

# Quantum Information Theory

Nihar Shah

May 19, 2025

# Quantum Information Theory

Nihar Shah

Guide: Professor Phani Motamarri



*Department of Computational and Data Science*  
*Indian Institute of Science*

## **Abstract**

The following content consists of the work done by me as a part of my Master's Thesis at Indian Institute of Science, Bengaluru. The work is done under the guidance of Prof. Phani Motamarri. This work includes notes made during the 1 year work from June 2024 to August 2025 and is based on the lectures, papers, books, and other resources that I have referred to during the course of my work. To a reader who is not at all familiar with Quantum Computing, this work will serve as a good starting point. I suggest starting with the Appendix to get a good grasp of the math used in the entire text. The work is divided into parts: The first part is the introduction to Quantum Computing, the second part is Quantum Linear Algebra, the third part is on Quantum Information Theory and at last the fourth part is on Quantum Error Correction. Along with the references, I have also included the code snippets that I have written during the course of my work. The references are included at the end of the document. I hope that this text helps the reader to get a good grasp on Quantum Computing. Enjoy reading!

## Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor, Professor Phani Motamarri, for the continuous support of my study and related, for his patience, and immense knowledge. His guidance helped me through the time of research and writing of this thesis.

Besides my advisor, I am also grateful to the faculty members and staff at the Department of Computational and Data Science of Indian Institute of Science for their assistance and support. Special thanks to MATRIX lab for providing a stimulating and collaborative research environment.

Last but not the least, I would like to thank my family: my parents, for supporting me spiritually throughout writing this thesis and my life in general.

# Contents

<b>Acknowledgements</b>	<b>ii</b>
<b>1 Introduction to Quantum Information</b>	<b>1</b>
1.1 Density Matrices . . . . .	2
1.2 Operations on Density matrices . . . . .	6
1.3 Postulates . . . . .	10
1.3.1 Postulate 1: The State Postulate . . . . .	10
1.3.2 Postulate 2: Evolution Postulate . . . . .	10
1.3.3 Postulate 3: Measurement Postulate . . . . .	10
1.3.4 Postulate 4: Composite Systems . . . . .	12
1.4 The Reduced Density Operator . . . . .	16
<b>Bibliography</b>	<b>19</b>

# Chapter 1

## Introduction to Quantum Information

So far, [1] this course has focused almost entirely on quantum algorithms. In the beginning of the course I described the model of quantum information that we have used up to this point. Specifically, the model describes states of qubits as unit vectors, and the allowed operations come from a fairly restricted set (unitary operations and measurements of a simple type). This description was sufficient for discussing quantum algorithms, so it has not been necessary to extend it until now. This is an alternate formulation using a tool known as the density operator or density matrix. This alternative formulation is mathematically equivalent to the state vector approach, but provides a much more convenient language for thinking about some commonly encountered scenarios in quantum mechanics.

To describe the model generally model of quantum information, some notation will be helpful. First, for any finite non empty set  $\Sigma$ , let  $\mathbb{C}(\Sigma)$  denote the vector space of all column vectors indexed by  $\Sigma$ . Just as before, elements of such spaces are denoted by kets, e.g.  $|\psi\rangle \in \mathbb{C}(\Sigma)$ , and scripted letters such as  $\mathcal{A}, \mathcal{B}, \mathcal{X}, \mathcal{Y}$ , etc., are generally used. For example, we might write  $\mathcal{X} = \mathbb{C}(\{0, 1\}^n)$  to indicate that the space  $\mathcal{X}$  is indexed by the set  $\{0, 1\}^n$ , and simply use the symbol  $\mathcal{X}$  to refer to that space from that point on. Although it will seldom be necessary, we may also write  $\mathbb{C}(\Sigma)^\dagger$  or  $\mathcal{X}^\dagger$  (for example) to refer to the space of all row vectors (or bra vectors)  $\langle\psi|$ , for  $|\psi\rangle \in \mathbb{C}(\Sigma)$  or  $|\psi\rangle \in \mathcal{X}$ . You will also see things written with a star  $*$  instead of a dagger  $\dagger$  sometimes, as in  $\mathcal{X}^*$ .

Also similar to before, when we consider a particular quantum system we assume that it has some finite set  $\Sigma$  of associated classical states. We will typically use the term register from now on to refer to abstract physical devices (such as qubits or

collection of qubits). Associated with any register having classical state  $\Sigma$  is the vector space  $\mathbb{C}(\Sigma)$ . It is sometimes helpful, but certainly not necessary, to use the same letter (in different fonts) to refer to a register and its associated space. For example, register  $X$  may have associated space  $\mathcal{X}$ .

## 1.1 Density Matrices

So far nothing is new except a little bit of notation. In order to describe what really is new, it is helpful to start with a special case. Suppose that in the “old” representation, a register  $X$  having classical state set  $\Sigma$  is in a quantum state  $|\psi\rangle \in \mathcal{X}$  for  $\mathcal{X} = \mathbb{C}(\Sigma)$ . The “new” way of representing this state will be:

$$|\psi\rangle \langle\psi|$$

We have seen objects like this before (in the analysis of Grover’s algorithm). It is effectively a matrix. For example, suppose  $\Sigma = \{0, 1\}$  and  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ . Then

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{and} \quad \langle\psi| = (\bar{\alpha} \quad \bar{\beta})$$

so

$$|\psi\rangle \langle\psi| = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} (\bar{\alpha} \quad \bar{\beta}) = \begin{pmatrix} \alpha\bar{\alpha} & \alpha\bar{\beta} \\ \bar{\alpha}\beta & \beta\bar{\beta} \end{pmatrix} = \begin{pmatrix} |\alpha|^2 & \alpha\bar{\beta} \\ \bar{\alpha}\beta & |\beta|^2 \end{pmatrix}$$

This is called a *density matrix* or *density operator*. (the term operation usually just means a linear map from a space to itself). Not all density operators have the special form  $|\psi\rangle \langle\psi|$  for some unit vector  $|\psi\rangle$ , as we will see. All of the quantum states we have considered so far in the course have actually been of this special kind of state.

