Physics 160 Lecture Notes

Professor: Mikhail Lukin
Notes typeset by Emma Rosenfeld and Mihir Bhaskar

February 24, 2021

Contents

1 Introduction

There is vast experimental evidence that quantum mechanics is complete and correct. Quantum theory is now the basis for many ubiquitous technologies, such transistors and lasers. We also understand why many objects around us don't behave quantum mechanically: measurement collapses quantum states into one observable outcome or another (for example, dead or alive in the famous case of Schrödinge's cat). In fact, the measurement doesn't have to be active. In practice, physical systems aren't isolated from their environment, and the interaction with the environment has the effect of making many measurements on the systems, such that the dynamics converge according to classical laws of physics.

At the same time, information is closely connected to physics - to acquire information one needs a physical device, to store information one needs a physical system, and to process it one executes physical operations on bits. Is there a limit to information and information processing? A *bit* of information in the classical world is either 0 or 1. As Moore's law continues, a transistor will become the size of one atom. What does that mean

for the classical bit? Can one encode a bit of quantum information in a quantum system, such as an atom, nucleus, or photon? Can we use quantum mechanical evolution to perform computation? In this case, what is the speed and performance? Is there a fundamental limit to Moore's law? Is this a new opportunity? These are the questions we seek to answer in the field of *quantum information processing*. It will rely on two central concepts:

- *Superposition*: the idea that a particle can be in multiple states at the same time. The canonical example is Schrödinger's cat, in which a cat is in a superposition of dead and alive.
- *Entanglement*: the idea that objects in superposition can be linked together and contain mutual information, even if they are physically separated from each other. Einstein argued in the famous EPR (Einstein-Podolsky-Rosen) paradox[1] that quantum mechanics is not complete, because affecting one particle at one side would affect the other, even if they are separated by an arbitrary distance.

1.1 Types of quantum technologies

Motivated by the questions defined above, physicists have identified an opportunity to use quantum physics to process and store information in three specific areas.

1. *Quantum Metrology*. Using the concepts of superposition and entanglement, one can make measurements that's much more precise than any classical device. However, these superpositions are generally fragile.
2. *Quantum Communication*. Quantum communication uses the concepts of superposition and entanglement to transmit information in a secure way. If a third party, an adversary, tries to measure the superposition, they leave a trace because the measurement affects the state. One can therefore detect eavesdroppers, enabling fundamentally secure communication channels.
3. *Quantum Computing*. Suppose one can prepare a register in a superposition of input states, and utilize them in a machine which can take the superposition as input, and then quantum mechanically perform quantum logic (for example, additions, subtractions, etc.). These different inputs interfere with each other throughout the algorithm. This provides the possibility of doing something which is called 'quantum parallelism', which uses interference to perform computations. These devices can *potentially* have much more computational capabilities than classical computers. They are also useful for understanding and simulating the behavior of quantum systems. For example, if one has a many body system from many interacting subsystems and lets it evolve, the resulting state and dynamics would be very challenging for classical computers to simulate. This is the basis for *quantum complexity*, and an opportunity of *quantum simulation*, which is a subset of *quantum computing*.

   However, the field faces a number of serious challenges:

   - Nobody knows how to build truly large scale quantum machines. The largest are only 10's of qubits, and how to bridge this gap is not clear. The main problem is that the qubits are coupling to environment. Theoretically, there are some concepts like quantum error correction and fault tolerance. We know theoretically we should be able to build these machines, but the necessary hardware is well beyond current imaginable experimental capability.
   - It isn't yet clear what quantum computers will be useful for. This is an algorithm challenge. It is possible to build *useful* algorithms which utilize a quantum advantage?

   Nevertheless, now is a special time. In several labs, quantum machines of increasing complexity are being built. For example, 50 or so qubits can maintain superposition and are programmable, similar to the times of 1940's and 1950's for classical computers. The physics of maintaining entangled states and utilizing them is also interesting in general, because it poses an exciting opportunity for science and technology. This course will aim to cover a foundation of quantum information science, and bring the student close to the forefront of the field. There will be three components:

[1] See Einstein, Podolsky, Rosen, *Physical Review* 47, 777 (1935) for the original paper, as well as the "resolution" to this paradox: Bell, *Physics Physique Fizika*, 1, 3 (1964).

(a) Quantum foundations, focusing on physics of open quantum systems; understanding how they evolve; how the dynamics are controlled; understanding entanglement.

(b) Building and using quantum machines, focusing on quantum computers; basics of quantum algorithms and how to implement them; discussion of how the most advanced quantum machines are built.

(c) Connections to practical quantum information systems; students will simulate quantum systems, to further understand why it's hard to do so on classical systems; explore near term quantum computing machines, which are now available as a service; employment of a combination of classical numerics and web-based quantum computer access.

Preskill's lecture notes will form the basis of the course, as a high-level undergraduate or introductory level graduate class. A little bit of programming experience will be helpful.

2 Quantum foundations

2.1 States, Measurements and Observables

How do we describe states, measurements and evolution? Here we review some basics of atomic physics and quantum mechanics.

References: Preskill's notes, Nielsen & Chaung, McMahon

2.1.1 States

States are a vector in an n-dimensional space, called a *Hilbert* space $\mathscr{H}$. We will use Dirac notation for a state with a *ket* by $|\psi\rangle$ in a given basis, specified by n complex amplitudes, $\{a_1, a_2, ..., a_n\}$:

$$|\psi\rangle = \sum_{k=1}^{n} a_k |k\rangle \tag{1}$$

Where $|k\rangle = (0, 0, 0, ..., 1, 0)^T$ has a 1 at the kth component, and zero elsewhere.

The *bra* is given by:

$$\langle\psi| = \sum_{k=1}^{n} a_k^* \langle k| \tag{2}$$

The inner product $\langle\phi|\psi\rangle$ is a complex number given by:

$$\langle\phi|\psi\rangle = \sum_{k=1}^{n} c_k^* a_k \tag{3}$$

With the important properties:$|\langle\phi|\psi\rangle| \leq 1$; and normalization: $\langle\psi|\psi\rangle = 1$.

2.1.2 Operations

Operations take states of physical systems and convert them to other states. The operator A is defined as:

$$\hat{A}|\psi\rangle = |\phi\rangle \tag{4}$$

Such that the following matrix $\hat{A}$:

$$\hat{A} = |\phi\rangle\langle\psi| \tag{5}$$

is a nxn-dimensional matrix if $|\phi\rangle$ and $|\psi\rangle$ are in $\mathscr{H}$.

2.1.3 Observables

An observable is a property of the physical system such that it at least in principle can be measured. It is represented by a *Hermitian* matrix A, which means that it is equal to its complex transpose.

For all Hermitian operators, there is a *spectral decomposition* - one can find a basis in the Hilbert space where this matrix will be diagonal:

$$\hat{A} = \sum_n a_n \hat{P}_n \tag{6}$$

Where a_n are eigenvalues and $\hat{P}_n = |n\rangle \langle n|$ are a complete set of Hermitian operators called *projectors* such that:

$$\hat{P}_n \hat{P}_m = \delta_{n,m} \hat{P}_n \tag{7}$$

The set of projects are a complete set such that they span the entire Hilbert space:

$$\sum_n \hat{P}_n = I. \tag{8}$$

2.1.4 Measurements

The outcome of measurements of the observable $\hat{A}$ for a system in an arbitrary state $|\psi\rangle$ will be one of the eigenvalues a_n with probability $p_n = \langle\psi| \hat{P}_n |\psi\rangle$ and state $\hat{P}_n |\psi\rangle / \sqrt{p_n}$. **The outcome of the measurement is generally probabilistic.**

Remarks (the following is true for any closed quantum system):

1. Repeated measurements will yield the same results as the first measurement, hence 'projection'.
2. One can measure the expectation value, or average value of the observer by repeated measurements after re-preparing the superposition and measuring many times:

$$\langle \hat{A} \rangle = \langle\psi| \hat{A} |\psi\rangle = \sum_n a_n p_n \tag{9}$$

3. If you are given one copy of the unknown state ψ, you will not be able to determine the state with a single measurement. A single measurement on the state does not reveal complete information about it. This is the basis of quantum cryptography.

2.2 Evolution of quantum states

Dynamics are given by the Schrodinger equation:

$$i\hbar \frac{d}{dt} |\psi\rangle = H |\psi\rangle \tag{10}$$

The solution can be obtained by integrating with respect to time (assuming the Hamiltonian is time independent):

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle , \tag{11}$$

where:

$$U(t) = e^{-iHt/\hbar} \tag{12}$$

The operator $U(t)$ is unitary since it is easy to see that $U(t)^{-1} = U(t)^\dagger$.

Remarks:

1. To clarify what we mean by an exponential of a matrix, we make explicit the following definition of operator functions

$$f(A) = \sum_n c_n A^n \tag{13}$$

For an operator A.

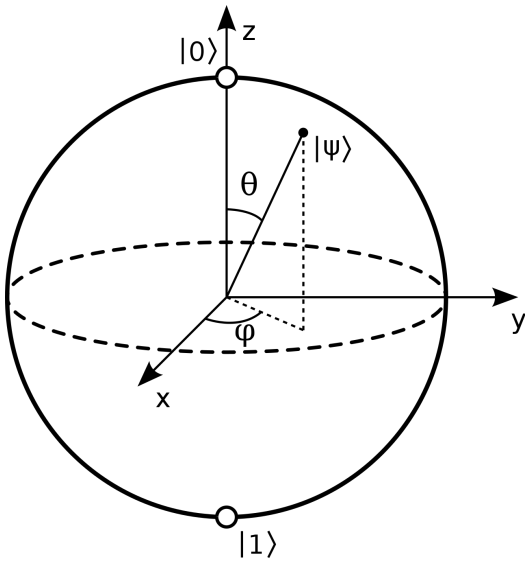


Figure 1: A qubit representation of a Bloch vector on the Bloch sphere. Picture from Wikipedia.

2. The unitary evolution preserves the norm of the state:

$$\langle\psi(0)|\psi(0)\rangle = \langle\psi(t)|\psi(t)\rangle = 1 \tag{14}$$

3. Unitary evolution is time reversible. By sending $t \to -t$, the state $|\psi(t)\rangle$ returns back to its original state.

4. Unitary evolution is both linear and deterministic. This is to be contrasted with measurement, which is fundamentally probabilistic.

2.3 Quantum bits

A qubit is composed of two state systems in a Hilbert space $\mathscr{H} \equiv \{|0\rangle, |1\rangle\}$. Free particles and harmonic oscillators cannot be approximated as quantum bits.

Let's discuss the example of a spin 1/2 system as a qubit. The most general state can be written:

$$|\psi\rangle = c_0 |0\rangle + c_1 |1\rangle \tag{15}$$

Where $|c_0|^2 + |c_1|^2 = 1$ for normalization. One can parameterize the qubit using two angles θ and ϕ, rewriting the state as:

$$|\psi(\theta, \phi)\rangle = \cos\theta/2 \, |0\rangle + e^{i\phi} \sin\theta/2 \, |1\rangle \tag{16}$$

The phase ϕ determines the phase between the two components, and the angle θ determines the probabilities of finding the states in $|0, 1\rangle$. This representation is convenient because one can represent it on the *Bloch sphere*. The Bloch vector is a unit vector in 3D, defined such that:

$$|0\rangle \equiv |\uparrow\rangle \, , \; |1\rangle \equiv |\downarrow\rangle \, . \tag{17}$$

Measurement along the z-basis corresponds to projection along the up or down directions. See figure 1 for a visualization.

The spin matrices $S_{x,y,z}$ are observables and are equal to the Pauli matrices (up to a factor of 1/2):

$$\sigma_{1,2,3} = \sigma_{x,y,z} = \hat{X}, \hat{Y}, \hat{Z} \tag{18}$$

Remarks:

1. Properties of Pauli matrices:
$$\sigma_\alpha^2 = \mathbb{1} \tag{19}$$
So the eigenvalues are ± 1.

2. The Pauli matrices are traceless:
$$\mathrm{Tr}\sigma_\alpha = 0 \tag{20}$$

3. One can write the Pauli matrices as:
$$\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1| \tag{21}$$
$$\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0| \tag{22}$$
$$\sigma_y = i|1\rangle\langle 0| - i|0\rangle\langle 1| \tag{23}$$

4. The eigenstates are: $Z : \{|0\rangle, |1\rangle\}$, $X : \{(|0\rangle + |1\rangle)/\sqrt{2}, (|0\rangle - |1\rangle)/\sqrt{2}\} \equiv \{|+_x\rangle, |-_x\rangle\}$, and $Y : \{(|0\rangle + i|1\rangle)/\sqrt{2}, (|0\rangle - i|1\rangle)/\sqrt{2}\} \equiv \{|+_y\rangle, |-_y\rangle\}$.
On the Bloch sphere, the states $|\pm_x\rangle$ correspond to vectors pointing along the $\pm x$ direction, and the same for y.

5. The commutation relation holds:
$$[\sigma_\alpha, \sigma_\beta] = 2i\epsilon_{\alpha\beta\gamma}\sigma_\gamma \tag{24}$$
such that the product of two Pauli matrices is another Pauli matrix.

6. The set $\{\sigma_x, \sigma_y, \sigma_z, \mathbb{1}\}$ forms a complete basis for the 2x2 matrices: any 2x2 operator can be expressed as a linear combination of these operators.

2.4 Quantum dynamics

Considering a general form of a Hamiltonian:
$$H = \frac{\hbar}{2}\sum_{i=1}^{3}\sigma_i \equiv \frac{\hbar\omega}{2}\sigma_{\vec{n}} \tag{25}$$
Where $\omega = \sqrt{(\sum h_i^2)}$, $n_i = \frac{h_i}{\omega}$ such that $\vec{n}$ is a unit vector. The vector $\vec{n}$ can be associated with the direction of a magnetic field, where the Bloch vector is the spin that precesses around the field.

Quantum mechanically, we can describe this precession through the unitary evolution:
$$U = e^{-iHt/\hbar} = \mathbb{1}\cos\frac{\omega t}{2} - i\sin\frac{\omega t}{2}\vec{n}\cdot\vec{\sigma} \tag{26}$$
For example, let $\vec{n} = \hat{z}$. Then $U = \mathbb{1}\cos\frac{\omega t}{2} - i\sin\frac{\omega t}{2}\sigma_z$. Multiplying an original state $|\psi(\theta,\phi)\rangle$ by this unitary, we can solve for its time evolution:
$$|\psi(t)\rangle = e^{-i\omega t/2}(\cos\theta/2|0\rangle + e^{i\phi+i\omega t}\sin\theta/2|1\rangle) \tag{27}$$
These dynamics correspond to the Bloch vector rotating about the z axis on the Bloch sphere, at frequency ω. More generally, the vector $\vec{n}$ is the direction of the magnetic field that the Bloch vector rotates around on the Bloch sphere.

Remarks

1. *Probability amplitude method.* Define the ansatz:
$$|\psi(t)\rangle \equiv a_0(t)|0\rangle + a_1(t)|1\rangle \tag{28}$$

The time dependence is fully encoded in the coefficients. By plugging into the Schrodinger equation, we recover a system of two linear differential equations for $a_{0,1}$:

$$\begin{aligned} i\frac{da_0}{dt} &= \frac{h_3}{2}a_0 + \frac{(h_1 - ih_2)}{2}a_1, \\ i\frac{da_1}{dt} &= \frac{-h_3}{2}a_1 + \frac{(h_1 + ih_2)}{2}a_0 \end{aligned} \tag{29}$$

For example, for $h_1 = h_2 = 0$, there is free precession about the z axis. For $h_3 = 0$, we get two equations:

$$\begin{aligned} \frac{d^2a_0}{dt^2} &= -i\Omega\frac{da_1}{dt} \\ \frac{d^2a_1}{dt^2} &= i\Omega^*\frac{da_0}{dt} \end{aligned} \tag{30}$$

Where $\Omega \equiv \frac{(h_1 - ih_2)}{2}$. The solutions are sines and cosines, which is what we call *Rabi oscillation* about the axis given by $\vec{n}$.

2. *Time dependent fields. Reference: AMO II, physics 285b lecture notes*

For a time dependent Rabi frequency, we can incorporate the time dependence as a retarded time using a change of variables:

$$d\tau \equiv dt\Omega \tag{31}$$

Then the solution is sines and cosines of the integral $\int \Omega d\tau$. To flip the state from up to down, using a so-called a π pulse, the time is set such that $\int \Omega(t)dt = \pi$.

Considering $|h_3| >> \Omega$, with $\Omega = \Omega_0 e^{i\nu t}$, with $\nu \sim h_3$, we define another change of variables:

$$\begin{aligned} a_0 &= \tilde{a_0}e^{i\nu t/2} \\ a_1 &= \tilde{a_1}e^{-i\nu t/2} \end{aligned} \tag{32}$$

Plugging these equations into (30), one derives for $a_{\{0,1\}}$:

$$\begin{aligned} \dot{\tilde{a_0}} &= i\frac{h_3 - \nu}{2}\tilde{a_0} + ... \\ \dot{\tilde{a_1}} &= i\frac{h_3 - \nu}{2}\tilde{a_1} + ... \end{aligned} \tag{33}$$

When $\nu = h_3$, then the drive is *resonant* with the atom and the $\tilde{a_0}$ and $\tilde{a_1}$ will rotate into each other. Such a system could be created with a two level atom driven by an electromagnetic wave at a frequency close to its energy splitting. Such a phenomenon is called *resonance*. Transforming the a_0 and a_1 components as in (32) can be viewed as transforming under a unitary:

$$U = e^{i\nu t/2\sigma_z}. \tag{34}$$

2.5 Quantum operations

Below we define some 'quantum operations' which are unitary:

$$\begin{aligned} \hat{n} = \hat{z}, \omega t/2 = \pi/2 \\ \rightarrow U = i\hat{Z} \end{aligned} \tag{35}$$

$$\begin{aligned} \hat{n} = \hat{x}, \omega t/2 = \pi/2 \\ \rightarrow U = i\hat{X} \end{aligned} \tag{36}$$

For a Hadamard gate H:

$$\begin{aligned} \hat{n} = \hat{x}/\sqrt{2} + \hat{z}/\sqrt{2}, \omega t/2 = \pi/2 \\ U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \end{aligned} \tag{37}$$

Such a unitary is called a *Hadamard gate*.

The *Solovay-Kitaev theorem* shows that with these gates as well as a $\pi/8$ rotation about Z, one can generate an arbitrary rotation on the Bloch sphere.

2.6 Tensor products

Reference: McMahon Consider a multi-partite system (also known as a composite system), with $N = 2$ qubits. Qubit A(B) has a Hilbert space $\mathscr{H}_{A(B)}$ of dimension $d_{A(B)}$. The dimension of the total Hilbert space $\mathscr{H}$ is $d_A d_B$. An example of a state in $\mathscr{H}$ can be written as:

$$|\psi_{AB}\rangle = |\phi_A\rangle \otimes |\chi_B\rangle \tag{38}$$

Where $|\phi_A\rangle$ is in H_A and $|\chi_B\rangle$ is in H_B.

Remarks:

1. If $\{|n_A\rangle\} \in \mathscr{H}_A, \{|n_B\rangle\} \in \mathscr{H}_B$, where $\{|n_{\{A,B\}}\rangle\}$ is an orthonormal basis for $\mathscr{H}_{A,B}$, then $\{|n_A\rangle \otimes |n_B\rangle\}$ is a basis for $\mathscr{H}$. The most general state can be written as a linear combination of these basis vectors:

$$|\psi_{AB}\rangle = \sum_{n_A, n_B} |n_A\rangle |n_B\rangle \langle n_A| \langle n_B|\psi_{AB}\rangle \tag{39}$$

   If the following is true:

$$|\psi_{A,B}\rangle \neq |\psi_A\rangle \otimes |\psi_B\rangle \tag{40}$$

   Then $|\psi_{A,B}\rangle$ is an *entangled state*.

2. If an operator A acts in H_A and B acts in H_B, then:

$$A \otimes B\,|\psi_{AB}\rangle = A\,|\phi_A\rangle \otimes B\,|\chi_B\rangle\,. \tag{41}$$

3. Considering states:

$$\begin{aligned} |\phi_A\rangle &= \begin{pmatrix} a \\ b \end{pmatrix} \\ |\chi_A\rangle &= \begin{pmatrix} c \\ d \end{pmatrix} \end{aligned} \tag{42}$$

   The definition of the product state is:

$$|\phi_A\rangle \otimes |\chi_B\rangle = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix} \tag{43}$$

   For example, consider two qubits in $|0\rangle_A |0\rangle_B$, with the notation e.g. $|01\rangle \equiv |0\rangle_A \otimes |1\rangle_B$.

4. We make the following definitions:

$$\begin{aligned} |\Phi^{\pm}\rangle &= \frac{1}{\sqrt{2}}(|00\rangle \pm 11) \\ |\Psi^{\pm}\rangle &= \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \end{aligned} \tag{44}$$

   These four states are called the *Bell states*, and they form a basis called the *Bell basis*.

2.7 Density operator

References: John Preskill's notes

Suppose that we have a quantum state of a composite system described by the following:

$$|\psi_{AB}\rangle = a\,|0\rangle_A\,|0\rangle_B + b\,|1\rangle_A\,|1\rangle_B\,, \tag{45}$$

such that $|a|^2+|b|^2= 1$. Suppose that one only can measure qubit A, and only cares about system A. We consider observables $M_A \otimes \mathbb{1}_B$. The expectation value is:

$$\langle\psi_{AB}|\,M_A \otimes \mathbb{1}_B\,|\psi_{AB}\rangle = |a|^2\langle 0|\,M_A\,|0\rangle_A + |b|^2\langle 1|\,M_A\,|1\rangle_B = \mathrm{Tr}M_A\rho_A \tag{46}$$

Where $\rho_A = |a|^2|0\rangle\langle 0|_A + |b|^2|1\rangle\langle 1|_B$. We call ρ_A the *density operator* for subsystem A. The physical meaning is the following: ρ_A is an ensemble of possible quantum states, each occuring with some probability. In this case, one example is preparing state $|0\rangle_A$ with probability $|a|^2$ and state $|1\rangle_A$ with probability $|b|^2$.

For the state $|\Phi^{\pm}\rangle$, $|a|^2= |b|^2= \frac{1}{2}$. In this case, $\langle\sigma_{x,y,z}\rangle = 0$. One can compare this with the single qubit state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, where $\langle\sigma_x\rangle \neq 0$ but $\langle\sigma_{y,z}\rangle = 0$. The results obtained for a two -qubit entangled state is much different from a single qubit state. For this reason, the density operator is described as a statistical mixture of pure states, and can describe a more general picture of quantum states than the pure states.

We can also change the basis using a unitary operation: $\rho_A \rightarrow U\rho_A U^{\dagger}$. For the case of the statistical mixture presented above ($|a|^2= |b|^2= 1/2$), we see that under any unitary transformation (in any basis), we have $\rho_A \rightarrow U\frac{\mathbb{1}}{2}U^{\dagger} = \mathbb{1}/2$. Physically, this means that a maximally mixed state is a statistical mixture *no matter what basis we measure in.*

This is in contrast with a pure state (for example $a = 1$ or $b = 1$), in which one can always find basis with a deterministic measurement outcome. For example, we can consider the pure state described by the density matrix $\rho = |+_x\rangle\langle +_x|$. If we measure along the Z-axis, we will of course get up and down with probability 1/2. However, along the x-axis, we will always measure +x, which is in contrast to the maximally mixed state.

Next we will look at how a state can evolve from a pure state (e.g. $|0\rangle$) into a mixed state, as described above. This clearly cannot happen as a result of unitary dynamics alone, and will thus require a new formalism. To start, we consider one large closed quantum system, which can be decomposed into two subsystems. In general, we can write the system in this way

$$|\psi_{AB}\rangle = \sum_{i,\mu} a_{i,\mu} \otimes |i\rangle_A\,|\mu\rangle_B\,. \tag{47}$$

To calculate the expectation value of some operator in subspace A, $\hat{M}_A$, this is equivalent to computing:

$$\begin{aligned} \langle\hat{M}_A\rangle &= \langle\hat{M}_A \otimes \mathbb{I}_B\rangle \\ &= \sum_{j,\nu,i,\mu} a^*_{j\nu}a_{i\mu}\,\langle j_A|\,\langle\nu_B|\,\hat{M}_A \otimes \mathbb{I}_B\,|i\rangle_A\,|\mu\rangle_B \\ &= \sum_{i,j,\mu} a^*_{j\mu}a_{i\mu}\,\langle j|_A\,\hat{M}_A\,|i\rangle_A = Tr_A\hat{M}_A\rho_A \end{aligned} \tag{48}$$

This is because we can insert identity $\mathbb{I}_A = \sum_k |k_A\rangle\langle k_A|$, and we have

$$\rho_A = \sum_{i,j,\mu} a_{i\mu}a^*_{j\mu}\,|i\rangle_A\,\langle j|_A \tag{49}$$

The physical interpretation is the following. The matrix $|\psi\rangle_{AB}\langle\psi|$ is the density operator for the full system A and B. But suppose we only care about the degrees of freedom of system A, and do not care at all about subsystem B. We can simply trace over all of the degrees of freedom of subsystem B to obtain the *reduced density operator* ρ_A, which contains all of the information we have about subsystem A. This density operator is a much more general description of a system than just a pure state, since it allows us to describe subsystem A even if it is part of a larger subsystem AB. There are some important properties of this density operator:

- Hermitian: $\rho_A^{\dagger} = \rho_A$
- ρ_A is positive

- $\langle \rho^2 \rangle \leq \langle \rho \rangle$
- tr $\rho_A = \sum_{i,\mu} |a_{i,\mu}|^2 = 1$

Since the density operator has eigenvalues that are real and positive, and should sum to one, we can write it in a diagonal form $\rho_A = \sum_\alpha p_\alpha |\psi_\alpha\rangle \langle\psi_\alpha|$ where $\sum_\alpha p_\alpha = 1$. The state described by this density matrix can be viewed as drawn from an ensemble of different quantum states $|\psi_\alpha\rangle$ which are each drawn with probability p_α. This gives us an intuition for why ρ_A must be positive: in order for this density matrix to have this physical meaning, the probability p_A of drawing a state $|\psi_A\rangle$ must be positive. In quantum mechanics, we can have amplitudes that are negative and complex. This is what makes quantum mechanics unique, and makes quantum computers potentially powerful. However, the underlying probabilities for measurement outcomes *must always be positive.*

From this result, we can make a few remarks.

- If and only if $p_\alpha = 1$ and all others are $0 \rightarrow \rho_A = |\psi\rangle \langle\psi|$ is pure
- In the ensemble interpretation, $\{|\psi_\alpha\rangle, p_\alpha\}$ is a mixed state as described above.
- The origin of mixed states arises from entanglement with the environment. In other words, our subsystem of interest (A) was entangled with subsystem (B, also known as the environment).
- The density matrix elements are:
$$\langle i| \rho |j\rangle = \sum_\mu a_{i,\mu} a^*_{j,\mu} \tag{50}$$
We can clearly see that the indices sum to 1: $\sum_{i=1}^d \rho_{ii}$. In addition, we see that the off diagonal elements obey the Hermiticity condition: $\rho^*_{ij} = \rho_{ji}$. These density matrix elements are an important description of quantum systems. the diagonal elements ρ_{ii}, are known as the *populations* (i.e. the probabilities to measure the system in a particular state), and ρ_{ij} are known as the *coherences*, which are nonzero when there is some well-defined phase between the states described by indices i and j.
- The density operator for a single qubit can be written as:
$$\begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \tag{51}$$
For a qubit in a pure state, we can write down the density matrix immediately from the state:
$$\rho = (c_0 |0\rangle + c_1 |1\rangle)(c_0^* \langle 0| + c_i^* \langle 1|) \tag{52}$$
and we see that the matrix elements will obey $\rho_{10}\rho_{01} = \rho_{11}\rho_{00}$. We can now write the density operator as a superposition of identity and all Pauli matrices:
$$\hat{\rho} = \frac{1}{2}(\mathbb{I} + \vec{P} \cdot \vec{\sigma}) \tag{53}$$
where
$$\det\rho = \frac{1}{4}(1 - \vec{P}^2) \geq 0 \tag{54}$$
because all eigenvalues are positive. It is easy to see that for $|\vec{P}| = 1$, it is clear that ρ describes a pure state. For general (not necessarily pure) states, $|\vec{P}| \leq 1$, where $|\vec{P}| = 0$ corresponds to a maximally mixed state. In fact, it turns out that $\vec{P}$ is a generalization of the *Bloch vector*, and that $\vec{P}^2$ is a measure of the degree of purity of the state in question.

2.8 Generalized evolution

We now have a generalized description of a subsystem A which is not necessarily in a pure state. Suppose initially we did prepare subsystem A in a pure state. How might the pure state describing subsystem A lose its purity?

This must occur via some dynamics involving interactions between A and the environment B. These dynamics will be described by a Hamiltonian of the general form:

$$H = H_A + H_B + H_{AB}, \tag{55}$$

where $H_A(H_B)$ is the Hamiltonian acting only on system A (B), and H_{AB} is the *interaction Hamiltonian* describing the interaction between system A and B.

Let us assume for simplicity that at initial $t = 0$, we can write down the density matrix in the form:

$$\rho_{AB} = \rho_A \otimes |E_B\rangle \langle E_B| . \tag{56}$$

At some later time t, what is ρ_{AB}? For an isolated total system AB, we can describe the system with a state vector and its dynamics with unitary evolution: $|\psi_{AB}\rangle \to U |\psi_{AB}\rangle$. For a density matrix $\rho = \sum_\alpha p_\alpha |\psi_\alpha\rangle \langle\psi_\alpha|$, we will have evolution $\rho \to U\rho U^\dagger$.

At this point, we are describing Schrodinger equation evolution to the full system AB. Now we take this description and trace over the degrees of freedom of subsystem B. This will transform the reduced density operator for subsystem A:

$$\begin{aligned} \rho_A \to \rho_A^{'} &= \mathrm{Tr}_B \rho_{AB}^{'} \\ &= \sum_\mu \langle\mu| U_{AB} |E_B\rangle \rho_A \langle E_B| U_{AB}^\dagger |\mu\rangle \end{aligned} \tag{57}$$

Let us define a new set of operators:

$$M_\mu = \langle\mu| U_{AB} |E_B\rangle \tag{58}$$

which act only on system A. These are known as *Kraus operators*, and they act in only subsystem A. Writing the evolution in the compact form:

$$\rho_A^{'} = \sum_\mu M_\mu \rho M_\mu^\dagger. \tag{59}$$

This expression is known as the Kraus operator sum representation, which is the *most general description of the evolution of quantum states*. There are a few interesting properties to note:

- First, the operators form a completee set:

$$\sum_\mu M_\mu^\dagger M_\mu = \hat{I}_A. \tag{60}$$

  Equation (60) follows from the fact that summing over all basis vectors in subsystem B yields the identity operator: $\sum_\mu |\mu\rangle \langle\mu| = \hat{I}_B$

- This map is linear, Hermitian ($\rho_A^{'\dagger} = \rho_A^{'}$), and preserves positivity. In other words, this map preserves all of the important properties of the density operator.

One can show that any linear map that preserves the trace and which is completely positive will always have this operator sum representation (59). It is also possible to show that if a system with an operator sum representation can be understood as dynamics of a pure state & coherent evolution on a larger, extended Hilbert space (see Preskill's note). Mixed states in general can be thought of as originating as a pure state on a larger system; the most general dynamics of quantum systems can be understood as unitary evolution on a larger Hilbert space. For example, subsystem A is a qubit, and subsystem B is an environment in the lab. Next, we will discuss specific examples of such non-unitary evolution on system A.

2.9 Examples - quantum channels.

These processes describe how qubits lose their coherence.

