Quantum Information and Computing

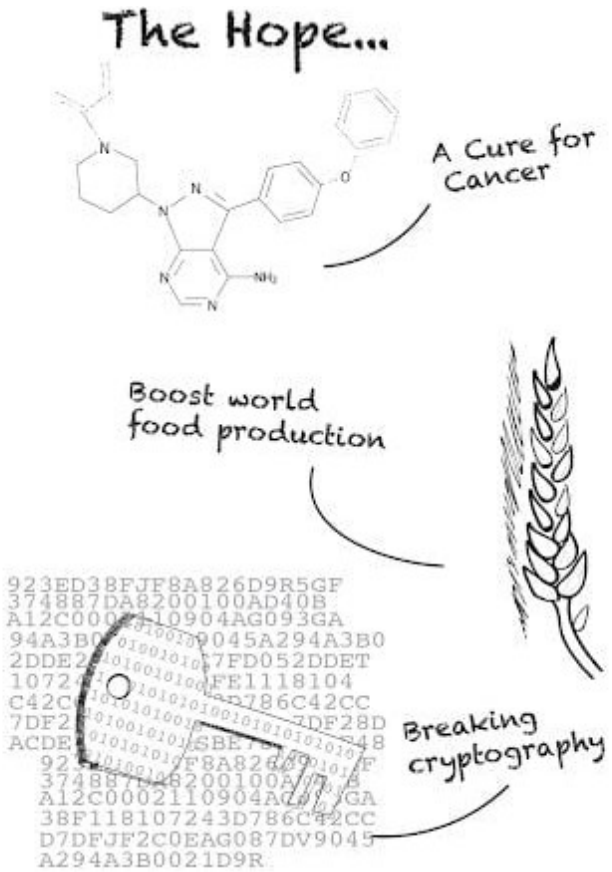
SoS Endterm Report-2022

Simran Sinha
Mentor: Aryaman Mihir Seth, Aditya Sriram
Indian Institute of Technology Bombay

Contents

Introduction

Modern Computing is based on laws of classical physics and mathematical logic. Traditional computers are designed for serial computation which means that for an algorithm, the logic flow takes place from point to other in terms of time, i.e. a particular process must be completed before another process is taken up. A quantum computer on the other hand exhibits inherent parallelism, that is, the same processor can perform operations on multiple inputs simultaneously. The state of a register can exist in a simultaneous superposition of different quantum states resulting in faster algorithms. Ultimately, quantum computers have the potential to provide computational power on a scale that traditional computers cannot ever match. Quantum computers will revolutionize drug research, material discovery and artificial intelligence by solving problems in a new way. To understand this, Simran Sinha under the guidance of mentors Aryaman Mihir Seth and Aditya Sriram tries to make a small attempt to review how quantum computers would work and solve problems.



Chapter 1

Postulates of Quantum Mechanics

An important distinction needs to be made between quantum mechanics, quantum physics and quantum computing. Quantum mechanics is a mathematical language, much like calculus. Just as classical physics uses calculus to explain nature, quantum physics uses quantum mechanics to explain nature. Just as classical computers can be thought of in boolean algebra terms, quantum computers are reasoned about with quantum mechanics. There are four postulates to quantum mechanics, which will form the basis of quantum computers:

1.1 Postulate 1: A Quantum Bit

(Nielsen and Chuang, page 80):
"Associated to any *isolated* physical system is a complex vector space with inner product (i.e. a Hilbert space) known as the state space of the system. The system is completely described by its state vector, which is a *unit vector* in the system's state space."

1.2 Postulate 2: Evolution of Quantum Systems

(Nielsen and Chuang, page 81):
"The evolution of a *closed* quantum system is described by a *unitary trans-*

formation. That is, the state $|\Psi\rangle$ of the system at time t_1 is related to the state of $|\Psi'\rangle$ of the system at time t_2 by a unitary operator U which depends only on times t_1 and t_2."

1.3 Postulate 3: Measurement

(Nielsen and Chuang, page 84):
"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 $|\Psi\rangle$ immediately before the measurement then the probability that result m occurs is given by:

$$p(m) = \langle\Psi|M_m^\dagger M_m|\Psi\rangle$$

and the state of the system after measurement is:

$$\frac{M_m|\Psi\rangle}{\sqrt{\langle\Psi|M_m^\dagger M_m|\Psi\rangle}}$$

The measurement operators satisfy the *completeness equation*:

$$\sum_m \langle\Psi|M_m^\dagger M_m|\Psi\rangle = I$$

The completeness equation expresses the fact that probabilities sum to one:

$$1 = \sum_m p(m) = \sum_m \langle\Psi|M_m^\dagger M_m|\Psi\rangle$$ "

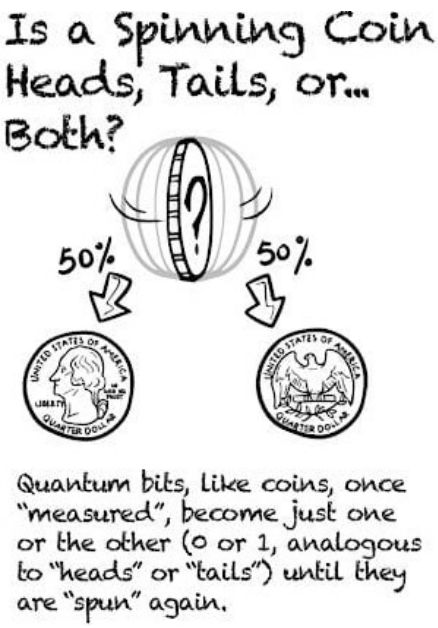
*[To be contd. in chapter6]

1.4 Postulate 4: Multi-qubit Systems

(Nielsen and Chuang, page 94):
"The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. [sic] e.g. suppose systems 1 through n and system i is in state $|\Psi_i\rangle$, then the joint state of the total system is $|\Psi_1\rangle \otimes |\Psi_1\rangle \otimes ... \otimes |\Psi_1\rangle$."

Chapter 2

Quantum States and Gates



A spinning coin is analogous to the behavior of a qubit in a superposition

2.1 Qubit-The smallest unit

- Any physical state (ray) can be associated with a direction in $\vec{r} = (x, y, z)$ in space for which $S_r = +\frac{1}{2}$.

- For spin $\frac{1}{2}$ representation of a qubit

$$\vec{S} = S_x\hat{\imath} + S_y\hat{\jmath} + S_z\hat{k}$$

$$S_x = \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = \frac{\hbar}{2}\sigma_x,\ S_y = \frac{\hbar}{2}\begin{pmatrix}0 & -\iota\\ \iota & 0\end{pmatrix} = \frac{\hbar}{2}\sigma_y,\ S_z = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = \frac{\hbar}{2}\sigma_z$$

where $\sigma_x, \sigma_y, \sigma_z$ are Pauli Matrices

2.2 Qubit-Bloch Sphere representation

- Pauli Matrix along direction

$$\hat{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

- $\sigma_n = \begin{pmatrix}\cos\theta & e^{-\iota\theta}\sin\theta\\ e^{\iota\theta}\sin\theta & -\cos\theta\end{pmatrix}$ has eigenvalue $\lambda = \pm 1$. The eigenvector for $\lambda = 1$ is $\begin{pmatrix}\cos\frac{\theta}{2}\\ e^{\iota\phi}\sin\frac{\theta}{2}\end{pmatrix}$

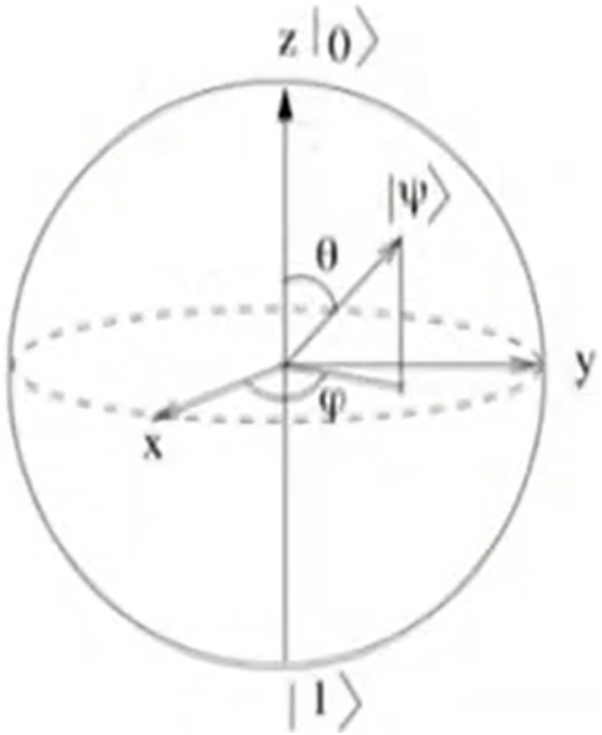


Fig: Bloch-Sphere

- All(pure) states lie on the surface of a unit sphere, position corresponding to (θ, ϕ)
- $|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{\iota\phi}\sin\frac{\theta}{2}|1\rangle$
- North Pole: $\theta = \phi == 0$: the state $|0\rangle$
- Sourh Pole: $\theta = \pi, \phi = 0$: the state $|1\rangle$
- Point where x-axis meets equator: $\theta = \frac{\pi}{2},\ \phi = 0$: the state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$

Qubit-How much information

- A qubit may contain infinite amount of information since the binary expansion of θ and ϕ may not be terminating. However, a measurement of the qubit will make the state collapse to $|\rangle\, 0$ or $|1\rangle$.
- Expectation value of $\langle\sigma_z\rangle$

$$\begin{pmatrix}\cos\frac{\theta}{2} & e^{\iota\phi}\sin\frac{\theta}{2}\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\cos\frac{\theta}{2}\\ e^{\iota\theta}\sin\frac{\theta}{2}\end{pmatrix} = \cos\theta$$

- We may determine $\langle\sigma_x\rangle$ similarly. These will determine $n_z,\ n_x$ but the sign of n_y remains undetermined.
- Determination of expectation value requires repeated measurements on identically prepared copies of the system.

specifically meant to visualise single qubit statevectors only while the Q-spheres are used to visualise combined states of multi-qubit circuits

2.3 Single Qubit States

- Quantum gates are analogous to classical logic gates, except that they must be implemented unitarily (and thereby reversibly)
- In classical computing, the only single bit gate is a NOT gate $0\leftrightarrow1$, which is reversible
- According to quantum postulate, quantum states evolve unitarily. The operator which can act on a single qubit state has to be unitary $U\dagger U = I$, which preserves norm

Unitary Operations on a single qubit

- Each state on Bloch sphere goes to another point on the Bloch Sphere
- Operations are rigid body rotations and reflections
- Matrix representation must be 2×2
- Any 2×2 unitary matrix can be written as

$$U = e^{\iota\alpha} exp(-\iota\theta\hat{n} \cdot \vec{\sigma}/2)$$

- Take $U = aI + b\sigma_x + c\sigma_y + d\sigma_z$

$$U^\dagger = a^*I + b^*\sigma_x + c^*\sigma_y + d^*\sigma_z$$

$UU^\dagger = I$ gives on equating coefficients

$$|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$$

$$ab^* + ba^* + \iota(cd^* - c^*d) = 0 \quad etc.$$

Choose, $a = e^{\iota\alpha}\cos\frac{\theta}{2}, \quad b = -\iota e^{\iota\alpha}\sin\frac{\theta}{2}n_x, \quad c = -\iota e^{\iota\alpha}\sin\frac{\theta}{2}n_y$

2.3.1 Unitary Gates

- NOT gate:(also called X-gate)

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

- Phase gate(A NOT gate in diagonal basis)

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{\iota\phi} \end{pmatrix}$$

$$Z(\phi = \pi) : \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$T - gate(\phi = \frac{\pi}{4}) : \begin{pmatrix} 1 & 0 \\ 0 & e^{\iota\frac{\pi}{4}} \end{pmatrix}$$

2.4 Single Qubit Gates

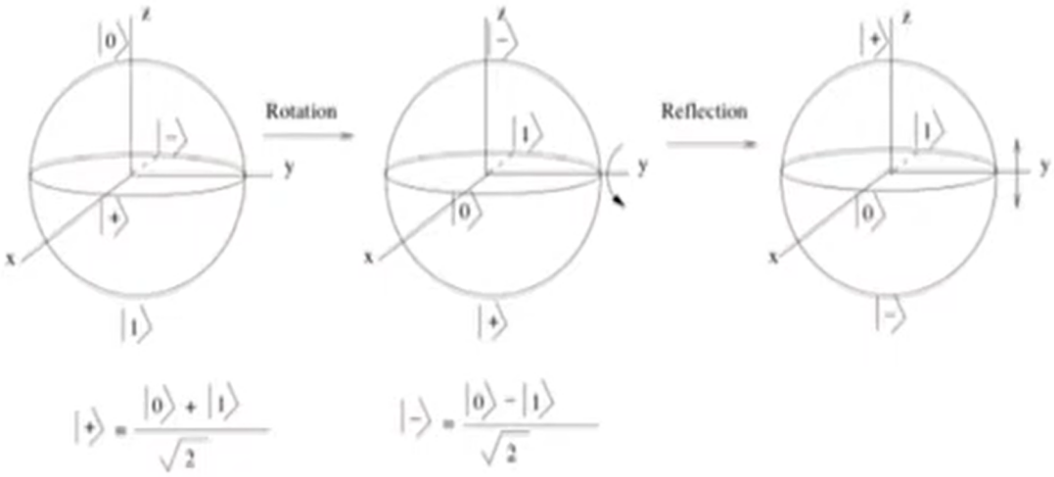
These gates can be applied on single qubits. Some important gates are as follows:

Pauli X-Y-Z gates

The X-gate does bit-flip ($|0\rangle$ to $|1\rangle$ and $|1\rangle$ to $|0\rangle$). The Y-gate is responsible for bit as well as phase flip ($|0\rangle$ to $-|1\rangle$). The Z-gate only does phase flip ($|0\rangle$ to $-|0\rangle$). Their corresponding matrices are as follows:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -\iota \\ \iota & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Hadamard Gate



Effect of Hadamard Gate is (1) Rotation by 90 about y–axis followed by (2) Reflection in x–y plane

A Hadamard gate represents a rotation of Π about the axis that is in the middle of the X-axis and Z-axis. It maps the basis state $|0\rangle$ to $(\frac{|0\rangle+|1\rangle}{\sqrt{2}})$, which means that a measurement will have equal probabilities of being in $|1\rangle$ or $|0\rangle$, creating a 'superposition' of states. This state is also written as $|+\rangle$. Hence, Hadamard gate is known to put any qubit in an equal superposition of two states. In this state, the qubit can be observed to have equal amplitudes of both states. The corresponding matrix of the Hadamard gate is: $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

Chapter 3

Multiple Qubit States and Entanglement

3.1 Introduction

Composite State
Orthonormality:

$$_{AB}\langle\alpha^{'},\beta^{'}|\alpha,\beta\rangle_{AB} = \delta_{\alpha,\alpha'}\delta_{\beta\beta'}$$

Define an operator in this composite space

$$\begin{aligned} M_A \otimes N_B \left|\psi,\phi\right\rangle &= M_A \left|\psi\right\rangle \otimes N_B \left|\phi\right\rangle \\ &= \sum_{\alpha,\beta} (M_A)_{\psi,\alpha} \left|\alpha\right\rangle (N_B)_{\phi,\beta} \left|\beta\right\rangle \end{aligned}$$

3.2 Two-Qubit States

- Classical 2-bits are: 00, 01, 10, 11
- Quantum 2 qubit state is a linear superposition

$$\alpha_{00}\left|00\right\rangle + \alpha_{01}\left|01\right\rangle + \alpha_{10}\left|10\right\rangle + \alpha_{11}\left|11\right\rangle$$

  Normalized

$$|\alpha_{00}|^2 + |\alpha_{01}|^2 + |\alpha_{10}|^2 + |\alpha_{11}|^2 = 1$$

3.2.1 CNOT Gate

The CNOT (or CX) gate is an important two-qubit gate that plays an important role in establishing entanglement in quantum experiments. This gate is a conditional gate that performs an X-gate on the second qubit(target), if the state of the first qubit (control) is $|1\rangle$ CNOT is short for Controlled-NOT gate. Its matrix is given by: $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Notice that we now use two qubits (q_0 and q_1) to invoke a quantum circuit, and CX gate takes two parameters (0,1). This means that the gate is controlled on q_0 (control qubit) and is applied on q_1 (target bit). It works according to the state value of q_0; if $q_0 = |0\rangle$, then there will be no change in q_1's state, but if q_0 is in state $|1\rangle$, then q_1's state will be flipped.

