Quantum circuits

Download the slides for this lesson.

Open the YouTube video for this lesson in a separate window.

Introduction

This lesson introduces the *quantum circuit* model of computation, which provides a standard way to describe quantum computations. It also introduces a few important mathematical concepts, including *inner products* between vectors, the notions of *orthogonality* and *orthonormality*, and

projections and *projective measurements*, which generalize standard basis measurements. Through these concepts, we'll derive fundamental limitations on quantum information, including the *no-cloning theorem* and the impossibility to perfectly discriminate non-orthogonal quantum states.

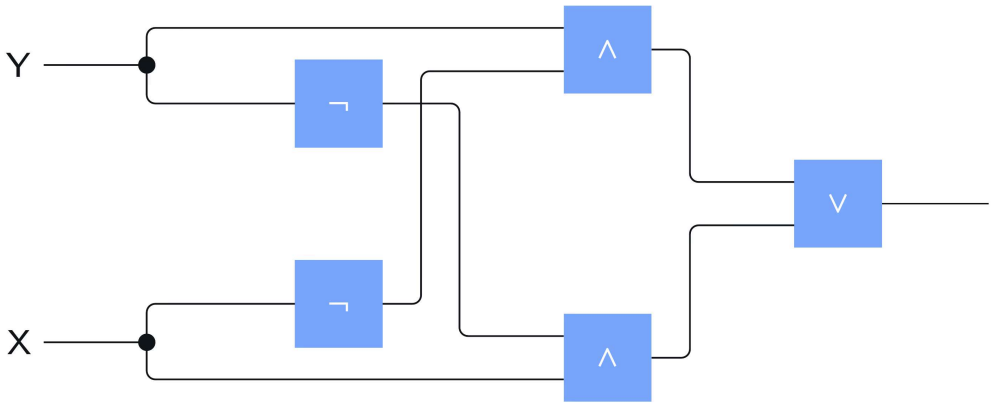
Circuits

In computer science, *circuits* are models of computation in which information is carried by wires through a network of *gates*, which represent operations on the information carried by the wires. *Quantum circuits* are a specific model of computation based on this more general concept.

Although the word "circuit" often refers to a circular path, circular paths aren't actually allowed in the circuit models of computation that are most commonly studied. That is to say, we usually consider *acyclic circuits* when we're thinking about circuits as computational models. Quantum circuits follow this pattern; a quantum circuit represents a finite sequence of operations that cannot contain feedback loops.

Boolean circuits

Here is an example of a (classical) Boolean circuit, where the wires carry binary values and the gates represent Boolean logic operations:



The flow of information along the wires goes from left to right: the wires on the left-hand side of the figure labeled X and Y are input bits, which can each be set to whatever binary value we choose, and the wire on the right-hand side is the output. The intermediate wires take whatever values are determined by the gates, which are evaluated from left to right.

The gates are AND gates (labeled $\wedge$), OR gates (labeled $\vee$), and NOT gates (labeled $\neg$). The functions computed by these gates will likely be familiar to many readers, but here they are represented by tables of values:

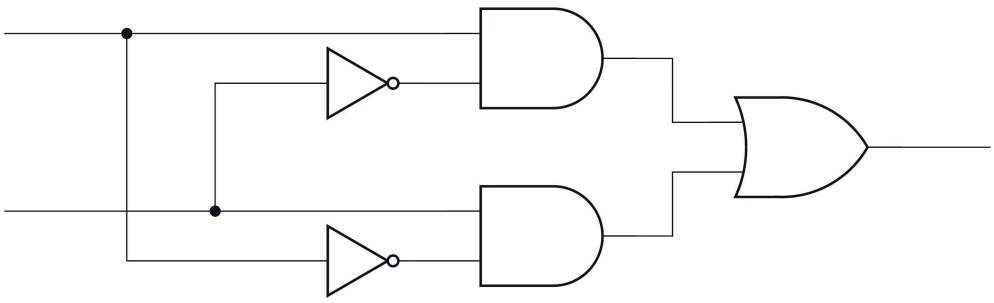
a	$\neg a$
0	1
1	0

ab	$a \wedge b$
00	0
01	0
10	0
11	1

ab	$a \vee b$
00	0
01	1
10	1
11	1

The two small, solid circles on the wires just to the right of the names X and Y represent *fanout* operations, which simply create a copy of whatever value is carried on the wire on which they appear, allowing this value to be input into multiple gates. Fanout operations are not always considered to be gates in the classical setting; sometimes they're treated as if they're "free" in some sense. When Boolean circuits are converted into equivalent quantum circuits, however, we do need to classify fanout operations explicitly as gates to handle and account for them correctly.

