1

Basic Ideas

1.1 Introduction

The state of a quantum mechanical system $S$ is described by a density operator $\hat{\rho}$ acting on a Hilbert space $\mathcal{H}$. The system in consideration specifies the dimension of the Hilbert space. The operator $\hat{\rho}$ satisfies the following three defining properties:

$$\hat{\rho} = \hat{\rho}^\dagger, \quad \text{Tr} \hat{\rho} = 1, \quad \hat{\rho} \geq 0.$$ (1.1)

A pure state is described by a normalised (unit) vector $|\Psi\rangle$ in the Hilbert space $\mathcal{H}$, and the density operator corresponding to $|\Psi\rangle$ is given by

$$\hat{\rho} = |\Psi\rangle\langle\Psi|.$$ (1.2)

Clearly, distinct vectors in the Hilbert space do not correspond to distinct states. All unit vectors in $\mathcal{H}$ which differ from one another by phase factors, represent one and the same state. In other words, states are represented by an equivalence class of unit vectors of the Hilbert space. It is clear that the state $\hat{\rho}$ in Eq. (1.2), satisfies the three defining requirements in Eq. (1.1).

The most general state of a quantum mechanical system $S$ is described by a ‘mixed’ state $\hat{\rho}$, which is a convex combination of pure states, i.e.,

$$\hat{\rho} = \sum_k p_k |\Psi_k\rangle\langle\Psi_k|, \quad p_k > 0, \quad \sum_k p_k = 1.$$ (1.3)

The quantum state space thus forms a convex set in which pure states correspond to the case when all but one of the $p_k$’s are zero. The pure states satisfy the additional
requirement
\[
\hat{\rho}^2 = \hat{\rho},
\] (1.4)

and correspond to the extremal points of the convex state space. They cannot be realised as nontrivial convex combinations of other states. While \(\text{Tr}\hat{\rho}^2 = 1\) for pure states, mixed states satisfy
\[
\text{Tr}\hat{\rho}^2 < 1,
\] (1.5)

and correspond to non extremal points of the convex state space.

The probabilities \(p_k\)'s and the ensemble realisation in Eq. (1.3) are in general associated with a preparation procedure. The nontriviality of the ensemble realisation arises from the fact that the \(|\Psi_k\rangle\)'s need not be orthogonal, or even linearly independent, and that the set of ensembles realising a given mixed state \(\hat{\rho}\) is a huge family [1]. A preparation procedure yields an ensemble realisation for \(\hat{\rho}\), but given a \(\hat{\rho}\) it is impossible to tell which preparation procedure it was derived from.

The natural objects of interest are the expectation values of observables. Observables in quantum theory are represented by hermitian operators \(\hat{\Omega}\),
\[
\hat{\Omega} = \sum_i \lambda_i |\Phi_i\rangle\langle \Phi_i| = \sum_i \lambda_i P_i, \quad \sum_i P_i = 1,
\] (1.6)

and \(P_i\)'s are projection operators obeying \(P_i P_j = \delta_{ij} P_i\). The \(\{\lambda_i\}\) are interpreted as the outcome or eigenvalue of an experiment corresponding to the observable \(\hat{\Omega}\), and \(\{|\Phi_i\rangle\}\) as the corresponding eigenstates. The eigenvalues \(\lambda_i\) are real, but can be negative. The expectation value of an observable \(\hat{\Omega}\) with respect to a pure state \(|\Psi\rangle\) is given by \(\langle \hat{\Omega} \rangle = \langle \Psi|\hat{\Omega}|\Psi\rangle\). In the case of a mixed state \(\hat{\rho}\), the expectation value is given by
\[
\langle \hat{\Omega} \rangle = \text{Tr}(\hat{\Omega}\hat{\rho}) = \sum_k p_k \langle \Psi_k|\hat{\Omega}|\Psi_k\rangle.
\] (1.7)

The expectation value \(\langle \hat{\Omega} \rangle\) is interpreted as the average value of the observable \(\hat{\Omega}\) over repeated trials of the experiment, with the same state \(\hat{\rho}\) prepared each time.

Though the average value of the outcome is calculated as in Eq. (1.7), a particular trial yields a particular eigenvalue \(\lambda_i\) as the outcome of the experiment. The probability of occurrence \(q_i\) of the \(i\)th outcome corresponding to the eigenvalue \(\lambda_i\) is given by
\[
q_i = \text{Tr}(P_i\hat{\rho}) = \sum_k p_k \langle \Psi_k|P_i|\Psi_k\rangle.
\] (1.8)
Given a particular outcome \( i \), the state of the system after measurement is no more represented by \( \hat{\rho} \), but collapses to the corresponding eigenstate of the observed eigenvalue,

\[
\hat{\rho}' = \frac{P_i \hat{\rho} P_i}{\text{Tr}(P_i \hat{\rho})}.
\] (1.9)

This is the Von Neumann collapse postulate. Since the given state has collapsed into a particular eigenstate of the observable, it is no more useful in the study of the properties of the original state. Thus for a new trial, one has to repeat the preparation procedure to obtain the \( \hat{\rho} \), and rerun the experiment. Such a measurement scheme is called as the Von Neumann projective measurement, and the probabilities \( q_i \) are called the Von Neumann projection valued measure.

To summarise, the density operator \( \hat{\rho} \) completely specifies all the properties of the system. All expectation values of all possible experimental observables \( \hat{\Omega} \) are captured by \( \hat{\rho} \).

### 1.2 Composite systems

Consider a bipartite system \( S \) which consists of subsystems \( A \) and \( B \). Let \( \mathcal{H}_A \) and \( \mathcal{H}_B \) denote the Hilbert spaces of the subsystems, then the Hilbert space of the total system is the tensor product \( \mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B \) of the Hilbert spaces of the subsystems. Let the dimension of \( \mathcal{H}_A \) be \( m \), and that of \( \mathcal{H}_B \) be \( n \). Let \( \{ |\psi_j\rangle \} \) form an ONB in \( \mathcal{H}_A \), and \( \{ |\phi_\alpha\rangle \} \) an ONB in \( \mathcal{H}_B \). Then any pure state \( |\Psi\rangle \) of the combined system can be written as

\[
|\Psi\rangle = \sum_{j,\alpha} c_{j\alpha} |\psi_j\rangle \otimes |\phi_\alpha\rangle. \tag{1.10}
\]

A pure state \( |\Psi\rangle \) of the bipartite system \( S \) is said to be a product state if and only if the expansion coefficients \( c_{j\alpha} \) have the product form \( c_{j\alpha} = x_j y_\alpha \), i.e., the \( m \times n \) coefficient matrix \( c \) is the outer product of two vectors. Any state \( |\Psi\rangle \) which cannot be written in the product form

\[
|\Psi\rangle \neq |\psi\rangle \otimes |\phi\rangle, \tag{1.11}
\]

is said to be entangled.

**Theorem 1.1** Given a state \( |\Psi\rangle \) in the tensor product space \( \mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B \), it can
always be written in the form [2]

\[ |\Psi\rangle = \sum_{j=1}^{r} \sqrt{\lambda_j} |\psi'_j\rangle \otimes |\phi'_j\rangle, \quad \text{where} \]
\[ r \leq \min(m, n), \quad \lambda_j > 0, \quad \sum_j \lambda_j = 1, \quad (1.12) \]

and \(|\psi'_j\rangle\) and \(|\phi'_j\rangle\) are vectors from an ONB in \(\mathcal{H}_A\) and \(\mathcal{H}_B\) respectively.

**Proof:** This can easily be seen from the singular value decomposition of the \(c\) matrix, i.e.,

\[ c \rightarrow c' = VcW^T, \quad (1.13) \]

where \(V\) and \(W\) correspond to independent local unitary change of basis in \(\mathcal{H}_A\) and \(\mathcal{H}_B\) respectively, i.e.,

\[ |\psi'_k\rangle = \sum_j V_{jk} |\psi_j\rangle, \quad |\phi'_\alpha\rangle = \sum_\beta W^{*}_{\alpha\beta} |\phi_\beta\rangle. \quad (1.14) \]

\(V\) and \(W\) are chosen such that \(c'\) is diagonal.

The integer \(r\) is known as the Schmidt rank of \(|\Psi\rangle\). The Schmidt rank \(r\), and the Schmidt coefficients \(\{\lambda_i\}\), constitute the local invariants of the state. It is clear that for product states the Schmidt rank is one. The Schmidt rank \(r\), of a given bipartite pure state \(|\Psi\rangle\), is thus an entanglement witness of the state.

The notion of entanglement for the case of mixed states is much more subtle. As we have already seen in Eq. (1.3), any mixed state of the bipartite system \(A + B\) can be written as

\[ \hat{\rho}_{AB} = \sum_k p_k |\Psi_k\rangle\langle \Psi_k|, \]
\[ p_k > 0, \quad \sum_k p_k = 1. \quad (1.15) \]

A state \(\hat{\rho}_{AB}\) is said to be *separable*, if there exists an ensemble realisation \(|\Psi_k\rangle, \ p_k\) of \(\hat{\rho}_{AB}\), such that all the \(|\Psi_k\rangle\)'s are product states \([3]\), i.e.,

\[ \hat{\rho}_{AB} = \sum_k p_k \hat{\rho}_{Ak} \otimes \hat{\rho}_{Bk}, \quad (1.16) \]
where the \( p_k \)'s are positive, and \( \hat{\rho}_{Ak} \)'s and \( \hat{\rho}_{Bk} \)'s are density operators of subsystems \( A \) and \( B \) respectively. Without loss of generality, these density operators can be chosen to correspond to pure states. Stated differently, any convex sum of product states is by definition a separable state, and separable states constitute a convex subset of the convex state space. Any state which cannot be written as a convex sum of product states is said to be entangled. The subtle part of the definition stems from the fact that one has to run, in principle, through all possible decompositions of a state to conclude if it is separable or not.

### 1.2.1 Partial trace

Consider a bipartite system in the state \( \hat{\rho}_{AB} \). Suppose we are interested in the subsystem \( A \) alone, i.e., we are interested in measurement of a local observable \( \hat{\Omega}_A \) which acts only on the Hilbert space \( \mathcal{H}_A \). The action of such an observable is described by the operator \( \hat{\Omega}_A \otimes \mathbb{1}_B \) on the extended Hilbert space \( \mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B \). Using the resolution of identity in the \( B \) subsystem with the choice of an ONB \( \{ |\phi_\alpha \rangle \} \in \mathcal{H}_B \), the expectation value of the \( A \) subsystem observable \( \hat{\Omega}_A \) is

\[
\text{Tr}(\hat{\Omega}_A \otimes \mathbb{1}_B \hat{\rho}_{AB}) = \text{Tr}_A \text{Tr}_B(\hat{\Omega}_A \otimes \mathbb{1}_B \hat{\rho}_{AB}) = \sum_{k,\alpha} \langle \psi_k | \otimes \langle \phi_\alpha | \hat{\Omega}_A \otimes \mathbb{1}_B \hat{\rho}_{AB} | \psi_k \rangle \otimes | \phi_\alpha \rangle = \text{Tr}_A(\hat{\Omega}_A \hat{\rho}_A), \quad \text{where} \\
\hat{\rho}_A = \text{Tr}_B \hat{\rho}_{AB} = \sum_\alpha \langle \phi_\alpha | \hat{\rho}_{AB} | \phi_\alpha \rangle.
\]

(1.17)

Clearly, \( \hat{\rho}_A \) is an operator on the Hilbert space \( \mathcal{H}_A \). The trace operation executed only on the \( B \) subsystem is called partial trace, and the resulting \( \hat{\rho}_A \) is called the reduced density operator of subsystem \( A \). It is clear that partial trace preserves the defining requirements on a density operator. The notions indicated above hold irrespective of whether \( \hat{\rho}_{AB} \) was pure or mixed. Every observable of subsystem \( A \) sees the state \( \hat{\rho}_{AB} \) as if it were the state \( \hat{\rho}_A \). Clearly, partial trace of pure bipartite states leads to either pure or mixed states of the subsystem. If we begin with a bipartite pure entangled state, the partial traced state of the subsystem is definitely mixed, and the matrix rank of the reduced state is the Schmidt rank of the initial bipartite pure entangled state. Thus, partial trace can be viewed as an entanglement witness for bipartite pure entangled states. The concept of partial trace also renders another explanation to the origin of mixed states, i.e., through the process of considering evolutions in composite systems and then discarding one of the subsystems, we are able to generate mixed states of a subsystem. The notion of mixed
states was earlier considered through the notion of a preparation procedure.

1.2.2 Positive Operator Valued Measure

The notion of Positive Operator Valued Measure (POVM) is the generalisation of the Von Neumann measurement scheme, and is easily understood in the context of measurement in composite systems.

Consider a system $A$ with Hilbert space $\mathcal{H}_A$, to be in the state $\hat{\rho}_A$, and an auxiliary system $B$ with Hilbert space $\mathcal{H}_B$, to be in the state $\hat{\rho}_B$. Then the state of the combined system is

$$\hat{\rho}_S = \hat{\rho}_A \otimes \hat{\rho}_B,$$

where

$$\hat{\rho}_A \otimes \hat{\rho}_B = (\hat{\rho}_A)_{mn}(\hat{\rho}_B)_{\alpha\beta}. \quad (1.18)$$

A Von Neumann measurement on the combined system system is represented by projection operators

$$P_i P_j = \delta_{ij} P_j, \quad \sum_i P_i = \mathbb{I}. \quad (1.19)$$

The probability of the $i$th outcome of such a test given that the state of the combined system is $\hat{\rho}_A \otimes \hat{\rho}_B$ is,

$$q_i = \text{Tr}[P_i (\hat{\rho}_A \otimes \hat{\rho}_B)] = \sum_{m\alpha, n\beta} (P_i)_{m\alpha, n\beta} (\hat{\rho}_A)_{mn} (\hat{\rho}_B)_{\alpha\beta}. \quad (1.20)$$

This can be equivalently written as

$$q_i = \text{Tr}_A (M_i \hat{\rho}_A), \quad \text{where} \quad (M_i)_{mn} = \sum_{\alpha\beta} (P_i)_{m\alpha, n\beta} (\hat{\rho}_B)_{\alpha\beta}, \quad (1.21)$$

and \{ $M_i$ \} are operators on the Hilbert space $\mathcal{H}_A$ of the $A$ subsystem. The hermitian nonnegative operators $M_i$ which need not commute clearly satisfy

$$\sum_i M_i = \mathbb{I}. \quad (1.22)$$

Each member of the set \{ $M_i$ \} is called a positive operator valued measure (POVM) [4, 5], since each $M_i$ is a positive operator by construction. The main difference between a Von Neumann type measurement and a POVM is that the $M_i$ are not projection operators.
and the number of outcomes is independent of the dimensionality of the Hilbert space \( \mathcal{H}_A \). The probability of the \( i \)th outcome is now given by

\[
q_i = \text{Tr}(M_i \hat{\rho}),
\]

as compared to Von Neumann \( \text{Tr}(P_i \hat{\rho}) \). We have removed the subscript \( A \) to indicate the comparison at system level. The Von Neumann collapse postulate holds, except that the state of the subsystem after the measurement is the partial trace of the collapsed state of the composite system.

### 1.3 Quantum Dynamics

Consider a system with Hilbert space \( \mathcal{H} \). The set of all density operators \( \hat{\rho} \) acting on \( \mathcal{H} \) is a subset of the set of all linear operators acting on \( \mathcal{H} \). The set of all linear transformations on \( \mathcal{H} \) forms a vector space. If \( \mathcal{H} \) is \( n \) dimensional, then this new vector space is clearly \( n^2 \) dimensional. Quantum evolutions are linear transformations on the linear operators acting on \( \mathcal{H} \). Linear transformations on this new vector space are called linear maps to distinguish them from linear operators on \( \mathcal{H} \). They are sometimes called super-operators.