3.2.2 CZ Gate

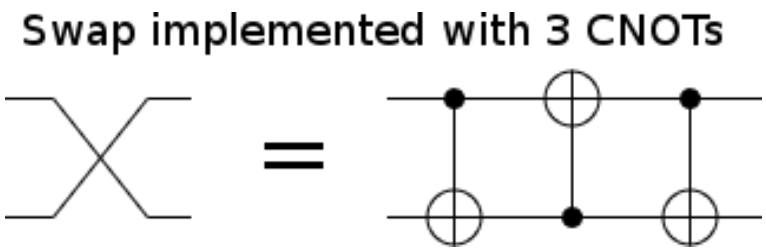
The job of the CZ (Controlled-Z) gate is to apply phase flip to the target qubit according to the state pf control qubit. The matric for this gate is as follows:

$$CZ = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

Case-1: When input is $|00\rangle$*:*

Case-2: When the control qubit (q_0) *is in state* $|1\rangle$*:*

3.2.3 Swap gate



Exchange two states without entanglement

$$U_S |\psi, \phi\rangle = U_S(|\psi\rangle \otimes |\phi\rangle) = |\phi, \psi\rangle$$

$$U_S = |00\rangle \langle 00| + |01\rangle \langle 01| + |10\rangle \langle 10| + |11\rangle \langle 11|$$

3.2.4 Controlled swap(Fredkin)

3.3 Three-Qubit States

There are gates that take three qubits to be applied and operate according to their states.

3.3.1 CCNOT (Toffoli) gate

The CCNOT gate is a three qubit gate that is controlled on two qubits and has one target qubit. It flips the target qubit based on the states of two controlled qubits.

3.4 Entanglement

Entanglement in 2 qubit states

- Not all two qubit states can be written as product of two single qubit states.
- Bell States:

$$|\psi_+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}} \qquad |\psi_-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

$$|\phi_+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \qquad |\phi_-\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}$$

Measurement in computational basis

- Computational basis: $(|0\rangle, |1\rangle)$. If one measures $\alpha |0\rangle + \beta |1\rangle$ in this basis, one can get information about $|\alpha|^2$ and $|\beta|^2$ but not about relative phase
- A measurement in diagonal basis can provide information about realtive phase as well

Quantum Entanglement
Quantum entanglement is the state where two systems are so strongly correlated that gaining information about one system will give immediate information about the other no matter how far apart these systems are. It is a quantum mechanical phenomenon in which the quantum states of two or more objects have to be described with reference to each other, even though the individual objects may be spatially separarted. This leads to correlations between observable physical properties of the systems. It is a property of a multi-qubit state space (multi-qubit system) and can be thought of as a resource. To explain entanglement we'll examine the creation and destruction of an EPR pair of qubits named after Einstein, Podolsky, and Rosen.

The phenomena of quantum entanglement comes useful to cut down on the time and computing power to process information transfer between qubits. Entanglement enables tasks such as quantum cryptography, superdense coding, and teleportation.

How is entanglement used in quantum computing? In quantum computers, changing the state of an entangled qubit will change the state of the paired qubit immediately. Therefore, entanglement improves the processing speed of quantum computers. Doubling the number of qubits will not necessarily double the number of processes since processing one qubit will reveal information about multiple qubits (i.e. the entangled qubits). According to research, quantum entanglement is necessary for a quantum algorithm to offer an exponential speed-up over classical computations.

Chapter 4

Quantum Protocols and Quantum Circuits

4.1 Quantum Circuit

A quantum circuit is a computational routine consisting of coherent quantum operations on quantum data, such as qubits, and concurrent real-time classical computation. It is an ordered sequence of quantum gates, measurements and resets, all of which may be conditioned on and use data from the real-time classical computation.

A set of quantum gates is said to be universal if any unitary transformation of the quantum data can be efficiently approximated arbitrarily well as a sequence of gates in the set. Any quantum program can be represented by a sequence of quantum circuits and non-concurrent classical computation.

4.2 Quantum No-Cloning Theorem

In classical computation, it is quite easy to copy information, even in a reversible manner. This leads to a natural question: given a qubit $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$, is it possible to copy $|\Psi\rangle$ so that two *unentangled* qubits each have state $|\Psi\rangle$? This would require the two qubits to be in a joint state of $|\Psi\rangle \otimes |\Psi\rangle$. To check, before the CNOT gate, the qubits are in the joint state of $(\alpha|0\rangle + \beta|1\rangle) \otimes (|0\rangle) = \alpha|00\rangle + \beta|10\rangle$. After the CNOT gate, the qubit are then in the joint state of $\alpha|00\rangle + \beta|11\rangle$. Unfortunately, this is not equal

to our desired to our desired state of

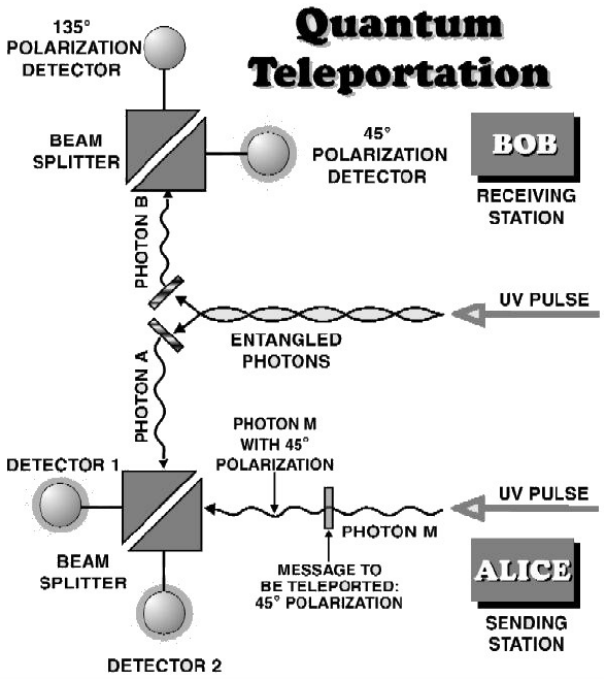
$$|\Psi\rangle \otimes |\Psi\rangle = \alpha^2 |00\rangle + \alpha\beta |01\rangle + \alpha\beta |10\rangle + \beta^2 |11\rangle$$

unless $\alpha = 1$ and $\beta = 0$ or $\alpha = 0$ and $\beta = 1$. that is, the circuit described above fails to copy an arbitrary quantum state into two unentangled copies. It turns out that there does not exist any quantum circuit which can take in an arbitrary qubit $|\Psi\rangle$ and produce the state $|\Psi\rangle \otimes |\Psi\rangle$, even when permitted to use ancillas in the input and garbage in the output. This result is called the *No-cloning Theorem*.

THe quantum No cloning Theorem says, it is not possible to find a unitary transformation which will duplicate a given state and write it on to a blank state. Basically, the xeroxing of a quantum state is not possible unless either the two states are identical or they are orthogonal, they cannot be arbitrary states.

The no cloning theorem is a result of quantum mechanics which forbids the creation of identical copies of an arbitrary unknown quantum state. If $|\Phi\rangle$ and $|\Psi\rangle$ are not orthogonal, then no measurement perfectly distinguishes them, and we always have some constant probability of error. However, if we could make many copies of the unknown state, then we could repeat the optimal measurement many times, and make the probability of error arbitrarily small. The no cloning theorem says that this isn't physically possible. Only sets of mutually orthogonal states can be copied by a single unitary operator.
(Can be proved from the norm preserving property of the inner product, or from the linearity of quantum mechanics)

4.3 Teleportation



Quantum teleportation is the process of exchanging quantum information such as photons, atoms, electrons, and superconducting circuits between two parties. Research suggests that teleportation allows QCs to work in parallel and use less electricity reducing the power consumption up to 100 to 1000 times.

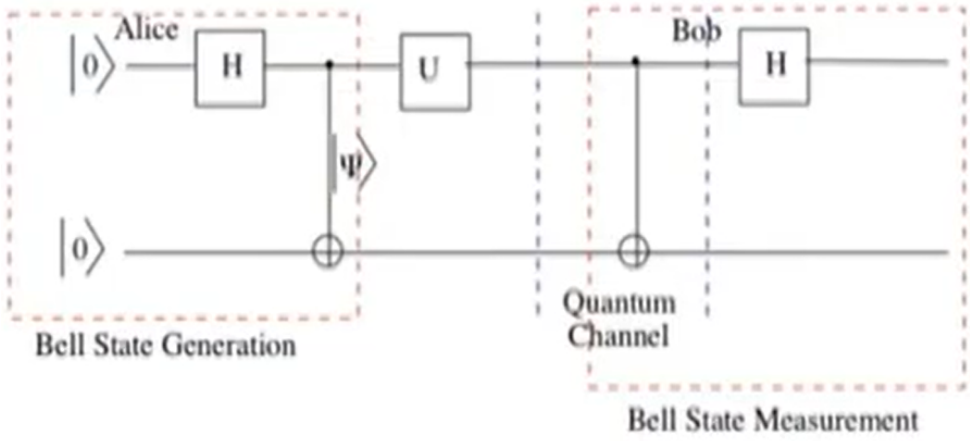
4.4 Super dense Coding

Reverse process of teleportation-ability to transmit to bits of classical information by transmitting one qubit of quantum information. Entangled state is used to achieve this. Starting point is one of the Bell states.
In simple words, superdense coding is the process of transporting 2 classical bits of information using 1 entangled qubit. Superdense coding can:

- Allow user to send ahead of time half of what will be needed to reconstruct a classical message ahead of time, which let's the user transmit at double speed until the pre-delivered qubits run out.
- Convert high-latency bandwidth into low-latency bandwidth by sending half of the information over the high latency channel to support the

information coming over the low latency channel.

- Double classical capacity in one direction of a two-way quantum channel (e.g. converting a 2-way quantum channel with bandwidth B (in both directions) into a one-way classical channel with bandwidth 2B).



Imagine a situation where two people (named Alice and Bob) are in different parts of the world. Alice has two bits: a and b. She would like to communicate these two bits to Bob by sending him just a single qubit. It turns out that there is no way they can can accomplish this task without additional resources. This is not obvious, but it is true—Alice cannot encode two classical bits into a single qubit in any way that would give Bob more than just one bit of information about the pair (a, b).

However, let us imaging that a long time ago, before Alice even knew what a and b are, that the two of them prepared two qubits A and B in the superposition

$$\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$$

Alice took the qubit A and Bob took the qubit B. We say that Alice and Bob share an e-bit or share an EPR pair in this situation. It is natural to view entanglement as a resource as we will see; and when Alice and Bob each have a qubit and the two of them are in the state above, it is natural to imagine that Alice and Bob share one unit (i.e., one entangled bit, or e-bit for short) of entanglement. Given the additional resource of a shared e-bit of entanglement, Alice will be able to transmit both a and b to Bob by sending just one qubit. Here is how:

Superdense coding protocol

1. If $a = 1$, Alice applies the unitary transformation

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

to the qubit A. (If $a = 0$ she does not.)
2. If $b = 1$, Alice applies the unitary transformation

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

to the qubit A. (If $b = 0$ she does not.)
3. Alice sends the qubit A to Bob. (This is the only qubit that is sent during the protocol.)
4. Bob applies a controlled-NOT operation to the pair (A, B), where A is the control and B is the target. The corresponding unitary matrix is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

5. Bob applies a Hadamard transform to A.
6. Bob measures both qubits A and B. The output will be (a, b) with certainty. To see that the protocol works correctly, we simply compute the state of (A, B) after the steps involving transformations:

4.5 Difference between Superdense coding and Quantum Teleportation

Quantum teleportation and superdense coding are closely related.
Quantum teleportation is a process by which the state of qubit ($|\Psi\rangle$) can be transmitted from one location to another, using two bits of classical communication and a Bell pair. In other words, we can say it is a protocol that destroys the quantum state of a qubit in one location and recreates it on a qubit at a distant location, with the help of shared entanglement.
Superdense coding is a procedure that allows someone to send two classical bits to another party using just a single qubit of communication.

The teleportation protocol can be thought of as a flipped version of the superdense coding protocol, in the sense that Alice and Bob merely "swap their equipment."

Chapter 5

Mixed States and Density Matrix

The postulates of quantum mechanics described earlier are valid for a pure, isolated quantum system. The systems that we study are usually ensembles, which are much bigger in which the quantum postulates (1)rays representing a state, (2)measurements being orthogonal projections, (3)Unitary evolution may not remain valid.

5.1 Quantum Mechanics of an Ensemble

Consider a 2-d system with the basis states $|x\rangle, |y\rangle$. Let the ensemble have only two types of states in it: $|\psi\rangle = \alpha|x\rangle + \beta|y\rangle$ with a probability p, and $|\phi\rangle = \gamma|x\rangle + \delta|y\rangle$ with a probability 1-p. If we measure the system in a computational basis, what would we get? When we pick up at random, if we happen to have got the state $|\psi\rangle$ then probability that the state collapses to $|x\rangle$ is given by $p|\alpha|^2$ but if happen to have picked up the state $|\phi\rangle$ then the probability of getting state $|x\rangle$ is $(1-p)|\gamma|^2$, so this is the product of two types of probability. p and (1-p) are classical probabilities and $|\alpha|^2$ and $|\gamma|^2$ are Born probability. Probability of state collapsing to state $|x\rangle$

$$p|\alpha|^2 + (1-p)|\gamma|^2$$

Probability of state collapsing to state $|y\rangle$

$$p|\beta|^2 + (1-p)|\delta|^2$$

So when we consider a general system which does not contain only one type of states then this is the extra considerations that comes in.

5.2 A Composite System-effect of environment coupling

Consider a composite system in $\mathcal{H}_A \otimes \mathcal{H}_B$.
If $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, then an operator O_A which only acts on the system A gives $\langle\psi_{AB}|O_A|\psi_{AB}\rangle = \langle\psi_A|O_A|\psi_A\rangle$ and system behaves like a pure state. If the system is entangled this factorization does not work and the combined system needs to be used to extract information about A. The system is said to be in a mixed state.

Let $|\psi\rangle_{AB} = a\,|0\rangle_A \otimes |0\rangle_B + b\,|1\rangle_A \otimes |1\rangle_B$. Consider an operator $M_A \otimes I_B$ which is a general measurement operator on subsystem A. What is $\langle\psi_{AB}|M_A \otimes I_B|\psi_{AB}\rangle$?

$$\langle\psi_{AB}|M_A \otimes I_B|\psi_{AB}\rangle = [a^*\,|0_A\rangle\,|0_B\rangle + b^*\,|1_A\rangle\,|1_B\rangle]|M_A \otimes I_B|[a\,|0_A\rangle\,|0_B\rangle + b(|1_A\rangle\,|1_B\rangle)]$$

$$= |a|^2\,\langle 0_A|M_A|0_A\rangle\,\langle 0_B|I_B|0_B\rangle + |b|^2\,\langle 1_A|M_A|1_A\rangle\,\langle 1_B|I_B|1_B\rangle$$

$$(\langle 0_B|I_B|0_B\rangle + |b|^2 = 1)$$

$$(\langle 1_B|I_B|1_B\rangle = 1)$$

$$= |a|^2\,\langle 0_A|M_A|0_A\rangle + \langle 1_A|M_A|1_A\rangle$$

$$(Since, a^*b\,\langle 0_A|\,M_A\,|1_A\rangle\,\langle 0_B|\,I_B\,|1_B\rangle\, and\ other\ terms = 0)$$

So this is sum of $|a|^2$ times the diagonal matrix element of M_A and $|b|^2$ times the diagonal matrix element of M_A in state 1.
Composite System: $\langle\psi_{AB}|M_A \otimes I_B|\psi_{AB}\rangle = tr(\rho_A M_A)$
where, $\rho_A = |a|^2\,|0_A\rangle\,\langle 0_A| + |b|^2\,|1_A\rangle\,\langle 1_A|$ is the density operator (density matrix) for the subsystem A.