Suppose we have a probability distribution  $(p_1, \dots, p_k)$ , for some positive integer  $k$ , as well as unit vectors  $|\psi_1\rangle, \dots, |\psi_k\rangle \in \mathcal{X}$ , where  $\mathcal{X} = \mathbb{C}(\Sigma)$  is the space corresponding to some register  $X$ . Somebody randomly chooses  $j \in \{1, \dots, k\}$  according to the probability distribution  $(p_1, \dots, p_k)$ , prepares the register  $X$  in the state  $|\psi_j\rangle$  for the chosen  $j$ , and then hands you  $X$  without telling you the value of  $j$ . The collection  $\{(p_1, |\psi_1\rangle), \dots, (p_k, |\psi_k\rangle)\}$  that describes the different possible states  $|\psi_j\rangle$  along with their associated probabilities is called a *mixture*. How do you represent this beyond specifying the mixture. in the “old” representation, there is no convenient way to do this beyond specifying the mixture. In the “new” representation, the density matrix corresponding to the above mixture is:

$$\rho = \sum_{j=1}^k p_j |\psi_j\rangle \langle\psi_j|$$

In other words it is meaningful to take a weighted average of the pure states  $|\psi_j\rangle\langle\psi_j|$ . More precisely, suppose a quantum system is in one of a number of states  $|\psi_i\rangle$ , where  $i$  is an index, with respective probabilities  $p_i$ . We shall call  $\{p_i, |\psi_i\rangle\}$  an ensemble of pure states.

**Example 1.1.1.** Suppose Alice has a qubit A. She flips a fair coin: if the results is heads she prepares A in the state  $|0\rangle$ , and if the result is Tails she prepares A in the state  $|1\rangle$ . She gives Bob the qubit without revealing the result of the coin-flip. Bob's knowledge of the qubit is described by the density matrix

$$\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

**Example 1.1.2.** Suppose Alice has a qubit A as in the previous example. As before she flips a fair coin, but now if the result is Heads she prepares A in the state  $|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ , and if the result is Tails she prepares A in the state  $|-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$ . Bob's knowledge of the qubit is now described by the density matrix

$$\frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -| = \frac{1}{2}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

Same as before!

The previous two examples demonstrate that different mixtures can yield precisely the same density matrix. This is not an accident, but rather is one of the strengths of the density matrix formalism. A given density matrix in essence represents a perfect description of the state of a quantum system - two different mixtures can be distinguished (in a statistical sense) if and only if they yield different density matrices. Whether one uses the density operator language or the state vector language is a matter of taste, since both give the same results; however it is sometimes much easier to approach problems from one point of view rather than the other.

Typically, lower case Greek letters such as  $\rho, \xi, \sigma$  and  $\tau$  are used to denote density matrices. The state of a system that corresponds to a given density matrix is called a mixed state.

A quantum system whose state  $|\psi\rangle$  is known exactly is said to be in a *pure state*. In this case density operator is simply  $\rho = |\psi\rangle\langle\psi|$ . Otherwise,  $\rho$  is in a *mixed state*; it is said to be a mixture of the different pure states in the ensemble for  $\rho$ . Note: sometimes people may use the term 'mixed' state as a catch-all to include both pure and mixed quantum states. The term 'pure state' is often used in reference to a state vector  $|\psi\rangle$ , to distinguish it from a density operator  $\rho$ .



Finally, imagine a quantum system is prepared in the state  $\rho_i$  with probability  $p_i$ . It is not difficult to convince yourself that the system may be described by the density matrix  $\sum_i p_i \rho_i$ . A proof of this is to suppose that  $\rho_i$  arises from some ensemble  $\{p_{ij}, |\psi_{ij}\rangle\}$  (note that  $i$  is fixed) of pure states, so the probability for being in the state  $|\psi_{ij}\rangle$  is  $p_i p_{ij}$ . The density matrix for the system is thus

$$\begin{aligned}\rho &= \sum_{ij} p_i p_{ij} |\psi_{ij}\rangle \langle \psi_{ij}| \\ &= \sum_i p_i \rho_i\end{aligned}$$

where we have used the definition  $\rho_i = \sum_j p_{ij} |\psi_{ij}\rangle \langle \psi_{ij}|$ . We say that  $\rho$  is a mixture of the state  $\rho_i$  with probabilities  $p_i$ . This concept of a mixture comes up repeatedly in the analysis of problems like quantum noise, where the effect of the noise is to introduce ignorance into our knowledge of the quantum state. A simple example is provided by the measurement scenario. Imagine that, for some reason our record of the result  $m$  of the measurement was lost. We would have a quantum system in the state  $\rho_m$  with probability  $p(m)$ , but would no longer know the actual value of  $m$ . The state of such a quantum system would therefore be described by the density operator

$$\begin{aligned}\rho &= \sum_m p(m) \rho_m \\ &= \sum_m \text{Tr}(M_m^\dagger M_m \rho) \frac{M_m \rho M_m^\dagger}{\text{tr}(M_m^\dagger M_m \rho)} \\ &= \sum_m M_m \rho M_m^\dagger\end{aligned}$$

General properties of density matrices: We now move away from this description to develop an intrinsic characterization of density operator that does not rely on an ensemble interpretation. This allows us to complete the program of giving a description of quantum mechanics that does not take as its foundation the state vector.

### 1. Trace Condition: The trace of a density matrix is 1

(The trace of a matrix is sum of its diagonal entries)

In fact, the diagonal entries describe precisely the probability distribution that would result if the system in question were measured (with respect to the restricted type of measurement we have so far considered).