2.9.1 Depolarization channel of the qubit

Suppose with probability $1-p$, the state of the system is preserved, and with probability $p/3$ the bit flips about the x axis ($\sigma_x |\psi\rangle$), with probability $p/3$ the phase flips ($\sigma_z |\psi\rangle$), and with probability $p/3$ both occur ($\sigma_y |\psi\rangle$). How can we think about these noise processes? We can use Krauss operators. In this case, they are:

$$\begin{aligned} M_0 &= \sqrt{1-p}\mathbb{1} \\ M_{1,2,3} &= \sqrt{\frac{p}{3}}\sigma_{x,y,z} \end{aligned} \tag{61}$$

Such that the density operator evolves as:

$$\rho \to \rho' = (1-p)\rho + \sum_{\alpha=1,2,3} (p/3)\sigma_\alpha \rho \sigma_\alpha \tag{62}$$

For example, a corresponding unitary evolution on the extended subspace is:

$$U_{AB} |\psi\rangle_A |0\rangle_B = \sqrt{1-p} |\psi\rangle_A |0\rangle_B + \sqrt{\frac{p}{3}} \sum_\alpha \sigma_\alpha |\psi\rangle_A |\alpha\rangle_B , \tag{63}$$

where all the $|\alpha\rangle$ are orthogonal to $|0\rangle$ as well as to each other. The environment keeps track of what the error was (the reason why this is not intuitive is because at this stage we have not specified what the interaction is between the system and environment, and we have not specified anything about the environment's degrees of freedom). Note that these states $\{|\alpha\rangle\}$ are not necessarily unique, but in general, there exists a unitary description using a larger Hilbert space that describes decoherence of the smaller system A.

2.9.2 Dephasing channel of the qubit

Suppose with probability $1-p$ the state is preserved, and with probability p there is just a phase flip ($\sigma_z |\psi\rangle$). In this case, the two Kraus operators are:

$$\begin{aligned} M_0 &= \sqrt{1-p}\mathbb{1} \\ M_1 &= \sqrt{p}\sigma_z \end{aligned} \tag{64}$$

Such that:

$$\begin{aligned} \rho \to \rho' &= (1-p)\rho + p\sigma_z \rho \sigma_z \\ &= \begin{pmatrix} \rho_{00} & (1-2p)\rho_{01} \\ (1-2p)\rho_{10} & \rho_{11} \end{pmatrix} \end{aligned} \tag{65}$$

When $p = 1$, the off diagonal components pick up a minus sign. The inherent randomness of the errors eventually destroys the coherence of the qubits. However, in this particular case, if $p = 1$, the system undergoes the dynamics where with unity probability the phase is flipped. This is coherent evolution because the phase is flipped deterministically. The coherent evolution of the larger Hilbert space for general p is:

$$\begin{aligned} & U_{AB}(\alpha |0\rangle_A + \beta |1\rangle_A) |0\rangle_B \\ = \alpha |0\rangle_A (\sqrt{1-p} |0\rangle_B + \sqrt{p} |1\rangle_B) + \beta |1\rangle_A (\sqrt{1-p} |0\rangle_B - \sqrt{p} |1\rangle_B) \end{aligned} \tag{66}$$

Note if the qubit starts in state $|0\rangle_A$ or $|1\rangle_A$, the dephasing channel has no effect. When $p = 1/2$, the two states of system B multiplying the possible states of A are orthogonal to each other - the environment keeps track of the error, consequently, decoherence is due to the environment measuring the state of the system and that when tracing over the environment, we lose this information.

2.10 Master equation for density operator

Consider again subsystems A and B which each evolve with their own Hamiltonian and have some interaction between them:

$$H = H_A + H_B + H_{AB} \tag{67}$$

Can we find the equations of motion for the density operator for subsystem A? What is $\dot{\rho}_A$? If $H_{AB} = 0$, and $\rho_A(0) = \sum_\alpha p_\alpha |\psi\rangle_A \langle\psi|_A$. the density operator for subsystem A evolves as:

$$\dot{\rho}_A = -\frac{i}{\hbar}[H_A, \rho_A] \tag{68}$$

This can be found from the Schrodinger equation. However, the case where $H_{AB} \neq 0$ is much more complicated. Now the dynamics depend on the interaction H_{AB} but it also depends on H_B. To derive the equation of motion for ρ_A, the so-called *master equation*, we first recall a necessary concept, Fermi's Golden Rule.

2.10.1 Fermi's golden rule

Let's consider one closed system which has many states. We assume that there is some Hamiltonian which couples the initial state to every state of a continuum, but not between states of the continuum (suppose the Hamiltonian is diagonal in the continuum subspace, for simplicity). Fermi's golden rule is used to describe a bound (well defined ground state) to continuum transition (many excited states $|1\rangle_k$). We a wave function description:

$$|\psi(t)\rangle = c_0(t)|0\rangle + \sum_k c_k(t)|k\rangle \tag{69}$$

The evolution equations are, from the Schrodinger equation:

$$\begin{aligned} \dot{c}_0 &= -i\frac{E_0}{\hbar}c_0 - i\sum_k \frac{\langle 0|H|k\rangle}{\hbar}c_k \\ \dot{c}_k &= -i\frac{E_k}{\hbar}c_k - i\sum_k \frac{\langle k|H|0\rangle}{\hbar}c_k \end{aligned} \tag{70}$$

First we make a transformation to eliminate diagonal elements:

$$c_j = \tilde{c}_j e^{-iE_j t/\hbar} \tag{71}$$

Substituting into the Schrodinger equation:

$$\begin{aligned} \dot{\tilde{c}}_0 &= -i\sum_k \frac{\langle 0|H|k\rangle}{\hbar}\tilde{c}_k e^{-i(E_k-E_0)t/\hbar} \\ \dot{\tilde{c}}_k &= -i\sum_k \frac{\langle k|H|0\rangle}{\hbar}\tilde{c}_k e^{-i(E_0-E_k)t/\hbar} \end{aligned} \tag{72}$$

Suppose the system starts in the state $|0\rangle, c_0 = 1$. If we assume $\tilde{c}_0 = 1$, and invoke perturbation theory, the first order correction will be:

$$\tilde{c}_k = \frac{\langle k|H|0\rangle}{E_0 - E_k}(e^{-i(E_0-E_k)t/\hbar} - 1) \tag{73}$$

Using this description, the total probability of the system leaving the state $|0\rangle$ is:

$$p = \sum_k |c_k|^2 = \sum_k \frac{|\langle k|H|0\rangle|^2 \sin^2((E_0 - E_k)t/\hbar)}{(E_0 - E_k)^2} \tag{74}$$

In short time, the components of each sum evolves quadratically in time, since $\sin((E_0 - E_k)t/\hbar) \approx (E_0 - E_k)t/\hbar$. In the limiting case where the states $\{|k\rangle\}$ is a very broad continuum, then we must replace the sum with an integral:

$$p = \int \frac{|\langle k|H|0\rangle|^2 \sin^2((E_0 - E_k)t/\hbar)}{(E_0 - E_k)^2}\rho(E_k)\, \mathrm{d}E_k \tag{75}$$

The $\rho(E_k)$ here is a density of states, namely, the number of states with energy E_k. This is a sharply peaked function, where only the states in which $E_0 \sim E_k$ contributes, such that the probability becomes:

$$p = \rho(E_0)|\langle k|H|0\rangle|^2 \int_{-\infty}^{+\infty} \frac{\sin^2(E_0 - E_k)t/\hbar}{(E_0 - E_k)^2}\, \mathrm{d}E_k \tag{76}$$

Making a change of variables to $(E_0 - E_k)t$, we find that:

$$p \sim \gamma t, \tag{77}$$

where $\gamma = \pi\rho(E_0)|\langle k| H |0\rangle|^2$. The probability of staying in zero after a short time t is $p_0 = 1 - \gamma t$. The probability leaks from state $|0\rangle$ to the continuum. Note that there is no return - the probability of staying in zero is monotonically decreasing. We can also describe these dynamics by an effective Hamiltonian evolution:

$$H_{eff} = (E_0 + i\gamma) |0\rangle \langle 0| \tag{78}$$

This imaginary term effectively describes departure of population from the initial state. We will describe this as a model for the *Markovian environment*. The environment is large (lots of states), dense in its spectrum (many modes around resonance) and it is featureless. If the information leaks from the system to the environment, it immediately disappears and it never comes back. In this case, this type of environment, a Markovian environment, is described by a Markov process that has no memory. The erasure of this memory is related to the linear dependence in time of the probability of the leakage in probability. If the dynamics are Markovian we can derive the master equation.

2.10.2 Density operator evolution for Markovian environments

Suppose the density operator $\rho(t)$ is given at some time t. Let us assume that the evolution is linear:

$$\rho(t + \delta t) = \rho(t) + \mathcal{O}(\delta t), \tag{79}$$

where $\mathcal{O}(\delta t)$ is linear in δt. We know that the most general description of evolution of the density matrix can be written using the Kraus operator sum, so we can also write

$$\rho(t + \delta t) = \sum_\mu M_\mu \rho M_\mu^\dagger. \tag{80}$$

We now investigate the question, what are the possible forms of M_μ?

- First, there must be one that acts like identity and are linear in δt.

  $$M_0 = \hat{I} + O(\delta t) = \hat{O} + (K - iH)\delta t, \tag{81}$$

  where K and H are Hermitian operators.

- Next, there must be operators that are proportional to $\sqrt{\delta t}$.

- We must preserve normalization:

  $$\sum_\mu M_\mu^\dagger M_\mu = \hat{I} \quad (\mu \neq 0). \tag{82}$$

  This tells us that we must have K to be

  $$K = -\frac{1}{2}\sum_{\mu>0} L_\mu^\dagger L_\mu. \tag{83}$$

Now we can evaluate the differential:

$$\dot{\rho} = \frac{\rho(t + \delta t) - \rho(t)}{\delta t} = (K - iH)\rho + \rho(K + iH) + \sum_\mu L_\mu \rho L_\mu^\dagger. \tag{84}$$

We can now rewrite the terms in a more illuminating fashion. Taking the first two terms, we recognize that the terms involving H can be rewritten in terms of the commutator, and we can rewrite K using equation (83):

$$(K - iH)\rho + \rho(K + iH) = -i[H, \rho] - \frac{1}{2}\sum_\mu (L_\mu^\dagger L_\mu \rho + \rho L_\mu^\dagger L_\mu). \tag{85}$$

The resulting equation with substitution given in (85) is the most general evolution of the density operator for a Markovian environment, and is called the *master equation*.

Remarks:

1. What is the meaning of K? K plays the role of an imaginary, non-Hermitian component of the Hamiltonian. In fact, it is sometimes called the "non-Hermitian" correction to the Hamiltonian. This is the term that results in the decay we expect, for example from H_{eff} given in equation (78). However, we know that probability should be preserved, since this is a general description of the evolution of ρ, and leaving the evolution only described by K would result in decaying total probabilities with time.

2. This brings us to L_μ, the so-called *quantum jump operators*. They *project* the density operator into certain states, as if the environment is measuring the subsystem described by ρ. By doing so, they also preserve the normalization of the state described by ρ. If we have a pure-state $\rho = |\psi\rangle\langle\psi|$, under application of the jump operators, we have a state $L_\mu \langle\psi| |\psi\rangle L_\mu^\dagger$, which preserves normalization.

In summary, the general description of quantum dynamics, the *master equation*, can be viewed as evolution under a non-Hermitian Hamiltonian ($K - iH$ term), along with quantum jumps ($\sum_\mu L_\mu \rho L_\mu^\dagger$), which preserve normalization.

2.10.3 Example: spontaneous emission

For example, consider an atom with two electronic states $|0\rangle$ and $|1\rangle$ (such as two orbitals of the Hydrogen atom). If the atom is in an excited state $|1\rangle$, we know the atom can emit a photon into the environment and undergo a transition $|1\rangle_A |0\rangle_B \to |0\rangle_A |1\rangle_B$, where $|\psi\rangle_A$ denotes the state of the atom and $|0\rangle_B$ and $|1\rangle_B$ denotes the absence or presence of a photon in the environment.

Note that in practice, although we are treating just two orbitals of a single atom, the total system is very complicated: it contains, in principle, all of the modes of the electromagnetic field in the entire universe! Our master equation formalism allows us to treat *just the degrees of freedom of the atom*, and describe its interaction with the environment. The jump operator $L = |0\rangle\langle 1| \sqrt{\gamma}$, which describes transitions from the excited to the ground state at rate γ, describes this interaction. Now using (83), we can also write down the dissipation term, $K = -\frac{\gamma}{2} |1\rangle\langle 1|$. Remarkably, this is exactly the non-Hermitian term we wanted to add in the first place, such as in our un-physical Hamiltonian (78). The key here, is that the jump operator L allows us to maintain normalization even in the presence of this non-Hermitian evolution.

Let's examine how these L operators by applying the quantum jump to an arbitrary density operator:

$$\rho \to L\rho L^\dagger = |0\rangle \rho_{11} \langle 0| . \tag{86}$$

The jump operators force the system into the state $|0\rangle$, which is what we expect when the atom undergoes a spontaneous emission event: after the jump, the system always ends in the state $|0\rangle$.

Suppose we could measure the environment, for example by putting detectors all around the atom. If at some point, the detector clicks, we know the photon was emitted. By recording a click, we know the atom *must* be in the completely pure, normalized $|0\rangle$ state, thus purifying a potentially mixed state through measurement.

Now let us describe the evolution of the density operator under spontaneous emission. To do so, we can examine how the density matrix elements $\langle i|\rho|j\rangle = \rho_{ij}$ evolve with time. We use the master equation to arrive at the so-called *optical Bloch equations*:

$$\begin{aligned} \dot{\rho}_{11} &= -\gamma\rho_{11} \\ \dot{\rho}_{00} &= \gamma\rho_{11} \\ \dot{\rho}_{01} &= -\frac{\gamma}{2}\rho_{01}. \end{aligned} \tag{87}$$

The first equation tells us how probability decays from state $|1\rangle$ due to spontaneous emission. Note however, that the second equation tells us that the population of $|0\rangle$ increases accordingly, since when population leaves $|1\rangle$ it *must be preserved*. The normalization condition $\rho_{00} + \rho_{11} = 1$ is preserved by these equations.

The final term is a little less intuitive. We can get a sense of what is happening by considering a pure state $\alpha|0\rangle + \beta|1\rangle$. If we now include the state of the environment, we see that spontaneous emission results in entanglement with the environment:

$$\alpha(\sqrt{1-p}\,|1\rangle |0\rangle_E + \sqrt{p}\,|0\rangle |1\rangle_E) + \beta\,|0\rangle |0\rangle_E . \tag{88}$$

The coherence (or relative phase) between $|0\rangle$ and $|1\rangle$ of the atom is decaying as a result of spontaneous emission, and this is described in the master equation for $\dot{\rho}_{10}$. Why is the decay rate of the coherence $\frac{1}{2}$ of that of

the population? Let us consider the description of the system with the non-Hermitian Hamiltonian (78). The equation of motion for amplitude for the state $|1\rangle$ is:

$$\begin{aligned} \dot{c}_1 &= -\frac{\gamma}{2} c_1 \\ \frac{d}{dt}|c_1|^2 &= -\gamma |c_1|^2 \end{aligned} \tag{89}$$

However, remember that this non-Hermitian Hamiltonian, does not couple to $|0\rangle$ and therefore $\dot{c}_0$ is unchanged. Considering the coherence term $\rho_{01} = c_0 c_1^*$, we can see this will only decay at a rate $\frac{\gamma}{2}$, unlike the probability $|c_1|^2$ itself, which decays at γ. Of course, the master equation also includes the jump terms which help preserve normalization, but these do not contribute to the rate at which population is transferred between $|0\rangle$ and $|1\rangle$, since that is only described by the non-Hermitian term K.

The master equation allows for treatment of coherent and incoherent dynamics simultaneously. For example, if a quantum computer has some finite interaction with the environment, this formalism will allow us to calculate, the coherent gate errors arising from incoherent dynamics. We did not yet discuss the coherent dynamics in our example master equation, but one can simply employ the relevant Hamiltonian in the master equation for H, calculate the commutation relations, and obtain the result. As a shortcut however, one can just write down the Schrödinger equation (which describes only the coherent evolution), and combine this with the simple equations governing the incoherent evolution given in equation (87). The two methods are equivalent, and yield the results:

$$\begin{aligned} \dot{\rho}_{11} &= -\gamma \rho_{11} + i\Omega \rho_{10} - i\Omega^* \rho_{01} \\ \dot{\rho}_{00} &= \gamma \rho_{11} - i\Omega \rho_{10} + i\Omega^* \rho_{01} \\ \dot{\rho}_{01} &= -\frac{\gamma}{2} \rho_{01} - i h_z \rho_{01} - i\Omega(\rho_{00} - \rho_{11}). \end{aligned} \tag{90}$$

Note that there is yet another equation for $\dot{\rho}_{10}$, but it is simply the complex conjugate of the third master equation: $\dot{\rho}_{10} = \dot{\rho}_{01}^*$. Let us examine the dynamics in general. In the case of $h_z = 0$, the coherent evolution is a solution of sines and cosines, allowing for Rabi oscillations between state $|0\rangle$ and $|1\rangle$. However, the incoherent dynamics will result in a decay of the oscillations to some steady-state values. In fact, we can calculate the steady state values by setting all of the time-derivatives to zero and solving the resulting system of linear equations. The steady state population ρ_{11} will generically be less than $\frac{1}{2}$, since spontaneous emission always favors putting us back to $|0\rangle$. Even in the case of a very strong drive, in the steady state, the population in $|1\rangle$ is at most $\rho_{11} = \frac{1}{2}$.

Remarks:

1. We can return to the intuition of a non-Hermitian Hamiltonian

$$H_{eff} = H + iK = H - i\frac{\gamma}{2} |1\rangle \langle 1| \,(+\text{jumps}). \tag{91}$$

   Here, we add these quantum jumps to keep the state normalization and allow the description to be physical. The quantum jumps are especially helpful since it allows us to incorporate the physics of measuring the system into the dynamics. For example, we can consider performing a measurement and then doing feedback to control the state of the system, and this is a growing tool in quantum information processing.

2. We can also consider the role of *dephasing*, rather than spontaneous decay. In this case, the jump operator will be $L = \sqrt{\gamma} Z$, and the non-Hermitian operator will be $K = -\frac{\gamma}{2}\hat{I}$. In this case, note that while K looks rather insignificant, since it is non-Hermitian, it must be treated with care as it contributes to a reduction in the purity of the state.

   We have found before that this dephasing will not affect the populations, but will destroy the coherence. The purely incoherent dynamics is as follows, given by the master equation:

$$\begin{aligned} \dot{\rho}_{00} &= 0 \\ \dot{\rho}_{11} &= 0 \\ \dot{\rho}_{01} &= -2\gamma \rho_{01}. \end{aligned} \tag{92}$$

3. For a physical picture of dephasing, suppose we have a qubit which evolves under a time-dependent magnetic field in the z direction, $h_z(t)$. The phase ϕ in the x-y plane undergoes evolution described by precession around the z-axis with frequency given by h_z. In this time-dependent case, we have, as seen before with a change of variables $\tau = h_z(t)dt$, that the phase between states $|0\rangle$ and $|1\rangle$ becomes at time t:

$$\phi(t) = \int_0^t h_z(t')dt'. \tag{93}$$

This evolution is deterministic, given an $h_z(t)$. However assume now that $h_z(t)$ is a random variable, drawn probabilistically from some distribution such that it has a random stochastic value. The phase will now undergo a *random* diffusion such that $\langle\phi\rangle = 0$, but the variance will grow such that $\langle\phi^2\rangle \sim \gamma t$, consistent with the master equation. Note that this is not a quantum mechanical average, but a statistical average over the random process of phase accumulation over a random $h_z(t)$: if we make a measurement of the qubit in the x-y plane, we will no longer have a well-defined result, but rather a random result. *This is what we mean when we say that a qubit has lost its coherence.* Since the density matrix represents a statistical average over each of these possible $h_z(t)$, the off-diagonal matrix elements, the coherences, go to zero as $t \to \infty$.

4. The picture defined above allows us to extend our intuition to a *non-Markovian process*. Now suppose that our field $h_z(t)$ is random, but changes very slowly. For example, each experiment, we may have a random $h_z(t)$. However, within an experiment, h_z may not change, or may change very slowly. In this case, the dynamics will not be described by a master equation (recall that our derivation started with an assumption of a Markovian environment, or a very dense continuum of states in Fermi's Golden Rule).

While this may appear unfortunate for our understanding of the dynamics of the system, these slow, non-Markovian dynamics can often be eliminated. Suppose after some evolution time t under a random (but nearly static h_z), we rotate the state of the qubit by π around the x-axis. Now, if the system evolves for another period t, the system will return to its original state, as if there were no stochastic h_z in the first place. This is the idea of so-called *dynamical decoupling*, which is a key tool in quantum information processing, also known as *spin-echo* in the NMR community.

2.11 Generalized measurements

Armed with a description of coherent and incoherent dynamics, we will now turn our discussion to measurements. How is the probabilistic, projective nature of measurements consistent with our quantum description thus far? How can we describe measurements in general? We will now begin to explore these questions.

As we have previously discussed, ideal projective measurements are described by projectors $\{P_i\}$ such that $P_iP_j = \delta_{ij}$, such that the density matrix is transformed to $\rho \to P_i\rho P_i$, yielding the pure state described by P_i. The measurement occurs with probability $p_i = \mathrm{Tr}_A P_i\rho$.

However, we must have a description of more general measurements. In practice, states are not pure after most measurements in the classical world; for example, touching a table and measuring its dimensions does not turn it into a pure quantum state. Moreover, we have not yet considered the measurement device and its effect on the system, for example, photo-detectors destroy the photon that they detect. In particular, now we will use the description of coupled system dynamics to consider a model of realistic measurements.

Suppose we want to measure system A. Suppose it couples to system B at time $t = 0$, and we let the two systems evolve. Then eventually, we measure system B at time t. We will assume that we measure system B completely quantum mechanically (e.g. a projective measurement). For concreteness, system B in this case could describe the detector with which the measurement is made, such as a photo-detector that clicks when a photon (system A) impinges on it.

Recalling the discussion of Kraus operators, we describe the most general dynamics:

$$|\phi\rangle_A |0\rangle_B \to U_{AB} |\phi\rangle_A |0\rangle_B \tag{94}$$

Inserting unity partitioned by the basis vectors of B:

$$\begin{aligned} |\phi\rangle_A |0\rangle_B &\to \sum_\mu |\mu\rangle_B \langle\mu|_B U_{AB} |\phi\rangle_A |0\rangle_B \\ &= \sum_\mu |\mu\rangle_B M_\mu |\phi\rangle_A , \end{aligned} \tag{95}$$

where ere we recognize M_μ as the Kraus operator, here called 'measurement operators'. Next, we perform a projective measurement of system B, in which we find the system B in one of the states $|\mu\rangle_B$. From equation (95), this occurs with probability:

$$p_\mu = \langle\phi|_A M_\mu^\dagger M_\mu |\phi\rangle_A , \tag{96}$$

and the state of system B is projected into the state $M_\mu |\phi\rangle_A / \sqrt{p_\mu}$. Note that these operators M_μ are not necessarily projectors - they are Kraus operators (the projection was another operator, acting on system B). Also, the operators M_μ do not have to sum to unity (in general $\sum_\mu M_\mu \neq \mathbb{1}$) and they are not necessarily Hermitian, since they are not generally projectors.

To formalize this description, we introduce another operator:

$$F_\mu \equiv M_\mu^\dagger M_\mu \tag{97}$$

The probability of projecting system B into state μ can now be written as $p_\mu = \mathrm{Tr}_A F_\mu \rho_A$. These operators F_μ have the following properties:

1. F_μ are Hermitian.
2. F_μ are positive, since they describe probability outcomes.
3. $\sum_\mu F_\mu^\dagger F_\mu = \mathbb{1}$; the operators F_μ are complete.

These set of operators are called **Positive Operator Value Measure** (POVM). They are a formalism that describes general measurements on a quantum system.

Remarks:

1. The measurement described by F_μ is very similar to conventional projective measurements, however they are not necessarily projectors, namely, $F_i F_j \neq \delta_{ij} F_i$ in general. If one specifies the POVM F_μ, one cannot specify the post measurement state uniquely (since that is defined by the *measurement operator* M_μ, and there are multiple M_μ that yield the same F_μ).
2. Any POVM on some Hilbert space on system A can be realized as projective measurements on a larger Hilbert space (in this case, the projective measurement on system A and B is equivalent to the projective measurement on system B, since we assume all states $\{|\mu\rangle\}$ are distinct and orthogonal). See 'Neumark's theorem' in Preskill's notes.

2.11.1 Example: POVM

In the context of distinguishing non-orthogonal states, POVMs can be utilized. Consider two parties Alice and Bob, where Alice sends to Bob one two possible (non orthogonal) states: $|\psi_1\rangle = |0\rangle$ and $|\psi_2\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Bob would like to perform a measurement that determines which state she sent.

Bob constructs a detector that implements the following POVM. Consider the POVM $F_1 = |1\rangle\langle 1| (\frac{\sqrt{2}}{1+\sqrt{2}})$ and $F_2 = \frac{\sqrt{2}}{1+\sqrt{2}} \frac{1}{2}(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)$; and lastly, necessarily, $F_3 = \mathbb{1} - F_1 - F_2$ to satisfy the completeness condition. If Bob measures F_1, he knows with certainty that $|0\rangle$ was not sent. Likewise, if Bob measures F_2, then he knows that ψ_2 was not sent. These choices would constitute a perfect set of measurements with no ambiguity if the POVM was only $\{F_1, F_2\}$. However, there must be some ambiguity since the states are not orthogonal: Bob can also measure F_3. The third POVM F_3 is important because it guarantees that the operators sum to unity, and allows for some ambiguity in trying to distinguish non-orthogonal states.

Remarks:

1. The post-measurement state is not defined by POVM. For example, consider the weak measurement POVM:

$$\begin{aligned} F_0 &= |0\rangle\langle 0| + (1-\epsilon)|1\rangle\langle 1| \\ F_1 &= \epsilon |1\rangle\langle 1| \end{aligned} \tag{98}$$

   If $\epsilon \to 0$, then F_0 becomes unity - the measurement only weakly perturbs the system. Naively, one may say that the post-measurement state is proportional to $F_\mu |\phi\rangle_A$. But, in our original description, the operator acting on the state is the Kraus operator M_μ, not necessarily F_μ. However, the M_μ are not unique for a particular F_μ! For example, $M_0 = |0\rangle\langle 0| + e^{i\phi}\sqrt{1-\epsilon}\,|1\rangle\langle 1|$ satisfies $F_0 = M_0^\dagger M_0$ for any ϕ.

2. One way to interpret these effects is in terms of open quantum systems and decoherence & depolarization. In this case, the environment measures the system and learns something about it. If one could keep track of all degrees of freedom of the environment, then we could learn exactly the information that it gains about the system. In practice the environment contains many degrees of freedom, and we must trace over them (add all outcomes). That is precisely the motivation for the operator sum representation and Kraus evolution: entanglement, measurement and decoherence are all very closely connected.

2.12 Entanglement

We will now move on from measurements and generalized dynamics, and formalize our definition of entanglement, which plays a role in all of the concepts we have discussed thus far. We define a state as *entangled* when:

$$|\psi\rangle_{AB} \neq |\psi\rangle_A \otimes |\psi\rangle_B \tag{99}$$

This is a description of real-world systems, and as we have seen, it can be thought of as an 'enemy' of purity when describing system-environment interactions and measurements, but it is also a resource for quantum information. In the following sections, we will discuss how entanglement is a resource.

2.12.1 Example: Bell states

After preparing the *Bell state* $|\phi^+\rangle_{AB} = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$, and measuring the state of each qubit in the computational basis, the outcomes of the measurements will be correlated. This correlation is sometimes called *non-local correlation*, because the qubits A and B can be physically separated from each other. This can be used for superdense coding - by sending one qubit from one location to another, two classical bits can be sent.

2.12.2 Example: multiple qubits

Consider multiple qubits and *quantum function evaluation* which consists of two registers:

$$|01....10\rangle\,|00...1\rangle \tag{100}$$

Where the first register is called x and the second is the function is the function evaluation $f(x)$. Consider the unitary that evaluates $f(x)$:

$$U\sum_i |x_i\rangle\,|0...0\rangle \rightarrow \sum_i |x_i\rangle\,|f(x_i)\rangle \tag{101}$$

For N qubits, there are 2^N combinations of basis states. To specify the state of this type, one requires 2^N complex numbers! This becomes intractable very quickly for a classical computer and is the basis of quantum simulation and quantum supremacy.

2.13 Properties of Bell states

The *Bell states* are special two-qubit entangled states, which are employed in many applications of quantum entanglement. They are described by the expressions:

$$\begin{aligned} |\Phi^\pm\rangle &= \frac{1}{\sqrt{2}}(\,|00\rangle \pm |11\rangle\,) \\ |\Psi^\pm\rangle &= \frac{1}{\sqrt{2}}(\,|01\rangle \pm |10\rangle\,). \end{aligned} \tag{102}$$

The notations of using $|\Phi^\pm\rangle$ and $|\Psi^\pm\rangle$ are quite widely used by the community: the Φ or Ψ describes whether the qubits are in the same or different states, respectively, and the $\pm$ describes whether or not there is a π phase between the two terms. Important properties about the Bell states include:

1. The Bell states form an orthogonal basis in $\mathcal{H}_A \otimes \mathcal{H}_B$.

2. They translate into one another via *local operations*, which are operations acting only on one of the qubits. For example, consider $X_A\,|\Phi^+\rangle$. The X operation flips the first qubit on both states $|00\rangle$ and $|11\rangle$ in the superposition $|\Phi^+\rangle$, so we clearly see that $X_A\,|\Phi^+\rangle = |\Psi^+\rangle$.

3. The expectation value of any single-particle observable is zero. Mathematically, $\forall \sigma_\alpha^{\{A,B\}}, \langle \sigma_\alpha^{\{A,B\}} \rangle = 0$.

4. The Bell states are *simultaneous eigenstates* of two-qubit operators $X_A X_B$, $Y_A Y_B$, $Z_A Z_B$. We can see this, for example, by considering $X_A X_B |\Phi^+\rangle$. Each X_i flips its respective bit, leaving us with $X_A X_B |\Phi^+\rangle = |\Phi^+\rangle$. This is in contrast to the case of a single qubit, where $[X, Y] = 0$ ensures that a single qubit cannot be a simultaneous eigenstate of X and Y. However, note that $[X_A X_B, Y_A Y_B] = 0$. This will have some very important consequences, for example, in the Bell inequalities as we will learn later on. While single qubit expectation values are all zero, the two-qubit expectation values $\langle \sigma_{A,i} \sigma_{B,i} \rangle$ are in general nonzero.

5. Consider a partial measurement, a measurement on just one of the qubits. For example, consider a measurement of only qubit A in the Z basis. The possible outcomes are as follows:

$$\begin{aligned} 0 &\to |0_A\rangle \langle 0_A|\Phi^+\rangle \to |0_A\rangle |0_B\rangle \\ 1 &\to |1_A\rangle \langle 1_A|\Phi^+\rangle \to |1_A\rangle |1_B\rangle . \end{aligned} \tag{103}$$

   Now consider also a measurement of the same state in the X basis, where we have the possible outcomes:

$$\begin{aligned} + &\to |+_{X_A}\rangle \langle +_{X_A}|\Phi_{AB}+\rangle = |+_{X_A}\rangle (|0_B\rangle + |1_B\rangle) = |+_{X_A}\rangle |+_{X_B}\rangle \\ - &\to |-_{X_A}\rangle \langle -_{X_A}|\Phi_{AB}+\rangle = |-_{X_A}\rangle (|0_B\rangle - |1_B\rangle) = |-_{X_A}\rangle |-_{X_B}\rangle . \end{aligned} \tag{104}$$

   If we measure one of the qubits, we will always get a well-defined outcome on the second state. This outcome will be determined by (1) the basis of the measurement and (2) the outcome of the measurement. The Bell states have some *non-classical correlation* where the correlations between qubits extends beyond a single basis and can be maintained simultaneously in multiple non-commuting bases.

6. If we trace over one of the qubits, we get a completely mixed state. Mathematically, the reduced density matrix is described by:

$$\rho_A = \mathrm{Tr}_B |\Psi_{AB}\rangle \langle \Psi_{AB}| = \frac{1}{2}\mathbb{1}. \tag{105}$$

   This has important consequences for classifying entangled states. If we are given an arbitrary state, we can ask (1) are the qubits entangled? And (2) how entangled are they? This could be for a general entangled state in a large Hilbert space, but we can get a good intuition from the two qubit case, where we see that tracing over one of the qubits in the *maximally entangled* Bell-states results in a maximally mixed state. This is powerful, because given an arbitrary quantum state, it is not always obvious whether or not it is entangled. For example, consider the two states:

$$\begin{aligned} |\psi\rangle &= \frac{1}{2}(\,|00\rangle + |01\rangle + |10\rangle + |11\rangle\,) \\ |\psi\rangle &= \frac{1}{2}(\,|00\rangle + |01\rangle + |10\rangle - |11\rangle\,). \end{aligned} \tag{106}$$

   These states seem similar, but the first state is a simple product (unentangled) state $|+_x +_x\rangle$, whereas the second state is a maximally entangled state. In this case, examining at the reduced density matrix could help us identify a lack of entanglement in the first state, and the presence of entanglement in the second.

