^2 + p^2$$

So by total expectation the mean number of the photons is $\langle Q^2 \rangle + \langle P^2 \rangle = \tilde{V}_{mod}$

After Bob's measurement is done, the next step is information reconciliation. In CVQKD, it is usual to perform *reverse* reconciliation, instead of direct reconciliation. This means that Bob‘s key is the original one, and Alice corrects her through one-way communication over an authenticated public channel, as usual in QKD. Finally they perform privacy amplification to obtain a smaller shared key but which is secret from Eve, that is the mutual information between Eve´s observations and the final key is almost zero).

3.2 Entanglement based scenario

In the EB scenario, we need to find an appropriate three modes state such that when Alice does a partial measurement on the first two modes, sends the second one to Bob who then does another measurement, will give us the same statistics (probability distribution) as in the PM scenario. The joint probability distributions of the three real random variables must be the same in both scenarios.

The right state to use is a two-mode squeezed vacuum state (TMSVS) with variance V. For our purpose we are only interested in the covariance matrix of the TMSVS, which since it is a Gaussian state, characterizes it (together with its mean vector which is null). A detailled review of Gaussian states can be found in Cariolaro [2015]

$$\begin{pmatrix} V\mathbb{1}_2 & \sqrt{V^2-1}\sigma_z \\ \sqrt{V^2-1}\sigma_z & V\mathbb{1}_2 \end{pmatrix}$$

In reality Alice wants to get both quadratures components, in full conformance with the PM scenario. Since this isn't naively possible as they are noncommuting operators, she will split her mode in half using a beamsplitter and will measure each quadrature on different modes(q on the first mode and p on the second mode). Since the beamsplitter rescales the eigenvalue of its coherent state input, this will affect the covariance matrix so we are left with the covariance matrix of a 3 mode state :

$$\begin{pmatrix} \frac{V+1}{2}\mathbf{1}_2 & \frac{V-1}{2}\mathbf{1}_2 & \sqrt{\frac{1}{2}(V^2-1)}\sigma_z \\ \frac{V-1}{2}\mathbf{1}_2 & \frac{V+1}{2}\mathbf{1}_2 & \sqrt{\frac{1}{2}(V^2-1)}\sigma_z \\ \sqrt{\frac{1}{2}(V^2-1)}\sigma_z & \sqrt{\frac{1}{2}(V^2-1)}\sigma_z & V1_2 \end{pmatrix}$$

The formal proof of equivalence of such protocol with the PM one is the object of the next section.

3.3 Security Analysis

Security analysis has been carried out in the asymptotic sense, that is, assuming that the block length goes to infinity. Although not realistic, it is still useful, at least in providing upper bounds to the maximum achievable security.

3.3.1 Key Rate

The (per symbol) key rate is defined as the average amount of secret bits Alice and Bob can extract from a single symbol. Here the symbol for Alice is a complex number, modulated into a coherent state on which Bob perform homodyne or heterodyne detection on to estimate one or both quadrature components. The performance of the protocol will also depend on the symbol frequency. Denote by f_{sym} the number of symbols sent per second, the performance of the protocol that is the average amount of secret bits generated per second is thus related to the rate r by

$$K = f_{sym}.r$$

The theoretical rate of the discussed protocol is given by the Devetak-Winter formula Devetak and Winter [2005]:

$$r = I_{AB} - \chi_{EB}$$

This formula is not too much surprising, as it is the difference of the number of shared bits between Alice and Bob, and the bits Eve has been able to find. Note that the Holevo information is calculated between Eve and Bob because of reverse reconciliation (it would be between Alice and Eve in direct reconciliation). It somewhat parallels the classic formula for the secret-key agreement:

$$r = I_{AB} - I_{AE}$$

This rate however is theoretical and will be lower in practice, due to various reasons.

First, the reconciliation efficiency is not perfect and depends on the chosen algorithm, so I_{AB} is replaced by βI_{AB} where $\beta \leq 1$ is the **reconciliation efficiency**.

Second, there is a proportion of frames that are either discarded because of errors, or disclosed so they do not participate in the reconciliation. Say for example half of the frames are not used, then the previous rate should be divided by 2 in the previous formula. In the general if the proportion is e then the rate should be multiplied by $(1-e)$

This gives us the realistic rate achieved by such a protocol:

$$r = (1 - FER)(1 - \nu)(\beta I_{AB} - \chi_{EB})$$

3.3.2 Calculation of the mutual information

The mutual information is not difficult in the PM scenario (unlike the Holevo information, because). Since we have a signal affected by a Gaussian noise, we can use the usual Shannon formula: $I_{AB} = \frac{\mu}{2}\log(1 + SNR)$ Here μ refers to the bandwith and is 1 for homodyne detection and 2 for heterodyne detection.

The variance of Bob´s outcome as shown before is V assuming no losses and noise.

3.3.3 Formal proof of equivalence of the PM and EB scenarios

We want to prove that the PM and EB scenario are equivalent, so we can use the Holevo information calculated in EB case in our analysis of attacks even if the actual protocol carried out is PM.

By equivalents, we mean that the random variables arising in both cases have the same joint distribution. Since they are Gaussian vectors, it suffices to show that they have the same covariance matrix, and mean vector.

Denote by (A_1, A_2, B) the random vector generated in the PM scenario. As explained before, A_1 and A_2 are generated classically by Alice, and B is the result of Bob's homodyne/hetrodyne measurement. Similarly denote by (A_1', A_2', B') the random vector generated in the EB scenario. Its components are generated by Alice's and Bob's measurements on a three mode Gaussian state. Specifically, Alice measures the first two modes and Bob the last one, the order of measurement is in principle arbitrary since they are commuting operators.

