# Quantum Computation and Quantum Communication for beginners

Quantum Troopers

Summer 2020

## Contents

# 1 Introduction and Overview

We, quantum troopers, found that the qiskit documentation and other relevant introductory textbooks on quantum computing are not quite approachable for beginners to the field. While quantum computing has become a growing research interest in the past decade, we aim at providing general readers with a step-to-step guide towards understanding the mechanisms behind quantum computation and quantum communication. The first couple chapters of this textbook are organized in a math-oriented way which informs readers of.... whereas the later chapters focus more on specific algorithms and applications...

# 2 Linear Algebra

Linear algebra is at the very core of the framework of quantum mechanics. From Schrodinger's equation to quantum gates in quantum computing, all of the mathematical models are constructed based on linear algebra. At the first glance this may be counter intuitive — why would the description of particles, wave functions and their evolution have anything to do with vectors and matrices?
To speak broadly, it is because matter waves, which are mathematically modeled by wave functions, and their evolution over time, display linearity. We'll come to the more precise definitions of linear spaces (vectors) and linear transformations (matrices) in this chapter.

But as a plain intuition, two waves with amplitude of $a$ and $b$ at a particular position can be composed into a single wave of amplitude $a + b$, this additive phenomenon called superposition is what allows wave functions that model matter waves to be treated as vectors. Similarly, it only makes sense that if a wave function $|\phi_0\rangle$ evolves over a small period of time to become $|\phi_1\rangle$, and that by adding to that wave function $|\psi_0\rangle$ which would evolve into $|\psi_1\rangle$, the collective effect should be that the wave function begins as $|\phi_0\rangle + |\psi_0\rangle$ and ends as $|\phi_1\rangle + |\psi_1\rangle$ due to superposition; this is the qualifying property that allows the evolution of wave functions to be expressed as linear transformations, which will be defined in more exact terms later.

## 2.1 Vectors & Complex Vector Space

Linear algebra is the general class of algebra that is concerned with properties of quantities that are "linear" in their nature. Such quantities, regardless of their presented form, are dubbed the name "vectors". To define linearity and consequently the kinds of quantities on which conclusions from linear algebra apply, a **vector space** $V$ over a field $F$ must satisfy:

1. For all vectors $x$ and $y$ in $V$, $x + y = y + x$.
2. For all vectors $x$, $y$ and $z$ in $V$, $(x + y) + z = x + (y + z)$.
3. There exists additive identity 0. (the additive identity isn't the numerical quantity 0, it's a shorthand notation for the additive identity that is unique to each space)
4. For each vector $x$, $1x = x$.
5. For each pair of element $a, b$ in $F$ and each element $x$ in $V$, $(ab)x = a(bx)$.
6. For each pair of element $a$ in $F$ and each pair of elements $x, y$ in $V$, $a(x+y) = ax+ay$. (distributivity)
7. For each pair of elements $a, b$ in $F$ and each element $x$ in $V$, $(a + b)x = ax + bx$. (distributivity)

The field $F$ mentioned in the above definition can be any set of numbers. For example, the set of real numbers $\mathbb{R}$ is a perfectly valid field. Another example would be the complex numbers $\mathbb{C}$, and so on. In quantum computing, the vector space that we concern ourselves with is $\mathbb{C}^n$, where $n$ is an integer. For example, a vector from the space $\mathbb{C}^3$ must have the form $(a, b, c)$, where $a, b$, and $c$ are all complex numbers. For any two vectors $(a_1, a_2, ..., a_n)$ and $(b_1, b_2, ..., b_n)$ from $\mathbb{C}^n$, and $x \in \mathbb{C}$, the operations are defined as such:

1. $(a_1, a_2, ..., a_n) + (b_1, b_2, ..., b_n) = (a_1 + b_1, ..., a_n + b_n)$.
2. $x(a_1, a_2, ..., a_n) = (xa_1, xa_2, ..., xa_n)$.

As you can see, these rules are exactly the same as those defined for geometric vectors, except that the entries are changed from real numbers to complex numbers. Further, most properties of "vectors" in the sense of geometric spaces are preserved in members of $\mathbb{C}^n$, so one can feel free to visualize them as regular vectors when learning the concepts. The following chapters will be focused on discussing $\mathbb{C}^n$, which is particularly useful for quantum computing, due to the limited "spaces" in this textbook.

### 2.1.1 Linear combination

For a set of $m$ vectors in $\mathbb{C}^n$, $v_i$ being the $i$-th vector of this set, and $a_i \in \mathbb{C}^n$, $v = a_1v_1 + a_2v_2 + ... + a_mv_m$ is a **linear combination** of the vectors $v_1$ to $v_m$. The set of all linear combinations of a set of vectors is also called the **span** of vectors. For example, the span of $(0, 1)$, $(1, 0)$ is the entirety of the $\mathbb{C}^2$ space.

### 2.1.2 Linear independence

For a set of $m$ vectors in $\mathbb{C}^n$, $v_i$ being the $i$-th vector of this set, the set is called **linearly independent** when the only way to make

$$a_1v_1 + a_2v_2 + ... + a_mv_m = 0$$

is by setting $a_1 = a_2 = a_3 = ... = a_m = 0$. Another way to conceptualize the idea of linear independence is that when a set of vectors are linearly independent, there is no way to express any vector in the set as linear combinations of the others.

### 2.1.3 Basis

A **basis** $\beta$ for a linear space $V$ is a set $S$ of linearly independent vectors that also span that whole space. You can think of it as a minimum skeleton of the linear space. Due to the **replacement theorem**, it turns out that the number of vectors required in any basis for the same linear space is always the same, no matter whether the vectors are orthogonal or skewed in their arrangement. The number of vectors in a basis $\beta$ for linear space $V$ is called the **dimension** of the space.

As an example, consider the set $\beta = \{(1, 0), (0, 1)\}$ from $\mathbb{C}^2$, the set is a basis of $\mathbb{C}^2$. To show $\beta$ as a basis, for any $(a, b) \in \mathbb{C}^2$, we can find $a, b \in \mathbb{C}$ such that $a(1, 0) + b(0, 1) = (a, b)$, thus it **spans** the entire space. Further, by setting $a(1, 0) + b(0, 1) = 0$, the only way for this to happen is when $a = b = 0$ as we can directly observe from that $a(1, 0) + b(0, 1) = (a, b) = (0, 0)$, satisfying the criteria of **linear independence**. Now that we have established $\beta$ as a basis of $\mathbb{C}^2$, it can be deduced that the dimension of $\mathbb{C}^2$ is 2. And from a similar line of argument, it can be deduced that for any Hilbert space $\mathbb{C}^n$, its dimension will be exactly $n$.

### 2.1.4 Coordinates

Once we define a basis $\beta$ on $V$ we can rewrite a vector $v$ in terms of the coordinates that it takes on the particular basis. For example, if $\beta = \{v_1, v_2, ...v_n\}$ is a set of basis then $v$ can be uniquely decomposed into $v = c_1v_1 + c_2v_2 + ... + c_nv_n$, in which $c_1, c_2, ..., c_n$ is what we call the coordinates of $v$ on $\beta$. We write

$$[v]_\beta = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

Such a decomposition is exists because $\beta$ spans the entire space $V$, and the coordinates are unique for each $v$ because $\beta$ is linearly independent. Typically to find the coordinates of a vector on a basis we need to solve a linear set of equations, but if the basis is orthogonal we can quickly find the coordinates by doing projections.

### 2.1.5 Vector representation: Dirac notation and dual vectors

A more common way to represent vectors is by writing them as column vector $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ and row vectors $\begin{pmatrix} a & b & c \end{pmatrix}$. There is nothing different in the quantity they represent, but in the later chapter of inner product and linear transformations, it will make sense why this format is advantageous. Meanwhile, the Dirac notation is convention in quantum computing and quantum mechanics to represent column vectors and row vectors. For a row vector, instead of writing $u = \begin{pmatrix} x & y & z \end{pmatrix}$, we write:

$$\langle u| = \begin{pmatrix} x & y & z \end{pmatrix},$$

and for a row vector, instead of $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, we write

$$|v\rangle = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

so that when we refer to $\langle u|$ and $|v\rangle$, we know for sure that we are representing a vector, and whether it is a column vector or a row vector. The first notation $\langle ...|$ is called a bra, and the second $|...\rangle$ a ket, so together they form a bracket $\langle ...|...\rangle$.
Another rule given by this notation is the "adjoint" of a vector. If we originally denote $|\psi\rangle = \begin{pmatrix} a_1 \\ a_2 \\ ... \\ a_n \end{pmatrix}$, then $\langle\psi| = \begin{pmatrix} a_1 \\ a_2 \\ ... \\ a_n \end{pmatrix}^\dagger = \begin{pmatrix} a_1 & a_2 & ... & a_n \end{pmatrix}^* = \begin{pmatrix} a_1^* & a_2^* & ... & a_n^* \end{pmatrix}$, essentially turning the column vector that $|\psi\rangle$ represents into a row vector, and taking the complex conjugate of each of its entry, and vice versa.

*The more complete definition of the ket notation would be that if $\psi$ is a vector from $V$, then $|\psi\rangle$ is the coordinate vector of $\psi$ on an otn basis of $V$.

## 2.2 Inner product space (Hilbert space)

### 2.2.1 Canonical inner product

For a vector space $V$, an inner product, denoted as $\langle u, v\rangle$, is a function that takes two vectors $|u\rangle, |v\rangle$ from $V$ and produces a complex number in $\mathbb{C}$, such that it satisfies:

1. linearity with respect to the second argument

$$\left\langle u, \sum_i^n a_i v_i \right\rangle = \sum_i^n a_i \langle u, v_i \rangle$$

2. conjugate commutativity

$$\langle u, v \rangle = \langle v, u \rangle^*$$

3.

$$\langle u, u \rangle \geqslant 0, \text{ and } \langle u, u \rangle = 0 \text{ only when } |u\rangle = 0$$

Any vector space with a particular inner product defined on it is called a inner product space, or equivalently a hilbert space. Deriving from the definition of inner product is the idea of orthogonality, as well as a way of quickly decomposing any vector into coordinates on an orthonormal basis.

For vectors $|u\rangle = \begin{pmatrix} u_1 \\ u_2 \\ ... \\ u_n \end{pmatrix}$ and $|v\rangle = \begin{pmatrix} v_1 \\ v_2 \\ ... \\ v_n \end{pmatrix}$ from the space $\mathbb{C}^n$, we can define inner product $\langle u, v \rangle$ as:

$$\langle u, v \rangle = \sum_i^n u_i^* v_i = u_1^* v_1 + u_2^* v_2 + ... u_n^* v_n.$$

This definition of the inner product is the cannonical inner product on $\mathbb{C}^n$. It can be easily verified that the cannonical inner product on $\mathbb{C}^n$ satisfies all the aforementioned properties of an inner product. Now if we have a row vector and a column vector $\langle u| = \begin{pmatrix} u_1^* & ... & u_n^* \end{pmatrix}$ and $|v\rangle = \begin{pmatrix} v_1 \\ ... \\ v_n \end{pmatrix}$, we can multiply them together directly as if they were matrices:

$$\langle u|v\rangle = \begin{pmatrix} u_1 & ... & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ ... \\ v_n \end{pmatrix} = \sum_i^n u_i * v_i,$$

which turns out to be exactly equal the canonical inner product of the two vectors (note that a bar is dropped between $\langle u| \, and \, |v\rangle$ for conciseness). Then $\langle u|v\rangle$ has the double meaning of multiplying the adjoint of $|u\rangle$ and $|v\rangle$ together as well as taking the inner product of the two.

### 2.2.2 Norm

The norm of a vector $|\psi\rangle$, $\||\psi\rangle\|$, is defined as

$$\||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle}.$$

From the previous section we have known that the inner product of a vector with itself is larger or equal to 0, so the norm of a vector is always guaranteed to be a real number. We can

observe that under the canonical definition, the norm of $|u\rangle = \sqrt{u_1^* u_1 + u_2^* u_2 + u_3^* u_3 + ... + u_n^* u_n}$, which is by Pythagorean theorem exactly the length of $|u\rangle$.

Further, it can be proven that $\langle u|v\rangle = \||u\rangle\|\||v\rangle\| \cos(\theta)$, where $\theta$ is the angle between vector $|u\rangle$ and $|v\rangle$, using vector analysis. This gives us the intuition about the inner product that it is multiplying the length of one vector with the vector projection of the other one. In other words, besides taking into account the length (norm) of both vectors, it also reflects how much of them overlaps spatially.

### 2.2.3 Orthogonality

We say that two vectors $u, v \in V$ are **orthogonal** if $\langle u, v\rangle = 0$.
In the case of $\mathbb{C}^n$, two vectors $|u\rangle, |v\rangle$ are orthogonal if $\langle u|v\rangle = 0$. From the previous property, we can know when both $|u\rangle, |v\rangle$ have real entries, then they being orthogonal implies that they are also geometrically normal to each other.

For two vectors $u$ and $v$, the **projection** of $u$ onto $v$ is defined as

$$proj_v(u) = \frac{\langle u, v\rangle}{\langle v, v\rangle} v;$$

$proj_{|v\rangle}(|u\rangle)$ and $|u\rangle - proj_{|v\rangle}(|u\rangle)$ uniquely decomposes $|u\rangle$ into a vector parallel to $|v\rangle$ and a vector orthogonal to $|v\rangle$. Algebraically, we can check that

$$\begin{aligned}
\langle u - proj_v(u), v\rangle &= \left\langle u - \frac{\langle u, v\rangle}{\langle v, v\rangle} v, v \right\rangle \\
&= \langle u, v\rangle - \frac{\langle u, v\rangle}{\langle v, v\rangle} \langle v, v\rangle \\
&= \langle u, v\rangle - \langle u, v\rangle = 0
\end{aligned}$$

### 2.2.4 Orthonormal basis

For a set $S$ of non-zero vectors in $V$, if each pair of elements from this set are orthogonal, then $S$ is linearly independent:
Suppose $S = \{v_1, v_2, ..., v_n\}$, then if we let $a_1 v_1 + ... + a_n v_n = 0$, by taking inner product with $v_i$ on both sides of the expression, we have

$$\begin{aligned}
&a_1 \langle v_i, v_1\rangle + ... + a_n \langle v_i, v_n\rangle = \langle v_i, 0\rangle \\
\Rightarrow &a_1 0 + a_2 0 + ... + a_i \langle v_i, v_i\rangle + ... + a_n 0 = 0 \\
\Rightarrow &a_i \|v_i\|^2 = 0. \\
\Rightarrow &a_i = 0
\end{aligned}$$

An **orthonormal basis** of $V$, or abbreviated as otn basis, is an orthogonal set of vectors of norm 1 that also span the whole $V$ space. From the **Gram-Schmidt orthogonalization process**, it can be shown that for all vector spaces $V$, there must exist sets of otn basis,

which can be directly constructed from any basis of that space.

One of the most important advantage of an orthogonal basis is that it allows us to decompose any vectors quickly. For an otn basis $\beta = \{v_1, v_2, ..., v_n\}$ on $V$ and a vector $v$, then the $i$-th coordinate of $v$ on $\beta$ is simply:

$$c_i = \langle v_i, v \rangle .$$

We can verify that if $v = c_1 v_1 + c_2 v_2 + ... + c_n v_n$, then

$$\langle v_i, v \rangle = c_1 \langle v_i, v_1 \rangle + ... + c_i \langle v_i, v_i \rangle + ... + c_n \langle v_i, v_n \rangle = c_i \|v_i\| = c_i.$$

### 2.2.5 Dirac's notation revisited

Now that the otn basis has been explained, it can be revealed that the full meaning of the ket $|\rangle$ notation is not just denoting a column vector. If we have a vector $\psi$ from an inner product space $V$, and an otn basis $\beta$ implicitly defined on $V$, then $|\psi\rangle$ actually represents the coordinates of $\psi$ on $\beta$, namely $|\psi\rangle = [\psi]_\beta$.

The elegance of this notation is that, for any two vectors $u, v \in V$, we have the relation $\langle u, v \rangle = \langle u|v \rangle$, namely the inner product of $u$ and $v$ is equal to the canonical product of $|u\rangle$ and $|v\rangle$. The proof will be abbreviated here, but the central idea is that $u$ and $v$ can be first decomposed on $\beta$, and then if we perform the inner product entry wise, the only terms that remains are those that form $\langle u|v \rangle$.

We can likewise perform all other operations such as linear combinations ($|u + v\rangle = |u\rangle + |v\rangle$), linear transformations ($|T(\psi)\rangle = A_T |\psi\rangle$), or taking coordinates on other bases, with $|V\rangle$, so that instead of having to do complicated algebra on linear space $V$, which can be polynomials, matrices, or the real function space, everything are reduced to entry-wise algebra on complex numbers.

### 2.2.6 Outer product

## 2.3 Linear transformations

If a function $T$ mapping from linear space $V$ to linear space $W$, satisfies that for $u, v \in V$ and $c \in \mathbb{C}$:

1. $T(cv) = cT(v)$
2. $T(u + v) = T(u) + T(v)$,

then it is considered to be a linear transformation. By induction we can quickly show that if $T$ is a linear transformation, then

$$T(c_1 v_1 + c_2 v_2 + ... + c_n v_n) = c_1 T(v_1) + c_2 T(v_2) + ... + c_n T(v_n)$$

### 2.3.1 Matrix representation: Matrix as operators

An important theorem establishing toward the matrix representation of linear transformations is that a linear transformation can be completely defined by the value it takes on a basis of its input space.
Namely, if we know that a linear space $V$ has basis $\beta = \{v_1, v_2, ..., v_n\}$, and $T$ a linear transformation on V, once we know the value of $T(v_1) = w_1, T(v_2) = w_2, ..., T(v_n) = w_n$, the value of $T(v)$ for all $v \in V$ can be obtained through linear combination:

$$\begin{aligned} T(v) &= T(c_1v_1 + c_2v_2 + ... + c_nv_n) \\ &= c_1T(v_1) + c_2T(v_2) + ... + c_nT(v_n) \\ &= c_1w_1 + c_2w_2 + ... + c_nw_n, \end{aligned}$$

where $c_i$ is the i-th coordinate of $v$ on $\beta$.

Due to this nature of linear transformations, we can rewrite a linear transformation mapping over the same space $T : V \to V$ by the value it takes on an otn basis $\beta = \{v_1, v_2, v_3, ..., v_n\}$ of $V$:

$$A_T = \begin{pmatrix} w_{11} & w_{12} & ... & w_{1n} \\ w_{21} & w_{22} & ... & w_{2n} \\ ... & & & ... \\ w_{n1} & w_{n2} & ... & w_{nn} \end{pmatrix} = \left( |w_1\rangle \ |w_2\rangle \ ... \ |w_n\rangle \right),$$

where

$$\begin{aligned} w_1 &= T(v_1) \\ &... \\ w_n &= T(v_n), \end{aligned}$$

and $|w_i\rangle$ representing the coordinates that $w_i$ takes on $\beta$ with $|w_i\rangle = \begin{pmatrix} w_{i1} \\ w_{i2} \\ ... \\ w_{in} \end{pmatrix}$. To compute the linear transformation $T(v)$, we take the coordinate vector of $v$ on $\beta$: $|v\rangle = \begin{pmatrix} c_1 \\ ... \\ c_n \end{pmatrix}$, and perform matrix multiplication

$$\begin{aligned} A_T |v\rangle &= \left( |w_1\rangle \ |w_2\rangle \ ... \ |w_2\rangle \right) \begin{pmatrix} c_1 \\ ... \\ c_n \end{pmatrix} \\ &= \begin{pmatrix} w_{11} & w_{12} & ... & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & ... & w_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ ... \\ c_n \end{pmatrix} \\ &= c_1 |w_1\rangle + ... + c_n |w_n\rangle \\ &= |c_1w_1 + ... + c_nw_n\rangle = |T(c_1v_1 + ... + c_nv_n)\rangle \\ &= |T(v)\rangle . \end{aligned}$$

As demonstrated by the derivation above, what we end up with by computing $A_T |v\rangle$ is not exactly $T(v)$, but its coordinates $|T(v)\rangle$ that are taken on $\beta$, which can be quickly recovered through a linear combination.
While matrix representation of $T : V \to V$ always has the same dimensions in the input space and the out space, it is always a square matrix. The definition of matrix representation of linear transformations described above is limited to $T$ mapping over the same space and an otn basis on $V$, in a more generalized definition, $A_T$ can represent $T : V \to W$ where linear spaces $V$ and $W$ can be different; further, the basis on which the matrix is defined doesn't have to be orthonormal. However the general definition of matrix as linear transformations wouldn't be expanded here in details since quantum mechanics has simplified the mathematics with some tricks to avoid its use.

### 2.3.2 The identity matrix

The identity transformation $I$ defined on $V$ is a linear transformation that maps a vector $v \in V$ to it self. Looking at the input and output vector by its basis, we notice that the i-th basis vector $v_i$ when put into $I$ produces exactly $v_i$, and hence we have the identity matrix representing identity transformations:

$$I = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

### 2.3.3 Eigenvectors & eigenspaces

Before we head into our main topic on Eigenvectors, we hope to take some time and put our thoughts together. Things have gone wild here. So the key idea is that we are able to form a corresponding matrix given the limited inputs and outputs of some vectors, or the basis, as you notice from the exercises. Linearity allows the matrix to perform our desired operations on the entire linear space. Clearly, something is going on here. The matrix is able to produce much more useful information compared with what we have endowed it. You might ponder, in a mathematical sense, it appears to be a manipulation of numbers. That's all. But when someone asks you what music means to you, will you answer a manipulation of notes? The fundamental mechanism of linear transformations lies in the mathematical structures of matrices, the algebraic structures as well as the geometric structures. To give a clear sense of what matrices are doing to those innocent vectors, we are to take advantage of some visualization tools. Let's start with R2, one of the simplest Euclidean space.

(1 0 0 1) Put the introduction of the identity matrix down here.

(2 0 0 3)

We can see from the graphs that all the horizontal and vertical lines remain horizontal and vertical after the transformation. That means the basis vectors are only stretched by some factor without experiencing any rotation. (2 0 0 3)(1 0) = (2 0) = 2* (1 0) (2 0 0 3)(0 1) = (0 3) = 3* (0 3) Imagine Alice is sitting at (0,0), and Bob is sitting at (0,1). Then the

transformation happen. Alice is still sitting at (0,0), while Bob now sits at (0,3). Bob moves three times as far from Alice as the unit distance along the same direction. Note that Bob is only SITTING there. He is not doing anything. The space itself is expanding.

(Another vis-example would be nice, say (2 0 0 -3))

There' something in common of all the cases. The geometric property of the linear transformations that we have demonstrated is associated with the fact that the off-diagonal elements of the matrix operator are all zero. We would like to call them diagonal matrices. Take a look at this funny matrix (1 1 0 1). (1 1 0 1)(1 0) = (1 0) (1 1 0 1)(0 1) = (1 1) While the horizontal lines stay still, the vertical lines rotate by an angle of $\pi/4$ due to the appearance of an off-diagonal element. You have to admit that you are not very comfortable with rotations. If you take a minute to experience the transformation process, you will soon get dizzy with all the angles and axes floating around. Nevertheless, those stretches put our mind at ease! And that is because we only need a scalar quantity to identify stretches whereas we need a vector to identify rotations. Of course, stretches are generally rare, thus more valuable. The vectors that only experience stretching can tell us something fundamental about the structure of the transformation. We define them as the "characteristic vectors", or the "eigenvectors", of the matrix, given:

M $|\psi\rangle = \lambda$ * $|\psi\rangle$, where $\lambda$ is the eigenvalue of that corresponding eigenvector $|\psi\rangle$

For generalized diagonal matrices, (Lumda1, lumda2, lumda3,... lumdaN), The eigenvectors alongside their eigenvalues are clear:

In our context, we only consider matrices that satisfy l1*l2*l3*... lN not equal 0. Otherwise, the space will lose dimensions of information. That contradicts to our assumption of a full-rank Matrix, or a Matrix that maps Vn to itself. under such transformation, the orthonormal basis will stay orthonormal to one another.