It would seem that any linear map acting on the density operator \( \hat{\rho} \), and preserves the three defining properties of density operators in Eq. (1.1), is a valid quantum evolution. This is not so! Further conditions beyond (1.1) arise from looking at composite systems. Suppose we were dealing with only closed systems, then any map which preserves the defining properties of the density operators would have appeared physical. Since we may be a part of a larger system but observing only our system locally, it becomes imperative that the map under consideration takes valid density operators of the extended system to valid density operators. Thus one is lead to consider what are called completely positive maps.

#### 1.3.1 Completely positive maps

Consider a bipartite system \( S \) with Hilbert space \( \mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B \). A linear map \( \Lambda \) acting on the space of operators acting on \( \mathcal{H}_A \) can be extended to act on the space of operators acting on \( \mathcal{H}_S \) through the definition of the map \( \Lambda \otimes \mathbb{1}_n \), where \( \mathbb{1}_n \) is the identity map on subsystem \( B \), i.e., the map \( \Lambda \) acts only on the \( A \) subsystem, but leaves the \( B \) subsystem as it is. Every possible choice of \( \mathcal{H}_B \) gives us a possible extension of the above kind for \( \Lambda \). A map \( \Lambda \) is said to be completely positive if it is positive for all such possible extensions. By positive we mean that it takes valid density operators acting on \( \mathcal{H}_S \) to valid density operators acting on \( \mathcal{H}_S \). The subtle part of the definition stems from the fact that one
has to in principle run over all possible extensions specified by the choices of $\mathcal{H}_B$, to conclude if a given positive map $\Lambda$ is completely positive or not. We have the following important result:

**Theorem 1.2** The action of any trace-preserving completely positive map $\Lambda$ on a density operator $\hat{\rho}$ can always be written in the following form:

$$\Lambda(\hat{\rho}) = \sum_{\alpha} W_{\alpha} \hat{\rho} W_{\alpha}^\dagger, \quad \sum_{\alpha} W_{\alpha}^\dagger W_{\alpha} = \mathbb{1}. \quad (1.24)$$

We don’t give the proof here, but a heuristic way to see this is as follows. Let $\{ | \psi_j \rangle \}$ and $\{ | \phi_\alpha \rangle \}$ be a set of ONB in $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. Suppose that the density operator of the bipartite system is initially in the state

$$\hat{\rho}_A \otimes |0\rangle_B \langle 0|, \quad (1.25)$$

where $|0\rangle_B$ denotes a pure state in the $B$ subsystem. Evolve the state unitarily in the combined system so that

$$\hat{\rho}_A \otimes |0\rangle_B \langle 0| \rightarrow U(\hat{\rho}_A \otimes |0\rangle_B \langle 0|)U^\dagger. \quad (1.26)$$

Now performing a partial trace over the $B$ system Hilbert space yields

$$\hat{\rho}'_A = \text{Tr}_B(U(\hat{\rho}_A \otimes |0\rangle_B \langle 0|)U^\dagger) = \sum_{\alpha} \langle \phi_\alpha | U(\hat{\rho}_A \otimes |0\rangle_B \langle 0|)U^\dagger | \phi_\alpha \rangle$$

$$= \sum_{\alpha} \langle \phi_\alpha | U | 0 \rangle_B \hat{\rho}_{AB} | 0 \rangle_B^\dagger \langle \phi_\alpha \rangle. \quad (1.27)$$

If we denote

$$W_{\alpha} = \langle \phi_\alpha | U | 0 \rangle_B, \quad (1.28)$$

then we can express $\hat{\rho}'_A$ as

$$\hat{\rho}'_A = \sum_{\alpha} W_{\alpha} \hat{\rho}_A W_{\alpha}^\dagger. \quad (1.29)$$
It follows from the unitarity of $U$ that

$$\sum_\alpha W^\dagger_\alpha W_\alpha = \sum_\alpha B\langle 0|U^\dagger|\phi_\alpha\rangle\langle\phi_\alpha|U|0\rangle_B$$

$$= B\langle 0|U^\dagger U|0\rangle_B = \mathbb{1}_A. \tag{1.30}$$

Thus we have two ways of picturing completely positive maps [6–9].

- Every trace preserving completely positive linear map $\Lambda$, has an operator sum representation as in Eq. (1.24).

- Every trace preserving completely positive linear map $\Lambda$, has an unitary representation as in Eq. (1.27).

What we have just demonstrated is that an unitary representation of a completely positive map can in fact be viewed as an operator sum representation. The nontrivial part of the theorems is that every trace-preserving completely positive map can be obtained in this manner. A simple way to understanding this aspect is by reasoning that any reasonable evolution should be accomplished as a unitary (hamiltonian) evolution on a larger system. It is useful to note that in the unitary realization one begins with a product state of the combined system, and a pure state for the $B$ subsystem proves sufficient. It is clear from the definition and Eq. (1.24) that the set of all trace preserving completely positive maps form a convex set.

### 1.4 Detecting entanglement

One of the foremost problems of quantum information theory has been the development of tools for detection of entanglement. Since a given density operator $\hat{\rho}_{AB}$ of a bipartite system $S$, has infinitely many decompositions [1], and since we cannot possibly run through all of them to see if $\hat{\rho}_{AB}$ is separable, it is imperative that we devise efficient methods of detecting entanglement. Bell inequalities provide us with sufficient criteria for entanglement, and entropy based inequalities can also detect entanglement in suitable cases. These are scalar manifestations of entanglement, which has its roots at the density operator level. Since we only measure scalar quantities in the laboratory, the scalar manifestations of entanglement are crucial from an experimental point of view. From a theoretical perspective, the scalar manifestation of entanglement is intimately connected to the theory of positive maps which has a direct bearing on the concept of entanglement. We briefly discuss these ideas below.
1.4.1 Entropic inequalities

Entropic inequalities originate from the observation that there is more information in an entangled state viewed as a whole than viewed as aggregate of information in the subsystems. A simple example such as a maximally entangled state in $2 \times 2$ dimensions illustrates this. The state when viewed in $2 \times 2$ dimensions as a whole is a pure state, but when we look at the state of either of the subsystems the state is a random mixture proportional to the identity operator. Thus from either of the subsystems we gain no knowledge of the state. Such a qualitative feature can be made quantitative through entropic inequalities such as

$$S(\hat{\rho}_A) \leq S(\hat{\rho}_{AB}), \quad S(\hat{\rho}_B) \leq S(\hat{\rho}_{AB}),$$

where $S(\hat{\rho}) = -\text{Tr}(\hat{\rho} \log_2 \hat{\rho})$ is the Von Neumann entropy of a state $\hat{\rho}$. Any separable state obey the inequalities, but entangled states need not [10–13]. The idea is easily generalised to entropic inequalities such as those based on Renyi quantum entropies.

1.4.2 Majorisation

Majorisation is a technique that helps us compare two vector quantities. In the context of classical probability theory, it becomes useful when we compare two discrete probability distributions. Based on majorisation, we may be able to conclude if one probability distribution is more ‘spread out’ than the other, or in other words if one probability distribution is more ‘disordered’ than the other. We now state the definition [14, 15].

Let $x = (x_1, x_2, \cdots, x_n)$ and $y = (y_1, y_2, \cdots, y_n)$ be two vector quantities, arranged in the nondecreasing order, i.e., $x_1 \leq x_2 \leq \cdots \leq x_n$ and $y_1 \leq y_2 \leq \cdots y_n$. Then we say $x < y$ ($x$ is majorised by $y$) if and only if

$$\sum_{j=1}^k x_j \geq \sum_{j=1}^k y_j, \quad \forall \quad k \leq n, \quad \text{and}$$

$$\sum_{j=1}^n x_j = \sum_{j=1}^n y_j.$$  

(1.32)

That majorisation captures the disorderliness is seen through its implication on entropy. Let the vectors $x$ and $y$ denote two probability distributions and let $x < y$, then $H(\{x_i\}) \geq H(\{y_i\})$, where $H(.)$ is the Shannon entropy [14, 16]. The majorisation relation is more fundamental in capturing disorderliness, in the sense that the entropic inequality can be seen to follow as a consequence of the majorisation relation.
As already mentioned, entangled states have more information when seen as a whole rather than in their parts. This statement is made mathematically precise through the following majorisation relation:

**Theorem 1.3** If a bipartite mixed state $\hat{\rho}_{AB}$ is separable and $\hat{\rho}_A$ and $\hat{\rho}_B$ are the reduced density matrices of the subsystems $A$ and $B$, then

$$\lambda(\hat{\rho}_{AB}) < \lambda(\hat{\rho}_A), \text{ and } \lambda(\hat{\rho}_{AB}) < \lambda(\hat{\rho}_B).$$

(1.33)

Here, $\lambda(\hat{\rho}_{AB})$, $\lambda(\hat{\rho}_A)$, and $\lambda(\hat{\rho}_B)$, constitute the eigenvalues of $\hat{\rho}_{AB}$, $\hat{\rho}_A$, and $\hat{\rho}_B$, arranged in the nondecreasing order [17].

The above majorisation relation can be violated if the state $\hat{\rho}_{AB}$ is entangled. Clearly, entropic inequalities such as Eq. (1.31) are implied by the majorisation relation in Eq. (1.33).

Majorisation relations are not only useful in detecting entanglement, but appear in more general scenarios such as when we are dealing with quantum evolutions. The Schur-Horn lemma plays a pivotal role in this [14, 15]. We will discuss more of the majorisation relations in Chapter 5, where we study compatibility relations for Gaussian states [18].

### 1.4.3 Bell’s inequalities

Bell inequalities arose initially from the study of quantum theory from the perspective of classical probability theory. One of the profound implications of quantum theory is that it gives rise to new possibilities in the correlation of distant events that cannot be explained by classical local models. Bell observed that there is an upper bound on the correlation of distant events as explained by a classical local model, and that quantum mechanics could violate it [19]. It was evident that any such violation was easily explained through entanglement. We now have the well established fact that violation of any of the Bell type inequalities is a clear manifestation of entanglement. Here we briefly discuss one such inequality, namely the Clauser-Horne- Shimony-Holt (CHSH) inequality [20].

The CHSH inequality refers to correlation experiments involving two dichotomic observables at two sites. The observed values of each of these observables can be taken to be $\pm 1$. Let the observables in the Alice’s side be denoted $A_1$ and $A_2$, and those on Bob’s side $B_1$ and $B_2$. The outcomes of the experiment in each trial is denoted by $a_1, a_2$ and $b_1, b_2$ respectively. Define the correlation function between two observables $A$ and $B$, $A$ on Alice’s side and $B$ on Bob’s side, as

$$E(A, B) = \langle ab \rangle,$$

(1.34)
which is the average value of the correlation over repeated trials. Then the CHSH inequality reads as

$$|\mathcal{B}| = |E(A_1, B_1) + E(A_1, B_2) + E(A_2, B_1) - E(A_2, B_2)| \leq 2.$$  \hspace{1cm} (1.35)

This can be easily seen from the fact that for a given trial \( \mathcal{B} \) is \( \pm 2 \), hence the average is always less than or equal to 2.

The quantum mechanical version of the CHSH inequality for \( 2 \times 2 \) dimensional systems is easily stated through the definition of the Bell-CHSH observable

$$\mathcal{B} = \hat{a}_1 \cdot \sigma \otimes (\hat{b}_1 + \hat{b}_2) \cdot \sigma + \hat{a}_2 \cdot \sigma \otimes (\hat{b}_1 - \hat{b}_2) \cdot \sigma.$$  \hspace{1cm} (1.36)

Here \( \hat{a}_1, \hat{a}_2, \hat{b}_1 \) and \( \hat{b}_2 \) are arbitrary unit vectors in \( R^3 \), \( \hat{a} \cdot \sigma = \sum_{i=1}^{3} a_i \sigma_i \), and \( \sigma_i \) are the Pauli matrices. Any \( \hat{a} \cdot \sigma \) corresponds to a spin observable with eigenvalues \( \pm 1 \). With this definition the CHSH inequality for a bipartite \( 2 \times 2 \) dimensional system reads as

$$\text{Tr}(\hat{\rho} \mathcal{B}) \leq 2.$$  \hspace{1cm} (1.37)

The constraint imposed by the above equation is not generally obeyed by quantum mechanical systems. For instance, for the choice of the various unit vectors, each separated by angle of 22.5 degrees and with the singlet chosen as the state, \( \text{Tr}(\hat{\rho} \mathcal{B}) = 2\sqrt{2} \) [5], clearly a violation of the CHSH bound implied by a local classical model. An upper bound on the maximum possible expectation value of the Bell-CHSH observable was obtained by Cirelson to be \( 2\sqrt{2} \) [21]. The remarkable aspect of the inequality is that experiments confirmed the violation [22], in complete agreement with quantum mechanical predictions, thus demonstrating entanglement. This inequality can be extended to more observables on each side [23] in a bipartite setup, and to more number of parties [24]. Violation in the latter case will indicate presence of multipartite entanglement. The Bell inequalities are only one of the ways to detect entanglement, and they are not very powerful in the sense that there are inseparable states obeying the Bell inequalities [3, 25]. That is, there is entanglement that is not detected through violation of Bell inequalities. See [26] for a review.

### 1.4.4 Positive maps

The theory of positive maps is an inescapable ingredient in the theory of entanglement. The first use of them was demonstrated by Peres in [27]. He observed that a separable state remains a state (positive) if subjected to partial transposition (PT). A crucial
observation by the Horodecki’s that partial transpose is a positive map but not a completely positive one, led to the exploration of the intimate connection between the theory of entanglement and positive maps [28]. Earlier in Eq. (1.24), we had introduced the notion of completely positive map. We noted that a linear map $\Lambda$ is completely positive if and only if the extended map $\Lambda \otimes \mathbb{1}_n$ is positive for all $n$. We now introduce the notion of positive but non completely positive maps. We say that a linear map $\Lambda$ is positive but not completely positive if $\Lambda$ takes density operators to density operators, but it has an extension which fails to do so.

Let the map $\Lambda$ act on operators on Hilbert space $\mathcal{H}_A$, and let its extension act on operators on the Hilbert space $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B$. Suppose the extended map given by $\Lambda \otimes \mathbb{1}_n$, does not take positive operators on $\mathcal{H}_S$ to positive operators on $\mathcal{H}_S$, but the map $\Lambda$ takes positive operators on $\mathcal{H}_A$ to positive operators on $\mathcal{H}_A$. Then the map $\Lambda$ is said to be positive but not completely positive. Such a map could be too a in detecting entanglement in bipartite systems specified by the Hilbert space $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B$. If such a map acts on a separable state $\hat{\rho}_{AB}$ on $\mathcal{H}_S$, we have the following to be true:

$$(\Lambda \otimes \mathbb{1}_n)(\hat{\rho}_{AB}) = (\Lambda \otimes \mathbb{1}_n) \sum_k p_k \hat{\rho}_{Ak} \otimes \hat{\rho}_{Bk}$$

$$= \sum_k p_k \Lambda(\hat{\rho}_{Ak}) \otimes \hat{\rho}_{Bk} \geq 0. \quad (1.38)$$

This follows from the linearity and positivity of the map $\Lambda$. On the other hand, if the extended map acted on an entangled state, it could lead to the following possibility:

$$(\Lambda \otimes \mathbb{1}_n)(\hat{\rho}_{AB}) = (\Lambda \otimes \mathbb{1}_n) \left( \sum_k p_k \hat{\rho}_{ABk} \right)$$

$$= \sum_k p_k (\Lambda \otimes \mathbb{1}_n)(\hat{\rho}_{ABk}) \geq 0. \quad (1.39)$$

This possibility arises because the map $\Lambda$ is not completely positive. Thus a positive but not completely positive map helps us detect entanglement! We have the following important theorem [28]:

**Theorem 1.4** Let $\hat{\rho}_{AB}$ act on the Hilbert space $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B$. Then $\hat{\rho}_{AB}$ is separable if and only if for every positive map $\Lambda$ on $\mathcal{H}_A$, the operator

$$(\Lambda \otimes \mathbb{1}_n)(\hat{\rho}_{AB}) \quad (1.40)$$

is positive.