2.14 Schmidt Decomposition

The Schmidt decomposition is a tool used to quantify and test for entanglement within a two-qubit pure state. The Schmidt decomposition can be formulated as a theorem as follows:

Theorem: $\forall$ pure state $|\psi_{AB}\rangle$ can be written as

$$|\psi_{AB}\rangle = \sum_i \sqrt{M_i}\, |u_i\rangle_A |v_i\rangle_B , \tag{107}$$

where $\{|u_i\rangle\}, \{|v_i\rangle\}$ are an orthonormal basis in $\mathcal{H}_A$ and $\mathcal{H}_B$. $M_i \geq 0$ are known as the Schmidt coefficients.

Proof: The reduced density operator for system A is $\rho_A = \mathrm{Tr}_B |\psi\rangle_{AB} \langle\psi|$. We choose the basis $\{|n_A\rangle\}$ for subsystem A such that ρ_A is diagonal. We can decompose the state as:

$$|\psi_{AB}\rangle = \sum_{n,\mu} a_{n,\mu} |n_A\rangle |\mu_B\rangle \equiv \sum_n |n_A\rangle \otimes |\tilde{n}_B\rangle \tag{108}$$

where we have defined $|\tilde{n}_B\rangle \equiv \sum_\mu a_{n,\mu} |\mu_B\rangle$, which denotes the change of basis for subsystem B for choosing the basis where ρ_A is diagonal. Now the reduced density matrix is:

$$\rho_A = \mathrm{Tr}_B \sum_{n,m} |n_A\rangle |\tilde{n}_B\rangle \langle m_A| \langle \tilde{m}_B| = \sum_{n,m} |n_A\rangle \langle m_A| \, |\langle \tilde{m}_B|\tilde{n}_B\rangle|^2. \tag{109}$$

Recall we chose the basis $\{|n_A\rangle\}$ such that ρ_A was diagonal, so it follows that $\langle \tilde{m}_B|\tilde{n}_B\rangle$ is proportional to $\delta_{\tilde{m},\tilde{n}}$, and specifically $\langle \tilde{m}_B|\tilde{n}_B\rangle \equiv M_n \delta_{\tilde{m},\tilde{n}}$. From this, (107) follows. □

Examining for example the state:

$$|\psi\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle), \tag{110}$$

it is clear this is not in its Schmidt form given by equation (107). Rather, the description of the state as $|\psi\rangle = |+_{A,x}\rangle |+_{B,x}\rangle$ is the Schmidt form, in which it is clearly not entangled.

Remarks:

1. Consider the concept of *purification of mixed states*. Any density operator $\rho_A = \sum_i M_i |u_i\rangle \langle u_i|$ can always be represented as a pure state $|\psi_{AB}\rangle$ in some extended Hilbert space given by (107). This is intimately connected to what we have studied previously: generalized evolution, and POVMs can always be understood in terms of unitary evolution in a larger Hilbert space.

2. Schmidt decomposition eq (107), tells us whether or not a state is an entangled state of two subsystems. In particular, $|\psi_{AB}\rangle$ is separable if and only if one $M_i = 1$ and all other $M_j = 0$. We can also see this in terms of the reduced density operator ρ_A: if ρ_A it is a pure state, then the state is separable.

   On the other hand, we can consider the opposite limit: the state is *maximally entangled* if all $M_i = \frac{1}{d}$ where $d = \min(\dim\mathcal{H}_A, \dim\mathcal{H}_B)$. We can therefore define a degree of entanglement as the purity of the reduced density operator. For example, if we want to quantify the degree of entanglement of two qubits, we can simply trace over one of the qubits and calculate the length of the remaining Bloch vector to understand the degree of entanglement.

3. *Entanglement entropy*, also known as the von Neumann entropy is defined as:

$$S(|\psi_{AB}\rangle) \equiv S(\rho_A) = -\mathrm{Tr}\rho_A \log_2 \rho_A = -\sum_x M_x \log_2 M_x. \tag{111}$$

   This quantity is the analogue of Shannon entropy in classical information theory. For example, if the state $|\psi_{AB}\rangle$ is pure, we will have $S(|\psi_{AB}\rangle) = 0$ since only one $M_i = 1$ and the rest are zero. For a maximally entangled state, we will have $S(|\psi_{AB}\rangle) = \log_2 d$.

   This metric is very useful in theory for classifying entanglement in large systems. However, it is generally not as useful experimentally. One reason is because the term $\log_2\rho$ is challenging to measure. There is a family of these entropy measures, for example Reyni entropy, which is proportional to ρ^2, which can be measured in some experimental systems. However, the main challenge is that we assume explicitly that we started with a pure state of the two subsystems $|\psi_{AB}\rangle$. In practice, no system is completely isolated from the environment, so it is very challenging to start in a pure state. Then, when we measure an impure reduced density matrix ρ_A, it is very challenging to ensure that any mixedness or impurity, arises from entanglement with subsystem B (generally desirable) rather than with the rest of the environment (which is undesirable).

2.15 Unitary evolution of entangled states

Now that we have defined entanglement more formally, we will discuss how entangled states evolve with time, or the dynamics of entangled states.

Suppose we have two systems which evolve under a Hamiltonian $H = H_A + H_B$. In this case, the unitary evolution can be written as a product of unitary evolution on the individual subsystems: $U = U_A \otimes U_B$. Can such evolution change the degree of entanglement? Examining the Schmidt decomposition, eq (107) shows that this unitary evolution will not change individual Schmidt numbers, but simply transform the underlying basis vectors

$|u_j\rangle_A |v_j\rangle_B$: in order to change the degree of entanglement, we will need some component of our Hamiltonian H_{AB} which acts on both components of our state. Practically, there must be an interaction between the qubits that governs and evolves their joint quantum state $|\psi_{AB}\rangle$ as a combination of qubits.

We can further examine this evolution by returning to the states described by eqs (106). To change the first state into the second, the sign of the $|11\rangle$ component only must be changed. This will correspond to the Hamiltonian:

$$\begin{aligned} H_{AB} &= g\,|1\rangle_A \langle 1| \otimes |1\rangle_B \langle 1| \\ &\equiv \frac{g}{2}(\mathbb{1}_A - Z_A)(\mathbb{1}_B - Z_B). \end{aligned} \tag{112}$$

This Hamiltonian is proportional to $Z_A Z_B$, which is a *product* of the Pauli operators on each system A and B. Such a Hamiltonian, H_{AB}, is required in order to change the degree entanglement the two-qubit system, since any other Hamiltonians, H_A and H_B, correspond to unitary operations on individual subsystems A and B. This Hamiltonian maps the basis states as follows:

$$\begin{aligned} |00\rangle &\to |00\rangle \\ |01\rangle &\to |11\rangle \\ |10\rangle &\to |10\rangle \\ |11\rangle &\to e^{-i\phi t}\,|11\rangle \end{aligned} \tag{113}$$

We can implement the desired $-$ sign on the $|11\rangle$ component by choosing the time $gt = \pi$, which result in the so-called *controlled-phase gate*, or *controlled-Z rotation*. We can see logically that this operation will apply a π phase flip on the second (target) qubit conditional on the state of the first (control) qubit.

One interesting property of the controlled-phase unitary is that the control and target qubits and the same result still occurs. This is a unique property of the controlled-phase gate, and is why we draw the diagram for this controlled-phase gate in a symmetric form.

Another important two-qubit gate is a *controlled-NOT (CNOT)* gate. This is a specific implementation of the broader class of the controlled-unitary gates. In these gates, conditioned on the state of the control qubit, we implement a unitary rotation U on the target qubit. For the CNOT gate, $U = X$, so we invert the target qubit state, conditional on the state of the first qubit (i.e., only if the first qubit is in state $|1\rangle$). The CNOT gate has the following mapping:

$$\begin{aligned} |00\rangle &\to |00\rangle \\ |01\rangle &\to |10\rangle \\ |10\rangle &\to |11\rangle \\ |11\rangle &\to |10\rangle\,. \end{aligned} \tag{114}$$

One can show that the CNOT operation can be obtained from the C-phase gate using two Hadamard transformations: $CNOT \equiv H_B CZ H_B$. One important limitation on these multi-qubit unitaries is the *no-cloning* theorem. This states that if we have a system in the state $|\phi_A\rangle |0_B\rangle$, where $|\phi_A\rangle$ is an arbitrary quantum state, there exists no unitary U_{AB} that clones the state $|\phi_A\rangle$ into the second register: there is no such U_{AB} that satisfies $U_{AB} |\phi_A\rangle |0_B\rangle = |\phi_A\rangle |\phi_B\rangle$ for general states $|\phi\rangle$. The no-cloning theorem can be proved by taking an inner product of two instances:

$$\langle \psi_A | \phi_A \rangle = \langle \psi_A U_{AB}^\dagger | U_{AB} \phi_A \rangle = \langle \psi_A | \phi_A \rangle \langle \psi_B | \phi_B \rangle\,, \tag{115}$$

And thus $|\phi\rangle$ and $|\psi\rangle$ cannot be general states. This is the basis for the field of *quantum cryptography (quantum key distribution)*. It is interesting to note that if cloning did in fact exist, one could actually use entanglement to communicate faster than the speed of light, as explored on the homework.

2.16 Information content of Bell States

.

As we have seen, single qubit measurements of Bell states always yield expectation values of zero. That begs the question, what is the information content is there in Bell states? There are actually *two* bits of information:

1. *Parity*: the parity is measured in the ZZ basis and distinguishes whether the Bell state is in the $\{\Phi\}$ or $\{\Psi\}$ manifold.

2. the *phase* is measured by XX and determines if the state is in the + or - manifold.

To convert from Bell states to classical computational basis states, and extract these two bits of information, we need to eliminate the entanglement which requires two-qubit unitaries U_{AB}. These are not always easy to implement physically - for example, the two qubits could be spatially separated. In this lecture, we will be given an entangled state and want to learn something about it just by measuring single qubit operators. The focus of the following sections will be to understand how we can leverage the correlations between qubits and entangled states of physically separated qubits.

Suppose Alice and Bob are spatially separated and share a singlet state $|\Psi^-\rangle$. Alice measures her qubit in the Z basis with the operator σ_Z^A with measurement operators $M_0^A = |0\rangle\langle 0|_A$ and $M_1^A = |1\rangle\langle 1|_A$. She measures each with 50% occurrence rates, by definition of the Bell states. The value that Alice measures determines Bob's state since they share a Bell pair. In particular, the following Kraus operators act in Bob's subspace: $M_0^B = |0\rangle\langle 0|_B$ and $M_1^B = |1\rangle\langle 1|_B$.

Let us consider Bob's qubit throughout the process thus far. Before Alice performs a measurement, what is the density matrix of Bob's qubit? He has the identity as his reduced density matrix, because they share a Bell pair at first. If Alice then measures her qubit and sends this information to Bob, what state is Bob's qubit in? Bob has a pure state given by Alice's outcome. Conceptually, this is interesting - a measurement in Alice's lab influences Bob's qubit, even though they are spatially separated. Einstein grappled with this concept over 60 years ago when he formulated his Principle of Locality.

2.16.1 Einstein's Principle of Locality & Hidden Variable Theory

.

Einstein postulated that a complete description of physical reality requires that an action on subsystem A must not instantaneously affect the description of B if they are spatially separated. His conclusion was that the quantum mechanical description of the wave function must not be complete: we cannot have a measurement of A affect the description of the state of B. This is also referred to as the *EPR paradox*.

To resolve the paradox, the most general way is to employ a local hidden variable λ. If we consider a description of the Bloch vector in a particular direction $\hat{n}$, and measure it in a direction $\hat{m}$, we know that we obtain the outgoing state $|\hat{m}\rangle$ with probability $p_0 = \cos^2\theta$ and the antiparallel state $|\hat{-m}\rangle$ with probability $1-\cos^2\theta$. The role of λ is to provide the measurement outcome - λ is a variable between 0 and 1, and in this case is equal to p_0 for $\hat{m}$ and $1-p_0$ for $\hat{-m}$. In this description, by only examining a single qubit, we cannot test or gain information about *every* degree of freedom at play, since our measurements is incomplete as λ is hidden, by definition.

2.16.2 Bell inequalities

.

Another resolution to the EPR paradox uses the so-called Bell inequalities. Consider the state $|\Psi^-\rangle$. The measurement outcomes of Alice and Bob will be perfectly anti-correlated, and likewise the phase measurements will be anti-correlated, if we measure in any basis. If Alice and Bob perform measurements in arbitrary bases with respect to each other, will the results be correlated, will they not be correlated at all or will they depend on the basis? The Bell inequalities quantify the answer to that question.

Suppose Alice measures along a direction $\hat{a}$ with projection operators $M_0^A = \frac{1}{2}(\mathbb{1}+\hat{a}\cdot\sigma)$ and $M_1^A = \frac{1}{2}(\mathbb{1}-\hat{a}\cdot\sigma)$, and Bob measures along a direction $\hat{b}$ with projection operators $M_0^B = \frac{1}{2}(\mathbb{1}+\hat{b}\cdot\sigma)$ and $M_1^B = \frac{1}{2}(\mathbb{1}-\hat{b}\cdot\sigma)$. Since, given the state $|\Psi^-\rangle_{AB}$, the results are completely anti-correlated, we can replace σ_B with $-\sigma_A$, such that: $M_0^B = \frac{1}{2}(\mathbb{1}+\hat{b}\cdot\sigma_A)$ and $M_1^B = \frac{1}{2}(\mathbb{1}-\hat{b}\cdot\sigma_A)$. The possible outcomes of the measurement are $00, 01, 10, 11$ for Alice and Bob's measurements, no matter what basis they measure in (these are the eigenvalues of the Pauli matrices). The probabilities of these outcomes according to quantum mechanics is:

$$\begin{aligned} P_{00}(a,b) &= \langle\Psi^-|\,\frac{1}{2}(\mathbb{1}+\hat{a}\cdot\sigma_A)(\mathbb{1}-\hat{b}\cdot\sigma_A)\,|\Psi^-\rangle \\ &= \frac{1}{4}(1+\langle\hat{a}\cdot\sigma_A\rangle - \langle\hat{b}\cdot\sigma_A\rangle - \langle\hat{a}\cdot\sigma_A\hat{b}\cdot\sigma_A\rangle \qquad (116) \\ = \mathrm{Tr}(\rho_A\hat{a}\cdot\sigma_A) - \mathrm{Tr}(\rho_B\hat{a}\cdot\sigma_A) - \frac{1}{4}(1-\cos\theta) &= \frac{1}{4}(1-\cos\theta) = P_{11}(a,b) \end{aligned}$$

Where θ is the angle between $\hat{a}$ and $\hat{b}$. If Alice and Bob measure along the same direction, they always have anti-correlated results, and if they measure in opposite directions, they will obtain correlated results. Additionally, Alice and Bob can make measurements with some small angle between their measurement basis, and still get some degree of correlation between them. Suppose Alice measured in one basis, and Bob measures in another. If Bob sends his result to Alice, she knows what her measurement result *would have been* if she had measured in Bob's basis, because of the perfect correlation between qubits. This might be surprising - to further understand the correlations between measurements, we turn to a classical picture, and compare it to the quantum one described above.

We can quantify a figure of merit S corresponding to measurements along three axis such that:

$$S = P_{same}(e_1^A, -e_2^B) + P_{same}(e_1^A, -e_3^B) + P_{same}(e_2^A, -e_3^B) \geq 1. \tag{117}$$

Classically, if the probabilities are already determined and given, at least one of these terms must be one, since there are only two options of the measurement outcomes, but there are three independent measurements.

To understand the output **quantum mechanically**, let's define $\hat{z} = e_1$, and e_2 and e_3 in a plane, with angle $2\pi/3$ between the vectors. We can examine the probabilities:

$$P_{same}(e_1^A, -e_2^B) = P_{same}(e_1^A, -e_3^B) = P_{same}(e_2^A, -e_3^B) = 1/4 \tag{118}$$

Since $\cos(2\pi/3) = -1/2$. The value of $S = 3/4$ is less than 1 - a violation of the classical result. By using the principles of quantum mechanics to calculate what correlations of the measured values we expect, we were able to get a result which violates the case where measurements are predetermined, and thus we have disproved local hidden variable theory.

This formulation of the Bell in equality is not unique, it is a particular example utilizing the singlet state. There are many different ways to construct the Bell inequalities, and the formulation depends on the specific state that one starts with and the vectors that are measured.

Remarks.

1. Correlations predicted by the quantum theory are incompatible with local hidden variable theory.
2. There is a larger class of Bell inequalities. For example, the $|\Phi^+\rangle$ state should be able to generate non-classical correlations (correlated measurements along *any* axis). The general version is the CHSH inequality. Four directions $a_{1,2}$ and $b_{1,2}$ are chosen, and Alice and Bob calculate the quantity $C = |\langle a_1 b_1\rangle + \langle a_2 b_1\rangle - \langle a_1 b_2\rangle + \langle a_2 b_2\rangle| \leq 2$.

2.16.3 Violation of Bell's inequalities

.

Why are we studying this in a quantum information course? What can we learn from Bell's inequality?

1. Bell's inequalities are an *entanglement witness*. If we measure a violation of Bell's inequality, it means that we must have an entangled state. For two qubits in a pure state, this might be trivially expected, but for a larger system where the degree of entanglement is less clear, the Bell inequalities become useful. In addition, for mixed states of two qubits, we can use Bell's inequalities to check if two qubits are entangled, as well as characterize it.
2. If a density matrix fails the violate Bell's inequality, we cannot draw the conclusion that the state is not entangled, namely, there are entangled states that do not violate Bell's inequalities. For example, the pure state $|\Psi\rangle = \sqrt{1-\epsilon}\,|01\rangle - \sqrt{\epsilon}\,|10\rangle$ does not violate a Bell inequality for small ϵ.
3. Maximal violation of Bell's inequalities occurs with the Bell states; the minimum $S = 3/4$ and the maximum $C = 2\sqrt{2}$ occur with the Bell states. The values of S and C change monotonically with the degree of entanglement, so they are a tool to characterize how entangled a state is which is experimentally feasible.

2.16.4 Loopholes

.

Throughout the literature, there have been various experimental attempts to measure Bell inequalities (with a recently successful measurement, see 2015 Hensen et. al, Nature). In 1985, Apsect performed an experiment with Cs decay. A Cs atom has a ground and excited state, and can decay by emitting a pair of photons, one horizontally and one vertically polarized. The state of the two photons can be written as a bell state, in particular, $|\Psi+\rangle = \frac{|VH\rangle+|HV\rangle}{\sqrt{2}}$. Aspect measured a Bell inequality and found it to violate the classical result. However, the field was not satisfied with this experiment result because there were some important experimental loopholes:

1. *Fair sampling*: in the experiment, many the results are not measured - photons in the experiment get absorbed by something that is not the detector, or are deflected before reaching the detector. Not detecting a significant fraction of photons could allow for a hidden variable to control which photons were eliminated. This was resolved in 2004 using trapped ions with better readout.

2. *Locality* is a more challenging loophole. Alice and Bob must be space-like separated states to show a true violation of Bell inequalities. If Alice's measurement requires a finite amount of time, such that Bob is within Alice's lightcone, then a violation of Bell's inequality would be consistent with Einstein's theory and would not result in a paradox. This was resolved in 2015 in Delft (Hensen et. al, Nature). Clearly, this field is a frontier of modern research – the current technical challenges involve the questions, how far apart can we separate entangled states and how quickly can we measure them?

2.16.5 Applications of Bell's inequalities to quantum information processing

Device-independent quantum communication is sometimes called quantum key distribution (QKD). In this section, we will explore QKD further. Suppose Alice and Bob are spatially separated and share a singlet state $|\Psi^-\rangle$. They perform measurements in randomly chosen basis as described by the CHSH inequality. Alice and Bob then publicly share the basis they performed their measurements in. The values for which the basis was the same constitutes a secret key. However, if the basis was at some angle with respect to one another, they can use those measurements to instead check the CHSH inequality (measure C). If $C \geq 2$, they share a genuinely entangled state, and there must have not been an eavesdropper measuring the key (which would destroy the entanglement). In other words, if one can prove $C \geq 2$, one can verify that the channel is secure. This is a device-independent technique, because it can be applied to any channel, independent of the detector and channel hardware.

2.17 Applications of entanglement

In the previous lecture, we discussed the EPR paradox, the Bell inequalities, and its application to quantum information processing. Now that we have shown that quantum mechanics is indeed complete, as well as touched on some of the applications of entanglement, we will discuss further what entanglement can be used for as a resource. The applications and impact of entanglement include the following:

1. Quantum key distribution is verifiably secure. As a review, in the Ekert protocol, Alice and Bob create an entangled pair, then measure in the z or x basis, and subsequently publicly announce the basis. If the basis is the same, then they generate a bit in their secret key. If the basis is different, they can use those bits to check the Bell inequality and confirm that the channel is secure.

2. In superdense coding, Alice and Bob share a Bell pair, for example $|\Phi^+\rangle_{AB}$. Alice applies four operators to her qubit, $\mathbb{1}, X, Z, XZ$. Depending on the operation, the state will change to a different Bell state. For example, applying Z yields the state $|\Phi^-\rangle_{AB}$. After encoding these four possibilities, she sends her qubit to Bob. If Bob makes a measurement in the Bell basis, he can decode one of these four outcomes, which is two classical bits.

3. In *quantum teleportation*, Bell pairs can be used to communicate quantum bits over a classical channel. Suppose Alice has a state $|\phi_1\rangle = \alpha|0\rangle + \beta|1\rangle$. Suppose she wants to send this qubit to Bob but only over a classical channel. The two parties have several options: option 1 is that Alice could measure her qubit and send Bob the result. Bob could then reconstruct the state to the best of his ability. It's clear that this reconstruction will not be perfect - in fact, the fidelity of reconstruction will be 2/3, as seen on a previous homework problem. However, there is option 2: if Alice and Bob share an entangled pair, she can send this

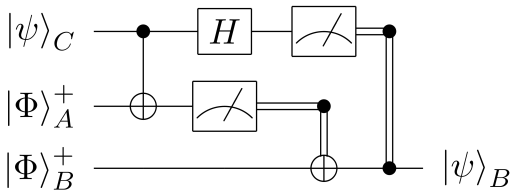


Figure 2: A circuit representation of quantum teleportation. Picture from Wikipedia.

state to Bob perfectly. Suppose that Alice has qubit 2 and Bob holds qubit 3, and they share a Bell pair in the state $|\Psi^-\rangle_{23}$. In that case, we can rewrite the three qubit state as:

$$|\psi\rangle_{123} = |\phi_1\rangle\,|\Psi_{23}^-\rangle = \frac{\alpha}{\sqrt{2}}(|001\rangle - |010\rangle + \frac{\beta}{\sqrt{2}}(|101\rangle - |110\rangle)$$
$$= \frac{1}{2}\Big(\,|\Psi_{12}^-\rangle\,(-\alpha\,|0\rangle_3 - \beta\,|1\rangle_3) + |\Psi_{12}^+\rangle\,(-\alpha\,|0\rangle_3 + \beta\,|1\rangle_3) + |\Phi_{12}^-\rangle\,(\alpha\,|1\rangle_3 + \beta\,|0\rangle_3) + |\Phi_{12}^+\rangle\,(\alpha\,|1\rangle_3 - \beta\,|0\rangle_3)\Big) \tag{119}$$

Rewriting the three-qubit state way suggests a strategy for transmitting the state $|\phi_1\rangle$. Suppose now that Alice performs a Bell basis measurement of her qubits (1 & 2). At this point, she will obtain one of four different outcomes and transmit the classical result to Bob. Then, when Bob learns which state she measures, he can apply a unitary transformation to his qubit to reconstruct the original state. For example, if Alice transmits to Bob that she measures the singlet state, Bob will do nothing. If Alice transmits that she measures $|\Psi^+\rangle$, then Bob needs to apply the Z operator to recover the qubit $|\phi_2\rangle$. To summarize the protocol, Alice makes a Bell basis measurement of her two qubits, sends Bob the result, and Bob applies a unitary operation to reconstruct the initial state.

Remarks about quantum teleportation:

(a) This result is consistent with the no cloning theorem because the state $|\phi_1\rangle$ must be destroyed to send it to Bob.

(b) Quantum teleportation does not violate causality. There is no instantaneous propagation of information, since Alice needs to transmit her measurement result for Bob to properly reconstruct the qubit.

(c) The quantum circuit for teleportation can be represented as shown in figure 2.
The first part of the circuit, which rotates between the product state basis and the Bell state basis, allows for the Bell state measurement between Alice's qubits. That is the CNOT and Hadamard gate. By performing a measurement of qubits $1, 2$, two classical bits of information are extracted and conditional on that result, a single qubit unitary is applied to Bob's qubit.

2.18 Mixed state entanglement

.

When we considered entangled states in quantum teleportation and the violation of Bell inequalities, we discussed the situation where Alice and Bob can share a Bell pair, and we assumed that the two party system was isolated, such that the Bell pair was a pure state. In general, there will be other degrees of freedom, for example in an environment C, which will interact with qubits A and B. In this case, the state of Alice and Bob's qubits is no longer pure and we must describe it with a density matrix.

Specifically, we have a two-qubit density matrix $\rho_{AB} = \sum p_\alpha\,|\psi_\alpha\rangle_{AB}\,\langle\psi_\alpha|_{AB}$. Let us define *uncorrelated states* as the joint density operator $\rho_{AB} = \rho_A \otimes \rho_B$. Secondly, *separable states* are defined as $\rho_{sep} = \sum_k p_k \rho_A^k \otimes \rho_B^k$, such that $\sum_k p_k = 1$. The separable state can be prepared if Alice generates a random number k which is a sample from the distribution $\{p_k\}$, and sends this random number to Bob, whereby Bob performs some unitary rotation that depends on the number. In this case, Alice and Bob have a-priori agreement in that Alice prepares ρ_A^k and Bob prepares his state in ρ_B^k. By using *local operations and a classical communication channel*, Alice and Bob

can prepare such separable states. The state is correlated, but *classically* correlated, or more technically, the state ρ_{sep} can be prepared by *local operations and classical communication* (LOCC). The intuition here is that Bell state have *non local* quantum correlations, so we cannot prepare a Bell state using LOCC. In light of these definitions, we will define *entangled states* as $\rho_{AB} \neq \sum_k p_k(\rho_A^k + \rho_B^k)$.

To consider this definition, we describe some quantities in further detail. A state of of two qubits can be written as:

$$\rho_{AB} = \sum_{i,j=0}^{3} \rho_{ij} \left|i\right\rangle_{AB} \left\langle j\right|_{AB}, \tag{120}$$

whereas, a Pauli representation of a two-qubit entangled state is more complicated but still manageable: $\{A_i\} = \{\mathbb{1}_{AB}, X_A\mathbb{1}_B, Y_A\mathbb{1}_B ...\}$ is a set of 16 operators that define a basis for the two-qubit density matrix. We could express a single qubit in terms of 4 operators, similarly, we can describe a two qubit state in terms of 16 operators. For example, consider the density operator:

$$\rho = f\left|\Phi^+\right\rangle\left\langle\Phi^+\right| + (1-f)\left|\Phi^-\right\rangle\left\langle\Phi^-\right|. \tag{121}$$

This is a statistical mixture of two states, one $|\Phi^+\rangle$ and one $|\Phi^-\rangle$. Is this state entangled for arbitrary f? If so, how much entanglement is there? Suppose that $f = 1/2$. In this case, there is equal probability of $|\Phi^+\rangle$ and $|\Phi^-\rangle$. The density operator will be:

$$\rho = \frac{1}{2}\left|\Phi^+\right\rangle\left\langle\Phi^+\right| + \frac{1}{2}\left|\Phi^-\right\rangle\left\langle\Phi^-\right| = \frac{1}{2}\left|00\right\rangle\left\langle 00\right| + \frac{1}{2}\left|11\right\rangle\left\langle 11\right|. \tag{122}$$

This matrix only has terms on the diagonal and is a statistical mixture of $|00\rangle$ and $|11\rangle$, two separable states, thus, it is not entangled. A signature of entanglement is non-zero off diagonal matrix elements in the 2 qubit density matrix, all of which in this case are zero. More specifically, the part of the density matrix responsible for entanglement are non-zero off diagonal terms in which both qubits change state between the ket and the bra part, for example, $|00\rangle\langle 11|$, and $|01\rangle\langle 10|$. These are akin to the coherences of the single-qubit density matrix. We will call these *multi-qubit coherences*, which can signify the presence of entanglement. (Note however, that just because a density matrix may have nonzero multi-qubit coherences, that does not mean it's entangled - for example, the pure and separable state $|+\rangle|+\rangle$ has nonzero two-qubit coherences). For the example above in equation (121), we see that in the case of $f = \frac{1}{2}$, these terms vanish.

Remarks about ρ_{AB}:

1. We can always define a reduced density operator $\rho_A = \mathrm{Tr}_B \rho_{AB}$, and this reduced density operator can sometimes help us to check if the qubits are entangled. However, if ρ_{AB} were not in a pure state to begin with (as is the case with mixed-state entanglement), this reduced density operator is not helpful in quantifying the entanglement between subsystems A and B. For example, consider the case $\rho_{AB} = \rho_A \otimes \rho_B$ where $\rho_A = \frac{1}{2}\mathbb{1}$ (completely mixed). In this case, even though the reduced density matrix ρ_A is completely mixed, we cannot conclude that it is entangled with subsystem B (and in fact it *is not*)! Only if we can certify that ρ_{AB} was pure to begin with can we use the reduced density operator purity as an entanglement witness.

2. *Positive Partial Transposition (PPT)* Suppose we have two subsystems, A and B with bases $\{|m\rangle\}$ and $\{|\mu\rangle\}$ in $\mathscr{H}_A$ and $\mathscr{H}_B$. We can write the density matrix elements as:

$$\rho_{m\mu,n\nu} = \left\langle m\right| \left\langle \mu\right|\rho\left|n\right\rangle \left|\nu\right\rangle. \tag{123}$$

   Partially transposing this density matrix (exchanging the indices μ and ν), we obtain the matrix $\rho_{AB}^{\intercal_B} = \rho_{m\nu,n\mu}$. For separable states, this corresponds to only transposing the reduced density operator for subsystem B:

$$\rho_{AB}^{\intercal_B} = \sum_p \rho_A^{(1)} \otimes \rho_B^{(1)\intercal}. \tag{124}$$

   Since $\rho_B^{(i)\intercal}$ is non-negative (based on the positivity of the density matrix $\rho_B^{(i)}$), $\rho_{AB}^{\intercal_B}$ will be nonnegative. The PPT criteria says exactly this: if ρ_{AB} is separable, then $\rho_{AB}^{\intercal_B}$ is nonnegative. For 2 qubits, PPT is a necessary and sufficient condition for entanglement. Note that this is not true for higher dimensional systems (e.g. qutrits).

3. *Entanglement witnesses.* An entanglement witness forms a hyperplane in a large Hilbert space, guaranteeing that everything outside this hyperplane is entangled. For an entanglement witness operator W, if ρ is separable, then $\langle W\rangle = \mathrm{Tr}\rho W > 0$. In other words, if we measure $\langle W\rangle < 0$, then ρ is entangled.

   One prominent example is called *entanglement fidelity.* Suppose we attempt to prepare the Bell state $|\psi\rangle = |\Phi^+\rangle$, but instead only prepare some mixed state ρ. We define the fidelity as:

$$F \equiv \langle\psi|\rho|\psi\rangle\,. \tag{125}$$

   In the case where $|\psi\rangle$ is one of the Bell states, showing that $F > \frac{1}{2}$ proves that the state is entangled. We can consider, for example, the state described by (121). We find that in the example of (121):

$$\begin{aligned} F_+ &= \langle\Phi_+|\rho_{AB}|\Phi_+\rangle = f \\ F_- &= \langle\Phi_-|\rho_{AB}|\Phi_-\rangle = 1-f, \end{aligned} \tag{126}$$

   as expected, the state is entangled as long as $f > \frac{1}{2}$ or $(1-f) > \frac{1}{2}$, meaning that it is only not necessarily entangled if $f = \frac{1}{2}$ exactly. However, for example we might consider instead the fidelity associated with $|\Psi^-\rangle$:

$$\langle\Psi_-|\rho|\Psi_-\rangle = 0, \tag{127}$$

   but this does not mean that the state described by (121) is not entangled! We can construct witnesses as described above (in this case the witness would be of the form):

$$W = -\,|\Phi^+\rangle\langle\Phi^+| + \frac{1}{2} \tag{128}$$

   such that $\langle W\rangle < 0$ for $f > \frac{1}{2}$). However, there exist many of these witnesses, and if *any* of them has an expectation value less than 0, then we can conclude our state is entangled. **The key is to choose an entanglement witness appropriate for the mixed state at hand.**

   Another example which we have already seen that uses entanglement witness is the violation of a Bell inequality (one just needs to appropriately modify the inequality to have the form $\langle W\rangle < 0$). We must carefully note that there are states which are entangled, such that there exists a witness that proves entanglement, but do not violate any Bell inequality.