Remember when we used to talk about vector spaces. The vectors that we deal with in this textbook are not necessarily the vectors built up in your head, with arrows, magnitudes, and stuff. Any mathematical quantities satisfying the eight gold rules could be considered vectors, and the laws of linear algebra could be applied to that whole new system. For instance, let's examine function space as well as its "derivative" operator:

Quite surprised, isn't it? Try and take a peak at this ordinary differential equation: $d/dt(f) = a * f$. Notice that the format is exactly the same as our definition of eigenvectors. Solving this differential equation means finding the eigenvectors of the operator matrix with eigenvalue $a$. The way we approach functions becomes totally different now.

Let's take a step back and examine our diagonal matrix closely. If m of the n eigenvalues are the same, say, $_0$ these corresponding eigenvectors form an orthonormal basis of an m dimensional linear sub-space within the entire vector space. All the vectors in the sub-space are all eigenvectors with eigenvalue $_0$. Observe:

*Exercises and Checkpoint

At this point, you might wonder that since an arbitrary operator could create any effect on the linear space, it's just normal that we encounter cases with the same geometric pattern but we made a wrong choice of our basis so that the matrix contains off-diagonal elements. By switching the basis, i.e. performing a separate linear transformation on the space, we could then apply the diagonal matrix to perform our desired transformation. Finally, we return back to our original basis by using the inverse matrix. See Appendix: Matrix Diagonalization. $M = T^(-1)DT$

### 2.3.4 Unitary matrices

A square matrix $A$ is **unitary** if $A^\dagger A = AA^\dagger = I$. In addition, all of the following are equivalent statements:

1. $AA^\dagger = A^\dagger A = I$
2. $\langle Ax|Ay\rangle = \langle x|y\rangle \, \forall x, y \in \mathbb{C}^n$
3. A transforms any otn basis into another otn basis
4. A transforms some otn basis into an arbitrary otn basis
5. $\|Ax\| = \|x\| \; \forall x \in \mathbb{C}^n$ (norm preserving).

A unitary operation is a linear transformation defined by a unitary matrix, it can be thought of as a "rotation" of the vector space. If a cube undergoes a unitary change, it will remain as a cube of the same size but in a different orientation. The rotation doesn't have to preserve symmetry though; for example a left hand can undergo a unitary change and become a right hand. **All matrices representing quantum gates are norm preserving, due to the necessity that the total probability of all possible states must add up to 1, and hence they are all unitary matrices.**
A convenient way to identify whether a matrix is unitary is by observing its column vectors. A matrix is unitary if and only if its column vectors form an orthonormal basis. (The proof shall be left as an exercise.)

### 2.3.5 Properties of Hermittian (Self-adjoint) matrices

## 2.4 Appendix

### 2.4.1 Matrix multiplication

### 2.4.2 Complex conjugate

*** First Section Complex vector space. ** Introduce the complex number, the complex plane, complex conjugate, the exponential form (norm and phase angle)

** Define the complex vector A vector $|v\rangle = a_1 |s_1\rangle + a_2 |s_2\rangle + \cdots + a_n |s_n\rangle \in V$ is a linear combination of vectors in set $S = \{|s_1\rangle, \ldots, |s_n\rangle\}$. **The set $S$ generates a complex vector space with even dimension. The real dimension of the vector space Hilbert n is 2*n-2.(will discussed later) $V$. **A set of vectors $B$ for which every element of $V$ can be written uniquely as a linear combination of vectors in $B$ is called a *basis* for $V$. (QM bases are usually assumed orthonormal). ** Hilbert Space
Define orthogonality and inner product space using the ad-joint operator: taking the complex conjugate of each component of the complex vector and then transposing the matrix.

Inner product $\langle v_1|v_2\rangle = \langle v_2|v_1\rangle$ is *non-negative*. Orthonormal $\langle \beta_1|\beta_2\rangle = 0$. Length $\langle v_1|v_1\rangle$.

Basis $\{|\beta_1\rangle, |\beta_2\rangle\}$. Ket $|v\rangle = a\,|\beta_1\rangle + b\,|\beta_2\rangle \doteq \begin{pmatrix} a \\ b \end{pmatrix}$. Kets (bras) are represented by column (row) matrices, with the relation $|v\rangle = \langle v|^\dagger$.

# 3 Single qubit

After a detailed overview of linear algebra, we are now ready to tackle the implementations of such math methods, revealing the building blocks of quantum computation. Unlike classical computers, quantum computers utilized quantum bits, or qubits, as fundamental unit of information processing. Their unique properties governed by modern quantum mechanics enable quantum systems to yield results in higher dimensions and solve problems with increasing computational complexity.

Let's first start working with the simplest case, the computation of Single Qubit. Even though the broader applications of single qubit are limited, important aspects that facilitate a quantum network will be discussed in this chapter. We are about to see something so interesting about qubits that inspires us as we move forward to more complicated cases of qutrits, qudits, etc.

## 3.1 Superposition

The fundamental difference between quantum systems and classical ones lies on the basic unit structure. While classical bits are generally described with discrete states, qubits exhibit a continuum of states, which physicists refer as "superposition".

To align with the computation methods, we choose to describe the systems using spin states. In a classic computer, bits are represented using binary (either 0 or 1), which is dependent on potential differences within a circuit. Low voltage translates to a 0, and higher voltage translates to 1. However, when looking at bits from a quantum perspective, a bit can be represented by either the spin-up state or a spin-down state of each electron, which corresponds to the value 0 and 1, respectively. Quantum systems can also be described with Quantum Spin States, a superimposed result of the two classical states. We now arrive at a critical point in this textbook where we address one of the major concerns for beginners: the realization of qubits in mathematical sense. The quantum spin state, as well as the classical spin states, is a vector sitting in a two dimension complex state space, $\mathbb{C}^2$. It is a linear combination of the two classical state vectors, which form a pair of orthonormal basis, denoted $|0\rangle$, $|1\rangle$, respectively.

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle,$$

where a and b are complex coefficients, and $|\psi\rangle$ is the superimposed state, or the wave function.

Something worth noticing, the descriptions here are pure math models. Like an old saying on electron spins, "imagine an electron spinning, except there is no electron and it's not spinning." Indeed, qubits are not real quanta. They could represent any quantum mechanical systems that can be modeled by the given method. Also, the spin is not the actual angular momentum quantity that characterizes rotation.

## 3.2 Checkpoint

Determine if the following are valid qubit states:

1. $|0\rangle$
2. $\frac{i}{\sqrt{2}}|0\rangle + \frac{-1}{\sqrt{2}}|1\rangle$
3. $\frac{i+1}{\sqrt{3}}|0\rangle + \frac{i-1}{\sqrt{3}}|1\rangle$

[1]Any device that measures a two-state quantum system must have two preferred states $\{|u\rangle, |u^{\perp}\rangle\}$. Must transform state to measurement basis.

## 3.3 Measurement

So far, we took a peak at the math models we used to describe qubits, but we haven't fully deciphered the secret of superposition. We know the classic states very well but what does it mean to be in a superimposed quantum state? The next important idea is one of the initial postulates that give birth to quantum mechanics. It bridges the gap between bits and qubits, and masks the classical vs. quantum nature of probability. The idea is associated with measurement and still remains debatable among modern quantum physicists.

Now let's talk math. Suppose we have a wave function $|\psi\rangle$ as a linear combination of an arbitrary basis pair $|0\rangle$ and $|1\rangle$. Note that there are infinite many choices of the basis pair that represent quantum states. What we do here is we take the wave function $|\psi\rangle$ and project it onto one of the basis pair, say, $|0\rangle$.

$$|\psi\rangle \rightarrow |0\rangle$$

This act of projection is in fact quantum measurement. It collapses the wave function $|psi\rangle$ and transforms it into classic states. Of course, we can go back in time and choose the other basis pair, $|1\rangle$, and make the projection.

$$|\psi\rangle \rightarrow |1\rangle$$

Something you might wonder earlier, since the qubit can take on any states in $\mathbb{C}^2$, does it mean that a single qubit could store a lot of classical information? In fact, the properties of quantum measurement *SEVERELY* restrict the amount of information that can be extracted. If a state $|psi\rangle = a'|u\rangle + b'|v\rangle$ is measured as $|u\rangle$ then the state $|psi\rangle$ changes to $|u\rangle$. It might be counter-intuitive that a single measurement will permanently change the state, making it impossible to determine the original state from any sequence of measurements. The fact that one cannot make two separate measurements on the original state of the qubit suggests that single quantum measurement yields at most one classical bit of information.

A critical idea here is that when we initiate a measurement of a quantum spin state, we have to select a particular set of orthonormal basis, along which the measurement is performed. Any measurement is worthless without identifying the basis of the measurement. Once we have the basis pair, we are ready to perform the measurement. However, as you might foresee, we are unable to predict whether the measurement result will be one or

[1]9/11/19

the other of the basis pair. The language we use in quantum measurement is probability. So now we are going to take advantage of our defined **inner product space** and extract useful information out of the measurement process. We are going to calculate the inner product between the vectors and the nice thing about inner products is that we obtain scalar quantities, which could be easily manipulated. $\langle|\, 0, |\, \psi|\, 0, |\, \psi =$ matrix multiplication

We associate the inner product result of the basis and $|\psi\rangle$ with the probability of the measurement result being that particular basis. Note that the inner product on $\mathbb{C}^2$ is non-commutable. To resolve this ambiguity, we rigorously define the probability to be the norm squared of the inner product results.

$$p(|\psi\rangle \to |0\rangle) = |\mathrm{a}|^2$$
$$p(|\psi\rangle \to |1\rangle) = |\mathrm{b}|^2$$

**The conservation of total probability** implies a hidden relationship between the arbitrary co-efficients $a$ and $b$:

$$|\mathrm{a}|^2 + |\mathrm{b}|^2 = 100 \text{ percent.}$$

This means that there is a normalization constraint on all the quantum spin state vectors in $\mathbb{C}^2$. Whether or not you have experience with complex number before reading this section, the complex format of this constraint equation is quite unsettling. It's just natural for us to put real numbers into context. We know that any complex numbers can be represented by two real numbers, which entitle two degrees of freedom. Accordingly, the vector $|\psi\rangle$ appears to have four degrees of freedom as we need four independent real variables to determine the state. Nevertheless, the normalization constraint above reduces the number to three. The math might not seem so clear at this point, but do hold on to that thought. It will all come to light once we start visualing the qubits.

## 3.4 Checkpoint

Let $|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{i}{\sqrt{2}}|1\rangle$. Keeping in mind that $\langle\psi|$ is the conjugate transpose of $|\psi\rangle$, and that $\langle\psi_0|\psi_1\rangle$ is the inner product between $\psi_1$ and $\psi_2$, calculate the following:

1. $\langle\psi|0\rangle$
2. $\langle\psi|1\rangle$
3. $\langle\psi|\psi\rangle$

Prove the basis pair vectors yield 0 inner product result, i.e. it's impossible to measure one of the basis to be the other./. Prove the basis paring is unique.

Example:

$$\left.\begin{array}{l}\text{Cat: } 0 \quad \text{or} \quad 1 \\ \text{Qubit: } a\,|0\rangle + b\,|1\rangle\end{array}\right\} \xrightarrow[\text{Measurement}]{} \left\{\begin{array}{l}p^2 \\ |a|^2\end{array}\right.$$

The exercises shed light on the relationship between different spin states, leading to our next idea of "independent states". So if a measurement on a wave function $|\psi\rangle$ along a basis pair yields equal probability of 1/2, then the wave function and the basis are referred

as independent states. It was inspired by the Heisenberg's Uncertainty Principle(footnote) regarding the measurement of position and momentum of an electron. As one of the states collapses, the uncertainty of the other state is infinity. Our intuition from real space is that we can only find one unique basis pair to form independent states with a given basis pair. However, the case appears differently in $\mathbb{C}^2$.

These are the preferred notation for the three defined basis pairs that intrigue us. A straight-forward calculation of inner products tells us that these basis states in $\mathbb{C}^2$ are mutually independent under our definition. In the next section, we will be introducing the Bloch's sphere visualization, which employs the method of mapping from $\mathbb{C}^2$ to $\mathbb{R}^3$. wThis is the key of understanding the realization of single qubit, and opens various possibilities to the computation and transformation of a single qubit.

## 3.5 Example: Photon polarization

Interaction of polaroid with a photon $\implies$ probabilistic outcomes *and* effect of the measurement on the state of the system.

Imagine you have a light source producing a beam of light on a projection screen. Now let's add a polaroid $A$ in between (let's assume polarization of the polaroid is horizontal)

$\implies$ intensity of light is reduced

Let's add a polaroid $C$ which is rotated so that its polarization is orthogonal (vertical) to polaroid $A$.

$\implies$ no light will reach the screen

Let's add polaroid $B$ in between $A$ and $C$. Surprisingly at most polarization angles of $B$ (maximal if set at $45^\circ$) there will be light.

- For a bright beam of light there is a classical explanation in terms of waves
- We can repeat this for only one photon at a time $\implies$ the results can only be described with quantum mechanics:
    1. A model of a photon's polarization
    2. A model of the interaction between a polaroid and a photon.

QM models a photon polarization state by a unit vector $|\updownarrow\rangle, |\leftrightarrow\rangle$.
$\implies$ Arbitrary vector: $|v\rangle = a\,|\updownarrow\rangle + b\,|\leftrightarrow\rangle$.
e.g. $|45^\circ\rangle = \frac{1}{\sqrt{2}}\left[|\updownarrow\rangle + |\leftrightarrow\rangle\right]$

## 3.6 Mapping from $\mathbb{C}^2$ to $\mathbb{R}^3$ (Principle behind Bloch sphere)

The assumptions under such mapping All mappings happen in linear space. The three basis pairs in C2 are mapped to the three main axis in R3 accordingly. X and Z-axes are real while y-axis is imaginary.

***What is counter-intuitive? The orthogonal vectors become parallel. A normalized complex state has three degrees of freedom, whereas a normalized real state only has two. What happened to the third parameter?

***Visualization: Bloch Sphere ** The building blocks of the Bloch Sphere Model:

### 3.6.1 Defining the global phase and the local phase

We know for an arbitrary state vector $|\psi\rangle \in \mathbb{C}^2$, we can represent it with $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$, $a, b \in \mathbb{C}$. Further, we can write $a$ and $b$ in their complex polar form $a = Ae^{i\alpha}, b = Be^{i\beta}$. Then $|\psi\rangle$ becomes:

$$\begin{aligned} |\psi\rangle &= Ae^{i\alpha}\,|0\rangle + Be^{i\beta}\,|0\rangle \\ &= e^{i\alpha}(A\,|0\rangle + Be^{i\beta-\alpha}) \end{aligned}$$

Because $|\psi\rangle$ is normalized, we can set $A = \cos(\delta)$ and $B = \sin(\delta)$. Meanwhile, we can use $\gamma = \beta - \alpha$ to re-express $|\psi\rangle$ as:

$$|\psi\rangle = e^{i\alpha}(\cos(\delta)\,|0\rangle + \sin(\delta)e^{i\gamma}\,|0\rangle)$$

Notice that to define such a state vector, we need at least three real parameters: $\alpha, \delta$, and $\gamma$. But $e^{i\alpha}$ is only a factor in front which gets preserved when we perform linear transformations on $|\psi\rangle$. Changing $\alpha$ also wouldn't alter the probability superposition which this vectors represents. Take any state $|x\rangle$, the probability of $|\psi\rangle$ being $x$ is

$$|\langle x|\psi\rangle|^2 = |e^{i\alpha}\langle x|...\rangle|^2 = |e^{i\alpha}|^2|\langle x|...\rangle|^2 = |\langle x|...\rangle|^2,$$

for that $|e^{i\alpha}|$ is always 1.

Since $e^{i\alpha}$ doesn't actually affect the probability of individual states, it can be reasonably discarded when we visualize $|\psi\rangle$ in a Bloch sphere on $\mathbb{R}^2$. Because $e^{i\alpha}$ exists as a common factor, we call $\alpha$ the global phase of $|\psi\rangle$, while $\gamma$ the local phase; both correspond to complex phases.

This suggests that the local(relative) phase will be the polar angle in R3. On the other hand, the global phase angle should not be represented on the Bloch sphere as it makes no observable difference to psi.

Hence the result, the real dimension of the vector space Hilbert n is 2*n-2.

** Need explicit statement then to draw conclusions as to why wave functions are written in this form:

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

### 3.6.2 The Pauli matrices and their eigenvectors

The matrices, or quantum gates, are linear transformations that act on complex 2D space, changing qubits into different states so they yield different measurement outcomes. Three common gates are:

Pauli-Z:

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

Pauli-X, the Not gate:

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

Pauli-Y:

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

## 3.7 Checkpoint

1. The eigenstates of the Pauli-Z gate are $|0\rangle$ and $|1\rangle$. Show this to be true.
2. The eigenstates of the Pauli-X gate are represented by $|+\rangle$ and $|-\rangle$. Solve for the eigenstates and match them to their labels.
3. The eigenstates of the Pauli-Y gate are represented by $|\circlearrowright\rangle$ and $|\circlearrowleft\rangle$. Solve for the eigenstates, and using knowledge of how gates transform vectors on the Bloch sphere, match the eigenstates to their labels.
4. Each gate represents a rotation by $\pi$ around some axis on the Bloch sphere. Either by using the last three problems or by visualization, figure out which axis this is for each gate.

** Visual accompaniment of what each gate does to a qubit on the bloch sphere would be good here, too.

### 3.7.1 The Mapping

*Define an index set $\sigma_x$, $\sigma_y$, $\sigma_z$ consisting of the pauli gates on the three directions. For each point on the Bloch sphere, we can mark it with coordinate vector $(x, y, z)$. The points along $x, y, z$ axis are not of particular interest. Each of them correspond to multiples of one of the Pauli-matrices. For an arbitrary point, we can construct the generalized matrix $S = x\sigma_x + y\sigma_y + z\sigma_z$. Bringing in the entries of the three matrices we have:

$$S = x\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + z\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} z & x - yi \\ x + yi & -z \end{pmatrix}$$

We have by now mapped the three coordinates in $\mathbb{R}^3$ (Bloch sphere) to a linear subspace of matrices. The subspace is spanned by pauli-x, pauli-y, pauli-z, the three linearly independent vectors in this matrix space. This 3D Pauli matrix space is a subspace of the 4D matrix space of all 2 by 2 matrices. Pauli S is given by a unique linear combination of paulix, pauliy, pauliz, with coefficients of the unit vector V in the x, y, z basis in R3!

To finish the mapping from C2 to R3, we find the (normalized) eigen vector of

$$S = \begin{pmatrix} z & x - yi \\ x + yi & -z \end{pmatrix}.$$

First to find the eigen values:

$$\begin{aligned} f_S(\lambda) &= (z - \lambda)(-z - \lambda) - (x - yi)(x + yi) \\ &= -z^2 + \lambda^2 - x^2 - y^2 \end{aligned}$$

Because $x^2 + y^2 + z^2 = 1$, the characteristic polynomial becomes:

$$\begin{aligned} f_S(\lambda) &= -z^2 + \lambda^2 - (1 - z^2) \\ &= -z^2 + \lambda^2 - 1 + z^2 \\ &= \lambda^2 - 1 = 0 \end{aligned}$$

Hence the eigen values of $S$ are 1 and -1. Further, we see that if we perform adjoint on $S = \begin{pmatrix} z & x - yi \\ x + yi & -z \end{pmatrix}$, we can still get $\begin{pmatrix} z & x - yi \\ x + yi & -z \end{pmatrix}$. That means $S$ is self-adjoint, which ensures it to be diagonalizable on an orthonormal basis. In consequence, we know that not only can we find a pair of eigen vectors to this matrix, they are also guaranteed to be orthogonal to each other! But in the end we want to obtain an one-to-one mapping between the Bloch sphere and the complex 2D space, so the eigen vector we take is the one corresponding to the eigen value 1 with norm 1.
You may still wonder, how is this one-to-one? After all, if $\mathbf{v}$ is an eigen vector of norm 1, $c\mathbf{v}$ is also an eigen vector of norm 1 as long as $|c| = 1$. And in complex field we know there is a "unit circle" of $c$ that we can pick from. But this goes back to our initial investigation of the degree of freedom in our representations of $|\psi\rangle$, and this $c$ is exactly the $e^{i\alpha}$ that we can discard.

Because $x^2 + y^2 + z^2 = 1$, if we look at each column of $S$, we can see that

$$\left| \begin{pmatrix} z \\ x + iy \end{pmatrix} \right| = z^2 + (x - iy)(x + iy) = z^2 + x^2 + y^2 = 1,$$

$$\text{and } \left| \begin{pmatrix} x - iy \\ z \end{pmatrix} \right| = (x + iy)(x - iy) + z^2 = z^2 + x^2 + y^2 = 1.$$

$$\text{further, } \left\langle \begin{pmatrix} z \\ x + iy \end{pmatrix} \middle| \begin{pmatrix} x - iy \\ z \end{pmatrix} \right\rangle = zx + iyz - zx - iyz = 0$$

.

This shows us that the columns of $S$ form an otn basis, and hence making it a unitary matrix. Besides the many nice properties of a unitary matrix, we know most fundamentally it transforms a vector of norm $n$ to another vector of norm $n$ — it preserves the size of vectors. *(Intuition about diagonalization)

Appendix: Matrix Operators

**Hadamard Gate from Z-basis to X-basis **Hadamard Gate from Z-basis to Y-basis **Hadamard Gate from X-basis to Y-basis

## 3.8 Single qubit: Bloch sphere

[2]Spin:

$$H_0 = \frac{e}{m} B_0 \hat{S}_z \tag{1}$$

[2]9/23/19

**Not enough introduction to dive in like this. Small definitions and visualization of spin states would benefit the reader here.
Energy level splitting due to $B$-field: $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. Rotation matrices:

$$R_x(\theta) = \begin{pmatrix} \cos\theta & -i\sin\theta \\ i\sin\theta & \cos\theta \end{pmatrix}, \qquad R_z(\theta) = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

On the bloch sphere, $|0\rangle$ and $|1\rangle$ are at the north and south poles respectively. The $\theta$ polar and $\varphi$ azimuthal angles follow typical definitions for spherical coordinates. Arbitrary state:

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

E.g. $\pi$-pulse, pulse length $\Delta t$ defined by

$$\omega\Delta t = \Delta\varphi = \pi \implies \Delta t = \frac{\pi}{\omega}.$$

Need to be able to do Bloch-sphere rotations about multiple axes. Introduce perturbation to spin Hamiltonian:

$$\hat{H} = \frac{e}{m}B_0\hat{S}_z + \frac{e}{m}B_1\cos\omega_0 t\hat{S}_x \quad \text{where} \quad B_1 \ll B_0$$

$$= \frac{e\hbar}{2m}\begin{pmatrix} B_0 & B_1\cos\omega_0 t \\ B_1\cos\omega_0 t & -B_0 \end{pmatrix}$$

** Explanation of why a factor of $\frac{\hbar}{2}$ should be added at beginning of chapter 3.
** An intermediate step should be added showing that the reason why the matrix takes this form is because of the pauli-x and pauli-z gates.