5.3 Mixed State

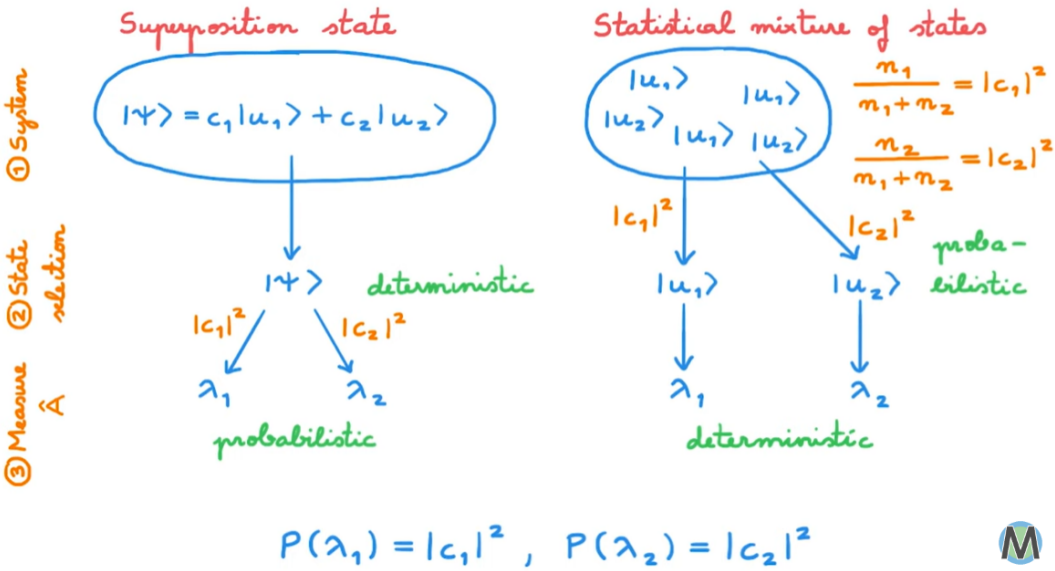
If we know everything there is to know about our quantum system, then the probability is a purely quantum effect that arises when we perform measurements. When this happens we say that we have a pure state.
eg: $|\psi\rangle = c_1\,|u_1\rangle + c_2\,|u_2\rangle$. So, a pure state is a perfectly known state.

If instead we only have partial information about our system, then probability enters in two places

- quantum mchanical probability associated with measurements

- the lack of knowledge about our system means that we must also use probability to describe which state we are actually dealing with in the first place, this has nothing to do with quantum mechanics and is something that also happens in the classical world

Mixed state in quantum mechanics is a state for which we only have partial information. For eg: the state could be $|\psi_1\rangle$ with probability p_1, $|\psi_2\rangle$ with probability p_2, ..., and so. Each of these states $|\psi_1\rangle$, $|\psi_2\rangle$,...., and so on are pure states so each has the quantum mechanical probability associated with it, when we perform a measurement, however we don't know in which of this pure state our system is, this is where the p probabilities come in they tell us how likely each pure state is, the p probabilities are independent of quantum mechanics they simply encode our lack of knowledge about the state of the system.



Two types of probability for mixed states

- Coherent superposition of states (quantum mechanical)
- A mixture of states with pre-determined (classical) probabilities

5.4 Density Matrix

5.4.1 Density operator for pure states

To introduce the density operator, let's considre a basis u of our state space and choose it to be orthonormal. We can write the general state vector $|\psi\rangle$ and expand it in the u basis as $|\psi\rangle = \sum_i c_i |u_i\rangle$ (A pure state may be expressed uniquely as a linear combination of basis states in its Hilbert space)
density operator $\rho = |\psi\rangle\langle\psi|$ (projection onto the state $|\psi\rangle$)
Density operator in u basis $\rho_{ij} = \langle u_i|\rho|u_j\rangle = \langle u_i|\psi\rangle\langle\psi|u_j\rangle = c_i c_j^*$
Notice that the state $|\psi\rangle$ and the density operator ρ contain the same info.

$$\hat{\rho} = |\psi\rangle\langle\psi| \qquad \rho_{ij} = c_i c_j^*$$

① Hermitian $\hat{e}^+ = \hat{e}$

$$\hat{e}^+ = (|\psi\rangle\langle\psi|)^+ = |\psi\rangle\langle\psi| = \hat{e}$$

② Idempotent $\hat{e}^2 = \hat{e}$

$$\hat{e}^2 = |\psi\rangle\underbrace{\langle\psi|\psi\rangle}_{=1}\langle\psi| = |\psi\rangle\langle\psi| = \hat{e}$$ pure state

③ No global phase ambiguity

$$|\psi\rangle \Longrightarrow \hat{e} = |\psi\rangle\langle\psi|$$

$$|\psi'\rangle = e^{i\theta}|\psi\rangle \Longrightarrow \hat{e}' = |\psi'\rangle\langle\psi'| = e^{i\theta}|\psi\rangle\langle\psi|e^{-i\theta} = |\psi\rangle\langle\psi| \Big\} \hat{e}' = \hat{e}$$

The diagonal elements of the diagonal matrix gives the Born probability of getting a state. The diagonal elements being $|c_i|^2$ are essentially non-negative quantity, so therefore the density operator is a positive operator, its eigenvalues are non-negative, thus, $\sum_i \lambda_i = 1$. The off-diagonal elements are the interference terms and depend on the phases between the states i and j.

Normalization

- State vector: $\langle\psi|\psi\rangle = 1$

$$\langle\psi|\psi\rangle = \left(\sum_i c_i^* \langle u_i|\right)\left(\sum_j c_j |u_j\rangle\right) = \sum_{i,j} c_i^* c_j \overbrace{\langle u_i|u_j\rangle}^{\delta_{ij}} = \sum_i |c_i|^2$$

- Density operator

$$\rho_{ii} = |c_i|^2$$

$$1 = \sum_i |c_i|^2 = \sum_i \rho_{ii} = \mathrm{Tr}(\hat{\rho})$$

$$1 = \langle\psi|\psi\rangle \iff 1 = \mathrm{Tr}(\hat{\rho})$$

Expectation values

$$\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle = \langle\psi|\mathbb{1}\hat{A}\mathbb{1}|\psi\rangle = \langle\psi|\left(\sum_i |u_i\rangle\langle u_i|\right)\hat{A}\left(\sum_j |u_j\rangle\langle u_j|\right)|\psi\rangle$$

$$= \sum_{i,j} \underbrace{\langle\psi|u_i\rangle}_{c_i^*} \underbrace{\langle u_i|\hat{A}|u_j\rangle}_{A_{ij}} \underbrace{\langle u_j|\psi\rangle}_{c_j} = \sum_{i,j} c_i^* c_j A_{ij}$$

$$\langle\hat{A}\rangle = \sum_{i,j} \langle\psi|u_i\rangle\langle u_i|\hat{A}|u_j\rangle\langle u_j|\psi\rangle = \sum_{i,j} \overbrace{\langle u_j|\psi\rangle\langle\psi|}^{\hat{\rho}} u_i\rangle\langle u_i|\hat{A}|u_j\rangle$$

$$= \sum_j \sum_i \langle u_j|\hat{\rho}|u_i\rangle\langle u_i|\hat{A}|u_j\rangle = \sum_j \langle u_j|\hat{\rho}\underbrace{\left(\sum_i |u_i\rangle\langle u_i|\right)}_{\mathbb{1}}\hat{A}|u_j\rangle$$

$$= \sum_j \underbrace{\langle u_j|\hat{\rho}\hat{A}|u_j\rangle}_{(\hat{\rho}\hat{A})_{jj}} = \mathrm{Tr}(\hat{\rho}\hat{A})$$

$$\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle = \mathrm{Tr}(\hat{\rho}\hat{A})$$

Time evolution

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

$$\frac{d}{dt}\hat{\rho}(t) = \frac{d}{dt}|\psi(t)\rangle\langle\psi(t)| = \overbrace{\left(\frac{d}{dt}|\psi(t)\rangle\right)}^{(*)}\langle\psi(t)| + |\psi(t)\rangle\overbrace{\left(\frac{d}{dt}\langle\psi(t)|\right)}^{(*)}$$

$$\frac{d}{dt}|\psi(t)\rangle = \frac{1}{i\hbar}\hat{H}(t)|\psi(t)\rangle \longleftrightarrow \frac{d}{dt}\langle\psi(t)| = -\frac{1}{i\hbar}\langle\psi(t)|\hat{H}(t)$$

$$\frac{d}{dt}\hat{\rho}(t) = \frac{1}{i\hbar}\hat{H}(t)\overbrace{|\psi(t)\rangle\langle\psi(t)|}^{\hat{\rho}(t)} - \frac{1}{i\hbar}\overbrace{|\psi(t)\rangle\langle\psi(t)|}^{\hat{\rho}(t)}\hat{H}(t)$$

$$= \frac{1}{i\hbar}\left(\hat{H}(t)\hat{\rho}(t) - \hat{\rho}(t)\hat{H}(t)\right) = \frac{1}{i\hbar}[\hat{H}(t), \hat{\rho}(t)]$$

$$\frac{d}{dt}|\psi(t)\rangle = \frac{1}{i\hbar}\hat{H}(t)|\psi(t)\rangle \Longleftrightarrow \frac{d}{dt}\hat{\rho}(t) = \frac{1}{i\hbar}[\hat{H}(t), \hat{\rho}(t)]$$

$$\longrightarrow \langle\hat{A}\rangle = \mathrm{Tr}(\hat{\rho}\hat{A})$$

Probability of measurement outcome

$$\hat{A}|a_m^i\rangle = a_m|a_m^i\rangle \qquad i = 1, \cdots, g_m$$

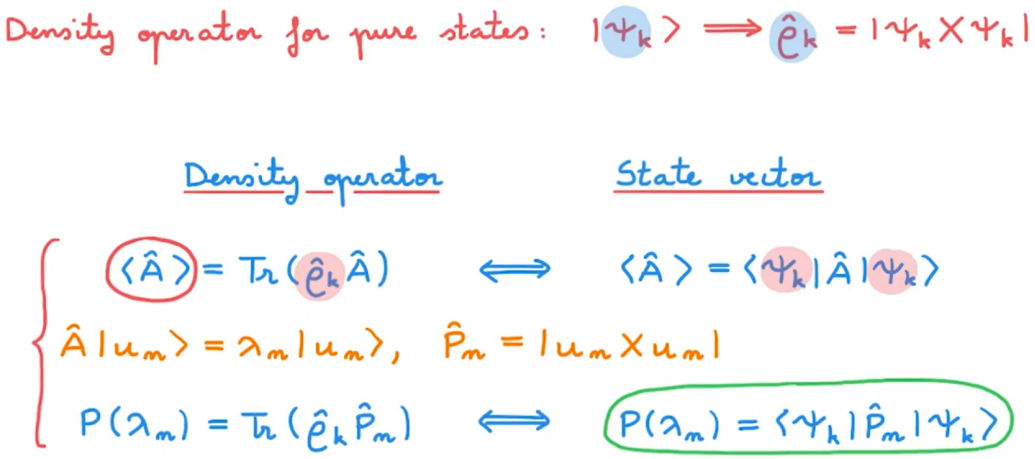
$$P(a_m) = \langle\psi|\hat{P}_m|\psi\rangle \qquad \hat{P}_m = \sum_{i=1}^{g_m}|a_m^i\rangle\langle a_m^i|$$

$$P(a_m) = \mathrm{Tr}(\hat{\rho}\hat{P}_m)$$

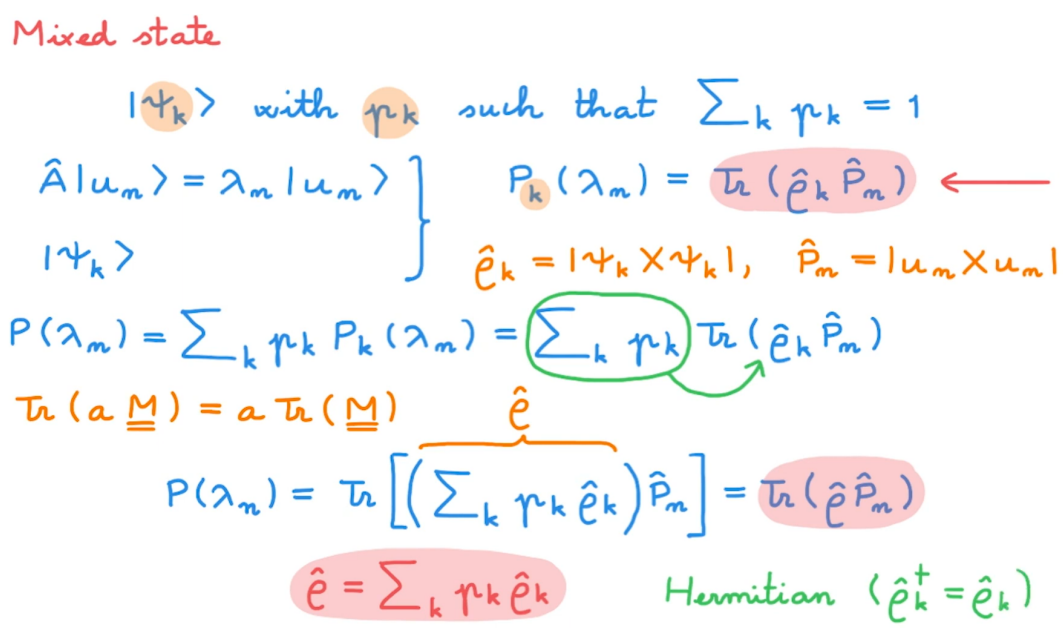
Overall, we have seen that we can do quantum mechanics for pure states in two equivalent ways, i.e, in terms of statevectors and in terms of the density operator. The density operator provides an equivalent language to quantum mechanics to that of the statevector. So, what is the point of introducing an alterbnative way of doing quantum mechanics.

5.4.2 Density Matrix for Mixed states

A mixed state does not have a statevector but is described only by a density matrix. It turns out that for mixed states it is more convenient to do quantum mchanics using the density operator is because in all the formulas the density opeartor appeared in a linear manner where the statevector can appear in a quadratic manner.



For example, the expression for the expectation value has the density operator only once but but the corresponding expression has a statevector twice.



