Quantum Computing: Foundations to Frontier Fall 2018

Lecture 1

Lecturer: Henry Yuen *Scribes: Dmitry Paramonov*

1 Overview

In the previous half of this lecture, we covered administrative details, as well as the motivation for quantum computing and the current state of quantum computing theory and implementation.

In this lecture, we covered the basic idea of a qubit, how they measure and evolve, as well as how they differ from classical probabilistic bits. We also covered the systems with multiple qubits, the No-Cloning Theorem and Holevo's Theorem.

In particular, a qubit is a complex unit vector in $\mathbb{C}^2$, specified as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

When measuring $|\psi\rangle$, we obtain outcome $|0\rangle$ with probability $|\alpha|^2$ and outcome $|1\rangle$ with probability $|\beta|^2$.

If a qubit is in the state $|\psi\rangle \in \mathbb{C}^2$, and another qubit is in the state $|\phi\rangle \in \mathbb{C}^2$, then the joint state of both qubits is described by the tensor product

$$|\psi\rangle \otimes |\phi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$$

Calculating the probability of measurement outcome $|i\rangle \otimes |j\rangle$ (where $i, j \in \{0, 1\}$) works the same way as before.

$$\Pr\left(|\psi\rangle|\phi\rangle \to |i\rangle|j\rangle\right) = |(\langle\psi| \otimes \langle\phi|)(|i\rangle \otimes |j\rangle)|^2 = |\langle\psi|i\rangle \cdot \langle\phi|j\rangle|^2$$

2 Quantum Information Theory

We wish to describe and design algorithms for a quantum computer. But to do this, we require knowledge and understanding of how a quantum computer works in the first place, as well as the math underlying its operation. Thus, we first need to specify exactly how quantum information is represented.

2.1 Qubits

In a classical system, we have a single *bit*. This is an item with two distinguishable states. It can be instantiated as a binary switch, electron spin, transistor voltage level or basically anything.

The simplest quantum system is a *qubit*, or a quantum bit. Mathematically, this is a complex unit vector in $\mathbb{C}^2$. We represent it as a vector $|\psi\rangle$ (which we call "ket psi"):

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \text{ such that } |\alpha|^2 + |\beta|^2 = 1$$

A qubit can represent classical 0 and 1 states as well.

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

These states are orthogonal. Thus, every qubit is a *superposition* of the classical states.

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

α and β are *amplitudes*. They are reminiscent of probabilities, albeit they can be negative or even complex.

The notation $|\psi\rangle$ is called *Dirac notation*, and is used to represent quantum states. Mathematically, it's a column vector. The dual of it is the row vector $\langle\psi|$ (which we call "bra psi"), which is the Hermitian conjugate of $|\psi\rangle$:

$$\langle\psi| = \left(\alpha^*, \beta^*\right) = \alpha^*\langle 0| + \beta^*\langle 1| = \alpha^*\left(1, 0\right) + \beta^*\left(0, 1\right).$$

Here, if $\alpha = a + ib$ for real numbers a, b, then α^* denotes the complex conjugate $a - ib$. Note: $\overline{\alpha}$ is traditionally used to denote complex conjugation, but we will use the star notation in this course.

We can thus represent the inner product between two states $|\psi\rangle, |\phi\rangle$ very cleanly, using the *bra-ket form*, $\langle\psi|\phi\rangle$. Let $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|\phi\rangle = \gamma|0\rangle + \delta|1\rangle$. Then

$$\begin{aligned}
\langle\psi|\phi\rangle &= \langle\psi||\phi\rangle \\
&= \left(\alpha^*\langle 0| + \beta^*\langle 1|\right)\left(\gamma|0\rangle + \delta|1\rangle\right) \\
&= \alpha^*\gamma\langle 0||0\rangle + \beta^*\gamma\langle 1||0\rangle + \alpha^*\delta\langle 0||1\rangle + \beta^*\delta\langle 1||1\rangle \\
&= \alpha^*\gamma \cdot 1 + \beta^*\gamma \cdot 0 + \alpha^*\delta \cdot 0 + \beta^*\delta \cdot 1 \\
&= \alpha^*\gamma + \beta^*\delta
\end{aligned}$$

Dirac notation lets us clearly distinguish column vectors, row vectors and scalars at a glance, just by seeing which types of brackets surround it.