We now give a brief discussion on positive maps. It is convenient to go to an indexed
notation to describe them. The three defining properties of a density operator in this notation are

\[ \rho_{rs} = \rho_{sr}^*, \quad \rho_{rs} x_r^* x_s \geq 0, \quad \rho_{rr} = 1. \]  

(1.41)

Summation over repeated indices is implied as usual. Any linear positive map \( \Lambda \) takes valid density operators to valid density operators. Such a map can be expressed as

\[ \Lambda : \hat{\rho} \rightarrow \hat{\rho}', \quad \rho_{r's'} = \Lambda_{r's'r,s} \rho_{rs}. \]  

(1.42)

The hermiticity requirement of \( \hat{\rho}' \) demands that

\[ \Lambda_{s'r',sr} = \Lambda_{r's'r,s}^*. \]  

(1.43)

trace preservation of \( \hat{\rho}' \) demands that

\[ \Lambda_{r'r',rs} = \delta_{rs}, \]  

(1.44)

and preservation of positivity implies that

\[ \rho_{r's'} x_r^* x_s' \geq 0 \quad \Rightarrow \quad \Lambda_{r's',rs} \rho_{rs} x_r^* x_s' \geq 0. \]  

(1.45)

From spectral resolution of \( \hat{\rho} \), it is sufficient to consider positivity requirement on any of its projectors, hence we only require

\[ \Lambda_{r's',rs} x_r^* x_s' y_r^* y_s \geq 0. \]  

(1.46)

Let us define a new matrix \( M \) by permuting the indices of \( \Lambda \):

\[ M_{r's',rs} = \Lambda_{r's'r,s}, \]  

(1.47)

The hermiticity condition in Eq. (1.43) now reads

\[ M_{r's',rs} = M_{r's'r,s}^*, \]  

(1.48)

i.e., the matrix \( M \) is hermitian. Hence we have the spectral resolution

\[ M_{r's',rs} = \sum_{\alpha} \lambda_\alpha \xi_{r'r}^*(\alpha) \xi_{s's}^*(\alpha). \]  

(1.49)
Rewriting the positivity requirement of $\hat{\rho}'$ in Eq. (1.45) in terms of the matrix $M$, we have

$$x_r^*y_r^* M_{r',s'}x_{s'}y_s,$$

which means that the matrix $M$ is positive on product vectors. With this, any positive linear map reads as

$$\rho'_{r',s'} = M_{r',s'} \rho_{rs}$$

$$= \sum_{\alpha} \lambda_{\alpha} \xi_{r',(\alpha)} \rho_{rs} \xi_{s',(\alpha)}^{\dagger} \Rightarrow$$

$$\hat{\rho}' = \sum_{\alpha} \lambda_{\alpha} \hat{\xi}(\alpha) \hat{\rho}^{\dagger}(\alpha).$$

The eigenvalues $\{\lambda_{\alpha}\}$ need not be positive, since our only requirement was positivity $M$ on product vectors. A completely positive map is a positive map as in Eq. (1.51) where all its eigenvalues, i.e., its $\{\lambda_{\alpha}\}$ are positive [29–31].

### 1.4.5 Entanglement Witness

The convex structure of the state space and the convex structure of the set of all separable states for a given bipartite system enable us employ the ideas known in the context of convex sets to distinguish entangled states from separable ones. The simple idea that a convex set and a point lying outside it can always be separated by a hyper plane is manifested in the context of state space as an entanglement witness. The following theorem due to Horodecki’s captures this idea [28].

**Theorem 1.5** A state $\hat{\rho}_{AB}$ acting on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ is separable if and only if

$$\text{Tr}(\hat{A}\hat{\rho}_{AB}) \geq 0$$

for every hermitian operator $\hat{A}$ satisfying $\text{Tr}(\hat{A}\hat{P} \otimes \hat{Q}) \geq 0$, where $\hat{P}$ and $\hat{Q}$ are projections operating on $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively.

The implication of the theorem is that if a state $\hat{\rho}_{AB}$ satisfies the inequality

$$\text{Tr}(\hat{A}\hat{\rho}_{AB}) < 0$$

for such hermitian $\hat{A}$ which is positive on product vectors, we definitely know that the state $\hat{\rho}_{AB}$ was entangled. Any $\hat{A}$ which is positive semidefinite will not serve our purpose.
in detecting entanglement, as such an operator would be positive on entangled states too. Thus the operator \( \hat{A} \) though positive on product vectors, has to be necessarily indefinite to detect entanglement of some states. Thus, if such a witness \( \hat{A} \) were to detect entanglement of a state \( \hat{\rho}_{AB} \), the state \( \hat{\rho}_{AB} \) is a point lying outside the convex set of separable states, and the numerical matrix determined by \( \hat{A} \) is a hyper plane separating this point from the convex subset of separable states. The observation that any entanglement witness is a hermitian operator which is positive on product vectors but not positive definite, reminds us of the defining property of a positive map in Eq. (1.51). In other words, we have an isomorphism between positive but not completely positive maps and entanglement witnesses [32]. The first use of an entanglement witness was in [3], where the flip operator \( \hat{V} \) served as an entanglement witness. It is easy to check for systems of \( 2 \times 2 \) dimensions that the flip operator defined in the standard basis is the \( M \) matrix corresponding to partial transpose. It is useful to note that any Bell observable \( \hat{B} \) as in Eq. (1.36) can also be viewed as an entanglement witness.

### 1.4.6 Partial Transpose

The partial transpose test is a classic example of the application of the theory of positive maps in detecting entanglement. Partial transpose was initially introduced by Peres in [27] to detect entanglement. It is defined as follows. Consider the matrix elements of a state \( \hat{\rho}_{AB} \) of a bipartite system specified by the Hilbert space \( \mathcal{H}_S \) of \( m \times n \) dimensions, i.e.,

\[
\hat{\rho}_{\alpha j, \beta} = \langle \psi_i | \otimes \langle \phi_a | \hat{\rho}_{AB} | \psi_j \rangle \otimes | \phi_\beta \rangle. \tag{1.54}
\]

The partial transposed matrix \( \hat{\rho}_{AB}^{PT} \) is defined through its matrix elements thus:

\[
\hat{\rho}_{\alpha j, \beta}^{PT} = \hat{\rho}_{\beta j, \alpha}. \tag{1.55}
\]

Such an operation is easily seen in its matrix form. The state \( \hat{\rho}_{AB} \) of the \( m \times n \) system can be written as

\[
\hat{\rho}_{AB} = \begin{pmatrix}
A_{11} & \cdots & A_{1m} \\
\vdots & \ddots & \vdots \\
A_{m1} & \cdots & A_{mm}
\end{pmatrix}, \tag{1.56}
\]

i.e., as an \( m \times m \) array of \( n \times n \) matrices \( A_{ij} \) acting on the second Hilbert space \( \mathcal{H}_B \). The whole matrix is defined by the matrix elements \( \{ A_{ij} \} \alpha \beta = \rho_{\alpha \beta} \). Then partial transpose
is simply the transposition of the \( \{ A_{ij} \} \) matrices.

\[
\hat{\rho}_{AB}^{PT} = \begin{pmatrix}
A_{11}^T & \cdots & A_{1m}^T \\
\vdots & \ddots & \vdots \\
A_{m1}^T & \cdots & A_{mm}^T
\end{pmatrix}.
\]

Under partial transpose a separable state goes to a valid state, i.e.,

\[
\hat{\rho}_{AB}^{PT} = \sum_k p_k \hat{\rho}_{Ak} \otimes \hat{\rho}_{Bk}^T = \sum_k p_k \hat{\rho}_{Ak} \otimes \hat{\rho}_{Bk}^* \geq 0.
\]

Since every \( \hat{\rho}_{Bk}^* \) is a valid state, \( \hat{\rho}_{AB}^{PT} \) is a valid state. An entangled state under partial transposition can result in a non positive operator. Thus partial transpose detects entanglement. Even though the operation is basis dependent, the conclusions are not. We have the following theorem for lower dimensional composite systems [28]:

**Theorem 1.6** A state \( \hat{\rho}_{AB} \) of a \( 2 \otimes 2 \) or \( 2 \otimes 3 \) system is separable if and only if its partial transposition is a positive operator.

### 1.4.7 Reduction Criterion

Yet another example of a positive map that is not completely positive is the reduction map. Consider the map

\[
\Lambda(\hat{\rho}) = \mathbb{1} - \hat{\rho}.
\]

This maps is clearly positive, since the eigenvalues of a density operator are positive and less than one. Now consider the extension of this map to composite systems. A separable state satisfies the following inequalities:

\[
\mathbb{1} \otimes \hat{\rho}_B - \hat{\rho}_{AB} \geq 0,
\hat{\rho}_A \otimes \mathbb{1} - \hat{\rho}_{AB} \geq 0,
\]

but entangled states need not. Here \( \hat{\rho}_A \) and \( \hat{\rho}_B \) are the partially traced versions of \( \hat{\rho}_{AB} \) for Bob’s and Alice’s subsystems respectively. The above two conditions are jointly called the reduction criterion [33]. The reduction inequalities in Eq. (1.60) imply the majorisation inequalities in Eq. (1.33), and consequently imply the entropic inequalities in Eq. (1.31) [34]. It is known that the reduction criterion is weaker than partial transpose test in detecting entanglement, i.e., there are entangled states that partial transpose can detect, but reduction criterion cannot [33]. Thus, entanglement criteria defined through the reduction map, majorisation, and entropic inequalities, are all weaker than
partial transpose in detecting entanglement. It is also known that more general entropic inequalities other than that stated in Eq. (1.31) are implied by the reduction criteria [35].

We now digress to introduce the important notion of distillability before continuing the discussion on positive maps. This digression is useful, as both the partial transpose map and the reduction criterion are intimately connected to the concept of distillation.

1.5 Distillation

A central aspect of quantum information theory is transmission of quantum information, i.e., transmission of quantum states through quantum channels. The whole idea rests largely on the teleportation protocol [36] which consumes a maximally entangled state in the process of transmitting a qubit from Alice’s side to Bob’s side. Anything less than a maximally entangled state results in the loss of fidelity of the transmission. Thus a primary resource for achieving a perfect teleportation is a shared maximally entangled state. Thus arises the need for creating maximally entangled states starting with lesser entangled states. The aim of a distillation or purification procedure is to extract from a large ensemble of low fidelity EPR pairs a smaller sub-ensemble of high fidelity EPR pairs, using only local operations and classical communication (LOCC); these may then be used for faithful teleportation [37–39].

We begin by illustrating a purification procedure in the context of pure states, and then extend it to mixed states, before considering formal definitions. We stick to $2 \times 2$ dimensional systems for the sake of simplicity.

Suppose Alice and Bob share two copies of an entangled state $|\psi\rangle$ which is not maximally entangled, i.e.,

$$ |\psi\rangle \otimes |\psi\rangle = (\alpha |0\rangle_A |0\rangle_B + \sqrt{1 - \alpha^2} |1\rangle_A |1\rangle_B) \otimes (\alpha |0\rangle_{A'} |0\rangle_{B'} + \sqrt{1 - \alpha^2} |1\rangle_{A'} |1\rangle_{B'}), \quad (1.61) $$

where $|0\rangle$ and $|1\rangle$ are the eigenstates of the $\sigma_z$ operator, and $A$ and $A'$ refer to the Alice’s side particles, and $B$ and $B'$ to Bob’s. Expanding the above, we have

$$ |\psi\rangle \otimes |\psi\rangle = \alpha^2 |0\rangle_A |0\rangle_B |0\rangle_{A'} |0\rangle_{B'} + (1 - \alpha^2) |1\rangle_A |1\rangle_B |1\rangle_{A'} |1\rangle_{B'} + \sqrt{2} \alpha \sqrt{1 - \alpha^2} \left[ |0\rangle_A |1\rangle_B |1\rangle_{A'} |0\rangle_{B'} + |1\rangle_A |0\rangle_B |0\rangle_{A'} |1\rangle_{B'} \right] \quad (1.62) $$

for the state of the two pairs. Now let Bob make a local measurement for the z-component of the spin. He can get either 2, -2, or zero. Suppose the result is zero, then Bob informs
Alice over phone and they decide by mutual consent to retain the state, else they decide to discard the state. The probability of such an outcome is \(2\alpha^2(1 - \alpha^2)\). The resultant state after the experiment is the term indicated in the square bracket:

\[
|\psi_r\rangle = \frac{\left|0\right>_A\left|0\right>_B\left|1\right>_A\left|1\right>_B' + \left|1\right>_A\left|1\right>_B\left|0\right>_A'\left|0\right>_B'}{\sqrt{2}}.
\] (1.63)

The subscript \(r\) denotes that the state is a result of local measurement. Suppose we do the following renaming, i.e., \(\left|0'\right>_A = \left|0\right>_A\left|1\right>_A'\), \(\left|1'\right>_A = \left|1\right>_A\left|0\right>_A'\), \(\left|0'\right>_B = \left|0\right>_B\left|1\right>_B'\), and \(\left|1'\right>_B = \left|1\right>_B\left|0\right>_B'\), the state that Alice and Bob share is then a singlet. Such procedure of generating the maximally entangled state is inefficient in the sense that we may end up losing part of the entanglement the parties initially shared through our very act of discarding some of the outcomes. But the method gets more and more efficient as Alice and Bob apply such a protocol collectively to more and more pairs. It was shown in [38] that Alice and Bob can obtain \(n\) singlets from \(k\) \(|\psi\rangle\)'s with the ratio approaching

\[
\lim_{n,k \to \infty} \frac{n}{k} = E(|\psi\rangle) = -\alpha^2 \log_2 \alpha^2 - (1 - \alpha^2) \log_2 (1 - \alpha^2).
\] (1.64)

\(E(|\psi\rangle)\) is the entropy of entanglement and equals the Shannon entropy of the squares of the Schmidt coefficients of \(|\psi\rangle\).