The fact that the trace is a linear function together with the useful formula  $Tr(AB) = Tr(BA)$  (which holds for any choice of matrices A and B for which the product AB is square) makes the above fact easy to verify for any given mixture:

$$Tr \left( \sum_{j=1}^k p_j |\psi_j\rangle \langle \psi_j| \right) = \sum_{j=1}^k p_j Tr(|\psi_j\rangle \langle \psi_j|) = \sum_{j=1}^k p_j Tr(\langle \psi_j| |\psi_j\rangle) = \sum_{j=1}^k p_j \langle \psi_j| \psi_j \rangle = 1$$

## 2. Positivity Condition: Every density matrix is positive semi-definite.

In general, a square matrix A is positive semi-definite if  $\langle \psi| A | \psi \rangle$  is a non-negative real number for every vector  $|\psi\rangle$ . An equivalent definition is that A is positive semi-definite if (i)  $A = A^\dagger$  (i.e., A is Hermitian), and (ii) all eigenvalues of A are non-negative real numbers.

Again this condition is easy to check for mixtures: if

$$\rho = \sum_{j=1}^k p_j |\psi_j\rangle \langle \psi_j|$$

and  $|\phi\rangle$  is any vector, we have

$$\langle \phi | \rho | \phi \rangle = \sum_{j=1}^k p_j \langle \phi | \psi_j \rangle \langle \psi_j | \phi \rangle = \sum_{j=1}^k p_j |\langle \phi | \psi_j \rangle|^2 \geq 0$$

You can interpret the above facts as the defining properties of density matrices: by definition, a density matrix is any matrix that is positive semi-definite and has trace equal to 1. Given such a matrix, it is possible to find a mixture that yields the given density matrix by letting  $p_1, \dots, p_k$  be the nonzero eigenvalues and  $|\psi_1\rangle, \dots, |\psi_k\rangle$ . Suppose  $\rho$  is any operator satisfying the trace and positivity conditions, Since  $\rho$  is positive, it must have a spectral decomposition

$$\rho = \sum_j \lambda_j |j\rangle \langle j|$$

where the vectors  $|j\rangle$  are orthogonal, and  $\lambda_j$  are real, non-negative eigenvalue of  $\rho$ . from the trace condition we see that  $\sum_j \lambda_j = 1$ . Therefore, a system in the state  $|j\rangle$  with probability  $\lambda_j$  will have density operator  $\rho$ . That is, the ensemble  $\{\lambda_j, |j\rangle\}$  is an ensemble of states giving rise to the density operator  $\rho$ . This theorem provides a characterization of density operators that is intrinsic to the operator itself: we can define a density operator to be a positive operator  $\rho$  which has trace equal to one. Making this definition allows us to reformulate the postulates of quantum mechanics in the density operator picture.

## 1.2 Operations on Density matrices

In our “old” description of quantum information, a unitary operation  $U$  applied to a state  $|\psi\rangle$  resulted in the state  $U|\psi\rangle$ . In the density matrix description, a unitary operation  $U$  applied to a pure state  $|\psi\rangle\langle\psi|$  results in the density matrix

$$U|\psi\rangle\langle\psi|U^\dagger$$

This is consistent with the observation that  $(U|\psi\rangle)^\dagger = \langle\psi|U^\dagger$ . More generally, applying  $U$  to the mixed state  $\rho$  results in the state  $U\rho U^\dagger$ . you can easily that that the two required conditions of density matrices are necessarily met by this new matrix.

We can, however, consider a much more general set of possible operations than just unitary operations. Any operation  $\Phi$  that can be written as

$$\Phi(\rho) = \sum_{j=1}^k A_j \rho A_j^\dagger$$

for some collection of matrices  $A_1, \dots, A_k$  satisfying

$$\sum_{j=1}^k A_j^\dagger A_j = I$$

represents an operation that can (in an idealized sense) be physically implemented. Such operations are called admissible operations. (There are several other names that are used as well, such as completely positive trace preserving operations and other variations on these words.)

If  $\rho$  is a matrix and  $\Phi$  is admissible, then  $\Phi(\rho)$  is also a density matrix. In fact a somewhat stronger property holds, which is that if  $\Phi$  is applied to just part of a larger system whose state is described by some density matrix, then the resulting state will also be described by a density matrix.

**Example 1.2.1.** Suppose we have a single qubit  $X$ . Consider the operations that corresponds to measuring the qubit and forgetting the result. An admissible operation that describes this process is given by

$$A_0 = |0\rangle\langle 0|$$

$$A_1 = |1\rangle\langle 1|$$

$$\Phi(\rho) = \sum_{j=0}^1 A_j \rho A_j^\dagger = |0\rangle \langle 0| \rho |0\rangle \langle 0| + |1\rangle \langle 1| \rho |1\rangle \langle 1|$$

First let us check that this is a valid admissible operation:

$$\sum_{j=0}^1 A_j A_j^\dagger = |0\rangle \langle 0|0\rangle \langle 0| + |1\rangle \langle 1|1\rangle \langle 1| = |0\rangle \langle 0| + |1\rangle \langle 1| = I$$

It satisfies the required property, so indeed it is admissible.

What does it do to the state  $\alpha |0\rangle + \beta |1\rangle$ , for example? First we need to represent  $|\psi\rangle$  as a density matrix:  $\rho = (|\psi\rangle \langle \psi|)$ . Now

$$\begin{aligned} \Phi(\rho) &= \Phi(|\psi\rangle \langle \psi|) \\ &= |0\rangle \langle 0| \psi\rangle \langle \psi|0\rangle \langle 0| + |1\rangle \langle 1| \psi\rangle \langle \psi|1\rangle \langle 1| \\ &= |\langle 0|\psi\rangle|^2 |0\rangle \langle 0| + |\langle 1|\psi\rangle|^2 |1\rangle \langle 1| \\ &= |\alpha|^2 |0\rangle \langle 0| + |\beta|^2 |1\rangle \langle 1| \end{aligned}$$

In terms of matrices:

$$\begin{pmatrix} |\alpha|^2 & \alpha\bar{\beta} \\ \bar{\alpha}\beta & |\beta|^2 \end{pmatrix} \xrightarrow{\Phi} \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}$$

In general, off diagonal entries get zeroed out.