Reminder. Since we're working with complex numbers, we have that

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*.$$

2.1.1 Measuring a Qubit

QM Postulate #1: The information in quantum states is not directly accessible; they have to be *measured*.

If you measure a qubit, you obtain a classical outcome $j \in \{0, 1\}$ probabilistically. In the case of $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, you get outcome 0 with probability $|\alpha|^2$ and you get outcome 1 with probability $|\beta|^2$.

This is known as the *Born Rule*, first formulated by physicist Max Born [1].

Let's say you start with a qubit $|\psi\rangle$, and you measure it, getting outcome j. If you then measure the qubit again, you will obtain j with probability 1. In other words, the qubit has *collapsed* to the classical state $|j\rangle$. After the measurement, $|\psi\rangle$ becomes one of $|0\rangle$ or $|1\rangle$. The act of measurement itself changes the state it is in. The superposition is destroyed.

2.1.2 Evolution of a Qubit

QM Postulate #2: In isolation (meaning that no measurements are being performed on the state), quantum states evolve via *unitary operations*. These are square matrices U that take unit vectors to unit vectors.

That means that if $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ were unitary and $|\psi\rangle$ were a unit vector, then we would get $U|\psi\rangle$ to be a unit vector.

Note that a unitary matrix also has several other necessary and sufficient conditions, such as that the inverse of a unitary matrix is its conjugate transpose, and that a unitary matrix preserves inner products:

1. For all $|\psi\rangle$ such that $\| |\psi\rangle \|^2 = 1$, we have $\| U|\psi\rangle \|^2 = 1$.
2. $U^{-1} = U^\dagger = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}$.
3. For all $|\psi\rangle, |\phi\rangle$, we have $\langle\psi'|\phi'\rangle = \langle\psi|\phi\rangle$ where $|\psi'\rangle = U|\psi\rangle$ and $|\phi'\rangle = U|\phi\rangle$.

A notational tip. If $|\phi\rangle = U|\psi\rangle$, then $\langle\phi| = \langle\psi|U^\dagger$. You need to take the Hermitian conjugate of U.

2.1.3 Quantum Weirdness

Now, how is a quantum bit different from a probabilistic bit? To see the difference, consider the following two experiments.

Experiment A: Let $|\phi\rangle = |0\rangle$ and let $U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$.

1. Apply U to the qubit.
2. Measure the qubit.
3. Apply U to the qubit again.
4. Measure the qubit.

What is the distribution of the final measurement outcome?

After the first application, we get $U|\phi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$. Thus, the first measurement has an even chance of resulting in either $|0\rangle$ or $|1\rangle$. Thus, we have two possible events. Call them (a) (corresponding to outcome $|0\rangle$) and (b) (corresponding to outcome $|1\rangle$).

In (a), we measured $|0\rangle$, so that is the qubit's state. After applying U again, it again becomes $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$. Thus, we again have an equal chance of measuring $|0\rangle$ and $|1\rangle$ in the last step of Experiment A.

In (b), its state became $|1\rangle$. After applying U, its state becomes $U|1\rangle = -\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$. Thus, there is an equal chance again of measuring $|0\rangle$ and $|1\rangle$.

Thus, after performing this experiment, we have an equal chance of measuring $|0\rangle$ and $|1\rangle$ during the second measurement. All this experiment does is randomize the bit.

Experiment B: Now, suppose that we did not measure the intermediate state.

1. Apply U to the qubit.
2. Apply U to the qubit again.
3. Measure the qubit.

What is the distribution of the final measurement outcome? Let us calculate the state that it will be in.