4. The *Werner states* are described by a statistical mixture of a Bell state, with the fully mixed state. For example:

$$\rho = (1-p)\,|\Phi^+\rangle\langle\Phi^+| + p\frac{1}{4}\mathbb{1}. \tag{129}$$

   One might think that naively the portion of the density operator $p\frac{1}{4}\mathbb{1}$ does not contribute at all to an entanglement fidelity with the state $|\Phi^+\rangle$. However, there is some contribution to the fidelity even from the component of the density matrix that is fully mixed, since we can write $\mathbb{1} = |\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-| + |\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-|$. In general, the Werner states are useful since they describe a statistical mixture of getting a particular entangled state some fraction $1-p$ of the time, and some completely random state some other fraction p of the time, which is true in many experiments.

5. Similar to our consideration of the Kraus operator evolution of the single qubit reduced density matrix, we can also examine similar generalized dynamics for subsystem AB:

$$\rho \to \rho' = \sum_\mu M_\mu^{AB}\rho M_\mu^{AB\dagger} + \sum_{\mu'} M_{\mu'}^{A}\rho M_{\mu'}^{A\dagger} + \sum_{\mu''} M_{\mu''}^{B}\rho M_{\mu''}^{B\dagger}. \tag{130}$$

   Here, we have separated the Kraus operator sum into terms that act on subsystem AB jointly, as well as terms that act only on subsystems A and B. These individual terms can describe individual coherent rotations of single qubits, but could also describe how the individual qubits interact with the environment and decohere (independent of the other qubits).

   As an example, if we consider Markovian dephasing of individual qubits, we will see that our master equation will have the form

$$\dot\rho = -i[H,\rho] + \mathcal{L}_A(\rho) + \mathcal{L}_B(\rho), \tag{131}$$

where $\mathcal{L}_i(\rho)$ describes the non-unitary evolution of subsystem i. We expect that the dynamics should be additive in the case where the qubits interact with the environment and dephase independently. This will give rise to density matrix evolution of the form

$$\dot{\rho}_{00,11} = -(\gamma_A + \gamma_B)\rho_{00,11} + ..., \tag{132}$$

such that in the example of an initial Bell state $\propto |00\rangle + |11\rangle$ undergoing dephasing, with single qubit dephasing rates given by γ_A (γ_B), the state undergoes a random flip Z_A OR Z_B, at rate $\gamma_A + \gamma_B$, *either* of which leaves us in the opposite Bell state $\propto |00\rangle - |11\rangle$. As such, the two qubit coherences decay at rate $\gamma_A + \gamma_B$.

As a final note, there are also decoherence processes which act collectively on multiple qubits at once, and these cannot be described by the individual terms acting on just subsystems A or B in (130). These are interesting, because they can in some cases result in entanglement between subsystems A and B. However, they are also potentially a major problem, since all modern theories of quantum error correction assume explicitly that errors are uncorrelated and system evolution can be described in a form such as the example in (132).

2.19 Multipartite entanglement

We will now consider the properties of multipartite entanglement amongst arbitrary numbers of qubits. First, let us consider how we might generate such an entangled state, starting with a pure initial state and using the operations we have already discussed. We remember that we can produce a Bell state with a circuit consisting of a Hadamard rotation followed by a CNOT gate. However, if we add additional qubits and additional CNOT gates, we can produce entangled states of multiple qubits. For two qubits, recall that the Hadamard and CNOT gate produces the state:

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}. \tag{133}$$

It is clear that repeating this process for N qubits we will obtain the N-partite entangled state:

$$\frac{|0\rangle^{\otimes N} + |1\rangle^{\otimes N}}{\sqrt{2}}. \tag{134}$$

2.19.1 GHZ states

This specific type of state in equation (134) is called the $|\text{GHZ}\rangle_N$ state. It is a Schrödinger cat-like state where the system is in a superposition of (sometimes many-qubit) different states $|00..0\rangle$ and $|11..1\rangle$. Some properties of these states include:

1. They are eigenstates of the product operators $Z_1Z_2\mathbb{1}_3...\mathbb{1}_N$, $\mathbb{1}_1Z_2Z_3\mathbb{1}_4...\mathbb{1}_N$, ... $\mathbb{1}_1...Z_{N-1}Z_N$, as well as $X_1X_2...X_N$ with eigenvalue $+1$. These operators which return the state to itself are often called *stabilizers* and are an important concept in quantum error correction.

2. Consider the case of $N = 3$. Using the fact that $ZX = iY$, one can show that the GHZ state is also an eigenstate of operators $Y_1Y_2X_3$, $X_1Y_2Y_3$, $Y_1X_2Y_3$ with eigenvalues -1. If we now consider the generalization of hidden variable theory to three qubits, we will find that $\langle X_1X_2X_3\rangle = -1$, which we know is not the case. In fact, $\rangle X_1X_2X_3\langle = 1$ as noted above. This is the basis for the so-called GHZM paradox, which is yet another violation of local realism and a generalization of Bell's theorem.

3. Partial measurements on $|\text{GHZ}\rangle$ states: consider a measurement of the first qubit in the computational basis. We will always measure a result 0 or 1, which will collapse the full state into either $|0\rangle^{\otimes N}$ or $|1\rangle^{\otimes N}$. Consider now a measurement of the first qubit in the $\{|+_x\rangle\,|-_x\rangle\}$ basis. The amplitude of the GHZ state in $|\pm_x\rangle_1$ is:

$$\frac{(\,\langle 0|_1 \pm \langle 1|_1\,)\,|\text{GHZ}_N\rangle}{\sqrt{2}}, \tag{135}$$

and the GHZ state is projected into the states:

$$\frac{|0\rangle^{\otimes N-1} + |1\rangle^{\otimes N-1}}{\sqrt{2}}$$
$$\frac{|0\rangle^{\otimes N-1} - |1\rangle^{\otimes N-1}}{\sqrt{2}} \tag{136}$$

By choosing the X basis to measure the first qubit, we don't reveal information about whether the remaining qubits are in $|0\rangle$ or $|1\rangle$. This is also a consequence of the fact that the initial GHZ state is an eigenstate of $X_1X_2...X_N$, such that the measurement of X_1 leaves the state in an eigenstate of $X_2...X_N$ with eigenvalue determined by the measurement outcome.

4. We can analyze the effect of decoherence on $|\text{GHZ}\rangle$. If we trace over even just one of the qubits, we obtain the state:
$$\text{Tr}_1 |\text{GHZ}\rangle \langle\text{GHZ}| = \frac{1}{2}(|00..\rangle \langle 00..| + |11..\rangle \langle 11..|). \tag{137}$$
This is a mixed state with no entanglement between qubits - there are no off-diagonal elements of this density matrix. This shows that GHZ states are very sensitive - decoherence of even a single qubit completely destroys the entanglement, which experimentally presents a major challenge in quantum information processing.

5. Quantum metrology, however, takes advantage of this sensitivity to precisely sense the environment. For example, if a magnetic field couples to a qubit through a Hamiltonian $H = hZ$, and we would like to measure the field h, we can use GHZ states for enhanced sensing. To see this, consider a single qubit prepared in the $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$, which will acquire a relative phase $\phi(t) = 2ht$ under precession of the field h. With a single measurement, we can distinguish between relative phases of $\phi = 0$ and $\phi = \pi$ (the two orthogonal states $|\pm_x\rangle$, defining a smallest possible detectable change in ϕ, a *sensitivity* of $\delta\phi \sim 1$. We can repeat this procedure with N independent spins (or repeat this with a single qubit N times), which will give an improvement in our phase sensitivity that scales as $\delta\phi \sim \frac{1}{\sqrt{N}}$.

   However, if we start with the GHZ state, under evolution under the Hamiltonian H it evolves into the state:
$$\frac{|0\rangle^{\otimes N} + e^{i\phi N} |1\rangle^{\otimes N}}{\sqrt{2}}. \tag{138}$$
   To detect the phase change in a single measurement, we get $\delta\phi N \sim \pi$, in other words, $\delta\phi \sim \frac{1}{N}$, corresponding to the two orthogonal states where $N\phi = 0$ and $N\phi = \pi$. This is known as *Heisenberg-limited sensitivity*, which is known to be the best possible scaling of the sensitivity with particle number, and is not possible without entanglement.

6. The fidelity $\mathcal{F} = \langle\Psi_{GHZ}| \rho |\Psi_{GHZ}\rangle$ is an entanglement witness for the GHZ state. The GHZ state is unique in that an entanglement witness can be defined, and it can also be measured - as it only depends on 4 elements of the density matrix (for most general N entangled states, we need to measure 2^N various elements of the density matrix). A fidelity $\mathcal{F} > 1/2$ is sufficient for entanglement.

2.19.2 W states

Another example of a multi-partite entangled state is the so-called *W state*:

$$|\Psi_W\rangle = \frac{1}{\sqrt{N}}(|100...0\rangle + |010...0\rangle + |001...0\rangle + ...) \tag{139}$$

For example for $N = 2$, we recover the Bell state $|\Psi^+\rangle$, which can be converted into the GHZ type Bell state $|\Phi^+\rangle$ with local operations. However, for $N > 2$, $|\Psi_W\rangle$ cannot be converted with local operations into a GHZ type state. Consider the situation where we trace over one of the qubits (e.g. qubit 1). Then:

$$\text{Tr}_1 |\Psi_W\rangle \langle\Psi_W| = \frac{1}{N} |000...0\rangle \langle 000...0| + \frac{N-1}{N} |\Psi_{W,N-1}\rangle \langle\Psi_{W,N-1}| \tag{140}$$

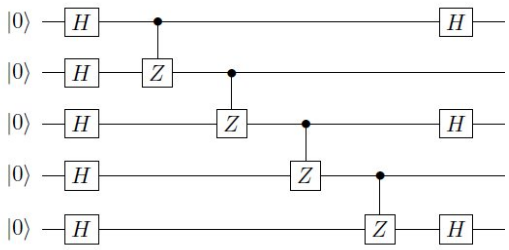


Figure 3: A circuit representation of the Cluster state.

This state is still *mostly* entangled, since the contribution from the first term is small. Unlike the GHZ state, if we trace over a single qubit, there is a significant amount of entanglement preserved, which makes these states more robust to noise. Additionally, the W state can be considered as a collective excitation in the many qubit system.

2.19.3 Cluster states

One last example of a multiparticle entangled state is the cluster state. We can introduce it with a circuit, as shown in figure 3. This state can be written as:

$$|\Psi_C\rangle = \frac{1}{2^{N/2}} \Pi_{a=1}^{N} (|0\rangle_a + Z_{a+1} |1_a\rangle) \tag{141}$$

Where we let $Z_{N+1} = \mathbb{1}$. A powerful way to represent cluster states is with a graph. If each vertex is a qubit prepared in $|+_x\rangle$ and each qubit is connected by an edge which corresponds to a CZ gate, then the qubits comprise a chain. One useful property of the cluster state is that it is an eigenstate of the operators $S_1 = X_1 Z_2$, $S_j = Z_{j-1} X_j Z_{j+1}$, ... $S_N = Z_{N-1} X_N$. The eigenvalues are all +1, such that these operators are *stabilizers* for the cluster state.

2.19.4 Remarks about multipartitle entanglement

1. Graph states can be generalized: for example, a central qubit connected to other qubits is equivalent to a GHZ state up to local operations. Graph theory can be used to see which states are entangled.

2. A 2D cluster states can be a very powerful resource. It allows for universal quantum computation just by measurements. These 2D cluster states contain enough entanglement to perform any quantum manipulation possible, provided that the state is large enough for the computation at hand. In fact, quantum supremacy and quantum computers can be defined in terms of their ability to create a 2D cluster state. Proofs of quantum supremacy are often based on a reduction to creating a 2D cluster state.

3. Degree of entanglement

   Another way to quantify entanglement in a many-body system is to divide the qubits into two subsections and trace over one of the subsections. For example, tracing over half of the qubits (system B) in the GHZ state gives the density matrix:

   $$\rho_A = \mathrm{Tr}_B |\Psi_{GHZ}\rangle \langle\Psi_{GHZ}| = |0..0\rangle_{N/2} \langle 0..0| + |1..1\rangle_{N/2} \langle 1..1| \tag{142}$$

   Calculating the entanglement entropy, we find $S(\rho_A) = 1$, which is small compared to the maximal entropy for N qubits of $N/2$. Despite the fact that the GHZ state is undeniably entangled, the degree of entanglement in terms of the entropy is actually quite small. Similar calculations for W and cluster states yield similar results - entanglement entropy on the order of 1 or a constant, as opposed to linear in N. Situations in which the entanglement entropy grows with the number of particles is sometimes called the *volume law* of the entanglement entropy: the entanglement grows as a function of the system size. Systems in which the entanglement entropy is constant is described as an *area law* entanglement. Intuitively, for area law entanglement, there exists a boundary such that in making a cut in the graph (tracing over the degrees of

freedom on one side), one breaks a single entangled pair bond in the graph, as opposed to order N entangled pairs. Note that making a cut for the 2D cluster state gives an entanglement entropy of $\sqrt{N}$. As expected, if subsystem A and B are all to all entangled, then there will be $\sim N$ bonds broken when tracing over half the qubits. In fact, because of the exponential scaling of the Hilbert space with the number of qubits, most states in the Hilbert space obey volume law entanglement.

4. In general, describing and quantifying entanglement is hard. This is the goal of quantum simulation, and quantum state tomography. These goals require many calculations and measurements, proportional to an exponential of the number of qubits N.

5. Do we need to study the zoo of entangled states that are volume law entangled? To describe the most general states, we require 2^N complex amplitudes. However, most desired states that we encounter in quantum information processing are separable, which have only $\sim N$ complex numbers. This question motivates the search for special classes of entangled states which is broader than separable states, but does not include everything, called *physically accessible states*. Is there a generic description for states like this? This subject is an area of current research and is relevant for quantum supremacy and quantum simulation. The class of states are sometimes called Multiparticle Entangled Renormalization Ansatz (MERA) states, and are created with a circuit. Another approach to this is called tensor networks, associated with the class of tensor network states.

2.19.5 Tensor network states

Tensor networks are a tool to represent a subset of many-body entangled states in an insightful way. Specifically, to start, represent a qubit as a pair of d-dimensional systems - extend the qubits' Hilbert space to d dimensions, then entangle these extra degrees of freedom in so-called bonds. Each bond, at site e, represents a maximally entangled state $|I_{e,e+1}\rangle \sim \sum_{i=1}^{d} |\alpha_e\rangle |\alpha_{e+1}\rangle$.

A 1D tensor network state is called a Matrix Product State (MPS). An MPS is a map on each cite which takes the pairs of the auxiliary indices and maps them into qubit states:

$$P_e = \sum_{i=1}^{2} \sum_{\alpha,\beta=1}^{D} A^e_{i,\alpha,\beta} |i_e\rangle \langle \alpha, \beta_e| , \tag{143}$$

such that the state is:

$$|\Psi_{MPS}\rangle = \Pi_{e=1}^{N} P_e |I_{12}\rangle |I_{23}\rangle ... |I_{N,1}\rangle = \sum_{i_1,...i_N}^{2} \mathrm{Tr} A^1_{i_1} ... A^N_{i_N} |i_1 ... i_N\rangle . \tag{144}$$

The state is specified by N dxd dimensional matrices $A^{\{s\}}_{ij}$. Suppose we introduce a cut and calculate the entanglement entropy. In this MPS case, specifically, we only cut one bond, so the maximal degree of entropy created is proportional to the dimension of the bond. This allows for a construction of a class of entangled state with a specific degree of entanglement, and allows for systematically keeping track of the degree of entanglement introduced. For example, GHZ states, W states, cluster states all have bond dimension equal to 2. These MPS states are useful both for analytics and numerics of studying many body entangled quantum systems.

3 Quantum algorithms

We have discussed some applications of quantum states in cryptography, metrology and communication. It is interesting to ask if these ideas can accelerate solving computational problems. In classical computers, the information is represented in a binary encoding which can represent numbers from 0 to 2^{n-1}, where n is the number of qubits. All of the operations of classical computers can be composed as a set of universal operations or universal gates. One set is NOT, AND, OR, and COPY. Conversely, in quantum computation, the dynamics are unitary and therefore reversible, and in general they have probabilistic outcomes. We will find that these properties can be leveraged to construct specific algorithms that have an advantage over classical computers, and we will work to quantify the advantage.

3.1 Simple quantum algorithms

We have managed to cover a wide foundational base of material. Coronavirus, SARS, communism, capitalism, republicans, democrats will come and go, but this understanding will stay. In this section, we will discuss some canonical examples of simple quantum algorithms.

3.1.1 Quantum parallelism example

Let's design a quantum unitary which implements the function:

$$(x_1, ..., x_n) \to (x_1, ...x_n, f(x)). \tag{145}$$

First, we define the unitary U_f on n input qubits to be:

$$U_f \left|x\right\rangle_n \left|y\right\rangle_m = \left|x\right\rangle_n \left|y \oplus f(x)\right\rangle_m, \tag{146}$$

where $\oplus$ is modulo 2 bitwise addition. Note that $U_f \left|x\right\rangle_n \left|0\right\rangle = \left|x\right\rangle_n \left|f(x)\right\rangle$, that U_f is unitary, and $U_f^{-1} = U_f$. By preparing a superposition of all the input states, and applying the unitary, we will evolve all of the superpositions. Also note that $H^{\otimes n} \left|0\right\rangle_n$ prepares the combination of all possible bitstrings, which we can express as $H^{\otimes n} \left|0\right\rangle_n = \frac{1}{2^{n/2}} \sum_{0<x<2^n} \left|x\right\rangle_n$. Exploiting the idea of quantum parallelism, we apply the unitary to this superposition of states:

$$U_f H^{\otimes n} \left|0\right\rangle_n \left|0\right\rangle_m = \frac{1}{2^{n/2}} \sum_n \left|x\right\rangle_n \left|f(x_n)\right\rangle_m \tag{147}$$

Although we have been able to evolve all of the inputs at once, in part showing how quantum computers can be powerful, the measurement and its outcome is probabilistic. If we measure the output, we always find just one value at random. In practice, there is no way to find all $f(x_n)$ without 2^n repetitions. **To overcome this, we would like to use additional operations to find some *relationship* between the $f(x_n)$ without actually revealing the values of $f(x_n)$.** The art of quantum algorithms is to use additional operations to find relationships between different $f(x_n)$ without revealing their values. Let us explore some examples.

3.1.2 Deutsch's problem

If we have a function $f : \{0, 1\} \to \{0, 1\}$ and want to find if the function is constant or not, we can use a quantum computer to reveal the answer in a single measurement. Classically, there is no choice except to evaluate the function for different x and determine if the answer is constant. Quantum mechanically, we can evaluate this question in a single function evaluation. The quantum algorithm evaluated on state $\left|x\right\rangle$ is as follows:

$$U_f \left|x\right\rangle_1 (\left|0\right\rangle_2 - \left|1\right\rangle_2) = \left|x\right\rangle_1 (\left|f(x)\right\rangle_2 - \left|1 \oplus f(x)\right\rangle_2) = \left|x\right\rangle_1 (-1)^{f(x)} (\left|0\right\rangle_2 - \left|1\right\rangle_2) \tag{148}$$

Now, preparing the input state in a superposition, we find:

$$U_f(\left|0\right\rangle_1 + \left|1\right\rangle_1)(\left|0\right\rangle_2 - \left|1\right\rangle_2) = ((-1)^{f(0)} \left|0\right\rangle_1 + (-1)^{f(1)} \left|1\right\rangle_1)(\left|0\right\rangle_2 - \left|1\right\rangle_2) \tag{149}$$

By subsequently measuring the first qubit in the X basis, we can evaluate if $f(0) = f(1)$ or not, namely, if the function f is constant.

Now we will consider the corresponding quantum circuits, which is generally a convenient and intuitive way to think about quantum algorithms. There are only 4 possible functions that are 1 to 1 that can be created:

1. $x \to 0$
2. $x \to x$
3. $x \to x \oplus 1$
4. $x \to 1$

The circuit representation that implements all of those functions are (see figure 4):

1. Do nothing on both qubits
2. Control X with target on second qubit

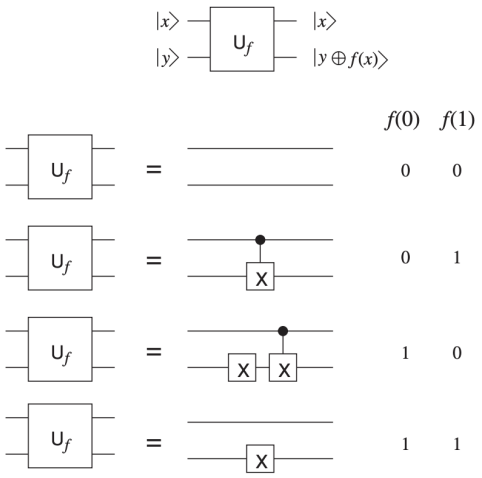


Figure 4: A circuit representation the possible functions in Deutsch's problem. Figure from Mermin.

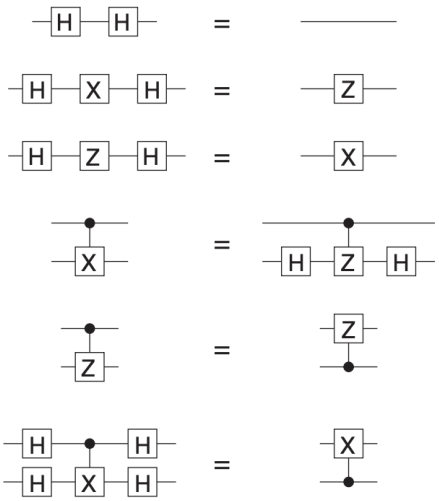


Figure 5: Various circuit simplifications. Figure from Mermin.

3. X on second qubit and control X with target as second qubit
4. X on second qubit

Inserting Hadamard gates before and after the unitary on both qubits, we can exploit quantum parallelism (figure 6). We start with the first qubit in $|0\rangle$ and the second qubit in $|1\rangle$. Various circuit simplifications can be made (see figure 5) to confirm that when f is constant, the output is 0, and if it is balanced the output is 1.

Now we can generalize this procedure to a slightly more complex case.

3.1.3 Deutsch-Jozsa algorithm

Suppose we have a function $f(x_n)$, of n bits. We are told that the function is drawn from the sample of two possible functions: $f(x_n) = c$ constant for all x_n, and balanced, which means $f(x_n) = 0$ for 50% of the cases, and 1 for the other 50% of cases. In one function evaluation, we can find out whether the function is constant or balanced. We again evaluate the function on a superposition of states:

$$U_f H^{\otimes n} |0\rangle_n (|0\rangle - |1\rangle)/\sqrt{2} = \frac{1}{2^{n/2}} \sum_x (-1)^{f(x)} |x\rangle (|0\rangle - |1\rangle)/\sqrt{2} \tag{150}$$

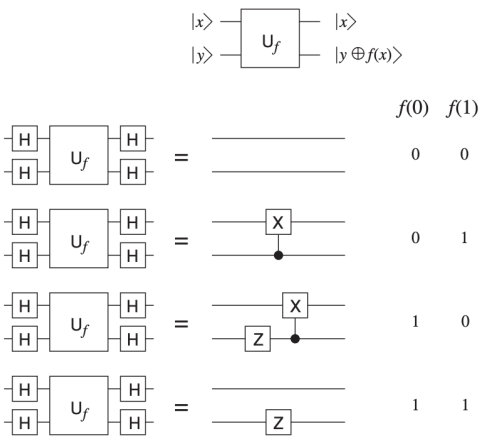


Figure 6: A circuit representation the possible functions in Deutsch's problem, including the Hadamard gates and corresponding circuit simplifications. Figure from Mermin.

We again apply Hadamard gates on each of the qubits:

$$H^{\otimes n}U_f H^{\otimes n}\left|0\right\rangle_n(\left|0\right\rangle-\left|1\right\rangle)/\sqrt{2}=H^{\otimes n}\frac{1}{2^{n/2}}\sum_x(-1)^{f(x)}\left|x\right\rangle(\left|0\right\rangle-\left|1\right\rangle)/\sqrt{2}. \tag{151}$$

Suppose $f(x)$ is constant. Then $(-1)^{f(x)}$ is a constant factor we can take out of the sum, and the H applied to all the qubits would return the zero qubit back on all the n registers, such that $\left|0\right\rangle_n \to \left|0\right\rangle_n$. For a constant function, with probability 1 we will get the same qubits returned. Similarly, when the function is balanced, the probability to get $\left|0\right\rangle_n$ is zero.

A mathematical remark: The Hadamard gate on a single qubit can be written as $H\left|x\right\rangle_1=\left|0\right\rangle+(-1)^x\left|1\right\rangle=\sum_{y=0}^1(-1)^{x_1\cdot y_1}\left|y_1\right\rangle$. Evaluating the Hadamard on many qubits, we have:

$$H^{\otimes n}\left|x\right\rangle_n=\frac{1}{2^{n/2}}\sum_{y_{n-1}=0}^{1}...\sum_{y_0=0}^{1}(-1)^{\sum_{j=0}^{n-1}x_jy_j}\left|y_{n-1}...y_0\right\rangle=\frac{1}{2^{n/2}}\sum_{y=0}^{2^n-1}(-1)^{x\cdot y}\left|y\right\rangle_n. \tag{152}$$

Here $x\cdot y=x_0y_0\oplus x_1y_1\oplus...$ is a sum of products of corresponding bits, modulo 2. The state in equation (151) can also be written as:

$$\left|\psi\right\rangle=\frac{1}{2^n}\sum_y\sum_x(-1)^{x\cdot y+f(x)}\left|y\right\rangle(\left|0\right\rangle-\left|1\right\rangle)/\sqrt{2}. \tag{153}$$

Again if $f(x)=c$, then the state is:

$$\left|\psi\right\rangle=\frac{1}{2^n}(-1)^c\Pi_{j=1}^n(\sum_{x_j=0}^{x}(-1)^{y_jx_j})\left|y\right\rangle(\left|0\right\rangle-\left|1\right\rangle)/\sqrt{2}=\delta_{y,0}\left|y\right\rangle(\left|0\right\rangle-\left|1\right\rangle)/\sqrt{2}, \tag{154}$$

such that the amplitude of the $y=0$ term will be 1, as expected. If $f(x)$ is balanced, we have $\sum_x(-1)^{f(x)}=0$ so the amplitude of the $\left|y\right\rangle=\left|0\right\rangle$ term is zero.

How does this compare to classical algorithms? Worst case, for a guaranteed answer, we would have to evaluate half of the values, plus one queries, to make sure that it is not balanced. Quantum mechanically, as shown, we can use a single function evaluation.

3.1.4 Bertstein-Vazirani problem

Another example of a quantum algorithm is the following. Consider the function $f(x)=a\cdot x$, with the goal to find a. Classically, if we want to isolate the mth bit of a, we need to multiply a to the number 2^m assuming a is written in binary, where here $2^m=(0,0,...1_m,...0)$. If we want to learn n bits, we can apply f to n values of

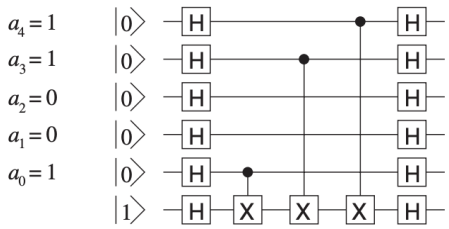


Figure 7: A circuit representation of the Bernstein Vazirani problem. Figure from Mermin.

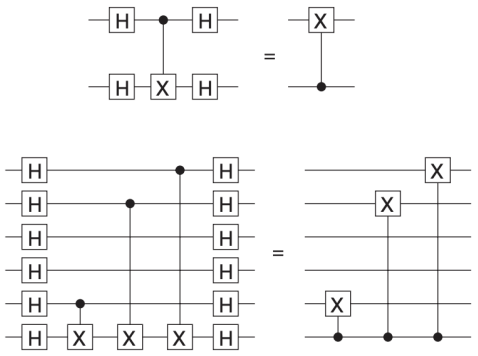


Figure 8: A circuit simplification for the Bernstein Vazirani problem. After this simplification, the result becomes more clear: the state of the input register is changed to a. Figure from Mermin.

$x = 2^m$ where $0 \leq m \leq n$. Quantum mechanically, we can do it in one step. The reason is similar to what we have discussed previously. Again using the same concepts and circuit, we end with the state in eq. (153), where here $f(x) = a \cdot x$. Performing the sum over x:

$$\sum_x (-1)^{x \cdot y + a \cdot x} = \Pi_{j=1}^n \sum_{x_j=0}^{1} (-1)^{(y_j + a_i) x_i} = \delta_{y,a} \tag{155}$$

We will receive all digits of a in a single measurement, since all bits will destructively interfere unless $y = a$. We can evaluate this in a circuit, as shown in figures 7 and 8. We can again think of this function $f(x)$ as an *oracle*.

3.1.5 Summary: simple quantum algorithms

We have now discussed several examples of toy quantum algorithms that illustrate how, in principle, quantum information can speed up computations.

- If we consider a function that takes some binary string $x_1 ... x_n \to (x_1 ... x_n, f(x))$, we defined a so-called *oracle* U_f which acts in the following way:

$$U_f \left|x\right\rangle_n \left|y\right\rangle_m = \left|x\right\rangle_n \left|y \oplus f(x)\right\rangle_m , \tag{156}$$

  where $\oplus$ denotes addition modulo 2.

- We constructed a superposition of all input states by applying

$$H^{\otimes n} \left|0\right\rangle_1 ... \left|0\right\rangle_n = \frac{1}{2^{n/2}} \sum_{x<2^n} \left|x\right\rangle_n . \tag{157}$$

- Just applying this transformation and measuring does not give rise to any sort of speedup, since in each measurement we only get the functional value for one of the inputs. Instead, we considered using the Oracle to get a final state $|\psi_f\rangle = H^{\otimes n} U_f$, which gave us the speedup in determining whether the function is constant or balanced.

- By speedup, we mean that the Deutsch algorithm requires 2 function evaluations classically, and only 1 quantum mechanically. On the other hand the Deutsch-Jozsa algorithm (for an n-bit function), we need only 1 quantum step, as opposed to $\sim 2^n$ (actually, polynomial in n to be more precise/generous to classical algorithms). In the Bernstein-Vazirani algorithm, as opposed to n steps, we only needed 1.

3.2 Simon's algorithm

Suppose we have a two-to-one function of n bits $f(x_1...x_n)$ where one value of this function corresponds to two input bit strings. This function reduces $n \to n-1$ bits. We consider the case that $f(x) = f(y)$ if and only if $x = y \oplus a$, or $x \oplus y = a$, such that $f(x \oplus a) = y$. This function is periodic with period a, and the goal of Simon's problem is to find this period.

Let us consider how many function evaluations it would take to solve this problem. In the classical case, we would need to feed the values x^1, x^2, x^3 (where here these are ordinary numbers, not binary digits), and find i, j such that $f(x^i) = f(x^j)$ and $a = x^i \oplus x^j$. Suppose we evaluate m different values of x, and we did not find this equality. Then we know that $a \neq x^i \oplus x^j$ for the m values evaluated. By evaluating m inputs, we tested at most $\frac{1}{2}m(m-1)$ values of a. In total, there are 2^n possible values of a. This means we need $m \sim 2^{n/2}$ calls of the classical function in order to find a on average! This shows (roughly) that the classical complexity of this problem is exponential.

Now, let us consider the quantum approach to solving this problem. As usual, we consider application of the oracle to a superposition of all inputs, yielding:

$$U_f \frac{1}{2^{n/2}} \sum_{x=0} |x\rangle |0\rangle = \frac{1}{2^{n/2}} \sum |x\rangle |f(x)\rangle . \tag{158}$$

Next, we will measure the function register, which gives an output to 1 out of 2^{n-1} possible values. Recall that there are two values of x that return the same $f(x)$. This measurement will therefore collapse the data register to the state which is a superposition of those two values

$$\frac{1}{\sqrt{2}}(|x_0\rangle + |x_0 \oplus a\rangle) \tag{159}$$

where x_0 is the value for which $f(x_0)$ agrees with the measurement result for f.

