INDIAN STATISTICAL INSTITUTE
KOLKATA

MASTER'S THESIS

Probabilistic and information theoretic interpretation of quantum measurement

---

Supervisor:
Prof. Subhamoy MAITRA

Author:
Shreyas GUPTA
CS1717

A thesis submitted in fulfillment of the requirements
for the degree of Master of Technology
in the
Computer Science

June 26, 2019

Declaration of Authorship

I, Shreyas GUPTA, declare that this thesis titled, "Probabilistic and information theoretic interpretation of quantum measurement" and the work presented in it, submitted to Indian Statistical Institute, Kolkata, is a bonafide record of the study carried out in the partial fulfilment for the award of the degree Master of Technology in Computer Science. I confirm that:

- No part of this thesis has previously been submitted for a degree or any other qualification at this University.
- I have acknowledged all main sources of help.

Signed:

Date:

Certificate

This is to certify that this thesis, "Probabilistic and information theoretic interpretation of quantum measurement" submitted by Shreyas GUPTAto Indian Statistical Institute, Kolkata, fulfills all the requirements of this institute

Signed:

Prof. Subhamoy Maitra
Professor (Higher Academic Grade)
Applied Statistics Unit
Indian Statistical Institute

"Cogito, ergo sum"

René Descartes

INDIAN STATISTICAL INSTITUTE KOLKATA

Abstract

Master of Technology

Probabilistic and information theoretic interpretation of quantum measurement

by Shreyas GUPTA

In this thesis, I have given a brief overview to the answers of two central questions in quantum computation. How much information can be encoded in quantum systems, and how efficiently can this information be extracted. Manipulation of quantum information through manipulating quantum states is relatively well studied topic. This is usually achieved by unitary transformations. In this thesis, after giving a brief overview of fundamentals of quantum computation, basic quantum information theory is briefly discussed. After that, the main question about how to estimate a quantum state has been looked into more carefully. Given a finite ensemble of a particular quantum state, say N copies, firstly a score is defined to measure how accurately the state is estimated. Then a bound for this score is calculate and is shown to be $\frac{N+1}{N+2}$, for a finite ensemble. In my work, with estimated state, I have tried to define quantum measurement in a more general fashion.

Acknowledgements

I would like to thank Prof. Subhamoy Maitra for helping me to think about this problem in a more structured manner. His suggestion about trying to estimate the state given n observations has been really helpful for to study in the direction of bounds on the amount of information that can be extracted from quantum states.

Preface

The main purpose of this thesis is to look more carefully at the question, how efficiently can information be extracted from a quantum states? I have very briefly reviewed the basics and postulates of quantum mechanics. One is assumed to be familiar with the basic theory of quantum computation. John Preskill's notes on the subject, especially the first 6 chapters, is good source to get acquainted with ths subject, which can be found here[1]. More advanced topics like POVM have been discussed in more detail.

In this thesis, I start with defining states and qubits. Postulates of quantum mechanics are discussed after that which dictate the laws that these states follow. After that, quantum measurement is discussed in detail and POVM is introduced. After a brief review of basic topics in quantum information theory, we turn to the main section on how accurately an unknown state can be estimated.

[1] http://www.theory.caltech.edu/people/preskill/ph229/

Contents

Chapter 1

Models of computation

1.1 Church Turing thesis and the nature of quantum model of computation

Quantum computers pose a serious challenge to the strong Church-Turing thesis, which states that every physically realizable computation device, can be simulated by a Turing machine with atmost polynomial slowdown. The famous Shor's algorithm is one example, where a quantum computer may be able to factorize numbers with exponential speed up as compared to any classical algorithm.

Very little knowledge of physics is needed to understand the central results of quantum computing. There are two results, which are more or less enough to explain the model. Firstly, the physical parameters of small enough systems seem to be quantized, i.e., they only take values from a discrete set. And secondly, the values of these physical parameters, at any moment in time, is not a unique value, but rather is a probability distribution on the set of all possible values. This distribution depends on the state of the system. The parameter only takes a definite value when it is measured by an observer. It is this last point that makes the quantum world interesting. Since the parameter dont have a unique value associated with it before observing, any changes done to the system has to give rise to a consistent probability distribution for the final value of the parameter. This results in the vast parallelism in the quantum computation. This vast parallelism is tightly regulated by the laws of quantum mechanics.

1.2 Qubits

Computation is essentially manipulation of information. This information is expressed, in case of turing machine, by an alphabet. This alphabet is often taken to be binary, consisting of only two letters, namely, 0 and 1. Since quantum computation is about manipulating state of a system, we will encode our information in these states of the system. The simplest type of system to consider again has two possible states, and is often called a two-level system. This state of a two level system is known as qubit, motivated from the classical case, where it is called a bit. But this qubit works in little different way. As we know, the state of the system is not uniquely determined, until it is observed. The state of the system is actually a probability distribution on the two possible state of the system that can be observed. This phenomena is known as superposition. Therefore, if we name our two possible states as $|0\rangle$ and $|1\rangle$, then a state of system is not simply any one of it, but a probability distribution over it.

1.3 Representation of qubit

A qubit is not usually just represented as a probabilistic function over the state space. This is due to many properties that can be measured for a system. All these properties can not usually be observed at the same time. Different properties of a system can have different probability distributions when the system is in a particular state. And we should be able to obtain all these probability distributions given the state. Hence we need a way to represent the state of a system, such that these probability distributions can be recovered from it. This can be achieved in a complex Hilbert space. A (pure) state of a quantum system is denoted by a vector $|\psi\rangle$ with unit length, i.e., $\langle\psi|\psi|\rangle = 1$ in a complex Hilbert space H. We call it a pure state to distinguish it from a more general type of quantum states, known as mixed states. Here it should be noted that $\langle\psi|$ is the dual vector of $|\psi\rangle$. Given vectors and the dual vectors, we can define operators in the form $O = |\psi\rangle \langle\phi|$.