Since the second term is treated as a perturbation, $x$-rotations are much slower than $z$-rotations. Time-dependent state:

$$|\psi(t)\rangle = \cos\frac{\omega_1 t}{2}|0\rangle + e^{i(\omega_0+\pi)t}\sin\frac{\omega_1 t}{2}|1\rangle$$

$$\text{where} \quad \begin{cases} \theta = \omega_1 t \\ \omega_1 = \frac{eB_1}{2m} \end{cases}$$

## 3.9 Arbitrary 2-level quantum system

**Understanding this chapter is contingent on there being a section at the beginning of chapter 3 defining spin states so that the reader doesn't get lost looking at symbols.
[Cohen-Tannoudji, Diu, Laloe; Quantum Mechanics Vol. 1] Most general $2 \times 2$ Hermitian matrix:

$$\hat{H} = \frac{1}{2}\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \quad \text{where} \quad H_{12}^* = H_{21}, \quad \text{and} \quad H_{11}, H_{22} \in \mathbb{R}$$

$$\text{Take} \quad H_{11} > H_{22} \quad \text{without loss of generality.}$$

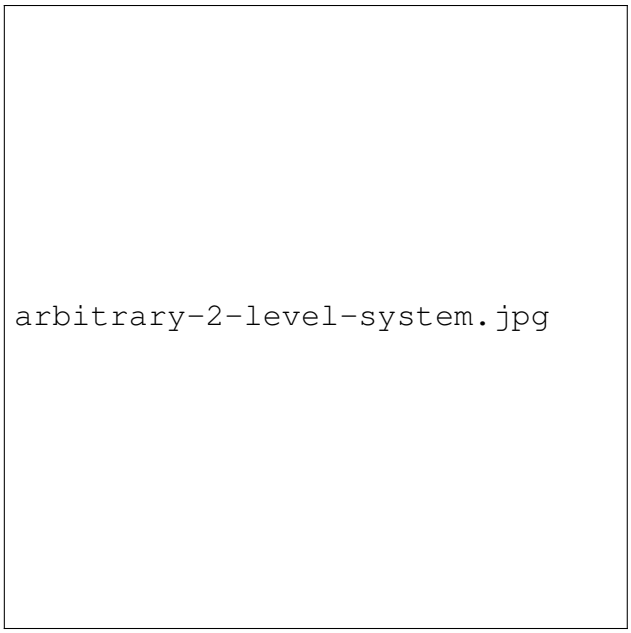



Figure 1: Eigensystem of a general 2-level quantum system on the Bloch sphere. $|\varphi_\pm\rangle$ are the eigenstates of the pure-diagonal Hamiltonian (e.g. the spin Hamiltonian (1)), while $|\psi_\pm\rangle$ are the eigenstates of the general Hamiltonian with off-diagonal elements.

Move zero of energy to midway between the levels:

$$\begin{aligned} H \to \hat{H} - \frac{H_{11}+H_{22}}{4}\otimes\mathbb{1} &= \frac{1}{2}\begin{pmatrix} \frac{H_{11}-H_{22}}{2} & H_{12} \\ H_{21} & \frac{-H_{11}+H_{22}}{2} \end{pmatrix} \\ &\equiv \frac{1}{2}\begin{pmatrix} \epsilon & \Delta - i\tilde{\Delta} \\ \Delta + i\tilde{\Delta} & -\epsilon \end{pmatrix} \\ &= \frac{\epsilon}{2}\hat{\sigma}_z + \frac{\tilde{\Delta}}{2}\hat{\sigma}_y + \frac{\Delta}{2}\hat{\sigma}_x \end{aligned}$$

Take $H_0 = \frac{1}{2}\begin{pmatrix} \epsilon & 0 \\ 0 & -\epsilon \end{pmatrix} = \frac{\epsilon}{2}\hat{\sigma}_z$ with eigenvectors $|\varphi_\pm\rangle$,

$$H_0 |\varphi_\pm\rangle = \pm\frac{\epsilon}{2}|\varphi_\pm\rangle,$$

as the north and south pole of the Bloch sphere (see Fig. 1. Then introduce the off-diagonals

$$H = H_0 + H_1, \quad \text{where} \quad H_1 = \frac{1}{2}\begin{pmatrix} 0 & \Delta - i\tilde{\Delta} \\ \Delta + i\tilde{\Delta} & 0 \end{pmatrix}.$$

Solve for the eigensystem:

$$\begin{gathered} H|\psi_\pm\rangle = E_\pm|\psi_\pm\rangle \\ \det\begin{pmatrix} \epsilon - 2E_\pm & \Delta - i\tilde{\Delta} \\ \Delta + i\tilde{\Delta} & -\epsilon - 2E_\pm \end{pmatrix} = 0 \end{gathered}$$

$$E_\pm = \pm\frac{1}{2}\sqrt{\epsilon^2 + \Delta^2 + \tilde{\Delta}^2}$$

$$\equiv \pm\frac{1}{2}\hbar\omega_q$$

$$|\psi_+\rangle = e^{-i\varphi/2}\cos\frac{\theta}{2}|\varphi_+\rangle + e^{i\varphi/2}\sin\frac{\theta}{2}|\varphi_-\rangle$$

$$|\psi_-\rangle = e^{-i\varphi/2}\sin\frac{\theta}{2}|\varphi_+\rangle + e^{i\varphi/2}\cos\frac{\theta}{2}|\varphi_-\rangle$$

This orients the $|\psi_\pm\rangle$ axis at polar angle $\theta$ and azimuthal angle $\varphi$ relative to the $|\varphi\rangle$ axis (see Fig. 1), where

$$\tan\theta = \frac{\sqrt{\Delta^2 + \tilde{\Delta}^2}}{\epsilon}$$
$$\tan\varphi = \frac{\tilde{\Delta}}{\Delta}$$

## 3.10 Spin in a generalized magnetic field

We can map the results of Sec. 3.9 to the spin problem as follows:

$$\begin{aligned}
|\uparrow\rangle &\longrightarrow |\varphi_+\rangle \\
|\downarrow\rangle &\longrightarrow |\varphi_-\rangle \\
|\uparrow\rangle_y &\longrightarrow |\psi_+\rangle \\
|\downarrow\rangle_y &\longrightarrow |\psi_-\rangle \\
\hbar\gamma B &\longrightarrow \hbar\omega_q \\
B_z &\longrightarrow \epsilon \\
B_x &\longrightarrow \Delta \\
B_y &\longrightarrow \tilde{\Delta}
\end{aligned}$$

$$H = -\frac{\gamma\hbar}{2}\begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix}$$

With a static field $\vec{B} = B_x\hat{x} + B_y\hat{y} + B_z\hat{z}$, a general state just precesses about the new eigenaxis (the $|\psi_\pm\rangle$ axis in Fig. 1). With a time-dependent perturbation in the field, $\vec{B} = B_z\hat{z} + B_1\cos(\omega_0 t)\hat{x}$ where $B_1 \ll B_0$, we can get fast precession about the $z$-axis and slow precession about the $x$-axis.

## 3.11 LC oscillator as a superconducting qubit

Frequency: $f = 1/\sqrt{LC} \sim \hbar\omega$.

$$\hat{H} = \frac{1}{2}\hbar\omega_a\hat{\sigma}_z + \hbar\Omega_R\cos(\omega_0 t + \Phi_R)\hat{\sigma}_x$$

LC oscillator has same spectrum as simple harmonic oscillator; evenly-spaced energy levels. Need to introduce anharmonicity via a nonlinear component to break this even spacing so that a single transition between 2 specific levels (qubit) can be addressed. This nonlinear component is the Josephson junction, characterized by a current-phase relationship (CPR).

## 3.12 Quantum LC circuit

** This section is needed to make chapter 4 a little easier to understand. The introductory section of wikipedia's article on this is actually pretty sufficient at supplying basic definitions here.

## 3.13 Example

Question:

Suppose that you have a system of N qubits. How many classical bits does it take to describe the state of the system? First, consider a quantum register containing three qubits, each of which is prepared in state

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

You perform a measurement that projects each qubit onto either $|0\rangle$ or $|1\rangle$ and record the result in a 3-bit classical register. List all possible outcomes.

Answer:

First we consider a system of 3 qubits, all of them in the state $(|0\rangle + |1\rangle)/\sqrt{2}$, so the combined wave function is:

$$|\psi\rangle = \frac{1}{2\sqrt{2}} \left(|000\rangle + |001\rangle + |010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle + |111\rangle\right).$$

If we measure each qubit in the $\{|0\rangle, |1\rangle\}$ basis, we can have 8 possible outcomes: $|000\rangle$, $|001\rangle$, $|010\rangle$, $|011\rangle$, $|100\rangle$, $|101\rangle$, $|110\rangle$, $|111\rangle$ all with probability 1/8.

In general, the Hilbert space of $N$ qubits can be expanded as the tensor product of the Hilbert spaces of each of its qubits. This allows us to define a basis in the space of all qubits as the tensor product of all the possible combinations of the basis elements of the single qubit states: $\{|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle, |101\rangle, |110\rangle, |111\rangle\}$ in the case of three qubits. In a system of $N$ qubits the number of basis elements is therefore $\mathbf{2^N}$. Therefore, to fully characterize the state of an $N$ qubit system we need to specify $2^N$ complex numbers (or, more precisely, $2^N - 1$ due to the normalization constraint and the ability to ignore an overall global phase, which removes two real parameters, or equivalently one complex parameter). Because the parameters specifying the state are continuous, they cannot be represented precisely/exactly with a finite number of bits, in general.

This exponential scaling of the size of the Hilbert space with the number of qubits is the reason why it is not possible to efficiently simulate quantum computers with classical ones.

A final note about representing a many-body quantum state with a finite number of bits: we are making a rounding error on each component of the wave function, therefore the wave function that we are representing is an approximation ($|\psi_{\text{Approx}}\rangle$) to the exact one ($|\psi\rangle$). We want to have a good approximation so: $\langle\psi_{\text{Approx}}|\psi\rangle \sim 1$. However, this approximation gets intrinsically worse as the number of qubits increases if the precision in the bits is not increased. This a consequence of something known as the **orthogonality catastrophe**. Imagine we can represent our single qubit states with an error $\varepsilon$:

$$\langle\psi^{(i)}_{\text{Approx}}|\psi^{(i)}\rangle = 1 - \varepsilon,$$

where $(i)$ labels qubit $(i)$. If we form a many-qubit state as the product of the approximate states, its overlap will decay exponentially with the number of qubits:

$$\langle\psi_{\text{Approx}}|\psi\rangle = (1-\varepsilon)^N \sim e^{-\varepsilon N}.$$

Therefore to accurately describe our quantum many-body states we need a lot of classical bits for each of the exponentially many wave function components.

## 3.14 The Quantum Zeno Effect

It has been previously discussed that measuring a qubit in the $\{|0\rangle, |1\rangle\}$ basis corresponds to projecting its state $|\psi\rangle$ onto one of the basis vectors. This projection implies a way to "preserve" the state of a quantum system [6]. Let's consider the simplest type of decay possible: suppose we have a qubit in the state $|0\rangle$ that constantly tries to decay to the state $|1\rangle$. In time $\Delta t$, our qubit will rotate very slightly, let's say by angle $\theta = \epsilon$.

If we let enough time elapse, as $\theta \to \pi/2$, the greater the chance of measuring $|1\rangle$ becomes, which would correspond to qubit decay. However, if we periodically measure the state after short time intervals $\Delta t$, the state will only have time to rotate by $\epsilon$. Then, the probability that the state $|\psi\rangle = \cos(\epsilon)|0\rangle + \sin(\epsilon)|1\rangle$ is measured to be $|0\rangle$ is $\cos^2(\epsilon) \approx 1 - \epsilon^2$, which, as $\epsilon$ becomes arbitrarily small, becomes closer and closer to certainty.

In the limiting case, as the period between measurements $\Delta t \to 0$, we arrive at an interesting result: if a quantum system is observed continuously, then it will *never* decay!

Experimentally, this has been verified multiple times. In the next section, we will see a result that encapsulates

## 3.15 The Elitzur-Vaidman Bomb

How can you eat your cake and have it too?

That is (essentially) the question Avshalom Elitzur and Lev Vaidman, two Israeli physicists, set out to solve. They proposed a thought experiment in which one could verify a bomb is functional without having to detonate it [2]. A crucial concept that this experiment demonstrates is that of *interaction-free measurement*.

Traditionally, the setup involves mirrors, detectors, and a light-sensitive bomb, which detonates if it interacts with even a single proton. Angled half-silvered mirrors are used

to place each photon in an equal superposition state of travelling upwards, and travelling through the bomb. From that, we can use the interference of these waves in a clever way to determine if a bomb is a dud or not, with exploding good bombs only half the time!

However, there is another setup more amenable to quantum circuits such that the proportion of bombs exploded can be made *arbitrarily small*. It is this setup that will be considered [1].

### 3.15.1 The Classical Setup

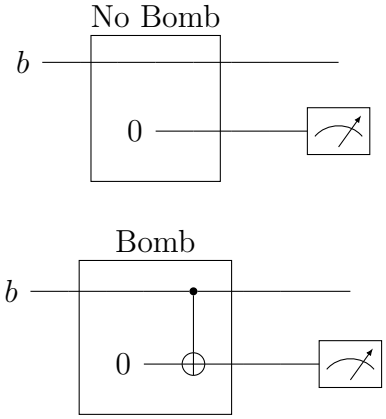
Suppose you are at an airport, and there is a suspicious piece of luggage left lying around. This suitcase may or may not contain a bomb; of course, if you open the suitcase, the bomb explodes.

The suitcase does, however, have an interface. It can be sent a classical bit,

$$b \in \{0, 1\}$$

where $b = 0$ means we don't make a query, and $b = 1$ means we do. This is a predicament; if we send 0, we don't learn anything, and if we send 1, we risk the bomb blowing up in our face!

For the sake of abstraction, the suitcase can be represented by a "black box," whose contents are unknown. There are two classical bits in the circuit: our input $b$ and one that gets measured. The way we denote an "explosion" is if we measure the second bit in the state 1. Note that the act of measuring 1 sets off the bomb, not that the output bit is 1. The corresponding circuits look like:




Think about what happens. If there is no bomb, and we send in $b = 0$, then we measure zero. If there is a bomb, and we send in $b = 0$, then we *also* measure 0. So sending in $b = 0$ doesn't give us any information about what's going on inside. If there is no bomb, and we send in $b = 1$, then we measure zero. We can say for certain, then, that there isn't a bomb. If there is a bomb, and $b = 1$, then the suitcase explodes. This is very unhelpful, and is equivalent to just opening up the suitcase to see for yourself.

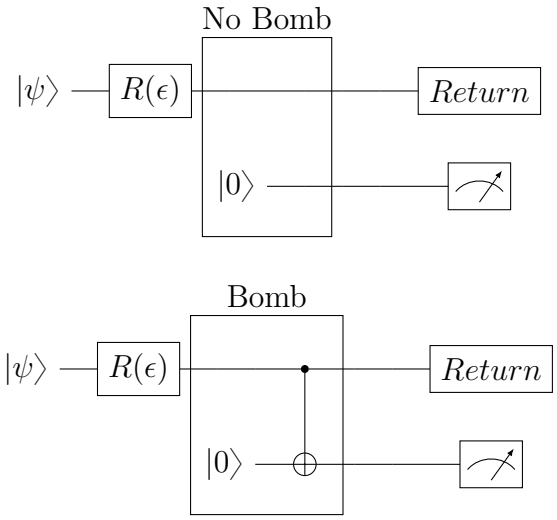
### 3.15.2 The Quantum Setup

Let's suppose the query system is a quantum one. That is, we are allowed to send in a quantum bit

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

to our suitcase. Let's assume that if there is no bomb, then $|\psi\rangle$ is returned. If there is a bomb, the bomb measures the second qubit in the basis $\{|0\rangle, |1\rangle\}$. If the outcome is $|0\rangle$, then because the states of the two qubits are entangled, our qubit $|\psi\rangle$ immediately collapses to $|0\rangle$ and therefore returns $|0\rangle$. If the output is $|1\rangle$, the bomb explodes (and then returns $|1\rangle$).

Our corresponding quantum circuits look like:




How do we go about testing if there is a bomb? The idea is to start in the state $|\psi\rangle = |0\rangle$. Next, we apply an arbitrarily small rotation of angle $\epsilon = \frac{\pi}{2N}$, where $N$ is the nubmer of times we will repeat this process, and is arbitrarily large. The matrix is

$$R(\epsilon) = \begin{pmatrix} \cos(\epsilon) & -\sin(\epsilon) \\ \sin(\epsilon) & \cos(\epsilon) \end{pmatrix}$$

giving us $|\psi\rangle = \cos(\epsilon)\,|0\rangle + \sin(\epsilon)\,|1\rangle$. If there is a bomb, the chance that it explodes is $\sin^2(\epsilon) \approx \epsilon^2$. With a high probabiltiy, we get back $|0\rangle$ and put that back into the black box.[3]

Now let's repeat this process $N$ times, and then we measure our $|\psi\rangle$ safely in our lab.

If there is no bomb, then our state is

$$R^N(\epsilon)\,|0\rangle = R(N\epsilon)\,|0\rangle = R(\pi/2)\,|0\rangle = |1\rangle$$

and we measure $|1\rangle$ in our lab.

[3]That is, unless you live in the unlucky branch of the multiverse where it explodes! The authors of [2] argue that the part of the superposition where the bomb explodes actually occurs, just not in our "world." Generally, this is known as the *Many-Worlds Interpretation* (MWI) of quantum mechanics. It has never been tested and is hotly debated among scientists.

If there is a bomb, however, each time we measure we "reset" the state of our qubit back to $|0\rangle$. Therefore, the total probability of setting off our bomb is only

$$N \sin^2(\epsilon) \approx N\epsilon^2 = \frac{\pi^2}{4N} = O\left(\frac{1}{N}\right).$$

In conclusion, if there is no bomb, we measure $|1\rangle$ in our lab after doing $N$ repetitions, and if there is, we measure $|0\rangle$ in our lab since the circuit keeps resetting the state of $|\psi\rangle$ to zero. If we do 100 repetitions, then our chance of exploding (if there is a bomb) is about 2.47%. I like those odds!

### 3.15.3 Qiskit Implementation

Below is the implementation of this in Qiskit.

```
N = 10
epsilon_ = np.pi/(2*N)
isBomb = False


def run_ev_circ(isBomb):
    qr = QuantumRegister(1, 'q')
    cr = ClassicalRegister(1,'c')
    qc = QuantumCircuit(qr,cr)
    if (isBomb):
        qc.ry(2*epsilon_,0)
        qc.measure(0,0)
    else:
        qc.ry(2*N*epsilon_,0)
        qc.measure(0,0)

    simulator = Aer.get_backend('qasm_simulator')
    job = execute(qc, simulator, shots=N*1000)
    result = job.result()
    counts = result.get_counts(qc)
    return counts
```

# 4 Multiple qubits

[4]Enormous difference in dimension between classical and quantum state space is due to a difference in the way the spaces combine.

The state of a classical system can be completely characterized by describing the state of each of its component pieces separately.

An unintuitive aspect of quantum systems is that often the state of the system cannot be described in terms of the state of its component pieces $\implies$ *entangled states*.

Feynman/Manin said classical systems cannot simulate entangled states in polynomial time (due to the scaling of the Hilbert space dimension with system size). $n$ qubits $\implies$ $2^n$ states, $2^n \times 2^n$ dimensional Hilbert space.

[4]10/16/19

Classical system of $n$ objects whose individual state can be described in 2D vector space $\implies 2n$. *Direct sum*, $\oplus$.

Quantum system of $n$ objects $\implies 2^n$. *Tensor product*, $\otimes$.

## 4.1 Direct sum of vector spaces

$$\mathbb{V} : A = \{|\alpha_1\rangle, |\alpha_2\rangle, \ldots, |\alpha_n\rangle\}$$
$$\mathbb{W} : B = \{|\beta_1\rangle, |\beta_2\rangle, \ldots, |\beta_n\rangle\}$$

Every element $|x\rangle \in \mathbb{V} \oplus \mathbb{W}$ can be written as $|x\rangle = |v\rangle \oplus |w\rangle$ for some $|v\rangle \in \mathbb{V}$ and $|w\rangle \in \mathbb{W}$.

Suppose the state of three objects $O_1, O_1, O_3$ is fully described by two parameters, the positions $x_i$ and the momenta $p_i$. Then the state of the system can be described by direct sum of the states of the individual objects,

$$\begin{pmatrix} x_1 \\ p_1 \end{pmatrix} \oplus \begin{pmatrix} x_2 \\ p_2 \end{pmatrix} \oplus \begin{pmatrix} x_3 \\ p_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ p_1 \\ x_2 \\ p_2 \\ x_3 \\ p_3 \end{pmatrix}$$

$\implies$ state of $n$ classical objects (each with 1 degree of freedom) has dimension $2n$.

## 4.2 Tensor products of vector spaces

The tensor product $\mathbb{V} \otimes \mathbb{W}$ of two vecor spaces $\mathbb{V}$ and $\mathbb{W}$ with bases $A = \{|\alpha_1\rangle, \ldots, |\alpha_n\rangle\}$ and $B = \{|\beta_1\rangle, \ldots, |\beta_m\rangle\}$ respectively is an $nm$-dimensional vector space with a basis consisting of $nm$ elements of the form $|\alpha_i\rangle \otimes |\beta_j\rangle$ where $\otimes$ is the tensor product. All elements of $\mathbb{V} \otimes \mathbb{W}$ can be represented:

$$a_{11} |\alpha_1\rangle \otimes |\beta_1\rangle + a_{12} |\alpha_1\rangle \otimes |\beta_2\rangle + \cdots + a_{nm} |\alpha_n\rangle \otimes |\beta_m\rangle$$

The standard basis for the vector space $\mathbb{V}_{n-1} \otimes \mathbb{V}_{n-2} \otimes \cdots \otimes \mathbb{V}_0$ for an $n$-qubit (i.e. 2-level, size 2 basis) system consists of $2^n$ vectors:

$$|0\rangle_{n-1} \otimes \cdots \otimes |0\rangle_1 \otimes |0\rangle_0$$
$$|0\rangle_{n-1} \otimes \cdots \otimes |0\rangle_1 \otimes |1\rangle_0$$
$$|0\rangle_{n-1} \otimes \cdots \otimes |1\rangle_1 \otimes |0\rangle_0$$
$$\vdots$$
$$|1\rangle_{n-1} \otimes \cdots \otimes |1\rangle_1 \otimes |1\rangle_0$$

Every $n$-qubit quantum state can be represented:

$$e^{i\theta}(a_0 |000\ldots0\rangle + \cdots + a_{2^n-1} |111\ldots1\rangle)$$

$\implies$ $2^n - 1$ complex numbers (global phase + normalization condition removes 2 real numbers).

If Alice controls 2 qubits and Bob the last three we may write a state as

$$\frac{1}{\sqrt{2}}\left(|00\rangle_A |101\rangle_B + |10\rangle_A |011\rangle_B\right)$$

where the subscripts indicate which qubits Alice controls and which Bob controls.

## 4.3 Meaning of entanglement

**Bell basis** for a two-qubit system:

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

The elements of the Bell basis are entangled. For instance $|\Phi^+\rangle$ cannot be described in terms of the state of its component qubits separately (i.e. in terms of a product state).