In general, admissible operation do not need to preserve the sizes of quantum systems. For example, one might consider the situation in which one of a collection of qubits is lost or somehow destroyed. More notation will help to discuss this issue in greater specificity.

Given two spaces  $\mathcal{X}$  and  $\mathcal{Y}$ , we let

$$L(\mathcal{X}, \mathcal{Y})$$

denote the set of all linear mappings from  $\mathcal{X}$  to  $\mathcal{Y}$ . The shorthand  $L(\mathcal{X})$  is used to mean  $L(\mathcal{X}, \mathcal{X})$ . Also let

$$D(\mathcal{X})$$

denote the set of all density matrices on  $\mathcal{X}$  (so that  $D(\mathcal{X}) \subset L(\mathcal{X})$ ). For example, if  $X$  is a quantum register with classical set  $\Sigma$  and  $\mathcal{X} = \mathbb{C}(\Sigma)$ , then any mixed state of the register  $X$  is represented by some element of  $D(\mathcal{X})$

Now, if we have a collection of mappings (or matrices)  $A_1, \dots, A_k \in L(\mathcal{X}, \mathcal{Y})$  that satisfy

$$\sum_{j=1}^k A_j^\dagger A_j = I$$

(the density matrix in  $L(\mathcal{X})$ ) then the admissible operation  $\Phi$  is defined by

$$\Phi(\rho) = \sum_{j=1}^k A_j \rho A_j^\dagger$$

maps elements of  $D(\mathcal{X})$  to elements of  $D(\mathcal{Y})$ .

**Example 1.2.2.** Let us suppose that we have two qubits:  $X$  and  $Y$ . We will consider the admissible operation that corresponds to discarding the second qubit. Thus, once the operation has been performed, we will be left with a single qubit  $X$ . The vector space corresponding to  $X$  will be  $\mathcal{X}$  and the space corresponding to  $Y$  will be  $\mathcal{Y}$ .

Now if  $\Phi$  is to describe the operation of discarding the second qubit, then we must have that  $\Phi(\rho) \in D(\mathcal{X})$  whenever  $\rho \in D(\mathcal{X} \otimes \mathcal{Y})$ . This means that if

$$\Phi(\rho) = \sum_{j=1}^k A_j \rho A_j^\dagger$$

for some choice of matrices  $A_1, \dots, A_k$  then these matrices must come from the set  $L(\mathcal{X} \otimes \mathcal{Y}, \mathcal{X})$ . this means that they must be  $2 \times 4$  matrices. Let

$$A_0 = I \otimes \langle 0|, \quad A_1 = I \otimes \langle 1|$$

where  $I$  denotes the identity operator on the first qubit, and define  $\Phi$  by

$$\Phi(\rho) = \sum_{j=0}^1 A_j \rho A_j^\dagger = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger$$

for all  $\rho$ . writing  $A_0$  and  $A_1$  in ordinary matrix notation gives

$$\begin{aligned} A_0 &= I \otimes \langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes (1 \ 0) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ A_1 &= I \otimes \langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes (0 \ 1) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

To see that  $\Phi$  is admissible, we compute:

$$\begin{aligned} A_0^\dagger A_0 + A_1^\dagger A_1 &= (I \otimes |0\rangle)(I \otimes \langle 0|) + (I \otimes |1\rangle)(I \otimes \langle 1|) \\ &= I \otimes |0\rangle \langle 0| + I \otimes |1\rangle \langle 1| \\ &= I \otimes (|0\rangle \langle 0| + |1\rangle \langle 1|) \\ &= I \otimes I \\ &= I_{\mathcal{X} \otimes \mathcal{Y}} \end{aligned}$$

(The subscript on the identity operator in the last line is just representing what space the identity is action on.)

Let us now consider the effect of this operation on couple of states. First suppose that  $(X, Y)$  is in the state  $\xi \otimes \sigma$  for two  $2 \times 2$  density matrices  $\xi$  and  $\sigma$ . Just as for vectors in the “old” description of quantum information, we view states of the form  $\xi \otimes \sigma$  as representing completely uncorrelated qubits. What do you expect the state to be if you discard the second qubit? Now let us check:

$$\begin{aligned}
 \Phi(\xi \otimes \sigma) &= A_0(\xi \otimes \sigma)A_0^\dagger + A_1(\xi \otimes \sigma)A_1^\dagger \\
 &= (I \otimes \langle 0|)(\xi \otimes \sigma)(I \otimes |0\rangle) + (I \otimes \langle 1|)(\xi \otimes \sigma)(I \otimes |1\rangle) \\
 &= \xi \otimes \langle 0|\sigma|0\rangle + \xi \otimes \langle 1|\sigma|1\rangle \\
 &= (\langle 0|\sigma|0\rangle + \langle 1|\sigma|1\rangle) \xi \\
 &= \text{Tr}(\sigma)\xi \\
 &= \xi
 \end{aligned}$$

Indeed this what you presumably expected.

Now we shall see that happens when this operation is applied to entangled qubits.