1. (After step 1):

$$\begin{aligned} U|\phi\rangle &= \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \\ &= \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \end{aligned}$$

2. (After step 2):

$$\begin{aligned}
U^2|\phi\rangle &= U\left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\right) \\
&= \frac{1}{\sqrt{2}}U|0\rangle + \frac{1}{\sqrt{2}}U|1\rangle \\
&= \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\right) + \frac{1}{\sqrt{2}}\left(-\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\right) \\
&= \frac{1}{2}|0\rangle + \frac{1}{2}|1\rangle - \frac{1}{2}|0\rangle + \frac{1}{2}|1\rangle \\
&= |1\rangle
\end{aligned}$$

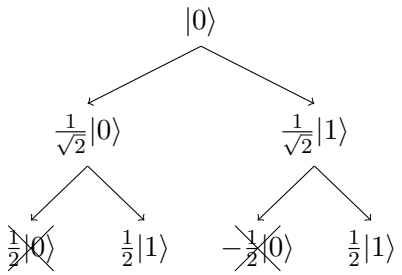
So after applying U two times, we deterministically get state 1. The final outcome is deterministic – very different from what happened in Experiment A!

What happened to the $|0\rangle$ state in Experiment B? Looking at the calculation above, it seemed to have *cancelled itself*. Here, we see an example of *interference*.

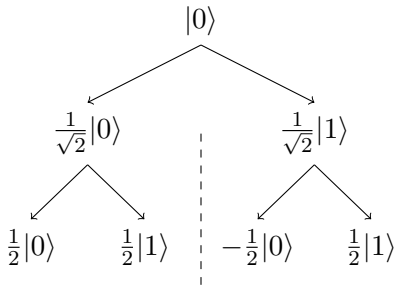
On the other hand, in Experiment A, this cancellation/interference didn't happen; this is because of the intermediate measurement step!

There are multiple takeaways from this. The first take away is: the sign of the amplitude matters! Without minus signs in the amplitudes, such interference could never happen.

The second takeaway is this: applying multiple unitary operations to a qubit gives rise to a "tree of paths" between states. Each path has an amplitude associated with it. But before you measure, you sum all the amplitudes among the leaves. The paths can constructively or destructively interfere with each other. This differs from classical probability, where probabilities can only add, and can't destructively interfere. Below is a diagram of the tree of paths in Experiment B.



But in Experiment A, our intermediate measurements destroy superpositions and prevent interference. Measurements make the qubits become classical. To see quantum effects, we need to delay measurements as long as possible. This is why it is hard to see quantum effects in everyday life. Measurement is constantly happening, which also makes a quantum computer hard to build.



2.2 Composite Quantum Systems

We've discussed the rules of quantum mechanics for a single qubit, which are pretty interesting, but things really start cooking when we talk about multiple qubits.

The state of a qubit lives in the vector space $\mathbb{C}^2$. This is called the *Hilbert space* of a qubit. For our purposes, this is a complex vector space with an inner product.

The Hilbert space of two qubits is the *tensor product space* $\mathbb{C}^2 \otimes \mathbb{C}^2$. $\mathbb{C}^2$ has an orthonormal basis $\{|0\rangle, |1\rangle\}$. The tensor product space $\mathbb{C}^2 \otimes \mathbb{C}^2$ is 4-dimensional (in fact it is isomorphic to $\mathbb{C}^4$), and has the following orthonormal basis:

$$\left\{ |0\rangle \otimes |0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, |0\rangle \otimes |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, |1\rangle \otimes |0\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, |1\rangle \otimes |1\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

We will use the shorthand notation $|ij\rangle = |i\rangle|j\rangle = |i\rangle \otimes |j\rangle$.

If we have two qubits, one of which is in the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \in \mathbb{C}^2$, and the other is in the state $|\phi\rangle = \gamma|0\rangle + \delta|1\rangle \in \mathbb{C}^2$, then the joint state of the two qubits together is the tensor product state

$$\begin{aligned} |\psi\rangle \otimes |\phi\rangle &= (\alpha|0\rangle + \beta|1\rangle) \otimes (\gamma|0\rangle + \delta|1\rangle) \\ &= \alpha\gamma|0\rangle \otimes |0\rangle + \alpha\delta|0\rangle \otimes |1\rangle + \beta\gamma|1\rangle \otimes |0\rangle + \beta\delta|1\rangle \otimes |1\rangle \\ &= \alpha\gamma|00\rangle + \alpha\delta|01\rangle + \beta\gamma|10\rangle + \beta\delta|11\rangle \\ &= \begin{pmatrix} \alpha\gamma \\ \alpha\delta \\ \beta\gamma \\ \beta\delta \end{pmatrix}. \end{aligned}$$

The tensor product operation $\otimes$ is like a multiplication operation, so it and the addition operation $+$ together obey the distributive and associative rules that we are familiar with.