[5]Example 4-qubit state:

$$\begin{aligned}
|\psi\rangle &= \frac{1}{2}(|0000\rangle + |0101\rangle + |1010\rangle + |1111\rangle) \\
&= \frac{1}{\sqrt{2}}(|0\rangle_1 |0\rangle_3 + |1\rangle_1 |1\rangle_3) \otimes \frac{1}{\sqrt{2}}(|0\rangle_2 |0\rangle_4 + |1\rangle_2 |1\rangle_4)
\end{aligned}$$

$\implies$ $|\psi\rangle$ is not entangled with respect to the system composition of a subsystem of the 1st & 3rd qubit and a subsystem of the the 2nd & 4th qubit.

### 4.3.1 Multi-qubit measurements

Two qubits:

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$$
$$|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$$

We wish to measure the 1st qubit with respect to the standard basis $\{|0\rangle, |1\rangle\}$. Let $\mathbb{V}_1, \mathbb{V}_2$ be the vector spaces of the 1st and 2nd qubits, respectively. Then $\mathbb{V} = \mathbb{V}_1 \otimes \mathbb{V}_2$ is the vecotr space of the composite system.

$$S_1 = |0\rangle \otimes \mathbb{V}_2, \qquad S_2 = |1\rangle \otimes \mathbb{V}_2$$

[5]10/21/19

$$|\psi_1\rangle = \frac{a}{c_1}|00\rangle + \frac{b}{c_1}|01\rangle \in S_1$$

$$|\psi_2\rangle = \frac{c}{c_2}|10\rangle + \frac{d}{c_2}|11\rangle \in S_2$$

$$|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle$$

$$c_1 + \sqrt{|a|^2 + |b|^2}, \qquad c_2 + \sqrt{|c|^2 + |d|^2}$$

$$\implies |\psi\rangle = |0\rangle \otimes (a|0\rangle + b|1\rangle) + |1\rangle \otimes (c|0\rangle + d|1\rangle)$$

Multi-bit measurement can be treated as a series of single-bit measurements in standard basis.

- Measure if two qubits have the same value without learning the actual value of the two qubits.
- Equivalent to unitary transformations followed by a standard measurement of individual qubits.

Measurement gives another way of thinking about entangled particles. Particles are not entangled if the measurement of one has no effect on the other. For example, the state $(|00\rangle + |11\rangle)/\sqrt{2}$ is entangled since the probability that the first bit is measured to be $|0\rangle$ is $1/2$ if the second bit has not been measured. However, if the second bit had been measured, the probability would be 0 or 1. On the other hand, the state $(|00\rangle + |01\rangle)/\sqrt{2} = |0\rangle \otimes (|0\rangle + |1\rangle)/\sqrt{2}$ is not entangled, since any measurement of the first qubit will yield $|0\rangle$ regardless of whether the second qubit was measured.

### 4.3.2 EPR paradox

Einstein, Podolsky, and Rosen proposed a thought experiment that uses entangled particles which seem to violate principles of relativity.

$$A \longleftarrow \text{EPR source} \longrightarrow B$$

The EPR source generates two maximally entangled particles,

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

$A$ and $B$ can be arbitrarily far apart.

Suppose $A$ measures her particle and observes state $|0\rangle$. $\implies$ means the combined state will now be $|00\rangle$. Now if B measures his particle, he will observe $|0\rangle$.

The change of the combined quantum state occurs instantaneously even though the particles may be arbitrarily far apart. It appears that this would enable $A$ and $B$ to communicate faster than the speed of light. (Actually, they can't.)

EPR proposed local hidden variable theory. Each particle has some internal state that completely determines what the result of the measurement will be, but these variables are hidden from us. For the moment, the best we can do is to make probabilistic predictions.

J. S. Bell showed that the EPR picture must satisfy an inequality (now known as Bell's inequality). This inequality has been experimentally violated $\implies$ EPR cannot explain the results of measurements with respect to a different basis.

Description of *cause and effect*. A measurement by Alice affects a measurement performed by Bob. It is possible to set up the EPR scenario so that one observer *sees* Alice measure first, while another observer *sees* Bob measure first. According to relativity, physics must explain equally well the observation of the first observer as the second. The experimental results should be explained equally. The symmetry shows that Alice and Bob cannot use their EPR pair to communicate faster than the speed of light. All that can be said is that Alice and Bob are observing the same random (though perfectly correlated) behavior.

# 5 Quantum gates

Quantum systems obey Schrödinger's equation. The dynamics must take states to other states in a way that preserves orthogonality (unitary dynamics). For a complex vector space, linear transformation, unitary.

Suppose matrix $M$ is unitary. $\implies$ $MM^\dagger = \mathbb{1}$. $\implies$ quantum transformations (e.g. gates) are unitary and therefore reversible. $\implies$ *Energy of computation*. Classical computation is *not* reversible.

## 5.1 Single-qubit gates

Computational basis $\{|0\rangle, |1\rangle\}$.

$$\mathbb{1} \quad \begin{matrix} |0\rangle \to |0\rangle \\ |1\rangle \to |1\rangle \end{matrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{identity}$$

$$X \quad \begin{matrix} |0\rangle \to |1\rangle \\ |1\rangle \to |0\rangle \end{matrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{negation}$$

$$Y \quad \begin{matrix} |0\rangle \to & -|1\rangle \\ |1\rangle \to & |0\rangle \end{matrix} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{note: not Pauli}$$

$$Z \quad \begin{matrix} |0\rangle \to & |0\rangle \\ |1\rangle \to & -|1\rangle \end{matrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{phase flip}$$

[6]Hadamard gate:

$$H \quad \begin{matrix} |0\rangle \to & |+\rangle \\ |1\rangle \to & |-\rangle \end{matrix} \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{Hadamrd gate}$$

Quantum transformations are unitary transformations:

- Invertible/reversible: $U^\dagger U = \mathbb{1} \implies U^\dagger = U^{-1}$.

[6]10/23/19

- Deterministic.
- Continuous: you can apply them in a time-continuous way.

Measurements break all three rules of unitary transformations:

- Irreversible: whatever information we didn't capture is lost.
- Probabilistic: QM is deterministic only until measurement.
- Discontinuous: measurement is assumed to project instantaneously.

Norms:

- $p$-norm: $|x|_p = (\sum_i |x_i|^p)^{1/p}$
- 1-norm: $|x| = \sum_i |x_i|$ (classical probability)
- 2-norm: $|x|_2 = \sqrt{\sum_i |x_i|^2}$ (QM)

Quantum circuit notion: (quantum circuit diagram).

## 5.2 Two-qubit gates

** Would be worth mentioning noise from CNOT gate (and others?) ** Subsection on phase kickback ** Should have subsections covering S-gate, T-gate, U-gates Controlled NOT (CNOT) gate. First qubit is control bit:

$$|00\rangle \to |00\rangle$$
$$|01\rangle \to |01\rangle$$
$$|10\rangle \to |11\rangle$$
$$|11\rangle \to |10\rangle$$

$$C_{\text{NOT}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & X \end{pmatrix}$$

$$C_{\text{NOT}} C_{\text{NOT}} = \mathbb{1}$$

The CNOT gate cannot be decomposed into a tensor product of two single qubit transformations.

(Matrix representations of 2-qubit gates constructed as tensor products of single-qubit gates. Connection between operator/Dirac notation and matrix/array representations.)

** A fun(?) exercise or optional subsection here should be made on proving universality. IBM's qiskit textbook does this in a pretty easy-to-follow way that isn't so rigorous if the authors don't care too much about mathematical rigor.

# 6 Quantum Communication

[7]App. A.

## 6.1 Example

Question :

Alice and Bob have two methods of communication: they can communicate classically (by text messages) and they can communicate quantumly (Alice can send Bob single photons over an optical fiber). It is possible for other people to hack into both of these methods of communication. They agree that Alice will encode a string of '0's and '1's in the polarization state of photons. If they can verify no one intercepted this string, they will use it as their secret key. They agree to use the following mapping:

$$|\leftrightarrow\rangle \rightarrow 1$$

$$|\mathbb{V}\rangle \rightarrow 0$$

Alice will send a photon with its polarization oriented either horizontally or vertically. She will choose the polarization randomly, with equal probability for either. She will also text Bob telling him that she sent a photon, without telling which polarization she sent. Bob will measure the photon and text her back to confirm that he received it, without telling her anything about how he measured the photon or the measurement outcome.

**a.** How can Bob measure the polarization of the photon that Alice sends? List two ways and explain why one would be better than the other.

**b.** If Bob measures a vertically polarized photon, with what percent certainty does he know which polarization Alice sent?

**c.** Now suppose there's an eavesdropper named Eve. She is able to intercept the photon Alice sends to Bob, detect it, and send Bob a replacement photon. If she overhears all of the discussions Bob and Alice have about their plan to share the secret key, is it possible for her to know the polarization state of every photon Alice sends with 100% certainty?

**d.** Assuming that Eve is able to send the replacement photon instantaneously, would Alice and Bob know that she has eavesdropped? Why or why not?

**e.** Explain why it is not secure for Alice to tell Bob in which basis she is sending photons.

**f.** In the BB84 protocol, Alice and Bob still have the same two ways of communicating (text messages and single photons), but Alice now has an additional choice of sending light polarized horizontally, vertically, or with one of the two diagonal polarizations. They agree to use the same mapping as in Equation 1 for horizontally or vertically polarized photons, and the following mapping for the diagonally polarized photons:

| Alice's bit | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| Alice's basis | + | × | × | + | × | + | + | × | × |
| Alice's polarization | ↔ | ↘ | ↗ | | | | | | |
| Bob's basis | × | × | + | + | + | × | + | × | + |
| Bob's measurement | ↘ | ↘ | 𝕍 | | ↗ | ↘ | | | ↔ |
| Same basis? | No | Yes | | | | | | | |
| Shared secret key | | | | | | | | | |

$$|\nearrow\rangle \rightarrow 0$$
$$|\searrow\rangle \rightarrow 1$$

Alice now randomly sends any of the four polarization states with equal probability. Bob doesn't know which basis Alice used, so he randomly chooses a basis (horizontal/vertical or diagonal) in which to measure the photon.

**g.** Fill in the empty spaces in the following table. What is the secret key that Alice and Bob generate?

Answer:

**a.** **1)** Bob can use a filter like an absorptive polarizer which lets light polarized in a particular direction pass while blocking light polarized in the orthogonal direction. If his filter is oriented to accept either horizontally or vertically polarized light, his single-photon detector will either detect a photon (which means the polarization of the photon was the same as the direction of the filter) or nothing (which means the polarization of the photon sent was perpendicular to the filter). Using this method, he will only detect photons of one particular polarization.

**2)** Alternatively, Bob can use a polarizing beam splitter, with a single-photon detector positioned at each arm. If the beam splitter is oriented to transmit vertically polarized light and reflect horizontally polarized light, Bob will be able to detect the arrival of all photons and determine each's polarization.

The second option is better because Bob is able to detect all photons using single-photon detectors and thereby determine each's polarization. In the first option, photons polarized perpendicular to the filter will be absorbed (i.e. annihilated), and it will be very difficult (though not physically impossible) to know that a photon was received at all.

**b.** The horizontal and vertical polarization states are orthogonal, i.e. $\langle\leftrightarrow|\mathbb{V}\rangle = \langle\mathbb{V}|\leftrightarrow\rangle = 0$. Therefore, if Alice prepares only $|\leftrightarrow\rangle$ or $|\mathbb{V}\rangle$ photons, and Bob measures a $|\mathbb{V}\rangle$ photon, he knows with 100% certainty that Alice prepared a $|\mathbb{V}\rangle$ photon.

| Alice's bit | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| Alice's basis | + | × | × | + | × | + | + | × | × |
| Alice's polarization | ↔ | ↘ | ↗ | ↔ | ↗ | 𝕍 | 𝕍 | ↘ | ↗ |
| Bob's basis | × | × | + | + | + | × | + | × | + |
| Bob's measurement | ↘ | ↘ | 𝕍 | ↔ | ↗[1] | ↘ | 𝕍 | ↘ | ↔ |
| Same basis? | No | Yes | No | Yes | No | No | Yes | Yes | No |
| Shared secret key | | 1 | | 1 | | | 0 | 1 | |

[1] This is incorrect in the given table.

Table 1: Table for problem 2 part g.

**c.** Yes. It is assumed that Alice and Bob use the classical communication channel to agree upon which measurement basis to use. Since Eve has access to this channel, she will also know the measurement basis. Therefore, in the same way Bob can determine with 100% certainty the polarization state that Alice sends, Eve can also use a polarizing beam splitter to transmit $|\mathbb{V}\rangle$ and reflect $|\leftrightarrow\rangle$ and thereby determine the polarization state with 100% certainty.

**d.** No. Once Eve knows the polarization state, she can prepare an identical copy (in the same way that Alice does) and relay it to Bob. Bob will then measure the same state that Alice prepared, which Eve reproduced. Since it is assumed that Eve can do this instantaneously, there is no delay in the transmission between Alice and Bob, and they cannot use timing records to detect any meddling by Eve. When Alice and Bob later compare notes over the classical communication channel, Bob's measurement outcomes will be 100% correlated with the states Alice prepared.

**e.** As already explained in the previous parts, it is not secure for Alice to tell Bob which measurement basis she is using (assuming *only one* basis is being used) because Eve has access to both the classical communication channel and the quantum communication channel, and therefore knows everything that Bob knows. If Alice tells Bob the basis she is using, Eve will also know the basis. She can then intercept the photon, measure the polarization, and relay an identical copy to Bob. In this way she can determine the states of all the photons Alice sends and not be detected.

Yet Bob *needs* to know the basis Alice is using so that he can appropriately orient his beam splitter. Clearly, this scheme involving one basis will not work. *We need to employ a second basis.*

**f.** This is just a description of the BB84 setup. The key is that Bob knows the two bases being used, but he doesn't know *which* basis Alice will use for each photon.

**g.** See table 1

# 7 Quantum Hardware

## 7.1 Josephson Junctions

[8]E.g. Al–$AlO_x$–Al, capacitor at room temperature.

Quantum resistance:

$$R_Q = h/e^2 \approx 25.8\,\mathrm{k}\Omega$$

Two coupled Schrödinger equations (one on each side of the junction):

$$i\hbar\frac{\partial\Psi_1}{\partial t} = \mu_1\Psi_1 + K\Psi_2$$

$$i\hbar\frac{\partial\Psi_2}{\partial t} = \mu_2\Psi_2 + K\Psi_1$$

where $K$ is the coupling across the barrier.

$$\Psi_1 = \sqrt{n_1}e^{i\theta_1}$$

$$\Psi_2 = \sqrt{n_2}e^{i\theta_2}$$

where $n_{1,2}$ are the density of Cooper pairs and $\theta_{1,2}$ are the phases.

$$\hbar\frac{\partial n_1}{\partial t} = -\hbar\frac{\partial n_2}{\partial t} = 2K\sqrt{n_1 n_2}\sin(\theta_2 - \theta_1)$$

$$-\hbar\frac{\partial}{\partial t}(\theta_2 - \theta_1) = \mu_2 - \mu_1$$

Then note current $\sim \frac{\partial n}{\partial t}$. Define $\phi = \theta_2 - \theta_1$, and $I_0 = 2K\sqrt{n_1 n_2}/\hbar$, yielding

$$I = \frac{\partial n_1}{\partial t} = I_0 \sin\phi$$

$$\frac{\partial\phi}{\partial t} = \frac{2e}{h}V$$

where $I$ is known as the supercurrent, first equation known as "current-phase relationship."

Harmonic oscillator, Schrödinger equation:

$$\left[\frac{-\hbar^2}{2m}\nabla^2 + \frac{1}{2}m\omega^2 x^2\right]\psi = E\psi$$

$H = T + V$. Solution:

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right)$$

$$\psi_n = A_n\left(\hat{a}_+\right)^n\psi_0(x)$$

$$a_\pm = \frac{1}{\sqrt{2\hbar\omega m}}(m\omega\hat{x} \mp i\hat{p})$$

[8]10/7/19

Introducing nonlinearity:

$$H = H_0 + \lambda x^4$$

where $H_0\psi_n = E_n\psi_n$. Then,

$$E_0 = E_0^0 + \lambda E_n^1.$$

Expand $x^4$ in terms of raising and lowering operators, $x \sim a_- + a_+$ to calculate

$$E_n^{(1)} = \langle \Psi_n | x^4 | \Psi_n \rangle = (6n^2 + 6n + 3) \left( \frac{\hbar}{2m\omega} \right)^2 \lambda$$

Each energy level shifts by a different amount, no longer even spacing between levels.

## 7.2 Josephson junction as qubit

Classical first:

$$H = T + V \qquad \mathcal{L} = T - V$$

Euler-Lagrange equations:

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial v_i} - \frac{\partial \mathcal{L}}{\partial x_i} = 0$$

$$\mathcal{L} = \frac{1}{2} m v^2 - v(x)$$

$$\dot{x} = \frac{\partial H}{\partial p} \qquad \dot{p} = -\frac{\partial H}{\partial x}$$

Oscillators:

$$\mathcal{L} = T - V = \frac{1}{2} C \dot{\Phi}^2 - \frac{1}{2L} \Phi^2$$

$$Q = \frac{\partial \mathcal{L}}{\partial \dot{\Phi}} = C\dot{\Phi} = CV$$

$$H = \frac{Q^2}{2C} + \frac{\Phi^2}{2L} \equiv \frac{p^2}{2m} + \frac{1}{2} k x^2$$

i.e. a harmonic oscillator with "mass" $C$ and "spring constant" $1/L$, natural frequency $\omega_c = 1/\sqrt{LC}$.

Now quantize Hamiltonian: $Q \to \hat{Q}$ and $\Phi \to \hat{\Phi}$, note $[Q, \Phi] = -i\hbar$. Operators $Q$, $\Phi$ can be written as raising and lowering operators:

$$\Phi = \Phi_{\text{z.p.}} (a_- + a_+) \tag{3}$$

$$Q = -iQ_{\text{z.p.}} (a_- - a_+) \tag{4}$$

where "z.p." = zero-point,

$$\Phi_{\text{z.p.}} = \sqrt{\frac{\hbar Z}{2}} \qquad Q_{\text{z.p.}} = \sqrt{\frac{\hbar}{2Z}}$$

where $Z = \sqrt{L/C}$ is characteristic impedance. If the characteristic impedance is on the order of quantum resistance $h/e^2$, then $Q_{\text{z.p.}}$ is on the order of the electron charge, and $\Phi_{\text{z.p.}}$ is on the order of the flux quantum.

Consider the input admittance of a parallel LC resonator:

$$Y(\omega) = i\omega C + \frac{1}{i\omega L} = \frac{1}{2Z}\left(\frac{\omega_c}{\omega}\right)\left(1 - \left(\frac{\omega}{\omega_c}\right)^2\right)$$

The Hamiltonian is given by $H = \hbar\omega_c a_+ a_-$ and physical observables flux $\hat{\Phi}$ and charge $\hat{Q}$ that are given by Eqs. (3), (4).

[9]Superconducting qubit: In order to go beyond the simple LC harmonic oscillator to create a qubit, we need a nonlinear element to produce anharmonicity in the spectrum. A number of qubit designs have been developed around Josephson junctions including:

- Cooper pair box based on charge
- flux qubit
- phase qubit
- fluxonium qubit.

Cooper pair box is topologically distinct from the other designs in that it has no wire closing the loop around the junction. It consists very simply of a small antenna whose two halves are connected by a Josphson junction.

$\implies$ Two sides of the junction are not shunted by an inductor.

$\implies$ # Cooper pairs transferred through the junction is a well-defined integer. $[Q, \Phi] = -i\hbar \implies$ periodicity of spectrum in charge.

$\implies$ charge-based qubits are sensitive to stray electric field noise.

### 7.2.1 Lagrangian for a Cooper pair box

Similar to the LC oscillator except unlike the inductor the energy stored in the Josephson junctions is not quadratic but periodic.

$$\mathcal{L} = \frac{1}{2}C_\Sigma \dot{\Phi}^2 + E_J \cos\phi$$

where $\phi$ is phase difference of superconducting order parameter across the JJ,

$$\phi = 2\pi\frac{\Phi}{\Phi_0} = \frac{2e}{\hbar}\Phi$$

with $\Phi_0 = h/2e$, and $C_\Sigma = C_{\text{geometry}} + C_J$, thus,

$$\mathcal{L} = \frac{1}{2}C_\Sigma\left(\frac{h}{2e}\right)^2 \dot{\phi}^2 + E_J \cos\phi.$$

[9]10/9/19


Figure 2: Family tree of superconducting qubits.

From Euler-Lagrange,

$$n = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \frac{\hbar^2}{4e^2} C_\Sigma \dot{\phi}$$

yielding the Hamiltonian,

$$\mathcal{H} = 4E_C(\hat{n} - n_g)^2 - E_J \cos\phi$$

where $E_C = e^2/2C_\Sigma$ is charging energy, $n = -i\,\partial/\partial\phi$ operator, and $n_g$ is the so-called offset charge representing external bias voltage and also containing random fluctuations due to stray charges jumping around.

In the limit $E_J \gg E_C$, the low-frequency noise in the offset charge can be neglected, resulting in very long coherence times $\implies$ "transmon qubit." With $E_J \gg E_C$ the "gravitational force" is very strong, and the "moment of inertia" is very large, so the phase angle undergoes very small quantum fluctuations.

$$H_0 = 4E_C\hat{n}^2 + \frac{E_J}{2}\hat{\phi}^2$$

$$V = E_J\left[-\hat{\phi}^4/4! + \cdots\right].$$

We see $H_0$ is the Hamiltonian of a simple LC harmonic oscillator with frequency $\hbar\Omega = \sqrt{8E_J E_C}$ and with leading order effect of the JJ being to play the role of a relatively large linear inductor. In the limit $E_J \gg E_C$, we are effectively ignoring the fact that $\hat{\phi}$ is an angular variable. It is this assumption which makes conjugate momenta $\hat{n}$ have continuous rather than integer values.

## Example

Q:

Using the current-phase and voltage-phase relationship of a single Josephson junction:

(a) derive the inductance of a Josephson junction?

(b) can it be zero? How about negative?

(c) calculate the Josephson energy.

(d) Now consider our junction is made out of aluminum leads separated by $d = 1.2nm$ ( at this thickness the resistivity is $5k\Omega/square$) with area A. An ideal junction satisfies: $I_c R_n = \Delta$ where $I_c$ is the supercurrent, $R_n$ is the normal resistance and $\Delta$ is the superconducting gap. What is the resonant frequency of this junction?

(e) Imagine we take thee junction and bias it with 10 micro volt of DC voltage. We then monitor the current with an ideal oscilloscope. What frequency do we detect?