Distillation procedures in the context of mixed states are more sophisticated. For the sake of simplicity we outline a procedure illustrated in [40]; it captures the essence of distillation in the context of mixed states. Suppose Alice and Bob share two copies of the mixed state

\[
\hat{\rho}_{AB} = f|\phi^+\rangle_{AB}\langle\phi^+| + (1 - f)|\psi^+\rangle_{AB}\langle\psi^+|,
\] (1.65)

where \(|\phi^+\rangle_{AB}\) and \(|\psi^+\rangle_{AB}\) are Bell states defined as

\[
|\phi^+\rangle_{AB} = \frac{|00\rangle_{AB} + |11\rangle_{AB}}{\sqrt{2}},
\]

\[
|\psi^+\rangle_{AB} = \frac{|01\rangle_{AB} + |10\rangle_{AB}}{\sqrt{2}}.
\] (1.66)

Unless \(f = \frac{1}{2}\), the state is inseparable. This can be seen for example through the partial transpose test. The aim here is to increase the fraction \(f\) of \(|\phi^+\rangle_{AB}\langle\phi^+|\) through some local operations and decisions taken through mutual consent arrived at through classical communication. To this end, they perform a bilateral local CNOT operation, i.e., CNOT operations performed on pairs \(AA'\) and \(BB'\). Such an operation is clearly local across
the Alice and Bob divide. The action of such an operation is summarised below.

\[
\begin{align*}
|\phi^+\rangle_A|\phi^+\rangle_B \rightarrow |\phi^+\rangle_A|\phi^+\rangle_{A'B'}, \\
|\psi^+\rangle_A|\phi^+\rangle_B \rightarrow |\psi^+\rangle_A|\psi^+\rangle_{A'B'}, \\
|\phi^+\rangle_A|\psi^+\rangle_B \rightarrow |\phi^+\rangle_A|\psi^+\rangle_{A'B'}, \\
|\psi^+\rangle_A|\psi^+\rangle_B \rightarrow |\psi^+\rangle_A|\phi^+\rangle_{A'B'}.
\end{align*}
\]

After this local operation the state Alice and Bob share is

\[
\hat{\rho}_{AB'BB'} = (f^2|\phi^+\rangle_A \langle \phi^+| + (1 - f)^2|\psi^+\rangle_A \langle \psi^+|) \otimes |\phi^+\rangle_{A'B'} \langle \phi^+| + f(1 - f)(|\phi^+\rangle_A \langle \phi^+| + |\psi^+\rangle_A \langle \psi^+|) \otimes |\psi^+\rangle_{A'B'} \langle \psi^+|.
\]\n
Now they perform local z-component measurements on the $A'$ and $B'$ particles. If they get the results to be correlated, they decide to retain the remaining pair, else they discard the remaining pair. The success probability is given by $f^2 + (1 - f)^2$, and they share the state

\[
\hat{\rho}_A = \frac{1}{f^2 + (1 - f)^2} (f^2|\phi^+\rangle_A \langle \phi^+| + (1 - f)^2|\psi^+\rangle_A \langle \psi^+|).
\]\n
The fraction \( f' = f^2/(f^2 + (1 - f)^2) > f \) for \( f > 1/2 \). The procedure is clearly lossy as in the case of pure states in that we may lose entanglement in some of the outcomes. But given a sufficiently large initial ensemble we can, with a finite probability at every step, tend towards a higher fidelity of the desired pure maximally entangled state by iterating this procedure. Alice and Bob can, at the end of the procedure, distill a smaller ensemble of pairs with entanglement fidelity \( f \) arbitrarily close to unity [40]. These pairs can then be used for faithful teleportation. Distillation procedures in the context of mixed states were initially described in [37].

A careful analysis of a general distillation procedure tells us that there are three aspects to it, namely, local general measurements, classical communication, and post-selection.

\textit{Local general measurements} : These are the most general possible measurements performed on the Alice’s and Bob’s side. They are described by two sets of operators $A_i$ and $B_j$ which satisfy the completeness relations

\[
\sum_i A_i^\dagger A_i = \mathbb{I}, \quad \sum_j B_j^\dagger B_j = \mathbb{I}.
\]

They can be realised by appending additional systems locally, evolving them together,
and then performing joint measurements locally. Under these actions, a shared state evolves as

$$\hat{\rho}_{AB} \rightarrow \sum_{ij} A_i \otimes B_j \hat{\rho}_{AB} A_i^\dagger \otimes B_j^\dagger.$$  \hspace{1cm} (1.71)

Such a procedure is clearly local.

*Classical communication*: This simply means that while performing the local general measurements, Alice’s and Bob’s actions can be correlated through mutual consent arrived at through communication over the phone, in which case, the evolution is described by

$$\hat{\rho}_{AB} \rightarrow \sum_i A_i \otimes B_i \hat{\rho}_{AB} A_i^\dagger \otimes B_i^\dagger.$$  \hspace{1cm} (1.72)

This describes a combination of both local general measurement and classical communication.

*Post-selection*: This is the procedure by which Alice and Bob choose to retain certain outcomes of the local general measurement subject to classical communication. Suppose they retained the state corresponding to the $i^{th}$ and $j^{th}$ local outcomes, the resultant state is

$$\frac{A_i \otimes B_j \hat{\rho}_{AB} A_i^\dagger \otimes B_j^\dagger}{\text{Tr}(A_i \otimes B_j \hat{\rho}_{AB} A_i^\dagger \otimes B_j^\dagger)}. \hspace{1cm} (1.73)$$

The denominator in Eq. (1.73) ensures the normalisation [41–44].

Any manipulation involving the above three procedures is succinctly called as local operation and classical communication (LOCC). Clearly any LOCC operation described by Eq. (1.71), is a completely positive map as in Eq. (1.49), but each of the $W$’s are in tensor product form $A_i \otimes B_j$. Thus, any LOCC can be thought of as a separable super operator acting on $\hat{\rho}_{AB}$. It is known that every LOCC can be represented by a separable super operator, but not every separable super operator is an LOCC [44].

To summarise, a purification or distillation procedure is essentially one of extracting singlets from multiple copies of shared entangled states through LOCC. A state is said to be distillable if one can actually distill, using a pre-agreed protocol, pure singlets from multiple copies of the state. However such a definition may appear imprecise from the implementation point of view, since the set the of all LOCC available to the two parties is truly enormous.

It was shown in [45] that every entangled state of a $2 \times 2$ dimensional system is distil-
lable. The idea is that even if an entangled state in $2 \times 2$ dimensions had singlet fraction less than $\frac{1}{2}$, it can always be converted by LOCC, i.e., a local filtering operation, to a state with singlet fraction greater than $\frac{1}{2}$ and, consequently, the recursion protocol described in [39], could be used for further distillation. A tangible criterion for distillability follows from this fact [45].

**Theorem 1.7** A state $\dot{\rho}_{AB}$ is distillable if and only if, for some two dimensional projectors $P$ and $Q$, and for some $n$, the state $P \otimes Q \rho_{AB}^{\otimes n} P \otimes Q$ is entangled.

Thus, projecting out the given $n$ copies of the state $\dot{\rho}_{AB}$ into a $2 \times 2$ dimensional subspace, and demonstrating entanglement in that subspace, amounts to showing the distillability of the state in contention.

An implication of this definition is that states which are positive under partial transpose (PPT) cannot be distilled by LOCC. This can for example be seen from the fact if a state is PPT, then $n$ copies of the state is also PPT, and LOCC doesn’t take PPT states outside the set of PPT states, i.e., LOCC is a PPT-preserving operation. Thus if

$$\dot{\rho}_{AB} = \sum_i A_i \otimes B_i \dot{\rho}_{AB} A_i^\dagger \otimes B_i^\dagger,$$

then

$$\dot{\rho}_{AB}^{\text{PT}} = \sum_i A_i \otimes B_i^T \dot{\rho}_{AB}^{\text{PT}} A_i^\dagger \otimes B_i^{TT}. \quad (1.74)$$

If $\dot{\rho}_{AB}^{\text{PT}}$ is positive, $\dot{\rho}_{AB}^{\text{PT}}$ is also positive, since it is derived by the action of a completely positive map on $\dot{\rho}_{AB}^{\text{PT}}$. Thus, negativity under partial transposition (NPT) is a necessary condition for distillability. It is known that any state that violates the reduction criteria in Eq. (1.60) is distillable [33].

Having introduced the useful concept of distillation or purification, and having shown that negativity under partial transposition is a necessary condition for distillability, an immediate question that arises is the following. Are there entangled states that are nondistillable? A immediate way of answering this question is by answering the following simpler question. Are there states that are PPT? The answer to this was given by the Horodecki’s in [45, 46]. There they constructed states that are entangled but PPT. Such states were called **bound entangled**, meaning one cannot distill any entanglement from them. To detect entanglement in states that are PPT is a nontrivial task, since we need to devise methods that are ‘stronger’ that partial transpose. This has led the search for stronger criteria in detecting entanglement. We discuss some of them below.
1.5.1 Range Criterion

The range criterion is a useful tool to detect entanglement, particularly in cases where partial transpose fails. The statement of the criterion is summarised in the following theorem.

Theorem 1.8 If a state $\hat{\rho}_{AB}$ acting on the space $\mathcal{H}_{AB}$ is separable, then there exists a family of product vectors $|\psi_j\rangle \otimes |\phi_\alpha\rangle$ such that,

(i) they span the range of $\hat{\rho}_{AB}$,

(ii) the vectors $\{|\psi_j\rangle \otimes |\phi_\alpha^*\rangle\}_{i=1}^k$ span the range of $\hat{\rho}_{AB}^{PT}$ (where $*$ denotes complex conjugation in the basis in which partial transposition was performed).

In [46], the range criterion was applied to detect entanglement in a PPT state. Consequently, it led to the elegant construction of UPB’s (Unextendible Product Basis) in [47]. These are a set of product orthogonal vectors in $\mathcal{H}_{AB}$ that has fewer elements than the dimension of the space, but there does not exist any product vector orthogonal to all of them. Thus, a UPB is a partial basis that cannot be completed into a product basis. A simple example of such a UPB for a $3 \otimes 3$ dimensional system is

$$
|\psi_0\rangle = \frac{1}{\sqrt{2}} |0\rangle \otimes (|0\rangle - |1\rangle), \quad |\psi_2\rangle = \frac{1}{\sqrt{2}} |2\rangle \otimes (|1\rangle - |2\rangle),
$$

$$
|\psi_1\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \otimes |2\rangle, \quad |\psi_3\rangle = \frac{1}{\sqrt{2}} (|1\rangle - |2\rangle) \otimes |0\rangle,
$$

$$
|\psi_4\rangle = \frac{1}{3} (|0\rangle + |1\rangle + |2\rangle) \otimes (|0\rangle + |1\rangle + |2\rangle). \quad (1.75)
$$

Given a UPB, it is easy to construct bound entangled states. Consider the projector

$$
P_{UPB} = \sum_{i=0}^4 |\psi_i\rangle \langle \psi_i|, \quad (1.76)
$$

now construct the state

$$
\hat{\rho}_{AB} = \frac{1}{4} (\mathbb{1} - P_{UPB}). \quad (1.77)
$$

The range of $\hat{\rho}_{AB}$ contains no product vectors, otherwise one would have been able to extend the product basis. Clearly $\hat{\rho}_{AB}^{PT}$ is positive, i.e., $\hat{\rho}_{AB}$ is PPT. The state $\hat{\rho}_{AB}$ is thus PPT entangled, and hence bound entangled. The range criteria as a tool was successfully extended to the case of continuous variables in [48], to detect bound entanglement.
1.5.2 Non-decomposable maps

The presence of PPT bound entanglement necessitates the definition of decomposable and non-decomposable positive maps. A positive map is said to be decomposable, if it can be represented in the form

$$\Lambda = \Lambda_{CP}^1 + \Lambda_{CP}^2 \circ T,$$  \hspace{1cm} (1.78)

where $\Lambda_{CP}^1$ and $\Lambda_{CP}^2$ are completely positive maps, $\circ$ denotes composition, and $T$ denotes the transposition map. Clearly, a decomposable map acting on a composite system is no stronger than partial transpose in detecting entanglement. It turns out that in $2 \otimes 2$ and $2 \times 3$ dimensional systems, every positive map is decomposable [49, 50]. Thus positivity under partial transpose turns out to be a necessary and sufficient condition for entanglement in these systems, as indicated earlier. However, in higher dimensional systems there are positive maps that cannot be decomposed as above. Any map that cannot be written as in Eq. (1.78), is said to be a non-decomposable map. Clearly, non-decomposable maps can potentially detect entanglement that partial transpose cannot. Thus the study of non-decomposable maps has turned out to be an integral part of entanglement theory. Examples of such maps can be found in [6, 51–54]. In [51], it was demonstrated that given a UPB one could construct a non-decomposable map.

There seems to be no simple universal way of showing a positive map to be non-decomposable. One possible way is to construct a PPT state, and then show that the map detects entanglement in the PPT state, thus establishing that the map is indecomposable. This particular route to demonstrating non-decomposability has led to search for systematically characterising PPT states [55, 56].

1.5.3 Uncertainty relations

The technique of detecting entanglement through uncertainty relations is based on the fact that separable states, in addition to obeying the general uncertainty principle which arise from non-commutativity of operators, have to obey additional constraints simply arising from the fact that they are separable. Such a technique, though initially introduced in the context of continuous variable systems [57, 58], has been successfully used to detect entanglement in finite dimensional systems [59–61]. Such a technique is powerful as it detects even bound entanglement [60, 61]. We will discuss these ideas later when we deal with continuous variable entanglement.
1.6 Measures of Entanglement

Since entanglement is a fundamental resource in quantum information science, it is important that one is able to quantify it. Any measure of entanglement has to meet some desirable physical requirements. Some reasonable requirements are enumerated below [37, 41].

(i) For any separable state as in Eq. (1.16), which does not have any entanglement, the measure of entanglement $E$ should be zero:

$$E(\hat{\rho}_{\text{sep}}^A) = 0.$$  \hspace{1cm} (1.79)

(ii) The second requirement concerns the invariance of the entanglement measure under simple local unitary transformations, since such a transformation amounts to a change of basis locally and hence can be undone locally in a deterministic manner. Any local unitary operation should not be able to change the quantity of entanglement shared:

$$E(\hat{\rho}_{AB}) = E(U_A \otimes U_B \hat{\rho}_{AB} U_A^\dagger \otimes U_B^\dagger).$$  \hspace{1cm} (1.80)

(iii) The third requirement states that the average entanglement should not increase under local operations, classical communications, and post-selection (Eq. (1.72)), since if the converse was possible, it would amount to creating entanglement from less entangled states through local operations alone. It is true that through appropriate post-selection, we are able to extract pure state singlets by LOCC as shown in Eq. (1.73), but we are able to do so only at the cost of discarding other outcomes. Thus, on the average entanglement does not increase. Suppose given the state $\hat{\rho}_{AB}$ we get the post-selected states $\hat{\rho}_{AB}^i$ with probability $p_i$, we then require

$$E(\hat{\rho}_{AB}) \geq \sum_i p_i E(\hat{\rho}_{AB}^i).$$  \hspace{1cm} (1.81)

(iv) Finally one would require that the measure of entanglement is additive. Given two entangled bipartite states $\hat{\rho}_1$ and $\hat{\rho}_2$ with the combined system in $\hat{\rho}_1 \otimes \hat{\rho}_2$ we would like to have

$$E(\hat{\rho}_1 \otimes \hat{\rho}_2) = E(\hat{\rho}_1) + E(\hat{\rho}_2).$$  \hspace{1cm} (1.82)

For bipartite pure states there is a unique measure of entanglement. Given a bipartite
pure state in the Schmidt form

$$|\Psi\rangle_{AB} = \sum_i \sqrt{\lambda_i} |\psi_i\rangle_A \otimes |\phi_i\rangle_B,$$  \hspace{1cm} (1.83)

the unique measure of its entanglement is

$$E(|\Psi\rangle_{AB}) = - \sum_i \lambda_i \log_2 \lambda_i,$$  \hspace{1cm} (1.84)

the Shannon entropy of the squares of the Schmidt coefficients. It is the Von Neumann entropy of the reduced density matrix of either subsystem. The uniqueness of the measure follows from the fact that given \(k\) copies of the state \(|\Psi\rangle\) one can distill \(n\) copies of singlets and, conversely, given \(n\) copies of singlets one can produce \(k\) copies of the given state. However, such a reversible transformation is to be understood in the sense of the limit \(k, n \to \infty\) [38]. With this interconvertibility available, we have

$$n E(|\Psi\rangle_{AB}) = k E(|S\rangle_{AB}),$$  \hspace{1cm} (1.85)

where \(|S\rangle_{AB}\) is the singlet state. The interconvertibility implies that

$$E(|\Psi\rangle_{AB}) = \lim_{n,k \to \infty} \frac{k}{n} E(|S\rangle_{AB}) = - \sum_i \lambda_i \log_2 \lambda_i.$$  \hspace{1cm} (1.86)

By convention, one takes \(E(|S\rangle_{AB})\) to be one [37, 62]. This measure which is called the entropy of entanglement or simply entanglement, ranges from zero for a product state to \(\log_2 d\) for maximally entangled states in \(d \otimes d\) dimensions.