1.4 Mixed states

The vectors, $|\psi\rangle$, can also be expressed projector operators that project onto the one dimensional space of the state vector. This is denoted by $\rho_\psi = |\psi\rangle\langle\psi|$, and is called the density matrix. Unless specified, the symbol ρ represents that we are talking about the physical state, rather than an arbitrary operator. Here the normalization condition becomes $tr(\rho) = 1$. These density matrices are used to define the more general type of states. A mixed state is a mixture of pure states. If ρ is a mixed state, it can be written as:

$$\rho = \sum_{k=1}^{N} p_k |\psi_k\rangle\langle\psi_k| \tag{1.1}$$

and the normalization condition means that $\sum\limits_{k=1}^{N} p_k = 1$. Intuitively, one may feel tempted to think that a mixed state represents a probabilistic pure state. But one must note that the summation form to represent the mixed state is not unique. To see this consider an ensemble of spin-$\frac{1}{2}$ particles, with half particles in the state $|0\rangle$ and other half in state, $|1\rangle$. Let us represent the density matrix of a particle, taken at random from this ensemble, by ρ. Then,

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \tag{1.2}$$

Now let us write this state in another basis. Let $|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$, and $|-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$. One can easily check that this forms another basis of the Hilbert space. Writing ρ, in this basis, we get

$$\rho = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -| \tag{1.3}$$

which gives a different probability distribution for the state ρ.

In next section, we will study the postulates of quantum mechanics, which describe how to represent a state of a system, how to modify the state of a system, and finally how to make an observation on a system in a particular state. With these postulates, we can manipulate the information encoded in a state of system or qubits.

1.5 Axioms of quantum mechanics

- Each physical quantum system is associated with a complex Hilbert space H with inner product $\langle\psi|\phi\rangle$. Rays (that is, subspaces of complex dimension 1) in H are associated with (pure) quantum states of the system. In other words, quantum states can be identified with equivalence classes of vectors of length 1 in H, where two vectors represent the same state if they differ only by a phase factor. A general (mixed) state is represented as described in the previous section.
- The Hilbert space of a composite system is the Hilbert space tensor product of the state spaces associated with the component systems. For a non-relativistic system consisting of a finite number of distinguishable particles, the component systems are the individual particles.
- Physical symmetries act on the Hilbert space of quantum states unitarily.
- Physical observables are represented by Hermitian matrices on H. A generalization of this will be discussed in the next chapter.
- The expectation value (in the sense of probability theory) of the observable A for the system in state represented by the unit vector $\psi \in H$ is $\langle \psi \mid A \mid \psi \rangle$
- By spectral theory, we can associate a probability measure to the values of A in any state ψ. We can also show that the possible values of the observable A in any state must belong to the spectrum of A. In the special case A has only discrete spectrum, the possible outcomes of measuring A are its eigenvalues. More precisely, if we represent the state ψ in the basis formed by the eigenvectors of A, then the square of the modulus of the component attached to a given eigenvector is the probability of observing its corresponding eigenvalue.
- More generally, a state can be represented by a so-called density operator, which is a trace class, nonnegative self-adjoint operator ρ

normalized to be of trace 1. The expected value of A in the state ρ is $\mathrm{tr}(A\rho)$

- If ρ_ψ is the orthogonal projector onto the one-dimensional subspace of H spanned by $\mid \psi\rangle$, then $\mathrm{tr}(A\rho_\psi) = \langle \psi \mid A \mid \psi\rangle$
- Density operators are those that are in the closure of the convex hull of the one-dimensional orthogonal projectors. Conversely, one-dimensional orthogonal projectors are extreme points of the set of density operators. Physicists also call one-dimensional orthogonal projectors pure states and other density operators mixed states.

Now that we know how states are represented and modified, let us study the next part, i.e., measurement in more detail.

Chapter 2

Quantum Measurement

2.1 What is quantum measurement operator?

Till now, almost all attention has been focused on discussing the state of a quantum system. As we have seen, this is most succinctly done by treating the package of information that defines a state as if it were a vector in an abstract Hilbert space. One of the most difficult and controversial problems in quantum mechanics is the so-called measurement problem. Opinions on the significance of this problem vary widely. At one extreme the attitude is that there is in fact no problem at all, while at the other extreme, the view is that the measurement problem is one of the great unsolved puzzles of quantum mechanics. The issue is that quantum mechanics only provides probabilities for the different possible out-comes in an experiment – it provides no mechanism by which the actual, finally observed result, comes about. Of course, probabilistic outcomes feature in many areas of classical physics as well, but in that case, probability enters the picture simply because there is insufficient information to make a definite prediction. In principle, that missing information is there to be found, it is just that accessing it may be a practical impossibility. In contrast, there is no 'missing information' for a quantum system, what we see is all that we can get, even in principle, though there are theories that say that this missing information resides in so-called 'hidden variables'. But in spite of these concerns about the measurement problem, there are some features of the measurement process that are commonly accepted as being essential parts of the final story.