The apparent difficulty lays in the fact that the co-variances appearing in the EB covariance matrix aren't a priori the covariances of random variables, but are formally defined by $\mathsf{Cov}(M, N) = \frac{1}{2}\{M, N\}$: What we show is that the covariances *of the r.v.s defined by Alice and Bob successive measurements* in the EB scenario are identical to those operator covariances. Then if the EB matrix is equal to the PM matrix, this means that effectively, the r.v.s arising in the PM and EB scenarios are the same, since two random Gaussian vectors with the same mean and covariance must have the same distribution.

Let us first explicit the covariance matrix in the PM case. This is a matrix of covariances in the classical sense. The variances were calculated before (which are $V_{mod}, V_{mod}, V_{mod}+1$), it remains to calculate the covariances which are non null between, which can be done using the fact that Bob's outcome is the sum of Alice's outcome and a $\mathcal{N}(0, 1)$:

$$\mathsf{Cov}(\hat{q}_A, \hat{q}_B) = \langle \hat{q}_A \hat{q}_B \rangle = \langle (\hat{q}_B - \hat{q}_A)\hat{q}_A \rangle + \langle \hat{q}_A^2 \rangle$$

Then using the fact that $(\hat{q}_B - \hat{q}_A)$ is a $\mathcal{N}(0, 1)$ independant from $\hat{q}_A$ all we are left with is $\langle \hat{q}_A^2 \rangle = V_{mod}$.

Now, we show that the covariances of (A_1', A_2', B') are those appearing the EB matrix.

It is clear that A_1' is a Gaussian variable whose mean is 0 and variance V, using the fact that its distribution can be recovered by integrating the Wigner function of the TMVS. Indeed the distribution of any quadrature of any given mode can be recovered by integrating the Wigner function over the rest of the $2N-1$ variables, and by definition of Gaussian states their Wigner function is Gaussian. For A_2' and B', this is also true even after Alice performs her measurement, since the quadratures of different modes are commuting (this is a basic property of quantum measurements, $M \otimes I$and $I \otimes N$ commute for any two operators M and N and the result extends to any number of modes) therefore even if Bob performs homodyne detection on the post-measurement state, the overall distribution of the second quadrature stays the same as if Alice didn't do anything.

To make more precise the above statement that the mean and variance of the EB random vector is the same as the PM one, we used the following result: Let M and N be two commuting observables and ρ a state. The distribution of N applied to ρ is the same as the distribution of the outcome one gets by applying M first, discarding the result and applying N on the post-measurement state. This result stays true for 3 observables as in our case.

Now we show that $Cov(A_1', B') = Cov(A_1, B)$. We know that $Cov(A', B') = \frac{1}{2}\{q_A, q_B\}$. We can use the following result: Let M, N be two commuting operators with zero mean. Then if we measure using M on a state ρ, then measure using N on the post-measurement state, the (classical) covariance of the outcomes is exactly the covariance of the operators that is $\frac{1}{2}\{q_A, q_B\}$. This is possible because two commuting operators can be considered a single operator(”we can measure both at the same time”) : We can find a common orthonormal basis $|c, d\rangle$ where $M|c, d\rangle = c|c, d\rangle$ and $N|c, d\rangle = d|c, d\rangle$. So this shows that the covariance matrix of the operators is equal to the covariance matrix of the corresponding measurements

Our EB matrix is identical to the PM one, except for a scale factor, so we can just rescale measurements:

$$\begin{pmatrix} \frac{V+1}{2}\mathbf{1}_2 & \frac{V-1}{2}\mathbf{1}_2 & \sqrt{\frac{1}{2}(V^2-1)}\sigma_z \\ \frac{V-1}{2}\mathbf{1}_2 & \frac{V+1}{2}\mathbf{1}_2 & \sqrt{\frac{1}{2}(V^2-1)}\sigma_z \\ \sqrt{\frac{1}{2}(V^2-1)}\sigma_z & \sqrt{\frac{1}{2}(V^2-1)}\sigma_z & V1_2 \end{pmatrix}$$

Since covariance is bilinear ($Cov(aX + Y, Z) = aCov(X, Z) + Cov(Y, Z)$, Alice only needs to rescale her measurements so that both matrices become identical.

$$\hat{q}_A \rightarrow \sqrt{\frac{2(V-1)}{V+1}}\hat{q}_A \tag{3}$$

$$\hat{q}_B \rightarrow \sqrt{\frac{2(V-1)}{V+1}}\hat{p}_A \tag{4}$$

3.4 Possible attacks

3.4.1 Classification of attacks

Attacks on a QKD are usually classified in three tiers depending on the attacker's capabilities:

- Individual attack: Eve's has access to independent ancillae each of which interact with a single optical pulse.
- Collective attack: Identical to the above, but Eve can measure her quantum state anytime, even after post-processing.
- Coherent attack: The i.i.d assumption is dropped. Eve may have a multi mode ancilla, possibly not separable, and it can interact with the joint pulses of the signal.

In general, analysis of attacks is made simpler by assuming that Eve holds a purification of Alice and Bob mutual state ρ_{AB}. The exact purification does not matter, as all purifications are related by a unitary transformation ($U\rho U^\dagger$) so the Holevo information, which is invariant by unitary transformation, is the same, and therefore the knowledge of Eve doesn't depend on her purification.

3.4.2 Entangling cloner attack

This attack is based on the fact that Alice and Bob are willing to tolerate a certain amount of noise (because their channel is assumed noisy), Eve can therefore replace the lossy channel between Alice and Bob by a noiseless channel, so Eve will get some ”room” to eavesdrop even if her intervention introduces noise, as long as the loss doesn't attain the threshold of the assumed lossy channel. More precisely, the intervention consist of a beamsplitter that will mix one of Eve's modes with Bob's mode. It is an example of a collective attack, since she can measure anytime after postprocessing and by computing the Holevo information she will measures state by state.