The previous example can be generalized to describe the operation that corresponds to discarding part of a system. Suppose  $X$  and  $Y$  are registers with corresponding spaces  $\mathcal{X}$  and  $\mathcal{Y}$ . A mixed state of these two registers is represented by some element of  $D(\mathcal{X} \otimes \mathcal{Y})$ . If, say the register  $Y$  is discarded, we will be left with some mixed state of  $X$  that is represented by some element of  $D(\mathcal{X})$ . Specifically, if  $\rho \in D(\mathcal{X} \otimes \mathcal{Y})$  is a density matrix representing the state of  $(X, Y)$  and  $Y$  is discarded, the resulting state of  $X$  is denoted by  $\text{Tr}_{\mathcal{Y}}(\rho)$  the partial trace. (It is called the partial trace because when extended by linearity to arbitrary matrices it satisfies  $\text{Tr}_{\mathcal{Y}}(X \otimes Y) = \text{Tr}(Y)X$  for all  $X \in L(\mathcal{X})$  and  $Y \in L(\mathcal{Y})$ .) We also refer to the action corresponding to this operation as tracing out the space  $\mathcal{Y}$ . to express  $\text{Tr}_{\mathcal{Y}}$  in the form of an admissible operation, let  $\Sigma$  denote the set of classical states of  $Y$ . Then

$$\text{Tr}_{\mathcal{Y}}(\rho) = \sum_{a \in \Sigma} (I \otimes \langle a|) \rho (I \otimes |a\rangle)$$

The partial trace is a particularly important admissible operation that we will continue discussing.

## 1.3 Postulates

It turns out that all the postulates of quantum mechanics can be reformulated in terms of the density operator language. The purpose of this section and the next is to explain how to perform this reformulation, and explain when it is useful.

### 1.3.1 Postulate 1: The State Postulate

Associated to any isolated physical system is a complex vector space with inner product (That is, a Hilbert space) known as the state space of the system. The system is completely described by its density operator, which is a positive operator  $\rho$  with trace one, acting on the state space of the system. If a quantum system is in the state  $\rho_i$  with probability  $p_i$ , then the density operator for the system is  $\sum_i p_i \rho_i$ .

### 1.3.2 Postulate 2: Evolution Postulate

The evolution of a closed system is described by a Unitary transformation. That is, the state  $\rho$  of the system at time  $t_1$  is related to the state  $\rho'$  of the system at time  $t_2$  by a unitary operator  $U$  which depends only on the times  $t_1$  and  $t_2$ .

Suppose evolution of a closed quantum system is described by the unitary operator  $U$ . If the system was initially in the state  $|\psi_i\rangle$  with probability  $p_i$  then after the evolution has occurred the system will be in the state  $U|\psi_i\rangle$  with probability  $p_i$ . Thus, the evolution of the density operator is described by the equation

$$\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i| \xrightarrow{U} \sum_i p_i U |\psi_i\rangle \langle \psi_i| U^\dagger = U \rho U^\dagger$$

### 1.3.3 Postulate 3: Measurement Postulate

Quantum measurements are described by a collection  $\{M_m\}$  of measurement operators. These are operators acting on the state space of the system being measured. The index  $m$  refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is  $\rho$  immediately before the measurement then the probability that the result  $m$  occurs is given by

$$p(m) = \langle \psi | M_m^\dagger M_m | \psi \rangle = \text{Tr}(M_m^\dagger M_m \rho)$$

and the state of the system after the measurement is

$$\frac{M_m |\psi\rangle \langle \psi| M_m^\dagger}{\sqrt{p(m)}} = \frac{M_m \rho M_m^\dagger}{\text{Tr}(M_m^\dagger M_m \rho)}$$

The measurement operators satisfy the completeness equation,

$$\sum_m M_m^\dagger M_m = I$$

Measurements are also easily described in the density operator language. Suppose we perform a measurement described by measurement operators  $M_m$ . If the initial state was  $|\psi_i\rangle$ , then the probability of getting result  $m$  is

$$p(m|i) = \langle \psi_i | M_m^\dagger M_m | \psi_i \rangle = \text{Tr}(M_m^\dagger M_m |\psi_i\rangle \langle \psi_i|)$$

Using the cyclic property of trace of a matrix. By the law of total probability (see Appendix) the probability of obtaining the result  $m$  is

$$\begin{aligned} p(m) &= \sum_i p(m|i) p_i \\ &= \sum_i p_i \text{Tr}(M_m^\dagger M_m |\psi_i\rangle \langle \psi_i|) \\ &= \text{Tr}(M_m^\dagger M_m \sum_i p_i |\psi_i\rangle \langle \psi_i|) \\ &= \text{Tr}(M_m^\dagger M_m \rho) \end{aligned}$$

The density operator of the system after the measurement result  $m$  is

$$|\psi_i^m\rangle = \frac{M_m |\psi_i\rangle}{\sqrt{\langle \psi_i | M_m^\dagger M_m | \psi_i \rangle}}$$

Thus, after a measurement which yields the result  $m$  we have an ensemble of states  $|\psi_i^m\rangle$  with respective probabilities  $p(i|m)$ . The corresponding density operator  $\rho_m$  is therefore

$$\rho_m = \sum_i p(i|m) |\psi_i^m\rangle \langle \psi_i^m| = \sum_i p(i|m) \frac{M_m |\psi_i\rangle \langle \psi_i| M_m^\dagger}{\text{Tr}(M_m^\dagger M_m |\psi_i\rangle \langle \psi_i|)}$$

But by elementary probability theory,  $p(i|m) = p(m, i)/p(m) = p(m|i)p_i/p(m)$ . Substituting, we obtain

$$\begin{aligned} \rho_m &= \sum_i p_i \frac{M_m |\psi_i\rangle \langle \psi_i| M_m^\dagger}{\text{Tr}(M_m^\dagger M_m \rho)} \\ &= \frac{M_m \rho M_m^\dagger}{\text{Tr}(M_m^\dagger M_m \rho)} \end{aligned}$$

Thus, the basic postulates of quantum mechanics related to unitary evolution and measurements can be rephrased in the language of density operators.



### 1.3.4 Postulate 4: Composite Systems

The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. moreover, if we have systems numbered 1 through  $n$ , and system number  $i$  is prepared in the state  $\rho_i$ , then the joint state of the total system is  $\rho_1 \otimes \rho_2 \otimes \dots \rho_n$ .

These reformulations of the fundamental postulates of quantum mechanics in terms of density operator are mathematically equivalent to the description in terms of the state vector. As a way of thinking about quantum mechanics, the density operator approach really shines for two applications: the description of quantum systems whose state is not known, and the description of subsystems of a composite quantum system.