A:

a) The Josephson effect describes the supercurrent $I_J$ that flows through the junction according to the classical equations

$$I_J = I_c sin\delta, \tag{5}$$

$$V = \frac{\Phi_0}{2\pi}\frac{d\delta}{dt} \tag{6}$$

where $\Phi_0 = \frac{h}{2e}$ is the superconducting flux quantum, $I_c$ is the critical-current parameter of the junction, and $\delta$ and $V$ are respectively the superconducting phase difference and voltage across the junction. The dynamical behavior of these two equations can be understood by first differentiating Eq. (5) and replacing $\frac{d\delta}{dt}$ with $V$ according to Eq.(6)

$$\frac{dI_J}{dt} = I_c cos\delta \frac{2\pi}{\Phi_0} V.$$

With $\frac{dI_J}{dt}$ proportional to V, this equation describes an inductor. By defining a Josephson inductance $L_J$ according to the conventional definition $V = L_J\frac{dI_J}{dt}$, one finds

$$L_J = \frac{\Phi_0}{2\pi I_c cos\delta}. \tag{7}$$

b) The magnitude of the inductance becomes large as $\delta \to \frac{\pi}{2}$ or $\frac{3\pi}{2}$, and is negative for $\frac{\pi}{2} < \delta < \frac{3\pi}{2}$. But it cannot be zero because $\cos\delta$ is bounded on $[-1, 1]$.

c) The energy stored in the junction is given by

$$\begin{aligned} U_J &= \int I_J V dt \\ &= \int I_c sin\delta \frac{\Phi_0}{2\pi}\frac{d\delta}{dt} dt \\ &= \frac{I_c\Phi_0}{2\pi}\int_0^\delta sin\delta' d\delta' \\ &= \frac{I_c\Phi_0}{2\pi}(1 - cos\delta) \end{aligned}$$

d) The resonant frequency of an LC circuit is

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

where the capacitance $C = \epsilon A/d$ with $d = 1.2\,\text{nm}$ contact separation, $A$ is the area of the interface between the superconducting aluminum contacts and the insulating medium between them, and $\epsilon$ is the dielectric constant of the insulator. The inductance $L = L_J$, as given in Eq. (7), depends on the critical current $I_c$, which can be obtained from the product

$$I_c R_n = \Delta \implies I_c = \Delta/R_n$$

where $\Delta$ is the superconducting gap, and the normal resistance of the junction $R_n = \rho d/A$, with $\rho$ the resistivity. Then

$$\omega_o = \frac{1}{\sqrt{\frac{\Phi_0}{2\pi cos\delta}\frac{\rho\epsilon}{\Delta}}}$$

Note there is some ambiguity in the resistivity provided in the problem. Normally, resistivity has units of [resistance·length], whereas the resistivity given in the problem has units of [resistance·length$^{-2}$]. If we assume $\rho = 5\,\mathrm{k}\Omega/\mathrm{length}^2$ as given in the problem, and take $R_n = \rho A$, then we obtain

$$\omega_o = \frac{1}{\sqrt{\frac{\Phi_0}{2\pi cos\delta}\frac{\rho A}{\Delta}\frac{\epsilon A}{d}}}$$

e) The bias voltage introduces a time-dependent phase,

$$V_{dc} = \frac{\Phi_0}{2\pi}\frac{d\delta}{dt} \rightarrow \delta = \frac{2\pi V_{dc}}{\Phi_0}t$$

Then, by the current-phase relationship of the junction,

$$I_J = I_c \sin\delta = I_c \sin\left(2\pi\frac{V_{dc}}{\Phi_0}t\right) \equiv I_c \sin(2\pi f_0 t)$$

which implies a frequency

$$f_0 = \frac{V_{dc}}{\Phi_0} = \frac{V_{dc}}{\frac{h}{2e}} = \frac{10\mu V}{\frac{4.135\times10^{-15}eVs}{2e}} = 4.8368\,\mathrm{GHz}$$

# 8 Superdense coding

Information theory (Shannon): by sending $n$ bits, you cannot communicate more than $n$ bits of information.

*Superdense coding*: Alice can send Bob 2 *classical* bits by sending him *only one qubit*. (Catch: A&B must share some entangled qubit ahead of time. With no prior entanglement, Alice can't send more than one bit per qubit.)

- Alice: $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$.
- Alice and Bob share a Bell pair in advance: $|\Psi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$.
- Alice is going to *manipulate* her half, then send her qubit to Bob, and Bob can measure both qubits and get 2 bits of information from Alice.

2 qubits → 4 states in basis. Will apply a NOT gate and phase flip. Alice can find 3 orthogonal states to the original Bell pair; these will be the other states of the Bell basis (2). Label the two qubits $x$ and $y$.

- If $x = 1$, she applies a NOT gate.
- If $y = 1$, she applies a phase gate.

For Bob to decode the information, he has to use the inverse transformation.

# 9 Density matrices

[10](This is from Scott Aaronson's Quantum Computation course [1, Lec. 6].) So far Basis states are defined: $\{|0\rangle, |1\rangle\} \to |i\rangle$. Pure states are superpositions of basis states: $|\psi\rangle = \sum \alpha_i |i\rangle$. Pure states exist in isolated quantum systems but you can also have quantum superposition layered together with regular, old probabilistic uncertainty. In some sense, mixed states are a distribution over quantum states, so

$$\{P_i, \Psi_i\} = P_1 |\Psi_1\rangle, \ldots, P_n |\Psi_n\rangle,$$

meaning that with probability $P_i$ the superposition is $|\Psi_i\rangle$. Mixed states (that is, the components $|\psi_i\rangle\langle\psi_i|$ of the mixed state) don't have to be orthogonal. Pure state is a degenerate case of a mixed state where all the probabilities are 0, except for one $P_i$ which is 1.

We write a mixed state as a *density matrix*:

$$\rho = \sum_i = p_i |\psi_i\rangle\langle\psi_i| .$$

(Examples of outer product in matrix representation.)

For any two orthgonal states $|v\rangle$ and $|w\rangle$,

$$\frac{|v\rangle\langle v| + |w\rangle\langle w|}{2} = \frac{\mathbb{1}}{2}$$

Measuring $\rho$ in the basis $|1\rangle, \ldots, |N\rangle$ gives outcome $|i\rangle$ with probability $\Pr(|i\rangle) = \rho_{ii} = \langle i|\rho|i\rangle$. A diagonal density matrix is another way of writing a classical probability distribution, e.g. $\begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}$.

Rank of a matrix: $\text{rank}(\rho)$ = the number of non-zero eigenvalues. The rank of an $n \times n$ matrix is at most $n$.

The matrix $\mathbb{1}/2$ is the even mixture of $|0\rangle$ and $|1\rangle$. It is also an even mixture of $|+\rangle$ and $|-\rangle$. This is a *maximally mixed state*. If we do a unitary transformation $U$ on the states $|\psi_i\rangle$ comprising $\rho$,

$$\rho' = \sum_i P_i (U |\psi_i\rangle) (U |\psi_i\rangle)^\dagger = \sum_i P_i U |\psi_i\rangle\langle\psi_i| U^\dagger = U\rho U^\dagger$$

i.e. the new density matrix is unitarily transformed. E.g. $|+\rangle\langle+| = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ in the $\{|0\rangle, |1\rangle\}$ basis, and the off-diagonals represent the interference between $|0\rangle$ and $|1\rangle$.

What are the constraints on $\rho$?

- Square $n \times n$.
- Hermitian $\rho^\dagger = \rho$.

[10] 10/28/19

- Probability $\sum_i \rho_{ii} = 1 = \mathrm{Tr}\{\rho\}$.
- Positive semidefinite: all eigenvalues are nonnegative (since they represent probabilities).

(Examples of pure vs. mixed states.)

Alice and Bob may share a bipartite state described by a density matrix. The state of each's qubit is described by the reduced density matrix obtained by taking a partial trace over the other's qubit space. The key points:

- A density matrix encodes all and only what is physically observable.
- Two quantum states will lead to different probabilities iff they have different density matrices.
- No communication theorem: If Alice and Bob share an entangled state, nothing $A$ chooses to do will have any effect on Bob's density matrix $\implies$ there is no observable effect on Bob's end.

Basis state $|i\rangle$. Pure state $|\psi\rangle = \sum_i \alpha_i |i\rangle$. Mixed state $\{P_i, |\psi_i\rangle\} \to \rho = \sum_i P_i |\psi_i\rangle\langle\psi_i|$, classical probability distributions over pure states.

What represents better the actual physical reality? Pure vs. mixed? When we look at part of an entangled state, a mixed state is the most complete representation possible that only describes the part that we are looking at.

# 10 Quantifying entanglement

[11]How do you quantify how much entanglement there is between two quantum systems? $\leftrightarrow$ measuring entanglement.

Entanglement is our resource $\leftrightarrow$ EPR $(|00\rangle + |11\rangle)/\sqrt{2}$. Entanglement of pure states; entanglement of mixed states.

## 10.1 Schmidt decomposition

Given $\sum_{ij} \alpha_{ij} |i\rangle_A |j\rangle_B$, how do we calculate how many EPR pairs it's worth?

Let $A = \Big(\alpha_{ij}\Big)$. In bipartite states, we can always find a change of basis on Alice's side and another on Bob's side that puts the state into the simpler form (Schmidt decomposition) $\sum_i \lambda_i |v_i\rangle |w_i\rangle$ where $\{|v_i\rangle\}$ is an orthonormal set, and $\{|w_i\rangle\}$ is an orthonormal set. This follows from the singular value decomposition of the matrix $A$: we can multiply $A$ by two unitary matrices, one on each side, to get a diagonal matrix $\Lambda = UAV$. Measuring in the $\{|v_i\rangle, |w_i\rangle\}$ basis would then yield the probability distribution

$$\Big(|\lambda_1|^2, \ldots, |\lambda_2|^2\Big)^T. \tag{8}$$

[11]10/30/19

## 10.2 Entropies

### 10.2.1 Shannon entropy

For classical probability distribution $D = (P_1, \ldots, P_n)$, the Shannon entropy is

$$H(D) = \sum_{i=1}^{n} P_i \log_2 \frac{1}{P_i} = -\sum_{i=1}^{n} P_i \log_2 P_i$$

E.g.

- $D = (0, 1) \implies H(D) = -0 \log_2 0 - 1 \log_2 1 = 0.$
- $D = (\frac{1}{2}, \frac{1}{2}) \implies H(D) - 2(\frac{1}{2} \log_2 \frac{1}{2}) = 1$

For the distribution (8), we can calculate the ordinary Shannon entropy of our system as

$$\sum_i |\lambda_i|^2 \log \frac{1}{|\lambda_i|^2}$$

in order to figure out how many Bell pairs our state is equivalent to.

### 10.2.2 Von Neumann entropy

Generalizes Shannon entropy from classical probability distributions to quantum mixed states. We say that the von Neumann entropy of a mixed state $\rho$ is

$$S(\rho) = \sum_{i=1}^{N} \lambda_i \log_2 \frac{1}{\lambda_i} = -\sum_i \lambda_i \log_2 \lambda_i.$$

The von Neumann entropy is the Shannon entropy of the vector of eigenvalues of the density matrix of $\rho$ (i.e. $\lambda_i$'s here are probabilities).

The von Neumann entropy of any pure state $|\psi\rangle\langle\psi|$ is zero because there is always some measurement basis that returns a definite outcome (i.e. the basis containing $|\psi_i\rangle$). Example: you could measure $|+\rangle$ in the $\{|0\rangle, |1\rangle\}$ basis where you have complete uncertainty and Shannon entropy of 1, but if you measure in the $\{|+\rangle, |-\rangle\}$ basis you have complete certainty with Shannon entropy of 0. In either case, the von Neumann entropy is 0.

Von Neumann entropy of a maximally mixed qubit state $\mathbb{1}/2$ is 1. The von Neumann entropy of a maximally mixed $n$-qubit state $\mathbb{1}/2^n$ is $n$.

Suppose Alice and Bob share a bipartite pure state $|\psi\rangle = \sum_{ij} \alpha_{ij} |i\rangle_A |j\rangle_B$. To quantify the entanglement entropy, we trace out either Alice or Bob's part of the system and calculate the von Neumann entropy of the reduced density matrix representing just Bob's are Alice's part, respectively (which we choose should not matter, $S(\rho_A) = S(\rho_B) = H(\lambda_i)$, Shannon entropy of $\lambda_i$ which you get by diagonalizing the Alice's or Bob's reduced density matrix.

# 11 Mixed vs. pure states

[12]Diagonal elements $\rho_{ii}$ of a density matrix represent the populations in the chosen basis:

$$\rho = \sum_i p_i \left|i\right\rangle\!\left\langle i\right|, \qquad \left|i\right\rangle = \sum_p c_p^{(i)} \left|p\right\rangle$$

$$\rho_{ii} = \sum_p = \left|c_p^{(i)}\right|^2 p_i \rightarrow \text{population in state } \left|i\right\rangle$$

The off-diagonal elements $\rho_{ij}$ of a density matrix provide information on interference of $\left|i\right\rangle$ and $\left|j\right\rangle \rightarrow$ they represent coherence in mixed states. When a qubit is coupled to an environment, this coupling can cause time-dependence of elements of density matrix $\rightarrow$ Decay of the off-diagonal elements is referred to as "dephasing" ($T_2$) while the decay of the diagonal elements is referred to as "relaxation" or "dissipation" ($T_1$).

What is the difference between the pure state $\frac{1}{\sqrt{2}}(\left|0\right\rangle + \left|1\right\rangle) = \left|+\right\rangle$ and mixed state $\frac{1}{2}(\left|0\right\rangle\!\left\langle 0\right| + \left|1\right\rangle\!\left\langle 1\right|) = \mathbb{1}/2$? Let's apply a Hadamard gate to both:

| | |
|---|---|
| pure state | $H\left\|+\right\rangle = \left\|0\right\rangle$ |
| mixed state | $H\mathbb{1}H/2 = \mathbb{1}/2$ |

In the first case the amplitudes can interfere and produce $\left|0\right\rangle$ while in the second case the states cannot interfere because they are in an "incoherent" mixture of the two states.

Open quantum systems can produce evolution which starts out in a pure state and evolves into a mixed state $\rightarrow$ such evolution can destroy our quantum computation $\rightarrow$ quantum error correction.

Recall,

$$\rho = \begin{pmatrix} \rho_{11} & & & \\ & \rho_{22} & & \\ & & \ddots & \\ & & & \rho_{nn} \end{pmatrix}, \qquad \operatorname{Tr}\rho = \sum_i \rho_{ii} = 1$$

A nice property of the trace operation is that a basis change leaves it invariant. Trace has a cyclical property,

$$\operatorname{Tr}(ABC) = \operatorname{Tr}(BCA) = \operatorname{Tr}(CAB).$$

Expectation value of any operator $\hat{O}$ in terms of projectors $\hat{P}_\psi$ can be given by:

$$\left\langle \hat{O} \right\rangle_\psi = \langle\psi|\hat{O}|\psi\rangle = \operatorname{Tr}\left(\hat{O}\hat{P}_\psi\right) = \operatorname{Tr}\left(\hat{P}_\psi\hat{O}\right).$$

Suppose we take pure states $\left|\pm\right\rangle$. The density matrices are

$$\rho_\pm = \frac{1}{2}\begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix},$$

[12]11/4/19

while the classical mixture is

$$\rho = \sum_{i=\pm} p_i \left|i\right\rangle\!\left\langle i\right| = \sum_{i=\pm} p_i \rho_i,$$

Statistical physics: statistical mixture of energy eigenstates in thermal equilibrium:

$$\rho = \sum_n p_n \left|n\right\rangle\!\left\langle n\right|, \quad \text{where} \quad p_n = \frac{e^{-E_n/kT}}{\mathcal{Z}}, \quad \text{and} \quad \mathcal{Z} = \sum_n e^{-E_n/kT}.$$

$\mathcal{Z}$ is the partition function. When the Hamiltonian does not depend on time, the mixture is time-independent.

Example: mixture where we prepare $|0\rangle$ with probability $\frac{1}{2}$ and $|+\rangle$ with probability $\frac{1}{2}$.

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$$

The eigenvalues are $\lambda_\pm = \frac{1}{2} \pm \frac{\sqrt{2}}{4}$, with eigenvectors $|\phi_\pm\rangle = \sqrt{\lambda_\pm}\,|0\rangle + \sqrt{1-\lambda_\pm}\,|1\rangle$. We can now view the mixed state as a mixture of two eigenstates $|\phi_\pm\rangle$ with weights equal to $\lambda_\pm$.

There are two simple ways to determine whether $\rho$ describes a pure state or not:

1. Pure state $\Leftrightarrow \rho^2 = \rho$. Mixed state $\Leftrightarrow \rho^2 \neq \rho$.
2. Pure state $\Leftrightarrow \operatorname{Tr}\rho^2 = 1$. Mixed state $\Leftrightarrow \operatorname{Tr}\rho^2 < 1$.

In fact, we can define the "purity of a state" $P \equiv \operatorname{Tr}\rho^2$. A state is pure if $P = 1$ and mixed otherwise. Example: $n$ qubits, dimension $d = 2^n \to \rho = \mathbb{1}/d$,

$$\implies P = \operatorname{Tr}\big((\mathbb{1}/d)^2\big) = \sum_1^d \frac{1}{d^2} = \frac{1}{d}.$$

Entropy $S(\rho) = -\sum_k \lambda_k \log_2 \lambda_k$ (note $\lim_{x\to 0} x \log x = 0$). We use log base 2 so that the unit of the entropy is the "bit." We can interpret the entropy as the missing information (in bits) about the state. Pure state: $S(\rho) = 0$. Mixed state: $S(\rho) > 0$.

## 11.1 Reduced density matrix

Consider a measurement just on system A, say of an observable $O_A$. The expectation value of that observable in terms of $\rho_{AB}$ is

$$\begin{aligned} \langle O_A \rangle = \operatorname{Tr}(\rho_{AB} O_A) &= \sum_{nm} \langle n_A m_B | \rho_{AB} O_A | n_A m_B \rangle \\ &= \sum_{nm} \langle n_A | \langle m_B | \rho_{AB} | m_B \rangle O_A | n_A \rangle. \end{aligned}$$

We then define the reduced density matrix for A as a partial trace over B,

$$\rho_A = \sum_m \langle m | \rho_{AB} | m \rangle \equiv \operatorname{Tr}_B \rho_{AB},$$

and then any observable in A can be expressed,

$$\langle O_A \rangle = \operatorname{Tr}(\rho_A O_A).$$

That is, $\rho_A$ describes the state of system A alone.

## 11.2 Quantifying entanglement in mixed states

[13](See [1, Lec. 11].) For bipartite mixed state there are two entanglement measures to consider:

- *Entanglement of formation:* $E_F(\rho_{AB})$ is the number of entangled qubits that A and B need to create one copy of the state $\rho_{AB}$ (in the limit where they are creating many copies of it, and assuming they are allowed unlimited local operations and classical communication (LOCC) for free).
- *Distillable entanglement:* $E_D(\rho_{AB})$ is the number of entangled qubits that A and B can extract per copy of $\rho_{AB}$ again in LOCC for free.

Clearly $E_F \geqslant E_D$ since if you could ever get out more entanlgment than you put it, it would give you a way to increase entanglement (entropy) arbitrarily using LOCC, which is impossible.

What about $E_F \gg E_D$? There exist bipartite states that take a lot of entanglement to make but you can only extract a small fraction of the entanglement.

A bipartite state $\rho_{AB}$ is separable if there is a way to write it as a mixture of product states:

$$\rho_{AB} = \sum_i p_i \left|v_i\right\rangle\!\left\langle v_i\right| \otimes \left|w_i\right\rangle\!\left\langle w_i\right|$$

A mixed state is called entangled iff it is *not* separable. It could happen that a density matrix looks entangled *but* there's some decomposition that shows it is separable.

Proven by Leonid Gurvits, 2003: If you are given a density matrix $\rho_{AB}$ for a bipartite state then deciding whether $\rho_{AB}$ represents a separable or entangled state is an NP-hard problem.

# 12 Some standard results, protocols, and other bits and bobs

## 12.1 No-cloning theorem

Assume that $U$ is a unitary transformation that clones in that $U\left|a0\right\rangle = \left|aa\right\rangle$ for all quantum states $\left|a\right\rangle$. Let $\left|a\right\rangle$ and $\left|b\right\rangle$ be orthogonal quantum states. Clearly $U\left|a0\right\rangle = \left|aa\right\rangle$ and $\left|b0\right\rangle = \left|bb\right\rangle$. But consider a superposition state, $\left|c\right\rangle = (\left|a\right\rangle + \left|b\right\rangle)/\sqrt{2}$:

$$U\left|c\right\rangle = \frac{1}{\sqrt{2}}(\left|aa\right\rangle + \left|bb\right\rangle) \neq \left|cc\right\rangle,$$

therefore $U$ is not a cloning operation.

It is possible to clone a known quantum state. The no-cloning theorem says we cannot reliably clone an *unknown* state. Nor is cloning possible by using measurements.

[13]11/6/19

## 12.2 Quantum teleportation

It is possible for Alice and Bob to use pre-shared entanglement plus classical communication to perfectly transmit a qubit.

Alice has a single qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. She also shares the EPR pair $(|00\rangle + |11\rangle)/\sqrt{2}$. Alice applies some transformation to $|\psi\rangle$ that entangles it with her half of the Bell pair. She then measures her qubit. Alice tells Bob the measurement outcomes over the phone. Bob then applies some transformation to his qubit of the entangled pair and "magically" will have $|\psi\rangle$.

(Example.) Alice applies CNOT with $|\psi\rangle$ as control and her half of the Bell state as target. Alice then applies Hadamard on $|\psi\rangle$. Alice measures both of her qubits. She tells Bob the result. Depending on the result, Bob applies a corresponding retrieval transformation on his qubit (which was part of the Bell pair), which deterministically reproduces $|\psi\rangle$ on Bob's end. $\implies$ they have transmitted a qubit without a quantum channel.

[14]Are we sending infinitely more information than two classical bits if we can send over enough information to perfectly describe an arbitrary qubit? In some sense we are sending a perfect qubit but Bob only obtains the information that he can measure which is significantly less.

We didn't make any assumption that $|\psi\rangle$ has to be unentangled from the rest of the world. You can have a protocol that has $|\psi\rangle$ to be half of another Bell pair. This protocol would entangle the 4th qubit to Bob's qubit. This suggests that it should be possible to teleport an arbitrary $n$-qubit entangled state by simply teleporting the qubits one at a time.

A consequence of this is that two qubits don't need to interact directly to become entangled. Consider a quantum circuit where $H$ acts on qubit 1, followed by a CNOT acting on qubit 2 conditioned on 1, followed by a SWAP gate between 2 and 3. Qubits 1 and 3 end up entangled even though there's never any "direct" contact between them $\rightarrow$ SWAP gate. *Entanglement swapping:* If Alice has two entangled qubits and also two EPR/Bell states shared with Bob, she can teleport both of her qubits to Bob whereupon they will be entangled on Bob's end, even though the two qubits on Bob's end were never in causal contact with each other.

## 12.3 GHZ states

There is a 3-qubit analogue of the Bell pair called a *GHZ state:*

$$\frac{|000\rangle\,|111\rangle}{\sqrt{2}}$$

Suppose three players share a GHZ state, where all three can see the state is entangled. What if Charlie is gone? Can Alice and Bob still see they are entangled? No. Charlie could have measured his qubit without Alice and Bob knowing. But if he did then Alice and Bob would have classical correlation only: either both 0s if Charlie measured zero or 1s if he measured one. Regardless of whether Charlie measured or not, A and B cannot know if they are entangled. (Like Borromean rings.)

[14]11/11/19

## 12.4 CHSH game

### 12.4.1 Classical version

We put Alice and Bob in separate rooms and each are given a challenge bit ($x$ and $y$ respectively) by a referee. The challenge bits are chosen uniformly at random and independently of each other. Then Alice sends an answer bit $a$ back to the referee and Bob sends back bit $b$. Alice and Bob win the game if

$$a + b = xy \mod 2 \iff \begin{cases} a = b & \text{if } x = 0 \text{ or } y = 0 \\ a \neq b & \text{if } x = y = 1 \end{cases}$$

Alice and Bob know the rules and are allowed to strategize together beforehand.