The above expression says that the density operator for a mixed states simply the average of the density opeartors of each individual pure state making up the statistical mixture.

$$\hat{\rho} = \sum_k p_k \hat{\rho}_k$$

Normalization

- Pure state: $\mathrm{Tr}(\hat{\rho}_k) = 1$

$$\mathrm{Tr}(\hat{\rho}) = \mathrm{Tr}\left(\sum_k p_k \hat{\rho}_k\right) = \sum_k p_k \underbrace{\mathrm{Tr}(\hat{\rho}_k)}_{=1}$$

$$= \sum_k p_k = 1$$

Expectation values

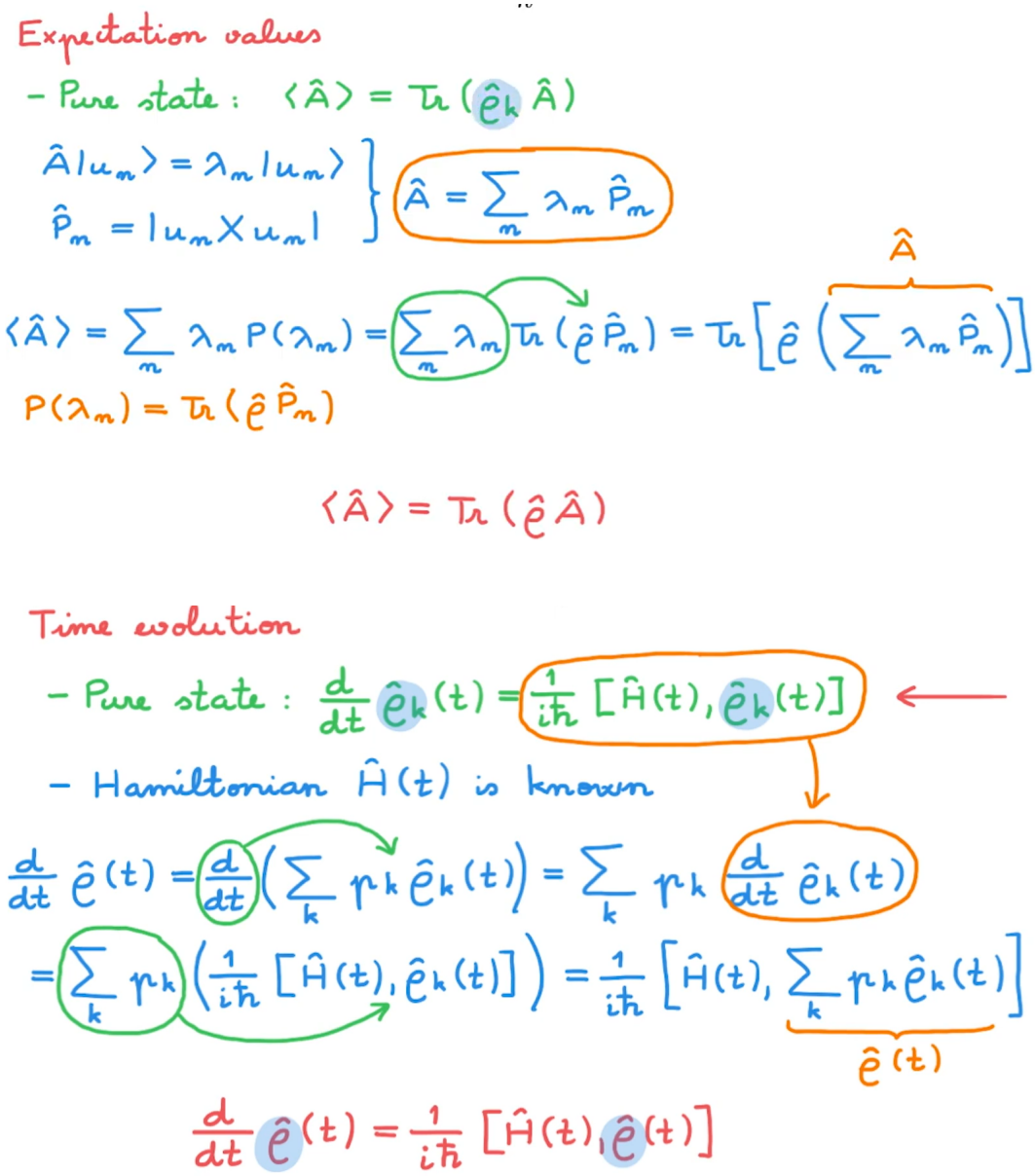
- Pure state: $\langle \hat{A} \rangle = \mathrm{Tr}(\hat{\rho}_k \hat{A})$

$$\left.\begin{aligned} \hat{A}|u_m\rangle &= \lambda_m |u_m\rangle \\ \hat{P}_m &= |u_m\rangle\langle u_m| \end{aligned}\right\} \quad \hat{A} = \sum_m \lambda_m \hat{P}_m$$

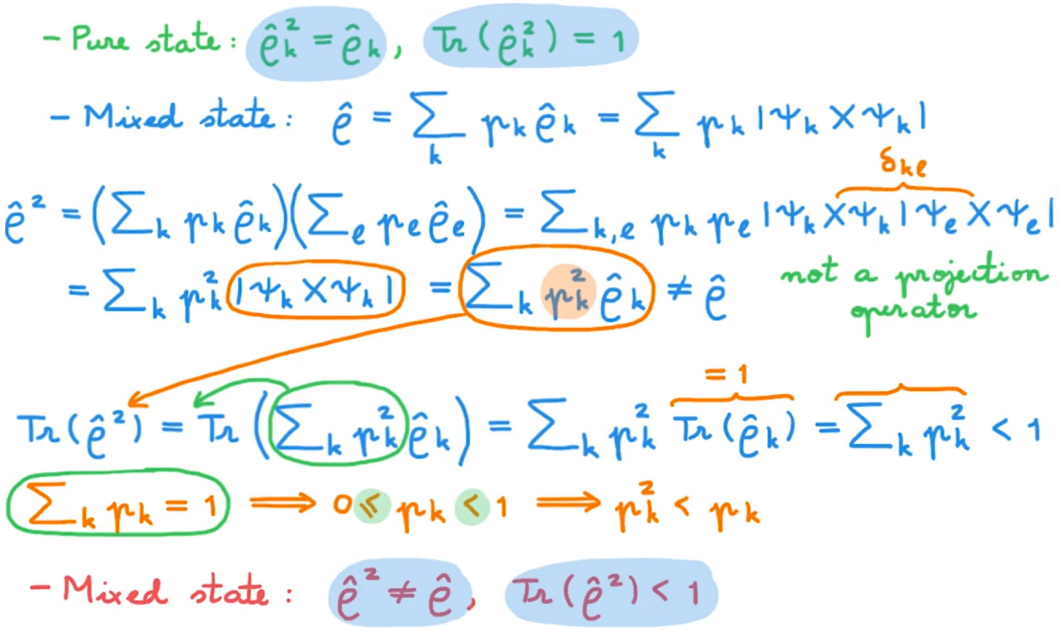
$$\hat{A}|\psi\rangle = \hat{A}\sum_m c_m |u_m\rangle = \sum_m c_m \underbrace{\hat{A}|u_m\rangle}_{\lambda_m |u_m\rangle} = \sum_m c_m \lambda_m |u_m\rangle$$

$$\sum_n \lambda_n \hat{P}_n |\psi\rangle = \sum_n \lambda_n |u_n\rangle\langle u_n| \sum_m c_m |u_m\rangle$$

$$= \sum_{n,m} c_m \lambda_n |u_n\rangle \underbrace{\langle u_n | u_m\rangle}_{\delta_{nm}} = \sum_m c_m \lambda_m |u_m\rangle$$



Density operator satisfies *Liouville Equation*. It does not satisfy Heisenberg equation of motion because, though it has mathematical structure of an operator, it does not represent a physical observable. The case with no off-diagonal elements and equal diagonal elements in density matrix is a maximally mixed state, otherwise it is a partially mixed state.



$\hat{\rho}_k = \lvert\psi_k\rangle\langle\psi_k\rvert$	$\hat{\rho} = \sum_k p_k \hat{\rho}_k$
$\mathrm{Tr}(\hat{\rho}_k) = 1$	$\mathrm{Tr}(\hat{\rho}) = 1$
$\langle\hat{A}\rangle = \mathrm{Tr}(\hat{\rho}_k \hat{A})$	$\langle\hat{A}\rangle = \mathrm{Tr}(\hat{\rho}\hat{A})$
$\frac{d}{dt}\hat{\rho}_k(t) = \frac{1}{i\hbar}[\hat{H}(t), \hat{\rho}_k(t)]$	$\frac{d}{dt}\hat{\rho}(t) = \frac{1}{i\hbar}[\hat{H}(t), \hat{\rho}(t)]$
$\hat{\rho}_k^2 = \hat{\rho}_k$	$\hat{\rho}^2 \neq \hat{\rho}$
$\mathrm{Tr}(\hat{\rho}_k^2) = 1$	$\mathrm{Tr}(\hat{\rho}^2) < 1$

Fig: The left column shows expressions for pure state while right shows for the mixed state

5.4.3 Bloch Sphere and Density Matrix

The state $|\psi\rangle$ corresponding to the eigenvalue $+1$ of σ_n

$$|\psi\rangle = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ e^{\iota\phi}\sin\left(\frac{\theta}{2}\right) \end{pmatrix}$$

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix} 1+\cos\theta & e^{-\iota\phi}\sin\theta \\ e^{\iota\phi}\sin\theta & 1-\cos\theta \end{pmatrix} = \frac{I}{2} + \frac{1}{2}\sin\theta\cos\phi\sigma_x + \frac{1}{2}\sin\theta\sin\phi\sigma_y + \frac{\cos\theta}{2}\sigma_z$$

$$= \frac{1}{2}[I + \hat{n}\cdot\vec{\sigma}] = \frac{1}{2}[I + \vec{a}\cdot\vec{\sigma}], \qquad |a| < 1$$

$$let\ a_z = \frac{1}{3}, \quad \rho = \frac{1}{2}[I + \frac{1}{3}\sigma_z] = \frac{2}{3}|0\rangle\langle 0| + \frac{1}{3}|1\rangle\langle 1|$$

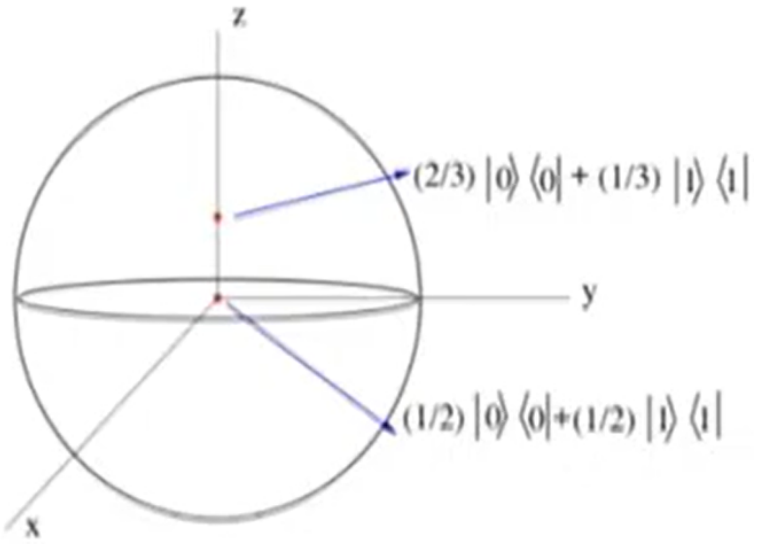


Fig: Bloch Ball

5.4.4 Reduced Density Matrix

- If the system A of interest is a part of a bigger system (A+B, B=environment), properties of A may be obtained by Taking partial trace over the environment B.
- $\rho_A = tr_B(\rho_{AB})$
- $tr_B[|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|]$

$$= \langle a_1|a_2\rangle \ \ tr(|b_1\rangle\langle b_2|) = \langle b_2|b_1\rangle\,|a_1\rangle\,|a_2\rangle$$

- Partial trace essentially averages out the effect of environment and extracts the properties of the system of interest.
- An entangled pure state, on being traced over one of the components may give a mixed state.

Chapter 6

Measurement Postulates

6.1 Projective Measurement

- If in addition to completeness, we require the measurement operators to be orthogonal projectors P_m, the measurement is called projective or Von Neumann measurement.
- $P_m^\dagger = P_m$ and $P_m P_m' = P_m \delta_{m,m'}$
- Since $P_m^2 = P_m$, $p(m) = \langle\psi|P_m|\psi\rangle$
- Expectation value of an operator has a spectral decomposition $M = \sum_m m P_m$ so that its expectation value is

$$E(M) = \sum_m m P_m = \sum_m m \langle\psi|P_m|\psi\rangle$$
$$= \langle\psi|\sum_m m P_m|\psi\rangle = \langle\psi|M|\psi\rangle$$

- Example: for one-qubit state, let measurement operators be $M_0 = |0\rangle\langle 0|$ and $M_1 = |1\rangle\langle 1|$. If a state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is measured in this basis, $p(0) = |\alpha|^2$ and the post measurement state is $\frac{\alpha|0\rangle}{|\alpha|} \equiv e^{\iota\phi}|0\rangle$, which is identical to state $|0\rangle$ as the global phase is immaterial.
- If the measurement operators are $\{|m\rangle\langle m|\}$ where $\{|m\rangle\}$ are the basis states, the measurement is in a basis. (All measurements don't have to be done in a basis, they could be general in terms of any orthogonal vectors in the Hilbert space)

6.1.1 Projective Measurement-Multiple Qubits

- Let $|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$
- Suppose we wish to only measure the first qubit, irrespective of what the state of the second qubit is. The corresponding measurement operator is $M_0 = M_{00} + M_{01}$ where $M_{00} = |00\rangle\langle 00|$ and $M_{01} = |00\rangle\langle 00|$

  $$M_0 = M_{00} + M_{01} = |00\rangle\langle 00| + |00\rangle\langle 00| = |00\rangle \otimes (\langle 00| + |01\rangle) = |00\rangle \otimes I_2$$

- Probability of measuring 1^{st} qubit as 0

  $$\langle\psi|M_0|\psi\rangle = (\alpha^*\langle 00| + \beta^*\langle 01|)|00\rangle \otimes I_2(\alpha\langle 00| + \beta|01\rangle) = |\alpha|^2 + |\beta|^2$$

6.2 Repeating Measurement

- Repetition of a measurement will give the state in which it collapsed in the first measurement
- Changimg a basis after a measurement:
  1. First measure in $\{0, 1\}$ basis then in $\{+, -\}$ basis
  2. First measure in $\{+, -\}$, then in $\{0, 1\}$ basis.
  The results are not same.
- Let $|\psi\rangle = \alpha|0 + \beta|1\rangle\rangle$. If we measure it in $\{0, 1\}$ basis, it gives 0 with probability $|\alpha|^2$. If we now measure it in $\{+, -\}$ basis, it gives + with $\frac{1}{2}$ probability. $P(0, +) = \frac{|\alpha|^2}{2}$
- If we first measure in $\{+, -\}$ and then in $\{0, 1\}$ it gives $\{+, 0\}$ with probability $\frac{|\alpha+\beta|^2}{2}$

6.3 POVM

- Positive operator-valued measurement
- It is non-projective. Projective measurements commute, POVM's need not
- Projective measurements are orthogonal $P_m P'_m = P_m \delta_{m,m'}$. POVM's are not necessarily so.

- In d dimensional space there are d number of projectors. In POVM these could be more than d.
- A collection of positive operators $\{E_i\}$ such that $\sum_i E_i = I$
- Since the operators are positive, $E_i = M_i^\dagger M_i$
- Probability of an outcome i is

$$p(i) = \langle\psi|E_i|\psi\rangle = Tr(\rho E_i)$$

- Post measurement state $\rho \longrightarrow \frac{M_i \rho M_i^\dagger}{Tr(\rho E_i)}$. If the state is not read $\rho \longrightarrow \sum_i M_i \rho M_i^\dagger$
- Example: $E_1 = \frac{1}{2}|0\rangle\langle 0|$, $E_2 = \frac{1}{2}|1\rangle\langle 1|$ and $E_3 = I - E_1 - E_2$

Chapter 7

Quantum Algorithms

Like any computer, quantum computer model has quantum registers corresponding to input and those corresponding to output. The difference is that while the classical register at a given time can only contain a particular state, the quantum register can contain a linear combination of states.
"State of the quantum register: linear combination of states"
Quantum Parallelism: Computation of a function for each of the states in the input register
and therein lies the great advantage of **Quantum Parallelism** because corresponding to the linear inputs combination of inputs, computer can compute functions corresponding to each one of these linear combination components and the output will be generated by what we call as a quantum oracle (**Oracle**: A black box computation, analogous to a classical subroutine).