This attack uses the following result: Let ρ be a Gaussian state with covariance matrix Σ. A beamsplitter with transmission T applied to two modes of a Gaussian state will transform them according to the following matrix:

$$\begin{pmatrix} \sqrt{T}\mathbb{1}_2 & \sqrt{1-T}\mathbb{1}_2 \\ \sqrt{1-T}\mathbb{1}_2 & \sqrt{T}\mathbb{1}_2 \end{pmatrix}$$

For a general N mode Gaussian state, we start with the identity matrix $\mathbb{1}_{2N}$and plug the above 4×4 matrix at the correct cells depending on the modes the beamsplitter is acting on so that the beamsplitter acting on modes 2 and 3 (i.e. Bob's mode and the first of Eve's modes) is:

$$\begin{pmatrix} \mathbb{1}_2 & 0 & 0 & 0 \\ 0 & \sqrt{T}\mathbb{1}_2 & \sqrt{1-T}\mathbb{1}_2 & 0 \\ 0 & \sqrt{1-T}\mathbb{1}_2 & \sqrt{T}\mathbb{1}_2 & 0 \\ 0 & 0 & 0 & \mathbb{1}_2 \end{pmatrix}$$

So when applied to the second and third mode of the 4-modes Gaussian state ρ, the covariance matrix will transform the total state to become:

$$\Sigma_{ABE_1E_2} = BS_{BE_1}\Sigma_{ABE_1E_2}BS^T_{BE_1}$$

$$= \begin{pmatrix} V\mathbb{1}_2 & \sqrt{T}\sqrt{V^2-1}\sigma_z & -\sqrt{1-T}\sqrt{V^2-1}\sigma_z & 0 \\ \sqrt{T}\sqrt{V^2-1}\sigma_z & (TV+[1-T]W)\mathbb{1}_2 & \sqrt{T(1-T)}(W-V)\mathbb{1}_2 & \sqrt{1-T}\sqrt{W^2-1} \\ -\sqrt{1-T}\sqrt{V^2-1}\sigma_z & \sqrt{T(1-T)}(W-V)\mathbb{1}_2 & ([1-T]V+TW)\mathbb{1}_2 & \sqrt{T}\sqrt{W^2-1}\sigma_z \\ 0 & \sqrt{1-T}\sqrt{W^2-1} & \sqrt{T}\sqrt{W^2-1}\sigma_z & \end{pmatrix}$$

So the covariance matrix between Alice and Bob (first two rows and first two columns) is:

$$\begin{pmatrix} V\mathbb{1}_2 & \sqrt{T}\sqrt{V^2-1}\sigma_z \\ \sqrt{T}\sqrt{V^2-1}\sigma_z & (TV+[1-T]W)\mathbb{1}_2 \end{pmatrix}$$

So if Eve chooses her variance to be related to the transmission as:

$$W = \frac{\xi}{1-T} + 1$$

Then the covariance between Alice and Bob is:

$$\begin{pmatrix} V\mathbb{1}_2 & \sqrt{T}\sqrt{V^2-1}\sigma_z \\ \sqrt{T}\sqrt{V^2-1}\sigma_z & (T[V-1]+1+\xi)\mathbb{1}_2 \end{pmatrix}$$

which corresponds to communication through a lossy channel with excess noise ξ without an eavesdropper.

Even if Eve's intervention is unnoticed, the attack is only useful if she can extract information, which prompts the calculation of her Holevo information: The von Neumann entropy depends on the symplectic eigenvalues of Eve's substate which is:

$$\begin{pmatrix} V\mathbb{1}_2 & \sqrt{T}\sqrt{V^2-1}\sigma_z \\ \sqrt{T}\sqrt{V^2-1}\sigma_z & (T[V-1]+1+\xi)\mathbb{1}_2 \end{pmatrix}$$

Denoting by a, b and c the coefficient of the identity matrix and Pauli Z matrix in the above matrix, direct calculation of the symplectic eigenvalues immediately leads to $\frac{1}{2}(z \pm [b-a])$ where $z = \sqrt{(a^2+b^2-4c^2}$. These eigenvalues can be plugged in the formula giving the von Neumann entropy as function of the symplectic eigenvalue of the covariance matrix (see Laudenbach et al. [2018] B.14)

$$H = \sum g(v_i)$$

with

$$g(v) = (\frac{v+1}{2})log_2(\frac{v+1}{2}) - (\frac{v-1}{2})log_2(\frac{v-1}{2})$$

and v_i being the symplectic eigenvalue of the covariance matrix of the state.

In order to calculate $H(E|B)_\rho$, we need to know the post-measurement state, which is a Gaussian state. **The covariance matrix of the post-measurement state doesn't depend on the measurement outcome**, it is similar to the formula for the conditional distribution of Gaussian vector:

$$\sigma_{E|B} = \Sigma_E - \frac{1}{V_B}(([1-T]V+TW)\mathbb{1}_2)$$

3.4.3 Purification

In Laudenbach et al. [2018], the question of an universal analysis of purification attacks was considered. That is, all we assume is that Eve holds an ancilla that purifies Alice and Bob's joint state ρ_{AB}.

This means that the total state is an element of the triple tensor product of Alice, Bob and Eve's Hilbert spaces, not being necessarily separable, i.e not necessarily being the product of some element of the tensor product of Alice and Bob's Hilbert spaces, with Eve's Hilbert space. However the following holds:

- The total state is pure: $\psi = |\psi\rangle\langle\psi|$
- The partial trace when tracing out Eve's space is ρ_{AB}. So the statistics of the state shared by Alice and Bob when they perform measurements, if Eve doesn't do anything on her ancilla, are guaranteed to be identical to those obtained in a scenario without any eavesdropper where their state would literally be ρ_{AB}.