Here's the same circuit illustrated in a style more common in electrical engineering, which uses conventional symbols for the AND, OR, and NOT gates:

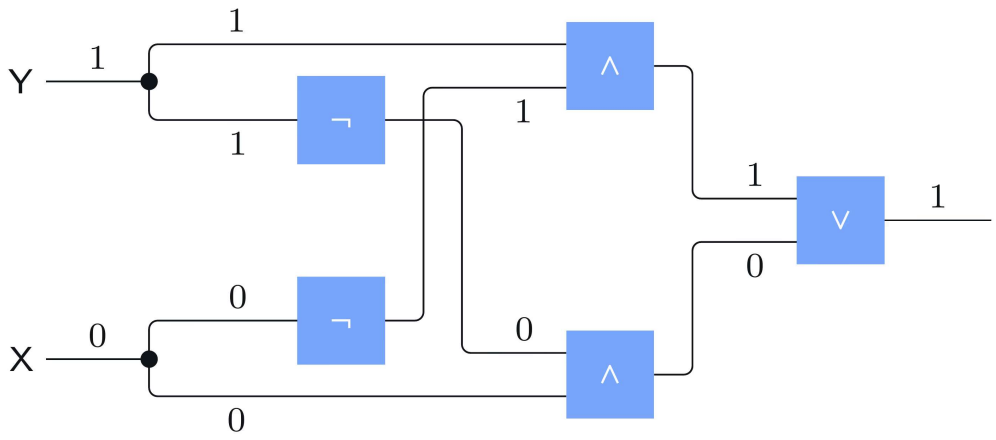


We won't use this style or these particular gate symbols further, but we will use different symbols to represent gates in quantum circuits, which we'll explain as we encounter them.

The particular circuit in this example computes the *exclusive-OR* (or XOR for short), which is denoted by the symbol $\oplus$:

ab	$a \oplus b$
00	0
01	1
10	1
11	0

In the next diagram we consider just one choice for the inputs: $\mathsf{X} = 0$ and $\mathsf{Y} = 1$. Each wire is labeled by value it carries so you can follow the operations. The output value is 1 in this case, which is the correct value for the XOR: $0 \oplus 1 = 1$.



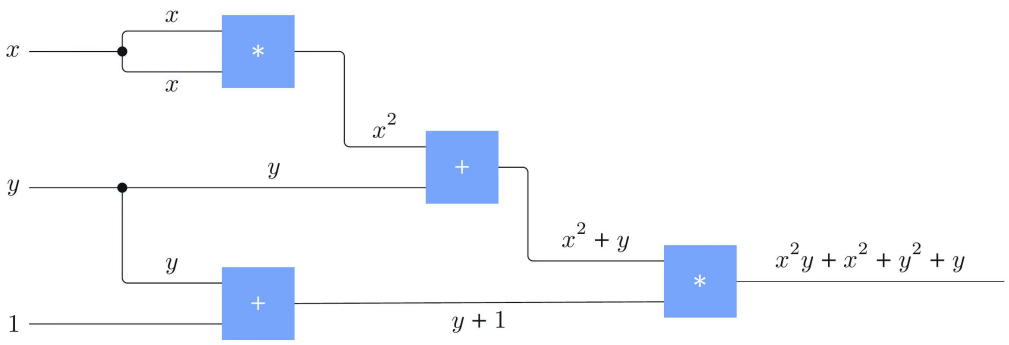
The other three possible input settings can be checked in a similar way.

Other types of circuits

As was suggested above, the notion of a circuit in computer science is very general. For example, circuits whose wires carry values other than 0 and 1 are sometimes analyzed, as are gates representing different choices of operations.

In *arithmetic circuits*, for instance, the wires may carry integer values while the gates represent arithmetic operations, such as addition and multiplication. The following figure depicts an arithmetic circuit that takes two variable input values (x and y) as well as a third input set to the value 1.

The values carried by the wires, as functions of the values x and y, are shown in the figure.

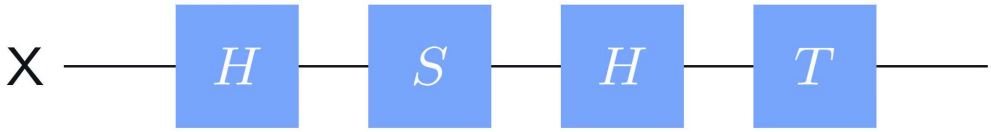


We can also consider circuits that incorporate randomness, such as ones where gates represent probabilistic operations.

Quantum circuits

In the quantum circuit model, wires represent qubits and gates represent operations on these qubits. We'll focus for now on operations we've encountered so far, namely *unitary operations* and *standard basis measurements*. As we learn about other sorts of quantum operations and measurements, we can enhance our model accordingly.