So therefore, a typical model of quantum computation consists of input registers, a target register (where the output will ultimately be stored), a black box or a quantum oracle (which will compute the type of functions that we want), this is basically the constitution of the essential computation. But after that comes the most important point of the computation, namely in any computing we need a read out or a print, so we ahve to extract the required information from the output. Now that we have pointed is a very tall order because even though the output is a linear combination of the results corresponding to the linear combination of the inputs, if we want to measure the output there will not be an information about every output corresponding to every input, but one of the random results will come up and the question always is that how do you use these random results to your advantage.

7.1 Deutsch Algorithm

Input: one qubit state which is either $|0\rangle$ or $|1\rangle$ or a linear combination
Output: $f(0)$ or $f(1)$; one qubit state, either $|0\rangle$ or $|1\rangle$
Type of functions:
• constant function: function whose value is independent of the input, $f(0) = f(1)$; $f(0) = f(1) = 0$ or $f(0) = f(1) = 1$
• Balanced function: $f(0) \neq f(1)$; either $f(0) = 0, f(1) = 1$ or $f(0) = 1, f(1) = 0$
Now supposing there is a very trivial programming task and were to do by a classical algorithm. How would we determine the function that we are calculating is a constant function or a balanced function, which is in a black box. If you are to do it classically, the only way is to first find out what is f(0) which is either 0 or 1, then find out what is f(1) and check whether the f(1) value is the same as the value of f(0). So, in other words, a classical computation to determine whether the function is a constant function or a balanced function requires two queries.
While in quantum computing, we can determine whether a function is constant or balanced by a single query.

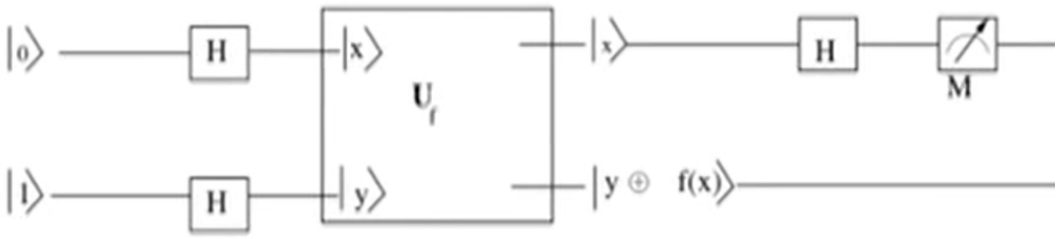


Fig: Deutsch Algorithm

• Oracle U_f to compute f(x)
• x is input register and y the target register
• $U_f : |x\rangle \otimes |y\rangle = |x\rangle \otimes |y \oplus f(x)\rangle$
• Input register $|x\rangle = |0\rangle \xrightarrow{H} \frac{|0\rangle+|1\rangle}{\sqrt{2}}$
• Target register $|y\rangle = |1\rangle \xrightarrow{H} \frac{|0\rangle-|1\rangle}{\sqrt{2}}$
• Input: $\frac{1}{2}(|00\rangle + |10\rangle - |01\rangle - |11\rangle)$
• Applications of oracle gives
Oracle: $\frac{1}{2}[|0, 0 \oplus f(0)\rangle + |1, 0 \oplus f(1)\rangle - |0, 1 \oplus f(0)\rangle - |1, 1 \oplus f(1)\rangle]$

$$= \frac{1}{2}[|0, f(0)\rangle + |1, f(1)\rangle - \left|0, \overline{f(0)}\right\rangle - \left|1, \overline{f(1)}\right\rangle] \quad (general)$$

(1) **Deutsch Problem-constant case**

- Output: $\frac{1}{2}(|0\rangle + |1\rangle) \otimes (f(0) - \overline{f(0)})$
- Except for a global $\pm$ factor, the output is

$$\frac{(|0\rangle + |1\rangle)}{\sqrt{2}} \otimes \frac{(|0\rangle - |1\rangle)}{\sqrt{2}}$$

- Passing 1^{st} register through Hadamard gate gives $|0\rangle$ in the 1^{st} register
- Finally $|0\rangle \otimes \frac{(|0\rangle - |1\rangle)}{\sqrt{2}}$

(2) **Deutsch Problem- Balanced Function**

- Balanced function: $\overline{f(0)} = f(1)$; $\overline{f(1)} = f(0)$
- Output: $\frac{(|0\rangle + |1\rangle)}{\sqrt{2}} \otimes (\pm)\frac{(|0\rangle - |1\rangle)}{\sqrt{2}}$

- First register, when passed through Hadamard gate gives $|1\rangle$. So, what it tells is that if we now make a measurement of the first register and the function is balanced, we get the result to be equal to $|1\rangle$. However, no information about the value of the function is obtained.

Summarizing the result, function constant's first register gives $|0\rangle$, function balanced's first register gives $|1\rangle$. Therefore, all that we need to do is to measure the first register, if we get a $|0\rangle$ it is a constant function and if we get $|1\rangle$ it is a balanced function.

7.2 Deutsch-Jozsa Algorithm

- An extension of Deutsch algorithm to the case of an n qubit input, $f : \{0,1\}^{\otimes n} \longrightarrow \{0,1\}$
- The function is either a constant or balanced, i.e, exactly half of them gives 0 while the other half of them evaluates to 1.
- Input is a uniform linear combination of the n qubit computational basis states $\{x_{n-1}, x_{n-2}, ..., x_1, x_0\}$ with $x_i = 0, 1$

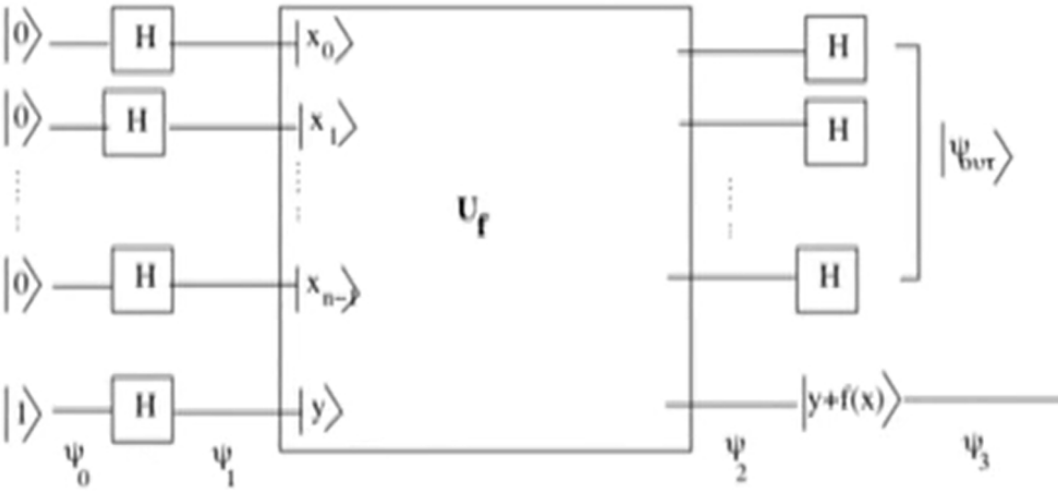


Fig: Deutsch-Jozsa Algorithm

- Input: $|0\rangle^{\otimes n} \xrightarrow{H^{\otimes n}} \frac{1}{\sqrt{2}} \sum_{x=0}^{2^n-1} |x\rangle$
- Target: $|1\rangle \xrightarrow{H} \frac{|0\rangle-|1\rangle}{\sqrt{2}}$
- Output(Entangled): $\frac{1}{\sqrt{2^{n+1}}} \sum_{x=0}^{2^n-1} |x\rangle \otimes \left|f(x) - \overline{f(x)}\right\rangle$
- $\frac{1}{2^n} \sum_{x=0}^{2^{n-1}} \sum_{k=0}^{2^{n-1}} (-1)^{f(x)+k\cdot x} |k\rangle \frac{|0\rangle-|1\rangle}{\sqrt{2}}$
- For balanced function, the coefficient of k=0 is $\sum_{x=0}^{2^{n-1}} (-1)^{f(x)} = 0$, because there are as many f(x)=1 as there are f(x)=0.
- First register=0 : Constant function
- First register anything else : balanced function

7.3 Bernstein-Vazirani Algorithms

- Given an n qubit input string

$$|x\rangle = |x_{n-1}x_{n-2}...x_1x_0\rangle$$

Find an unknown string

$$|a\rangle = |a_{n-1}a_{n-2}...a_1a_0\rangle$$

such that $f(x) = a \cdot x$, the bitwise product $a \cdot x = a_{n-1}x_{n-1} + \cdots + a_1x_1 + a_0x_0 \pmod 2$ is a given single qubit output. This is a special case of Deutsch-Jozsa with $f(x) = a \cdot x$

- Classical search requires n queries

$$f(10000....0) = x_{n-1}$$

$$f(01000....0) = x_{n-2}$$

$$...$$

$$f(00000....1) = x_0$$

A single query is enough in quantum computer

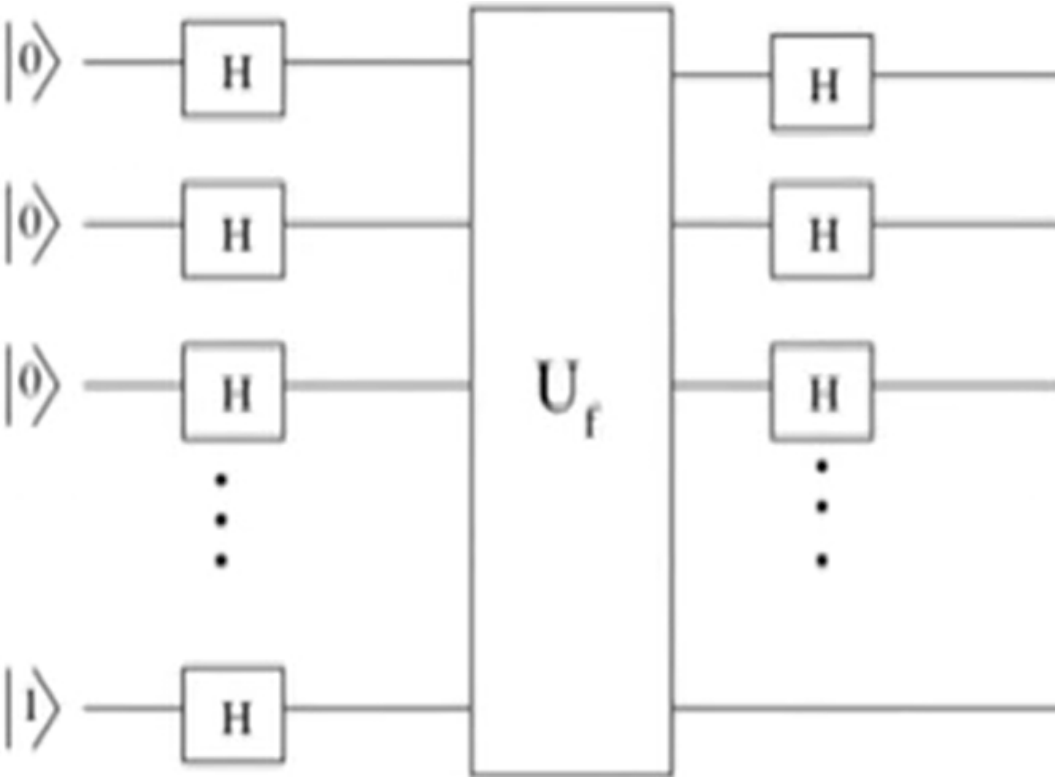


Fig: Bernstein-Vazirani Algorithm

- Using Deutsch-Jozsa (replacing $f(x) = a \cdot x$)
- Output: $\frac{1}{2^n}\sum_{x=0}^{2^{n-1}}\sum_{k=0}^{2^{n-1}}(-1)^{a\cdot x+k\cdot x}\,|k\rangle\,\frac{|0\rangle-|1\rangle}{\sqrt{2}}$

- Perform sum over x gives

$$\sum_{x=0}^{2^{n-1}}(-1)^{a\cdot x+k\cdot x} =$$

$$\prod_{i=1}^{n}\sum_{x_i=0}^{1}(-1)^{x_i(a_i+k_i)} = \prod_{i=1}^{n}[1+(-1)^{a_i+k_i}]$$

the other things remains the same.

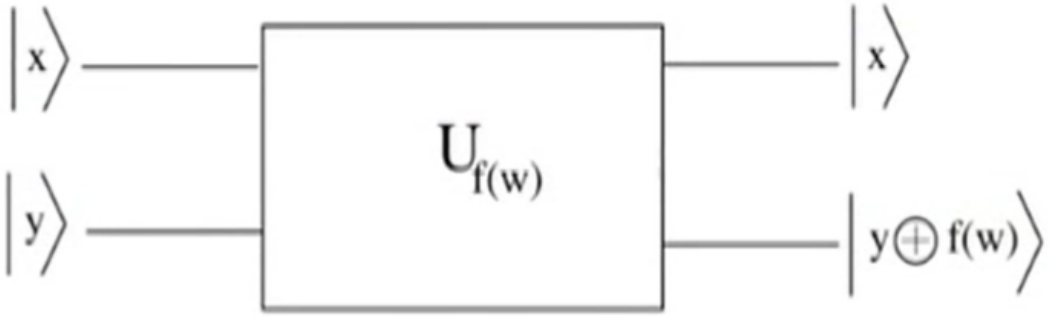
7.4 Grover's Search Algorithm

Finding a particular element in an unstructured database of N elements requires $O(N)$ searches. Grover's Algorithm speeds up the search quadratically,i.e. it requires $O(\sqrt{N})$ searches.

Let $N = 2^n$ so that our input is an n qubit data. An oracle defines a function defined over integers $x,\ 0 \leq k \leq 2^n - 1$ such that $f(x) = 0, \forall\ x \neq w$, for which $f(w) = 1$.

7.4.1 Grover Oracle

$$f_w(x) = \begin{cases} 0 & x \neq w \\ 1 & x = w \end{cases}$$



If $y = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$, the output can be written as $(-1)^{f_w(x)}\,|x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$. The second register is unaltered but the first register has a phase depending on $f_w(x)$. If $f_w(x) = 1$, the sign of the first register is flipped.