In this case, the Schmidt decomposition of a pure bipartite state turns out to be helpful. More precisely we see the total state as bipartite between the tensor product of Alice and Bob Hilbert space, that is the first space, and Eve's Hilbert space, that is the second space.

$$|\psi\rangle = \sum \sqrt{\lambda_i}|i\rangle_{AB}|i\rangle_E$$

The partial trace, once such decomposition is obtained, is trivial as we just eliminate Eve's kets and square the components. This comes from the fact that we need to transform the above expression to a density matrix since the partial trace is defined for operators not state vectors hence the induced squaring of the Schmidt components.

One might ask what would happen if we used another Schmidt decomposition, since the λ_i will be different and hence the Von Neumann entropy too, a priori. In fact, all purifications are related through a Unitary transform, and the Holevo information is invariant under such transforms. Therefore the Holevo information $H(E)$ is just $H(A, B)$

Finally we need to calculate the second part of the Holevo information which is the Von Neuman entropy after Bob performs a measurement. We consider only homodyne detection for simplicity. This can be obtained using the general formula giving the covariance matrix of the remaining mode(s) of a Gaussian state, after the last one has been measured. Note that this post-measurement covariance matrix doesn't depend on the outcome unlike the mean vector which actually does), a property of Gaussian states. Denoting by a, b and c the coefficient of the identity matrix and Pauli Z matrix in the covariance matrix of the shared state between Alice and Bob, the symplectic eigenvalue can be expressed as:

$$\sqrt{a(a - \frac{c^2}{b})}$$

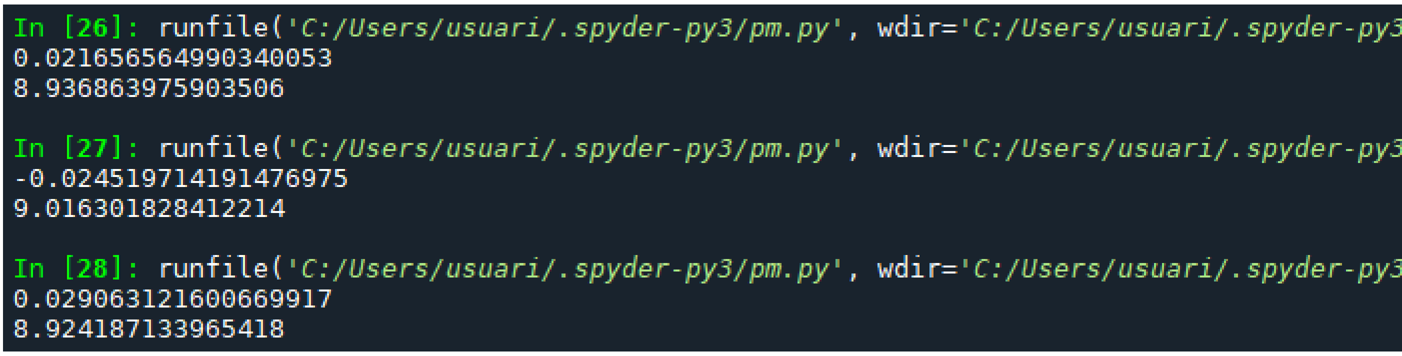
3.5 Basic numerical verification

Using StrawberryFields[Killoran et al., 2019] and [Bromley et al., 2020], we can experimentally assess some of the previous formulas:

3.5.1 Variance and Mean in the PM scenario

We simply generate samples from a Gaussian distribution with mean 0 and variance $\tilde{V}_{mod}$, which represent Alice's q quadrature outcomes. Then each outcome gives rise to a coherent state (the other quadrature of this state is a random unused p).

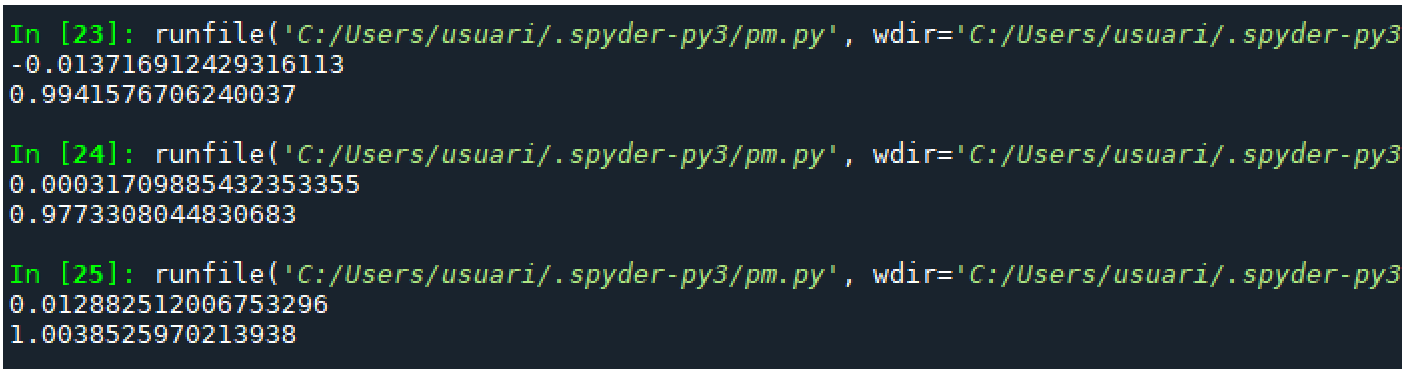
Bob then measures, always using $\hat{q}$ (In the actual protocol, we should choose one of the two quadratures at random). The mean should be 0 and the variance should be $4\tilde{V}_{mod}+1$ which is the case here after 10000 trials where we have set $\tilde{V}_{mod} = 2$:



```
In [26]: runfile('C:/Users/usuari/.spyder-py3/pm.py', wdir='C:/Users/usuari/.spyder-py3
0.021656564990340053
8.936863975903506

In [27]: runfile('C:/Users/usuari/.spyder-py3/pm.py', wdir='C:/Users/usuari/.spyder-py3
-0.024519714191476975
9.016301828412214

In [28]: runfile('C:/Users/usuari/.spyder-py3/pm.py', wdir='C:/Users/usuari/.spyder-py3
0.029063121600669917
8.924187133965418
```

3.5.2 Error on a fixed coherent state

According to the theory, Bob's mean outcome is zero (it is a centered random variable). We check this. We also check that the error (difference between Alice's quadrature component and Bob's estimation of it), is a $\mathcal{N}(0,1)$ Since the expectation of $\hat{p}$ is actually $2p$, we subtract this value from Bob's outcome, i.e we calculate $B-2q$ where B is the r.v representing Bob's outcome.