This seems like an incredible development, since we immediately create a superposition of two states separated by the period of interest. In practice, however, this is of limited utility. Of course, we can measure the data register and we will get either x_0 or $x_0 \oplus a$ at random. If we repeat this procedure, remember that we will get a *different* random x_0, so we will not actually get any speedup (we get exponentially unlucky, it turns out).

The solution, again, is not to measure this at random. We instead apply the Hadamard transform to the entire data register, to obtain:

$$\begin{aligned} |\psi_{\text{out}}\rangle &= \frac{1}{\sqrt{2}} H^{\otimes n}(|x_0\rangle + |x_0 \oplus a\rangle) \\ &= \frac{1}{2^{(n+1)/2}} \sum ((-1)^{x_0 \cdot y} + (-1)^{(x_0 \oplus a)\cdot y}) |y\rangle \end{aligned} \tag{160}$$

Physically, this is a transformation back to the original basis, with phase factors now given by the function evaluations. Note that we can rewrite:

$$(-1)^{(x_0 \oplus a).y} = (-1)^{x_0 \cdot y} (-1)^{a \cdot y}. \tag{161}$$

If $a \cdot y = 1$, then the coefficient in front of the corresponding state $|y\rangle$ goes to 0. Therefore, we are left with:

$$|\psi_{\text{out}}\rangle = \frac{1}{2^{(n-1)/2}} \sum_{a \cdot y = 0} (-1)^{x_0 \cdot y} |y\rangle . \tag{162}$$

Now, once we measure *any* $|y\rangle$, suppose we measure $|y^j\rangle$, we immediately know that $a \cdot y^j = 0$ for the y^j value received, and we can make the statement that:

$$\sum_{i=0}^{n-1} y_i^j a_i = 0 \mod 2. \tag{163}$$

Simon's algorithm then repeats this process $O(n)$ times, such that with high probability, we get n linear independent equations for bits a_i, which we can use to reconstruct the number $a \equiv a_0 a_1 ... a_n$. A few remarks:

1. One can show that with $n + x$ repetitions, the success probability becomes $p_s > 1 - \frac{1}{2^{x+1}}$ (see Mermin appendix for the derivation).

2. We have shown a quantum algorithm which accomplishes a certain task in $O(n)$ whereas the best known classical algorithm required $2^{n/2}$, getting a a true exponential speedup. Of course, this is again a bit of a contrived problem since it requires knowing the oracle (which in turn, requires knowledge of a in the first place).

3.3 Quantum search algorithm

We will now consider *Grover's search algorithm*, another oracle-based algorithm, but one that can now potentially be made useful. Suppose we have a large, unsorted database with N items. Our goal is to search for one specific item, labeled w.

To formulate this in terms of the oracle, we can consider a function $f(x)$ where $x \in \{0, ..., N-1\}$, where $f(x = w) = 1$ and $f(x \neq w) = 0$. Here, the problem is to find w. Classically, we must apply $f(x)$ to all values of x until it returns 1. On average, this requires $N/2$ function evaluations to find the solution ($p_s = 1/2$). In the worst case, if we are particularly unlucky, we will need to check $N - 1$ elements before success ($p_s = 1$).

In the quantum approach, we again define the quantum oracle:

$$U_f |x\rangle |y\rangle \to |x\rangle |y \oplus f(x)\rangle \tag{164}$$

where $|x\rangle$ is the data register, and $|y\rangle$ is the function evaluation register. In this case, we start with the function register in the superposition state $|y\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. The quantum oracle yields (for this particular y):

$$U_f |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}} = (-1)^{f(x)} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{165}$$

Since the oracle applies identity to the function register, we have that:

$$U_f |x\rangle = (-1)^{f(x)} |x\rangle . \tag{166}$$

The oracle changes the phase for $x = w$, so we must have the oracle:

$$U_f = \mathbb{1} - 2 |w\rangle \langle w| . \tag{167}$$

Even though we can express the oracle in this way, we don't necessarily know this w. Mathematically, U_f is a huge matrix, and if we think of the oracle as a black box, it is nontrivial to find this w.

However, in the following we will show that by using the oracle, we will only need $\sqrt{N}$ evaluations of the function to find w with probability close to 1. This is the essence of Grover's search algorithm. In summary, we do not know the value of w, a vector in $\mathscr{H}$, and the goal is to find it. The steps, in summary, of the Grover Search Algorithm are:

1. Prepare the superposition state $|s\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle$

2. Construct a unitary $\hat{U}_s = 2 |s\rangle \langle s| - \mathbb{1}$. This corresponds to a reflection of any vector in $\mathscr{H}$ perpendicular to $|s\rangle$, about the vector $|s\rangle$ (the sign will be flipped with respect to the direction $|s\rangle$).

3. Apply the Grover iteration:

$$R = U_s U_f \tag{168}$$

Where the matrix R is given by:

$$R^m |s\rangle \approx |w\rangle , \tag{169}$$

and where $m \sim \sqrt{N}$. The relationship between $|s\rangle$ and $|w\rangle$ is simply that they have nonzero overlap $1/\sqrt{N}$. The state $|s\rangle$ does not favor in any way the state $|w\rangle$ - it is simply required to have some small nonzero overlap. Let's consider the Grover iteration acting on state $|w\rangle$:

$$U_s U_f |w\rangle = (\mathbb{1} - 2|s\rangle\langle s|)(\mathbb{1} - 2|w\rangle\langle w|)|w\rangle = |w\rangle - 2|s\rangle\langle s|w\rangle = 2|s\rangle\langle s|)(\mathbb{1} - 2|w\rangle\langle w|)|w\rangle = |w\rangle - 2|s\rangle/\sqrt{N}. \tag{170}$$

To simplify the expression, we use another vector $|r\rangle \equiv \frac{1}{\sqrt{N-1}} \sum_{x \neq w} |x\rangle = \sqrt{\frac{N}{N-1}}|s\rangle - \frac{1}{\sqrt{N-1}}|w\rangle$, such that:

$$U_s U_f |w\rangle = |w\rangle - \frac{2}{N}|w\rangle - \frac{2\sqrt{N-1}}{N}|r\rangle = \cos\theta |w\rangle - \sin\theta |r\rangle . \tag{171}$$

Moreover, we can express θ in terms of N: we have $\theta = 1 - \frac{2}{N}$ or $\theta \sim 2/\sqrt{N}$ for large N, and $U_s U_f |w\rangle = \sin\theta |w\rangle + \cos\theta |r\rangle$.

We note that $U_s U_f$ **operates within the** $\{|w\rangle, |r\rangle\}$ **subspace only**:

$$U_s U_f \begin{pmatrix} |w\rangle \\ |r\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} |w\rangle \\ |r\rangle \end{pmatrix} . \tag{172}$$

We can interpret the entire evolution geometrically. The vector $|s\rangle$ is an angle in the $|r\rangle$, $|w\rangle$ plane, by some particular angle $\theta/2$ from axis $|r\rangle$ to $|w\rangle$. Then, U_f rotates $|s\rangle$ counterclockwise by angle θ. Lastly, U_s flips this vector about the original vector $|s\rangle$, such that $U_s U_f |s\rangle$ is a vector in the $|r\rangle$, $|w\rangle$ plane, by angle θ from $|r\rangle$ to $|w\rangle$. Therefore, the vector $|s\rangle$ has been rotated in total by angle $\theta/2$ towards $|w\rangle$ in this 2D space.

Applying the Grover iteration k times:

$$(U_s U_f)^k = \begin{pmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{pmatrix} \tag{173}$$

For an initial state $|s\rangle \sim |r\rangle$, then after k rotations, such that $k\theta = \pi/2$, then $(U_s U_f)^k |s\rangle \approx |w\rangle$. *The number of steps will be* $k = \frac{\pi}{2}\frac{1}{\theta} = \frac{\pi}{4}\sqrt{N}$*, such that the Grover algorithm requires* $\sim \sqrt{N}$ *iterations to find the element* $|w\rangle$. Recall that the classical result requires N, so there is a polynomial speedup achieved with Grover's search algorithm using a quantum computer.

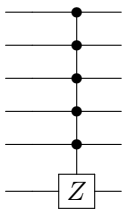
Remarks:

1. With the geometric expression of the algorithm, it is clear that the root of the speed up is from interference. The interference in the N-dimensional $\mathscr{H}$ space is between *amplitudes* (e.g, $\langle w|s\rangle$), not intensities (e.g, $|\langle w|s\rangle|^2$): the amplitude of the overlap of the two states $|r\rangle$ and $|w\rangle$ is changed as the vector rotates with each Grover iteration. When the amplitudes add, the intensity squares, hence the quadratic speedup. Additionally, because of the algorithm's reliance on intererence, we can view this quantum computation as an interferometer - we do not even need to consider qubits for this process! Considering classical waves, or even many states of a single atom can both exhibit the same interference that is required to execute the search properly. In fact, when this algorithm was first introduced, there were multiple experimental groups taking this approach. However, one needs to be extremely careful about scaling of other resources: the cost of other resources that are needed, such as the measurement, can negate the speedup. For example, using optics (and the interference of light waves), N beamsplitters are needed, and when N is large, the relative quadratic speedup is a moot point.

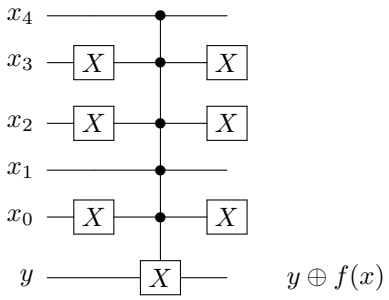
2. However, the search algorithm can be implemented efficiently using qubits and using a quantum circuit. Specifically, if we have entries which iterate from $x \in \{0, ..., N-1\}$, we can encode them using $n = \log_2 N$ qubits. The steps are as follows:

   - Prepare initial state $|s\rangle = H^{\otimes n} |0\rangle ... |0\rangle$

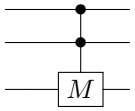
- We can express $U_s = 2\left|s\right\rangle\left\langle s\right| - \mathbb{1}$ as $H^{\otimes n}(2\left|00...0\right\rangle\left\langle 00...0\right| - \mathbb{1})H^{\otimes n}$. The quantity $(2\left|00...0\right\rangle\left\langle 00...0\right| - \mathbb{1})$ provides a phase of -1 for all states, except for $|00...0\rangle$, for which the phase is unity. To implement this term, we define an n qubit controlled phase rotation C^nZ as:



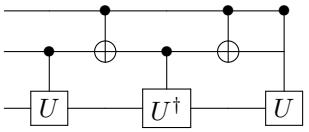
Adding X gates on any qubit allows for conditioning for a phase flip if the qubit is in $|0\rangle$. For example, if the oracle uses the state $|w\rangle = |10010\rangle$, then the circuit would be as follows:



3. How do we implement an n qubit controlled gate using only two-qubit gates? For $n = 2$, we have already CZ and CNOT gates, etc. Now, consider $n = 3$. A controlled unitary squared gate (where $M = U^2$) is:

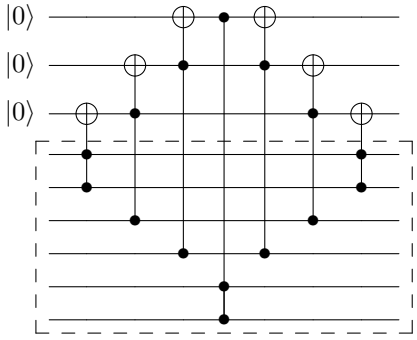


This gate can be implemented as a sequence of two-qubit operations as follows:



(one can see this by iterating through the different states of the inputs of the top two qubits, and confirming that U^2 is applied only if these are in state $|11\rangle$). If we want to implement for example C^2Z, then we construct the unitary such that $U^2 = Z$, namely a $\pi/2$ phase, and apply the conditional operations as given in the figure above. Another gate that we have already seen in the homework of this type is the Toffoli gate. Converting many-qubit controlled gates to two-qubit gates is polynomial overhead. This is generally not a problem, but for near-term quantum devices, it can be quite challenging. The idea of *co-design* is to develop hardware informed by the particular algorithm that one wants to execute.

4. Suppose we would like to implement the gate C^nZ where n is the number of the control qubits and Z is the phase flip. To do this using only two-qubit or three-qubit gates, we can introduce $n - 2$ *ancilla* helper qubits. See, for example, the case of $n = 5$, where three ancillary qubits are used to implement a C^5Z gate on the 6 qubit register below:



In general, this can be performed with $2(n-2)C^2X$ gates, or $\sim 10n$ CNOT gates. Using this qubit implementation, Grover's search can be implemented with $\sim C\sqrt{N}\log N$ steps, where here $\log N$ is the cost of implementing multi-qubit gates and allocating ancillary qubits (recall that the number of qubits is $n = \log N$).

5. Grover's algorithm is optimal: the $\sqrt{N}$ speedup is the best possible speedup. (See Preskill's notes for the proof).

6. Grover's search algorithm is probably not *actually* useful for a database search, because it requires a special kind of database where the oracle is already known, and provided to us. However, the algorithm is important because it *solves a problem without any internal structure.* All of the previous problems had internal structure (e.g. periodicity) which we could reveal by operating on the qubits and performing collective measurements. The power of these quantum algorithms is extended by Grover's search because it is applied to a problem where there is no particular requisite structure. For this reason, Grover's search can be applied to many different kind of problems: many complex optimization problems (including some NP complete ones) can be cast as a search problem, with $\sim \sqrt{N}$ speedup.

3.4 Quantum Fourier Transform

The Quantum Fourier Transform (QFT, not to be confused with quantum field theory) is an example of an algorithm that is immediately useful, unlike the previous examples. The QFT is similar to Simon's algorithm, in that it finds a if $f(x \oplus a) = f(x)$. It is motivated by the idea that the problem of period finding has some structure, and we can use quantum interference to reveal that structure. Recall that the key operation of Simon's algorithm was performing Hadamard gates on all qubits:

$$H^{\otimes n} |x\rangle = \sum_{y=0}^{N} (-1)^{x \cdot y} |y\rangle \tag{174}$$

where we used multiplication modulo two bitwise, $x \cdot y$.

The question we would like to study now is **can quantum computers find a period in conventional algebra**? Suppose for example we have a function that is periodic:

$$f(x) = f(x + mr), \tag{175}$$

where m is an arbitrary integer. The goal is to find the period r for such a function. For intuition on interference, it is helpful to turn to optics. Consider impinging a beam of light on a surface. If we have a rough surface and you send a beam of light towards it, the "grating" modulates the phase of the reflected light,giving it an effective momentum kick, leading to a reflection at a different angle. The light will constructively interfere only at specific points, depending on the periodicity of the grooves of grating, which determines the angle corresponding to constructive interference precisely. This is similar to the kind of interference that allows us to find the period of a function.

3.4.1 Discrete Fourier transform

First, we review the discrete Fourier transform. Suppose we have a certain function of integers $f(x)$. The discrete Fourier transform is

$$f(x) \to g(y) = \sum_{x=0}^{N-1} e^{2\pi i x y / N} f(x). \tag{176}$$

Here, the function $f(x)$ is defined on the interval $[0, ..., N-1]$, such that x and y are integers. Note that in this case xy is a *conventional* product. If the $f(x)$ is a 1-1 function, then the discrete Fourier transform is an NxN matrix that takes $f(x) \to g(y)$. Classically, calculating the Fourier transform requires N^2 operations (one for each element of the matrix). This can be sped up using the 'Fast Fourier Transform,' which calculates the discrete Fourier transform in $N \log N$ operations.

3.4.2 Quantum Fourier Transform definition

The definition of the QFT is analogous to the discrete Fourier transform:

$$\text{QFT}\,|x\rangle \to \frac{1}{\sqrt{N}} \sum e^{2\pi i x y/N}\,|y\rangle\,. \tag{177}$$

To implement this unitary transformation, we will make use of the n qubits to represent the N vectors in the Hilbert space efficiently, such that $N = 2^n$ as previously stated. However, we must first clarify some notation.

Mathematical remark

We will use the binary representation, such that $x = x_0 2^0 + x_1 2^1 + ..., x_{n-1}2^{n-1} \equiv (x_{n-1}, ..., x_0)$. The *binary fraction* is defined as $0.x_{n-1}....x_0 \equiv \frac{x_{n-1}}{2^{n-1}} + \frac{x_{n-2}}{2^{n-2}} + ..., \frac{x_0}{2^0}$. The reason why this representation is useful can be understood as follows: we need to write the phase factor in front of each $|y\rangle$ for arbitrary integers, and we can use similar algebra as Simon's problem if the conventional multiplicative number xy is considered in this binary representation. Consider the phase $\Phi = \frac{2\pi xy}{2^n} = \frac{2\pi}{2^n}(y_0 2^0 + ..., y_{n-1}2^{n-1})(x_0 2^0 + ..., x_{n-1}2^{n-1})$. Let's further examine these terms:

$$\Phi = 2\pi y_{n-1}(\frac{x_0}{2} + x_1 + 2x_2 + ...) + 2\pi y_{n-2}(\frac{x_0}{4} + \frac{x_1}{2} + x_2 + ...) + ... \tag{178}$$

All of the terms that give rise to 2π phase shifts are irrelevant (for example, the ones proportional to x_1 and $2x_2$ in the first sum, recalling that each x_i, y_i is 0 or 1). The sums above are reminiscent of the binary fraction:

$$\Phi = 2\pi(y_{n-1}0.x_0 + y_{n-2}0.x_1x_0 + y_{n-3}0.x_2x_1x_0 + ...) + ... \tag{179}$$

It becomes evident that the phase in the discrete Fourier transform can be written as a product of x and y written as binary representation. With this mathematical remark understood, we can write the QFT as:

$$\begin{aligned}\text{QFT}\,|x_{n-1}...x_0\rangle &= \frac{1}{2^{n/2}}(|0\rangle + e^{2\pi i 0.x_0}\,|1\rangle) \otimes (|0\rangle + e^{2\pi i 0.x_1x_0}\,|1\rangle) \\ &\quad \otimes (|0\rangle + e^{2\pi i 0.x_2x_1x_0}\,|1\rangle)...(|0\rangle + e^{2\pi i 0.x_{n-1}...x_0}\,|1\rangle) \\ &= \sum_{y=0}^{N-1} A_y\,|y_{n-1}...y_0\rangle\end{aligned} \tag{180}$$

This identity can be understood as a series of phase factors:

$$A_y = e^{2\pi i(y_{n-1}0.x_0 + y_{n-2}0.x_1x_0 + ... + y_0 0.x_{n-1}...x_0)}. \tag{181}$$

One can verify using this identity that the QFT does what we want it to do.

Remarks

1. The QFT takes a product state into another product state. It is not an entangling operation. However, there is some subtlety: the phase of one qubit depends on the input values of other qubits.

2. The operation we have created is a generalization of H^n acting on the block of n qubits. Indeed, the Hadamard operation acting on the kth qubit belongs to this family:

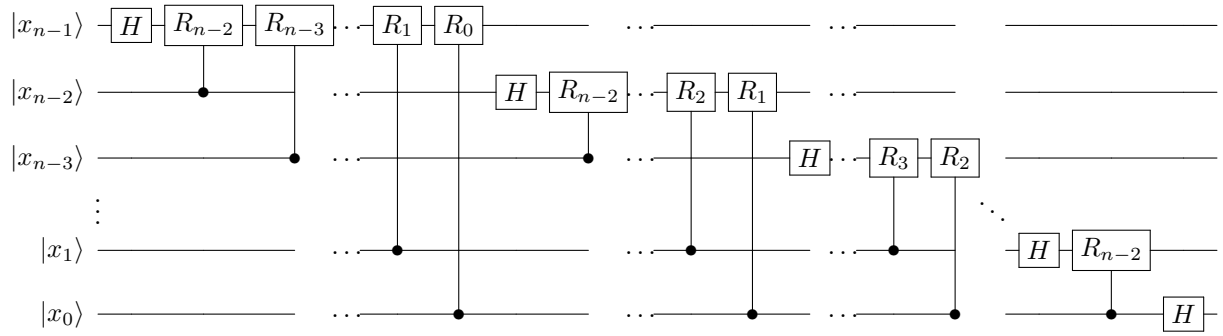
$$H\,|x_k\rangle = \frac{1}{\sqrt{2}}(|0\rangle_k + e^{2\pi i 0.x_k}\,|1\rangle_k) \tag{182}$$

   When H acts on $|0\rangle_k$, it produces $|+\rangle$, but when it acts on $|1\rangle_k$, it produces $|-\rangle$.

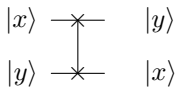
3. The phase operator R_d will imprint a phase on a qubit:

$$R_d = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^{n-d}} \end{pmatrix}. \tag{183}$$

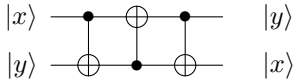
4. We can put together the QFT from Hadamard and R_d gates. The QFT quantum circuit is:



The phase on qubit 4, Φ_4, has a contribution of $2\pi\frac{x_3}{2}$ from qubit 3, a contribution of $2\pi\frac{x_2}{2^2}$ from the second qubit, and so on. Note that the circuit is not complete. In order to ensure consistency in the ordering of the bits (most significant bit at the top, down to least significant bit at the bottom), we need to swap different qubits between each other, using the operation SWAP $|x\rangle\,|y\rangle \to |y\rangle\,|x\rangle$. This is typically denoted in circuit form as



This can be constructed as a sequence of three CNOTs, as shown here,



which can be proven rigorously as an exercise. Note that we can swap the qubits either before or after the conditional rotations - it does not matter, it just reorders the qubits to ensure consistency between most-significant and least-significant bits.

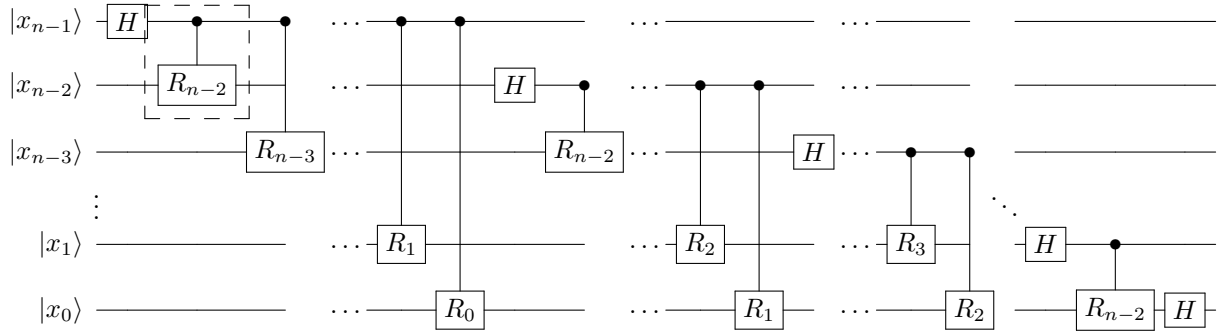
5. We can count the total number of steps, and see that we need at most

$$\frac{n(n+1)}{2} + \frac{n}{2} \sim n^2 \tag{184}$$

operations, where the first term corresponds to the conditional rotations, and the second term accounts for the swap operations. We see that this is an **exponential speedup compared to the FFT**, which is $O(n2^n)$ (remember that $n = \log_2 N$).

6. If we look at these conditional rotations R_d in detail, in particular for small d, we see that these rotations implement a very small phase $\sim 2^{-n}$. The circuit we have drawn is an exact quantum Fourier transform, but we can make an approximation by neglecting gates with distance $> m$, which introduces an error at most $n2^{-m}$.

7. One can also simplify the circuit by noting that controlled-phase gates are symmetric with respect to exchanging the control and target, since the gate implements $|11\rangle \to |11\rangle\, e^{i\phi}$. In other words, it does not matter which qubit is the control, and which qubit is the target, regardless of the phase accumulated. We can then replace the role of the control and target qubits in our original circuit:



Suppose we simply measure the qubits after the QFT. Now, if we start by measuring the first qubit

$|y\rangle_0 \to y_0$, we know exactly how to apply the phases to the next qubits, without having done any two-qubit operations. In particular, instead of applying the conditional R_{n-2} rotation to the next qubit, assume we measure the top qubit, giving us the value y_0. Now we can just apply the single qubit rotation

$$|x\rangle_{n-2} \to H(R_{n-2})^{y_0} |x\rangle_{n-2} = |y\rangle_1 . \tag{185}$$

Next, we can use this result to show that the next qubit will rotate in the following way

$$|x\rangle_1 \to H(R_{n-2})^{y_1}(R_{n-3})^{y_0} = |y_2\rangle . \tag{186}$$

We can continue this up the chain, and what we find is that we do NOT need any two-qubit operations! This is in line with our intuition that the QFT operation does not actually entangle the qubits. Remember, it takes a product state and transforms it into another product state, so it is not surprising that this is possible. All that is required is single qubit manipulations, as well as the feed-forward procedure required to implement gates of the form of (185) and (186).

3.5 Quantum phase estimation

Quantum phase estimation is an application of the quantum Fourier transform, that was developed by Kitaev. It is used for, e.g., finding eigenvalues of a Hamiltonian. It was developed after Shor's algorithm, but it is useful to understand prior to Shor's algorithm, since Shor's algorithm, at it's core, relies on it.

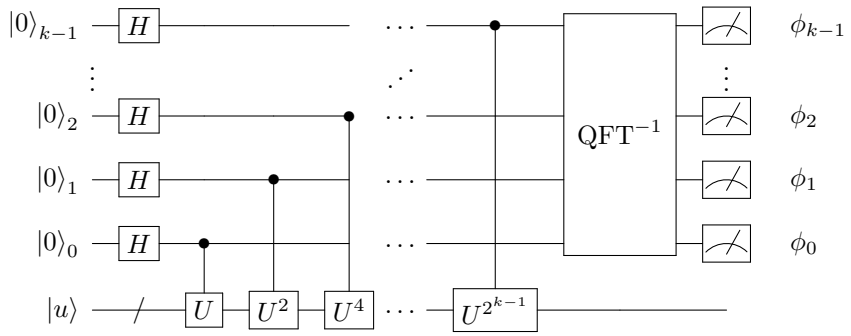
Consider a unitary U with an eigenstate $|u\rangle$. For example, this unitary could result from Hamiltonian evolution,

$$U = e^{iHt/\hbar}, \tag{187}$$

where the eigenstates of the Hamiltonian H will be eigenstates of U, and the eigenvalues of the Hamiltonian (energies) will be proportional to the acquired phases

$$U|u\rangle = e^{2\pi i\phi}|u\rangle . \tag{188}$$

Remember: if we write down the energy spectrum of the Hamiltonian $H|u\rangle = E|u\rangle$ we accrue a phase of $2\pi\phi = \frac{Et}{\hbar}$ under free evolution. In general, the eigenvalues of U can always be written as $e^{i\phi}$, since we know that $U^\dagger U = \mathbb{1}$, such that the eigenvalues of $U^\dagger$ will be $e^{-i\phi}$. The goal of quantum phase estimation is to find the phase ϕ, which can be approximated using k bits in binary notation $0.\phi_{k-1}\phi_{k-2}...\phi_k$, given a unitary matrix U and eigenvector $|u\rangle$. The quantum phase estimation algorithm is given by the circuit,



where the slashed wire next to state $|u\rangle$ indicates that it is possibly a multi-qubit register (and thus, U can be a multi-qubit gate).

We prepare our k input qubits in the superposition state $H^{\otimes k}|0\rangle$. These k qubits act as control qubits implementing these controlled unitaries on our data register, which is in the subspace of U. Next, we will use the kth qubit to implement conditional $U^{2^{k-1}}$. Consider the case where the data register is initialized in $|u\rangle$. The unitary U will result in accumulation of the phase $e^{2\pi i\phi}$. Thus, we can describe the effect solely on the state of the control qubit, which transforms

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \to \frac{|0\rangle + e^{2\pi i\phi}|1\rangle}{\sqrt{2}}. \tag{189}$$

Similarly, the second qubit will transform as

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \rightarrow \frac{|0\rangle + e^{2\pi i\phi(2)}|1\rangle}{\sqrt{2}}, \tag{190}$$

continuing to the kth qubit which picks up the phase

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \rightarrow \frac{|0\rangle + e^{2\pi i\phi(2^{k-1})}|1\rangle}{\sqrt{2}}. \tag{191}$$

Clearly, the data register is completely unchanged since it is an eigenstate of U^{2^j}. Going back to our Hamiltonian evolution, we can see that this evolution is obtained by evolving for times $t, 2t, 4t, ..., 2^{k-1}t$ conditionally on the state of the relevant data register qubit.

At this point, we need to implement the inverse quantum Fourier transform. To see why this is useful, we can explicitly consider the output of these conditional unitaries:

$$|\psi_I\rangle = \frac{1}{2^{k/2}}(\,|0\rangle + e^{i2\pi\phi 2^{k-1}}|1\rangle\,) \otimes ... \otimes (\,|0\rangle + e^{i2\pi\phi 2^0}|1\rangle\,) \otimes |u\rangle\,. \tag{192}$$

Now, we can rewrite this expression using binary fractions. Consider, for simplicity, the special case where $\phi = 0.\phi_{k-1}...\phi_0$ exactly. In other words, $\phi < 1$ can be expressed with exactly k digits. Looking at the first phase factor in (192), we see that

$$2\pi 2^{k-1}\phi = 2\pi(\phi_{k-1}2^{k-2} + ... + \frac{\phi_0}{2}) = 2\pi\frac{\phi_0}{2} + 2\pi m = 2\pi(0.\phi_0) + 2\pi m,\ m \in \mathbb{Z}. \tag{193}$$

We see that only the ϕ_0 term contributes, since the rest are integers and are multiplied by 2π, and can be neglected. We can continue this analysis for all terms in (192). For example, considering the next term, we will have contributions from the last two digits

$$2\pi 2^{k-2}\phi = 2\pi(0.\phi_1\phi_0) + 2\pi m,\ m \in \mathbb{Z}. \tag{194}$$

Now we can rewrite (192) in a form that is extremely familiar from our definition of the quantum fourier transform:

$$|\psi_I\rangle = \frac{1}{2^{k/2}}(\,|0\rangle + e^{i2\pi(.\phi_0)}|1\rangle\,) \otimes ... \otimes (\,|0\rangle + e^{i2\pi(.\phi_{k-1}...\phi_0)}|1\rangle\,), \tag{195}$$

where we have now dropped the irrelevant register $|u\rangle$. We can see that this is exactly the quantum fourier transform of the register $|\phi_{k-1}...\phi_0\rangle$, in other words,

$$\mathrm{QFT}^{-1}|\psi_I\rangle = |\phi_{k-1}...\phi_0\rangle\,. \tag{196}$$

Now, all we have to do is measure the k qubits in the computational basis, and we get all digits of ϕ at once.

Remarks

1. Consider the case of a general phase ϕ which cannot be exactly represented in terms of k digits. Quantum phase estimation gives a very good approximation to ϕ, with very high probability. We can rewrite the intermediate state prior to the inverse QFT

$$|\psi_I\rangle = \frac{1}{2^{k/2}}\sum_{x=0}^{2^k-1} e^{i2\pi\phi x}|x\rangle\,, \tag{197}$$

where we have now transformed to decimal integer value $x \in [0, 2^{k-1}]$. Now we can perform the inverse quantum fourier transform and get

$$\mathrm{QFT}^{-1}|\psi_I\rangle = \frac{1}{2^k}\sum_y\sum_x e^{i2\pi(\phi x - \frac{xy}{2^k})}|y\rangle\,. \tag{198}$$

We can now consider the probability of getting a certain outcome y,

$$P(y) = |a_y|^2 = \frac{1}{2^{2k}} \left| \sum_{x=0}^{2^k-1} e^{i2\pi(\phi - \frac{y}{2^k})x} \right|^2. \tag{199}$$

This is a geometric series which can be evaluated to

$$P(y) = \frac{1}{2^{2k}} \left| \frac{r^{2^k} - 1}{r - 1} \right|^2, \tag{200}$$

where r is

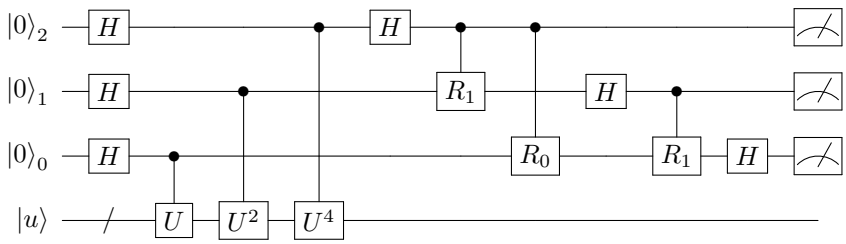
$$r = e^{i2\pi(\phi - \frac{y}{2^k})}. \tag{201}$$

Now we can write y and ϕ in binary notation as

$$\begin{aligned} \frac{y}{2^k} &= \frac{y_0}{2^k} + ... + \frac{y_{k-1}}{2} \\ \phi &= \frac{\phi_{k-1}}{2} + ... + \frac{\phi_0}{2^k} + \delta, \end{aligned} \tag{202}$$

where δ is the truncation error from expressing ϕ with only k bits. Now, looking at (200), we see that this is a sharply peaked function $\phi_i = y_i$ and $\delta \to 0$. Formally, one can show that $P(y) \sim 1$ requires that $\delta \ll \frac{1}{2^k}$. In other words, on average we will, with probability exponentially approaching 1, get the outcome that is exponentially close to the correct phase.