2.2 Projective measurement

Let us begin to review the theory of measurement in quantum mechanics. In the formalism where a state is denoted by a unit vector in a Hilbert space, the very first notion of measurement was associated to a self adjoint operator in the Hilbert space and the outcome of measurement is associated with one of the eigenvalue of this hermitian operator. It is often referred to as a projective measurement, or von Nuemann measurement. A self adjoint operator can be represented as follows:

$$\hat{M} = \sum_{m} m\hat{P}_m \tag{2.1}$$

where $\{m\}$ are eigenvalues of $\hat{M}$, and $\{\hat{P}_m\}$ are the orthogonal[1] projector on the corresponding eigenspace. There are two reasons why $\hat{M}$ is chosen to be hermitian. All the eigenvalues of a hermitian operator are real and its eigenvectors form an orthonormal basis of the Hilbert space. So the following are true:

$$\forall m, \hat{P}_m \text{ is hermitian} \tag{2.2}$$

$$\forall m, m', \hat{P}_m\hat{P}_{m'} = \delta_{m,m'}\hat{P}_m \tag{2.3}$$

And the probability of measuring m, when the state of the system is $|\psi>$ is:

$$p_\psi(m) =< \psi|\hat{P}_m|\psi > \tag{2.4}$$

and the state of the system soon after measuring m is:

$$\frac{\hat{P}_m|\psi>}{\sqrt{<\psi|\hat{P}_m|\psi>}} \tag{2.5}$$

Applying total probability theorem to 2.4, we get:

$$\sum_{m} \hat{P}_M = \mathbb{I} \tag{2.6}$$

[1]Eigenvectors of a hermitian operator form an orthnormal basis of the corresponding vector space

The expectation value of $\hat{M}$, for the system in state $|\psi>$ is:

$$\begin{aligned} E_\psi(\hat{M}) &= \sum_m m p_\psi(m) \\ &= \sum_m <\psi|\hat{P}_m|\psi> \\ &= <\psi| \left(\sum_m m\hat{P}_m \right) |\psi> \\ &= <\psi|\hat{M}|\psi> \end{aligned}$$

2.3 POVM

The notion of measurement has been generalized, just as the notion of states. Here, we define the most general form of measurement used, positive operator valued measurement. We consider the set of operators $\{\hat{E}_m\}$. The above description of projective measurement turns out to be too restrictive. There are measurements that can be performed on a system that cannot be described within this formalism.
A generalised measurement in quantum mechanics is described by a collection of positive operators $E_i \geq 0$ that satisfy $\sum_i E_i = \mathbb{I}$. We denote such a measurement as $M = \{E_i\}$. Each E_i is associated with an outcome of the measurement and since $E_i \geq 0$, it has a decomposition $E_i = M_i^\dagger M_i$. For a state ρ, the probability of obtaining the result associated with E_i is:

$$Pr(i) = tr(\rho E_i) \tag{2.7}$$

And the state of the system, soon after obtaining result associated with E_i is:

$$\rho \rightarrow \frac{M_i \rho M_i^\dagger}{tr(\rho E_i)} \tag{2.8}$$

It should be noted here that the projective measurements defined above is a special case of POVM, where $M_i = M_i^\dagger = P_i$, and $E_i = P_i^\dagger P_i = P_i$. The only difference being that POVM elements do not have to be orthogonal. In the next chapter, we will see how these POVM's can be physically realized.

Chapter 3

Implementing POVM

3.1 Motivation

How do we actually implement POVMs? The physical realisation of POVMs is guaranteed be the Neumark's theorem. Neumrak's theorem states that the non orthogonality of the different outcomes of a POVM can be lifted by an operator of the form $V^{\dagger}(.)V$ to a projective measurement in a higher dimensional Hilbert space. Let us see this with the help of an example. Suppose M is a POVM, in three dimensional Hilbert space, with following E_i's:

$$\begin{aligned}
E_1 &= \frac{\sqrt{2}}{\sqrt{3}}|0\rangle + \frac{1}{\sqrt{3}}|2\rangle \\
E_2 &= -\frac{1}{\sqrt{6}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{3}}|2\rangle \\
E_3 &= -\frac{1}{\sqrt{6}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle - \frac{1}{\sqrt{3}}|2\rangle
\end{aligned}$$

And let us take our state $|\psi\rangle$ to be in the two dimensional subspace of this Hilbert space, spanned by $|0\rangle$ and $|1\rangle$.
The probability of the obtaining E_1 is

$$\left| \left(\frac{\sqrt{2}}{\sqrt{3}}\langle 0| + \frac{1}{\sqrt{3}}\langle 2| \right) |\psi\rangle \right|^2 \tag{3.1}$$

or

$$\left| \frac{\sqrt{2}}{\sqrt{3}} \langle 0|\psi\rangle \right|^2 \tag{3.2}$$

since $|0\rangle$ is orthogonl to $|2\rangle$. If we define the unnormalized quantum states:

$$\begin{aligned}
|e_1\rangle &= \sqrt{\frac{2}{3}}\,|0\rangle \\
|e_2\rangle &= -\sqrt{\frac{1}{6}}\,|0\rangle + \sqrt{\frac{1}{2}}\,|1\rangle \\
|e_3\rangle &= -\sqrt{\frac{1}{6}}\,|0\rangle - \sqrt{\frac{1}{2}}\,|1\rangle
\end{aligned} \tag{3.3}$$

Then we can immediately see that probability of obtaining the result corresponding to E_i is

$$|\langle e_i|\psi\rangle|^2 \tag{3.4}$$

3.2 Physical realisation of POVM

Suppose we have a number of these unnormalized vectors $|e_i\rangle$, when can the above rule for choosing probabilities form a measurement? A necessary condition is that the total probabilities add up to 1.

$$\sum_{i=1}^{k} |\langle e_i|\psi\rangle|^2 = 1 \tag{3.5}$$

$$\Rightarrow \sum_{i=1}^{k} \langle\psi|e_i\rangle\langle e_i|\psi\rangle = 1 \tag{3.6}$$

$$\Rightarrow \langle\psi| \left(\sum_{i=1}^{k} |e_i\rangle\langle e_i| \right) |\psi\rangle = 1 \tag{3.7}$$

It turns out that this is necessary and sufficient condition for a collection of un-normalised vectors $|e_i\rangle$ to be a POVM (with all elements of rank 1).
In the above discussion, we worked out in reverse direction by reducing the dimension of a higher dimensional Hilbert space. Next I will show that if we have a collection of such $|e_i\rangle$'s, we can achieve the above outcome probabilities by a projective measurement in higher dimensional Hilbert space.