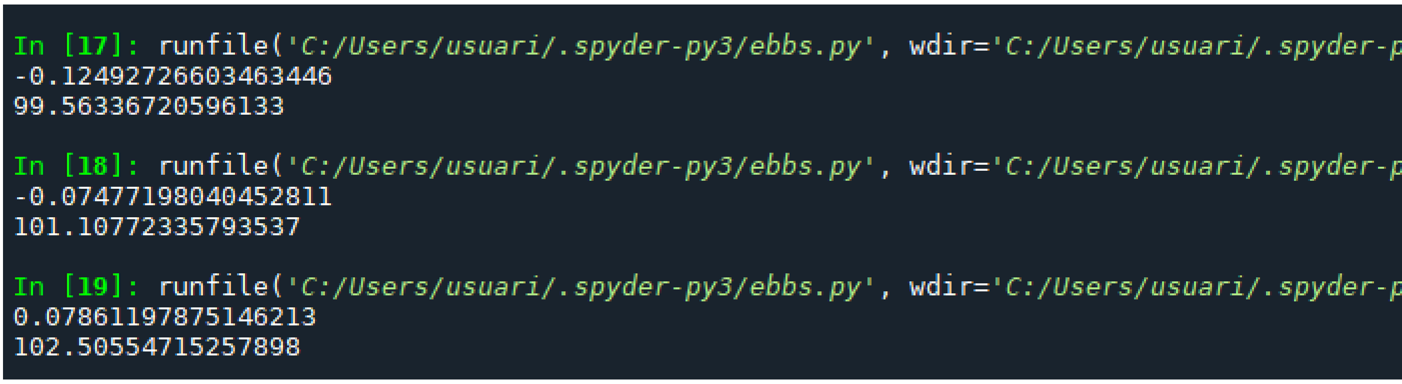
```
In [23]: runfile('C:/Users/usuari/.spyder-py3/pm.py', wdir='C:/Users/usuari/.spyder-py3
-0.013716912429316113
0.9941576706240037

In [24]: runfile('C:/Users/usuari/.spyder-py3/pm.py', wdir='C:/Users/usuari/.spyder-py3
0.00031709885432353355
0.9773308044830683

In [25]: runfile('C:/Users/usuari/.spyder-py3/pm.py', wdir='C:/Users/usuari/.spyder-py3
0.012882512006753296
1.0038525970213938
```

3.5.3 Variances in the EB scenario

We have seen in 3.2 that the EB scenario is based on the preparation of a 3 mode Gaussian state by Alice, who measures the first two modes to get her two classical outcomes, and sends the last mode to Bob who measures it. The outcomes will have the same joint distribution as in the PM scenario, if Alice rescales her measurements.

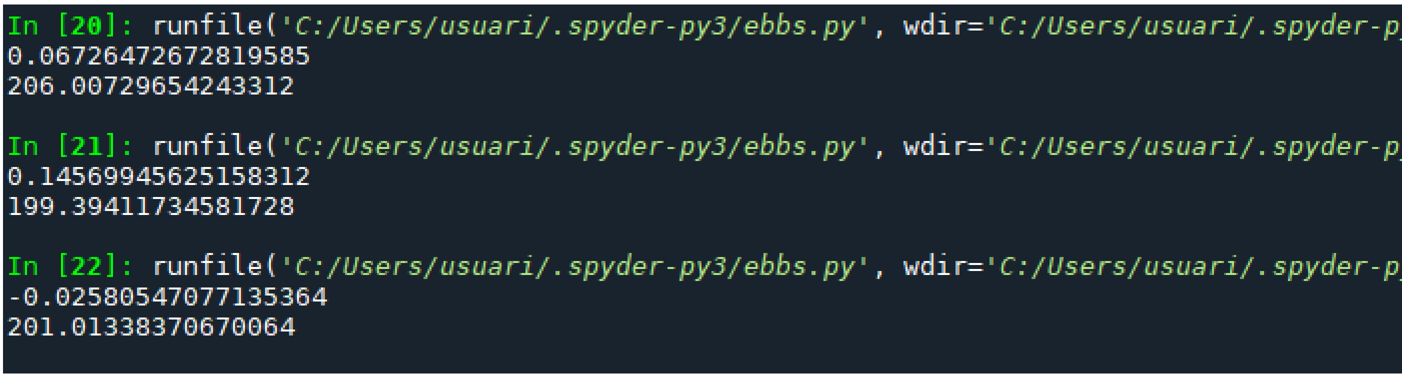
We prepare the required 3 mode Gaussian state in the following way: We start with three vacuuas. We then apply a two modes squeezing gate (S2 gate) with parameter r to modes