2. Consider the example of $k = 3$, shown in the following circuit.



Looking at the top qubit, we see that we prepare a superposition state, and apply the controlled phase rotation

$$U^4 |u\rangle = e^{i2\pi(4)(\frac{\phi_2}{2} + \frac{\phi_1}{4} + \frac{\phi_0}{8})} = e^{i2\pi\frac{\phi_0}{2}}. \tag{203}$$

(Note that the overall operation looks a bit like an interferometer, where we prepare a superposition, apply a phase, and then return back to the original basis.) Now, if we measure this phase, we see that we will measure $|0\rangle$ if $\phi_0 = 0$, and $|1\rangle$ if $\phi_0 = 1$. So this circuit simply measures the least significant digit of ϕ.

This is a bit surprising - we would expect to begin by measuring the most significant digit. But this is the clever aspect of quantum phase estimation! Now, we can move onto the next qubit, which will now be sensitive to both the least significant digit, and the second least significant digit:

$$U^2 = e^{i2\pi 2(\frac{\phi_2}{2} + \frac{\phi_1}{4} + \frac{\phi_0}{8})} = e^{i2\pi(\frac{\phi_1}{2} + \frac{\phi_0}{4})}. \tag{204}$$

However, the controlled rotation from the first qubit, as part of the inverse QFT, exactly compensates for the value of the least significant digit, allowing the second qubit to measure the second-least-significant digit precisely. Continuing along this line of reasoning, we can see how the QFT is used to measure each subsequent digit of the phase very efficiently in this unintuitive order.

3. This algorithm is very useful for estimating energies of eigenstates of a Hamiltonian.

4. However, in general, controlling the Hamiltonian evolution required for this algorithm is challenging to experimentally implement.

5. Because of the sequential nature of the measurement, often times we don't need many qubits to estimate a phase with high precision. It turns out, with only $k = 1$, we can already measure many digits with high precision. The idea here is that we can reuse this qubit, since we are able to perform operations in a sequential manner. We first use the qubit to measure ϕ_0 by using a long time evolution. Then we measure ϕ_1 by cutting the measurement time in half, and so on.

3.6 Order finding and factoring

The order-finding and subsequent factoring algorithm is a canonical (and rare) example of an application of quantum computers. Certain cryptography protocols (e.g., public key cryptography) leverage factoring, a computationally challenging problem, to generate a secure key. We will show below that quantum computers can be used to factor large numbers more efficiently than classical computers, and quantify the speedup. Order finding and factoring is also another application of the QFT. In this section, we will describe the order finding and factoring algorithms in detail.

Consider two positive integers a and N such that $a < N$, and which have no common factors. We call r an *order* of $a \mathrm{mod} N$ if r is the smallest integer such that $a^r = 1 \mathrm{mod} N$. This is equivalent to

$$a^r = bN + 1, \tag{205}$$

for $b \in \mathbb{Z}$. As a simple example, consider $a = 5, N = 44$. We can brute force check $r = 2, 3, ..$ to check which r is the order. In this case, we find that $r = 5$ is the order, since $5^5 = 3125$, and we see that $3125 - (71 \times 44) = 1$, satisfying (205). For large numbers, this procedure is computationally challenging. We can get some intuition for why in our original approach for finding the order r, the simple example above, where we "brute-force" checked the values of r in ascending order.

Mathematical remarks

Order finding is very closely related to factoring. In the factoring problem, we are given the large number N, and we need to find its factors p, q where $N = pq$. Of course, if we know p or q, finding the other factor is easy. However, if neither is known, this is believed to be hard classically. This problem serves as the basis for public key cryptography (such as RSA, which is widely used today). The basic idea of RSA is that Alice can choose two values p and q, send N to Bob, and to decrypt the message, one must know either p or q. A few statements are in order:

- The fastest known classical algorithm for factoring scales with $O(N^{1/3})$, where N is exponential in the number of digits, or number of bits required to encode the number.
- It is actually believed that faster classical algorithms exist, but they are not yet known.
- In fact, the problem itself is not computationally hard on a quantum computer, since we will see that Shor's algorithm, reducing factoring to order finding, can solve the problem in a number of steps that scales with $\log N$.

In fact, in the following we will describe an algorithm which shows that order finding is equivalent to factoring. We consider a large number N to be factored. The steps are as follows:

1. Pick some $a < N$, and check if the greatest common denominator is 1. This can be done using the Euclid algorithm, where if we assume $N = pq$, and $a = pz$, and that an integer b_1 exists such that:

$$N = b_1 a + r_1, \tag{206}$$

   then r_1 is *also* divisible by p.

2. We can continue the Euclidean algorithm another step:

$$a = b_2 r_1 + r_2, \tag{207}$$

   where we find that by the same logic r_2 is also divisible by p if N and a are divisible by p. Recall that for factoring, we need to find these two numbers which have a common factor different than 1. Therefore, if after many steps we find $p \neq 1$, our problem is solved.

3. However, if $p = 1$, we must find r such that (an order-finding problem) a^r is

$$a^r = bN + 1. \tag{208}$$

If r is odd, we have to repeat this procedure from the beginning (picking another a).

4. If r is even, we can write this expression as

$$\begin{aligned} &(a^{r/2} + 1)(a^{r/2} - 1) \\ &= bN \\ &= bpq. \end{aligned} \tag{209}$$

At this point, one should check if $(a^{r/2} \pm 1)$ is divisible by N. If it is, then we must start from the beginning.

5. However, if not, $(a^{r/2} \pm 1)$ shares a common divisor with N. At this point, one can use the efficient Euclidean algorithm to find the divisor.

In other words, if we can efficiently find r that satisfies (205), we can use the Euclidean algorithm to find the factors of N efficiently. So the problem of factoring indeed reduces to order finding.

Now, with this mathematical remark aside, let us consider the quantum algorithm for order finding. Our approach will be based on quantum phase estimation. We begin by constructing a unitary

$$U_a \left|y\right\rangle = \left|ay \mathrm{mod} N\right\rangle, \ 0 \leq y \leq N - 1. \tag{210}$$

Consider the eigenstates of U_a, which can be formally written in the following form

$$\left|u_s\right\rangle = \frac{1}{\sqrt{r}} \sum_{m=0}^{r-1} \exp\left(- \frac{2\pi i s m}{2} \right) \times \left|a^m \mathrm{mod} N\right\rangle, \ 0 \leq s \leq N - 1 \tag{211}$$

where r is the order as seen above. We can show formally that this is true by acting $U_a \left|u_s\right\rangle$, which will simply lead to

$$U_a \left|u_s\right\rangle = \frac{1}{\sqrt{r}} \sum_{m=0}^{r-1} \exp\left(- \frac{2\pi i s m}{2} \right) \times \left|a^{m+1} \mathrm{mod} N\right\rangle, \tag{212}$$

where we can now rewrite the terms in the sum using the fact that r obeys $a^r = 1$ from (205), so that the last term $\left|a^r \mathrm{mod} N\right\rangle$ can be rewritten as $\left|1\right\rangle$. We can rewrite (212) as

$$U_a \left|u_s\right\rangle = \exp\left(\frac{2\pi i s}{2} \right) \left|u_s\right\rangle, \tag{213}$$

so we see that $\left|u_s\right\rangle$ are all eigenstates of U_A with eigenvalue given by $\exp(\frac{2\pi i s}{2})$. At this point, we can implement order finding using the following steps:

1. Prepare some eigenstate (or superposition of eigenstates)

2. Perform QPE to find s/r. With k auxiliary qubits, we can determine with high accuracy k digits of the fraction s/r.

3. Extract s and r, which are integers, from the ratio. This is a bit subtle, but suppose we know the fractional number s/r exactly (e.g. 0.153 in decimal). We can easily write this as $\frac{153}{1000}$, and simplify to get s and r. The subtlety arises from the fact that we have an approximation to s/r, but it turns out since we know this with exponential accuracy, we can use so-called continued fractions to obtain very good estimates of s and r.

4. Once we have a vaule of r, we can easily check that it is an order.

Remarks

1. How do we prepare $|u_s\rangle$? Let us consider the superposition

$$\frac{1}{\sqrt{r}}\sum_{s=0}^{r-1}|u_s\rangle = \frac{1}{r}\sum_{m=0}^{r-1}\sum_{s=0}^{r-1}\exp\Big(-\frac{2\pi i s m}{r}\Big)|a^m \mathrm{mod} N\rangle\,. \tag{214}$$

If $s \neq 0$, each term will have a phase, these terms will interfere destructively and add to 0. So we can rewrite (214) as

$$\sum_{m=0}^{r-1}\delta_{m,0}\,|a^m, \mathrm{mod} N\rangle = |1\rangle\,. \tag{215}$$

This means that we can actually produce the superposition in (214) by simply preparing the binary state $|1\rangle$, and run this phase estimation. Then, for different instances of phase estimation, we will sample different eigenstates for U_A.

To summarize this approach, we will start in $|1\rangle$ and run phase estimation several times. In the first instance, we will estimate $\frac{s_1}{r}$, then $\frac{s_2}{r}$, etc. Knowing a few s_i/r will help us determine the simplified expression $\frac{s}{r}$.

2. It remains to show that we can implement QPE efficiently. This is done using modular exponentiation. Recall that we need to implement e.g.

$$\begin{aligned} U_a^2\,|y\rangle &= |a^2 y \mathrm{mod} N\rangle \\ U_a^4\,|y\rangle &= |a^4 y \mathrm{mod} N\rangle \\ &\cdots \end{aligned} \tag{216}$$

Consider the register $|x\rangle\,|y\rangle$, where $|x\rangle$ is the auxiliary register we use to implement QPE, and $|y\rangle$ stores the superposition of eigenvectors to U_a. We want to implement the transformation

$$\begin{aligned} |x\rangle\,|y\rangle &\to |x\rangle\,U_a^{x_k 2^{k-1}}...U_a^{x_0 2^0}\,|y\rangle \\ &= |x\rangle\,|a^{x_k 2^{k-1}+x_0 2^0} y \mathrm{mod} N\rangle \\ &= |x\rangle\,|a^x y \mathrm{mod} N\rangle. \end{aligned} \tag{217}$$

If we can implement $|1\rangle \to |a \mathrm{mod} N\rangle$, then $|1\rangle \to |a^2 \mathrm{mod} N\rangle$, etc, then the entire procedure will only require k steps to implement the unitary to the power 2^k. What remains is to show how to implement this modulo square. This is similar to conventional multiplication and requires only n^2 operations. This step is a bit cumbersome, but is related to finding a classical multiplication algorithm/circuit and making it reversible. At this point, one can implement it using e.g. Toffoli gates. In total, this requires $kn^2 \sim n^3 \sim (\log N)^3$.

3. We can consider starting with

$$H^{\otimes k}\,|0...0\rangle\,|1\rangle = \sum |x\rangle\,|1\rangle\,, \tag{218}$$

and after the modulo square and subsequent exponentials, we get

$$\sum |x\rangle\,|a^x \mathrm{mod} N\rangle\,. \tag{219}$$

Note that r is exactly the period of the function (a^x mod N). At this point we can perform QFT. We can see the effect of this by writing the expression in (219)

$$\sum_m \Big(\sum_j |m + jr\rangle\Big)\,|a^m \mathrm{mod} N\rangle\,. \tag{220}$$

Now we see that QFT will exactly reveal this period r. (More precisely, QFT will give us a very precise estimate of r.)

3.7 Implementing quantum algorithms

Errors are kind of like a virus - they spread exponentially. We wonder if there is any hope to build a quantum computer with many qubits that is of practical use, in the presence of errors. There are some ideas developed theoretically that allow for this. The cycle is development of algorithms, to building, to fighting noise, and then designing algorithms that can work in the presence of noise - *co-design* is the idea of thinking of this problem with all of these steps together. If we start coming up with new algorithms, what kind of means do you need in a lab to implement them? An important concept in this field is *universality*. It turns out that any quantum computation can be built from a discrete set of operations. The simplest and best-known approach is based on quantum circuits, which we have already considered. Today, we will show that any quantum computation can be built from a finite set that is a so-called universal set of quantum gates.

3.7.1 Universality theorem

To begin, we will prove the theorem. The theorem states that any U on a dxd dimensional system can be decomposed into $U = U_1 U_2 ... U_k$ with U_i acting only on two states, and where $k \leq d(d-1)/2$. The idea of the proof (see Nielsen & Chuang) is to construct $U_k^\dagger U_{k-1}^\dagger ... U_1^\dagger U = \mathbb{1}$, and find the U_i. Suppose we have a $3x3$ matrix:

$$U = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & j \end{pmatrix} \tag{221}$$

We find $U_1^\dagger U = \mathbb{1}$, such that $b \to b' = 0$, such that:

$$U_1^\dagger = \frac{1}{\sqrt{|a|^2+|b|^2}} \begin{pmatrix} -a^* & b^* & 0 \\ b & -a & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{222}$$

Next we find $U_2^\dagger$ such that $U_2^\dagger U_1^\dagger U = \mathbb{1}$

$$U_2^\dagger = \begin{pmatrix} 1 & 0 & 0 \\ 0 & ... & ... \\ 0 & ... & ... \end{pmatrix} \tag{223}$$

For a d-dimensional matrix, we use the same procedure and find that order d^2 steps are needed. However, there is a problem: d scales exponentially with the number of qubits, such that the total number of operations would scale exponentially with the number of qubits. Next we will like to construct two-level matrix corresponding to arbitrary states from single and two qubit gates. Suppose we want to construct a unitary between two states $|101001\rangle$ and $|110011\rangle$. Note that the operations should not influence any other states, such that this is only influences the two states chosen. Our method is the following: we would like to take the first state and flip bits one by one in such a way that only one bit changes at a time, conditional on the other bits. At each step there is a controlled flip so that only each bit is flipped if the state corresponds to the chosen starting and ending states. That way we can stay *within* the subspace of the two states chosen. We can therefore get arbitrary rotation among these two many-body states using these controlled rotations. This requires only *linear* in the number of qubits n, so we can implement an arbitrary 2-level unitary using only $\sim n$ steps.

For example, recall that the gate $C^n Z = 2(n-2)C^2 Z$ which is linear in n. Thus each two level U is at most n^2 single and two qubit gates.

3.7.2 Discrete set of universal operations

The following set of gates is universal: H, CNOT, and T, where we have the T gate:

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/8} \end{pmatrix} \tag{224}$$

The rotation by $\pi/8$ is important to get a rotation away from one of the principal axis. It turns out that this is sufficient to approximate any single-qubit rotation, as will be proved on the homework. The angle of rotation $\cos\theta/2 = \cos^2 \pi/8$ can be generated, which is an irrational fraction of π, and that can be used to generate any rotation on the Bloch sphere. The *Solovay-Kitaev* theorem shows that the number of rotations is actually

logarithmic in $1/\epsilon$, where ϵ is the error. However, for an arbitrary unitary, this will require an exponential number of gates. The art of quantum algorithms is to find algorithms such that they scale polynomially in the number of qubits n. This is the basis of quantum complexity classes.

3.7.3 Other approaches to quantum computation

The gate model is universal, but it is by no means unique: there are other approaches to implementing quantum algorithms and building quantum computers. One approach is to build a machine with a complicated range of interactions that implements the Hamiltonian that corresponds to the desired unitary U. If we succeed in buidling this Hamiltonian, it is relatively easy to implement the desired evolution, but generally it is not easy to build the desired Hamiltonian. In general, the Hamiltonian will have products of k qubits, with k running from 1 to n. Such terms are called 'k-local' terms, and the Hamiltonian becomes very complicated. If we can do this, what will be the power of this machine? Can it be universal? We will discuss this point.

Quantum computing by measurement is another approach. If you start with a given state, such as a 2D cluster state, and perform sequential measurements on it. For example, by performing measurements and single qubit operations on columns of a 2D cluster state, effective operations are performed on the remaining qubits. If the measurements are started to the left, the physics of the effective operations on the remaining bits is like time in the circuit-based model propagating to the right. This approach is also universal.

Examples

1. Quantum simulations. Suppose you are given a Hamiltonian and you would like to make some predictions: what is the ground state? What is the energy of the ground state? What are the quantum dynamics after turning on the Hamiltonian and starting in a particular state (called a *quantum quench*)? These problems are hard classically, because it requires solving 2^n differential equations, corresponding to the dimension of the Hilbert space. This is quite hard for e.g. $n > 50$. Note that if there are no interactions (or more precisely, specific types of interactions), the number of equations can be linear in n and this can become tractable, but in general, it is not. There are two ways to answer these questions. One is to prepare the Hamiltonian and the state in the desired system and let it run the dynamics under the Hamiltonian, then perform measurements. However, if you want to measure a more subtle quantity like the energy of the ground state, you can combine this approach with quantum phase estimation. Current quantum computers which have a number of qubits on the order of 50 are already useful for this application. This simulation idea is how the field of quantum computers started - Richard Feynman stated that we better use quantum systems to simulate quantum systems!

2. Adiabatic algorithms can be best formulated in terms of solving combinatorial optimization problems such as NP complete problems like the Traveling Salesman problem, MaxCut, and Maximum Independent Set (MIS). These problems attempt to find a minimum of a cost function of a certain number of variables. Suppose these variables are bits (classical problem). You can write this cost function in terms of the qubit states and solve them using a quantum computer. Often, terms in the cost function depend on products of multiple states, which correspond to a many-qubit interaction. Some of these problems can be described using a graph - suppose that we have a number of vortices and they are connected by links. The maximum independent set is the maximum number of vertices that are not directly connected by the links. The problem of MIS involves finding the maximal number of these vertices, without violating this independent set constraint. This problem is very easy to formulate and check for a given possible solution, but it is an NP complete problem (for some examples it is even harder than NP complete). Even finding an approximate solution is quite hard. One approach where quantum computers can help is to design a cost function $H(z_1, ...z_N)$ as a Hamiltonian corresponding to the cost function which has a form depending on the clauses (constraints), for example $H_p = aZ_1Z_2 + bZ_2Z_3Z_4 + cZ_4Z_5 + ...$ with penalization $a, b, c, ...$ corresponding to the clauses. The ground state corresponds to the lowest value of h, the cost function - which is solving the goal of the problem: $H_p |Z_1...Z_N\rangle = h(Z_1, ...Z_N) |Z_1...Z_N\rangle$, simulating the ground state and measuring it guarantees that we solve this problem. This makes a correspondence between hard problems in computer science to finding ground states of complex spin models and statistical mechanics. Now, with this corresponding established, we can try to design a new family of algorithms, called adiabatic

quantum algorithms. Consider a situation where the Hamiltonian depends on time, such that at $t = 0$, the Hamiltonian has a ground state that is easy to prepare, for example $a \sum_i x_i$. Once we prepare the state $|-\rangle$, we know we are in the ground state. By changing the Hamiltonian to the desired Hamiltonian H_p over a long time T, we are guaranteed to be in the ground state, as long as the change is performed slowly enough. How slow is slow enough? The time T must take longer than the energy gap, such that the Fourier limit of this time is smaller than the energy gap. However, the gap in the spectrum generically closes exponentially for complicated Hamiltonians including any arbitrary interactions, such as a spin glass. The adiabatic quantum computer can be shown to be universal, meaning it is equivalent to the circuit model with polynomial overhead.

3. Variational quantum algorithms. Two examples we will consider are called Quantum approximate optimization algorithm (QAOA) and Variational Quantum Eigensolver (VQE). QAOA is targeted for classical optimization problems, whereas VQE is designed for solving hard quantum mechanical problems.

   The idea behind variational quantum algorithms is a generalization of the adiabatic algorithm. Recall that the adiabatic Hamiltonian is

$$\gamma H_p + \beta \sum_i X_i. \tag{225}$$

   Similarly, we begin by defining unitaries

$$\begin{aligned} U_p(H_p, \gamma) &= e^{-i\gamma H_p} \\ U_x(\beta) &= \prod_{i=1}^{n} e^{-i\beta_i X_i} \end{aligned} \tag{226}$$

   and preparng qubits in the symmetric superposition

$$|s\rangle = \frac{1}{\sqrt{2^n}} \sum_z |z\rangle \,, \tag{227}$$

   which is obtained by performing $H^{\otimes n} |0\rangle^{\otimes n}$. This implies that $X_i |s\rangle = |s\rangle$, in other words $|s\rangle$ is an eigenstate of each individual Pauli X operator with eigenvalue 1.

   In QAOA, we apply $U_p U_x$ k times, giving us

$$|\vec{\gamma}, \vec{\beta}\rangle = U_{x_k}(\beta_k) U_{p_k}(\gamma_k) ... U_{x_1}(\beta_1) U_{p_1}(\gamma_1) |s\rangle \,. \tag{228}$$

   Here, we measure the cost function and find new $\vec{\gamma}, \vec{\beta}$, and optimize over these vectors (trajectories) in order to minimize the cost function. Here, k is an effective circuit depth, and as $k \to \infty$, this converges to the adiabatic path. However, this is a more general approach since we can in principle reach the cost function minimum faster (in a fininte number of states).

   Remarks

   (a) QAO (adiabatic limit), QAOA are examples of heuristic algorithms. This means that we do not yet have any guarantee that the algorithm will actually provide a speedup. The general belief in the field is that we must now build quantum computers to test how they will perform. This may sound incredible, but this was certainly the case for numerous classical algorithms on ordinary computers. One prominent example of this are so-called Deep Neural Networks, which solve hard problems quickly, yet noone knows why they work so well.

   (b) QAOA has close connections with the gate model of quantum computing, since $U_p(\gamma)$ can be built from 2-qubit gates.

   The current frontier of QAOA is interested in tackling the following problems

   - How expensive is it to encode a given, useful Hamiltonian? What is the overhead in implementing the optimization?

- Can we make use of hardware implemented n-qubit gates? Does this map effectively to n-local models?
- Does the performance depend on the hardware implementation (e.g. gate connectivity)
- How can we design the algorithm specific to the hardware implementation (co-design)

4 Implementation of quantum computers

In the early 2000's, David DiVincenzo set forth general criteria for a quantum computer, based on the idea of universality. These are known as *DiVincenzo criteria*:

- A set of well-defied qubits with long coherence times
- The ability to initialize all qubits in a simple state
- The ability to individually control and measure qubits
- The ability to perform two-qubit operations, including the ability to turn the required interactions on and off
- Need to be able to scale qubits, such that adding qubits can be done with resources $\sim O(n)$

It turns out that almost any non-harmonic quantum system can be used as a quantum computer! In the 2000's, this resulted in a lot of creative ideas and proposals for building quantum computers (in fact, even some based on Harmonic systems).

In practice, DiVincenzo's criteria turned out to be only a guide. In today's modern research, we often see the terms *NISQ* (Noisy intermediate-scale quantum era - the era we currently live in), as opposed to *FASQ* (Fault-tolerant application quantum era - hopefully a near-future era). The issues which have emerged is that individual qubits (or a few qubits) can be controlled and operated with exquisite precision. However, often even individual gate fidelities scale with n - in other words, it becomes harder to build a high-fidelity machine when there are more qubits to control. Physical systems may employ collective gates on large registers, which can be advantageous in some situations, but can also be a hindrance in implementation of simple single and two-qubit gates. The ability to perform two-qubit operations between certain qubits (i.e. connectivity) is yet another limitation. Finally, the ability to control and *readout* all qubits in parallel with high fidelity is a major challenge. A summary (by no means comprehensive, or up-to-date!) is presented in Fig. 9

4.1 Neutral atoms and ions: background

These are individual atoms or ions suspended in vacuum "tubes" or chambers, which have very good quantum coherence when individually isolated and cooled. These, in fact, form the best timekeepers (atomic clocks) which are the basis for time standards and global positioning systems (GPS). A major challenge is to isolate and manipulate single atoms in a way that can be scaled up to many atoms. A second major challenge is to engineer the required controllable interaction to make two-qubit interactions.

Summary of atomic physics

A generic atomic level structure can be understood by simply studying the quantum mechhanical problem of the Hydrogen atom: a single negatively charged electron orbiting around a single positively charged proton. By solving the Schrödinger equation for the potential between these particles, we obtain an energy spectrum of discrete levels, which are conventionally described by several quantum numbers $|n, l, s, I, m_l, m_s, m_I\rangle$, where n is the principle quanutm number (electronic orbital level), l is the angular momentum quantum number, s is the electronic spin degree of freedom, I is the nuclear spin sublevel, and the m_i are the individual values for the various angular momenta.

We can see that this results in a fairly complex spectrum! Each principle orbital will be split by the various internal spin degrees of freedom into several levels. In the case of Hydrogen, we have a spin-1/2 electron and spin-1/2 nucleus, hybridization of which (via hyperfine interaction) results in the formation of triplet and singlet states (a total of 4 magnetic sublevels). The hyperfine interaction is a type of dipole-dipole interaction, which primarily results from the contact term (overlap between the wavefunctions of electron and nucleus). The further splitting of levels is a result of the Zeeman interaction with an external magnetic field. Once the spectroscopy

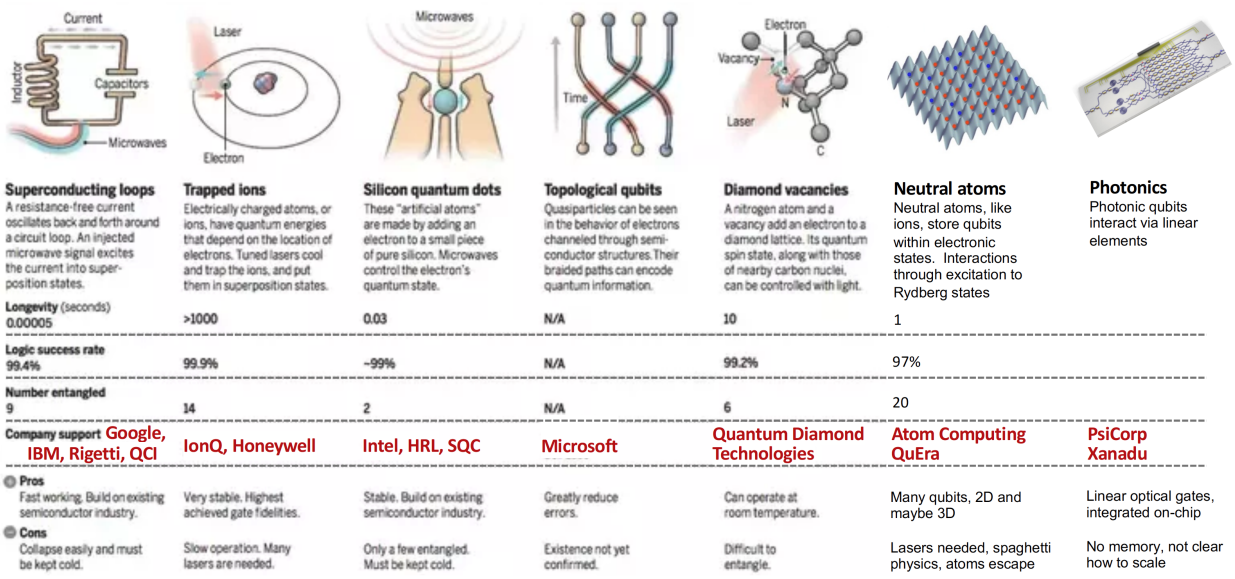


Figure 9: High-level overview of some leading quantum computing systems, adapted from Monroe et al, *Science* (December 2016).

of the individual atom is known, we generally choose a qubit consisting of two "spin" states which have the best coherence between them (for example, two states in the electronic ground state with different F or m_f, which can be manipulated using e.g. radio-frequency fields).

The main tool for manipulation of atoms and qubits therein is the laser. These can be used to implement coherent rotations of qubits, perform cooling and trapping of the atoms themselves, as well as preparation in well-defined states via optical puumping.

Examples:

1. Optical pumping. The atom starts at room temperature, which means that it is with equal probability in all of the ground states. By using laser light with a circular polarization σ^+, which always increases the magnetic quantum number by $+1$, we can ensure that that on average, the magnetic quantum number of the atoms is increasing under the laser field. The atom will continually be excited to a state with $+1$ angular momentum, and decay probabilistically to a state with either ± 1 or the same angular momentum quantum number. After a few excitation and spontaneous emission cycles, the atom is stochastically polarized into the level which is no longer coupled to the laser field, since all of the sublevels that were coupled to the laser field were "optically pumped." Once polarized with high probability, we have effectively cooled the internal degrees of freedom of the atom to very low temperature ($T \to 0$). Where does the entropy go?

2. The physics of atomic motion in laser light can be understood in a similar spirit. Intuitively, the absorption and emission of a photon changes the atomic momentum by $\pm\hbar k$. Mathematically, we can describe the emission of a photon using the Pauli lowering operator between the ground $|g\rangle$ and excited $|e\rangle$, $\sigma^- = |g\rangle\langle e|$. The running wave can be descibed in space as $\Omega(x) = \Omega_0 e^{ikx}$. The Hamiltonian that arises will be given by
$$H = \hbar\Omega_0 e^{ikx}\sigma^+ + \hbar\Omega_0 e^{-ikx}\sigma^-, \tag{229}$$
where the two terms correspond to stimulated absorption or emission of a photon from the classical laser

field. By rewriting the exponential in terms of a quantized position operator $\hat{x}$

$$\begin{aligned} e^{ik\hat{x}} &= \int dpdp' \, |p\rangle \langle p|e^{ikx'}|p'\rangle \langle p'| \\ &= dp \, |p + \hbar k\rangle \langle p| \, , \end{aligned} \tag{230}$$

where we have recognized that $\langle p|e^{ikx'}|p'\rangle = \delta(p - p' + \hbar k)$. So we can rewrite the Hamiltonian very simply as

$$H = \hbar\Omega_0 \int dp \, |p + \hbar k\rangle \langle p| \, \sigma^+ + \text{h.c.} \tag{231}$$

where h.c. denotes the Hermitian conjugate. Hence, an atom can transition between $|p\rangle |g\rangle \leftrightarrow |p + \hbar k\rangle |e\rangle$.

Remarks

- When this process is accompanied by spontaneous emission, which results in irreversible momentum transfer, we can achieve either cooling or heating. Depending on the configuration of the laser field and the atoms, this is the basis for so-called *Doppler cooling*.
- In the off-resonant case, we can think of the interaction in terms of dipole forces. Recall that we will have an effective Hamiltonian (obtained via perturbation theory in small parameter Ω_0/Δ, where Δ is our detuning from the optical transition

$$H_{\text{eff}} = \frac{|\Omega|^2}{\Delta} |g\rangle \langle g| \tag{232}$$

  where Ω depends on the position of the atom! This creates a potential and can in principle be used to trap the atom at the location of the optical field maximum, for example in the central focal point of a tightly focused "tweezer" beam.

Remark: simple harmonic oscillators (SHO)

The simple harmonic oscillator is a good approximaton for the motion of trapped ions and atoms, assuming the atoms are cold and exploring the bottom of their trapping potential (which is usually quadratic, to leading order). Additionally, the SHO captures the quantum description of electromagnetic fields, down to the level of single photons.