Suppose we have k un-normalised quantum states $|e_i\rangle$ in n-dimensions, $k \geq n$, such that:

$$\sum_{i=1}^{k} |e_i\rangle \langle e_i| = \mathbb{I} \tag{3.8}$$

Therefore, we have a $k \times n$ matrix as follows:

$$M = \begin{bmatrix} \langle 1|e_1\rangle & \langle 1|e_2\rangle & \dots & \langle 1|e_k\rangle \\ \langle 2|e_1\rangle & \langle 2|e_2\rangle & \dots & \langle 2|e_k\rangle \\ \dots & \dots & \dots & \dots \\ \langle n|e_1\rangle & \langle n|e_2\rangle & \dots & \langle n|e_k\rangle \end{bmatrix} \tag{3.9}$$

Now consider the scalar product of row i and j:

$$\sum_{m=1}^{k} \langle i|e_m\rangle \langle e_m|j\rangle = \langle i| \left(\sum_{m=1}^{k} |e_m\rangle \langle e_m| \right) |j\rangle = \langle i|j\rangle = \delta_{i,j} \tag{3.10}$$

Therefore, we have n orthonormal rows in k dimensional Hilbert space. Using Gram-Schmidt, we can extend this to a set of k orthonormal rows. Now any square matrix whose rows are orthonormal is a unitary matrix, hence has orthonormal columns, therefore, the columns of this new $k \times k$ matrix correspond to a projective measurement. If this measurement is restricted to act on the n dimensional subspace given by the first n basis vectors, this becomes a POVM.

This procedure is usually realised by coupling the system to be measured with ancilla, and then doing projective measurement of the ancilla. To formally see this, consider system to be measured in a pure state $|\psi\rangle$. If this system is coupled with an ancilla, in state $|\phi\rangle$, and the whole system of state to be measured and the ancilla is evolved with U_{SA}, the resulting state of whole system is $U_{SA} |\psi\rangle_s \otimes |\phi\rangle_A$. Now assume we do a projective measurement of this system on the ancilla in the basis $|m_i\rangle_A \langle m_i|$, for $i = 1, 2, \dots$, the probability of getting result i is:

$$p_i = \left(\langle\psi|_S \otimes \langle\phi|_A \right) U_{SA}^{\dagger} \left[\mathbb{I} \otimes |m_i\rangle_A \langle m_i| \right] U_{SA} \left(|\psi\rangle_s \otimes |\phi\rangle_A \right) \tag{3.11}$$

And the state of the whole system, after recording the measurement of i on the ancilla, is:

$$|\psi\rangle_S \otimes |\phi\rangle_A \rightarrow \frac{(M_i |\psi\rangle_S) \otimes |m_i\rangle}{\sqrt{p_i}} \tag{3.12}$$

where M_i is the operator acting on the system to be measured only, that takes the form:

$$M_i |\psi\rangle_S \equiv \langle m_i| U_{SA} (|\psi\rangle_s \otimes |\phi\rangle_A) \tag{3.13}$$

This operator depends on $|\phi\rangle_A$, $|m_i\rangle$ and U_{SA}, and defines a generalised measurement on the system. Note that this generalised measurement can be tuned by choosing three things:

- The initial state of the ancilla, $|\phi\rangle_A$
- The unitary operation that couples the system and the ancilla, U_{SA}
- The basis that ancilla is measured in $\{|m_i\rangle\}$

Chapter 4

Information in quantum states

4.1 How many bits are in a qubit?

The next question we will try to answer is exactly how much accessible information can be encoded in a qubit. Suppose there are two parties, Alice and Bob, and Alice wants to send a string $x \in \{0,1\}^n$. Suppose Alice does some computation, on x, to create $|\psi_x\rangle \in \mathbb{C}^d$. Note that the scheme only works perfectly if $d \geq 2^n$, otherwise there will exist two non-orthogonal vectors $|\psi_x\rangle$ and $|\psi_y\rangle$, with $x \neq y$
To answer the question about how much information is actually transmitted, we need to be more careful. Firstly, we need to quantify how much Bob already knows about the string x. To do this, let us assume that Alice samples $x \in \{0,1\}^n$ with probability $p(x)$. Also, Alice need not send a pure state, but instead she could send a mixed state. Lastly, Bob need not perform a projective measurement, he could perform the more general POVM. To be specific, let us consider the following scenario:

- Alice samples $X \in \Sigma \subseteq \{0,1\}^n$, where $X = x$ with probability $p(x)$
- Alice sends $\sigma_X \in \mathbb{C}^d$
- Bob picks POVM $\{E_y\}_{y \in \Gamma}$, where $\Gamma \subseteq \{0,1\}^n$
- Bob measures σ_X, and receives output $Y \in \Gamma$, where $Y = y$ given $X = x$ with probability $tr(\sigma_x E_y)$

4.2 Accessible information

Now Bob wants to infer X from Y. Let us first see how much information is encoded in X. To measure this, we will use shannon entropy.