**Optimal strategy:** Alice and Bob always return $a = 0$ and $b = 0$. They win with probability 3/4. (They win on $(x, y) \in \{(0, 0), (0, 1), (1, 0)\}$ and lose on $(x, y) = (1, 1)$.)

### 12.4.2 Quantum version

If Alice and Bob can share a Bell pair, they can win with probability $\cos^2 \frac{\pi}{8} \approx 85\%$

[15] App. B.

[16] Mid-project presentations.

[17] The strategy involves A and B measureing their respective qubits in *different bases* depending on whether their input bits $x$ and $y$ are 0 or 1, and then outputting bits $a$ and $b$ respectively based on the outcomes of those measurements. Let's consider the case where A gets $x = 0$ and she measures $|0\rangle$, she will output $a = 0$ and A and B will win iff Bob outputs $b = 0$. Given that A measured her qubit already Bob's qubit has collapsed to the state $|0\rangle$. First suppose $y = 0$. Bob measures $|0\rangle$ in a basis rotated by $\pi/8$. He outputs 0 and they win. Probability of Bob outputting zero is $\cos^2\left(\frac{\pi}{8}\right) \approx 85\%$.

### 12.4.3 Einstein-certified randomness

In crypto, one of the most important tasks is to create random numbers.

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle \iff n \text{ assumptions}$$

Once we have quantum mechanics, you might think that the solution is trivial. After all, you get a random number by measuring the $|+\rangle$ state in the computational basis. For cryptography, though, this is insufficient. The Bell inequality lets us certify numbers as being truly random under very weak assumptions, namely, that information cannot travel faster than the speed of light.

Given a black box device that generates random numbers, we use one to generate a bit $x$ and another to generate a bit $y$, and then just play the CHSH game. If the win probability is $\cos^2\left(\frac{\pi}{8}\right)$, it's perfectly random. Anything above 3/4 is certifiably random

[15] 11/13/19

[16] 11/18/19–11/20/19

[17] 11/25/19

to some degree. Now suppose you have two boxes, designed by your enemies, that share quantum entanglement. Referee sends random $x, y$, boxes return $a, b$. Play CHSH game: if $\Pr(\text{success}) = \cos^2\left(\frac{\pi}{8}\right)$, we declare the string truly random.

Shannon entropy: $\sum p_x \log_2 \frac{1}{p_x}$ where $p_x$ is the probability that string $x$ will be in general less than $n$. Problem: we need the input bits to be uniformly random in order to play CHSH game. I.e. $x, y$ truly random, $a, b$ not truly random.

Pironio et al.: You don't have to give Alice and Bob perfectly random bits every time. Instead, you can input $x = y = 0$ most of the time, and occasionally input some random $x$'s and $y$'s to prevent Alice and Bob from figuring out the pattern. This is because Alice and Bob have no way of knowing whether each round is for testing or randomness generation.

How much entropy can we get out per bit of entropy put in?

- Colbeck et al.: You can get $cn$ bits out per $n$ bits in $(c > 1)$.
- Pironio et al.: $n^2$.
- Vazirani et al.: $\exp(\sqrt{n})$.

How many random numbers do I need to jump start this? One naive way is to feed Alice and Bob the same true random bits that they themselves create → it is not random to them. Doesn't work.

- Courdon & Yuan: Suppose we only have a fixed number of A and B players. (e.g. suppose $A, B \to$ referee $\to C, D$).
- Aaronson & Gross: Few tens of thousands (upper bound).

Reichardt et al. (2012): As long as we have two quantum computers, and as long as they are entangled with each other but unable to exchange messages, we can use CHSH game to *verify* that the computers are behaving as expected.

# 13 Computational complexity

[18]The study of how many resources are required to solve various problems. Divide problems into "easy" and "hard." Easy problems are those for which the resources needed to solve them grow only like some polynomial in input size. E.g. multiplication of $n$-bit integer → add & shift → $n \times 2n$ bitwise addition, which we say is $\mathcal{O}(n^2)$.

Now consider factoring $n$-bit number, $x$. The simplest algorithm is to check all the numbers up to $\sqrt{x}$. This requires $2^{n/2}$ checks → exponentially difficult in $n \implies$ hard.

$$n \to \begin{cases} \text{factoring} & 2^{n/2} \to \text{hard} \\ \text{multiplication} & n^2 \to \text{easy} \end{cases}$$


Figure 3: Potential relationship between the complexity classes.

## 13.1 Complexity classes

See Fig. 3.

- P (polynomial time): Class of decision problems which we can solve in polynomial time ($\implies$ polynomial memory).
- NP (nondeterministic polynomial): Class of problems which we *may* not be able to solve but we can *verify* a solution in polynomial time (and $\therefore$ polynomial memory) if it is given to us. $P \subseteq NP$.
- NP-hard: Problems which are at least as hard as *anything* in NP. That is, if someone were to give you a box that efficiently solved an NP-hard problem, you could use it to solve any problem in NP in polynomial time.
- NP-complete: Problems which are both NP-hard and are themselves in NP.
- NP-intermediate: Problems which are in NP but are not in P or in NP-complete.
- BQP (bounded-error quanutm polynomial): Class of decision problems solvable in polynomial time on a quantum computer with a probability of error less than 1/3. It seems BQP overlaps with NP-intermediate (e.g. Shor's algorithm).
- PSPACE (polynomial space): Class of problems solvable using polynomial memory but perhaps requiring exponential time.

[18] 12/2/19

### 13.2 3-SAT

See Ref. 7.

3-SAT is a very simple NP-complete problem. We are given a boolean epression which is a big AND ($\wedge$) of clauses,

$$E = C_0 \wedge C_1 \wedge \cdots \wedge C_{m-1},$$

where each clause $C_i$ is an OR ($\vee$) of three literals,

$$E = \underbrace{(a \vee b \vee c)}_{C_0} \wedge \underbrace{(\overline{a} \vee \overline{b} \vee \overline{c})}_{C_1} \wedge \underbrace{(a \vee \overline{b} \vee \overline{c})}_{C_2} \cdots$$

Is there an assignment of the variables so that $E$ is true?

Independent set decision problem (ISDP): Given a graph $G$ and a number $k$, can we find a set of $k$ vertices in $G$ such that there are no edges between any two of the vertices? We want to use 3-SAT to prove that ISDP is NP-complete. First note if you are given a set of $k$ vertices you can easily check to verify that there are no edges between two nodes in the set. Second, map 3-SAT to ISDP:

1. Turn each clause into 3 nodes and label the nodes with their literals.
2. For every pair of nodes with the same, but negated literals (e.g. $b$ and $\overline{b}$), add an edge between that pair of nodes.
3. Any independent set of size $k$ (corresponding to the number of clauses in the 3-SAT expression) will correspond to an assignment of the literals for which the 3-SAT expression is true.

# A Python & Qiskit: Installation & Introduction

To get started with Python and Qiskit, have a look at Ref. 10.

# B Guest lecture: Quantum algorithms

Guest lecture by Ali Javadi, IBM.

The hidden shift problem, also known as Simon's problem, is a member of a broader class of so-called "hidden subgroup" problems. Given a function $f$ defined over all binary strings of length $n$ such that

$$f(x) = f(x \oplus s) \quad \forall x,$$

where $\oplus$ denotes bitwise addition, the challenge is to find the value of the bitstring $s$, known as the "hidden shift." While the function $f$ is unknown, you are provided with an oracle that can evaluate $f$. In other words, you can ask the oracle the value of $f(x)$ for any $x$.

Classically, this problem requires $\mathcal{O}(2^{n/2})$ queries to the oracle in order to deduce $s$ with high probability. Using a quantum algorithm (such as Simon's algorithm), this problem may be solved in $\mathcal{O}(n)$ queries, an exponential speedup.

Ref. 5 demonstrates how to solve for the hidden shift of a "bent" function in Qiskit.

# C Additional Problems

## Q1

Any measurement performed in qiskit is always done along the z-axis of the Bloch sphere, it is however not always the relevant axis in practice. Using adequate operation on the qubit before the measurement, can you figure out how deduce the projection of the qubit state on all three axis of the Bloch Sphere (x, y, z) ?

The answer should rely on the use of the qasm simulator which reflects the actual behavior of the quantum hardware. However feel free to use the state vector simulator to figure out the impact of operations on the qubit.

## Solution

It is possible to measure the qubit in any of the three axes of the Bloch sphere by performing a rotation or change of basis right before the measurement. First, the matrix for the change of basis from the $\sigma_z$ basis to the $\sigma_x$ basis is the Hadamard gate:

$$U(\sigma_z \to \sigma_x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

The matrix of change of basis from the $\sigma_z$ basis to the $\sigma_y$ one is:

$$U(\sigma_z \to \sigma_y) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

We want to prove that is the case using Qiskit. To do so we need a quantum circuit that firstly takes the initial quantum state $|0\rangle$ and puts it in any desired state $|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$, then performs either rotation $U(\sigma_z \to \sigma_x)$ or $U(\sigma_z \to \sigma_y)$ and then performs a measurement. To perform both the state preparation and then the rotation we will use the physical gate:

$$U_3(\alpha, \beta, \lambda) = \begin{pmatrix} \cos(\alpha/2) & e^{-i\lambda}\sin(\alpha/2) \\ e^{i\beta}\sin(\alpha/2) & e^{i(\lambda+\beta)}\cos(\alpha/2) \end{pmatrix}.$$

In this language the state preparation is done by applying $U_3(\alpha = \theta, \beta = \phi, \lambda = 0)$ and the rotations as: $U(\sigma_z \to \sigma_x) = U_3(\alpha = -\pi/2, \beta = \pi, \lambda = 0)$ and $U(\sigma_z \to \sigma_y) = U_3(\alpha = \pi/2, \beta = 0, \lambda = \pi/2)$:

In order to test our circuit, we will run it with the qasm simulator of Qiskit (collecting $2 \cdot 10^3$ measurements for the statistics) and then compare the output probabilities with the analytical ones. Consider we start with state $|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$, the probabilities of measuring it in the states of the basis of $\sigma_x$ and $\sigma_y$ are:

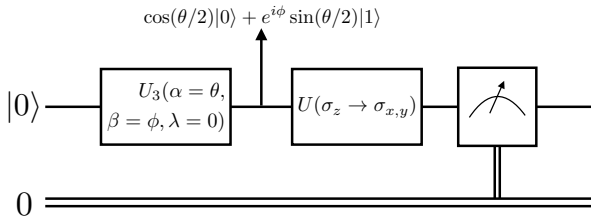



Figure 4: Quantum circuit used.

$$
\begin{aligned}
P\left(|+\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) &= \frac{1}{2}\left(1 + \sin(\theta)\cos(\phi)\right); \\
P\left(|-\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right) &= \frac{1}{2}\left(1 - \sin(\theta)\cos(\phi)\right); \\
P\left(|+\rangle_y = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)\right) &= \frac{1}{2}\left(1 - \sin(\theta)\sin(\phi)\right); \\
P\left(|-\rangle_y = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)\right) &= \frac{1}{2}\left(1 + \sin(\theta)\sin(\phi)\right).
\end{aligned}
$$

In order to compare the analytical probabilities versus the simulated ones, with the circuit in Figure 4 we generated 200 random initial states (random uniformly distributed values of $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$) and then measured them either in the $\sigma_x$ or $\sigma_y$ basis. Figure 5 shows simulated versus analytical probabilities when measuring the random states in $\sigma_x$ and $\sigma_y$ basis. It shows that, up to small statistical deviations from the simulated probabilities, the circuit and analytical values exactly agree in all cases. This proves that indeed we can measure the states in any basis we want by rotating the state before the measurement.

## Q2

Coherent state is nothing more than a displaced vacuum state:

$$|\alpha\rangle = U(\alpha)\,|0\rangle$$

Find the transformed operators in terms of original ones (a and a dagger):

$$\tilde{a} = U_\alpha^\dagger a U_\alpha$$

$$\tilde{a}^\dagger = U_\alpha^\dagger a^\dagger U_\alpha$$

Remember:

$$a\,|\alpha\rangle = \alpha\,|\alpha\rangle$$


Figure 5: Probabilities simulated from circuit in Figure 4 versus the analytical expected values. Blue dots indicate the probabilities and the dashed orange line is for visual guidance. **a)** Measuring the initial random states in the $\sigma_x$ basis.**b)** Measuring the initial random states in the $\sigma_y$ basis.

## Solution

A coherent state $|\alpha\rangle$ may be defined as the unique eigenstate of the annihilation operator $a$ with corresponding eigenvalue $\alpha \in \mathbb{C}$,

$$a\,|\alpha\rangle = \alpha\,|\alpha\rangle\,.$$

The formal solution to this eigenvalue problem is the vacuum state $|0\rangle$ displaced in phase space by $\alpha$:

$$|\alpha\rangle = D(\alpha)\,|0\rangle\,, \quad \text{with} \quad D(\alpha) = e^{\alpha a^\dagger - \alpha^* a}, \tag{9}$$

where $D(\alpha)$ is the displacement operator.[19] It is easy to derive the former relations by

[19] $D(\alpha) \equiv U_\alpha$, the notation given in the problem statement.

writing $|\alpha\rangle = \sum_n \psi_n |n\rangle$ and plugging into the eigenvalue problem:

$$a\,|\alpha\rangle = \alpha\,|\alpha\rangle \implies \sum_n \psi_n a|n\rangle = \sum_n \alpha\psi_n|n\rangle \implies \sum_{n=1}^{\infty} \psi_n\sqrt{n}|n-1\rangle = \sum_{n=0}^{\infty} \alpha\psi_n|n\rangle$$

which implies the recursion relation

$$\psi_{n+1}\sqrt{n+1} = \alpha\psi_n \implies \psi_n = \frac{\alpha^n}{\sqrt{n!}}\psi_0.$$

Therefore, the state $|\alpha\rangle$ can be written as:

$$|\alpha\rangle = \psi_0 \sum_n \frac{\alpha^n}{\sqrt{n!}}\,|n\rangle = \psi_0 \sum_n \left(\frac{(\alpha a^\dagger)^n}{n!}\right)|0\rangle = \psi_0 e^{\alpha a^\dagger}\,|0\rangle\,.$$

In principle, we could think of writing $D(\alpha) = e^{\alpha a^\dagger}$, however this operator is not unitary and therefore it does not preserve the norms of states. In the previous derivation, the normalization of the wave function is encoded in $\psi_0$, but we can just take $D(\alpha)$ to be unitary and therefore we will not have to worry about $\psi_0$. The easiest way to make that operator unitary is to multiply by the inverse of its complex conjugate, $e^{-\alpha^* a}$. The inverse of the complex conjugate will not change its action over $|0\rangle$ as $e^{-\alpha^* a}\,|0\rangle = |0\rangle$. Therefore:

$$|\alpha\rangle = e^{\alpha a^\dagger - \alpha^* a}\,|0\rangle \implies D(\alpha) = e^{\alpha a^\dagger - \alpha^* a}$$

Finally, we wish to compute $\tilde{a} = D^\dagger(\alpha) a D(\alpha)$ and $\tilde{a}^\dagger = D^\dagger(\alpha) a^\dagger D(\alpha)$. Using the expression for the displacement operator (9), we have

$$\begin{aligned} \tilde{a} &= e^{\alpha^* a - \alpha a^\dagger} a e^{\alpha a^\dagger - \alpha^* a} \\ \tilde{a}^\dagger &= e^{\alpha^* a - \alpha a^\dagger} a^\dagger e^{\alpha a^\dagger - \alpha^* a}. \end{aligned} \tag{10}$$

These are easily computed using the Baker-Hausdorff lemma [9, Eq. 2.3.47],

$$e^{iG\lambda} A e^{-iG\lambda} = A + i\lambda[G,A] + \frac{(i\lambda)^2}{2!}[G,[G,A]] + \cdots + \frac{(i\lambda)^n}{n!}[G,[G,[G,\ldots[G,A]]]\ldots] + \cdots, \tag{11}$$

for some operator $A$, where $G$ is Hermitian and $\lambda$ is a real parameter. Identifying $iG = \alpha^* a - \alpha a^\dagger$ and $\lambda = 1$, we compute the commutators

$$\begin{aligned} [G,a] &= -i\left[\alpha^* a - \alpha a^\dagger, a\right] &&= i\alpha\left[a^\dagger, a\right] = -i\alpha \\ \left[G,a^\dagger\right] &= -i\left[\alpha^* a - \alpha a^\dagger, a^\dagger\right] &&= -i\alpha^*\left[a, a^\dagger\right] = -i\alpha^* \end{aligned}$$

where we have used $\left[a, a^\dagger\right] = 1$. As these are each proportional to the identity, all subsequent commutators in (11) vanish, so (10) reduces to

$$\begin{aligned} \tilde{a} &= a + \alpha \\ \tilde{a}^\dagger &= a^\dagger + \alpha^*. \end{aligned}$$

## Q3

Consider a phenomenological Hamiltonian:

$$H_T = -\frac{E_J}{2}\sum_N \left[|N+1\rangle\langle N| + |N\rangle\langle N+1|\right]$$

as a representation of the microscopic tunneling of Cooper pairs across the JJ, where state $|N\rangle$ represents the integer N number of excess Cooper pairs in the "box". Let's also define a set of phase states $|\delta\rangle$ which are the Fourier dual to the number states $|N\rangle$, $|\delta\rangle = \sum_N e^{iN\delta}|N\rangle$. (Note that the phase $\delta$ lives on the compact space $[0\ 2\pi]$ as the Cooper pair number N on the "island" is a discrete integer).

(a) Show that the states $|\delta\rangle$ are eigenstates of $\hat{H}$.

(b) Find the associated eigenvalues.

c) Find the group velocity:

$$v_g \equiv \frac{1}{\hbar}\frac{\partial\lambda_{T,\delta}}{\partial\delta}.$$

## Solution

The tunneling of Cooper pairs into or out of a Cooper pair box via a Josephson junction may be represented by the Hamiltonian

$$H_T = -\frac{E_J}{2}\sum_N \left[|N+1\rangle\langle N| + |N\rangle\langle N+1|\right] \tag{12}$$

where $E_J$ is the Josephson energy, $|N\rangle$ is the state of the box containing an excess of $N$ Cooper pairs relative to the number of Cooper pairs outside the box,[20] and the sum ranges over all integers $N \in \mathbb{Z}$. The phase states $|\delta\rangle$ are Fourier dual to the number states $|N\rangle$,

$$|\delta\rangle = \sum_N e^{iN\delta}|N\rangle. \tag{13}$$

The action of the Hamiltonian (12) on the phase states (13) yields

$$\begin{aligned}
H_T|\delta\rangle &= -\frac{E_J}{2}\sum_N\sum_{N'} e^{iN'\delta}\left\{|N+1\rangle\langle N|N'\rangle + |N\rangle\langle N+1|N'\rangle\right\} \\
&= -\frac{E_J}{2}\sum_N\sum_{N'} e^{iN'\delta}\left\{|N+1\rangle\,\delta_{N,N'} + |N\rangle\,\delta_{N+1,N'}\right\} \\
&= -\frac{E_J}{2}\sum_N \left\{e^{iN\delta}|N+1\rangle + e^{i(N+1)\delta}|N\rangle\right\} \qquad (14) \\
&= -\frac{E_J}{2}\sum_N \left\{e^{i(N-1)\delta}|N\rangle + e^{i(N+1)\delta}|N\rangle\right\} \qquad (15) \\
&= -\frac{E_J}{2}(e^{-i\delta} + e^{+i\delta})\sum_N e^{iN\delta}|N\rangle \\
&= -E_J\cos\delta\,|\delta\rangle
\end{aligned}$$

[20] For $N < 0$, this represents a *deficit* of $N$ Cooper pairs in the box, relative to the number outside the box.

where, in going from (14) to (15), we were able to take $N \to N-1$ in just the first term because the sum ranges from $-\infty$ to $\infty$. Therefore, the phase states $|\delta\rangle$ are eigenstates of the tunneling Hamiltonian $H_T$ with corresponding eigenvalues $-E_J \cos\delta$, for $\delta \in [0, 2\pi]$.

Assuming $\lambda_{T,\delta} = -E_J \cos\delta$, i.e. the energy of state $|\delta\rangle$ associated with tunneling, the group velocity is

$$v_g \equiv \frac{1}{\hbar}\frac{\partial \lambda_{T,\delta}}{\partial \delta} = \frac{E_J}{\hbar}\sin\delta.$$

## Q4

Can you always take the square root of a unitary matrix? That is, given a unitary $U$, does there always exist a $\sqrt{U}$ such that $\sqrt{U}\sqrt{U} = U$?

## Solution

Yes, but the square root is not unique. By definition, a complex square matrix $U$ is unitary iff its conjugate transpose is also its inverse:

$$U \text{ is unitary} \iff U^\dagger = U^{-1}, \tag{16}$$

where the inverse of $U$, denoted $U^{-1}$, is defined as the matrix that, when left- or right-multiplied by $U$, produces the identity:

$$U^{-1} \text{ is the inverse of } U \iff UU^{-1} = U^{-1}U = \mathbb{1}. \tag{17}$$

It immediately follows that $U$ is normal. By definition, a complex square matrix $N$ is normal iff it commutes with its conjugate transpose:

$$N \text{ is normal} \iff NN^\dagger = N^\dagger N. \tag{18}$$

Therefore, by definitions (16) and (17),

$$UU^\dagger = \mathbb{1} = U^\dagger U \implies U \text{ is normal.}$$

Furthermore, a matrix is normal iff it is unitarily diagonalizable.[21] Consider the forward direction. By the Schur decomposition, any complex square matrix $U$ can be written

$$U = VDV^\dagger,$$

where $V$ is unitary and $D$ is upper-triangular. Then $UU^\dagger = V(DD^\dagger)V^\dagger$ by unitarity of $V$, and likewise $U^\dagger U = V(D^\dagger D)V^\dagger$. If $U$ is normal, then $DD^\dagger = D^\dagger D$ and $D$ is also normal. Hence $D$ is in fact diagonal, since a normal upper-triangular matrix can only have nonzero elements along the diagonal (this follows trivially from the definition (18) and upper-triangularity). Thus $U$ is unitarily diagonalizable.

[21] This is a version of the spectral theorem.

Given that $U$ can be unitarily diagonalized and given the fact that $U^\dagger U = 1$, this implies that the eigenvalues of $U$ can be written as:

$$\lambda_j = e^{i\phi_j},$$

where $\lambda_j$ designates the $j^{th}$ eigenvalue of $U$. Then,

$$D = \begin{pmatrix} \ddots & & \\ & \lambda_j & \\ & & \ddots \end{pmatrix}$$

In the basis where $U$ is diagonal, its square root is obtained simply by taking the square root of the eigenvalues, i.e.

$$D^{1/2} = \begin{pmatrix} \ddots & & \\ & \sqrt{\lambda_j}_{\pm} & \\ & & \ddots \end{pmatrix}.$$

However, for each eigenvalue, there are two possible values of the square root:

$$\sqrt{\lambda_j}_+ = e^{i\phi_j/2} \text{ and } \sqrt{\lambda_j}_- = e^{i(\phi_j/2+\pi)}.$$

If matrix $U \in \mathbb{C}^{N\times N}$, then there are $2^N$ possible combinations of $\sqrt{\lambda_j}_+$ and $\sqrt{\lambda_j}_-$. Therefore, a unitary matrix has a square root, but there are $2^N$ different $\sqrt{U}$. By construction, in the original basis

$$\sqrt{U} = V D^{1/2} V^\dagger.$$

## Q5

Prove the following points:

a) Two quantum states will lead to different probabilities iff they have different density matrices.

b) No-communication theorem: if Alice and Bob share an entangled state, nothing Alice chooses to do will have any effect on Bob's density matrix → There is no observable effect on Bob's end.