### 1.6.1 Entanglement of Formation

Having defined a unique measure of entanglement for bipartite pure states, it is now possible to extend this measure to define a good measure for the case of bipartite mixed states. As indicated earlier, the set of ensembles \(\varepsilon = \{p_i, \hat{\rho}_{AB}^i\}\) realising a given mixed state \(\hat{\rho}_{AB}\) is an infinitely huge family [1]. Keeping this fact in mind, the definition for the entanglement measure for the case of bipartite mixed states is defined as

$$E(\hat{\rho}_{AB}) = \min_{\varepsilon} \sum_i p_i E(\hat{\rho}_{AB}^i).$$  \hspace{1cm} (1.87)

The minimisation has to be carried out with respect to the possible ensemble decompositions. Clearly, any separable state has an entanglement measure zero by definition. The above measure is called the Entanglement of Formation (EOF). The EOF satisfies the
first three requirements required of a good measure. The regularised version of the EOF is defined as [63, 64]

\[ E_c(\hat{\rho}_{AB}) = \lim_{n \to \infty} \frac{E(\hat{\rho}_{AB}^\otimes n)}{n}, \]  

(1.88)

the entanglement cost of \( \hat{\rho}_{AB} \). By definition the entanglement cost of a state is expected to be less than its EOF, since a minimisation over a much larger ensemble set is to be carried out. In case the EOF of a state is additive, then its entanglement cost \( E_c \) is equal to its EOF. It is now known that EOF is in general not additive [65, 66].

The EOF of a state is in general very hard to evaluate. It is thus remarkable that Wootters was able to provide a closed form expression for the EOF of an arbitrary two qubit state [67, 68]. The other cases where closed form expressions have been found are those states with a high degree of symmetry [69–72]. We will come back to this measure when we discuss the EOF for two-mode Gaussian states.

### 1.6.2 Distillable entanglement

This is the amount of pure state entanglement that can be distilled from multiple copies of a given entangled mixed state through LOCC. There are no hard and fast rules as to how one goes about it. Distillable entanglement, as a measure of entanglement, is protocol dependent. Given \( n \) copies of the shared entangled mixed state \( \hat{\rho}_{AB} \), where \( n \) is very large, if one is able to extract \( m \) copies of pure state singlets by LOCC, the distillable entanglement is defined as

\[ D(\hat{\rho}_{AB}) = \lim_{n \to \infty} \frac{m}{n}. \]  

(1.89)

Fundamental prototypes of various protocols were initially introduced in [37]. As indicated earlier, not all entangled states are distillable, thus for bound entangled states distillable entanglement is zero. Distillable entanglement as a measure of entanglement is however useful in the sense that it actually gives an estimate of the useful entanglement one can harvest from multiple copies of a given state through LOCC. This measure being protocol dependent, is not unique. As an example in the case of two qubits, the one way hashing protocol gives a finite yield with

\[ \frac{m}{n} \approx 1 - S(\hat{\rho}_{AB}), \]  

(1.90)

where \( S(\hat{\rho}_{AB}) \) is the Von Neumann entropy of the state \( \hat{\rho}_{AB} \). In comparison, the recursion protocol similar to the one discussed in Section 1.5, gives almost a zero yield [37]. It
is known that any distillable entanglement is a lower bound on the entanglement of formation [37].

1.6.3 Distance based measures

This category of measures arise from the convex structure of state space. We know that the set all states in a given bipartite system form a convex set, so does the set of all separable states of that system. And clearly, the set of all separable states form a subset in the state space. Thus one can talk of distance between the given entangled state and a separable state. Let $S$ be the set of all separable states. The measure of entanglement for a given state $\hat{\sigma}$ is defined as

$$E(\hat{\sigma}) = \min_{\hat{\rho} \in S} D(\hat{\sigma} || \hat{\rho}),$$

(1.91)

where $D$ is any measure of distance (not necessarily a metric) between the two density operators $\hat{\sigma}$ and $\hat{\rho}$, such that it satisfies the requirements put forth above. One measure which satisfies the first three requirements is based on the quantum relative entropy defined as

$$D(\hat{\sigma} || \hat{\rho}) = S(\hat{\sigma} || \hat{\rho}) = \text{Tr} [\hat{\rho} \ln \hat{\sigma} - \ln \hat{\rho}].$$

(1.92)

The quantum relative entropy is not symmetric and does not satisfy the triangle inequality required of a metric. One could have defined the measure with $\hat{\sigma}$ and $\hat{\rho}$ interchanged, however in such a case, pure entangled states have infinite measure of entanglement, which is not desirable. It is known that for pure bipartite states, the relative entropy of entanglement is the Von Neumann reduced entropy [41, 42]. The relative entropy of entanglement has been evaluated for Bell diagonal states, however a closed form expression is not known in the case of general two qubit mixed states. It is known that the relative entropy of entanglement is not additive [71], that it is always less than the entanglement of formation [42], and that it is an upper bound on the distillable entanglement, thus a lower bound on the entanglement of formation [42, 73]. While in the case of pure states one can distill as much entanglement as is present, the mixed state scenario is different. One cannot distill all the entanglement that is present. The existence of bound entanglement is a manifestation of this aspect. This has led to consideration of thermodynamical analogies in the context of mixed state entanglement [42, 62, 74]. Other distance measures such as the Bures metric have also been analogously extended to define entanglement measures [42].
1.6.4 Negativity

Most of the measures of entanglement discussed above are hard to compute for an arbitrary mixed entangled state. Especially, the entanglement of formation and the relative entropy of entanglement require an optimisation procedure, which is in general not easy. Thus it is useful to have a practical measure which is easily calculable. The negativity or the logarithmic negativity is one such measure. The negativity for a bipartite state \( \tilde{\rho}_{AB} \) is defined as

\[
N(\tilde{\rho}_{AB}) = \frac{\|\tilde{\rho}_{AB}^{PT}\|_1 - 1}{2},
\]

where \( \|\cdot\|_1 \) denotes the trace norm. The trace norm for a hermitian operator is the sum of the absolute values of its eigenvalues. Thus \( N(\tilde{\rho}_{AB}) \) is simply the sum of negative eigenvalues of \( \tilde{\rho}_{AB}^{PT} \). Given negativity, we can define yet another measure of entanglement, the logarithmic negativity which is defined as

\[
E_N(\tilde{\rho}_{AB}) = \log_2\|\tilde{\rho}_{AB}^{PT}\|_1.
\]

The logarithmic negativity satisfies the last three of the requirements required of a good measure of entanglement. It is local unitary invariant, does not increase under LOCC, and is additive. However, since it is based on partial transpose, it fails to measure entanglement in bound entangled states [75]. In a similar fashion, one can define calculable measures of entanglement based on negativity based on other positive maps [76, 77].

1.7 Continuous variables

In the earlier Sections, our concern was systems described by finite dimensional Hilbert spaces. We now consider extension of some of these ideas to the case of continuous variable systems. Such a study is necessitated by the very nature of the current available technology. One approach towards practical implementation of the fundamental ideas of quantum information processing has been through the currently available quantum optical technology. Many fundamental aspects of quantum information theory have already been demonstrated in various labs, essentially through quantum optical technology [78–111]. Entanglement or EPR correlation was initially demonstrated in [78, 79]. Teleportation in the continuous variable context was initially discussed in [82] and then demonstrated in [83]. It was subsequently reported in other experiments [85, 86, 88–90, 92, 93]. Continuous variable dense coding has been reported in [96, 103–105]. Entanglement of the polarisation degree of freedom of photons was established in [84–87].
In [85] two qubit entanglement was demonstrated, and in [87] three qubit entanglement. In [86] teleportation of a single photon’s polarisation state was reported. Coherent state based key distribution was reported in [97–100]. Other experiments that report continuous variable entanglement are found in [80, 81, 91, 94, 95, 101, 102, 107, 108, 111]. In [94] atomic ensemble entanglement was reported, and in [95, 101, 111] multipartite entanglement was reported. In [109, 110] teleportation between light and matter was reported.

A fundamental attribute to most of the above mentioned experiments has been the use of nonclassical resources. It is well known that states of radiation, naturally divide into classical and nonclassical types [112], nonclassicality being more fundamental than entanglement in the sense that it is a prerequisite for entanglement [113–115]. A primary reason for the possibility of many experiments in the context of quantum information theory has been the availability of nonclassical resources, primarily in the form of quadrature squeezing [116–135]. It is well known that this kind of nonclassicality, when processed appropriately with the help of passive devices like beamsplitters, can create entanglement, particularly in Gaussian states [113–115, 136–140]. Squeezing has been, and continues to be studied as a resource in the context of continuous variable entanglement.

Bearing in mind this technological scenario, a study of entanglement from the perspective of quantum states of radiation becomes desirable. Any such study requires one to deal with a system of quantum harmonic oscillators, systems whose Hilbert spaces are infinite dimensional. The potential role of such systems in quantum information theory have been explored in [141–163]. In [141] the issue of encoding a qubit in an oscillator was discussed, and the possibility of using coherent states to carry out quantum computational tasks was discussed in [164]. Such a study showed the possibility of using continuous variable set up to do quantum information tasks carried out on finite dimensional systems. In [142] the possibility of key distribution using squeezed states was discussed. In [143, 147, 148] quantum cryptography with Gaussian states was analysed. Multipartite entanglement and its potential use in quantum networks were considered in [144, 145, 150, 154–156]. In [151] universal quantum computation based on continuous variable cluster states using linear optics and homodyne measurement was explored. In [152] the problem of quantum state engineering was considered. In [146, 149, 157–159] continuous variable Gaussian channels were studied. In [161, 162, 165, 166] the possibility of generating entanglement in nanomechanical oscillators was explored; such systems are also modelled as a system of quantum harmonic oscillators. Thus it is natural to undertake the study of entanglement in the context of continuous variables.

The simplest study of entanglement in the context of bipartite continuous variable systems is the study of two-oscillator systems. For this bipartite continuous variable
system \( S = A + B \), subsystems \( A \) and \( B \) consists of a single quantum harmonic oscillator each. The total system \( S \) is specified by respective annihilation and creation operator pairs \( \hat{a}^\dagger, \hat{a} \), and \( \hat{b}^\dagger, \hat{b} \) acting on the Hilbert spaces \( \mathcal{H}^a \) and \( \mathcal{H}^b \). Their only non-vanishing commutators are

\[
[\hat{a}, \hat{a}^\dagger] = [\hat{b}, \hat{b}^\dagger] = \mathbb{I}. \tag{1.95}
\]

The Fock or photon number states for the two modes provide a natural set of ONB’s for \( \mathcal{H}^a \) and \( \mathcal{H}^b \) respectively. They are given as,

\[
|n\rangle = (n!)^{-1/2}(\hat{a}^\dagger)^n|0\rangle_a, \\
|m\rangle = (m!)^{-1/2}(\hat{b}^\dagger)^m|0\rangle_b, \\
n, m = 0, 1, 2, \ldots \tag{1.96}
\]

Then the product states \(|n, m\rangle \equiv |n\rangle_a \otimes |m\rangle_b\) form an ONB for \( \mathcal{H}^a \otimes \mathcal{H}^b \) [167]. This description is easily extended to multipartite continuous variable systems through introduction of additional creation and annihilation operator pairs.

### 1.7.1 Quasi-probabilities

When dealing with quantum mechanics of continuous variables, i.e., radiation field modes, it is useful to go over to their description in the language of quasi-probabilities. A state of a quantum mechanical system described by a density operator \( \hat{\rho} \) can be equivalently described in the language of their \( s \)-ordered quasi-probabilities in a complete manner. The \( s \)-ordered quasi-probabilities capture all the statistical information present in the density operator \( \hat{\rho} \). In this set up a density operator \( \hat{\rho} \) is mapped into a quasi-probability, and the observables to corresponding ordered functions in phase space. Considering a single-mode for simplicity, the \( s \)-ordered quasi-probability corresponding to a state \( \hat{\rho} \) is defined as

\[
W_{\rho}(\alpha, s) = \text{Tr}(\hat{\rho} \hat{T}(\alpha, s)), \quad \text{where},
\]

\[
\hat{T}(\alpha, s) = \frac{1}{\pi} \int D(\xi, s) \exp(\alpha \xi^* - \alpha^* \xi) d^2 \xi, \quad \text{and}
\]

\[
D(\xi, s) = \exp(\hat{a}^\dagger \xi^* - \hat{a} \xi + \frac{1}{2} s |\xi|^2). \tag{1.97}
\]

\( D(\xi, s) \) is the displacement operator corresponding to a given \( s \), and \( \alpha \) is the phase space variable which is denoted as \( \alpha = \frac{q + i p}{\sqrt{2}} \). The parameter \( s \) takes the values \(-1 \leq s \leq 1\).

The three defining properties of a density operator given in Eq. (1.1) transcribe into the
following requirements on the $s$-ordered quasi-probability:

$$\text{Tr}(\hat{\rho}) = 1 \iff \int W_\rho(\alpha, s)d^2\alpha = 1,$$

$$\hat{\rho}_t = \hat{\rho} \iff W_\rho(\alpha, s) = W_\rho^s(\alpha, s),$$

$$\hat{\rho} \geq 0 \iff \text{Tr}(\hat{\rho}\hat{\rho}') \geq 0 \iff \int W_\rho(\alpha, s)W_\rho(\alpha, -s)d^2\alpha \geq 0. \quad (1.98)$$

The first requirement demands that the $s$-ordered quasi-probability is normalised to one with respect to integration over the phase space, while the second demands that it be real. For the third requirement to be satisfied one has to in principle check for positivity of the trace inner product of the given density operator $\hat{\rho}$ with respect to all valid density operators $\hat{\rho}'$, which amounts to checking for positivity of the overlap integral of the given $s$-ordered quasi-probability with respect to all valid $(-s)$-ordered quasi-probabilities. The defining requirements at the density operator level doesn’t enforce pointwise positivity of the $s$-ordered quasi-probability, further, a valid $s$-ordered quasi-probability can be highly singular.

Objects of interest are the expectation values of observables with respect to the given density operator $\hat{\rho}$. Given an observable $\hat{\Omega}$, assume that it possesses an $s$-ordered power series expansion

$$\hat{\Omega} = \sum_{n,m=0}^{\infty} \Omega_{n,m}(s)\{\hat{a}^\dagger^n \hat{a}^m\}_s,$$  \quad (1.99)

where $\{\hat{a}^\dagger^n \hat{a}^m\}_s$ is the $s$-ordered product given by

$$\{\hat{a}^\dagger^n \hat{a}^m\}_s = \frac{1}{\pi} \int \hat{T}(\alpha, s)\alpha^n\alpha^m d^2\alpha,$$  \quad (1.100)

then the expectation value $\langle \hat{\Omega} \rangle$ is given by

$$\text{Tr}(\hat{\rho}\hat{\Omega}) = \int \Omega(\alpha, -s)W_\rho(\alpha, s)d^2\alpha, \quad \text{where}$$

$$\Omega(\alpha, -s) = \sum_{n,m=0}^{\infty} \Omega_{n,m}(-s)\alpha^n\alpha^m. \quad (1.101)$$

Evaluating expectation values of $s$-ordered operators with respect to a density operator $\hat{\rho}$ corresponds to evaluating the function $\Omega$ which is obtained from $\hat{\Omega}$ simply by replacing $\{\hat{a}^\dagger^n \hat{a}^m\}_s$ by $\alpha^n\alpha^m$, and then evaluating its overlap integral with the $s$-ordered quasi-probability $W_\rho(\alpha, s)$ corresponding to the state $\hat{\rho}$. Conversely, any observable $\hat{\Omega}$ can be
written as
\[
\hat{\Omega} = \frac{1}{\pi} \int \Omega(\alpha, s) \hat{T}(\alpha, s) d^2 \alpha.
\] (1.102)

The \( s \)-ordered quasi-probabilities enables us to recast quantum mechanics in the language of statistical mechanics. Given an observable \( \hat{\Omega} \), we have an \( s \)-ordered weight function \( \Omega(\alpha, -s) \), which captures the content of \( \hat{\Omega} \) for each \(-1 \leq s \leq 1\).