Definition *[shannon entropy] The shannon entropy of a random variable X, distributed on a set Σ is*

$$H(X) = -\sum_{x\in\Sigma} p(x)\log p(x)$$

where $p(x) = Pr[X = x]$

In general, if we have two random variables X and Y supported on the sets Σ and Γ respectively with joint distribution $p(x,y) = Pr[X = x, Y = y]$, we have:

$$H(X,Y) = -\sum_{x\in\Sigma, y\in\Gamma} p(X,Y)\log p(X,Y) \tag{4.1}$$

Here, it is easy to note that if the two random variables X and Y are independent, then $p(X,Y) = p(X)\times p(Y)$ and hence $H(X,Y) = H(X) + H(Y)$. And, if the two random variables are perfectly correlated, then $p(X,Y) = p(X) = p(Y)$ and hence $H(X,Y) = H(X) = H(Y)$

Definition *[Mutual information] The mutual information $I(X,Y)$ between two random variables X and Y is*

$$I(X,Y) = H(X) + H(Y) - H(X,Y)$$

This mutual information is the best Bob can do. This is the most amount of accessible information for Bob.

Definition *[Accessible information] The accessible information is*

$$I_{acc}(\sigma,p) = \max_{\textit{over all POVMs}\{E_y\}_{y\in\Gamma}} I(X,Y)$$

This the best Bob can do given Alice's choice of the σ_x. The overall best that both can achieve is:

$$\max_{\{\sigma_x\}_{x\in\Sigma}} I_{acc}(\sigma,p)$$

This is upper bounded by $H(X) \leq \log|\Sigma|$. Now we will relate this accessible information to the amount of quantum information in the σ's. Suppose we have a mixed state:

$$\begin{cases} |\psi_1\rangle \text{ with probability } p_1 \\ |\psi_2\rangle \text{ with probability } p_2 \\ \vdots \end{cases} \tag{4.2}$$

In this case, defining H by taking p to be distribution over $|\psi_i\rangle$'s is not well defined. This is because the representation of a mixed state is not unique. The correct analogue of classical entropy in this case was given by von Neumann, which is following

Definition *[Quantum Entropy] Given a mixed state, with density matrix ρ, we define:*

$$H(\rho) = -\sum_{i=1}^{d} \alpha_i \log \alpha_i$$

where α_i's are the eigenvalues of ρ. This can also be represented as:

$$H(\rho) = -tr(\rho \log \rho) \tag{4.3}$$

Chapter 5

Detection of quantum information

5.1 A general definition of measurement

Measurement, in quantum mechanics, is a function that takes a quantum state $|\psi\rangle$ and a measurement operator M as an input, and gives the outcome of M on ψ which is in accordance with the postulates of quantum mechanics.

$$f : B \times S \longrightarrow \{\lambda_m\} \tag{5.1}$$

where, B is the Bloch sphere, the state space in quantum mechanics, S is the space of measurement operators, and $\{\lambda_m\}$ is the set of eigen values of the measurement operator.
One section of this measurement function is defined in the postulates of quantum mechanics, namely, when $|\psi\rangle \in B$ is fixed.

$$f_{|\psi\rangle} : S \longrightarrow \{\lambda_m\} \tag{5.2}$$

Here, I will try to define the other section of this function, namely

$$f_M : B \longrightarrow \{\lambda_m\} \tag{5.3}$$

where $M \in S$. And hence I will try to define the function f.

5.2 Measurement for one qubit

In this section, I will consider the state space to be the space of one qubit space, namely the Bloch Sphere, and the space of measurement operators to be the space of self adjoint operators acting on Bloch Sphere.In this specific case of one dimensional Hilbert space of a qubit, space of measurement is relatively easy to define.

The measurement function is a probability density function, and it gives a probability of each outcome, i.e. the eigen value of the measurement operator. But according to the axioms of quantum mechanics, this probability density function only depends on the eigen vectors of the measurement operator and not on the specific eigen values.

Let $\sim$ be a relation between two measurement operators. I will call two measurement operators, M_1 and M_2, related by '$\sim$' if they have same set of eigen vectors.

Theorem:'$\sim$' is an equivalance relation.

Proof:1. Reflexive: M and M have same set of eigen vectors.

2. Symmetric: Let $M_1 \sim M_2$, therefore M_1 and M_2 have same set of eigen vectors. Hence $M_2 \sim M_1$.

3. Transitive: Let $M_! \sim M_2$ and $M_2 \sim M_3$. Therefore, M_1, M_2 have same set of eigen vectors and M_2, M_3 have same set of eigen vectors. Therefore, M_1, M_3 have same set of eigen vectors. Hence $M_1 \sim M_3$.

Therefore, $\sim$ partitions S, the space of measurement operators. For our purpose here, it is enough to study $S\sim$. Also note that for a one dimensional Hilbert space of single qubit system, each state vector has a unique orthogonal state vector. Therefore, to represent a measurement operator in $S\sim$, it is enough to know one of its eigen vector. We therefore have $S\sim$ is homeomorphic to B, the bloch sphere, i.e. $S\sim \cong B$.

So, to study measurement in general, I will take a general eigen state $|\psi\rangle$ and a general measurement operator M, and will define this probability density according to the postulates.