## Solution

a) Suppose we measure a mixed state $(p_i, |\psi_i\rangle)$ in an orthonormal basis $|\beta_k\rangle$. Let's calculate the outcome:

We denote the probability of measuring $|\beta_k\rangle$ by $\Pr[k]$. Then

$$\begin{aligned}\Pr[k] &= \sum_j p_j \mid \langle\psi_j|\beta_k\rangle \mid^2 \\ &= \sum_j p_j \langle\beta_k|\psi_j\rangle\langle\psi_j|\beta_k\rangle \\ &= \langle\beta_k| \sum_j p_j \, |\psi_j\rangle\langle\psi_j| \, |\beta_k\rangle \\ &= \langle\beta_k| \, \rho \, |\beta_k\rangle\end{aligned}$$

Therefore different probabilities (LHS) will lead us different density matrices (RHS) and vice versa.

b) Consider an arbitrary bipartite state,

$$\rho = \sum_{ijk\ell} \rho_{ijk\ell} \underbrace{|i\rangle\langle j|}_{\text{A}} \otimes \underbrace{|k\rangle\langle\ell|}_{\text{B}},$$

where we take the bases in each subspace to be orthonormal without loss of generality. The reduced density matrix for Bob is[22]

$$\begin{aligned}\rho_B &= \mathrm{Tr}_A \, \rho \\ &= \sum_m \langle m|\rho|m\rangle && \text{note:} \quad |m\rangle \in \mathcal{H}_A \\ &= \sum_{ijk\ell m} \rho_{ijk\ell} \langle m|i\rangle\langle j|m\rangle \, |k\rangle\langle\ell| && \text{note:} \quad \langle m|i\rangle = \delta_{mi}, \ \langle j|m\rangle = \delta_{jm} \\ &= \sum_{k\ell m} \rho_{mmk\ell} \, |k\rangle\langle\ell| \, . && (19)\end{aligned}$$

Suppose Alice performs a local unitary operation, $U : \mathcal{H}_A \to \mathcal{H}_A$, on her subsystem. The new bipartite state becomes

$$\begin{aligned}\rho' &= U\rho U^\dagger \\ &= \sum_{ijk\ell} \rho_{ijk\ell} U \, |i\rangle\langle j| \, U^\dagger \otimes |k\rangle\langle\ell| \, ,\end{aligned}$$

[22]The trace is basis-independent. For simplicity and without loss of generality, we take the trace with respect to the same orthonormal basis in which Alice's subspace is decomposed.

yet Bob's reduced density matrix remains invariant:

$$\begin{aligned}
\rho_B &= \mathrm{Tr}_A\, \rho' \\
&= \sum_{ijk\ell m} \rho_{ijk\ell} \langle m|U|i\rangle \langle j|U^\dagger|m\rangle |k\rangle\langle\ell| && \text{commute scalar matrix elements} \\
&= \sum_{ijk\ell m} \rho_{ijk\ell} \langle j|U^\dagger|m\rangle \langle m|U|i\rangle |k\rangle\langle\ell| && \text{remove identity} \quad \sum_m |m\rangle\langle m| = \mathbb{1} \\
&= \sum_{ijk\ell} \rho_{ijk\ell} \langle j|U^\dagger U|i\rangle |k\rangle\langle\ell| && \text{by unitarity,} \quad U^\dagger U = \mathbb{1} \\
&= \sum_{ijk\ell} \rho_{ijk\ell} \langle j|i\rangle |k\rangle\langle\ell| && \text{by orthonormality,} \quad \langle j|i\rangle = \delta_{ji} \\
&= \sum_{jk\ell} \rho_{jjk\ell} |k\rangle\langle\ell| \,. && (20)
\end{aligned}$$

In either case, Bob's reduced density matrix (19) or (20) is the same.

In particular, even if $\rho$ is maximally entangled (and therefore pure), there is no operation Alice can perform on her subsystem that will be observable by Bob. Therefore, entanglement does not enable faster-than-light communication.

## Q6

Can different probability distributions over pure states give rise to the same mixed state?

## Solution

Yes. Consider a uniform distribution over the $\sigma_z$ eigenstates, $\left\{(\frac{1}{2}, |0\rangle), (\frac{1}{2}, |1\rangle)\right\}$. This gives rise to the mixed state

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

in the $\sigma_z$ eigenbasis. However, the identity operator $\mathbb{1}$ has the same representation, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, in every orthonormal basis. For instance, consider a uniform distribution over the $\sigma_x$ eigenstates, $\left\{(\frac{1}{2}, |+\rangle), (\frac{1}{2}, |-\rangle)\right\}$. This gives rise to the mixed state

$$\rho = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -| = \frac{1}{2}\mathbb{1},$$

now in the $\sigma_x$ eigenbasis. But, noting $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$,

$$\begin{aligned}
\mathbb{1} &= |+\rangle\langle +| + |-\rangle\langle -| \\
&= \frac{1}{2}\left(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|\right) \\
&\quad + \frac{1}{2}\left(|0\rangle\langle 0| - |0\rangle\langle 1| - |1\rangle\langle 0| + |1\rangle\langle 1|\right) \\
&= |0\rangle\langle 0| + |1\rangle\langle 1| = \mathbb{1}.
\end{aligned}$$

This is obvious if we consider the unitary transformation $U$ connecting the two bases, whereby $U\mathbb{1}U^\dagger = UU^\dagger = \mathbb{1}$.

Although one might think that a uniform distribution over $\{|0\rangle, |1\rangle\}$ is physically distinct from a uniform distribution over $\{|+\rangle, |-\rangle\}$—corresponding to, say, an ensemble of spins with half aligned along $+\hat{z}$ and the other half along $-\hat{z}$, versus an ensemble with half along $+\hat{x}$ and half along $-\hat{x}$—in fact, there is no measurement which can distinguish between the two.

## Q7

Find a maximally mixed state of a 2-qubit system.

## Solution

Given the computational basis for a two qubit system, $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, the maximally mixed state is an ensemble mix of all the states with equal probability:

$$\rho = \frac{1}{4}(|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10| + |11\rangle\langle 11|)$$

It is trivial to see that the rank of the density matrix is 4, as for a maximally entangled state. In general, for a system of $n$ qubits, with computational basis elements $\{|i\rangle : i = 0, 1, ..., 2^n - 1\}$, the maximally mixed state is: $\rho = \mathbb{1}/2^n$, which has rank $2^n$. The von Neumann entropy of these states is maximal and follows a volume law:

$$S = -\operatorname{Tr}(\rho \log_2 \rho) = -\sum_{i=0}^{2^n-1} \frac{1}{2^n} \log_2 \frac{1}{2^n} = n.$$

## Q8

Implement superdense coding in Qiskit. Attach the code and results.

## Solution

In this case Alice has two classical bits she wants to send to Bob by just sending one quantum bit. In the following implementation we will map all 4 possible combinations of the two bits to integers from 0 to 3: 00 → 0, 01 → 1, 10 → 2 and 11 → 3.

The algorithm proceeds as follows:

1. Alice decides which combination of two classical bits she wants to send between possibilities 0 to 3. In the qiskit implementation this is stored in variable "bit". In this example Alice wants to send 11 to Bob:

```
#Usual imports
import qiskit
import numpy as np

bit = 3

```

2. Alice and Bob share a maximally entangled qubit pair: $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0_A 0_B\rangle + |1_A 1_B\rangle)$. In this notation the most significant bit corresponds to Alice. The initial state of the qubits is $|00\rangle$. Therefore, to put the system into state $|\psi_0\rangle$ we apply a Hadamard gate on Alice's qubit and a CNOT(A→B) gate where the control qubit is Alice's and the target one is Bob's ($|\psi_0\rangle = \text{CNOT(A} \rightarrow \text{B})H(A)\,|00\rangle$):

```
#Define the quantum circuit with two classical
#registers and two qubits
qc = QuantumCircuit(2,2)

#Create the entangled pair shared by Alice and Bob
qc.h(0) #Hadamard on 0
qc.cx(0,1) #CNOT where 0 is the control and 1 is target

```

3. Alice performs one operation on her qubit depending on what pair of classical bits she wants to send:

$$
\begin{aligned}
00 &\rightarrow \mathbb{1}_A \otimes \mathbb{1}_B\,|\psi_0\rangle = |\psi_{00}\rangle\,\frac{1}{\sqrt{2}}(|0_A 0_B\rangle + |1_A 1_B\rangle)\\
01 &\rightarrow X_A \otimes \mathbb{1}_B\,|\psi_0\rangle = |\psi_{01}\rangle\,\frac{1}{\sqrt{2}}(|0_A 1_B\rangle + |1_A 0_B\rangle)\\
10 &\rightarrow Z_A \otimes \mathbb{1}_B\,|\psi_0\rangle = |\psi_{10}\rangle\,\frac{1}{\sqrt{2}}(|0_A 0_B\rangle - |1_A 1_B\rangle)\\
11 &\rightarrow ZX_A \otimes \mathbb{1}_B\,|\psi_0\rangle = |\psi_{11}\rangle\,\frac{1}{\sqrt{2}}(|0_A 1_B\rangle - |1_A 0_B\rangle).
\end{aligned}
$$

   Which in Qiskit is implemented as follows:

```
if bit == 1: #Sending 01
    qc.x(0)
if bit == 2: #Sending 10
    qc.z(0)
if bit == 3: #Sending 11
    qc.x(0)
    qc.z(0)

```

4. Alice sends her qubit to Bob over a quantum channel. Nothing to do on Qiskit.

5. Bob, who now has the two qubits, applies a CNOT(A→B) gate and then applies a Hadamard gate to the qubit that used to belong to Alice. Bob obtains the following set of orthogonal states depending on what qubit Alice sent him:

$$
\begin{aligned}
00 &\rightarrow H_A \otimes \mathbb{1}_B \cdot \text{CNOT(A} \rightarrow \text{B})\,|\psi_{00}\rangle = |00\rangle\\
01 &\rightarrow H_A \otimes \mathbb{1}_B \cdot \text{CNOT(A} \rightarrow \text{B})\,|\psi_{01}\rangle = |01\rangle\\
10 &\rightarrow H_A \otimes \mathbb{1}_B \cdot \text{CNOT(A} \rightarrow \text{B})\,|\psi_{10}\rangle = |10\rangle\\
11 &\rightarrow H_A \otimes \mathbb{1}_B \cdot \text{CNOT(A} \rightarrow \text{B})\,|\psi_{11}\rangle = |11\rangle\,.
\end{aligned}
$$

Bob finally measures the two qubits in the computational basis, obtaining precisely the two classical bits Alice wanted to send. This final step with the simulation of the whole process is implemented in Qiskit:

```
    qc.cx(0,1) #Applying the CNOT
    qc.h(0) #Applying the Hadamard to Alice's qubit
    qc.measure(0,0) #Measuring qubit that used to belong to Alice
    qc.measure(1,1) #Measuring qubit that always belonged to Bob

    #Simulation using qasm simulator
    Measurements = 2000 #Number of measurements
    emulator = Aer.get_backend('qasm_simulator')
    job = execute( qc, emulator, shots = Measurements )
    hist = job.result().get_counts() #results of the measurement
    print(hist) #Print results

```

The code will be attached in the email as a "jupyter notebook". For the results, even using the "qasm" simulator, feeding the program with "bit" = 0 gave output 00 with probability exactly 1. Feeding the program with "bit" = 1 gave output 10 with probability exactly 1. Feeding the program with "bit" = 2 gave output 01 with probability exactly 1, and feeding the program with "bit" = 3 gave output 11 with probability exactly 1. Note that the output for "bit" = 1 and "bit" = 2 are swapped. This is because Qiskit takes the zeroth qubit to be the least significant one. However, the results are correct.

## Q9

What choice of $\lambda_1, \ldots, \lambda_n$ maximizes the entropy,

$$S(\rho) = -\sum_k \lambda_k \log_2 \lambda_k,$$

where $\sum_k \lambda_k = 1$.

## Solution

The von Neumann entropy expressed in units of bits,

$$S(\rho) = -\sum_k \lambda_k \log_2 \lambda_k, \tag{21}$$

may be straightforwardly extremized over all diagonal density operators $\rho$ with eigenvalues $\lambda_k$, subject to the probability conservation constraint,

$$\sum_k \lambda_k = 1, \tag{22}$$

by the method of Lagrange multipliers. We form the Lagrange function

$$\mathcal{L}(\{\lambda_k\}, \eta) = S(\rho) - \eta \left( \sum_k \lambda_k - 1 \right),$$

and then vary each parameter $\{\lambda_k\}, \eta$ independently to find the stationary point in the parameter space. Variation with respect to $\eta$ recovers the probability constraint (22), while variation with respect to each $\lambda_j \in \{\lambda_k\}$ yields

$$\frac{\partial \mathcal{L}}{\partial \lambda_j} = -(\log_2 \lambda_j + 1) - \eta = 0, \quad \forall\ j$$

which must vanish at the stationary point. This implies

$$\lambda_j = 2^{-(\eta+1)}, \quad \forall j,$$

i.e. all $\lambda_j$'s are equal to the same constant. Enforcing the probability constraint (22) determines $\eta$ such that

$$\lambda_j = 2^{-n} = \frac{1}{D}, \quad \forall j$$

for $n$ qubits, or Hilbert space dimension $D = 2^n$. Therefore, the von Neumann entropy is maximized[23] when $\{\lambda_k\}$ is a uniform distribution, i.e. when $\rho$ is a ($\sim$ classical) maximally mixed state. Moreover, this state is unique.

In the above, we assumed $\rho$ is diagonal, but note that any $\rho$ can be diagonalized and the von Neumann entropy is basis-independent, as

$$\begin{aligned} & \rho = U\rho_d U^\dagger \\ \implies S &= -\operatorname{Tr}(\rho \ln \rho) \\ &= -\operatorname{Tr}\big(U\rho_d U^\dagger \ln[U\rho_d U^\dagger]\big) \\ &= -\operatorname{Tr}\big(U\rho_d U^\dagger [U \ln(\rho_d) U^\dagger]\big) \\ &= -\operatorname{Tr}\big(U\rho_d \ln(\rho_d) U^\dagger\big) \\ &= -\operatorname{Tr}(\rho_d \ln \rho_d) \end{aligned}$$

where the final equality follows the cyclicity of the trace.

For the sake of completeness, note that since $0 \leqslant \lambda_k \leqslant 1\ \forall k$, the von Neumann entropy $S(\rho) \geqslant 0$ according to (21). This lower bound is attained for $\lambda_k = 1$ for some particular $k$ and $\lambda_k = 0$ for all others. This occurs for any $\rho$ representing a pure state. So we conclude

$$0 \leqslant S(\rho) \leqslant \ln D$$

where the upper bound is attained for the (unique) maximally mixed state $\rho = \mathbb{1}/D$, while the lower bound is attained for every (i.e. not unique) pure state $\rho$.

## Q10

1D harmonic oscillator:

$$H\,|n\rangle = \hbar\omega(n + \frac{1}{2})\,|n\rangle$$

$$a\,|\alpha\rangle = \alpha\,|\alpha\rangle$$

$$|\alpha\rangle\,, (|\alpha|, n)$$

[23] One can verify this extremum is a maximum using a bordered Hessian test.

a) Time evolution of coherent state.

b) Calculate:

$$\langle x(t)\rangle = \langle \alpha(t)|x|\alpha(t)\rangle$$
$$\langle p(t)\rangle = \langle \alpha(t)|p|\alpha(t)\rangle$$

c) Consider the mixed state:

$$\rho = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\phi \, ||\alpha|\exp(i\phi)\rangle\langle|\alpha|\exp(i\phi)| \,,$$

a mixture of coherent states with random phase and fixed amplitude $|\alpha|$. Show this is equivalent to a mixture of coherent states with a Poisson distribution with $\bar{n} = |\alpha|^2$.

## Solution

a) We first expand the coherent state $|\alpha\rangle$ in the Fock basis (derived in our solution to HW2 Problem 2), as

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \,. \tag{23}$$

The time evolution of the coherent state is then obtained by time-evolving the individual Fock-state components. These are energy eigenstates of the harmonic oscillator with energy $E_n = \hbar\omega(n + \frac{1}{2})$. Their time evolution follows from the Schrödinger equation as

$$|n\rangle \longrightarrow e^{-in\omega t} e^{-\frac{i\omega t}{2}} |n\rangle \,,$$

which yields the time-evolved coherent state

$$|\alpha(t)\rangle = e^{-\frac{i\omega t}{2}} e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{(\alpha e^{-i\omega t})^n}{\sqrt{n!}} |n\rangle \tag{24}$$

If we define

$$\alpha(t) = \alpha e^{-i\omega t}, \tag{25}$$

then $|\alpha(t)\rangle$ is given by Eq. (23) with $\alpha \to \alpha(t)$, times a global phase factor $\exp(-i\omega t/2)$.

b) We can write the position and momentum operator in terms of creation and annihilation operators:

$$x = \sqrt{\frac{\hbar}{2m\omega}}(a^\dagger + a), \qquad p = i\sqrt{\frac{m\hbar\omega}{2}}(a^\dagger - a).$$

Note that the eigenvalue equation for the coherent state holds at all times,

$$a\,|\alpha(t)\rangle = \alpha(t)\,|\alpha(t)\rangle \,,$$

where the eigenvalue is given by Eq. (25). It is then straightforward to compute

$$\langle\alpha(t)|x|\alpha(t)\rangle = \sqrt{\frac{\hbar}{2m\omega}}(\alpha^* e^{i\omega t} + \alpha e^{-i\omega t}) \ = x_0\cos(\omega t - \phi)$$
$$\langle\alpha(t)|p|\alpha(t)\rangle = i\sqrt{\frac{m\hbar\omega}{2}}(\alpha^* e^{i\omega t} - \alpha e^{-i\omega t}) = -m\omega x_0\sin(\omega t - \phi),$$

where we have defined $\alpha = |\alpha|e^{i\phi}$ and $x_0 = |\alpha|\sqrt{2\hbar/m\omega}$.

c) In this case we wish to compute the density matrix for the time evolution of state $|\alpha(t)\rangle$ given by Eq. (24). By definition, we wish to compute:

$$\rho = \frac{1}{2\pi}\int_0^{2\pi} d\phi \left||\alpha|e^{i\phi}\rangle\langle|\alpha|e^{i\phi}\right| = \frac{e^{-|\alpha|^2}}{2\pi}\int_0^{2\pi} d\phi \sum_{m,n} \frac{\left(|\alpha|e^{-i\phi}e^{i\omega t}\right)^n \left(|\alpha|e^{i\phi}e^{-i\omega t}\right)^m}{\sqrt{n!m!}} |m\rangle\langle n| =$$
$$= \sum_{n,m} \frac{e^{-|\alpha|^2}|\alpha|^{n+m}e^{-i\omega t(m-n)}\frac{1}{2\pi}\int_0^{2\pi} d\phi e^{i\phi(m-n)}}{\sqrt{n!m!}} |m\rangle\langle n|$$

Now, taking into account that $\frac{1}{2\pi}\int_0^{2\pi} e^{i\phi(m-n)} = \delta_{m,n}$, we end up obtaining:

$$\rho = \sum_{n=0}^{\infty} \frac{e^{-|\alpha|^2}\left(|\alpha|^2\right)^n}{n!} |n\rangle\langle n|$$

Finally recalling that Poisson distribution of average $\lambda$ is $f(\lambda, n) = \frac{e^{-\lambda}\lambda^n}{n!}$, we see that our density matrix is poisson distributed with average $|\alpha|^2$.

## Q11

a) Design your own quantum bank that produces quantum money.

b) what are the odds for hacking your quantum money?

## Solution

a) In this scheme, a piece of quantum money comprises an $n$-qubit "bill" state and a classical serial number. The quantum bank issues money by generating a bill state and assigning it a unique serial number. The description of each issued bill state and its corresponding serial number are stored in a secret ledger maintained by the bank. A vendor, upon receiving a piece of quantum money from a customer, sends the bill state and its serial number to the bank for verification. The bank uses the serial number to look up the bill state description in the ledger and compare it to the bill state received from the vendor. If they match, the transaction is approved.

b) The security and performance of this scheme hinges on how the bank produces the $n$-qubit bill state. One potential method of state generation follows. For each qubit, the bank randomly selects a basis, say $\sigma_x$ or $\sigma_z$, and randomly assigns an eigenstate of

that basis to the value of the qubit, i.e. $|0\rangle$ or $|1\rangle$ for $\sigma_z$ and $|+\rangle$ or $|-\rangle$ for $\sigma_x$. The $n$-qubit bill state can then be prepared by single-qubit rotations, with a typical bill state looking something like,

$$\underbrace{|+\ 0\ 1 - +\ +\ -0 + -1 - 0\ 0\ 1\ 0 + +\ 0\ 0 \cdots\rangle}_{n \text{ qubits}}.$$

Note that the space complexity for storing the bill description in the ledger is $\mathcal{O}(n)$, since for each qubit we need to store 1 bit specifying the basis and 1 bit specifying the eigenstate. Furthermore, forging bill states is prohibited by the no-cloning theorem, since the counterfeiter does not know the basis used for each qubit. Therefore, possession of a bill state does not allow duplicates to be produced.

Given the serial number, the only counterfeiting strategy is to attempt to guess the bill state. Assuming the counterfeiter knows the basis set $\{\sigma_x, \sigma_z\}$, the probability of guessing the correct basis is $1/2$, and the probability of guessing the correct eigenstate within that basis is also $1/2$, so the probability of getting the state of any qubit exactly correct is $1/4$. However, if the counterfeiter guesses the incorrect basis, it is still possible that the bank measures the correct eigenstate during the verification process. Since each wrong-basis eigenstate is a uniform superposition of the right-basis eigenstates, this occurs with probability $1/2$. The probability for the bank to erroneously approve a counterfeit 1-qubit bill state is then

$$\underbrace{\Pr(\text{right eigenstate}|\text{right basis})}_{1}\underbrace{\Pr(\text{right basis})}_{1/2} + \underbrace{\Pr(\text{right eigenstate}|\text{wrong basis})}_{1/2}\underbrace{\Pr(\text{wrong basis})}_{1/2} =$$

For an $n$-qubit bill state, every qubit must be approved individually, with a single wrong outcome identifying a counterfeit. The total probability for the bank to erroneously approve a counterfeit $n$-qubit bill state is then $(3/4)^n$ and is shown in Fig. 6. This scheme is therefore exponentially secure in the number of qubits used for the bill state.[24]

## Q12

There exist bipartite states that take a lot of entanglement to make but you can only extract a small fraction of the entanglement. Find an example and show $E_f \gg E_D$ can be achieved by extension for larger qubit systems.

## Solution

This question is a very active field of research in the moment and it is actually very complicated to answer. Here we will give a simple toy model answer and describe some of the recent works about this matter.

[24]Note that this security analysis is identical to that of the BB84 protocol. Our quantum money scheme is essentially the same as BB84, with the counterfeiter playing the role of the eavesdropper.