In general, a two qubit state $|\psi\rangle$ is a unit vector in $\mathbb{C}^2 \otimes \mathbb{C}^2$:

$$|\psi\rangle = \sum_{i,j} \alpha_{i,j}|ij\rangle \text{ such that } \sum_{i,j} |\alpha_{i,j}|^2 = 1$$

In general, a two-qubit state $|\psi\rangle$ cannot be written as a tensor product state, meaning that we cannot find a single-qubit states $|\phi\rangle \in \mathbb{C}^2$ and $|\theta\rangle \in \mathbb{C}^2$ such that $|\psi\rangle = |\phi\rangle \otimes |\theta\rangle$. States that cannot be written in this product form are called *entangled*. Otherwise, they are unentangled.

A notable such state is the state $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$ – this is called the *EPR pair*, or *Bell pair*, or *maximally entangled state*. This cannot be expressed as the tensor product of two single qubit states.

The inner products in $\mathbb{C}^2 \otimes \mathbb{C}^2$ works as follows: let $|a\rangle, |b\rangle, |c\rangle, |d\rangle \in \mathbb{C}^2$, then

$$(\langle a| \otimes \langle b|)(|c\rangle \otimes |d\rangle) = \langle a|c\rangle \cdot \langle b|d\rangle.$$

Measurements. Measuring a two-qubit state gives you a classical outcome, as expected. Suppose that $|\psi\rangle = \sum_{i,j\in\{0,1\}} \alpha_{i,j}|ij\rangle$ is our two-qubit state. Then when we measure $|\psi\rangle$, we could obtain outcome $|00\rangle$, $|01\rangle$, $|10\rangle$ or $|11\rangle$. And the probability is again proportional to the square magnitude of the quantum amplitude associated with the outcome:

$$\Pr\left(|\psi\rangle \to |ij\rangle\right) = |\langle\psi|ij\rangle|^2 = |\alpha_{i,j}|^2$$

Again, once you measure and obtain outcome (i, j), then the qubits have collapsed to the state $|i, j\rangle$ (i.e. the first qubit is in state $|i\rangle$, the second qubit is in the state $|j\rangle$).

But what happens if you only want to measure the first qubit, and leave the second qubit alone? This is called a *partial measurement*. We will get outcome $i \in \{0, 1\}$ with probability

$$p_i = \sum_{j\in\{0,1\}} |\alpha_{i,j}|^2$$

Then, after we measure this state, we get the *post-measurement state*

$$|\phi_i\rangle = |i\rangle \otimes \frac{1}{\sqrt{p_i}} \sum_j \alpha_{i,j}|j\rangle.$$

In other words, the first qubit has collapsed to the state $|i\rangle$, and the second qubit has collapsed to the state

$$\frac{1}{\sqrt{p_i}} \sum_j \alpha_{i,j}|j\rangle.$$

The normalization factor $\frac{1}{\sqrt{p_i}}$ is important, because we need normalized unit vectors at the end.

As an exercise to the reader, prove that doing a partial measurement on the first qubit and then doing a partial measurement on the second yields the same distribution of outcomes as one full measurement.

Unitary evolution. Isolated evolution of two qubits still undergoes evolution via unitary operations on $\mathbb{C}^2 \otimes \mathbb{C}^2$. These can be represented by 4×4 matrices satisfying the same definition(s) of unitarity we had above (preserve norm, inverse is conjugate transpose, preserve inner products).