**Example 1.3.1.** Let  $\rho$  be a density operator. Show that  $Tr(\rho^2) \leq 1$ , with equality if and only if  $\rho$  is a pure state.

Recall that  $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$  Now,

$$\begin{aligned}
 Tr(\rho^2) &= Tr\left(\left(\sum_i p_i |\psi_i\rangle \langle \psi_i|\right) \left(\sum_j p_j |\psi_j\rangle \langle \psi_j|\right)\right) \\
 &= Tr\left(\sum_i \sum_j p_i p_j |\psi_i\rangle \langle \psi_i| \psi_j\rangle \langle \psi_j|\right) \\
 &= \sum_i \sum_j p_i p_j Tr(|\psi_i\rangle \langle \psi_i| \psi_j\rangle \langle \psi_j|) \\
 &= \sum_i \sum_j p_i p_j Tr(\langle \psi_j| \psi_i\rangle \langle \psi_i| \psi_j\rangle) \\
 Tr(\rho^2) &= \sum_i \sum_j p_i p_j |\langle \psi_i| \psi_j\rangle|^2
 \end{aligned}$$

Now using the Cauchy-Schwarz Inequality, we know,  $|\langle \psi_i| \psi_j\rangle|^2 \leq \langle \psi_i| \psi_i\rangle \langle \psi_j| \psi_j\rangle$ , we

get,

$$\begin{aligned}
\sum_{i,j} p_i p_j |\langle \psi_i | \psi_j \rangle|^2 &\leq \sum_{i,j} p_i p_j \langle \psi_i | \psi_i \rangle \langle \psi_j | \psi_j \rangle \\
&\leq \sum_{i,j} p_i p_j \\
&\leq \left( \sum_i p_i \right) \left( \sum_j p_j \right) \\
&\leq 1 \\
\text{Tr}(\rho^2) &\leq 1
\end{aligned}$$

Clearly, the equality holds here if and only if  $|\psi\rangle = |\psi_j\rangle \quad \forall i, j$  upto a global phase factor i.e. there is only one  $|\psi\rangle$  that occurs with probability 1.

It is tempting (and surprisingly common) fallacy to suppose that the eigenvalues and eigenvectors of a density matrix have some special significance with regard to the ensemble of quantum states represented by that density matrix. For example, one might suppose that a quantum system with density matrix

$$\rho = \frac{3}{4} |0\rangle \langle 0| + \frac{1}{4} |1\rangle \langle 1|$$

must be in the state  $|0\rangle$  with probability  $3/4$  and in the state  $|1\rangle$  with probability  $1/4$ . However, this is not necessarily the case. Suppose we define

$$|a\rangle = \sqrt{\frac{3}{4}} |0\rangle + \sqrt{\frac{1}{4}} |1\rangle$$

$$|b\rangle = \sqrt{\frac{3}{4}} |0\rangle - \sqrt{\frac{1}{4}} |1\rangle$$

and the quantum system is prepared in the state  $|a\rangle$  with probability  $1/2$  and in the state  $|b\rangle$  with probability  $1/2$ . Then it is easily checked that the corresponding density matrix is

$$\rho = \frac{1}{2} |a\rangle \langle a| + \frac{1}{2} |b\rangle \langle b| = \frac{3}{4} |0\rangle \langle 0| + \frac{1}{4} |1\rangle \langle 1|$$

That is, two different ensembles of quantum states give rise to the same density matrix. In general, the eigenvectors and eigenvalues of a density matrix just indicate

one of many possible ensembles that may give rise to a specific density matrix, and there is no reason to suppose it is an especially privileged ensemble.

That is, two different ensembles of quantum states give rise to the same density matrix. In general, eigenvalues and eigenvectors of a density matrix indicate one of many possible ensembles that may give rise to a specific density matrix, and there is no reason to suppose it is an especially privileged ensemble.

A natural question to ask in the light of this discussion is what class of ensemble does give rise to a particular density matrix? The solution to this problem, which we now give, has surprisingly many applications in quantum computation and quantum information, notably in the understanding of quantum noise and quantum error-correction. For the solution it is convenient to make use of vectors  $|\tilde{\psi}_i\rangle$  which may not be normalized to unit length. We say the set  $|\tilde{\psi}_i\rangle$  generates the operator  $\rho = \sum_i |\tilde{\psi}_i\rangle \langle \tilde{\psi}_i|$ , and thus the connection to the unusual ensemble picture of density operators is expressed by the equation  $|\tilde{\psi}_i\rangle = \sqrt{p_i} |\psi_i\rangle$ . When do two sets of vectors,  $|\tilde{\psi}_i\rangle$  and  $|\tilde{\varphi}_j\rangle$  generate the same operator  $\rho$ ? The solution to this problem will enable us to answer the question of what ensemble gives rise to a given density matrix.

**Theorem 1.3.1. (*Unitary freedom in the ensemble for density matrices*)**  
*The sets  $|\tilde{\psi}_i\rangle$  and  $|\tilde{\varphi}_j\rangle$  generate the same density matrix if and only if*

$$|\tilde{\psi}_i\rangle = \sum_j u_{ij} |\tilde{\varphi}_j\rangle$$

where  $u_{ij}$  is a unitary matrix of complex numbers, with indices  $i$  and  $j$ , and we ‘pad’ whichever set of vectors  $|\tilde{\psi}_i\rangle$  or  $|\tilde{\varphi}_j\rangle$  is smaller with additional vectors 0 so that the two sets have the same number of elements. As a consequence of the theorem, note that  $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i| = \sum_j q_j |\varphi_j\rangle \langle \varphi_j|$  for normalized states  $|\psi_i\rangle, |\varphi_j\rangle$  and probability distributions  $p_i$  and  $q_j$  if and only if

$$\sqrt{p_i} |\psi_i\rangle = \sum_j u_{ij} \sqrt{q_j} |\varphi_j\rangle$$

for some unitary matrix  $u_{ij}$ , and we may pad the smaller ensemble with entries having probability zero in order to make the two ensembles the same size. Thus, the theorem characterizes the freedom in the ensembles  $\{p_i, |\psi_i\rangle\}$  giving rise to a given density matrix  $\rho$ . indeed, it is easily checked that our earlier example of a density matrix with two different decomposition, arises as a special case of this general result.