The more familiar \( s \)-ordered quasi-probabilities are Sudarshan’s \( \phi \) function corresponding to \( s = 1 \), the Wigner function corresponding to \( s = 0 \), and the Husimi or \( Q \) function corresponding to \( s = -1 \). The notion of \( s \)-ordered quasi-probability can be easily extended to the multi-mode case as well [167-169].

### 1.7.2 Sudarshan’s \( \phi \) function

It turns out that any density operator corresponding to a single mode of radiation field can always be expanded as
\[
\hat{\rho} = \int \frac{d^2z}{\pi} \phi_\rho(z) |z\rangle \langle z|,
\] (1.103)

where \( \phi_\rho(z) = W_\rho(z, 1) \), and \(|z\rangle\) is the coherent state. This is called as the diagonal representation. It is remarkable that every density operator can be expanded diagonally in the coherent state basis. This has been possible because of the over completeness of the coherent state basis [112]. This representation enables one to evaluate ensemble averages of normally ordered operators. Normal ordering corresponds to shifting all the creation operators to the left and all the annihilation operators to the right. The useful concept that arises out of this representation is the notions of classicality and nonclassicality. If \( \phi_\rho(z) \) corresponding to a density operator \( \hat{\rho} \) is pointwise nonnegative in the complex plane, then the density operator is a convex combination of coherent states. Since the coherent states are the most elementary of all quantum mechanical states exhibiting classical behaviour, any state that can be written as a convex combination of these elementary states is deemed classical. We have,
\[
\phi_\rho(z) \geq 0 \text{ for all } z \in \mathcal{C} \Rightarrow \hat{\rho} \text{ is classical.}
\] (1.104)

Any state which cannot be written so is deemed nonclassical. This definition is readily extended to the multi-mode case. Not all states are classical, and in fact the \( \phi \) function can be highly singular. For the Fock states \(|n\rangle\), the diagonal function \( \phi_\rho(z) \) turns out to be \( n \)-th derivative of the delta function. We will make extensive use of nonclassicality in
the context of entanglement in bipartite continuous variable systems.

1.7.3 Nonclassicality

One of the primary concerns in quantum optics over the years has been nonclassicality. This notion has been explored in various ways and its manifestations detailed [170–192]. Physical manifestations such as photon number oscillations have been explored in [170, 175, 178, 181, 183, 193]. More recently, nonclassicality has been viewed as a resource for entanglement [113–115, 137–139, 171, 194].

An important concern in the study of nonclassicality is in respect of the methods employed to detect nonclassicality. We briefly describe one of the techniques. Suppose we have the operator

\[
F(\hat{a}^\dagger, \hat{a}) = \sum_{ij} c_{ij} \hat{a}^j \hat{a}^i. \tag{1.105}
\]

Define the positive operator \( F^\dagger F \) and normal order it to obtain \( :F^\dagger F : \). Then the expectation value of this normal ordered operator in a state \( \hat{\rho} \) is

\[
\langle : F^\dagger F : \rangle = \text{Tr}(\hat{\rho} : F^\dagger F :) = \int \frac{d^2z}{\pi} \phi_{\rho}(z)|F(z^*, z)|^2. \tag{1.106}
\]

Calculating \( \langle : F^\dagger F : \rangle \) is equivalent to evaluating the phase space average of a pointwise positive function \( |F(z^*, z)|^2 \) with respect to \( \phi_{\rho}(z) \). Thus if the expectation value of a positive normal ordered operator with respect to a state turns out to be negative, then we can surely conclude that the state was nonclassical. However if the expectation value turns out to be positive, we cannot conclude that the state was classical. In principle, one has to run over all possible positive functions over the phase space to test if a state is classical or not.

The method outlined above has been discussed and used in [170, 182, 186, 189–191]. The idea can be further extended through the use of positive polynomials which are not sum of squares of other polynomials [191]. It is useful to note that there is no conclusive test for nonclassicality, of an arbitrary mixed state. There are however two exceptions, namely single-mode Gaussians [185] and single-mode phase invariant states [170].

Another important aspect in the study of nonclassicality has been its quantification. Measures of nonclassicality have been discussed in [171, 173, 174, 177, 188, 195–198]. Distance based measures were discussed in [173, 188, 195, 196], where the measure of nonclassicality was defined as the least ‘distance’ of a nonclassical state with respect to the set of all classical states. The distance based measures are in general hard to evaluate.
because of the inherent minimisation procedure involved in the definition. In [174, 177] a measure based on the smoothening of the \( \phi \) function was defined. One can go from the \( \phi \) function to any \( s \)-ordered quasi-probability through a Gaussian convolution parametrised by \( \tau \). The point \( \tau \), at which the resultant quasi-probability is just about a true probability gives us a measure of nonclassicality. In [197, 198] a measure of nonclassicality based on Hudson’s theorem was defined. Hudson’s theorem says that the only pure state whose Wigner function is positive is a Gaussian state [199]. This fact indicates that any measure that captures the negativity of the Wigner function is a measure of nonclassicality of the state. More recently, in [171] a measure of nonclassicality inspired from an entanglement point of view was defined. Nonclassicality of a single-mode state was defined as the maximal bipartite entanglement it can produce when coupled with additional classical resources and passed through a 50:50 beamsplitter.

1.7.4 The Wigner function

The Wigner function is the \( s \)-ordered quasi-probability corresponding to \( s = 0 \). It is obtained by convoluting the \( \phi \) with a particular Gaussian weight function. Thus, in principle, it is not as singular as the \( \phi \) function, but nevertheless it can be pointwise non-positive in phase space. In this description, density operators are put in correspondence with real valued functions over the phase space through the rule

\[
W_\rho(q, p) = \frac{1}{\pi} \int dq' \langle q' | \hat{\rho} | q \rangle \exp(2iqp).
\]  

(1.107)

We recall that \( \alpha = \frac{q + ip}{\sqrt{2}} \). We may write \( W_\rho(q, p) \) as \( W_\rho(\xi) \) for convenience [200, 201], where \( \xi \) is the pair \( \{ q, p \} \). The Wigner description offers several advantages, from the theoretical and experimental perspectives. It turns out to be convenient in the context of unitary evolution of states under the action quadratic hamiltonians. Given the unitary operator \( U(S) \) corresponding to a quadratic hamiltonian which is captured by a symplectic group element \( S \in Sp(2n, R) \), a state evolves as

\[
\rho \rightarrow \rho' = U(S)\rho U(S)\dagger.
\]  

(1.108)

This transformation takes a very simple form in the Wigner description:

\[
S: \; \rho \rightarrow \rho' = U(S)\rho U(S)\dagger \Leftrightarrow W_\rho(\xi) \rightarrow W_{\rho'}(\xi) = W_\rho(S^{-1}\xi).
\]  

(1.109)

That is, \( W_\rho(S\xi) = W_\rho(\xi) \) for every canonical transformation \( S \in Sp(2n, R) \). That is, the Wigner function transforms as a \( Sp(2n, R) \) scalar field [202]. This is also true
of inhomogeneous linear canonical transformations involving phase space translations as well.

The Wigner function is also important from an experimental perspective, as it can be directly measured in a lab. The Wigner function by definition, is the expectation value of the displaced parity operator with respect to the given density operator [169], and thus can be observed in a laboratory. Indeed, the Wigner function of various quantum states of radiation have been reported [203–215].

1.7.5 The Q function

The $Q$ function corresponds to the quasi-probability with order parameter $s = -1$ For a state $\hat{\rho}$ of a single-mode of radiation field it is defined as

$$Q_\rho(\alpha) = \langle \alpha | \hat{\rho} | \alpha \rangle.$$ (1.110)

It is thus pointwise nonnegative in the complex plane, being the expectation value of a coherent state on a positive semidefinite density operator $\hat{\rho}$. Thus it is a true probability distribution. However it should be noted that not all valid probability distributions are valid $Q$ functions. The quantum mechanical uncertainty principle places severe restrictions on probability distributions which qualify to be valid $Q$ functions. The advantage of working with them is that one will deal only with true probabilities. We will consider them in more detail in Chapter 7 where we discuss non-Gaussianity [216].

1.8 Detection of Entanglement

Detecting entanglement in continuous variables has been an important pursuit, as methods employed in the case of finite dimensions cannot always be extended in a naive manner. Among tests based on positive maps, only the partial transpose test and reduction criterion have been extended [57, 217]. However, inseparability criteria (inequalities) defined through EPR-like operators based on the quadratic moments, initially introduced in [58], are effective in detecting entanglement in continuous variables. These inequalities are derived from first principles as consequences of separability. The uncertainty principle places certain restrictions on the moments, and all states have to obey the uncertainty principle [218]. However, separable states have to obey further constraints, which can be taken as inseparability criteria. Simple examples of such constraints are the Bell type inequalities, violation of which have been observed in entangled Gaussian states [219, 220].
1.8.1 Bell’s inequalities

The Bell type inequalities as in Eq. (1.36) are extended to the continuous variable case through the definition of the parity operator. The parity operator takes the role of the spin observable, and the role of ‘direction’ of the spin observable is taken by displacement in phase space. The parity operator $\Pi(\xi)$ for $N$ modes is given as

$$\Pi(\xi) = \bigotimes_{i=1}^{N} \Pi_i(\xi_i) = \bigotimes_{i=1}^{N} D_i(\xi_i)(-1)^{\hat{n}_i}D_i^\dagger(\xi_i),$$

(1.111)

where $D_i(\xi_i)$ is the phase space displacement operator of the $i^{th}$ mode, defined in Eq. (1.97). It happens that the expectation value of the parity operator $\Pi(\xi)$ with a given state $\hat{\rho}$ gives its Wigner function [168, 169]:

$$W_\rho(\xi) = \text{Tr}(\hat{\rho}\Pi(\xi)).$$

(1.112)

Hence, the expectation value of the parity operator on simple states such as Gaussian states is easily evaluated. In [219] the Bell operator

$$\mathcal{B} = \Pi(0, 0) + \Pi(0, \beta) + \Pi(\alpha, 0) - \Pi(\alpha, \beta)$$

(1.113)

was considered, where $\alpha$, and $\beta$ are phase space displacements on the two modes. Separable continuous variable states have to satisfy the inequality $|\mathcal{B}| \leq 2$. It was demonstrated in [219] that the two-mode squeezed vacuum clearly violated this inequality. This idea was further generalised to the multi-mode case in [221].

Another equivalent approach towards Bell inequalities in continuous variable systems was given in [220]. This was done through the extension of the spin operators to infinite dimensions through a direct sum of the spin-$\frac{1}{2}$ Pauli operators. In this set up, the Pauli spin operators $\hat{\sigma}_z$, $\hat{\sigma}^+$, and $\hat{\sigma}^-$ are given by

$$\hat{\sigma}_z = (-1)^{\hat{n}}, \quad \hat{\sigma}^+ = \hat{\sigma}^- = \sum_{n=0}^{\infty} |2n\rangle\langle 2n+1|.$$  

(1.114)

It is easy to see that the pseudo-spin operators defined above, satisfy the $SU(2)$ algebra required of spin-$\frac{1}{2}$ operators. It is now easy to extend the Bell-CHSH inequalities in Eq. (1.36) to the continuous variable case. In [220] such an extension was done, and the entanglement in two-mode squeezed vacuum was demonstrated. The role of Bell’s inequalities in detecting entanglement is also discussed in [222–225].
1.8.2 Partial transpose

We have already seen in Eq. (1.109) that the Wigner function transforms as a scalar field under the action of symplectic group elements corresponding to evolution under quadratic hamiltonians. The transpose map and the partial transpose map also take a geometric form in the Wigner description. For a single mode of radiation field, it follows from the definition of the Wigner function that transpose operation on the density operator is equivalent to complex conjugation of the elements of the density matrix in position representation, which transcribes faithfully into momentum reversal operation in the Wigner description:

\[ T : W_\rho(q, p) \rightarrow W'_\rho(q, p) = W_\rho(q, -p) = W_\rho(\Lambda \xi), \]

\[ \Lambda = \text{diag}(1, -1) = \sigma_3. \] (1.115)

This amounts to a mirror reflection which inverts the \( p \) coordinate, leaving the \( q \) coordinate unchanged. The transpose map takes density operators to density operators, but is unphysical as seen in Section 1.4.6. For a bipartite system of two modes of radiation field, partial transposition amounts to inverting the \( p \) coordinate for one of the subsystems. Its action on the Wigner function is given by

\[ \text{PT} : W_\rho(q_a, p_a, q_b, p_b) \rightarrow W'_\rho(q_a, p_a, q_b, p_b) = W_\rho(q_a, p_a, q_b, -p_b). \] (1.116)

Though the partial transpose, as a map, is given by an elegant transformation with regard to the Wigner function, its usefulness as an entanglement witness is manifest only when we are able to test a phase space distribution for its Wigner quality. That is, one has to answer the question as to when is a function in phase space a Wigner function. Such a question was initially raised in the limited context of Gaussian Wigner functions in [201], and fully answered in [226] in that context. This was possible because of the Williamson's theorem [226]. This consequently led to the demonstration of partial transpose as an effective entanglement criterion in the case of two-mode Gaussians [57]. In general, one can test for entanglement in a limited manner through the manifestation of partial transpose on moments.

Given a state \( \hat{\rho}^{(ab)} \) of a bipartite system of two modes of radiation field, a test for entanglement in \( \hat{\rho}^{(ab)} \) through partial transpose would be to test for negativity of \( \hat{\rho}^{(ab)PT} \). A direct approach would be to test for violation of positivity of the diagonals of \( \hat{\rho}^{(ab)PT} \) in some orthogonal basis pertaining to the composite system Hilbert space \( \mathcal{H}_a \otimes \mathcal{H}_b \). Such an effort may be tedious while one is dealing with continuous variable systems. A simpler
practical approach, but limited in scope, would be to test for violation of positivity of $\hat{\rho}^{(ab)PT}$ by taking its expectation value with a positive operator. Such a test may not in principle capture the negativity of $\hat{\rho}^{(ab)PT}$ in its entirety, but may prove to be useful in certain contexts. An immediate requirement of the procedure demands the systematic construction of positive operators acting on $\mathcal{H}^a \otimes \mathcal{H}^b$. Given an operator

$$\eta = \sum_{jklm} c_{jklm} \hat{a}^j \hat{a}^k \hat{b}^l \hat{b}^m,$$

(1.117)

the operator $\eta^{\dagger} \eta$ is positive by construction. A simple test for violation of positivity of $\hat{\rho}^{(ab)PT}$ is to then check if

$$\text{Tr}(\hat{\rho}^{(ab)PT}(\eta^{\dagger} \eta)) = \text{Tr}(\hat{\rho}^{(ab)}(\eta^{\dagger} \eta)^{PT}) < 0.$$

(1.118)

To this end, we need to know how partial transpose acts on an ordered monomial $(\hat{a}^j \hat{a}^k \hat{b}^l \hat{b}^m \hat{b}^r \hat{b}^s)$. Since the creation and annihilation operators $\hat{b}^\dagger$ and $\hat{b}$ are real, transposition on Bob's side alone amounts to hermitian conjugation of operators on Bob's side. We have

$$(\hat{a}^j \hat{a}^k \hat{b}^l \hat{b}^m \hat{b}^r \hat{b}^s)^{PT} = (\hat{a}^j \hat{a}^k \hat{b}^l \hat{b}^m \hat{b}^r \hat{b}^s)^{PT} = (\hat{a}^j \hat{a}^k \hat{b}^l \hat{b}^m \hat{b}^r \hat{b}^s)^{PT} = (\hat{a}^j \hat{a}^k \hat{b}^l \hat{b}^m \hat{b}^r \hat{b}^s)^{PT}.$$

(1.119)

As a simple example demonstrating the effectiveness of the procedure occurs in the case of two-mode Gaussians. The violation of partial transpose at the level of variance matrix corresponding to $\eta^{\dagger} \eta$ being quadratic in the annihilation and creation operators of the two modes, turns out to be both necessary and sufficient test for entanglement [57].