A general state on a Bloch sphere is as follows:

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle \tag{5.4}$$

Now note that in order to consider a general measurement operator, we only need to consider another general state on the Bloch sphere. Let M be a general measurement operator with one of the eigen state, $|m\rangle$ as follows:

$$|m\rangle = \cos\frac{\zeta}{2}|0\rangle + e^{i\xi}\sin\frac{\zeta}{2}|1\rangle \tag{5.5}$$

Also, Let λ_m be the eigen value of M corresponding to the eigen vector $|m\rangle$ of the measurement operator M. With these assumptions, the probability of getting the outcome λ_m is as follows:

$$\begin{aligned}
p(\lambda_m) &= ||\cos\frac{\theta}{2}\cos\frac{\zeta}{2} + e^{i(\phi-\xi)}\sin\frac{\theta}{2}\sin\frac{\zeta}{2}||^2 \\
&= ||\cos\frac{\theta}{2}\cos\frac{\zeta}{2} + \cos(\phi-\xi)\sin\frac{\theta}{2}\sin\frac{\zeta}{2} + i\sin(\phi-\xi)\sin\frac{\theta}{2}\sin\frac{\zeta}{2}||^2 \\
&= \cos^2\frac{\theta}{2}\sin^2\frac{\zeta}{2} + \cos^2(\phi-\xi)\sin^2\frac{\theta}{2}\sin^2\frac{\zeta}{2} + 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}\sin\frac{\zeta}{2}\cos\frac{\zeta}{2}\cos(\phi-\xi) \\
&\qquad + \sin^2(\phi-\xi)\sin^2\frac{\theta}{2}\sin^2\frac{\zeta}{2} \\
&= \cos^2\frac{\theta}{2}\cos^2\frac{\zeta}{2} + \sin^2\frac{\theta}{2}\sin^2\frac{\zeta}{2} + 2\sin\frac{\theta}{2}\cos\frac{\zeta}{2}\sin\frac{\zeta}{2}\cos\frac{\zeta}{2}\cos(\phi-\xi) \\
&= \cos^2\frac{\theta}{2}\cos^2\frac{\zeta}{2} + (1 - cos^2\frac{\theta}{2})\sin^2\frac{\zeta}{2} + \frac{1}{2}\sin\theta\sin\zeta\cos(\phi-\xi) \\
&= \sin^2\frac{\zeta}{2} + \cos^2\frac{\theta}{2}(\cos^2\frac{\zeta}{2} - \sin^2\frac{\zeta}{2}) + \frac{1}{2}\sin\theta\sin\zeta\cos(\phi-\xi) \\
&= \sin^2\frac{\zeta}{2} + \frac{\cos\zeta}{2} + \frac{\cos\zeta}{2}\cos\theta + \frac{1}{2}\sin\zeta\sin\theta\cos(\phi-\xi)
\end{aligned}$$

Therefore, we see that the measurement function that we defined earlier, is a function of 4 parameters:

$$f(\theta,\phi,\zeta,\xi) = \sin^2\frac{\zeta}{2} + \frac{\cos\zeta}{2} + \frac{\cos\zeta}{2}\cos\theta + \frac{1}{2}\sin\zeta\sin\theta\cos(\phi-\xi) \tag{5.6}$$

Now, given n-copies of an unknown state vector, we want to estimate the state with as much accuracy as possible by making n-observations to it. Here, the assumption will be that the measurement operator, and hence ζ and ξ are known, and we want to estimate θ and ϕ.The function f can be thought of as a surface and with each observation, we get the value of f at particular ζ_i and ξ_i. Given these n-points on the surface, we want to estimate the values of θ and ϕ.

5.3 State estimation

The question of how well the state, $|\psi\rangle$, of a physical system can be estimated is an important one. A state contains all the information about the quantum state and hence the probability distributions of any measurements can be calculated. But can we reconstruct a quantum state from a set of probability distributions?
It is definitely possible if we are given infinitely many identical copies of the system. Ours basic assumption in quantum mechanics is that if an infinite ensemble of identically prepared quantum states is given, then it can be determined exactly. But in practice, we never have such an infinite ensemble. Given a finite ensemble, the state can be known only approximately. How much knowledge can be extracted from a finite ensemble? What strategies furnish the maximum knowledge? We will try to answer this question in the rest of the thesis.

5.4 Ensemble of identically prepared states

It is well known that a composite system of non interacting particles can possess non local properties. A composite system can exhibit correlations which can not yet be explained by any theoretical model that involves only variables belonging to each subsystem separately.
Let us define a simple quantum game to formalize this problem. The game consists of many rounds. In each round, a player receives N qubits with same state. The player knows that the state of all N qubits is same. The player is allowed to do any measurement that he wants and is finally required to guess the state. The score of each round is $\cos^2(\alpha/2)$, where α is the angle between the original state and the guessed state. As the game has been defined, the score is the number between 0 and 1. If no measurement is performed and the state is measured randomly, the expected score obtained is 1/2. Therefore, the improvement over 1/2 represents the gain in information.
Let us denote the state of the system to be measured by $|\psi\rangle$, and we will follow the POVM measurement procedure, described in the previous

chapter. So, we will couple this system with an ancilla. Let the initial state of the ancilla be $|\phi\rangle$. Then,

$$|\psi\rangle_S |\phi\rangle_A \to U_{SA} |\psi\rangle_S |\phi\rangle_A \tag{5.7}$$

Let us describe the action of U_{SA}. Let $\{|i\rangle\}$ denote an orthonormal basis of the Hilbert space of system to be measured. Then, let the following denote the action of U_{SA} on this orthonormal basis:

$$|i\rangle |\phi\rangle_A \xrightarrow{U_{SA}} \sum_{i,f} |f\rangle \left|\phi_f^i\right\rangle_A \tag{5.8}$$

where, $\left|\phi_f^i\right\rangle = U_f^i |\phi\rangle$, when $U_{SA} = \sum\limits_{i,f}(|f\rangle \langle i|) \otimes U_f^i$. Therefore, we have:

$$|\psi\rangle_s |\phi\rangle_a \to \sum_{i,f} \langle i|\psi\rangle |f\rangle \left|\phi_f^i\right\rangle \tag{5.9}$$