Let's say we wanted to apply a single-qubit unitary $U \in \mathbb{C}^{2\times2}$ to one qubit, and a single-qubit unitary $V \in \mathbb{C}^{2\times2}$ to another qubit. The overall unitary describing what is happening to both

qubit is given by the tensor product $U \otimes V$, which is defined as:

$$U = \begin{pmatrix} u_{1,1} & u_{1,2} \\ u_{2,1} & u_{2,2} \end{pmatrix}$$
$$V = \begin{pmatrix} v_{1,1} & v_{1,2} \\ v_{2,1} & v_{2,2} \end{pmatrix}$$
$$U \otimes V = \begin{pmatrix} u_{1,1}V & u_{1,2}V \\ u_{2,1}V & u_{2,2}V \end{pmatrix}.$$

Applying $U \otimes V$ to a general two-qubit state $|\psi\rangle = \sum \alpha_{i,j}|i,j\rangle$ behaves as follows:

$$(U \otimes V)|\psi\rangle = (U \otimes V) \sum \alpha_{i,j}|i,j\rangle = \sum \alpha_{i,j} U|i\rangle \otimes V|j\rangle.$$

But once again, most two-qubit unitaries cannot be expressed as a tensor product of two single-qubit unitary operators. Most two-qubit unitaries are *entangling*; meaning given a tensor product state $|\psi\rangle \otimes |\phi\rangle$ the output will often be an entangled state.

How do we describe doing *nothing* to a qubit? Let's say we want to apply some unitary U to the first qubit, and leave the second qubit alone. This corresponds to applying the *identity* unitary $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. The overall unitary is $U \otimes I$.

2.2.1 No-Cloning Theorem

Classical bits are easily copied. But quantum information is different. This is known as the *No-Cloning Theorem*. Informally, it states that "there is no quantum Xerox machine".

Formally, this Theorem states that there is no unitary U acting on two qubits such that for any one-qubit state $|\phi\rangle$,

$$U\left(|\phi\rangle \otimes |0\rangle\right) = |\phi\rangle \otimes |\phi\rangle$$

Here's an intuition, any quantum Xerox machine must also act as a classical Xerox machine. However, a classical Xerox machine will fail to copy certain entangled states.

In more detail: Let U be a two-qubit classical copying unitary that acts as follows: for $x \in \{0,1\}$

$$U\left(|x\rangle \otimes |0\rangle\right) = |x\rangle \otimes |x\rangle$$
$$U\left(|x\rangle \otimes |1\rangle\right) = |x\rangle \otimes |x \oplus 1\rangle$$
$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

This unitary has a special name: It is the *CNOT* unitary.

Let's try to copy the following $|\phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.

$$\begin{aligned} U(|\phi\rangle \otimes |0\rangle) &= U\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle\right) \\ &= \frac{1}{\sqrt{2}}(U(|0\rangle \otimes |0\rangle) + U(|1\rangle \otimes |0\rangle)) \\ &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\ &\neq |\phi\rangle \otimes |\phi\rangle \end{aligned}$$

The desired copying unitary doesn't actually work.

Note that the above only proves that the CNOT unitary doesn't copy the data. A more general proof showing that such a unitary cannot exist is left as an exercise to the reader.

2.2.2 Exponentiality of many qubits

Let's go beyond two qubits. Consider two Hilbert spaces $\mathbb{C}^r$ and $\mathbb{C}^s$, with orthonormal bases $\{|\psi_1\rangle, \ldots, |\psi_r\rangle\}$ and $\{|\phi_1\rangle, \ldots, |\phi_s\rangle\}$. These have dimension r and s, respectively.

The tensor product Hilbert space $\mathbb{C}^r \otimes \mathbb{C}^s$ has dimension rs. This has orthonormal basis

$$\{|\psi_i\rangle \otimes |\phi_j\rangle\}_{1 \leq i \leq r, 1 \leq j \leq s}$$

The inner product is again defined by taking independent components. If $\psi_1, \psi_2 \in \mathbb{C}^r$ and $\phi_1, \phi_2 \in \mathbb{C}^s$, then

$$(\langle\psi_1| \otimes \langle\phi_1|)(|\psi_2\rangle \otimes |\phi_2\rangle) = \langle\psi_1|\psi_2\rangle \cdot \langle\phi_1|\phi_2\rangle$$