Recall the conventional Hamiltonian for the SHO

$$H = \frac{p^2}{2m} + k\frac{x^2}{2}. \tag{233}$$

We can quantize this Hamiltonian by introducing quantum mechanical operators $p \to \hat{p}$ and $x \to \hat{x}$ which do not commute, such that $[\hat{x}, \hat{p}] = i\hbar$. More conveniently, we can write superpositions of position and momentum in terms of the so-called annihilation and creation operators

$$\begin{aligned} \hat{a} &= \frac{1}{\sqrt{2m\hbar\nu}}(m\nu\hat{x} + i\hat{p}) \\ \hat{a^\dagger} &= \frac{1}{\sqrt{2m\hbar\nu}}(m\nu\hat{x} - i\hat{p}) \end{aligned} \tag{234}$$

where we can clearly write

$$\begin{aligned} \hat{x} &= \sqrt{\frac{\hbar}{2m\nu}}(a^\dagger + a) \\ \hat{p} &= \sqrt{\frac{\hbar}{2m\nu}}\frac{a^\dagger - a}{i} \end{aligned} \tag{235}$$

X IN TERMS OF A, Adag

Now we can rewrite the Hamiltonian in terms of the creation and annihilation operators, instructively as

$$H = (a^\dagger a + \frac{1}{2})\hbar\nu, \tag{236}$$

which clearly has eigenstates of the form $a^\dagger a \left|n\right\rangle = \left|n\right\rangle$, where $\left|n\right\rangle$ are known as *Fock states*. This Hamiltonian has a ground state ($n = 1$) with nonzero energy $E_1 = \frac{\hbar\nu}{2}$, and a ladder of *harmonically spaced* energy levels separated by $\hbar\nu$. This is clearly different from the case of a qubit, since if we try to drive one transition, we drive all transitions because of their equal energy spacing. So it is clear that we cannot simply treat this system the way we treat a qubit, under any approximation.

The purpose of the creation and annihilation operators are also clear in this picture. If we apply them to the Fock states, we see that they "raise" and "lower" the Fock state respectively by exactly 1 quanta:

$$\begin{aligned} a\left|n\right\rangle &= \sqrt{n}\left|n-1\right\rangle \\ a^\dagger\left|n\right\rangle &= \sqrt{n+1}\left|n+1\right\rangle \end{aligned} \tag{237}$$

Note that these prefactors are important since, application of $a^\dagger a$ should obey the eigenvalue equation $a^\dagger a\left|n\right\rangle = \left|n\right\rangle$.

4.2 Trapped Ion Quantum Computer

This is an important sytstem to study, as it was one of the first systems that convinced the field that quantum computers could actually be built in the laboratory. Additionally, it remains at the frontier of the field of quantum computing even today.

This model of quantum computing relies on charged particles trapped in vacuum. Since the particle is charged, we can use electric fields to confine it. We would like to engineer an electrostatic potential which effectively pushes the ion from all directions to a single stable point, leading to confinement in three dimensions.

However, basic electromagnetic theory tells us that this is not straightforwardly possible! This is impossible because of Gauss's law

$$\delta\dot{E} = 0, \tag{238}$$

where we see that if we make a sphere around our ideal configuration, we have violated (238). The best that we can hope for in reality is confinement in two dimensions, and anticonfinement in three dimensions (i.e. a saddle point in potential space). We know however, that the particle will be unstable with respect to perturbations in this third direction.

The solution is in fact not to use a static potential, but a dynamic one. By rotating the saddle, we can obtain an effective time averaged potential that confines the ion in all directions as desired. This is known as a *Paul trap*. In order to undrstand this, consider a classical 1D oscillator described by

$$\begin{aligned} \ddot{x} = (k^2\cos\Omega t)x = 0 \\ k^2 = \frac{eV_0}{md^2}, \end{aligned} \tag{239}$$

where e is the ion charge, m is the ion mass, V_0 is the RF voltage amplitude, and d is the trap-size. We can recognize (239) as a *Matheiu equation*, where $x(t)$ is bounded for $k \ll \Omega$.

The trajectory contains *secular* oscillations at the trap-frequency $\omega_{\text{trap}} \approx \frac{k^2}{\Omega}$, typically in the MHz frequency range, as well as *micromotion* oscillations at fast frequnecies Ω in the 100 MHz range. The ion feels an averaged potential, and assuming the RF frequency Ω is high enough, this enables confinement of the ion in all three dimensions

Many modern ion traps use the configuration consisting of four rods, with positive voltage applied to two opposing rods, and negative on the alternate two. Using RF voltages, this leads to confinement in the plane of the rods. The two ring electrodes are used for confinement along the direction of the rods. Ths is typically used to generate a chain of ions along the axis of the rods.

The result of these techniques is the trapping of single ions in a confined potential, with relatively high overall trapping depth, even larger than room temperature. If we can cool the ion towards the bottom of the potential, to a very good approximation, this leads to simple-harmonic oscillator physics describing its motion:

$$H_T = \frac{p^2}{2M} + \frac{M}{2}\nu_T^2 x^2 \equiv \hbar\nu_T(a^\dagger a + \frac{1}{2}), \tag{240}$$

where the effective trapping frequency ν_T after all averaging over the RF trap is usually a few MHz (compared to the RF frequency of the order 100 MHz). Quantized motional excitations in this trap are known as *phonons*.

Lasers are now used to manipulate the motion of the ion in this trap, in a way that is different from the free-space case. The intuition for this is that the trap can now absorb the momentum kicks from the light, so momentum does not need to be conserved, leading to inelastic atom-photon collisions. As such, light can be directly used to modify n, the number of phonons (motional excitations) in the trap. Mathematicallly, application of the laser results in the atom-field interaction Hamiltonian

$$H_{A-F} = \hbar\Omega_0\sigma^+ e^{ik\hat{x}} + \text{h.c.}, \tag{241}$$

where we can rewrite the position operator as

$$\hat{x} = \sqrt{\frac{\hbar}{2M\nu_T}}(a^\dagger + a) = a_0(a^\dagger + a), \tag{242}$$

where a_0 is the ground state position of the simple harmonic oscillator. From (242), it is clear that by applying the laser field we can also change the positional state of the ion by either adding or subtracting phonons.

In the special case where the ion is strongly confined, such that $a_0 k \ll 1$, or $a_0 \ll \lambda$ (the zero-point motion is much smaller than the wavelength), we can expand the Hamiltonian in leading order

$$\begin{aligned} H_{A-F} &= \hbar\Omega_0\sigma^+ + \hbar\Omega_0 a_0 ki(a + a^\dagger)\sigma^+ + ... + \text{h.c.} \\ &\equiv H_{\text{Rabi}} + H_1 + H_2. \end{aligned} \tag{243}$$

This is known as the *Lamb-Dicke limit*. The first term just corresponds to Rabi oscillations (elastic scattering) between the ground and excited state of the atom, without affecting the motional state of the ion.

The other terms in (243) correspond to the case where the atom interacts with the field and also changes its internal motional state. Taking the terms shown proportional to $\sigma^\dagger$: these correspond to excitation of the atom, either with absorbtion of an additional phonon ($a^\dagger$), known as the *blue sideband*, or removal of a phonon (a), known as the *red sideband*.

Remark: physics of simple harmonic oscillator coupled to qubit

1. Consider the reduced Hamiltonian from (243) in the form

$$H_1 = \hbar g(a\sigma^\dagger + a^\dagger\sigma). \tag{244}$$

   This Hamiltonian corresponds to quantum excitation exchange between the qubit and the simple harmonic oscillator.

   Here, we can consider some detuning ΔE between the oscillator and the qubit, such that $\Delta E = E_e - E_g - \hbar\nu$, for all n. Manipulating this detuning will allow for direct removal of a phonon by laser driving. Furthermore, we see that the coupling rate will actually grow with n as $\langle n-1|a|n\rangle \sim \sqrt{n}$! The simple ground state $|g\rangle|0\rangle$ is clearly an eigenstate. This will be crucial to the idea of cooling to the ground state. However, due to the interaction in (244), the higher levels containing product state of qubit and oscillator are no longer eigenstates, for strong enough coupling g.

   As an aside, the atom-*photon* excitation exchange (as opposed to vibrational *phonons*) is known as the Jaynes-Cummings model of cavity quantum electrodynamics, and is a crucial component in several primitives of quantum information processing.

2. Consider the second term

$$H_2 = \hbar g(a^\sigma + a^\dagger\sigma^\dagger), \tag{245}$$

   in which excitations are created in pairs. This, is known as *parametric amplification* and can be used to generate entanglement between excitations within the Harmonic oscillator (i.e. quantum states of motion, also known as *squeezed* states.

With this in mind, one can now consider spontaneous emission which has not been included in the Hamiltonian treatment above. In general, this has been bad for quantum manipulation, but we have seen that it is crucial for cooling and initial state preparation. If we consider the trap frequencies $\nu_T \sim 100$ kHz - 10 MHz, we see that we require extremely low temperatures, necessitating such cooling schemes.

Considering that the excited atom emits in three-dimensions and a particular excited state $|e\rangle|n\rangle$ decays with the following pathways

$$\begin{aligned} &\to |g\rangle|n\rangle \ \text{rate}\gamma \\ &\to |g\rangle|n+1\rangle \ \text{rate}\gamma(ka_0)^2 \\ &\to |g\rangle|n-1\rangle \ \text{rate}\gamma(ka_0)^2, \end{aligned} \tag{246}$$

showing that the inelastic transitions will be suppressed in spontaneous emission in the Lamb-Dicke limit.

We can now perform laser-cooling of ions by exciting with a red detuned laser, effectively driving the system from the state $|g\rangle|n\rangle \to |g\rangle|n-1\rangle$, reducing the motional quantum number by 1. Now, when we spontaneously decay, we know we will most likely decay to the state $|g\rangle|n-1\rangle$, leaving us in a lower occupation state than we started with! This process will continue down the Harmonic ladder until we are in $|g\rangle|0\rangle$, at which point our laser will be detuned from all transitions.

This is the idea of so-called phonon-sideband cooling. Photon absorbtion leads to transitions of the form

$$|g\rangle|n\rangle \to |e\rangle|n-1\rangle, \tag{247}$$

whereas spontaneous emission predominantly leads to transitions of the form

$$|e\rangle|n-1\rangle \to |g\rangle|n-1\rangle. \tag{248}$$

Remarks

1. Cooling is most efficient when $\gamma \ll \nu_T$, which is known as the sideband resolved regime, since the linewidth of the transition given by γ is smaller than the separation between the sideband and the carrier (direct) transition. In the other limit, $\gamma \gg \nu_T$, cooling can still be performed, and this is known as Doppler cooling as discussed before.

2. The maximum rate of cooling is proportional to $\gamma\hbar\nu_T$. This goes back to the question of where the entropy goes in this process. The answer is that spontaneous emission carries the entropy away with the emitted photon. This is crucial: we need a combination of coherent laser driving *and dissipation* via spontaneous emission in order to perform cooling.

We can also use light to perform measurement of ion internal states. If we consider two ground states of the ion which can be used as a qubit, We typically have a *cycling* transition associated with one of the states $|1\rangle, such that application of the laser excites between |1\rangle$ and $|e\rangle$ in a cyclic fashion (the population stays within this manifold). By applying this laser and detecting some fraction of the emitted photons, we can determine that we were in the state $ket1$, since $|0\rangle$ would not have resulted in a scattered photon. Thus, in ideal conditions, detection of a photon corresponds to a projective measurement of the qubit in the state $|1\rangle$. Realistically, we can plot histograms of the number of photons detected on average from the two qubit states, and as long as they are well resolved, we can perform effective projective measurements.

Next, we can perform qubit operations on the long-lived hyperfine ground states (often known as "spin states"). This can be done with radio-frequency illumination to drive Rabi oscillations. However, these rotations are often done using so-called *Raman transitions*, which are two-photon transitions using an intermediate excited state. By driving the two transitions simultaneously with Rabi frequencies Ω_1, Ω_2, with a large detuning from the excited state Δ, we can obtain effective single-qubit rotations at a Rabi frequency $\Omega_R \sim \frac{\Omega_1\Omega_2}{\Delta}$. It turns out the coherence properties of ions prepared in superpositions in such a fashion are very good, now on the order of 10 seconds.

Next, in order to perform two-qubit operations, we can rely on the Coulomb interaction between the charged ions trapped in a chain. This is done by coupling the spin state to the common motional mode of ion vibrations.

By trapping many ions, we see they will arrange with a particular spacing due to the repulsive potential between ions. Now, ions will vibrate about their equilbrium positions in a way that is coupled to one another (since each ion feels the presence of the other ions via the Coulomb interaction). This can be treated as a system of coupled oscillators - we can solve this similarly to the classical case by finding the normal modes.

We consider the Hamiltonian consisting of N coupled Harmonic oscillators

$$H = \frac{1}{2}\sum_{i=1}^{N} M\nu_i^2 + \nu_T^2 M x_i^2 + \sum_{i,j} \frac{e^2}{x_i - x_j}. \tag{249}$$

This describes N coupled simple harmonic oscillators near equilibrium, and has N normal modes. The lowest energy mode consists of center of mass motion of all ions

$$x_{CM} = \sum_{i=1}^{N} \frac{x_i}{N}, \tag{250}$$

with frequency $\nu = \nu_T$ and effective mass $M \times N$. The next lowest mode will have a different ferquency $\sqrt{3}\nu_T$, and so on.

Thus, because each mode has a different frequency, motional modes can selectively be excited with a laser independently. For example, with $N = 2$, we can show that we have modes

$$x_{CM} = \frac{x_1 + x_2}{2} \tag{251}$$

for the center of mass, and

$$x_{rel} = x_1 - x_2, \tag{252}$$

corresponding to relative motion. This can be captured in a diagonalized Hamiltonian

$$H = \frac{1}{2}(M_{CM}\nu_{CM}^2 + \nu_T^2 M_{CM} x_{CM}^2) + \frac{1}{2}(\mathcal{M}\nu_{rel}^2 + \nu_T^2 \mathcal{M} x_r el^2), \tag{253}$$

where $\mathcal{M}$ can be shown to be $\sqrt{3}M$.

One challenge is that for larger and larger ion chains, the overall trapping frequency needs to be reduced in order to keep the system stable, in other words, to be able to resolve the modes from one another, which becomes more and more challenging and complex as we add ions and create additional motional modes.

4.2.1 Approach to two-qubit operations

We have seen before that the ion-phonon interaction is:

$$H_{af} = \hbar\Omega e^{ikx_i}\sigma_i^+ + h.c. \tag{254}$$

We can write the position as:

$$x_i = x_{CM} + rest \tag{255}$$

If the laser is tuned into resonance with the COM motion, the rest can be ignored. The leading order term is then:

$$H_{af}^i = H_{Rabi} + \Omega\frac{ka_0}{\sqrt{N}}(a_{CM} + a_{CM}^\dagger)\sigma_i^+ + h.c. \tag{256}$$

The Mossbauer effect which yields the $1/\sqrt{N}$ dependence corresponds to the fact that if we have a big object with many degrees of freedom and with mass, the entire object absorbs the momentum. Since the photon couples to the momentum of the entire chain, we have a collective effect which can be used to generate entanglement between the ions.

For example, if the laser is tuned to $|n\rangle|g\rangle \to |n-1\rangle|e\rangle$, the effective dyanmics is:

$$H_{JC} = \hbar\Omega\frac{ka_0}{\sqrt{N}}\sigma_i^+ a_{CM} + h.c. \tag{257}$$

This is equivalent to a TLS interacting with a cavity photon field, and allows for the quantum state of the ions to be exchanged with intertial (qubit) states and the motional (phonon) state of the ions! In practice this can be done with two-photon stimulated Raman transitions (add pic from lecture notes). If the two photon transition is detuned by the trapping frequency, then the ion will exchange internal and motional states. In particular, for a single ion, we have:

$$(\alpha |0\rangle + \beta |1\rangle) \otimes |0\rangle_m \rightarrow |0\rangle \otimes (\alpha |0\rangle_m + \beta |1\rangle_m) \tag{258}$$

If there are more than one ions, then this motional degree of freedom includes all of the ions. The idea is to exchange motional excitations between ions, which come from their internal degrees of freedom, *using the ionic motion as a data bus*. For further reading see Cirac and Zoller, PRL **74**, 4091, 1995.

State of the art and challenges

- The ion control has been perfected to a high degree. The gate fidelities are better than 99.9%, way ahead of every other system.
- However, scaling the control to a larger system is challenging. The ions heat up due to the coupling to the environment, and at large N thre is a large number of modes which can be difficult and slow and isolate a single one.
- Modern approaches include off-resonant motional excitation, which moves the ion conditionally on which state it is in gently. For example, if it only pushes the state $|1\rangle$, then it creates a dipole between two ions. This approach is robust since you don't need the ion to be perfectly in the ground state.
- Geometric gates are also used (see Molmer-Sorensen gate) which are also relatively insensitive to the motional states.
- 1D traps have achieved up to 50 qubits, but 2D traps is challenging. There has been some effort for a decade or so to extend to 2D, but this is also challenging - there are more degrees of motional freedom.
- Another approach is to cool the traps cryogenically to reach the ground state, but that can also be technically challenging.
- Some approaches for scaling up involve a 'quantum CCD' which includes storage zones and interaction zones, and shuttling the ions around between these zones. The challenges of this is that moving the ions tends to heat them, especially around the corners.
- Another approach involves entanglement by measurement - using photons and measurement to carve out entangled states of multiple ions, and using other ions as registers.

For further reading, see D. Liebfried et. al., RMP 75 281 (2003); R. Blatt & D. Wineland, Nature **453** 2008; I. Cirac & P. Zoller, Physics Today, March 2008.

4.3 Neutral atom quantum computer

Neutral atoms are like ions except they do not have a net charge. Why would we want to use them? They are extremely coherent in terms of internal degrees of freedom (information can be stored for a long amount of time - the best atomic clocks use neutral atoms), and there are lots of neutral atoms around, so it is relatively easy to create identical arrays of ions. The key challenges is that it is hard to isolate and control individually, and the interactions are weak (which are needed for multi-qubit gates). The approach is as follows:

- Optical tweezers are traps that use focused laser beams or standing waves. There is a point where the electric field is maximized at the center. The light, if detuned properly, will induce an electric dipole moment which pushes the atom to the point of maximum intensity. The dipole potential will be $|\Omega|^2/\Delta$. In this way, an atom can be trapped and held for a reasonably long time in vacuum.
- Additionally, we can start with an array of tweezer traps, and load them. However, the trap frequencies are only 10-100 MHz, which means that the atoms must be cooled before loading (unlike the case with ions).

- For interactions, we have two possibilities: (1) bring the atoms on top of each other to create 'cold collisions', but this is a slow, delicate process; and (2) excite atoms into 'Rydberg states'. These are highly excited states with principle number $n >> 1$. When this number is high, the size of the atom becomes large, on the order of microns. This induces dipole-dipole interactions which scale as n^4, but the Van de Waals interaction (a result of virtual excitations in a dipole-dipole interaction) scales as n^{11}, which is quite a large number. These Rydberg states have quite a long lifetime, since they cannot undergo spontaneous emission easily, and can last for over 100μs. This combination allows one to do very fast, high fidelity quantum operations.

- However, neutral atom trapping is a bit more delicate than trapped ions - the traps are relatively shallow, and the limitation is given by the vacuum (collisions with background atoms). Depending on how well the vacuum is created, the atom lasts in the trap for 10s - 10s of minutes. The solution is to reload the atoms often.

- In general, different atomic states will have different trapping potentials, since the interaction of the laser with the atom depends on its state. For example, the excited state in an optical tweezer is an anti-trapped state. This kind of trapping would generally entangle the motional and atomic degreees of freedom, which is a source of decoherence. The solutions for this involve 'magic wavelength trapping', frequencies such that the potential is the same for all states, as well as making the gates very fast, so that the atoms do not have time to move - the time spent in the excited state is very small.

In principle, we can directly use the Rydberg interaction to implement e.g. a control-phase gate. However, there is a more refined technique for implementing gates based on Rydberg interactions known as *Rydberg Blockade*. If atoms are far apart, and we excite the Rydberg transition resonantly, they will simply undergo Rabi oscillations individually. However, as they come closer together, at some point, the energy shift $U_{VDW} \sim \frac{n^1 1}{r^6}$ will become very large, and the transition will not be in resonance anymore for both atoms. We can still excite one of the atoms, but not both because of the significant additional energy cost of the Rydberg interaction, which leads to a shifted energy of the doubly excited level.

Thus, the simultaneous excitation of two atoms will be blockaded at a particular distance on the order of the *blockade radius* R_b few μm, with variations between 1 and 100μm by careful choice of Rydberg state and the Rabi frequency of Rydberg excitation lasers. One important aspect is that once the atoms are closer than the blockade radius, their exact separation is unimportant, since the blockade will be in effect. This means that high fidelity operations can be implemented in a way that is relatively insensitive to exact atomic position and their motional state

The atoms in the Rydberg states are highly excited, meaning they can undergo decay process. However, the timescale for this decay is usually quite long (several hundred μs), and operations can typically be performed much more rapidly. Additionally, typical trapping frequencies for atoms are in the MHz or even sub-MHz range. If we can excite and de-excite atoms very quickly, the presence of atoms in Rydberg levels does not actually impact the trapping of the atom (even though Rydberg states are typically *anti-trapped* under the typical trapping fields!). By simply turning off the traps during Rydberg interaction, and turning the trap back on once complete, we cna continue to trap the atoms, since they do not have time to fly out of the trap.

Practical Approaches

In principle, one can choose any atom (or even a molecule)! However, we generally want to choose atoms with relatively simple structure, such as Hydrogen-like atoms (Alkali atoms) such as Rb, Na, Cs, which are quite straightforard to control. At the next level of complexity, Alkaline-earth atoms such as Sr and Yb can be used.

Even after the atom is chosen, there are usually multiple choices of qubit encoding. The most straightforward approach is to use $|g\rangle$ and $|r\rangle$ as computational basis states, and manipulations are quite straightforward in general. Although note here, we cannot necessarily access the full 2^n dimensional subspace due to the blockade. Another approach is to use two ground levels split by the hyperfine interaction, with a laser carefully selected to only excite one of the qubit levels to $|r\rangle$. This has the advantage that the ground states can have excellent coherence properties. A final option is to even use various different Rydberg atoms (e.g. of the s and p character).

As an example of control and entanglement, we can consider $|g\rangle$ and $|r\rangle$ as the qubit basis states. The idea is to use the Blockade to prepare a Bell state of two atoms in nearby tweezers (sitting within R_b). If the atoms start in $|gg\rangle$, and we drive the atoms to $|r\rangle$, we will be able to excite at most 1 atom. Now we will undergo oscillations between $|gg\rangle$ and a state where only one atom is in the Rydberg state. However, since we cannot fundamentally

determine *which* of the atoms is excited, we effectively drive $|gg\rangle |W\rangle \propto |gr\rangle + |rg\rangle$, where the phase is set by the relative phase of the laser between the two atoms.

This is partially evident in the oscillations in the probability to detect 0 and 1 atom in the Rydberg state. However, such oscillations are not sufficient to tell whether or not we have created an entangled state. In order to show that we have entanglement, we need to be able to measure the relative phase between the components, not just the overall populations. This is equivalent to distinguishing $|\Psi^+\rangle$ from $|\Psi^-\rangle$. In order to do this in the experiment, we can add a differential phase shift on one atom relative to the other by application of an additional laser field (via the AC stark effect). If this phase is applied when the atoms are in the entangled state, we transform to a state like $|ge\rangle + e^{i\phi}|eg\rangle$. Now, for example if we are in the state $|D\rangle \propto |ge\rangle - |eg\rangle$, the laser will no longer have the correct phase to de-excite back to $|gg\rangle$. And thus, oscillations in the signal with respect to $|\phi\rangle$ directly probe this phase, and allow one to extract entanglement fidelity.

This procedure creates the entangled state desired, but does not show us how to implement universal gates on this two-qubit register. This can be seen by the fact that $|rr\rangle$ cannot be accessed by such excitation. The solution to this is to change the qubit basis. The basis $|g\rangle, |r\rangle$ is not ideal for a few reasons. First, it has a finite lifetime $\sim 200\mu$s for the Rubidium 70S level, for example. Additionally, Rydberg atoms are not trapped, leading to some loss. And finally, as already mentioned, we cannot access the full Hilbert space.

Instead, if we use qubits based on hyperfine spin states, we can solve these problems. Now, we can use microwave fields to fully manipulate the ground state qubit, which has excellent coherence times. Then, to perform multi-qubit gates we need to perform atom specific light-shifts via the Rydberg levels. One natural approach is to perform a CZ-type gate of the form

$$\begin{aligned} |00\rangle &\to |00\rangle \\ |01\rangle &\to |01\rangle e^{i\phi} \\ |10\rangle &\to |10\rangle e^{i\phi} \\ |11\rangle &\to |11\rangle e^{i(2\phi-\pi)}. \end{aligned} \tag{259}$$

Here, by choosing now a non-zero detuning for the Rabi excitation pulse, we can pick the detuning and pulse duration such that the atom undergoes a full oscillation on the Bloch sphere (not necessarily by passing through $|W\rangle$, since we now have some detuning. Now, the area enclosed by the trajectory will be the phase acquired by the state. By choosing Δ appropriately, we can implement (259). By combining this technique with individual light-shifts to impart $sigma_z$ rotations, one can perform quantum logic operations on small registers. Furthermore, the same idea can be used on three qubits within R_b simultaneously to implement three-qubit gates directly with just three atoms - no need for auxiliary qubits. This is helpful, since it allows one to directly implement complicated multi-qubit unitaries without overhead, related to the idea of $C^N Z$ gates discussed earlier. Finally, these operations can be performed on several sets of atoms in parallel. For details see H. Levine et al, *PRL* (2019).

Quantum simulations

The beauty of the neutral atom approach is that we can control many atoms in parallel. This is a nice setting for applications involving quantum simulation. Historically, this idea is the origin of the field of quantum information. Richard Feynman pointed out that simulating quantum systems is hard, and that we should use quantum hardware with programmable interactions in order to simulate the physics.

This can be done in the neutral atom case by assembling atoms in configurations equivalent to effective-spin systems, and simulate the dynamics. We can understand this by looking at the many-body Hamiltonian given by

$$H = \sum_i \frac{\hbar\Omega}{2}\sigma_x^i - \sum_i \hbar\Delta n_i + \sum_{i<j} V_{i,j} n_i n_j \tag{260}$$

for $n_i = 0, 1$ are Rydberg atoms on each cite i. This is very faimilar to the Ising Hamiltonian, except for this term adding an energy $V_{i,j}$. Note that we can effectively replace $n = \frac{\mathbb{1}+Z}{2}$. The middle term, proportional to Δ, the laser detuning, can determine whether all atoms are in the ground state or in the excited Rydberg state, in the absence of other terms. If we now add itneraction such that we have nearest-neighbor blockade, we can only access states with so-called Z_2 ordering, with every other atom excited. By increasing interaction range, we see differently ordered state Z_3, Z_4, etc (every third, every fourth atom excited, etc). Here Z_i refers to the broken symmetry in the configuration of atoms in the ground and Rydberg states.

To explore these phases, you could start with all atoms in the ground state, and slowly change the detuning to enter the relevant phase adiabatically. By doing this for various interaction ranges, one can map the relevant phase

diagram. The important concept is that the system undergoes a *phase transition*. This is special, since we are dealing with an isoltaed system with effective temperature of zero. This means that the phase transition is driven not by thermal fluctuations (as with typical statistical mechanical phase transitions), but rather by *quantum fluctuations*, making these *quantum phase transitions*. In the intermediate detunings, the system struggles to decide which phase to be in. This type of experimental system can probe snapshots of the physics of these quantum phase transitions.

This can now be done in larger arrays. Once we are in the final phase, we typically see long, ordered domains, with *domain walls* or breaks in the individual domains. This is akin to a crystal boundary or dislocation. By looking at the domain wall density, one sees that the mean of the domain wall density varies smoothly. However, the variance of this domain walls is sharply peaked at the phase transition, and this peak signifies the quantum phase transition, as it is a result of the frustration of the system between the two phases. Finally, one can look at the probability for a given microstate to occur (e.g. the ideal phase with no domain walls). This is a sort of measure of how good the system is - not as quantitative as a fidelity - but as a measure of how probable we are to reach the exact ground state. For details, see H. Bernien et al, *Nature* (2017).

This is now a frontier of quantum computing and simulation. Extending these results to arrays of two-dimensional atoms with improved coherence using hyperfine qubits with individual atom control are the crucial next steps. By using a combination of a spatial light modulator and an acousto-optic deflector to make arbitrary two-dimensional patterns, 2D experiments can now load on the order of 300 sites. One can verify the analogous phase diagram for the 2D Ising model. The first phase, for example, will be the checkerboard phase, the first signatures of which can now be seen.

The two dimensional patterns of atoms may also enable implementation of QAOA in a regime that cannot be compared classically otherwise. This is closely related to the idea of co-design, using a specific hardware system to efficiently encode a problem. In this case, the problem is to find the maximum independent set (MIS) of vertices on a graph such that no elements of the set are neighboring on the graph. An example of this, suited to the Rydberg array, is the MIS on a unit disk graph where vertices are connected if they are within a given distance. This is still an important problem in designing networks, and is still an NP complete problem. The graph now corresponds to the ground state of the system, since we are trying to excite as many Rydberg atoms as possible, given the constraint of the Rydberg blockade. Thus, by finding the ground state, e.g. by QAOA, one can solve this problem. Of course, QAOA is a heuristic algorithm, so there is no guarantee that it will work well. Now, the challenge is to benchmark algorithms, such as QAOA, against the best classical algorithms.

Outlook: There are now a handful of atom-array experiments with different flavors. Simultaneous, individual control corresponds to having large numbers of independent laser beams with rapid modulation control is a major challenge. Increasing both atom number and circuit depth is also a formidable challenge, since individual errors on the circuit or atomic level become exponentially important as the system size is scaled up. To solve some of these problems, techniques such as Rydberg atom trapping and control of spontaneous emission may be necessary. Additionally, making *fast* measurements is key to atomic and ionic approaches to quantum computing. These emeasurements are currently the slowest part of the computation, so this limits the amount of time required to perform an algorithm. Finally, identifying the useful algorithms and approaches is an open question. Building error-correct systems that can perform arbitrary quantum algorithms is also an outstanding goal.

4.4 Superconducting quantum computer

Our description of superconducting qubits goes back to harmonic oscillators, this time of a circuit made of a capacitor and an inductor. LC circuits have a resonance frequency at $\nu = 1/\sqrt{LC}$, coming from the differential equation describing the charge. The energy stored in the circuit is $E = \frac{1}{2}LI^2 + \frac{1}{2}\frac{Q^2}{C}$. We consider the flux through the conductor $\Phi = LI$ as a dynamical variable such that $E = \frac{1}{2}\frac{\Phi^2}{L} + \frac{1}{2}\frac{Q^2}{C}$. We can therefore express flux and charge in terms of creation and annihilation operators, with zero point motion $Q_{zpf} = \frac{\hbar}{2Z}$, where $Z = \sqrt{\frac{L}{C}}$ is the 'impedance' which has units of resistance. This almost comes out to be 50-100 ohms. Comparing this Hamiltonian with harmonic oscillators, we can define operators Φ, Q as conjugate variables analogous to x and p such that $[\Phi, Q] = -i\hbar$ and $H = \hbar\nu(a^{dagger}a + 1/2)$, where $QiQ_{zpf}(a - a^\dagger)$ and $\Phi = \phi_{zpf}(a + a^\dagger)$. Typical parameters involve $\nu = 10$ GHz and $kT << \hbar\nu$, $V_{zpf} \sim \mu V$ and $I_{zpf} \sim nA$. Two problems emerge: 1. the resistance introduces loss, and the LC circuit is harmonic (linear).

The approach to solve these problems is:

- Use superconducting elements that have $R \to 0$ when the temperature $T < T_c$.
- To introduce a non-linearity we use the Josephson effect, which acts as a nonlinear inductor.