1 and 3. The result of this gate is a TMSVS of variance $V = \cosh(2r)$ in modes 1 and 3. Then a beamsplitter is inserted between modes 1 and 2 (which is still a vacuum). According to the theory, the variance of the first mode should be $\frac{V+1}{2}$ Here $r = 3$, so $V = \cosh(6) \approx 202$, so the variance should be 101



```
In [17]: runfile('C:/Users/usuari/.spyder-py3/ebbs.py', wdir='C:/Users/usuari/.spyder-p
-0.12492726603463446
99.56336720596133

In [18]: runfile('C:/Users/usuari/.spyder-py3/ebbs.py', wdir='C:/Users/usuari/.spyder-p
-0.07477198040452811
101.10772335793537

In [19]: runfile('C:/Users/usuari/.spyder-py3/ebbs.py', wdir='C:/Users/usuari/.spyder-p
0.07861197875146213
102.50554715257898
```

The variance of the 3rd mode (Bob's mode) which according to 3.2 should be V is also checked:

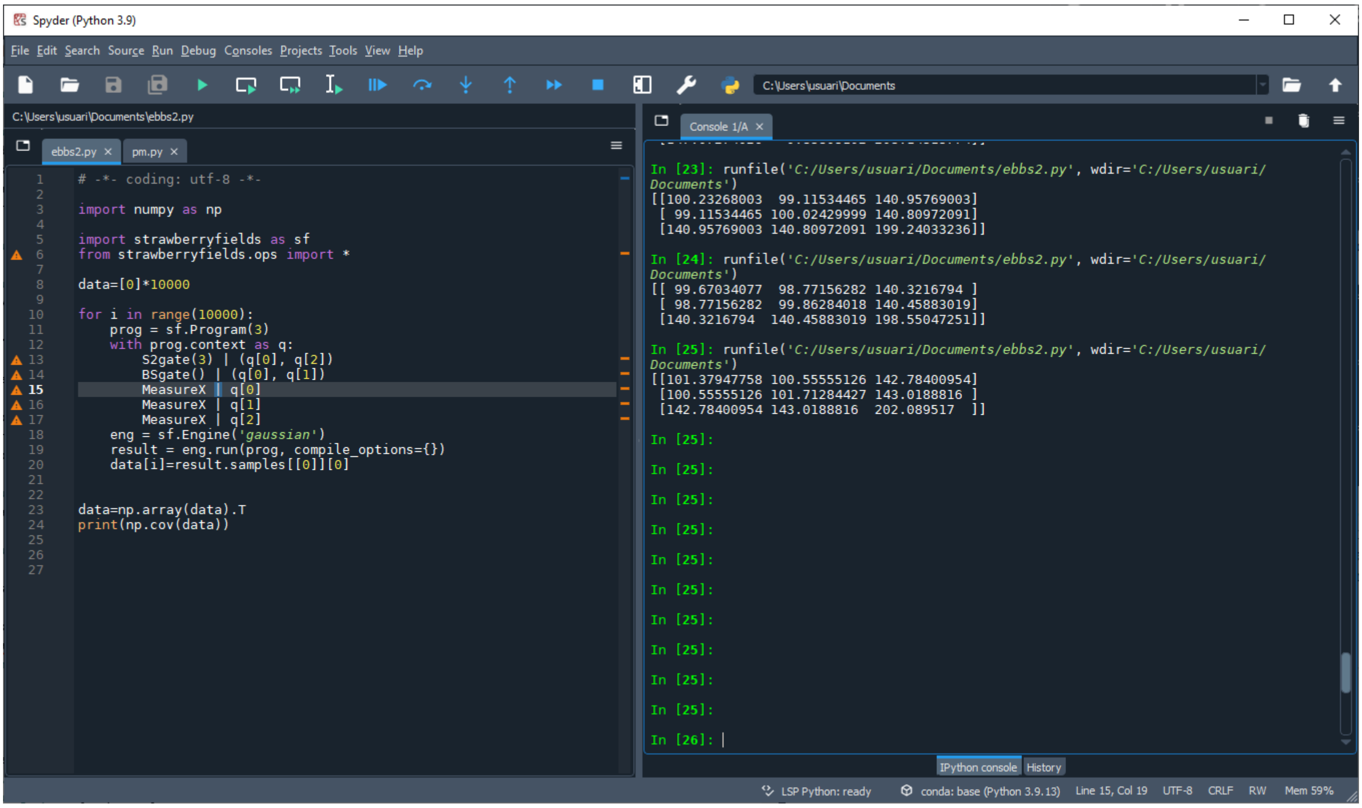


```
In [20]: runfile('C:/Users/usuari/.spyder-py3/ebbs.py', wdir='C:/Users/usuari/.spyder-p
0.06726472672819585
206.00729654243312

In [21]: runfile('C:/Users/usuari/.spyder-py3/ebbs.py', wdir='C:/Users/usuari/.spyder-p
0.14569945625158312
199.39411734581728

In [22]: runfile('C:/Users/usuari/.spyder-py3/ebbs.py', wdir='C:/Users/usuari/.spyder-p
-0.02580547077135364
201.01338370670064
```

In a similar fashion, we could also check the whole covariance matrix 3.3.3 (interpreted as a covariance matrix of random variables) by calculating the empiricial covariance. It is a 6×6 matrix(two quadratures for each of the three modes), but since we cannot measure both quadratures, we consider the following 3×3 submatrices:

- Covariances of the $\hat{q}$ quadrature on all 3 modes.
- Covariances of $\hat{q}$, $\hat{p}$ and $\hat{q}$ (which is what is done in the EB protocol)

The result are shown below. Note that the only line of code that changes is the 16th. The results are consistant with 3.3.3 (when taking $V = \cosh(6) \approx 202$)