Parital measurements. In general, partial measurements in tensor product spaces work like this. Let's say we have a quantum state $|\psi\rangle = \sum \alpha_{ij}|i\rangle|j\rangle \in \mathbb{C}^r \otimes \mathbb{C}^s$, where the index i runs over $\{1, \ldots, r\}$ and the index j runs over $\{1, \ldots, s\}$. Then we can always write

$$|\psi\rangle = \sum_{i=1}^{r} c_i |i\rangle \otimes |\varphi_i\rangle$$

where

$$|\varphi_i\rangle = \frac{1}{c_i} \sum_j \alpha_{ij}|j\rangle$$

with c_i being defined as

$$c_i = \left\| \sum_j \alpha_{ij}|j\rangle \right\| = \sqrt{\sum_j |\alpha_{ij}|^2}$$

i.e., the Euclidean norm of $\sum_j \alpha_{ij}|j\rangle$. Note that c_i is a nonnegative number, and the collection of quantum states $\{|\varphi_i\rangle\}$ are normalized unit vectors in $\mathbb{C}^s$, but are not necessarily orthogonal.

If we do a partial measurement on the $\mathbb{C}^r$ space, then we obtain outcome $|i\rangle$ with probability $|c_i|^2$, and the post-measurement state is

$$|i\rangle \otimes |\varphi_i\rangle.$$

Reminder. Note that when multiplying tensor products of vectors, you match up the tensor factors (the "slots"). But you do not do this when adding.

$$(|a\rangle \otimes |b\rangle) + (|c\rangle \otimes |d\rangle) \neq (|a\rangle + |c\rangle) \otimes (|b\rangle + |d\rangle)$$

QM Postulate #3: If the states of quantum systems A and B are $|\psi\rangle \in \mathbb{C}^r$ and $|\phi\rangle \in \mathbb{C}^s$ respectively, then the joint state of AB is $|\psi\rangle \otimes |\phi\rangle \in \mathbb{C}^r \otimes \mathbb{C}^s$.

This means that the joint state of N qubits is represented as a vector in $\left(\mathbb{C}^2\right)^{\otimes N}$, which is isomorphic to the space $\mathbb{C}^{2^N}$. Such a vector would be written in the form

$$|\psi\rangle = \sum_{x \in \{0,1\}^N} \alpha_x |x\rangle$$

Each additional qubit doubles the dimensionality of the Hilbert spaces. Thus, applying a unitary U to an N-qubit state appears to be doing exponentially many computations in parallel, since we appear to be applying a $2^N \times 2^N$ matrix to a vector of length 2^N.

Nature is thus doing an incredible amount of work for us. However, this extravagance is hidden behind the veil of measurement. We can only access this exponential information in a limited way.

We thus have a tradeoff between exponentiality and fragility of quantum data. Even as we can express more information in a quantum state, this information is hard to access without damaging the state.

2.2.3 Holevo's Theorem

This tradeoff between exponentiality and fragility is sharply illustrated with this following important result in quantum information theory. Naively, you might expect n qubits of quantum information to be able to store 2^n bits of classical information. However, because of the fragility of quantum states, one cannot reliably extract this exponential information.

This is known as *Holevo's Theorem*, postulated by Alexander Holevo [2]. Informally, it says that you cannot reliably store n bits of classical information in less than n qubits.

Formally, let $m \leq n$. Alice gets a uniformly random string $X \in \{0,1\}^n$, and sends an m qubit state $|\phi_X\rangle$ to Bob, who applies a measurement to obtain a string $Y \in \{0,1\}^n$. For any enconding/decoding strategy of Alice and Bob,

$$I(X : Y) \leq m$$

Here, $I(X : Y)$ denotes the *mutual information* between two random variables. Roughly, it measures the number of bits of correlation between X and Y. So if $m \ll n$, then it says that Bob's measurement outcome Y contains much fewer than n bits of information about X, meaning that Bob won't be able to figure out what X is.

References

[1] M. Born, *Quantenmechanik der Stoßvorgänge*, Zeitschrift für Physik, 37(12):863–867, 1926.

[2] A. S. Holevo, *Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel*, Problems of Information Transmission. 9(3):177–183, 1973.