What are each of these things physically? An example is to use a superconducting (sc) island separated from ground (also sc) via a thin insulating film. The island has charge Q. Provided that the insulator is thin enough, what happens is that there is a non-negligible probability for charge from the superconductor to tunnel across the potential barrier. Classically, the barrier has infinite resistance, but there will be some probability of tunneling. The physics is the following. The fundamental reason why sc have zero resistance is the energy gap. Inside sc, electrons which are usually fermions effectively attract each other and create cooper pairs, at least at long range. They form these objects from two fermions (2e), but the new pairs are actually bosons, such that they can condense in the ground state. These cooper pairs occupy energy states which are gapped from the rest of the spectrum, where the gap is related to the binding energy of the electrons in the cooper pairs. Only finite energy is required to break these cooper pairs (corresponding to T_c), but below that, they can condense into this state and behave like a condensate. In this state they can undergo motion without friction at all, namely super-fluidity, and that is what gives rise to superconductivity. These electrons travel in pairs, such that when charge tunnels through the barrier it happens in pairs. As long as the gap is much larger than the thermal energy, only the bottom state can be occupied on both sides of the insulator, such that the pairs can tunnel across the barrier *coherently*. The tunneling Hamiltonian is given by:

$$H_T = -\frac{1}{2}E_J \sum_m |m\rangle \langle m+1| + h.c. \tag{261}$$

where $n = N + m$ and n is the number of cooper pairs in the island. Note that the eigenstates are not Fock states $|n\rangle$! The eigenstates are $|\phi\rangle = \sum_{k=-\infty}^{+\infty} e^{ik\phi} |k\rangle$, namely:

$$H |\phi\rangle = -E_j \cos\phi |\phi\rangle \tag{262}$$

and the physics is that ϕ is the phase of the sc wave function on the island. The cooper pairs all occupy a macroscopic quantum state which has a phase and amplitude, and we have written ϕ as the phase here. In particular:

$$n |\phi\rangle = -i\frac{d}{d\phi} |\phi\rangle \tag{263}$$

such that in some way, n the number of particles acts as the momentum degree of freedom for the position, and the Hamiltonian is:

$$H = \frac{Q^2}{2C} - E_j \cos\phi \tag{264}$$

When we write this in terms of the phase of the wave function, we see that the tunneling energy is a nonlinear function of the phase. The cosine is $\approx E_j\phi^2/2$ for small phase ϕ, and we recover the harmonic oscillator Hamiltonian. Following the canonical commutation procedure, we have:

$$\begin{aligned} H &= \frac{E_j}{2}\phi^2 + 8E_c\frac{n^2}{2} \\ [\phi, n] &= -i \\ H &= \hbar_j(a^\dagger a + 1) \\ \omega_j &= \sqrt{\frac{E_j 8E_c}{\hbar}} \\ \phi &= \phi_{zpf}(a + a^\dagger) \\ \phi_{zpf} &= \sqrt{\frac{2E_c}{E_j}} \end{aligned} \tag{265}$$

the an-harmonicity is $H_{ah} = -E_j\phi^4/4! = -\frac{E_c}{2\hbar}(a^\dagger a^\dagger aa + 2a^\dagger a)$. We call E_C the charging energy. The spectrum of the system is two lowest state separated by the plasma frequency ω_j which is a few GHz, and then the *third* excited state is at a separation of $\omega_j + E_C$, where the anharominc term $H_{ah} \sim E_C$. Fundamentally what gives rise to this term is the coulomb interaction. You can think about this problem as a linear oscillator + nonlinear Josephson Junction, OR we can say let's neglect all interactions and look at linear motion of the particle. Tunneling is in principle a non-interacting particle (linear) effect. If you neglect the Q^2 term, all we will have is linear motion of the quasi-particles. *Then* adding the charging term, there are corrections from the charging energy $\sim Q^2$ which introduce a non-linearity.

Remarks

1. The derivation above is kind of sloppy. In one way, ϕ is NOT an operator, but $e^i\phi$ is. The phase should be periodic. The way we have derived things above loses the periodicity. We can do it the way we did above since we assumed that the phase fluctuations are small.

2. The effect of an external E field by applying a voltage across the gap is that we. must add another term to the Hamiltonian: $H \to H - V_0Q = 4e^2(n-n_0)^2/2C \equiv E_C(n-n0)^2$. The $n_0 = CV_0/2e$ is the equilibrium charge sensitive to V_0. This term is both good news and bad news. On one hand, we can use voltage to control the system, which is nice (we control classical computers with voltages!). However it turns out the fluctuations introduced by this term is parasitic. In practice, we would like to *minimize* the sensitivity of the qubit to charge noise by minimizing n_0 without losing nonlinearity.

3. We would like to work in the optimal regime where $E_C << E_J$ but finite E_C for nonlinearity. This is the idea of the *transmon* qubit charge island such that $E_C/E_J \sim 1/50$ or $1/100$.

4. The typical best coherence time is $T_1 \sim$ ms and $T_2 \sim 10\mu$s, limited by acharge in purities.

5. By adding a magnetic field, youc an efefectively control the background flux. In analogy with the simple harmonic oscillator, we can re-write the phase ϕ as the flux $\phi = \frac{2e}{\hbar}\Phi = 2\pi\frac{\Phi}{\Phi_0}$ and Φ_0 is the flux quantum. The physics of the eJJ is nonlinear inductance, such that adding a magnetic field controls the effective inductance L. Many variation of sc qubits exist and it is an ongoing field, but in all of them we use microwave fields to manipulate the qubits.

6. the superconductors are made out of niobium or aluminimum and the tunnel barrier is aluminum oxide (oxidized aluminum). The Schoelkopf group at Yale attached a large capacitance to one of the cooper pair boxes which acts as an antenna. The antenna is then used to control the qubits.

7. to operate the circuit quantum mechanically, the circuit must be cooled well below the critical temperature (typically to 0.01 K). Then the circuit is isolated from the environment by only connecting the circuit electrically to the microwave line. Then we are left with a nonlinear oscillator where the qubit can be encoded in $n = 0$ and $n = 1$ states.

The remaining issue, is how do you coule qubits and perform multi-qubit gates? The idea is to employ coupled oscillators. One qubit is coupled capacitively to a linear LC circuit which is used as a data bus. Then qubit 2 is coupled to the same linear LC circuit. The physics is that of two coupled oscillators such that $H_{Q-res} = V_g 2en$ wher V_g is the voltage on the coupling capacitor created by the LC circuit. The voltage acts as a driving voltage for the qubit, such that the two oscillators are coupled through normal electric potential Vq. Therefore, the quantum degrees of freedom of the linear circuit can be coupled to the qubit, when we write $V_g = V_{zpf}(b + b^\dagger)$. One popular linear resonator that people use is a stripline about 25 mm long with a 2μm gap. The excitations of the stripline are microwave photons inside the circuit, which are confined to a volume $\sim \lambda d^2$ where d is the gap in conductors of the waveguide and λ is the wavelength of the microwave field and the length of the waveguide.

the caivty and circuit QED systeem is dseecribed by a TLS and SHO. For $g << \nu$:

$$H = \hbar g(b + b^\dagger)(\sigma + \sigma^\dagger) \approx \hbar g(b\sigma^+ + b^\dagger\sigma_-) \tag{266}$$

such that we arrive at the same JC Hamiltonian as we saw before with trappeed ions. The physics of g is:

- Suppose that the qubit is a dipole moment d (the big antenna with the capacitor attached)

- The coupling g is the dipole moment $dE_0/\hbar$ where E_0 is the electric field of one microwave photon in hte resonator.
- The electric field from one photon is given by $\hbar\nu = \epsilon_0 \int |E_0|^2 dV = \epsilon_0 |E_0|^2 V$ where the volume V is the volume of the e/m field.
- therefore $g = d\sqrt{\frac{\hbar\nu}{\epsilon_0 V}}/\hbar$. The quantity g then decribes how much one single ephton in the resonator affects the qubit. The typical values are $g > 100MHz$.

To use the total circuit (circuit QED) we tune the resonator into resonace with teh qubit which allows for coherent exchange of excitations, and generate nonclassical states of MW radiation. However the most common way to use it is to introduce a detuning between teh resoantr and the qubit that is much larger than g. In this case, using second order perturbation theory we see that:

$$H_{eff} = \frac{\hbar\nu}{2} Z + \hbar\omega_r b^\dagger b + \hbar \frac{g^2}{\nu - \omega_r} b^\dagger b Z \tag{267}$$

We can think about this Hamiltonian in two ways:

1. The qubit has a refractive index which the resonator experiences, and can be used for the qubit readout.
2. There is a photon dependent shift in the qubit frequency. The application for that is a two-qubit gate. For example, for two qubits coupled to the resonator, we can map the state of qubit 1 to the resonator mode and then it will be sensed by qubit 2.

For qubit readout, there is a shift in the frequency of the microwave resonator, such that by detecting transmitted or reflected photons the qubit state can be read out. There are many groups in academia and industry pursuing these approaches. There are several challenges and opportunities:

- The lifetime and coherence times of the qubits is limited to 10-100μs for almost a decade because of charge defects in the JJ, as well as imperfections in the sc. This is because even at low temperature, sc still have some quasiparticles (people don't understand where it comes from and it's an area of current research).
- Using transmons that are 3D and much larger cavities have much longer coherence times than qubits, but they are hard to scale. Each qubit is a centimeter scale!
- All qubits are different - there is some inhomogeneity between qubits from fabrication imperfections, which dramatically increases the complexity of the control needed.

These motivate new approaches for scaling (for example, using bosonic degrees of freedom to store information), as well as new approaches to quantum error correction. For further references see S. Girvin's lecture notees and arXiv 1302.5842; M. Devoret & Shoelkopf Science 339, 1169.

In summary, for implementation of QC's there are contradictory requirements: isolation from the environment, but control of qubits and their interactions. The state of the art of gate fidlities is $F = 0.99 - 0.995$. For an N qubit machine, running a p-circuit depth algorithm, then you find that $\mathscr{F}^{NP} << 1$ for $NP \sim 100 - 1000$. There are two approaches to solve this problem: increase fidelities, or apply quantum error correction.

5 Quantum error correction

We have shown that most of the quantum channels we've discussed can be represented as random Pauli matrices acting on the qubits. These are the kind of errors we need to correct. In practice, there are two types of errors: (1) memory errors, where the quantum state is not maintained due to coupling to the environment which can be addressed with QEC, and (2) operation errors, which will be reduced to memory errors.

5.0.1 Classical error correction: review

One simple way to correct classical error is to use redundant coding. Suppose that we have a single classical bit we want to store for time t. However, a bit flip error occurs with probability p. The idea of EC is to encode one logical bit into several physical bits. For example, we can encode 1_L as 111 and 0_L as 000. Then, consider the bits 0_L after time t. The outcomes will be 0_L with probability $(1-p)^3$, but there will be one bit flipped with probability $3p(1-p)^2$, two bits are flipped with probability $3p^2(1-p)$ and three bits will be flipped with probability p^3. How can we correct this error? If the $p << 1$, then the errors will be dominated by a single bit flip nature. We can then measure all qubits after time t and if the majority of them are in 0, we assign 0_L and flip the bit in 1. This is called *majority voting*.

Remark

1. This procedure will work only if a single bit is flipped. For two or three errors, the procedure will fail.
2. The probability of a correct outcome is then $(1-p)^3 + 3p(1-p)^2$, such that we have removed the linear term in p, such that this procedure is helpful if $p < 1/2$. This makes sense: if $p > 1/2$, most of the time two or three bits flip.
3. To store for a long amount of time, divide the total time into shorter amounts of time τ, and do EC after each $\tau = t/N$. For large time N, the probability of the bits being in the correct state after error correction is $p \approx (1 - 3(ct/N)^2)^N$, such that $N >> 3(ct)^2$

5.1 QEC key idea

We can use redundant encoding in quantum states such that $|0_L\rangle = |000\rangle$ and $|1_L\rangle = |111\rangle$. There are two problems: (1) how do we measure without destroying superposition? And (2) the errors are continuous, such that one superposition state is changed to another.

To start we consider single qubit errors: the bit flip $\sigma_x^i = X^i$, similar to the classical case, as well as a phase error $\sigma_z^i = Z^i$. Let's treat them individually.

5.1.1 Bit flip error

Consider three physical qubits such that $|0_L\rangle = |000\rangle$ and $|1_L\rangle = |111\rangle$ and a general state $|\psi\rangle = c_0 |0_L\rangle + c_1 |1_L\rangle$. The outcomes are $|\psi\rangle \to |\psi\rangle$ with probability $(1-p)^3$ a single bit is flipped with error $3p(1-p)^2$, etc.

For a correction, we consider a *collective* measurement of four operators:

$$\begin{aligned} P_0 &= |0_L\rangle\langle 0_L| + |1_L\rangle\langle 1_L| \\ P_i &= \sigma_x^i P_0 \sigma_x^i \end{aligned} \tag{268}$$

where i runs over the Pauli matrices. This has the following properties: $P_0 |\psi\rangle = |\psi\rangle$, $P_i |\psi\rangle = 0$, $P_i \sigma_x^j |\psi\rangle = \delta_{ij} \sigma_x^i |\psi\rangle$, such that P_i projects into $\sigma_x^i |\psi\rangle$. The logical state is an eigenstate of the first projector P_0, and the other projectors identify which bit has been flipped! With the measurement result, one can then proceed to fix the error. Also, this procedure works *regardless* of c_0 and c_1: it doesn't matter what the superposition is. *The key is to do a measurement on the redundantly encoded state, without revealing what the state is.*

The procedure is to measure P_0 and if it is 1, the state is fine, if it is 0, then measure P_1. If that measurement is 1, apply σ_x^1 and if 0, move on to P_2.

Remarks

- This works only if a single bit has flipped. If two or three bits was flipped, then there is a problem.
- After the corretion we have $|\psi\rangle$ with probability $(1 - 3p^2 + 2p^3)$ such that the linear term is gone.
- MISSED THIS REMARK
- What if the error was coherent? For example what if $|\psi\rangle \to |\psi\rangle + \eta\sigma_x^i |\psi\rangle$? In this case, measuring P_i still works! For example, if $P_i = 1$, then it forces the system into $\sigma_x^i |\psi\rangle$! A measurement of $P_0 = 1$ projects the state into $|\psi\rangle$. The measurement *digitizes* the coherent, continuous errors! This is very important.

- The physics is that the measurement projects into different subspaces of the Hilbert space. It either projects you into the logical subspace H_L or out of it, into an orthogonal subspace, which can then be corrected - you know how these states are related to the original states in H_L.
- How do we measure these P operators in practice? We want to distinguish e.g. $|000\rangle$ from $|100\rangle$. One way to do this is to measure Z_1Z_2 and Z_2Z_3. If we measure both operators on the logical states, we always get an eigenvalue of 1: these operators compare the first and second qubit, and the second and the third qubits. However, if the measurement is on the state $|100\rangle$ or a state where one bit was flipped, this will no longer be the case. In particular a measurement of Z_iZ_j compares the i and jth qubit.

5.1.2 Phase errors

In quantum mechanics, there are also phase errors! How can this procedure identify those? They correspond to acting the Z Pauli matrix. We can recognize that a bit flip error in one basis is a phase flip error in another, and vice versa. We can use X_iX_j operators therefore to correct phase flip errors. For a state $|\psi\rangle = c_0|0_L\rangle + c_1|1_L\rangle$, the phase flip error brings it to $|\psi\rangle = c_0|0_L\rangle - c_1|1_L\rangle$. We can now measure in the X basis, noting that $\sigma_z|+\rangle = |-\rangle$ and $\sigma_z|-\rangle = |+\rangle$. To protect from Z errors we can encode $|0_L\rangle = |+++\rangle$ and $|1_L\rangle = |---\rangle$.

However, what if we have both errors - for example, a depolarization channel?

5.1.3 Shor's 9 qubit code

The trick is to combine protection from σ_x and σ_z errors by using the states:

$$|0_L\rangle = \frac{1}{2^{3/2}}(|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \, |1_L\rangle = \frac{1}{2^{3/2}}(|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle) \tag{269}$$

How does this work?

- if there's a spin flip, e.g. σ_x^1, then a measurement of $\sigma_z^1\sigma_z^2$ will be -1, and a measurement of $\sigma_z^2\sigma_z^3$ is 1.
- For a phase flip $\sigma_z^i|\psi\rangle$ we can measure $X_1X_2X_3X_4X_5X_6$, $X_4X_5X_6X_7X_8X_9$. These measure if the *phase* is the same on the first three set of qubits as on the second three sets of qubits, and likewise on the second three set of qubits on the last three set of qubits. In this case, if the phase of the first qubit flips, then the phase of the first block will not be the same on the second block, etc. Then you can apply a phase gate to any of the qubits in the block that has the error. The interesting thing here is that the same effect happens if the phase of the second or the third qubit in the first block has flipped and is corrected.

Why does this work? The GHZ states are eigenstates of the X operators! Namely, $\sigma_x^i\sigma_x^j\sigma_x^k(|000\rangle \pm |111\rangle) = \pm(|000\rangle \pm |111\rangle)$, such that a measurement of $X_1X_2X_3$ is kind of like a measurement Z_1 and $X_4X_5X_6$ is like a measurement Z_2 in a rotated basis where the $\pm$ states are the Z basis states.

The bottom line is that if we perform measurements of these six operators as well as pairs of the Z_iZ_j's (6 products of two operators), we can correct ALL single qubit errors! However, still we cannot correct two qubit errors or more. The resulting total error after QEC is the following: for a bit flip error it is $3p^2x3$ blocks which is $9p^2$, and for a phase flip error it's $3p^2$ probability of error, such that the total error is $12p^2$. This procedure only improves if $12p^2 < p$. The reason why this error grows is that we had to use more qubits in a more entangled and complicated state.

5.1.4 Implementing QEC

How to measure $\sigma_x^1 ... \sigma_x^6$ The basic idea is to use an additional qubit 'ancilla' $|\psi_a\rangle$ and prepare it in a superposition $|0_a\rangle + |1_a\rangle$. Then perform CNOT gates with the ith qubit: $U_c^{ai}|\psi_a\rangle|\psi_i\rangle = \frac{1}{\sqrt{2}}(|0\rangle_a|\psi_i\rangle + |1_a\rangle\sigma_x^i|\psi_i\rangle$. Hence after 6 CNOTs we have the state $|0_a\rangle + \Pi_{i=1}^6\sigma_x^i|1_a\rangle$ on the ancilla qubit, which we can then measure in the x basis. You can also use the same procedure to measure products of Z. In total, we need 24 CNOT gates - this is a lot!

Additionally, the operations on the encoded states must be collective. For example, to flip from $|0_L\rangle$ to $|1_L\rangle$ - such that these operations must be collective.

Remarks

- The generalization of Shor's code is the stabilizer formalism. For example, a 3bit code is $S_1 = Z_1 Z_2$, and $S_2 = Z_2 Z_3$. This divides the Hilbert space into logical segments, namely no error, and each of the single qubit bit flip errors. One can generalize this using abelian subgroups S with some elements $[S_i, S_j] = 0$, such that they simultaneously diagonalize the whole Hilbert space $\mathscr{H}_{2^n}$. One can say that $|\psi\rangle$ is stabilized by S if for all S_i in S, $S_i |\psi\rangle = |\psi\rangle$. For S with l generators, each subspace labeled by $S_1, ...S_l$ has dimensions 2^{n-l} which requires $n - l$ qubits. For further reading see Nielsen & Chuang.

- Operational errors introduce two kinds of problems: the error propagate and multiply, for example, a CNOT gate with an error on a control qubit brings the error to the target qubit. Moreover, if the CNOT gates themselves are imperfect, they can also introduce errors, in principle on both of the qubits simultaneously.

- The bad news is that if a circuit creates two errors $\sigma_x^i \sigma_x^j$ there will be no recovery (the code only corrects single qubit errors).

- The good news is that operational errors can be treated as memory errors - we can assume the gates are perfect, but that after each gate we need to apply some memory error correction. This bit of good news can be used to develop the idea of fault-tolerant quantum computing.

The idea of FTQC is to build circuits in a clever way such tath in each step in the circuit you introduce at most one error into each logical encoding block. If this happens, those errors can be corrected at hte next step. You exchange QEC step with logical gates to correct these. If p is the probailiy of failure of each individual component CNOT, then the evolution can be constructed carefully enough such taht the probaility oto intrudce two errors ias at most p^2.

5.1.5 Example: FTQC

A FT CNOT gate is to start with two logical qubits (each having 9 qubits). First do QEC on teh input control states, then do CNOT, then QEC on the output states. What are the possiblities when 2 errors are introduced in one qubit?

1. TEH QECC code error simultaneous on 1 and 2: 24 CNOTs for 1 step of QEC if one fails with probability p_1, the total probabiltiy of error in both upper and lower logical qubit is $(24p_1)^2$.

2. If there is an error in one block and a failed CNOT, that will also introduce 2 errors. The total error probability will be $2(24p_1)(100p_1)$ where 100 comes from the number of CNOTs from a block CNOT.

3. One error correction step fails has probaiblity $(24p_1)^2$ this will also intrduce two errors.

The total probaiblity of having two errors after QEC and CNOT is $10^3 - 10^4 p_1^2$. We require that this error is smaller than the probaility fo a gate error in one qubit. If this is true, then the two error rates do not increase overall and $p_1 < p_{th} = 10^3 - 10^4$. The probability p_{th} is the threshold error. This is the basis for FTQC.

REmarks:

1. Very important result: can keep the error from magnifying if $p_1 < p_{th}$!

2. The perfect decoding at teh end of computation – the resulting error will be p_1^2

3. Cascaded encoding can be used to reduce error even further to do p^{2m} for m levels of encoding or number of cascades. See Nielsen & Chuang or Preskill's notes for further reading.

4. The error threshold is a pecularіar thing! Specifically, it depends on the details of QECC, and it assumes that I only have errors in (GET THIS)? It also assumes that the measurements are efficiently made with time, and that it assumes the errors are completely uncorrelated on all qubits - the noise is uncorrelated.

Teh types of coupling e.g. nearest-neighbor or all-to-all matters. The exact p_{th} is not known, and current reasonable estimates guess that $p_{th} \sim 10^{-3} - 10^{-6}$.

It is extremely unlikely that any practical QC can be built using these existing methods in Misha's opinion. Directions to explore include:

1. More efficient codes lower p_{th} + 'natural topology'

2. Realistic situations and certain kinds of errors dominate (bias-preserving oeprations) - certain operations are difficult.

3. Other approaches include e.g. qubits protected by physics and topology. See for example A. Kitaev's papers, and MS project.

4. More efficient approaches for specific architecture: cat codes, colective gates etc - this is a current frontier of current research! This is the idea of co-design.

6 Quantum complexity theory

This will be a brief single guest lecture by Boaz Barack. Two great references to learn more: *Math and Computation* by Avi Wigderson and *Quantum Computing since Democritus* by Scott Aaronson.

The first question in discussing computational complexity we must ask is what do we want to solve? The two tasks we will discuss is the computation of a function with input x and we want to produce an output $f(x)$. If f is boolean, we sometimes call this "deciding a language." The second will be to sample from a distribution. Given an input p,, whcih is some description of a distribution, the algorithm may use some internal randomness but the output $x \in_R p$. is sampled from the distribution.

In the case of the compute function, we can consider an input string x with size n, we perform a number of operations $S(n)$ that will grow with the size of the input, and we output $f(x)$. Note that we can formally define (using Turing machines or boolean circuits) the function

$$f : \{0,1\}* \to \{0,1\} * \text{computableein} S(n) \text{steps}. \tag{270}$$

Examples

1. Multiply: $a, b \to a\dot{b}$
2. Factor: $m \to p_1, ..., p_k$ s.t. $P_1\dot{p}_2....p_k = m$
3. Min cut: $G \to S$ to mimimize the number of edges
4. Max cut: $G \to S$ to maximize the number of edges.

We can discuss the best known algorithm for each of these algorithms. Note that here we want the algorithm to work on any inputs, which includes the "worst case" inputs which may be the computationally hardest. For example, multiplication takes roughly n^2 using standard algorithm (that we learn in grade school). Karatsuba's algorithm can get $n^{1.6}$, and the best is $O(n\log n)$ in a way similar to the fast Fourier transform (since the FFT is an efficient way of effectively multiplying polynomials).

For factoring, the naive algorithm is to check all of the numbers (say, up to square root of the number to factor). So it takes $O(\sqrt{m})$. However, if the number of digits of m is n, then this will take roughly $2^{n/2}$ steps. There are more clever algorithms such as the quadartic sieve and the nonlinera sieve, which lead to $2^{n^{1/3}}$, which is better but still exponential in n. Min cut can be done in $O(n^2)$, and max cut is actually $\approx 2^n$, with no better known algorithm. This is interesting: it often happens that the complexity of minimizing an objective is much easier than maximizing the same function. These can be very different tasks.

In complexity theory, we can distinguish between problems of varying difficulty. The largest class are problems that can be solved in exponential time *EXP* (e.g. $2^n, 2^{\sqrt{n}}$. We call the problems solvable in polynomial time P. Formally,

$$P : \{f\} st S(f) \leq a.n^b \text{forconst} a, b \tag{271}$$

This is the smallest class, and is a subset of of EXP. We call these problems ones that "we can solve efficiently." Of course, depending on the details of the scaling it may or may not actually be efficient, but comparatively, we cna solve these problems since they do not scale out of control with the problem size. Next, we call problesm $P \subseteq NP \subseteq EXP$, in terms of set-theory. An example of NP is max cut. These are problems that we know we can *verify* in polynomial time. NP is in some sense, the problems that we actually want to solve, since we could usefully verify the answer once we have a solution. Note that NP does not mean "not polynomial" - it only means that the solution can be verified in polynomial time.

It is possible that $P = NP \subsetneq EXP$, but it is also possible that $P \subsetneq NP = EXP$. We don't really know. Of course, it is also possible (and strongly believed) that $P \subsetneq NP \subsetneq EXP$. In fact, this is usually widely believed in computer science and used as a working hypothesis, but we do not have a working proof of this.

So where does quantum fit in? It fits in the class BQP:

$$BQP : \{f | x \to_{\text{quantumcircuit}} f(x)\} \text{ in} \leq a.n^b \text{ gates} \tag{272}$$

Since we know that classical gates can be performed with polynomial overhead on a quantum computer, we know that $P \subseteq BQP$. Since we know that classically we can simluate a quantum computer we know that $\subseteq EXP$.

Again, it is not known strictly whether these are subsets or equalities, however it is believed $P \subsetneq BQP \subsetneq EXP$. For example, we know that factoring is in BQP, but it is not known whether factoring is in P. Note that the quantum fourier transform is the workhorse of problems that are known to be in BQP but are believed not to be in P. There may be other ways to get exponential speedups, but it is not necessarily known. Also, it may in fact be that BQP extends into EXP - this is also not known.

There is a notion of problems such as 3 SAT and MAX CUT which are known as NP-complete (NPC). It is conjectured that none of them are in P nor BQP (not certain, but widely believed). One reason for this is that it can be shown that if even a single problem is in P, all problems in NP would be in P (or in BQP as well). The formal statement of the theorem is that if $MAXCUT \in P \rightarrow f \in P \; \forall \; f \in NP$. The way of proving this is so-called reduction, where we can take an algorithm for MAXCUT and translate it into an algorithm for f. In other words, if we have a magic box that solves MAXCUT, by using this magic box we could solve f. Note that this theory (and overall description) does not necessarily apply to best or average cases of problems, but to worst case implementation.

The canonical way of thinking of a problem in NP is by considering a circuit which has input $p_1...p_k$, and the output is 1 if $p_1...p_k = m$, and 0 otherwise. In fact, we can rephrase the factoring problem in terms of this circuit. This circuit can be taken and mapped onto a graph such that if there was an input that made the output 1, there would be a set that cuts $\geq 0.9n$ of the edges. So we can reduce the task of this generic circuit into finding the MAXCUT.

The correllary of this is that if we have a polynomial time algorithm for MAXCUT, we get a polynomial time algorithm to solve any efficiently verifiable task. For example, suppose someone gave you an encryption taht takes plaintext and a key, and outputs cipher text. We now the algorithm and the ciphertext and the plaintext. However, we don't know the secret key. Of course, if someone told you the secret key you could verify (e.g. write a circuit that outputs 1 for the correct answer). In other words, if you could solve MAXCUT, you could break *all* crypto-systems. This is to be distinguished to known quantum algorithms like Shor's algorithm which breaks factoring based crypto-systems, but not necessarily all crypto-systems. In fact, we believe there are crypto systems that take an exponential amount of time even on a quantum system, implying that $BQP \subsetneq NP$. The conjecture is that any quantum algorithm requires $2^{n/2}$ operations, as opposed to 2^n for classical systems. An exmaple of this is Grover's algorithm, which is actually known to be the best. One can make a corrolary to this that an NP problem will take of the order $\geq 2^{\epsilon n}$ operations.

So how do we cope with NP hard problems? Recall that some problems may be NP hard in the *worst case*. If the input is some $(x_1, y_1)...(x_n, y_n)$, and the goal is to find $f \in NN(d, S) \; \forall_i \; f(x) = y_i$. Training a neural net NN is NP hard in the worst case (but not always). However, we know that we can do this efficiently for some useful problems. In the context of graph theory, we sometimes pick special graphs such as bounded graphs, graphs on the edge of a genus, that can be solved with specific algorithms in polynomial time. These are algorithms for special instances. It is also possible for random instances (not necessarily well known in advance, but likely to occur). Finally, sometimes we change the problem and get more data by using less parameters and modifying the problem until we can actually solve it. For example, we may want to solve equations on integers but could solve them on real numbers and use rounding.

Now, onto quantum supremacy. It is a slightly different computational task, since we are given a distribution and we want to sample from that distribution. It turns out that this is very closely related to the problem of counting, where the input is a function $f : \{0,1\}^n \rightarrow \{0,1\}$, and the output is to count $\sum_{x \in \{0,1\}^n} f(x)$. For example, instead of MAXCUT you could ask *how many cuts* of a certain size do we have. One example of a distribution we may want to sample from is p over an $n \times n$ random matrix A such that $p(A) \propto \text{perm}(A)^2$, referring to the *permanent*, which is defined as

$$perm(A) := \sum_{\tau \in S_n} \prod_{i=1}^{n} A_{i,\tau(i)}. \tag{273}$$

There exists a conjecture that no polynomial time randomized algorithm exists for computing the permanent.

The ideology behind tackling these problems on a quantum computer is the following. Building a full-fledged fault tolerant quantum computer is hard. Perhaps it is not as hard to build a quantum computer that can solve some computational task which is hard for classical computers. Informally, the hope is that we have a quantum computer the size of roughly a refrigerator, that can perform a certain task that would take a classical data center the size of a small city. The hope would be that then as we scale up the size of the problem, the time and resources for solving the problem on a quantum machine would scale manageably, and not exponentially like with the case of the classical machine. Here, we give up on the idea that the task is *useful*, just focusing

on the fact that the task is *hard*, as an experimental demonstration that this technology can potentially provide exponential speedups.

The example for this is p described by a quantum circuit C. here we have

$$p_C(x) \propto (\langle x|C|0^{\otimes n}\rangle)^2. \tag{274}$$

The dream quantum supremacy experiment takes in a circuit, programs it onto the device, and we get out some $x \sim p_C$ (a sample from the distribution generated by the circuit). In reality, the best we can hope is that given C, we get $x \sim \delta p_c + (1-\delta)J$ where $J \approx (1-\epsilon)^{\text{numgates}}$. (Here, J stands for *junk*). In practice, δ will be small, but hopefully not so small that after running the experiment a manageable number of times wen have a statistically meaningful number of samples that are not junk. In the case of the Google experiment, this was on the order of a fraction of a percent.

There is strong evidence that the ideal sampling problem is hard classically, but not necessarily in the case of the imperfect ($\delta < 1$) circuits. In practice, another problem is that in the real device

$$p(x) = (1+\delta)2^{-n}, \tag{275}$$

comparing with $p(x) = 2^{-n}$ for $x \sim \{0,1\}^n$ and $p(x) = 3 \times 2^{-n}$ for $x \sim p_c$. We would ideally have to verify the runtime as 2^n classically to show that this is hard.

If this supremacy proof is true, the implications are as follows. We can actually provide quantum speedups over classical algorithms. We have known this thanks to Shor's theorem, but it is nice to see this experimentally. We can also think of this as an extension of Bell's inequalities, which tell us that the universe cannot be simulated by some classical machine with hidden variables.

Finally, consider the PCP theorem. Assume we have an efficient verifier which runs in polynomial time. However, we have some provers that are unbounded. The provers can compute $f(x)$ and tell the verifier that $f(x) = 1$. However, there should be some way of checking that the provers cannot cheat. If there is a barrier between the provers and we randomly interrogate the provers, even if f is computable only in exponential time, we can still verify in polynomial time. The point of the barrier is to force the prover to be non-adaptive. In other words, we can now ask the provers subsequent questions which do not depend on the previous questions and answers. So we can think of the provers responses as completely pre-defined (all 2^n responses), and we query at random and are guaranteed that we will get the right answer.

Suppose there exists a function $H : \{0,1\}^n \to \{0,1\}$. The question is does there exist an x such that $H(x) = 0$? The PCP theorem tells us that it is possible to transform $H(x)$ into a more robust form. If there exists x such taht $H(x) = 0 \to$, there exists x such that the expectation over $i,j,kH_{i,j,k}(x) = 0$. In other words, for all x such that $H(x) > 0 \to \forall\ x$ the expectation value $> 1/100$.

This is related to recent work $MIP* = RE$