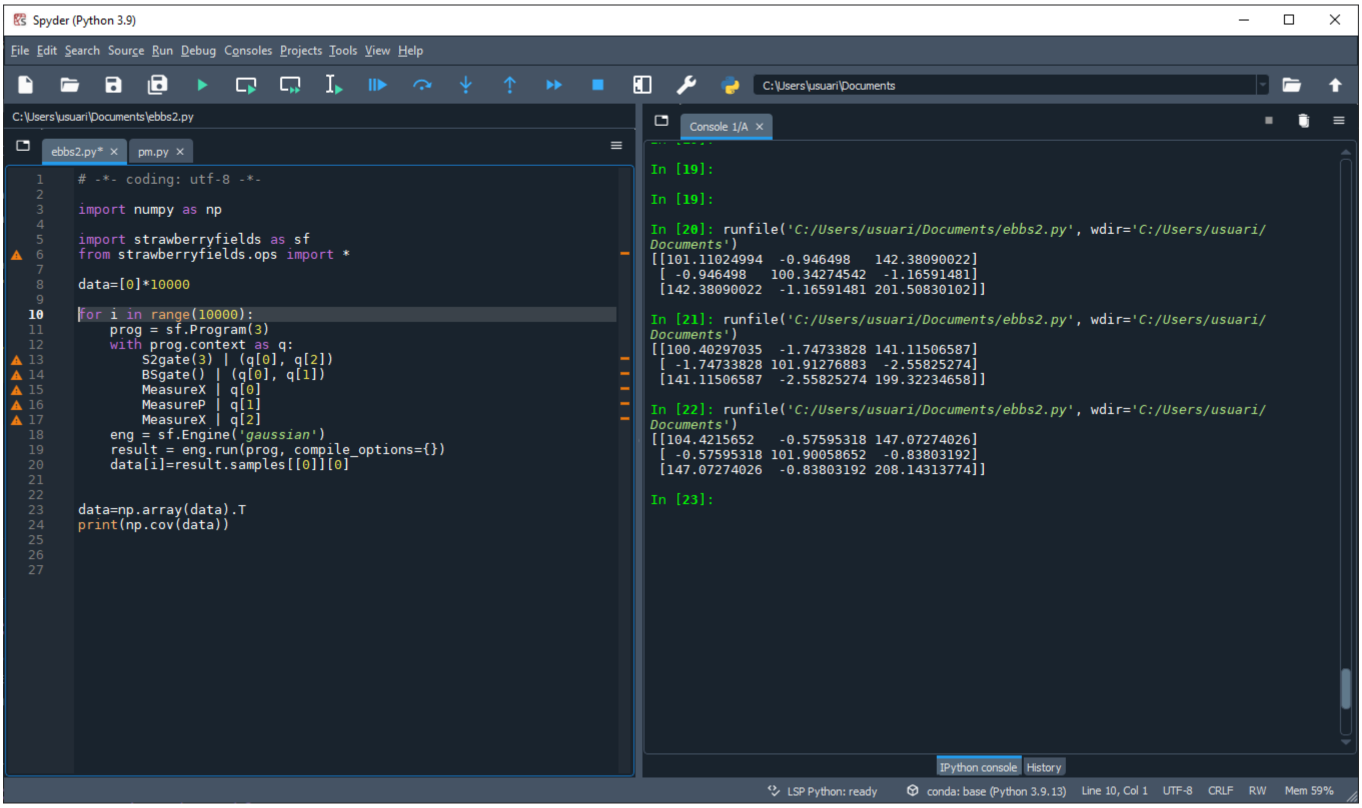


```
# -*- coding: utf-8 -*-

import numpy as np

import strawberryfields as sf
from strawberryfields.ops import *

data=[0]*10000

for i in range(10000):
    prog = sf.Program(3)
    with prog.context as q:
        S2gate(3) | (q[0], q[2])
        BSgate() | (q[0], q[1])
        MeasureX | q[0]
        MeasureX | q[1]
        MeasureX | q[2]
    eng = sf.Engine('gaussian')
    result = eng.run(prog, compile_options={})
    data[i]=result.samples[0][0]


data=np.array(data).T
print(np.cov(data))
```

```
In [23]: runfile('C:/Users/usuari/Documents/ebbs2.py', wdir='C:/Users/usuari/
Documents')
[[100.23268003  99.11534465 140.95769003]
 [ 99.11534465 100.02429999 140.80972091]
 [140.95769003 140.80972091 199.24033236]]

In [24]: runfile('C:/Users/usuari/Documents/ebbs2.py', wdir='C:/Users/usuari/
Documents')
[[ 99.67034077  98.77156282 140.3216794 ]
 [ 98.77156282  99.86284018 140.45883019]
 [140.3216794  140.45883019 198.55047251]]

In [25]: runfile('C:/Users/usuari/Documents/ebbs2.py', wdir='C:/Users/usuari/
Documents')
[[101.37947758 100.55555126 142.78400954]
 [100.55555126 101.71284427 143.0188816 ]
 [142.78400954 143.0188816  202.089517  ]]

In [25]:

In [25]:

In [25]:

In [25]:

In [25]:

In [25]:

In [25]:

In [25]:

In [25]:

In [25]:

In [26]: 
```

Spyder (Python 3.9)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\usuari\Documents
C:\Users\usuari\Documents\ebbs2.py
ebbs2.py
pm.py
Console 1/A
IPython console
History
LSP Python: ready
conda: base (Python 3.9.13)
Line 15, Col 19
UTF-8
CRLF
RW
Mem 59%