We shall see later how this procedure can be effective in more general contexts [227–229].

1.8.3 Inseparability criteria through uncertainty relations

This method of detecting entanglement in continuous variable systems is based on the fact that expectation values of nonlocal operators pertaining to the composite system have to obey additional constraints for separable states in addition to the usual uncertainty principle. As a simple example, consider the pair of EPR like operators $\hat{x}_a - \hat{x}_b$ and $\hat{p}_a + \hat{p}_b$. The sum of the variances of these two operators goes to zero for maximally entangled states. Such a state may appear unphysical though, but for the two-mode squeezed vacuum the total variance rapidly tends to zero with increasing degree of squeezing. And this state approaches the maximally entangled state as the squeeze parameter grows.
However for a separable state, assuming that the first moments were zero, we have

\[
\begin{align*}
\langle (\hat{x}_a - \hat{x}_b)^2 \rangle + \langle (\hat{p}_a + \hat{p}_b)^2 \rangle &= \\
\sum_i p_i \langle \hat{x}_a^2 \rangle_i + \sum_i p_i \langle \hat{x}_b^2 \rangle_i - 2 \sum_i p_i \langle \hat{x}_a \rangle_i \langle \hat{x}_b \rangle_i + \\
&+ \sum_i p_i \langle \hat{p}_a^2 \rangle_i + \sum_i p_i \langle \hat{p}_b^2 \rangle_i + 2 \sum_i p_i \langle \hat{p}_a \rangle_i \langle \hat{p}_b \rangle_i \\
&= \sum_i p_i \langle (\hat{x}_a^2) \rangle_i - \langle (\hat{x}_a) \rangle_i^2 + \sum_i p_i \langle (\hat{p}_a^2) \rangle_i - \langle (\hat{p}_a) \rangle_i^2 \\
&+ \sum_i p_i \langle (\hat{x}_b^2) \rangle_i - \langle (\hat{x}_b) \rangle_i^2 + \sum_i p_i \langle (\hat{p}_b^2) \rangle_i - \langle (\hat{p}_b) \rangle_i^2 + \\
&\sum_i p_i \langle (\hat{x}_a) \rangle_i - \langle (\hat{x}_b) \rangle_i^2 + \sum_i p_i \langle (\hat{p}_a) \rangle_i + \langle (\hat{p}_b) \rangle_i^2 \geq 2. & (1.120)
\end{align*}
\]

Such a criterion was effectively used in [58, 230] to demonstrate entanglement in two-mode Gaussian states. The EPR uncertainty relation in Eq. (1.120) also becomes useful in evaluating the EOF of two-mode symmetric Gaussian states. This method was extended to the multi-mode scenario in [231]. Similar techniques, based on the uncertainties on moments were used in the case of finite dimensional systems to detect entanglement [59–61], and later extended to continuous variable systems in [232].

**Positive maps on matrix of moments:** This technique was recently introduced in [233]. It enables us to apply the positive maps familiar from the context of finite dimensional systems to detect entanglement in continuous variable systems through the matrix of moments. Any moment matrix generated by tensoring operators belonging to the individual systems is separable on separable states, thus inseparability of the moment matrix implies entanglement for the state. As a simple example, for two pairs of operators \(\hat{f}_1, \hat{f}_2\) on Alice’s side, and \(\hat{g}_1, \hat{g}_2\) on Bob’s side, the matrix of moments formed by the tensored set \((\hat{f}_1 \hat{g}_1, \hat{f}_1 \hat{g}_2, \hat{f}_2 \hat{g}_1, \hat{f}_2 \hat{g}_2)\), i.e.,

\[
M = \begin{pmatrix}
\langle \hat{f}_1^2 \hat{g}_1^2 \rangle & \langle \hat{f}_1^2 \hat{g}_1 \hat{g}_2 \rangle & \langle \hat{f}_1 \hat{f}_2 \hat{g}_1^2 \rangle & \langle \hat{f}_1 \hat{f}_2 \hat{g}_1 \hat{g}_2 \rangle \\
\langle \hat{f}_1^2 \hat{g}_1 \hat{g}_2 \rangle & \langle \hat{f}_1^2 \hat{g}_2^2 \rangle & \langle \hat{f}_1 \hat{f}_2 \hat{g}_2^2 \rangle & \langle \hat{f}_1 \hat{f}_2 \hat{g}_1 \hat{g}_2 \rangle \\
\langle \hat{f}_1 \hat{f}_1 \hat{g}_1^2 \rangle & \langle \hat{f}_1 \hat{f}_1 \hat{g}_1 \hat{g}_2 \rangle & \langle \hat{f}_1 \hat{f}_2 \hat{g}_2^2 \rangle & \langle \hat{f}_1 \hat{f}_2 \hat{g}_1 \hat{g}_2 \rangle \\
\langle \hat{f}_1 \hat{f}_2 \hat{g}_2^2 \rangle & \langle \hat{f}_2 \hat{g}_1 \hat{g}_2 \rangle & \langle \hat{f}_2 \hat{g}_2 \hat{g}_1 \rangle & \langle \hat{f}_2 \hat{g}_2 \hat{g}_2 \rangle \\
\end{pmatrix}, & (1.121)
\]

is separable for separable states. This method enables us to indirectly (directly) use the theory of positive maps to detect entanglement in continuous variable systems.
1.9 Gaussian states

Gaussian states form an integral part of quantum information theory in the context of continuous variables. Initially they were studied in the context of radiation fields in quantum optics [185, 201, 202, 226, 234–240]. Much of the study was devoted to characterising them through the variance matrix and the exploration of nonclassicality of these states which was primarily in the form of squeezing. But with the development of quantum information theory the focus has shifted to the study of entanglement in these states [57, 58, 70, 136, 138, 140, 143, 144, 147–150, 156–159, 217, 241–271]. Teleportation, the fundamental protocol of quantum information theory has been achieved using them [83, 88–90, 96, 97, 103]. This has lead to the exploration of Gaussian states in a major way.

We review, very briefly, some of these developments. The potential role of Gaussian states in quantum information processes was initially realised in [82] in the form of teleportation. The entanglement involved in this protocol was first pointed out in [105]. Subsequently, the detection of entanglement of these states was completely characterised in [58, 272]. The effective use of the variance matrix formalism in the context of bipartite entanglement was initially carried out in [57], and this was possible because Gaussian states are completely specified by their variance matrix. The primary tool is the effective use of uncertainty principle in characterising them [192, 202, 218, 226]. The Williamson theorem has a fundamental role to play in this context [192, 202, 226, 272, 273].

In [241] a family of bound entangled Gaussian states was constructed, and in [217] the issue of distillability was solved: it was shown that every NPT Gaussian state is distillable. Purification of Gaussian states has been discussed in [259, 268, 271]. In [242], the issue of bipartite separability in the multi-mode case was completely solved. The various possible situations that could arise with regard to separability in tripartite systems was discussed in [243]. Entanglement in multipartite Gaussians was studied in [249, 250]. It was found that in special cases Gaussian multipartite entanglement can be reduced to two-mode Gaussian entanglements using local operations alone.

The role of squeezing in generating entanglement in these states was discussed in [136–138, 140], and some of the special properties of Gaussian states were discussed in [244, 246, 252, 253]. In [244, 246] it was found that Gaussian states cannot be distilled with Gaussian operations alone, and in [252] it was shown that the optimal cloning of Gaussian states required additional non-Gaussian resources. Some of the extremal properties of Gaussian states were discussed in [253]. One particular such extremal property helped solve the problem of determining the entanglement of formation in symmetric Gaussian states [70]. We will have more to say on this later in the thesis. Lower bounds
on the entanglement of formation of general bipartite Gaussian states was obtained in [262].

The determination of entanglement in bipartite Gaussian states through purity measurements was discussed in [248, 261, 263]. Such a study brought out the possibility of characterising Gaussian states through both global and local purities. Properties such as the loss of entanglement in evolution was discussed in [137, 260, 274].

Gaussian channels have been discussed in [149, 157, 269]. A definition for the most general Gaussian channel was given in [157]. The quantum channel capacity defined through coherent information was evaluated for a certain class of Gaussian channels in [158, 159].

The possibility of assessing the entanglement in two-mode Gaussian states using local parity measurements and classical communication was discussed in [247]. It was shown that given sufficiently large number of copies of a Gaussian state, its entanglement and the state itself can be characterised completely by LOCC!

The concept of entanglement monogamy [275, 276] was extended to the Gaussian case in [254, 255], and monogamy relations were established. In [251] the variance matrix set up was discussed from the perspective of convex sets, and numerical routines were set up to generate entanglement witnesses. More detailed reviews on Gaussian states can be found in [256–258].

We now describe the basic formalism for handling Gaussian states and their transformation, with particular attention to nonclassicality and entanglement.

1.9.1 Nonclassicality in Gaussians

Nonclassicality in Gaussian states primarily occurs in the form of quadrature squeezing. This quantum optical concept was initially explored in [185, 192, 237–240, 277, 278]. An elegant definition for squeezing through the variance matrix formalism was given in [192]. We briefly review this work.

Since zero-mean Gaussian states are completely specified by their variance matrix, we begin by giving the basic setup of variance matrix, first from the perspective of nonclassicality, and later we give the description from the entanglement perspective.

Consider an $n$ mode quantum system with annihilation operators $\hat{a}_j, \hat{a}_j^\dagger, j = 1, 2, \ldots, n$, obeying the standard boson commutation relations

$$[\hat{a}_j, \hat{a}_k^\dagger] = \delta_{jk}, \quad [\hat{a}_j, \hat{a}_k] = [\hat{a}_j^\dagger, \hat{a}_k^\dagger] = 0,$$

(1.122)
or equivalently

\[
\begin{align*}
[\hat{q}_j, \hat{p}_k] &= i\delta_{jk}, \\
[\hat{q}_j, \hat{q}_k] &= [\hat{p}_j, \hat{p}_k] = 0, \\
\hat{a}_j &= \frac{\hat{q}_j + i\hat{p}_j}{\sqrt{2}}, \\
\hat{a}_j^\dagger &= \frac{\hat{q}_j - i\hat{p}_j}{\sqrt{2}}.
\end{align*}
\] (1.123)

It is convenient to arrange the hermitian \(\hat{q}_j, \hat{p}_j\) and the \(\hat{a}_j, \hat{a}_j^\dagger\) in \(2n\) component column vector forms:

\[
\hat{\xi}^{(r)} = \begin{pmatrix}
\hat{q}_1 \\
\vdots \\
\hat{q}_n \\
\hat{p}_1 \\
\vdots \\
\hat{p}_n
\end{pmatrix}, \quad \hat{\xi}^{(c)} = \begin{pmatrix}
\hat{a}_1 \\
\vdots \\
\hat{a}_n \\
\hat{a}_1^\dagger \\
\vdots \\
\hat{a}_n^\dagger
\end{pmatrix}.
\] (1.124)

The ‘vectors’ \(\hat{\xi}^{(c)}\) and \(\hat{\xi}^{(r)}\) are related by a fixed numerical unitary matrix \(\Gamma\)

\[
\hat{\xi}^{(c)} = \Gamma \hat{\xi}^{(r)}, \quad \text{where} \quad \Gamma = \frac{1}{\sqrt{2}} \begin{pmatrix}
\mathbb{I} & i\mathbb{I} \\
\mathbb{I} & -i\mathbb{I}
\end{pmatrix}.
\] (1.125)

The canonical commutation relations among the mode operators can now be succinctly written as

\[
\begin{align*}
[\hat{\xi}^{(r)}_j, \hat{\xi}^{(r)}_k] &= i\beta_{jk}, \\
[\hat{\xi}^{(c)}_j, \hat{\xi}^{(c)}_k] &= i\Sigma_{jk}, \\
i, j &= 1, 2, \ldots, 2n,
\end{align*}
\] (1.126)

where the \(2n \times 2n\) dimensional matrices \(\beta\) and \(\Sigma\) are given in block form by

\[
\beta = \begin{pmatrix}
0 & \mathbb{I} \\
-\mathbb{I} & 0
\end{pmatrix}, \quad \Sigma = \begin{pmatrix}
\mathbb{I} & 0 \\
0 & -\mathbb{I}
\end{pmatrix}.
\] (1.127)
We define the $2n \times 2n$ real variance matrix $V^{(r)}$ for a state $\hat{\rho}$ by

$$V^{(r)}_{jk} = \frac{1}{2} \text{Tr}(\hat{\rho}(\hat{\xi}_j \hat{\xi}^{(r)T}_k)), \quad (1.128)$$

where $\{.,.\}$ is the anticommutator. We can alternatively arrive at the variance matrix by taking the expectation value of the positive operator $\hat{\xi}^{(r)} \hat{\xi}^{(r)T}$, i.e.,

$$\langle \hat{\xi}^{(r)} \hat{\xi}^{(r)T} \rangle = \text{Tr}(\hat{\rho} \hat{\xi}^{(r)} \hat{\xi}^{(r)T}) = V^{(r)} + \frac{i}{2} \beta.$$

We can write the variance matrix in an $n \times n$ block form as

$$V^{(r)} = \begin{pmatrix} V_1 & V_2 \\ V_2^T & V_3 \end{pmatrix},$$

$$(V_1)_{jk} = \langle \hat{q}_j \hat{q}_k \rangle, \quad (V_2)_{jk} = \frac{1}{2} \langle \{\hat{q}_j, \hat{p}_k\} \rangle, \quad (V_3)_{jk} = \langle \hat{p}_j \hat{p}_k \rangle,$$

$$j, k = 1, 2, \ldots, n. \quad (1.129)$$

The matrix $V_1$ gives the correlations among the $\hat{q}$’s, $V_3$ gives those among the $\hat{p}$’s, and $V_2$ the correlations between the $\hat{q}$’s and $\hat{p}$’s. For a state $\hat{\rho}$ with non zero mean ($\langle \hat{\xi}^{(r)} \rangle \neq 0$, the variance matrix is defined by simply replacing $\hat{\xi}^{(r)}$ by $\Delta \hat{\xi}^{(r)} = \hat{\xi}^{(r)} - \langle \hat{\xi}^{(r)} \rangle$. Such a replacement corresponds to a rigid translation in phase space by amount $-\langle \hat{\xi}^{(r)} \rangle$, implemented by the displacement operator $D(-\langle \hat{\xi}^{(r)} \rangle)$.