Here, note that we have no restriction on the space of ancilla or the interaction. The wavefunctions $\left|\phi_f^i\right\rangle$ are not necessarily normalized, nor orthogonal. The only restriction they obey is:

$$\sum_f \left\langle \psi_f^i \middle| \psi_f^{i}{}' \right\rangle = \delta^{i,i'} \tag{5.10}$$

The next step is a projective measurement on ancilla. Let the orthonormal basis corresponding to this projective measurement on ancilla be $|m_\xi\rangle$. The probability of outcome being ξ, the eigenvalue corresponding to $|m_\xi\rangle$, when the initial state of the system to be measured is $|\psi\rangle$ is:

$$p_\xi(\psi) = \sum_{i,i',f} \langle\psi|i'\rangle \langle i|\psi\rangle \left\langle \phi_f^{i}{}' \middle| m_\xi \right\rangle \left\langle m_\xi \middle| \phi_f^i \right\rangle \tag{5.11}$$

When the observation of ξ has been recorded, some information is obtained about the state. This information could be expressed as a function $S(\xi, \psi)$. The average value of S is:

$$S = \sum_\xi \int \mathcal{D}\psi p_\xi(\psi) S(\xi, \psi) \tag{5.12}$$

We need to maximize S for the game we had defined initially. The combined Hilbert space of this combined system can be written as direct sum of different subspaces with total spin = N/2, N/2-1, ...A detailed discussion of this Hilbert space can be found in appendix. Now, we have n copies of same state, therefore it will belong to the subspace with the highest spin. So we will specify measuring interaction in this subspace. A basis of this subspace is $|m\rangle$, $m = -N/2, \ldots, N/2$. Let the unitary evolution of the N particles and the ancilla be specified as below:

$$|m\rangle\,|\phi\rangle \xrightarrow{U_{SA}} |v^m\rangle = \sum_{f=1}^{2^N} |f\rangle\,\left|\phi_f^m\right\rangle \tag{5.13}$$

Here $\{|f\rangle\}$ is the complete basis of the Hilbert space of n-qubits. The probability to obtain the result ξ is:

$$p_\xi(\psi) = \sum_{m,m'=-N/2}^{N/2} \sum_{f-1}^{2^N} \langle\psi|m\rangle \left\langle \phi_f^m \middle| e_\xi \right\rangle \times \left\langle e_\xi \middle| \phi_f^{m'} \right\rangle \langle m'|\psi\rangle \tag{5.14}$$

Now, after measuring the ancilla in state ξ, suppose we make a guess, $|\psi_\xi\rangle$, for the state, such that the score is $S(|\psi\rangle, |\psi_\xi\rangle)$. And, we know that $|\psi\rangle$ lives on Bloch sphere, hence, we can rewrite eqn (5.7) for this case as:

$$S_N = \sum_\xi \int \frac{\sin\theta d\theta d\phi}{4\pi} p_\xi(|\psi\rangle) S(|\psi\rangle, |\psi_\xi\rangle) \tag{5.15}$$

With $|\psi\rangle = \cos(\theta/2)\,|0\rangle +^{i\phi} \sin(\theta/2)\,|1\rangle$ and $|\psi_\xi\rangle = \cos(\theta_\xi/2)\,|0\rangle +^{i\phi_\xi} \sin(\theta_\xi/2)\,|1\rangle$. and the constaints:

$$\left\langle v^m \middle| v^{m'} \right\rangle = \sum_\xi \sum_{f=1}^{2^N} \left\langle \phi_f^m \middle| e_\xi \right\rangle \left\langle e_\xi \middle| \phi_f^{m'} \right\rangle = \delta^{mm'} \tag{5.16}$$

It convenient to work with only the following constraint:

$$\sum_{m=-N/2}^{N/2} \langle v^m|v^m\rangle = \sum_{m=-N/2}^{N/2} \sum_f \sum_\xi \left\langle \phi_f^m \middle| e_\xi \right\rangle \left\langle e_\xi \middle| \phi_f^m \right\rangle = N+1 \tag{5.17}$$

Upon adding to S_N the constraint in eqn (5.12) multiplied by the Lagrange multiplier λ and varying with respect to $\left\langle \phi_f^m \middle| e_\xi \right\rangle$, one obtains the following equation:

$$\sum_{m'} \left\langle e_\xi \middle| \phi_f^{m'} \right\rangle \left[M_{mm'}(\theta_\xi, \phi_\xi) - \lambda \delta_{mm'} \right] = 0 \tag{5.18}$$

where

$$M_{mm'}(\theta_\xi, \phi_\xi) = \int \frac{\sin\theta d\theta d\phi}{4\pi} \langle \psi | m \rangle \langle m' | \psi \rangle \times S(|\psi\rangle, |\psi_\xi\rangle) \tag{5.19}$$

Solving these equations, we get $S_{Nextreme} = \lambda(N+1)$. Eqn (4.13) has non-trivial solution only if λ is an eigenvalue of $M(\theta_\xi, \phi_\xi)$. It follows that eigenvalues of $M(\theta_\xi, \phi_\xi)$ are independent of θ_ξ, ϕ_ξ, with largest eigenvalue $\lambda = 1/(N+2)$. Therefore, the maximum score for this problem is $(N+1)/(N+2)$.