*Proof.* Suppose  $|\tilde{\psi}_i\rangle = \sum_j u_{ij} |\tilde{\varphi}_j\rangle$  for some unitary  $u_{ij}$ . Then

$$\begin{aligned} \sum_i |\tilde{\psi}_i\rangle \langle \tilde{\psi}_i| &= \sum_{ijk} u_{ij} u_{ik}^* |\tilde{\varphi}_j\rangle \langle \tilde{\varphi}_k| \\ &= \sum_{jk} \left( \sum_i u_{ki}^\dagger u_{ij} \right) |\tilde{\varphi}_j\rangle \langle \tilde{\varphi}_k| \\ &= \sum_{jk} \delta_{kj} |\tilde{\varphi}_j\rangle \langle \tilde{\varphi}_k| \\ &= \sum_j |\tilde{\varphi}_j\rangle \langle \tilde{\varphi}_j| \end{aligned}$$

which shows that  $|\tilde{\psi}_i\rangle$  and  $|\tilde{\varphi}_j\rangle$  generate the same operator.

Conversely, suppose

$$A = \sum_i |\tilde{\psi}_i\rangle \langle \tilde{\psi}_i| = \sum_j |\tilde{\varphi}_j\rangle \langle \tilde{\varphi}_j|$$

Let  $A = \sum_k \lambda_k |k\rangle \langle k|$  be a decomposition for  $A$  such that the states  $|k\rangle$  are orthonormal, and the  $\lambda_k$  are strictly positive. Our strategy is to relate the state  $|\tilde{\psi}_i\rangle$  to the states  $|\tilde{k}\rangle = \sqrt{\lambda_k} |k\rangle$ , and similarly relate the states  $|\tilde{\varphi}_j\rangle$  to the states  $|\tilde{k}\rangle$ . Combining the two relations will give the result. Let  $|\psi\rangle$  be any vector orthonormal to the space spanned by  $|\tilde{k}\rangle$ , so  $\langle \psi | \tilde{k} \rangle \langle \tilde{k} | \psi \rangle = 0$  for all  $k$ , and thus we see that

$$0 = \langle \psi | A | \psi \rangle = \sum_i \langle \psi | \tilde{\psi}_i \rangle \langle \tilde{\psi}_i | \psi \rangle = \sum_i |\langle \psi | \tilde{\psi}_i \rangle|^2$$

Thus  $\langle \psi | \tilde{\psi}_i \rangle = 0$  for all  $i$  and all  $|\psi\rangle$  orthonormal to the space spanned by the  $|\tilde{k}\rangle$ . It follows that each  $|\tilde{\psi}_i\rangle$  can be expressed as a linear combination of the  $|\tilde{k}\rangle$ ,  $|\tilde{\psi}_i\rangle = \sum_k c_{ik} |\tilde{k}\rangle$ . Since  $A = \sum_k |\tilde{k}\rangle \langle \tilde{k}| = \sum_i |\tilde{\psi}_i\rangle \langle \tilde{\psi}_i|$  we see that

$$\sum_k |\tilde{k}\rangle \langle \tilde{k}| = \sum_{kl} \left( \sum_i c_{ik} c_{il}^* \right) |\tilde{k}\rangle \langle \tilde{l}|$$

The operators  $|\tilde{k}\rangle \langle \tilde{l}|$  are easily seen to be linearly independent, and thus it must be that  $\sum_i c_{ik} c_{il}^* = \delta_{kl}$ . This ensures that we may append extra columns to  $c$  to obtain a unitary matrix  $v$  such that  $|\tilde{\psi}_i\rangle = \sum_k v_{ik} |\tilde{k}\rangle$ , where we have appended zero vectors to the list of  $|\tilde{k}\rangle$ . Similarly, we can find a unitary matrix  $w$  such that  $|\tilde{\varphi}_j\rangle = \sum_k w_{jk} |\tilde{k}\rangle$ . Thus  $|\tilde{\psi}_i\rangle = \sum_j u_{ij} |\tilde{\varphi}_j\rangle$ , where  $u = vw^\dagger$  is unitary.  $\square$

**Example 1.3.2. (Bloch sphere for mixed states)** The Bloch sphere picture for pure states of a single qubit

## 1.4 The Reduced Density Operator

Perhaps the deepest application of the density operator is as a descriptive tool for sub-systems of a composite quantum systems. Such a description is provided by the reduced density operator, which is the subject of this section. The reduced density operator is so useful as to be virtually indispensable in the analysis of composite quantum systems.

Suppose we have physical systems  $A$  and  $B$ , whose states are described by a density operator  $\rho^{AB}$ . The reduced density operator for system  $A$  is defined by

$$\rho^A \equiv \text{tr}_B(\rho^{AB})$$

where  $\text{tr}_B$  is a map of operators known as the partial trace over system  $B$ . The partial trace is defined by

$$\text{tr}_B(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|) \equiv |a_1\rangle\langle a_2| \text{tr}(|b_1\rangle\langle b_1|)$$

where  $|a_1\rangle$  and  $|a_2\rangle$  are any two vectors in the state space of  $A$ , and  $|b_1\rangle$  and  $|b_2\rangle$  are any two vectors in the state space of  $B$ . The trace operation appearing on the right hand side is the usual trace operation for system  $B$ , so  $\text{tr}(|b_1\rangle\langle b_1|) = \langle b_1|b_1\rangle$ . We have defined the partial trace operation only on a special subclass of operators  $AB$ ; the specification is completed by requiring in addition to the above equation that the partial trace be linear in its input.