The complex form of the variance matrix $V^{(c)}$ is generated by taking the expectation value of the positive operator $\hat{\xi}^{(c)} \hat{\xi}^{(c)†}$, i.e.,

$$\langle \hat{\xi}^{(c)} \hat{\xi}^{(c)†} \rangle = V^{(c)} + \frac{1}{2} \Sigma, \quad \text{where} \quad V^{(c)} = \Gamma V^{(r)} \Gamma^†. \quad (1.130)$$

Writing in $n \times n$ block form, we have

$$V^{(c)} = \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix},$$

$$A_{jk} = A_{kj}^* = \frac{1}{2} \langle \{\hat{a}_j, \hat{a}_k^†\} \rangle, \quad B_{jk} = B_{kj} = \langle \hat{a}_j \hat{a}_k \rangle. \quad (1.131)$$

By definition, $V^{(r)}$ is symmetric positive definite, this also implies that $V^{(c)}$ is complex hermitian positive definite. It also implies that $A$ is hermitian, and $B$ is symmetric. We have the following relations between the blocks of the variance matrix in its real and
complex forms:
\[
A = \frac{1}{2} (V_1 + V_3 + i(V_2^T - V_2)),
\]
\[
B = \frac{1}{2} (V_1 - V_3 + i(V_2^T + V_2));
\]
\[
V_1 = \frac{1}{2} (A + A^* + B + B^*),
\]
\[
V_2 = \frac{i}{2} (A - A^* - B + B^*),
\]
\[
V_3 = \frac{1}{2} (A + A^* - B - B^*). \tag{1.132}
\]

**Uncertainty principle**: Given a real symmetric positive definite \(2n \times 2n\) matrix \(V^{(r)}\), it has to satisfy additional requirements to qualify as a variance matrix derived from a quantum mechanical state. The additional requirements are succinctly captured by the uncertainty principle [192]:

\[
V^{(r)} + \frac{i}{2} \beta \geq 0. \tag{1.133}
\]

The above condition is both necessary and sufficient to validate a given symmetric positive definite matrix \(V^{(r)}\) as the variance matrix of some quantum state. The necessity of the condition follows from definition but the sufficiency part, which is nontrivial, follows from the use of Williamson theorem [272, 273]. In the case of Gaussian probability distributions this uncertainty principle on the variance matrix is both necessary and sufficient condition to qualify the probability as a Wigner distribution[226], since in this case the variance matrix completely specifies the quantum state. This is no longer true in more general contexts [279].

**Unitary evolution**: Unitary evolution of the mode operators under a quadratic hamiltonian corresponds to a symplectic transformation on the column vector \(\hat{\xi}^{(r)}\):

\[
U = \exp(-iH) \Rightarrow U^\dagger \hat{\xi}^{(r)} U = \hat{\xi}^{(r)} H, \quad \text{where} \\
H = \sum_{j,k} h^{(r)}_{jk} \hat{\xi}_j \hat{\xi}_k, \quad \text{and} \\
S^{(r)}_H \in Sp(2n, R), \text{i.e., } S^{(r)}_H \beta S^{(r)}_H^T = \beta. \tag{1.134}
\]

Under such a unitary evolution, the vector \(\hat{\xi}^{(c)}\) transforms as

\[
\hat{\xi}^{(c)} \rightarrow \hat{\xi}^{(c)} = S^{(c)}_H \hat{\xi}^{(c)}, \quad S^{(c)}_H = \Gamma S^{(r)}_H \Gamma^\dagger. \tag{1.135}
\]
Given a state $\hat{\rho}$ with variance matrix $V^{(r)}$, under time evolution by a quadratic hamiltonian the variance matrix evolves through a congruence by the symplectic transformation corresponding to the hamiltonian:

$$
\hat{\rho} \rightarrow \hat{\rho}' = U(S^{(r)})\hat{\rho}U(S^{(r)})\dagger \Rightarrow 
V^{(r)} \rightarrow V'^{(r)} = S^{(r)}V^{(r)}S^{(r)\dagger}, \quad S^{(r)} \in Sp(2n, R). \tag{1.136}
$$

We have removed the subscript $H$ for brevity. Similarly, the complex form of the variance matrix evolves as

$$
V^{(c)} \rightarrow V'^{(c)} = S^{(c)}V^{(c)}S^{(c)\dagger}, \quad S^{(c)} = \Gamma S^{(r)}\Gamma\dagger. \tag{1.137}
$$

**The symplectic group:** The defining property of matrices comprising the symplectic group $Sp(2n, R)$ is

$$
S^{(r)} \in Sp(2n, R) \Leftrightarrow S^{(r)\beta S^{(r)\dagger}} = \beta. \tag{1.138}
$$

The group is characterised by $n(2n + 1)$ generators, of which $n^2$ generators are compact generators and total photon number. The remaining $n(n + 1)$ generators are noncompact, and do not conserve the total photon number. The hermitian quadratic hamiltonians which correspond to the compact generators can be taken to be

$$
\begin{align*}
\frac{1}{4}(\hat{a}^+_j\hat{a}_j + \hat{a}^+_j\hat{a}_j), & \quad j = 1, 2, \ldots, n; \\
\frac{1}{4}(\hat{a}^+_j\hat{a}_k + \hat{a}^+_k\hat{a}_j), \\
\frac{i}{4}(\hat{a}^+_j\hat{a}_k - \hat{a}^+_k\hat{a}_j), & \quad j < k = 2, \ldots, n.
\end{align*} \tag{1.139}
$$

They commute with the total photon number

$$
\hat{N} = \sum_{j=1}^{n} \hat{a}^+_j\hat{a}_j, \quad \tag{1.140}
$$

and the unitary operators generated by them correspond to passive systems which preserve classicality. The remaining $n(n + 1)$ linearly independent hermitian quadratic noncompact generators can be given by

$$
\begin{align*}
\frac{1}{4}(\hat{a}^+_j\hat{a}^+_k + \hat{a}_k\hat{a}_j), \\
\frac{i}{4}(\hat{a}^+_j\hat{a}^+_k - \hat{a}^+_k\hat{a}_j), & \quad j \leq k = 1, 2, \ldots, n.
\end{align*} \tag{1.141}
$$
The maximal compact subgroup of \( Sp(2n, R) \) denoted by \( K(n) \) is generated by the \( n^2 \)
compact generators, and is given by

\[
S^{(r)}(X, Y) = \begin{pmatrix} X & Y \\ -Y & X \end{pmatrix},
\]

where \( X \) and \( Y \) are real \( n \times n \) matrices obeying

\[
XX^T + YY^T = 1, \quad XY^T = YX^T.
\]

It is easy to see that

\[
S^{(r)}(X, Y)S^{(r)T}(X, Y) = 1, \quad S^{(r)}(X, Y)\beta S^{(r)T}(X, Y) = \beta.
\]

Such a matrix is both orthogonal and symplectic in \( 2n \) dimensions. Going over to the complex form, we have

\[
S^{(c)}(X, Y) = \Gamma S^{(r)}(X, Y)\Gamma^\dagger = \begin{pmatrix} U & 0 \\ 0 & U^* \end{pmatrix} = S^{(c)}(U),
\]

\[
U = X - iY, \quad UU^\dagger = 1,
\]

That is \( X - iY \) (and hence \( X + iY \)) is a unitary matrix. Thus the maximal compact
subgroup \( K(n) \) of \( Sp(2n, R) \) is isomorphic to \( U(n) \). In other words,

\[
K(n) = SO(2n) \cap Sp(2n, R) = U(n).
\]

An interesting property of \( K(n) \) is that it acts transitively on the phase space unit sphere
\( S^{(2n-1)} \).

\textbf{Theorem 1.9} A variance matrix is squeezed if and only if

\[
l(V^{(r)}) < \frac{1}{2},
\]

where \( l(V^{(r)}) \) is the least eigenvalue of the variance matrix \( V^{(r)} \) \cite{192}.

One is able make this statement because of the transitive action of \( K(n) \) on the unit
sphere \( S^{(2n-1)} \), though one cannot in general diagonalise a given \( V^{(r)} \) by \( K(n) \) rotations
\cite{192}.
1.9.2 Entanglement in Gaussians

We now revise our notation into one which is useful in describing entanglement in multi-mode Gaussian states. The earlier notation was useful from the perspective of passive devices, and hence nonclassicality as the action of any passive device is compactly described as a unitary transformation on the annihilation operators. With entanglement in perspective, it is useful to make a modewise grouping of the canonical variables.

Consider a bipartite system consisting of $n$ modes, with $m$ modes in possession of Alice and the remaining $n-m$ modes in Bob’s possession. We introduce the following notation:

$$\hat{\xi}^{(a)} = (\hat{q}_1, \hat{p}_1, \hat{q}_2, \hat{p}_2, \ldots, \hat{q}_m, \hat{p}_m);$$

$$\hat{\xi}^{(b)} = (\hat{q}_{m+1}, \hat{p}_{m+1}, \hat{q}_{m+2}, \hat{p}_{m+2}, \ldots, \hat{q}_n, \hat{p}_n);$$

$$\hat{\xi} = (\hat{\xi}^{(a)}, \hat{\xi}^{(b)}).$$

The commutation relations are given in a compact form as

$$[\hat{\xi}_\alpha, \hat{\xi}_\beta] = i\Omega_{\alpha\beta}, \quad \alpha, \beta = 1, 2, \ldots, 2n,$$  \hfill (1.149)

where

$$\Omega = \begin{pmatrix}
J & 0 & \ldots & 0 \\
0 & J & \ldots & 0 \\
\vdots & \ddots & \ddots & \ddots \\
0 & 0 & \ldots & J
\end{pmatrix}, \quad J = \begin{pmatrix}
0 & 1 \\
-1 & 0
\end{pmatrix}.$$  \hfill (1.150)

Clearly, a canonical transformation $S$ satisfies

$$S\Omega S^T = \Omega.$$  \hfill (1.151)

Suppose that the state has a nonzero mean. Then we can define the translated operators

$$\Delta \hat{\xi} = \hat{\xi} - \langle \hat{\xi} \rangle,$$

in which case, the variance matrix is defined as

$$\langle \{\Delta \hat{\xi}_\alpha, \Delta \hat{\xi}_\beta\} \rangle = \text{Tr}(\{\Delta \hat{\xi}_\alpha, \Delta \hat{\xi}_\beta\} \hat{\rho}) = V_{\alpha\beta}.$$  \hfill (1.152)
The uncertainty principle is now stated as

\[ V + \frac{i}{2} \Omega \geq 0. \]  

(1.153)

We have the following theorem with regard to bipartite separability.

**Theorem 1.10** A necessary and sufficient condition for separability of a Gaussian state described by the variance matrix \( V \) is [280],

\[
V \geq \frac{1}{2} S_a S_a^T \oplus S_b S_b^T,
\]

\[
S_a \in Sp(2m, R), \quad S_b \in Sp(2n - 2m, R).
\]  

(1.154)

The inequality simply states that the variance matrix of a separable state is always greater than that of a pure product Gaussian state. In other words, a separable state can be made classical by local unitary canonical transformations. Such a thing is impossible if the Gaussian state was entangled across the \( a-b \) divide [26].

**Partial transpose:** As stated already in Section 1.8.2, partial transpose acts on the Wigner function as momentum reversal on Bob’s side. Under partial transpose, the variance matrix of a multipartite state across a bipartite cut, undergoes the change \( V \rightarrow \tilde{V} = \Lambda' \Lambda' \), where

\[
\Lambda = \text{diag}(1, 1, \cdots, 1, m, 1, m+1, -1, m+1, \cdots, 1, -1, n).
\]  

(1.155)

We can now implement partial transpose on the variance matrix and test for entanglement. In addition to the uncertainty principle, the variance matrix of a separable state has to obey

\[
\tilde{V} + \frac{i}{2} \Omega \geq 0, \quad \tilde{V} = \Lambda' \Lambda', \quad \Omega = \begin{pmatrix} \Omega_A & 0 \\ 0 & \Omega_B \end{pmatrix}.
\]  

(1.156)

This can be alternatively stated as

\[
V + \frac{i}{2} \bar{\Omega} \geq 0, \quad \bar{\Omega} = \Lambda' \Omega \Lambda', \quad \bar{\Omega} = \begin{pmatrix} \Omega_A & 0 \\ 0 & -\Omega_B \end{pmatrix}.
\]  

(1.157)

A variance matrix is said to be PPT, if it satisfies the above inequality.

Of particular interest are two-mode Gaussian states, whose variance matrix in block
form is given by

\[
V = \begin{pmatrix}
A & C \\
C^T & B
\end{pmatrix},
\]

(1.158)

where \(A\), \(B\), and \(C\) are \(2 \times 2\) matrices. Such variance matrices can be brought to the following canonical form using local symplectic transformations, i.e., using the action of \(S_{\text{local}} \in Sp(2, R) \oplus Sp(2, R)\):

\[
V \rightarrow V_0 = \begin{pmatrix}
a & 0 & c_1 & 0 \\
0 & a & 0 & c_2 \\
c_1 & 0 & b & 0 \\
0 & c_2 & 0 & b
\end{pmatrix},
\]

(1.159)

It turns out that in the two-mode case, PPT is a necessary and sufficient criterion for separability, and is stated in a local invariant form as [57]

\[
det A \det B + \left( \frac{1}{4} - |\det C| \right)^2 - \text{Tr}(AJC^TBJ + J) \geq \frac{1}{4}(\det A + \det B)
\]

\[
\Leftrightarrow 4(ab - c_1^2)(ab - c_2^2) \geq (a^2 + b^2) + 2|c_1c_2| - \frac{1}{4}.
\]

(1.160)

It is useful to note that Gaussian states with \(\det(C) \geq 0\) are separable.

In the case of symmetric two-mode Gaussian states specified by parameters \(a, b = a, c_1\) and \(c_2\), it is possible to solve for the entanglement of formation. It is given by [70]

\[
f(\Delta) = c_+(\Delta) \log[c_+(\Delta)] - c_-(\Delta) \log[c_-(\Delta)], \quad \text{where}
\]

\[
C_{\pm} = \left( \Delta - \frac{1}{2} \pm \frac{1}{2} \right)^2,
\]

\[
\Delta = 2\sqrt{(a - c_1)(a + c_2)}.
\]

(1.161)

The function \(f\) is a convex and decreasing function of \(\Delta\). We will study the EOF of general two-mode Gaussian states in Chapter 4.

### 1.9.3 Gaussian completely positive maps—Gaussian channels

A Gaussian completely positive map (or Gaussian channel) is any completely positive map which takes Gaussian states to Gaussian states. It was discussed initially in [149, 157], and more recently in [158, 159, 281-288]. In [158, 159, 281, 284, 286] single-mode Gaussian channels were discussed and classified, and issues regarding their channel capacity were studied. More recently multi-mode Gaussian channels have been classified in [287, 288].
Gaussian channels can be realised very much in the form of Eq. (1.26) and Eq. (1.27). But now one has to be careful to maintain the Gaussian character of the state. Given a Gaussian state $\hat{\rho}$, a simple way to generate the most general Gaussianity preserving map is to couple the given state to an auxiliary or ancilla Gaussian state, then evolve them together with a unitary evolution corresponding to the most general Gaussian preserving unitary, i.e., the unitary operator in this case consists of unitaries generated from canonical transformations corresponding to a symplectic group element and an arbitrary phase space translation, then trace away the auxiliary system. Clearly, all the operations done are canonical Gaussianity preserving operations. Thus the resultant state is Gaussian for every Gaussian input. The variance matrix of the resultant state is related to the variance matrix of the input state in the following way:

$$V \rightarrow X^T V X + Y,$$
$$X, Y \in \mathbb{R}^{2n \times 2n}. \quad (1.162)$$

Clearly, $X$ acts on the variance matrix through congruence and $Y$ is the additional noise matrix. Here $X$ and $Y$ are chosen so that the uncertainty principle is respected. We will have more to say on them in Chapter 6, where we derive the Kraus representation for a class of single-mode Gaussian channels.