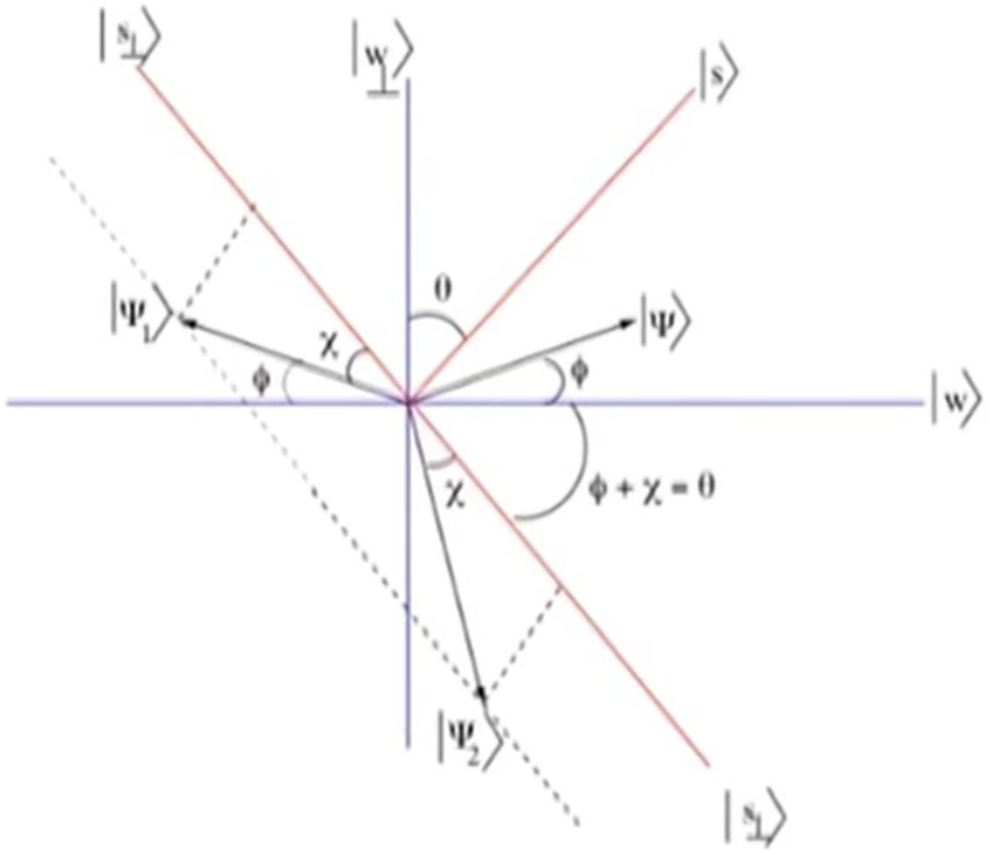
7.4.2 Grover Operators

$$U_w = I - 2\left|w\right\rangle\left\langle w\right|$$
$$U_s = 2\left|s\right\rangle\left\langle s\right| - I$$

where the "standard state" $|s\rangle$ is the uniform linear combination of the computational basis

$$|s\rangle = \frac{1}{\sqrt{N}}\sum_{0}^{N-1}|x\rangle \qquad (\langle w|s\rangle = \frac{1}{\sqrt{N}})$$

Grover Rotation Operator $R_G = U_s U_w$, Rotation has a simple geometrical interpretation with the angle in the rotation matrix being twice the angle of initial probability amplitude.



Arbitrary state resolved along w and perpendicular to w: $|\psi\rangle = |\psi_\perp\rangle + |\psi_\parallel\rangle$. Application of U_w flips the sign of parallel component taking it to $|\psi_1\rangle$ while the perpendicular component is unchanged. Resolve $|\psi_1\rangle$ along

s and perpendicular to it. Application of R_s flips the sign of perpendicular component taking it to $|\psi_2\rangle$.

Angle between $|w\rangle$ and $|s_\perp\rangle$ = angle between $|s\rangle$ and $|w_\perp\rangle = \theta = \phi + \chi$. Thus, angle of rotation = angle between $|\psi\rangle$ and $|\psi_2\rangle$ is 2θ.

Arbitrary state in computational basis $|\psi\rangle = \sum_x a_x |x\rangle$ and standard state $|s\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle$, where a_x is the amplitude of $|x\rangle$ in the state $|\psi\rangle$

$$\langle s|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{xx'} a_x \langle x|x'\rangle = \frac{1}{\sqrt{N}} \sum_x a_x \quad (\langle x|x'\rangle = \delta_{xx'})$$

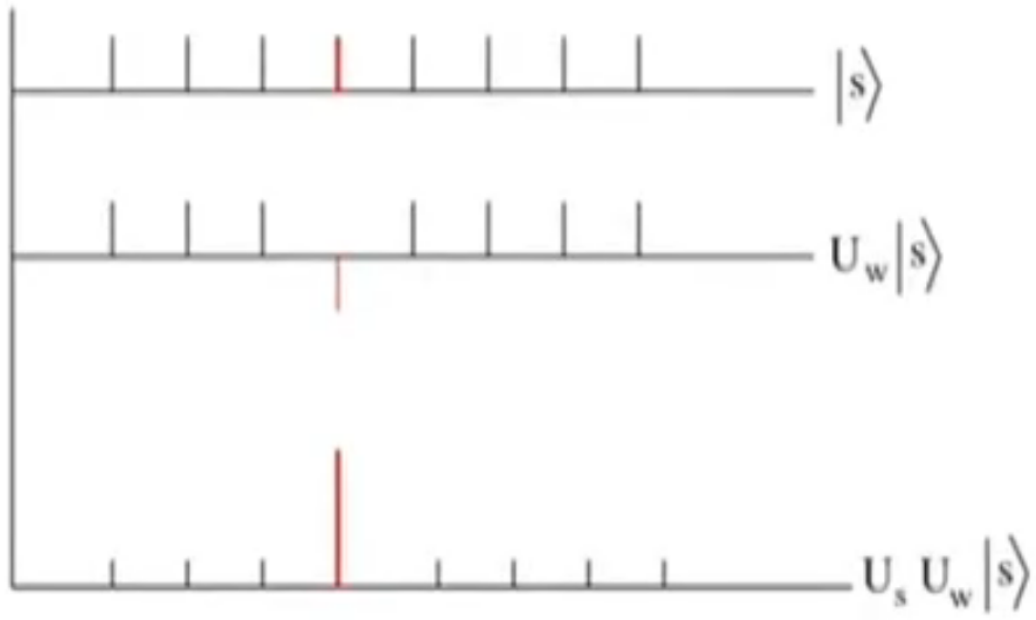
$\bar{a} = \frac{1}{N} \sum_x a_x$ Mean amplitude of $|\psi\rangle$ in a computational basis
$\langle s|\psi\rangle = \sqrt{N}\bar{a}$; U_s acting on state $|\psi\rangle$

$$U_s |\psi\rangle = (2 |s\rangle \langle s| - 1) \sum_x a_x |x\rangle = 2 |s\rangle \langle s|\psi\rangle - |\psi\rangle$$

$$= \sqrt{N}\bar{a} |s\rangle - |\psi\rangle = \sum_x (2\bar{a} - a_x) |x\rangle$$

Before application of U_s, the amplitude of $|x\rangle$ in state $|\psi\rangle = a_x$; Initial amplitude w.r.t mean = $a_x - \bar{a}$ and the final amplitude = $2\bar{a} - a_x$; final amplitude w.r.t mean = $\bar{a} - a_x$
Application of U_s on an arbitrary state $|\psi\rangle$: Amplitude w.r.t the mean of an arbitrary state on application of U_s gets inverted.
(Note: U_w inverts the amplitude of w(marked state) alone)

Selective Amplification of Amplitude
Amplitude of unmarked state after application of $U_s U_w$ is $2\bar{a} - a_x$
while that of marked state is $2\bar{a} - (-a_x) = 2\bar{a} + a_x$;
which means the corresponding probability density of marked state has amplified, called as selective amplification.



Maximum Number of iterations
angle between $|s\rangle$ and $|w\rangle$ is $\frac{\pi}{2} - \theta$. Since, each Grover rotation is by 2θ, after m iterations: $m \times 2\theta = \frac{\pi}{2} - \theta \implies m = \frac{\pi}{4\theta} - 2$
As $\theta \approx \sin\theta = \frac{1}{\sqrt{N}}$, $m \approx \pi\sqrt{N}/4$

Amplitude of $|w\rangle$ after m iterations
After m iterations, angle between $|s\rangle$ and $|w\rangle$ becomes $\frac{\pi}{2} - (2m+1)\theta$. Thus the amplitude of $|w\rangle$ in $|s\rangle$ is

$$\sin(2m+1)\theta = \sin\left[\left(\frac{\pi\sqrt{N}}{2} + 1\right)\frac{1}{\sqrt{N}}\right]$$

$$= \sin\left(\frac{\pi}{2} + \frac{1}{\sqrt{N}}\right) = \cos\frac{1}{\sqrt{N}} \cong 1 - \frac{1}{2N}$$

7.4.3 Diffusion Operator

Diffusion (Operator) or matrix D is defined as

$$D_{ij} = \begin{cases} -1 + \frac{2}{N} & i = j \\ \frac{2}{N} & i \neq j \end{cases}$$

The matrix D has the representation $D = -I + \frac{2J}{N}$ where J is a $N \times N$ matrix with each element as unity and $\frac{J}{N}$ is a projection operator $\left(\frac{J}{N}\right)^2 = \frac{J}{N}$.

D acting on state $(a_1, a_2, ..., a_N)^T$, first is $-I$, we get $a_1, a_2, ...a_N$ with a negative sign and J is a matrix whose every element is 1, so this gives 2 times the average.

$$\left(-I + \frac{2J}{N}\right) \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_N \end{pmatrix} = \begin{pmatrix} -a_1 + 2\bar{a} \\ -a_2 + 2\bar{a} \\ . \\ . \\ . \\ -a_N + 2\bar{a} \end{pmatrix}$$

$$D\left|\psi\right\rangle = \sum_x (2\bar{a} - a_x)\left|x\right\rangle$$

Diffusion operator may be implemented by $D = WRW$, where W is Walsh Hadamard transform and R is selective phase rotation

$$H^{\otimes n}\left|x\right\rangle = \frac{1}{\sqrt{N}} \sum_x (-1)^{x \cdot y}\left|y\right\rangle$$

$$W_{ij} = \frac{1}{\sqrt{N}}(-1)^{i \cdot j}$$

$$R_{ij} = (2\delta_{i0} - 1)\delta_{ij}$$

7.4.4 The Algorithm

1. Generate the standard state $\left|s\right\rangle = \frac{1}{\sqrt{N}} \sum \left|x\right\rangle$
2. $(n+1)$th qubit initialized to $\frac{\left|0\right\rangle - \left|1\right\rangle}{\sqrt{2}}$
3. Loop m times:
   (a) T(inverts the sign of the marked state w): Apply the oracle(initial $a_x = \frac{1}{\sqrt{N}}$)
   $$\left|s\right\rangle\left|y\right\rangle \longrightarrow \sum_{x=0}^{N-1} a_x \left|x\right\rangle (-1)^{f(x)}\left|y\right\rangle$$
   (b) Apply diffusion operator $D = WRW$
   (c) Apply the steps (a) and (b), $O(\sqrt{N})$ number of times
4. Measure the first register. With a high degree of probability it will identity the marked state w.

5. The algorithm may fail with probability $O\left(\frac{1}{N}\right)$ in which case go back to step 1.

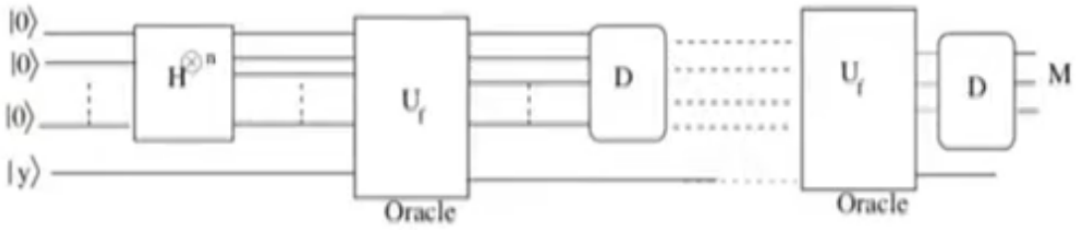


Fig: circuit for Grover's Algorithm

Algorithm Quality

Let $|u\rangle$ denote the linear combination of all states for which $f(x) = 0$, i.e.

$$|u\rangle = \frac{1}{\sqrt{N-1}} \sum_{x \neq w} |x\rangle$$

The standard state is then

$$|s\rangle = \frac{1}{\sqrt{N}} |w\rangle + \sqrt{\frac{N-1}{N}} |u\rangle$$

The marked state (w) is inverted by operation T. On application of D, the amplitude of the marked state increases. Suppose after the j-th iteration, the amplitude of each unmarked state(u, states other than state w) is u_j (recall: unmarked states are all at par, so whatever is the amplitude in the computational basis of one of the states, the same is true for all the unmarked states) while that of the marked state is w_j. Now apply D on this state to get the amplitudes of the (j+1)th iteration.

$$\begin{pmatrix} u_{j+1} \\ u_{j+2} \\ \dots \\ w_{j+1} \\ \dots \\ u_{j+1} \end{pmatrix} = - \begin{pmatrix} u_j \\ u_j \\ \dots \\ w_j \\ \dots \\ u_j \end{pmatrix} + \frac{2}{n} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} u_j \\ u_j \\ \dots \\ w_j \\ \dots \\ u_j \end{pmatrix}$$

$$w_{j+1} = \left(\frac{2}{N} - 1\right) w_j + \frac{2}{N}(N-1)u_j$$

$$u_{j+1} = \frac{2}{N} w_j + \frac{N-2}{N} u_j$$

Since before application of D, w was inverted by operation T, so therefore, in order to relate to the previous value of w_j, we need to invert the w_j in the above two equations

$$w_{j+1} = \left(1 - \frac{2}{N}\right) w_j + \frac{2}{N}(N-1)u_j$$

$$u_{j+1} = -\frac{2}{N} w_j + \frac{N-2}{N} u_j$$

let $c_j = \sqrt{N_1}\, u_j$

$$\begin{pmatrix} w_{j+1} \\ c_{j+1} \end{pmatrix} = \begin{pmatrix} 1 - \frac{2}{N} & \frac{2}{N}\sqrt{N-1} \\ -\frac{2}{N}\sqrt{N-1} & 1 - \frac{2}{N} \end{pmatrix} \begin{pmatrix} w_j \\ c_j \end{pmatrix}$$

$\frac{1}{\sqrt{N}} = \sin\theta$ and $\cos\theta = \sqrt{1 - \frac{1}{N}}$; $\sin 2\theta = \frac{2}{N}\sqrt{N-1}$ and $\cos 2\theta = 1 - \frac{2}{N}$

$$\begin{pmatrix} w_{j+1} \\ c_{j+1} \end{pmatrix} = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{pmatrix} \begin{pmatrix} w_j \\ c_j \end{pmatrix}$$

$$\begin{pmatrix} w_{j+1} \\ u_{j+1} \end{pmatrix} = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{pmatrix} \begin{pmatrix} w_j \\ u_j \end{pmatrix}$$

$$= \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{pmatrix}^j \begin{pmatrix} w_1 \\ u_1 \end{pmatrix}$$

(That's how j iterations may be done)

$$\begin{pmatrix} w_{j+1} \\ u_{j+1} \end{pmatrix} = \begin{pmatrix} \sin(2j+1)\theta \\ \cos(2j+1)\theta \end{pmatrix} \qquad \left(Since,\ \begin{pmatrix} w_1 \\ u_1 \end{pmatrix} = \begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix} \right)$$

(The diffusion and the operator T which inverts the initial state leads to selective amplification of a marked state and this allows to identify the marked state)

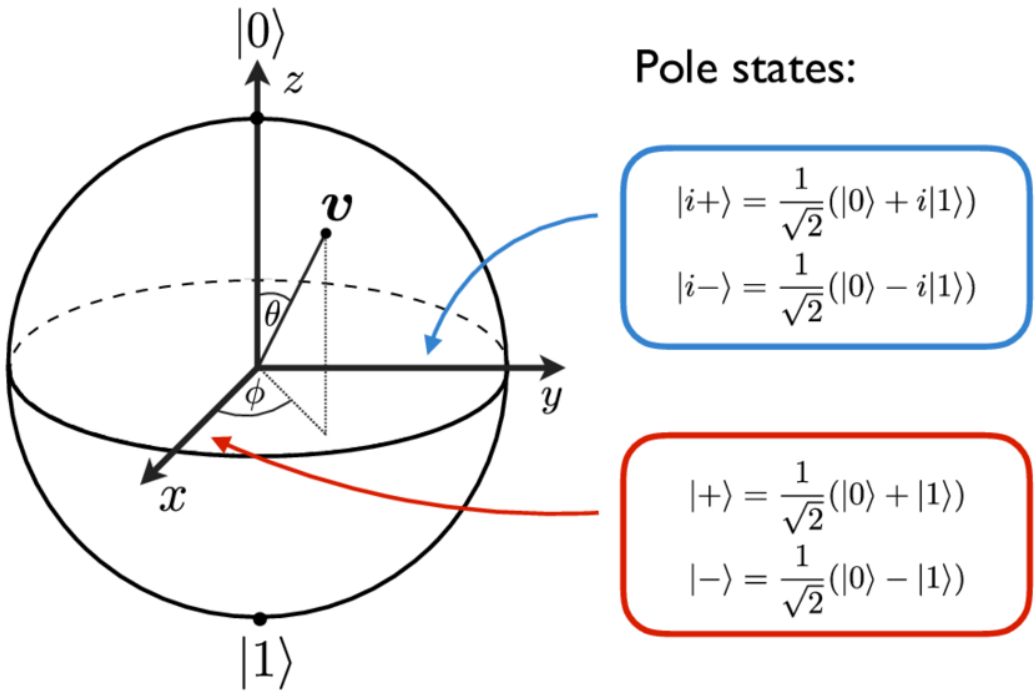
Amplitude of marked state

Thus measurement of first register will give marked state with a probability amplitude $\sin(2m+1)\theta$, which means the probability is given by $\sin^2(2m+1)\theta$. It is important to note that since we have an oscillating function, so if we start with the small value of θ and gradually increase the $\sin^2$ function till it has approached one or is just about to cross 1, after that increasing the number of iterations will not help, in fact, make it worse and that is why one must have a priori idea of number of iterations which is approximately $\frac{\pi}{4}\sqrt{N}$.