Appendix A

Combined Hilbert space of N spin-$\frac{1}{2}$ particles

A.1 Angular Momentum

Classically, the angular momentum vector $\vec{L}$ is defined by the cross product of of the position $\vec{r}$, and momentum $\vec{p}$

$$\vec{L} = \vec{r} \times \vec{p} \tag{A.1}$$

In quantum mechanics, for every observable there is an operator, we have the operators, $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ for each component of the angular momentum vector. But these operators do not commute with each other and hence can not be simultaneously observed.

$$[\hat{L}_i, \hat{L}_j] = i\hbar\varepsilon_{ijk}\hat{L}_k \tag{A.2}$$

The operator for the square of the magnitude of the angular momentum, $\hat{L}^2$ commutes with each operator for a particular component and hence $\hat{L}^2$ and one of the components, say z, $\hat{L}_z$ can be simultaneously observed. This means that the eigenbasis of $\hat{L}_z$ is also an eigenbasis of $\hat{L}^2$. The eigen values of $\hat{L}^2$ are $l(l+1)\hbar^2$ and eigenvalues of $\hat{L}_z$ are $m = -l, \ldots, l$, and let the common eigenvectors be represented by $|l, m\rangle$. The spin angular momentum is an intrinsic property, that is not due to motion in position space but it has been observed to follow same commutation relations. Let us denote the spin operators by S, then, for spin half particles, the following are true:

$$\hat{S}_x = \frac{\hbar}{2}\sigma_x \qquad \hat{S}_y = \frac{\hbar}{2}\sigma_y \qquad \hat{S}_z = \frac{\hbar}{2}\sigma_z \tag{A.3}$$

where σ's are the Pauli matrices. Similarly, we will denote the common eigenvectors of $\hat{S}^2$ and $\hat{S}_z$ by $|s, m_s\rangle$. Now suppose we have N spin-$\frac{1}{2}$ particles.

A.2 Hilbert space of combined system

We would like to think how we could describe our system of N particles, each of spin $\frac{1}{2}$, in way that emphasizes the composite system rather than the individual particles. For each individual subsystem, we have a total two possible states, for $m_s = \pm\frac{1}{2}$, therefore for N particles, we have 2^N such states. Let us see this with an example. Let $N = 2$. Then the basis vectors of the four dimensional product space in the uncoupled representation are $\left|m_{s_1} = \frac{1}{2}\right\rangle \otimes \left|m_{s_2} = \frac{1}{2}\right\rangle$, $\left|m_{s_1} = \frac{1}{2}\right\rangle \otimes \left|m_{s_2} = -\frac{1}{2}\right\rangle$, $\left|m_{s_1} = -\frac{1}{2}\right\rangle \otimes \left|m_{s_2} = \frac{1}{2}\right\rangle$ and $\left|m_{s_1} = -\frac{1}{2}\right\rangle \otimes \left|m_{s_2} = -\frac{1}{2}\right\rangle$. On the other hand, if we can use the coupled representation and find the total spin quantum number of the two particles together, then s is either $\frac{1}{2} + \frac{1}{2} = 1$ or $\frac{1}{2} - \frac{1}{2} = 0$. When the total spin quantum number is 1, the quantum number m_s can be $-1, 0, 1$. And when the total spin quantum number is 0, m_s can only be 0. Therefore the basis for the system in this coupled representation is $|s = 1, m_s = -1\rangle$, $|s = 1, m_s = 0\rangle$, $|s = 1, m_s = 1\rangle$ and $|s = 0, m_s = 0\rangle$. Therefore, if $\mathcal{H}^n$ denote the Hilbert space for spin-n particle, then $\mathcal{H}^{n_1} \otimes \mathcal{H}^{n_2} = \bigoplus_{k=|n_1-n_2|}^{n_1+n_2} \mathcal{H}^k$

Bibliography

Derka, R.ek, family$_i$ $= u., given = V., given_i = V.,,,,, And A.K. Ekert$ (1998). “Universal Algorithm for Optimal Estimation of Quantum States from Finite Ensembles via Realizable Generalized Measurement”. In: *Phys. Rev. Lett.* 80 (8), pp. 1571–1575. DOI: 10.1103/PhysRevLett.80.1571. URL: https://link.aps.org/doi/10.1103/PhysRevLett.80.1571.

Kubitzki, Marcus (2003). “State and Parameter Estimation in Quantum Theory”. MA thesis.

Latorre, J. I., P. Pascual, and R. Tarrach (1998). “Minimal Optimal Generalized Quantum Measurements”. In: *Phys. Rev. Lett.* 81 (7), pp. 1351–1354. DOI: 10.1103/PhysRevLett.81.1351. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1351.

Massar, S. and S. Popescu (1995). “Optimal Extraction of Information from Finite Quantum Ensembles”. In: *Phys. Rev. Lett.* 74 (8), pp. 1259–1263. DOI: 10.1103/PhysRevLett.74.1259. URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.1259.

Peres, Asher (1990). “Neumark’s theorem and quantum inseparability”. In: *Foundations of Physics* 20.12, pp. 1441–1453. ISSN: 1572-9516. DOI: 10.1007/BF01883517. URL: https://doi.org/10.1007/BF01883517.

– (1993). *Quantum theory : concepts and methods*. English. "Reprinted with corrections" [1995?]–Verso t.p. paperback reprint. Dordrecht ; Boston : Kluwer Academic. ISBN: 0792325494 (acid-free paper).

Peres, Asher and William K. Wootters (1991). “Optimal detection of quantum information”. In: *Phys. Rev. Lett.* 66 (9), pp. 1119–1122. DOI: 10.1103/PhysRevLett.66.1119. URL: https://link.aps.org/doi/10.1103/PhysRevLett.66.1119.

Index