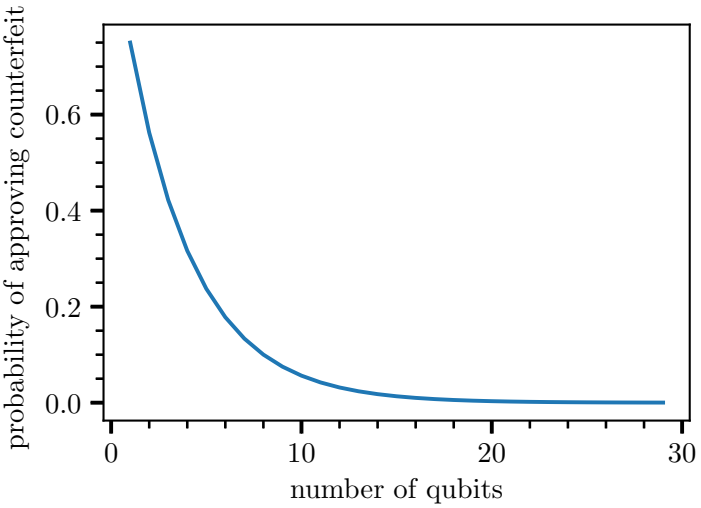



Figure 6: The probability of approving a counterfeit is $(3/4)^n$ where $n$ is the number of qubits.

One simple case of such bipartite states that take more entanglement to make than the entanglement you can measure is for example a system with 3 qubits. Suppose that subsystem A is qubit 0 and subsystem B is qubits 1 and 2. Now, imagine that qubits 1 and 2 are in a Bell state:

$$|\psi\rangle_{1,2} = \frac{|11\rangle_{1,2} + |00\rangle_{1,2}}{\sqrt{2}}$$

and qubit number 0 is for example in state $|\psi\rangle_0 = |0\rangle_0$. Then state of the whole system is:

$$|\psi\rangle_{0,1,2} = |0\rangle_0 \otimes \frac{|11\rangle_{1,2} + |00\rangle_{1,2}}{\sqrt{2}}.$$

In this case, the entanglement of formation is 1 as we created a Bell pair in qubits 1 and 2, however, the entanglement between subsystem A and subsystem B is zero and therefore we cannot extract any entanglement out of this partition of the state. Therefore $E_D < E_f$.

In the case of having many qubits, we can choose to split the system in a certain way obtaining subsystems A and B. Now imagine that qubits are laid out in a 1D chain and we form a Bell state of nearest neighbours in subsystem A and do the same in subsystem B, but we don't form a Bell pair between the qubits in the boundaries of A and B. In this case, the total number of Bell pairs we needed to form the state was of the order of the number of qubits, but subsystem A is not entangled with subsystem B and therefore the amount of entanglement I get out of this partition is zero and therefore $E_f >> E_D$.

Recent studies focus on the entanglement of formation and distillation from a more systematic point of view. Given $N$ copies of a bipartite state $\rho_{AB}$, we can imagine extracting $M$ copies of a Bell state $|\psi^-\rangle\langle\psi^-|$ using only local operations and classical communication (LOCC),

$$\rho_{AB}^{\otimes N} \xrightarrow{\text{LOCC}} \left|\psi^-\rangle\langle\psi^-\right|^{\otimes M}.$$

This process is known as entanglement concentration or distillation. The *distillable entanglement* $E_D(\rho_{AB})$ is the largest value of $M/N$ attainable as $N \to \infty$. This can be understood as the largest number of Bell states that can be distilled from the given bipartite state.

Conversely, we can imagine starting with $M$ copies of a Bell state and mixing them into $N$ copies of a given bipartite state using LOCC,

$$\left|\psi^-\right\rangle\!\left\langle\psi^-\right|^{\otimes M} \xrightarrow{\text{LOCC}} \rho_{AB}^{\otimes N}.$$

This process is known as entanglement dilution. The *entanglement cost* $E_C(\rho_{AB})$ is the *smallest* value of $M/N$ attainable as $N \to \infty$. It is therefore the asymptotic limit of the entanglement of formation. This can be understood as the smallest number of Bell states required to form the bipartite state.

Since any measure of entanglement cannot increase under LOCC, we must have $E_C(\rho_{AB}) \geqslant E_D(\rho_{AB})$. Intuitively, one might expect that entanglement distillation and dilution are asymptotically reversible, i.e.

$$\left|\psi^-\right\rangle\!\left\langle\psi^-\right|^{\otimes M} \xleftrightarrow{\text{LOCC}} \rho_{AB}^{\otimes N},$$

such that the bound is saturated as $N \to \infty$. Indeed, for pure states, this is in fact the case, and $E_C = E_D = S(\rho_A)$, the von Neumann entropy of the reduced state of $\rho_{AB}$.

For mixed states or multipartite entangled states, however, entanglement distillation/dilution is irreversible in the asymptotic limit, and consequently $E_C > E_D$. These processes can therefore be thought of as a type of lossy compression/decompression algorithm. Horodecki et al. showed the existence of so-called bound entangled states which require entanglement to create, but from which no entanglement can be distilled. Vidal and Cirac generalized this result to the asymptotic limit. Subsequently, Vollbrecht et al. showed that mixtures of maximally entangled states can be diluted and distilled reversibly if and only if they realize the lower bound of the Heisenberg uncertainty principle, demonstrating that irreversibility is a generic and common phenomenon. This remains an active area of research. For a more detailed overview of the various entanglement measures and their associated theorems, see Ref. [4].

## Q13

Qiskit: Make a circuit to generate a random number between 0 and $2^3$-1. What is the entropy? repeat for 0 and $2^{15}$-1. how much does the entropy change?

## Solution

In this case, we will design the simplest random number generator possible, even though in the literature one can find more refined algorithms that take into account that the quantum hardware can be biased (for cryptography applications). These algorithms are all based on the CHSH game with certain variations.

Here we will consider that our quantum hardware is not biased (the NSA has not included a back door) and that therefore we have a perfect quantum device that we have fabricated ourselves. In this case, the task is fairly simple, if we want a random 1-bit generator, we just

need to put one qubit in state $|+\rangle$ and measure in the computational basis outputs $|0\rangle$ or $|1\rangle$ with complete randomness and probability 1/2. Therefore, the protocol is clear if we want to generate a random number between 0 and $2^n - 1$: take $n$ qubits in the computational basis, put all of them in state $|+\rangle$ and measure them in the computational basis. Then, the result of your measurement represents the desired number in binary representation. Note that the random numbers between 0 and $2^n - 1$ will be uniformly sampled with probability $1/2^n$. In this sense, the quantum entropy is 0 as the state of the system is separable, but the classical entropy of the random number probability distribution is maximal: $S = -\sum_x p_x \log_2 p_x = n$, as $p_x = 1/2^n$ for every random number $x$ generated.

This algorithm can also be done with one qubit, where you prepare it to be in the state $|+\rangle$ and then measure it in the computational basis. The process is repeated $n$ times and the output of the measurements is ordered to generate the desired random number in binary representation as depicted in Figure 7 a). This approach might be better in terms of accuracy in NISQ devices where we can have one qubit with better fidelity and decay times than the rest and therefore the errors will be smaller by just using this 1-qubit device. This is the reason why we will implement this approach in Qiskit, jut using a circuit with 1 qubit. It also has the advantage that it can be simulated in a classical computer fairly fast. Otherwise it would be pretty hard to simulate the case with 15 qubits.

The Qiskit code we wrote is fairly simple, it is a fucntion that takes as an input the number of random bits to be generated $n$ and outputs the random distribution as a numpy array. In this case, the quantum circuit described above is simulated using the 'qasm' simulator:

```
##### generate random numbers between 0 and 2^n-1#####
#need to pass n to the function
def RandomNumberGen(n):
    num = np.zeros(n)
    for i in range(n):
        qc = QuantumCircuit(1,1)
        qc.h(0) #get + state
        qc.measure(0,0) #measure in computational basis

        #Simulate and get results
        emulator = Aer.get_backend('qasm_simulator')
        job = execute( qc, emulator, shots = 1 )
        hist = job.result().get_counts()
        if hist.get('0') == None:
            num[i] = 1
        if hist.get('1') == None:
            num[i] = 0
    return num

```

In order to test our random number generator put it under two standard tests: the fist one is analyzing if the distribution is uniformly sampled and the second one is making sure that samples are not correlated to each other. By samples, bits corresponding to the same random number are not correlated to each other.

Figure 7 b) and c) show the probability of generating random number $x$ between 0 and $2^n$ for $n = 3$ (panel c)) and for $n = 15$ (panel c)). Note this probability distributions should be flat and have a probability of $1/2^n$. The error bars correspond to $1/\sqrt{samps}$ according to

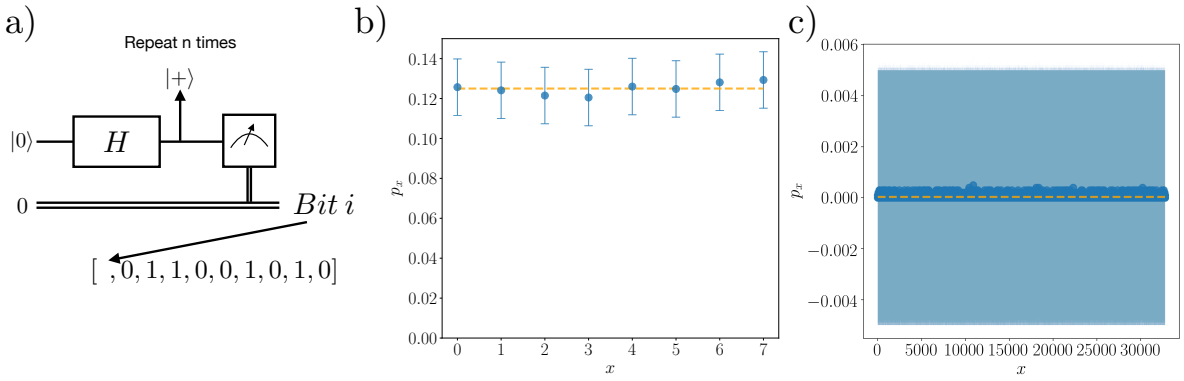



Figure 7: **a)** Schematic representation of the procedure followed to generate random numbers. **b)** Probability distribution of obtaining number $x$ with our random number generator if $n = 3$. Orange dashed line represents the expected value for the probability of obtaining each number: $1/2^n$. **c)** Probability distribution of obtaining number $x$ with our random number generator if $n = 15$. The blue cloud are the error bars. Orange dashed line represents the expected value for the probability of obtaining each number: $1/2^n$

the central limit theorem, where $samps = 10^3$ is the number of random numbers generated. Note that for $n = 15$ the error bars are much larger than the actual probability. This is due to the fact that the number of possible outputs of the random number generator is exponentially large and in order to do a good sampling of the probability distribution we would need to collect exponentially many samples, which we cannot do with our laptops. However, both for $n = 3$ and $n = 15$, the simulated values of the probabilities agree with the uniform probability distribution within the error bars.

The second check our random number generator needs to pass is that the samples are not correlated to each other. Pseudo-random number generators satisfy this condition, so this does not prove pure randomness (because it is impossible to test in a classical computer) but it is a sanity check random numbers should satisfy. To test it, within the same random number, we will compute the correlation between two bits and average over the number of samples taken:

$$C_{i,j} = \langle (x_i - x_j)^2 \rangle - \langle x_i^2 \rangle - \langle x_j^2 \rangle$$

where $x_i$ and $x_j$ are the bits in relevant positions $i$ and $j$ of random number $x$. Figure 8 shows the average correlations for $n = 3$ and $n = 15$ averaged over $samps = 10^3$ random numbers. It proves that within the error bars, the correlations are zero and therefore our random number generator does not generate correlated bits. **All the code to generate the correlation maps and to collect the statistics is attached as a Jupyter Notebook.**

## Q14

Here we prove the optimal quantum strategy to the CHSH game. We assume the following rules:

1. Alice and Bob agree on strategy. Thereafter, they can perform local operations, but they cannot communicate classically.

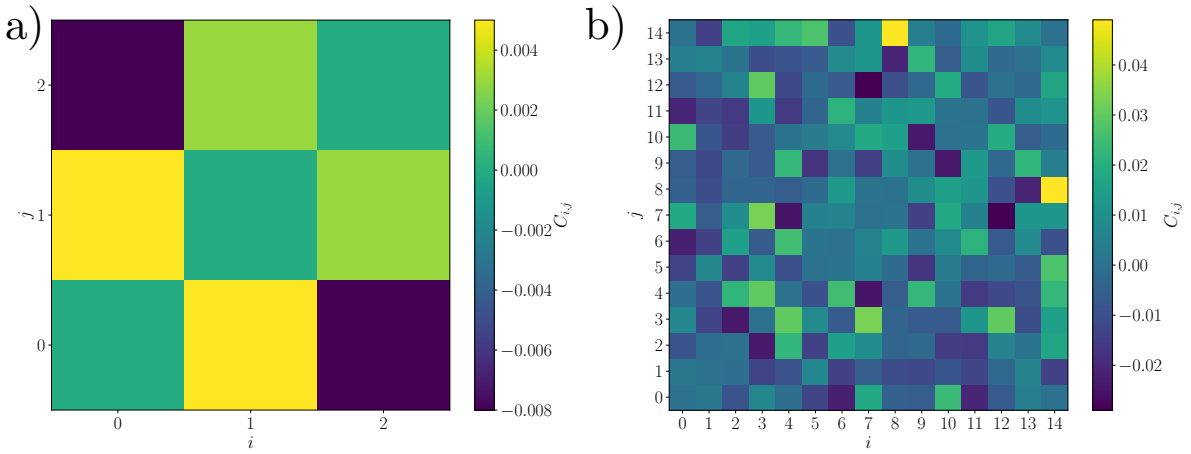



Figure 8: **a)** $n = 3$ **b)** $n = 15$. Correlation between bits $i$ and $j$ of a given random number, averaged over all the random numbers generated $samps = 10^3$

2. Alice and Bob respectively receive a random bit $x$ and $y$ from the referee.

3. Alice and Bob respectively return bits $a$ and $b$ to the referee.

4. Alice and Bob win iff

$$a + b = xy \mod 2. \tag{26}$$

We first consider the classical case. In the win condition, $a + b = 1 \Leftrightarrow a \neq b$. Similarly, $xy = 1 \Leftrightarrow x = y = 1$. If all bits are chosen at random, the former condition occurs with probability 1/2, while the latter occurs with probability 1/4. The following truth table then enumerates all possible outcomes and their probabilities:

| $a \neq b$ | $x = y = 1$ | Win | Prob. |
|---|---|---|---|
| T | T | T | 1/8 |
| T | F | F | 3/8 |
| F | T | F | 1/8 |
| F | F | T | 3/8 |

(27)

Therefore, if Alice and Bob choose $a$ and $b$ uniformly at random, they win with probability 1/2. The optimal classical strategy is for Alice and Bob to agree beforehand to always choose $a = b$ (either both 0 or both 1, always). This eliminates the first two rows of the table (where they are likely to lose), and they win with probability 3/4. We show that this is optimal by considering Alice's point of view (Bob's is symmetric) in the absence of any strategy. Note that the value of $x$ contains no information as to how she should choose $a$: if $x = 1$, the probabilities in each row are unaffected (since she doesn't know $y$), while if $x = 0$, only the second and fourth rows remain, where they still win and lose with equal probability. Therefore, Alice needs knowledge of $b$ so she can choose the fourth row in the case $x = 0$. If she has *perfect* knowledge of $b$, they will always win in this case (by choosing $a = b$), while

they can still only win 1/2 of the cases where $x = 1$. Since $x = 0$ (always win) and $x = 1$ (win 1/2) are equally likely, they win with probability 3/4, and any uncertainty in $b$ will only reduce their win percentage in the $x = 0$ case.

The optimal quantum strategy is obtained as follows. While Alice and Bob are strategizing before the game begins, they share a Bell state

$$\left|\Psi^+\right\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}. \tag{28}$$

They also define the orthonormal single-qubit measurement basis

$$\begin{aligned} |\mu_0(\theta)\rangle &= \cos\theta\,|0\rangle + \sin\theta\,|1\rangle \\ |\mu_1(\theta)\rangle &= \sin\theta\,|0\rangle - \cos\theta\,|1\rangle \end{aligned} \tag{29}$$

and agree to choose their angles based on the respective values of the referee bits $x$ and $y$:

$$\begin{aligned} \text{Alice:} &\quad \begin{cases} x = 0 \implies \theta = \alpha_0 \\ x = 1 \implies \theta = \alpha_1 \end{cases} \\ \text{Bob:} &\quad \begin{cases} y = 0 \implies \theta = \beta_0 \\ y = 1 \implies \theta = \beta_1. \end{cases} \end{aligned} \tag{30}$$

Finally, they agree to choose their respective answer bits $a$ and $b$ by measuring the Bell state (28) in the measurement basis (29):

$$\begin{aligned} \text{Alice:} &\quad \begin{cases} \mu_0 \implies a = 0 \\ \mu_1 \implies a = 1 \end{cases} \\ \text{Bob:} &\quad \begin{cases} \mu_0 \implies b = 0 \\ \mu_1 \implies b = 1. \end{cases} \end{aligned} \tag{31}$$

We now analyze the above strategy. First consider the particular case $(x, y) = (0, 0)$. According to the angle selection rule (30), Alice uses the measurement basis $\{|\mu_0(\alpha_0)\rangle, |\mu_1(\alpha_0)\rangle\}$ while Bob uses $\{|\mu_0(\beta_0)\rangle, |\mu_1(\beta_0)\rangle\}$. According to the win condition (26), they win if $(a, b) \in \{(0, 0), (1, 1)\}$. According to the answer-bit selection rule (31), these correspond to measurement outcomes $(\mu_a, \mu_b) \in \{(\mu_0, \mu_0), (\mu_1, \mu_1)\}$, which occur with probability,

$$\Pr(\text{win}|x, y = 0, 0) = \left|\left\langle \mu_0(\alpha_0)\mu_0(\beta_0)\middle|\Psi^+\right\rangle\right|^2 + \left|\left\langle \mu_1(\alpha_0)\mu_1(\beta_0)\middle|\Psi^+\right\rangle\right|^2$$

In more generality, for the cases $(x, y) \in \{(0, 0), (0, 1), (1, 0)\}$ corresponding to rows of the truth table (27) where $x = y = 1$ is False (denote this condition $C$), they win if $a = b$, which occurs with probability

$$\Pr(\text{win}|C = \text{False}) = \sum_{a=b} \left|\left\langle \mu_a(\alpha_x)\mu_b(\beta_y)\middle|\Psi^+\right\rangle\right|^2.$$

In the remaining case, $(x, y) = (1, 1)$ (i.e. $C$ is True), they win if $a \neq b$, which occurs with probability

$$\Pr(\text{win}|C = \text{True}) = \sum_{a\neq b} \left|\left\langle \mu_a(\alpha_1)\mu_b(\beta_1)\middle|\Psi^+\right\rangle\right|^2.$$

Using (28) and (29) we compute these probabilities explicitly as

$$\begin{aligned}
\Pr(\text{win}|x, y = 0, 0) &= \cos^2(\alpha_0 - \beta_0) \\
\Pr(\text{win}|x, y = 0, 1) &= \cos^2(\alpha_0 - \beta_1) \\
\Pr(\text{win}|x, y = 1, 0) &= \cos^2(\alpha_1 - \beta_0) \\
\Pr(\text{win}|x, y = 1, 1) &= \sin^2(\alpha_1 - \beta_1)
\end{aligned}$$

Since $x$ and $y$ are chosen uniformly at random, the total win probability is

$$\Pr(\text{win}) = \frac{1}{4}\left(\cos^2(\alpha_0 - \beta_0) + \cos^2(\alpha_0 - \beta_1) + \cos^2(\alpha_1 - \beta_0) + \sin^2(\alpha_1 - \beta_1)\right).$$

Extremizing with respect to each angle yields the conditions,

$$\begin{aligned}
\frac{\partial \Pr}{\partial \alpha_0} = 0 &\implies \sin 2(\alpha_0 - \beta_0) + \sin 2(\alpha_0 - \beta_1) = 0 \\
\frac{\partial \Pr}{\partial \alpha_1} = 0 &\implies \sin 2(\alpha_1 - \beta_0) - \sin 2(\alpha_1 - \beta_1) = 0 \\
\frac{\partial \Pr}{\partial \beta_0} = 0 &\implies \sin 2(\alpha_0 - \beta_0) + \sin 2(\alpha_1 - \beta_0) = 0 \\
\frac{\partial \Pr}{\partial \beta_1} = 0 &\implies \sin 2(\alpha_0 - \beta_1) - \sin 2(\alpha_1 - \beta_1) = 0,
\end{aligned}$$

which are satisfied for $\alpha_0 = 0$, $\alpha_1 = \pi/4$, $\beta_0 = \pi/8$, and $\beta_1 = -\pi/8$. These choices for the angles then yield the optimal strategy, and the win probability reduces to

$$\Pr(\text{win}) = \cos^2\left(\frac{\pi}{8}\right) \approx 85\%.$$

# References


[1] Scott Aaronson. Introduction to quantum information science, UT Austin. URL https://www.scottaaronson.com/qclec/6.pdf. URL https://www.scottaaronson.com/qclec/11.pdf.

[2] Avshalom C. Elitzur and Lev Vaidman. Quantum mechanical interaction-free measurements. *Foundations of Physics*, 23(7):987–997, Jul 1993. ISSN 1572-9516. doi: 10.1007/bf00736012. URL http://dx.doi.org/10.1007/BF00736012.

[3] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature? *Phys. Rev. Lett.*, 80:5239–5242, Jun 1998. doi: 10.1103/PhysRevLett.80.5239. URL https://link.aps.org/doi/10.1103/PhysRevLett.80.5239.

[4] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009. doi: 10.1103/RevModPhys.81.865. URL https://link.aps.org/doi/10.1103/RevModPhys.81.865.

[5] Ali Javadi. Hidden shift iPython notebook and Qiskit code. URL https://github.com/ajavadia/qiskit-terra/blob/Demo/demo/Hidden_Shift_Bent.ipynb. URL https://github.com/ajavadia/qiskit-terra/blob/Demo/demo/python/hidden_shift.py.

[6] B. Misra and E. C. G. Sudarshan. The zeno's paradox in quantum theory. *Journal of Mathematical Physics*, 18(4):756–763, 1977. doi: 10.1063/1.523304. URL https://doi.org/10.1063/1.523304.

[7] James S. Plank. University of Tennessee CS302 lecture notes. URL http://web.eecs.utk.edu/~jplank/plank/classes/cs302/Notes/NP/.

[8] Eleanor G. Rieffel and Wolfgang H. Polak. *Quantum computing: A gentle introduction*. MIT Press, 2014.

[9] J. J. Sakurai and Jim Napolitano. *Modern quantum mechanics*. Addison-Wesley, second edition, 1994. ISBN 9780805382914.

[10] Javad Shabani. Getting started with Python and Qiskit. URL https://wp.nyu.edu/shabanilab/courses/qiskit/.

[11] G. Vidal and J. I. Cirac. Irreversibility in asymptotic manipulations of entanglement. *Phys. Rev. Lett.*, 86:5803–5806, Jun 2001. doi: 10.1103/PhysRevLett.86.5803. URL https://link.aps.org/doi/10.1103/PhysRevLett.86.5803.

[12] Karl Gerd H. Vollbrecht, Reinhard F. Werner, and Michael M. Wolf. Irreversibility of entanglement distillation for a class of symmetric states. *Phys. Rev. A*, 69:062304, Jun 2004. doi: 10.1103/PhysRevA.69.062304. URL https://link.aps.org/doi/10.1103/PhysRevA.69.062304.