Chapter 8

Quantum Fourier Transform and its applications

QFT is effectively a change of basis from the computational basis to Fourier basis. Example: for one-qubit state, $|+\rangle$ is the fourier basis for the computational basis $|0\rangle$ and $|-\rangle$ is for $|1\rangle$; the one which does this is the Hadamard gate.



8.1 Discrete Fourier Transform

Let $n \in \mathbb{N}$, the set of natural numbers and S_n a set of $N = 2^n$ integers $0, 1, ...N-1$. Consider two discrete variables x and y such that $x, y \in S_n$. Discrete Integral transform $\tilde{f}$ of the variable is

$$\tilde{f}(y) = \sum_{x=0}^{N-1} K(y,x) f(x)$$

$K(x,y)$(kernel) is bivariate(in general, complex) function of the variables x and y. as x and y are discrete, this is a matrix equation with f and $\tilde{f}$ as N-component column vector and $K(y,x)$ as $N \times N$ matrix. If K is invertible(in particular, unitary) an inverse transformation also exists

$$f(x) = \sum_{y=0}^{N-1} K_\dagger(x,y) \tilde{f}(y)$$

Extension to Hilbert Space $(\mathbb{C}^2)^{\otimes n}$
Let $|x\rangle = |x_{n-1}, x_{n-2}, ...x_0\rangle$ be basis vector in n qubit space $(\mathbb{C}^2)^{\otimes n}$ with $x_i \in 0, 1$

$$|\tilde{x}\rangle = U|x\rangle = \sum_{y=0}^{N-1} |y\rangle \langle y| (U|x\rangle) = \sum_{y=0}^{N-1} U(y,x)|y\rangle$$

Compare with $\tilde{f}(y) = \sum_{x=0}^{N-1} K(y,x) f(x)$, it is seen that if U is a unitary matrix such that

$$U|x\rangle = \sum_{y=0}^{N-1} K(x,y)|y\rangle$$

U computes DIT. By quantum Parallelism, it computes DIT of any linear combination as well.

U computes DIT of any state

$$U \sum_x f(x)|x\rangle = \sum_x f(x) U(x) = \sum_x f(x) \sum_y K(x,y)|y\rangle$$

$$= \sum_y \left(\sum_x K(x,y) f(x) \right) |y\rangle = \sum_y \tilde{f}(y)|y\rangle$$

8.2 Quantum Fourier Transform(QFT)

U computes DIT of 2^n states in one go. Fourier Transform is a special case of DIT with a specified kernel

$$K(x,y) = \frac{1}{\sqrt{N}} e^{2\pi\iota xy/N} = \frac{1}{\sqrt{N}} \omega_n^{xy}$$

$\omega_n = e^{2\pi\iota/N}$, here x, y are usual decimal numbers and xy is usual multiplication(and not bitwise multiplication)
Example: $n = 1, N = 2$

$$K(x,y) = \frac{1}{\sqrt{2}} e^{2\pi\iota xy/2} = \frac{1}{\sqrt{2}} (-1)^{xy}$$

$x = 0, y = 1$

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

which is the matrix for Hadamard Gate. Thus Hadamard Gate is a QFT in $\mathbb{C}^2$. QFT of $\alpha |0\rangle + \beta |1\rangle$ is its Hadamard Transform.
K is unitary.

$$\tilde{f}(x) = \frac{1}{\sqrt{N}} \sum_y e^{2\pi\iota xy/N} f(y)$$

$$f(y) = \frac{1}{\sqrt{N}} \sum_x e^{-2\pi\iota xy/N} \tilde{f}(x)$$

$$\langle x| KK^\dagger |y\rangle = \sum_{z=0}^{N-1} \langle x| K |z\rangle \langle z| K^\dagger |y\rangle = \sum_{z=0}^{N-1} K(x,z) K^\dagger(z,y) = \frac{1}{N} \sum_{z=0}^{N-1} e^{2\pi\iota z(x-y)}$$

for $x \neq y$, this is a geometric series

$$= \frac{1}{N} \frac{e^{2\pi\iota(x-y)} - 1}{e^{2\pi\iota(x-y)/N} - 1} = 0$$

for $x = y$, each term of the sum is 1 and there are N terms in the series, so that $\langle x| KK^\dagger |y\rangle = \delta_{x,y}$

Implementing QFT
For n qubits: N=2^n basis states; $|x\rangle$ is in the computational basis and $|\tilde{x}\rangle$ is in the fourier basis

$$|\tilde{x}\rangle \equiv QFT |x\rangle = \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} e^{2\pi\iota xy/N} |y\rangle$$

Example: 1-qubit state, N=2

$$QFT\,|x\rangle = \frac{1}{\sqrt{2}}\sum_{y=0}^{2-1} e^{2\pi\iota xy/2}\,|y\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{\iota\pi x}\,|1\rangle\right)$$

$$QFT\,|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \equiv |+\rangle$$

$$QFT\,|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \equiv |-\rangle$$

Let $y = [y_1, y_2, ...y_n]$ be in binary representation, $y = \sum_{k=1}^{n} y_k 2^{n-k}$

$$|x\rangle = |x_1x_2x_3....x_n\rangle = |x_1\rangle \otimes |x_2\rangle \otimes |x_3\rangle \otimes .... \otimes |x_n\rangle$$

$$|\tilde{x}\rangle = \frac{1}{\sqrt{N}}\sum_{y=0}^{N-1} e^{2\pi\iota x\sum_{k=1}^{n} y_k/2^k}\,|y_1y_2...y_n\rangle = \frac{1}{\sqrt{N}}\sum_{y=0}^{N-1}\prod_{k=1}^{n} e^{2\pi\iota x y_k/2^k}\,|y_1y_2...y_n\rangle$$

$$= \frac{1}{\sqrt{N}}\left(|0\rangle + e^{2\pi\iota x/2^1}\,|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi\iota x/2^2}\,|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi\iota x/2^3}\,|1\rangle\right) \otimes .... \left(|0\rangle + e^{2\pi\iota x/2^n}\,|1\rangle\right)$$

Each qubit went from $|x_k\rangle$ to $\left(|0\rangle + e^{2\pi\iota x/2^k}\,|1\rangle\right)$

Two ingredients:

1.Hadamard operator, $H\,|x_k\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{2\pi\iota x_k/2|1\rangle}\right)$

2.Unitary Rotation, $UROT_k\,|x_j\rangle = e^{2\pi\iota x_j/2^k}\,|x_j\rangle$

$$UROT_k = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi\iota/2^k} \end{pmatrix}$$

applies phase $e^{2\pi\iota/2^k}$ for state $|1\rangle$

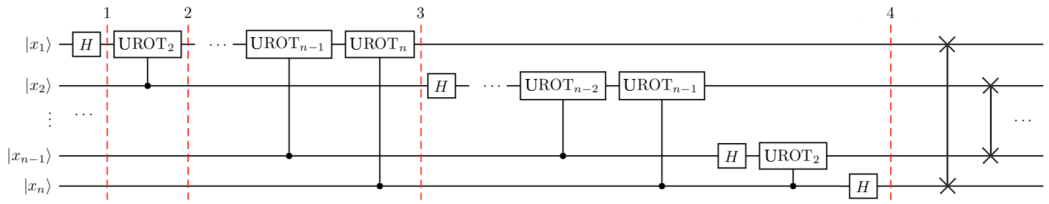


Fig: Quantum Circuit that implements QFT

Step 0: $|x_1x_2x_3....x_n\rangle$

Step 1: $\left(|0\rangle + e^{\frac{2\pi\iota x_1}{2^1}}\,|1\rangle\right) \otimes |x_2x_3....x_n\rangle$

Step 2: $\left(|0\rangle + e^{\frac{2\pi\iota x_2}{2^2}} e^{\frac{2\pi\iota x_1}{2^1}}\,|1\rangle\right) \otimes |x_2x_3....x_n\rangle$

Step 3: $\left(|0\rangle + e^{\frac{2\pi\iota x_n}{2^n}} e^{\frac{2\pi\iota x_{n-1}}{2^{n-1}}} ....e^{\frac{2\pi\iota x_1}{2^1}} |1\rangle\right) \otimes |x_2 x_3 ....x_n\rangle$
$= \left(|0\rangle + e^{\frac{2\pi\iota x}{2^n}} |1\rangle\right) \otimes |x_2 x_3 ....x_n\rangle \quad (x = 2^{n-1}x_1 + 2^{n-2}x_2 + ... + 2^1 x_{n-1} + 2^0 x_n)$
Step 4: $\left(|0\rangle + e^{\frac{2\pi\iota x}{2^n}} |1\rangle\right) \otimes \left(|0\rangle + e^{\frac{2\pi\iota x}{2^{n-1}}} |1\rangle\right) \otimes .... \left(|0\rangle + e^{\frac{2\pi\iota x}{2^1}} |1\rangle\right)$
QFT implemented in reverse order of qubits. Hence, finally reverse the circuit.

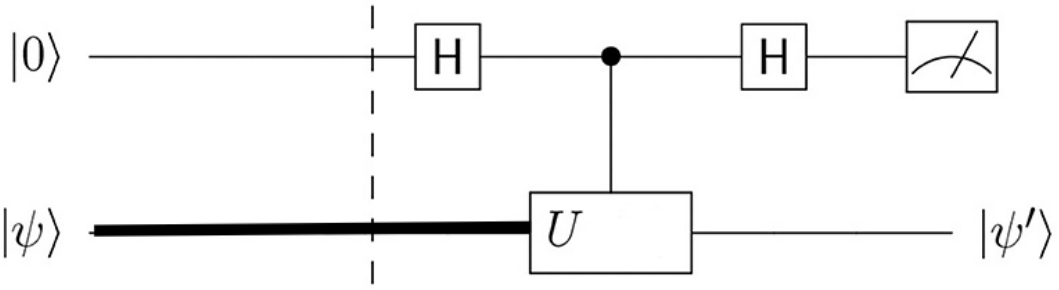
8.3 Quantum Phase Estimation

Quick refresher: A unitary matrix has eigenvalues of the form $e^{\iota\theta}$(i.e.,applying the unitary operator is effectively applying a phase onto its eigenvectors as it preserves the norm) and that it has eigenvectors that form an orthonormal basis.

$$U |\psi\rangle = e^{\iota\theta_\psi} |\psi\rangle$$

Problem: Can we extract θ_ψ(phase) given the ability to prepare state ψ and the ability to apply U?(Basically what it means to measure the global phase?)
Solution: Yes, use Quantum Phase Estimation(QPE)
Why do we care? Hamiltonian evolution (time evolution of real systems) are unitary which has implications in quantum simulations.

QPE trick
Assumption: We have the ability to prepare the state ψ and apply U (U is the controlled-$UROT_k(CROT_k)$: as phase is qubit-dependent and need to add more components to phase with more 1's)



Step 0:$|0\rangle |\psi\rangle$
Step 1:$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) |\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle |\psi\rangle + |1\rangle |\psi\rangle)$
Step 2:$\frac{1}{\sqrt{2}}(|0\rangle |\psi\rangle + |1\rangle e^{\iota\theta_\psi} |\psi\rangle)$
Step 3:$\frac{1}{\sqrt{2}} \left((\frac{|0\rangle+|1\rangle}{\sqrt{2}}) |\psi\rangle + e^{\iota\theta_\psi} (\frac{|0\rangle-|1\rangle}{\sqrt{2}} |\psi\rangle) \right) = \frac{1}{2} \left(|0\rangle (1 + e^{\iota\theta_\psi}) + |1\rangle (1 - e^{\iota\theta_\psi}) \right) |\psi\rangle$

In the difference between the probability of measuring 0 or 1 you have encoded the phase. In other words, we have taken the phase information and turned it into an amplitude that can be measured. This experiment is very painful as it uses only one-qubit to measure θ_ψ(the information(the amplitude) is coming from one qubit in order for us to understand what the phase is). Better idea: get more precision by using more qubits

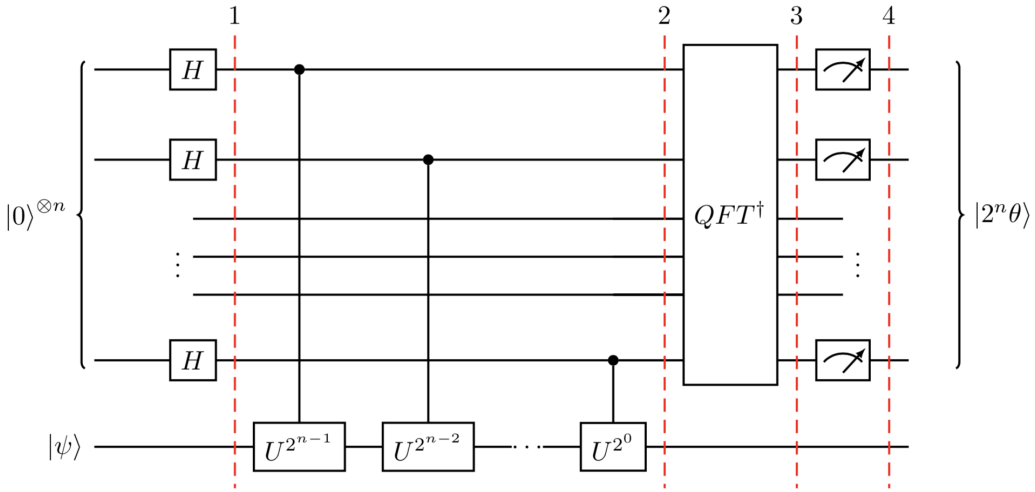


Fig: general quantum circuit for QPE

Given: $U|\psi\rangle = e^{2\pi\iota theta}|\psi\rangle$
Step 0: $|0\rangle^{\otimes n}|\psi\rangle$
Step 1: $\left(\frac{1}{\sqrt{2}}\right)^n(|0\rangle+|1\rangle)^{\otimes n}|\psi\rangle$
Step 2: $\left(\frac{1}{\sqrt{2}}\right)^n(|0\rangle+e^{\iota\theta_\psi 2^{n-1}}|1\rangle)\otimes(|0\rangle+e^{\iota\theta_\psi 2^{n-2}}|1\rangle)\otimes....(|0\rangle+e^{\iota\theta_\psi 2^0}|1\rangle)|\psi\rangle$
Compare it with QFT $|\tilde{x}\rangle = \frac{1}{\sqrt{N}}(|0\rangle+e^{2\pi\iota x/2^1}|1\rangle)\otimes(|0\rangle+e^{2\pi\iota/2^2}|1\rangle)\otimes....(|0\rangle+e^{2\pi\iota/2^n}|1\rangle)$,i.e., QPE is same as QFT except $\theta_\psi \longrightarrow \frac{2\pi x}{2^n}$
Step 3: Inverse QFT Step 4: state $|2^n\theta\rangle$ (where n is the no. of qubits used to estimate θ_ψ)
Earlier when there was one-qubit in first register, there wasn't enough precision to do the measurement, while in this case the phase got multiplied by 2^n which gives more precision(it is more easier to tell the difference between the probabilities of states 0 and 1 in this case)

Quantum Counting
grover's operator instead of Controlled-UROT

8.4 Shor's Factorization Algorithm

Let $N = pq$, where p, q are primes. Given N, choose $m < N$ such that m is coprime with N. Define $f : N \longrightarrow N$ such that

$$F_N(a) = m^a \ mod \ N$$

The smallest $r \in N$ for which $m^r \equiv 1 \ (mod \ N)$ is called the period of the function.
If r is even, factorise $(m^{r/2}+1)(m^{r/2}-1) = 0 \ mod \ N$, $(m^{r/2}-1) \neq 0$ because r is the smallest integer by definition of a period. If $(m^{r/2}+1) \neq 0 \ mod \ N$, $\{p, q\}$ is contained in atleast one of $\{gcd(m^{r/2}+1, N), gcd(m^{r/2}-1, N)\}$; else, the algorithm fails, start all over choosing a different m.
If r is odd, the algorithm fails, start over again.