Figure 3: Covariances of the $\hat{q}$ quadrature on all 3 modes.



```
# -*- coding: utf-8 -*-

import numpy as np

import strawberryfields as sf
from strawberryfields.ops import *

data=[0]*10000

for i in range(10000):
    prog = sf.Program(3)
    with prog.context as q:
        S2gate(3) | (q[0], q[2])
        BSgate() | (q[0], q[1])
        MeasureX | q[0]
        MeasureP | q[1]
        MeasureX | q[2]
    eng = sf.Engine('gaussian')
    result = eng.run(prog, compile_options={})
    data[i]=result.samples[0][0]


data=np.array(data).T
print(np.cov(data))
```

```
In [19]:

In [19]:

In [20]: runfile('C:/Users/usuari/Documents/ebbs2.py', wdir='C:/Users/usuari/
Documents')
[[101.11024994  -0.946498   142.38090022]
 [ -0.946498   100.34274542  -1.16591481]
 [142.38090022  -1.16591481 201.50830102]]

In [21]: runfile('C:/Users/usuari/Documents/ebbs2.py', wdir='C:/Users/usuari/
Documents')
[[100.40297035  -1.74733828 141.11506587]
 [ -1.74733828 101.91276883  -2.55825274]
 [141.11506587  -2.55825274 199.32234658]]

In [22]: runfile('C:/Users/usuari/Documents/ebbs2.py', wdir='C:/Users/usuari/
Documents')
[[104.4215652   -0.57595318 147.07274026]
 [ -0.57595318 101.90058652  -0.83803192]
 [147.07274026  -0.83803192 208.14313774]]

In [23]:
```

Spyder (Python 3.9)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\usuari\Documents
C:\Users\usuari\Documents\ebbs2.py
ebbs2.py*
pm.py
Console 1/A
IPython console
History
LSP Python: ready
conda: base (Python 3.9.13)
Line 10, Col 1
UTF-8
CRLF
RW
Mem 59%

Figure 4: Covariances of $\hat{q}_1$, $\hat{p}_2$ and $\hat{q}_3$

.

4 Conclusions

CV-QKD is without doubt a strong alternative to the conventional discrete QKD, despite its relatively young age. We have seen that using standard telecommunication equipment such as photodiodes, beamsplitters and fibers is sufficient to implement a CVQKD protocol, a major advantage when compared to discrete QKD.

We have reviewed the most known protocol based on coherent states and homodyne detection with Gaussian modulation, GG02, proven the equivalence of both scenarios (Prepare-and-Measure and Entanglement Based), numerically assessed some properties related to this equivalence, and reviewed two general attacks.

Some experimental challenges remain, such as increasing the key rate and the communication distance. Among the current focus of research, new paradigms for security proofs such as composable security seek to establish security proofs that are more general.

Although more effort is required to understand the security of CVQKD better, the existing proofs encompass a broad range of attack models. The Gaussian modulated protocol we reviewed has the property that its security against coherent attacks can only be proven asymptotically, whereas proofs of security against attacks are only possible if attacks are restricted to be collective, a weaker kind of attack. New directions focus on finite-size security. We mention the recent paper by Matsuura et al. [2021] that establish a provably secure finite-size CVQKD protocol. Another development in the continuous variable realm consist of cryptography primitives such as entanglement certification, where we mention the recent paper byAbiuso et al. [2021].

Bibliography

Paolo Abiuso, Stefan Bäuml, Daniel Cavalcanti, and AntonioAcí n. Measurement-device-independent entanglement detection for continuous-variable systems. *Physical Review Letters*, 126(19), may 2021. doi: 10.1103/physrevlett.126.190502. URL `https://doi.org/10.1103%2Fphysrevlett.126.190502`.

Matthieu Bloch and João Barros. *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011. doi: 10.1017/CBO9780511977985.

Thomas R Bromley, Juan Miguel Arrazola, Soran Jahangiri, Josh Izaac, Nicolá s Quesada, Alain Delgado Gran, Maria Schuld, Jeremy Swinarton, Zeid Zabaneh, and Nathan Killoran. Applications of near-term photonic quantum computers: software and algorithms. *Quantum Science and Technology*, 5(3):034010, may 2020. doi: 10.1088/2058-9565/ab8504. URL `https://doi.org/10.1088%2F2058-9565%2Fab8504`.

G. Cariolaro. *Quantum Communications*. Signals and Communication Technology. Springer International Publishing, 2015. ISBN 9783319156002. URL `https://books.google.co.ma/books?id=gyv3BwAAQBAJ`.

Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 461(2053):207–235, jan 2005. doi: 10.1098/rspa.2004.1372. URL `https://doi.org/10.1098%2Frspa.2004.1372`.

Nathan Killoran, Josh Izaac, Nicolás Quesada, Ville Bergholm, Matthew Amy, and Christian Weedbrook. Strawberry Fields: A Software Platform for Photonic Quantum Computing. *Quantum*, 3:129, March 2019. ISSN 2521-327X. doi: 10.22331/q-2019-03-11-129. URL `https://doi.org/10.22331/q-2019-03-11-129`.

Fabian Laudenbach, Christoph Pacher, Chi-Hang Fred Fung, Andreas Poppe, Momtchil Peev, Bernhard Schrenk, Michael Hentschel, Philip Walther, and Hannes Hübel. Continuous-variable quantum key distribution with gaussian modulation-the theory of practical implementations. *Advanced Quantum Technologies*, 1(1):1800011, jun 2018. doi: 10.1002/qute.201800011. URL `https://doi.org/10.1002%2Fqute.201800011`.

U. Leonhardt, P.L. Knight, and A. Miller. *Measuring the Quantum State of Light*. Cambridge Studies in Modern Optics. Cambridge University Press, 1997. ISBN 9780521497305. URL `https://books.google.co.ma/books?id=wmsJy1A_cyIC`.

Anthony Leverrier and Philippe Grangier. Simple proof that gaussian attacks are optimal among collective attacks against continuous-variable quantum key distribution with a gaussian modulation. *Physical Review A*, 81(6), jun 2010. doi: 10.1103/physreva.81.062314. URL `https://doi.org/10.1103%2Fphysreva.81.062314`.

Takaya Matsuura, Kento Maeda, Toshihiko Sasaki, and Masato Koashi. Finite-size security of continuous-variable quantum key distribution with digital signal processing. *Nature Communications*, 12(1), jan 2021. doi: 10.1038/s41467-020-19916-1. URL `https://doi.org/10.1038%2Fs41467-020-19916-1`.