It is not obvious that the reduced density operator for system  $A$  is in any sense a description for the state of system  $A$ . The physical justification for making this identification is that the reduced density operator provides the correct measurement statistics for measurements made on system  $A$ . This is

**Example 1.4.1. (Block sphere for mixed states)** The Bloch sphere description has an important generalization for mixed states as follows:

1. Show that an arbitrary density matrix for a mixed state qubit may be written as

$$\rho = \frac{I + \mathbf{r} \cdot \boldsymbol{\sigma}}{2}$$

where  $\mathbf{r}$  is a real three-dimensional vector such that  $\|\mathbf{r}\| \leq 1$ . This vector is known as the Bloch vector for the state  $\rho$ .

2. What is the Bloch vector representation for the state  $\rho = I/2$ .
3. Show that a state  $\rho$  is pure if and only if  $\|\mathbf{r}\| = 1$ .

4. Show that for pure state the description of the Bloch vector we have given coincides with the previous definition.

Since  $\{I, X, Y, Z\}$  form an basis the vector space of single-qubit linear operators, we can write (for any  $\rho$ , regardless of whether it is a density operator or not):

$$\rho = a_1 I + a_2 X + a_3 Y + a_4 Z$$

for constants  $a_1, a_2, a_3, a_4 \in \mathbb{C}$ . Since  $\rho$  is a Hermitian operator, we find that each of these constants are actually real, as

$$a_1 I + a_2 X + a_3 Y + a_4 Z = \rho = \rho^\dagger = a_1^* I^\dagger + a_2^* X^\dagger + a_3^* Y^\dagger + a_4^* Z^\dagger = a_1^* I + a_2^* X + a_3^* Y + a_4^* Z$$

Now, we require that  $\text{tr}(\rho) = 1$  for any density operator, hence:

$$\text{tr}(\rho) = \text{tr}(a_1 I + a_2 X + a_3 Y + a_4 Z) = a_1 \text{tr}(I) + a_2 \text{tr}(X) + a_3 \text{tr}(Y) + a_4 \text{tr}(Z) = 2a_1 = 1$$

from which we obtain that  $a_1 = \frac{1}{2}$ . Note that in the second equality we use the linearity of the trace, and in the third equality we use that  $\text{Tr}(I) = 2$  and  $\text{Tr}(\sigma_i) = 0$  for  $i \in \{1, 2, 3\}$ . Calculating  $\rho^2$ , we have that:

$$\begin{aligned} \rho^2 &= \frac{1}{4} I + \frac{a_2}{2} X + \frac{a_3}{2} Y + \frac{a_4}{2} Z + \frac{a_2}{2} X + a_2^2 X^2 + a_2 a_3 XY + a_2 a_4 XZ \\ &\quad + \frac{a_3}{2} Y + a_3 a_2 YX + a_3^2 Y^2 + a_3 a_4 YZ + \frac{a_4}{2} Z + a_4 a_2 ZX + a_4 a_3 ZY + a_4^2 Z^2 \end{aligned}$$

Now, using that  $\{\sigma_i, \sigma_j\} = 0$  for  $i, j \in \{1, 2, 3\}, i \neq j$  and that  $\sigma_i^2 = I$  for any  $i \in \{1, 2, 3\}$ , the above simplifies to:

$$\rho^2 = \left( \frac{1}{4} + a_2^2 + a_3^2 + a_4^2 \right) I + a_2 X + a_3 Y + a_4 Z$$

Taking the trace of  $\rho^2$  we have that:

$$\text{tr}(\rho^2) = \left( \frac{1}{4} + a_2^2 + a_3^2 + a_4^2 \right) \text{tr}(I) + a_2 \text{tr}(X) + a_3 \text{tr}(Y) + a_4 \text{tr}(Z)$$

From the previous exercise we know that  $\text{tr}(\rho^2) \leq 1$ , so:

$$2 \left( \frac{1}{4} + a_2^2 + a_3^2 + a_4^2 \right) \leq 1 \implies a_2^2 + a_3^2 + a_4^2 \leq \frac{1}{4} \implies \sqrt{a_2^2 + a_3^2 + a_4^2} \leq \frac{1}{2}$$

Hence we can write:

$$\rho = \frac{I + r_x X + r_y Y + r_z Z}{2} = \frac{I + r \cdot \sigma}{2}$$

with  $\|r\| \leq 1$ . The Bloch sphere representation for the state  $\rho = \frac{I}{2}$  is the above form with  $r = 0$ . This vector corresponds to the center of the Bloch sphere, which is a maximally mixed state  $\text{tr}(\rho^2)$  is minimized, with  $\text{tr}(\rho^2) = \frac{1}{2}$ .

For the previous calculations we know that for any  $\rho$

$$\text{tr}(\rho^2) = 2 \left( \frac{1 + r_x^2 + r_y^2 + r_z^2}{4} \right) = \frac{1 + r_x^2 + r_y^2 + r_z^2}{2}$$

if  $\|r\| = 1$ , then  $r_x^2 + r_y^2 + r_z^2 = 1$ . Hence,  $\text{tr}(\rho^2) = 1$  and  $\rho$  is pure. Conversely, suppose  $\rho$  is pure. Then,  $\text{tr}(\rho^2) = 1$ , so:

$$\frac{1 + r_x^2 + r_y^2 + r_z^2}{2} = 1 \implies r_x^2 + r_y^2 + r_z^2 = 1 \implies \|r\| = 1$$

For the states that lie on the surface of the Bloch sphere, which we parameterized as

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\varphi} \sin\left(\frac{\theta}{2}\right) |1\rangle$$

Calculating the density operator corresponding to  $|\psi\rangle$ , we have:

$$\rho = |\psi\rangle \langle\psi| =$$

# Bibliography

- [1] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.