Implementation
Choose l such that $N^2 < 2^l < 2N^2$. Let $Q = 2^l$. Initialize two l-qubit registers to null state. $|\psi_0\rangle = |0\rangle \otimes |0\rangle$, apply QFT on first register(H gates) to get a uniform linear combination of basis states

$$|\psi_1\rangle = \frac{1}{\sqrt{Q}} \sum_{x=0}^{Q-1} |x\rangle \, |0\rangle$$

Choose a random number $m < N$. Apply the oracle to compute $m^a mod \ N$ and store the result in the second register.

$$|\psi_2\rangle = \frac{1}{\sqrt{Q}} \sum_{x=0}^{Q-1} |x\rangle \, |m^x mod \ N\rangle$$

Measure the second register, this will result in any of the states which are in the second register, say $p = m^s mod \ N$(where p will be in between $m^0 mod \ N$ and $m^r mod \ N$). There are $\left(M = \frac{Q-s}{r} + 1\right)$ states in the first register starting at s, s+r,s+2r.... and ending in Q.

$$|\psi_3\rangle = \frac{1}{\sqrt{M}} \sum_{d=0}^{M-1} |s + d \cdot r, k\rangle$$

where M is the index of the period(number of states which had the measured value p in the second register), s was the first state in the linear combination for which the measurement of the second register gave p.

$$m^{x_0+d \cdot r} = m^{x_0} = k mod \ N$$

Apply QFT on the first register (over $\mathbb{Z}_Q$)

$$|\psi_4\rangle = \frac{1}{\sqrt{MQ}} \sum_{y=0}^{Q-1} \sum_{d=0}^{M-1} e^{2\pi\iota\cdot(s+rd)y/Q} |y,k\rangle$$

$$= \frac{1}{\sqrt{MQ}} \sum_{y=0}^{Q-1} e^{2\pi\iota sy/Q} \left(\sum_{d=0}^{M-1} e^{2\pi\iota rdy/Q} \right) |y,k\rangle$$

$$= \frac{1}{\sqrt{MQ}} \sum_{y=0}^{Q-1} e^{2\pi\iota sy/Q} \left(\sum_{d=0}^{M-1} z^d \right) |y,k\rangle \quad (z = e^{2\pi\iota ry/Q})$$

$$\sum_{d=0}^{M-1} z^d = \frac{1-z^M}{1-z}$$

Since z is unimodular,

$$|\sum_{d=0}^{M-1} z^d| = |\frac{z^{M/2}(z^{-M/2} - z^{M/2})}{z^{1/2}(z^{-1/2} - z^{1/2}} | = |\frac{sin\left(\frac{\pi yrM}{Q}\right)}{sin\left(\frac{\pi yr}{Q}\right)}|$$

Measure the first register. A particular $|y\rangle$ will be measured with probability

$$Pr(y_i) = \frac{1}{MQ} \left(\frac{sin\left(\frac{\pi yrM}{Q}\right)}{sin\left(\frac{\pi yr}{Q}\right)} \right)^2$$

The probabilities show peaks when the argument of sine function in the denominator is a multiple of π, i.e. when $\frac{yr}{Q} = n$, an integer. When $\frac{yr}{Q} = n$, the probability becomes

$$Pr(y_i) = \frac{M^2}{MQ} = \frac{M}{Q}$$

Repeated measurement of the first register can yield information about period.

Chapter 9

Entropy and Information

9.1 Classical Information Theory

9.1.1 Information

Information content in a statement is a measure of uncertainty associated with an event. The amount of uncertainty needs to be quantitatively defined. Two distinct approaches to define information content in a message:

1. Shannon: Number of bits required to be transmitted in order to select the correct answer from a list of previously agreed choices. This gives rise to decision tree.

2. Kolmogorov and Chaitin: Number of bits required to compress a given message

9.1.2 Shannon Entropy

Let us consider a simple statistical experiment of observing a random variable X, which takes one of the values $x_1, x_2, ..., x_M$ with respective probabilities $p_1, ..., p_M$ ($p_i > 0$ for all i and $\sum_i p_i = 1$). When we observe X we gain some information because the uncertainty regarding its value is eliminated. So the information gained is the uncertainty eliminated. We wish to have a mathematical model which gives us a measure of this information gained. A function which measures this information gained or the uncertainty associated with a statistical experiment must depend only on the probabilities p_i and it should be symmetric. This is based on the intuition that changing the

names of the outcomes does not change the uncertainty associated with the random variable X. Let $H(p_1, p_2, ..., p_M)$ be average uncertainty associated with the event $X = x_i$.

The desirable properties of a function H which measures the uncertainty associated with a statistical experiment are:

1. For each fixed M, $H(p_1, p_2, ..., p_M)$ is a nonnegative symmetric function of $p_1, p_2, ..., p_M$.

2. Let $f(M)$ be average uncertainty associated with M equally likely events $f(M) = H(\frac{1}{M}, \frac{1}{M}, ..., \frac{1}{M})$.

   - If $M > M' \implies f(M) > f(M')$, i.e, $f(M)$ is a monotonic function of its argument M.

   - $f(1) = 0$ because for only 1 event there is no uncertainty associated, it is a certainty.

3. $H(p_1, p_2, ..., p_M) = 0$ if and only if one of the p_i's is 1. This corresponds to the case when there is no uncertainty in the outcome of the experiment.

4. Let X and Y be two independent statistical experiments. Let XY denote the experiment where the experiments X and Y are performed together and the output is the ordered pair of the outcomes of X and Y. Then $H(XY) = H(X) + H(Y)$.

5. $H(p_1, p_2, ..., p_M)$ attains its maximum when $p_i = \frac{1}{M}$, for all $i \in 1, 2, ..., M$. That is, we gain maximum information when all possible outcomes are equally likely.

6. $H(p_1, p_2, ..., p_M, 0) = H(p_1, p_2, ..., p_M)$

7. $H(p_1, p_2, ..., p_M)$ is continuous in $p_1, ..., p_M$. This is a natural condition because we would like to say that, if two statistical experiments have the same number of possible outcomes and their associated probabilities are *close*, then the information contained in each of them should also be *close*.

8. Grouping theorem: If a random variable X has M possible outcomes of which group A has outcomes $x_1, x_2, ..., x_r$ and group B has outcomes

$x_{r+1}, x_{r+2}, ..., x_M$, then

$$H(p_1, p_2, ...p_M) - H\left(\sum_{i=1}^{r} p_i, \sum_{i=r+1}^{M} p_i\right)$$

$$= \sum_{i=1}^{r} p_i H\left(\frac{p_1}{\sum_{i=1}^{r} p_i}, ..., \frac{p_r}{\sum_{i=1}^{r} p_i}\right) - \sum_{i=r+1}^{M} p_i H\left(\frac{p_{r+1}}{\sum_{i=r+1}^{M} p_i}, ..., \frac{p_M}{\sum_{i=r+1}^{M} p_i}\right)$$

A function $f(M) = c\ log\ M$ satisfies all these properties. In information theory, c=1 and the base of logarithm is taken to be 2. The uncertainty does not depend on the value that the random variable takes but depends on the probabilities, that is on the uncertainty associated with an event. From grouping theorem,

$$H\left(\frac{1}{s}, \frac{1}{s}, ...\frac{1}{s}\right) - H\left(\frac{r}{s}, \frac{s-r}{s}\right) = \frac{r}{s} H\left(\frac{1}{r}, ...\frac{1}{r}\right) + \frac{s-r}{s} H\left(\frac{1}{s-r}, ...\frac{1}{s-r}\right)$$

$$f(s) = H\left(\frac{r}{s}, \frac{s-r}{s}\right) + \frac{r}{s} f(r) + \frac{s-r}{s} f(s-r)$$

Substitute $f(M) = log\ M$

$$H(p, 1-p) = -[p\ log\ r + (1-p) log(s-r) - log\ s]$$

$$= -[p\ log\ r - p\ logs + p\ log\ s + (1-p) log\ (s-r) - log\ s] = -p\ log\ p + (1-p) log(1-p)$$

$$H(\{p_i\}) = -\sum_{i=1}^{M} p_i\ log_2\ p_i$$

where $H(\{p_i\})$ is the uncertainty function, M is the total number of possible events and the i^{th} event has probability p_i.

Shannon entropy gives a lower bound on the uncertainty associated with an event. Statistical entropy is a measure of disorder in the system, more the disorder more is the entropy. Shannon(Information) entropy is a measure of uncertainty associated with events which occur with different probabilities. More the uncertainty is, higher the Shannon entropy.

9.1.3 Shannon Noiseless Coding Theorem

For uniquely decipherable codes where the letter x_i occurs with a probability p_i, the average length of a word which consists of several letters has a maximum compression given by the entropy function $H = -\sum_{i=1}^{M} p_i\ log_2\ p_i$, i.e,

if the length of the code for the letter x_i is n_i, then

$$\sum_i n_i p_i \geq H$$

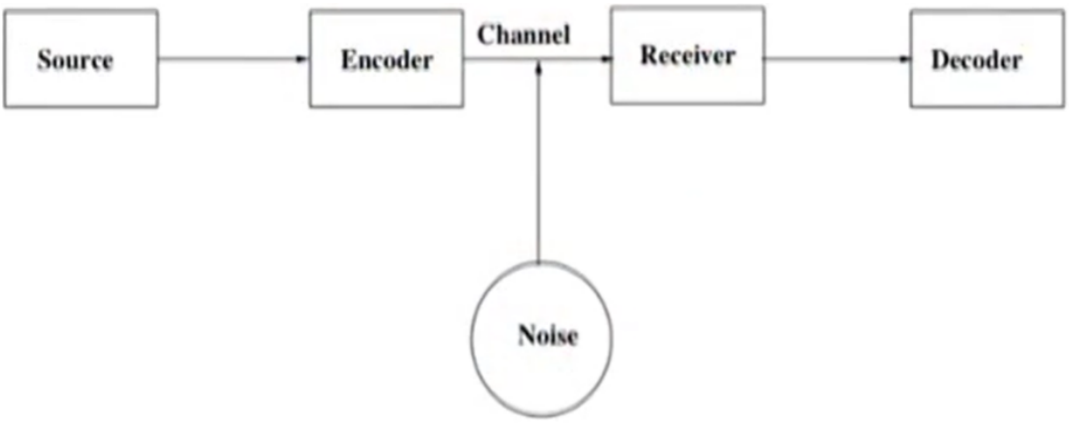


Fig: Communication System

Minimum bits per letter is given by Shannon entropy(Maximum possible compression)

9.2 Quantum Information Theory

9.2.1 Von Neumann Entropy

Like Shannon entropy characterizes uncertainties in classical events, Von Neumann entropy provides a measure of our ignorance of quantum systems.

- An ensemble of quantum system is described by a density matrix ρ, in terms of which $\langle A \rangle = Tr(A\rho)$ where trace is a weighted sum of expectation values of A in all pure states in the ensemble.
- Von Neumann entropy $s(\rho)$ provides a measure of degree of mixedness of the ensemble.
- For a well ordered state $S(\rho) = 0$
- Von Neumann entropy $S(\rho) = -Tr(\rho\, log_2\rho)$
- Entropy must be independent of the basis. Hence, $S(\rho) = -\sum_i \lambda_i\, log_2\lambda_i$
- $S(\rho) \geq 0$ and the equality implies a pure state. Since eigenvalues $\lambda_i \in [0, 1]$ and the value of $log_2\lambda_i \leq 0$, positivity is established. Further

$\lambda_i \, log_2 \lambda_i = 0 \iff \lambda_i = 0, 1$. For a pure state only one eigenvalue is 1, others 0.

- If there are D non-zero eigenvalues of ρ, since $0 \leq \lambda_i \geq 1$ (as ρ is a positive operator with unit trace), maximum value of $S(\rho)$ is $log_2 D$ which corresponds to maximally mixed state with each eigenvalue =1/D
- For a bipartite system AB, $S(\rho_{AB}) \leq S(\rho_A) + S_{\rho_B}$ with equality valid for the case where $\rho_{AB} = \rho_A \otimes \rho_B$. For pure states the eigenvalues of density matrices are identical, with one of the values being 1 and all others 0(follows from $Tr(\rho^2) = 1$. Thus ρ_A and ρ_B have the same eigenvalues, proving the case for equality

Chapter 10

EPR and Bell's Inequalities

10.1 Entangled States

EPR(Einstein Podolsky Rosen) pairs are maximally entangled, e.g. the two-qubit state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is maximally entangled. Take a reduced trace over either the first qubit or the second, we get a maximally mixed state $\rho = \frac{1}{2}$. On making a local measurement on either the first qubit or the second, we get no information about how the system was prepared because a measurement of the first qubit collapses the entangled qubit as well.

10.2 Bell's (gedanken) Experiment

Consider Bell state: $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$. It can be verified that $(\sigma_i^A + \sigma_i^B)\,|\psi^-\rangle = 0$ so that one can replace σ_i^B acting on particle 2 by $-\sigma_i^A$ acting on particle 1.

$$\langle\psi|\,(\sigma^A \cdot \hat{a})(\sigma^B \cdot \hat{b})\,|\psi\rangle = -\,\langle\psi|\,(\sigma^A \cdot \hat{a})(\sigma^A \cdot \hat{b})\,|\psi\rangle = -\sum_{i,j}\langle\psi|\,a_i b_j \sigma_i^A \sigma_j^A\,|\psi\rangle$$

10.3 CHSH Inequality

An entangled state $|\psi^-\rangle$ is prepared and shared by Alice and Bob. Each can measure two properties of the particle she/he holds, Alice measures P_Q, P_R while Bob measures P_S, P_T, each of which can take values +1 or -1. Example: Suppose Alice and Bob have been given a supply of T-shirts. Alice observes the color P_Q(which my be Red(+1) or Blue(-1)) and size P_R(which may

be large(+1) or medium(-1)). Bob measures the price P_S(which may be high(+1)or low(-1)) and quality of its fabric P_T(which may be good(+1)or bad(-1)). The example given is classical but they actually measure some quantum properties having two possible states.

To measure a property each tosses a coin and perform measurement of that property on a random sample in her/his collection as decided by the toss. The measurements are not casually related(each measures a property at a particular time). After a series of measurements, the measurements are tabulated. Out of the table, only those where Alice measure P_Q whereas Bob measured P_S at the same time are filtered and collected. Likewise for pairs P_Q and P_T, P_R and P_S, and P_R and P_T.

Consider the quantity $QS+RS+RT-QT = (Q+R)S+(R-Q)T = \pm 2$, this assumes a local hidden variable, i.e an object has a property independent of its observation. Since each term is ± 2, the average under the assumption of existence of hidden variables is $\langle QS\rangle + \langle RS\rangle + \langle RT\rangle - \langle QT\rangle \leq 2$ [CHSH Inequality]