Here's a simple example of a quantum circuit:



In this circuit, we have a single qubit named X, which is represented by the horizontal line, and a sequence of gates representing unitary operations on this qubit. Just like in the examples above, the flow of information goes from left to right — so the first operation performed is a Hadamard operation, the second is an S operation, the third is another Hadamard operation, and the

final operation is a T operation. Applying the entire circuit therefore applies the composition of these operations, $THSH$, to the qubit X.

Sometimes we may wish to explicitly indicate the input or output states of circuits. For example, if we apply the operation $THSH$ to the state $|0\rangle$, we obtain the state $\frac{1+i}{2}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$. This can be indicated as follows:

Quantum circuits often start out with all qubits initialized to $|0\rangle$, as we have in this case, but there are also situations where the input qubits are initially set to different states.

Now let's see how we can specify this circuit in Qiskit. We'll start with a version check along with the imports needed for the remainder of the lesson.

```
from qiskit import __version__
print(__version__)
```

Output:

```
1.3.1
```

```
from qiskit import QuantumCircuit, QuantumRegister,
from qiskit.quantum_info import Operator
from qiskit_aer import AerSimulator
from qiskit.visualization import plot_histogram
```

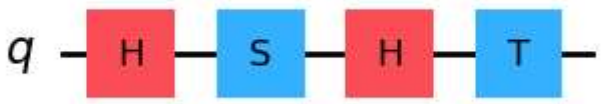
No output produced

We've seen some of these imports in the two previous lessons, but others are new. For now, let's just highlight that we will be using the Aer simulator to simulate quantum circuits.

To begin, we can build the circuit above as follows, by defining a quantum circuit with one qubit and sequentially adding gates from left to right.

```
circuit = QuantumCircuit(1)
circuit.h(0)
circuit.s(0)
circuit.h(0)
circuit.t(0)

display(circuit.draw(output="mpl"))
```

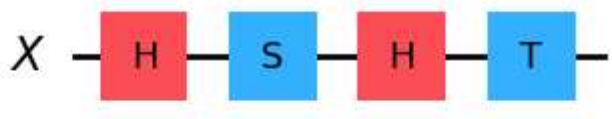
Output:



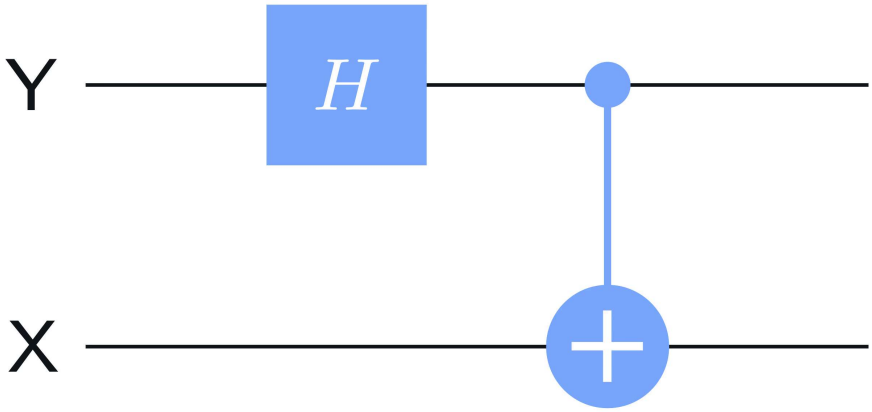
The default names for qubits in Qiskit are $\mathbf{q}_0$, $\mathbf{q}_1$, $\mathbf{q}_2$, etc., and when there's just a single qubit, like in our example, the default name is $\mathbf{q}$ rather than $\mathbf{q}_0$. If we wish to choose our own name we can do this using the `QuantumRegister` class, which allows us to name a collection of qubits as treat it as a single object. Here we're doing this with just a single qubit.

```
X = QuantumRegister(1, "X")
circuit = QuantumCircuit(X)
circuit.h(X)
circuit.s(X)
circuit.h(X)
circuit.t(X)

display(circuit.draw(output="mpl"))
```

Output:



Here's another example of a quantum circuit, this time with two qubits:



As always, the gate labeled H refers to a Hadamard operation, while the second gate is a *controlled-NOT* operation: the solid circle represents the *control qubit* and the circle resembling the symbol $\oplus$ denotes the *target qubit.*

Before examining this circuit in greater detail and explaining what it does, it is imperative that we first clarify how qubits are ordered in quantum circuits. This connects with the convention that Qiskit uses for naming and ordering systems that was mentioned briefly in the previous lesson.

> **Qiskit's qubit ordering convention for circuits**
>
> In Qiskit, the *topmost* qubit in a circuit diagram has index 0 and corresponds to the *rightmost* position in a tuple of qubits (or in a string, Cartesian product, or tensor product corresponding to this tuple). The

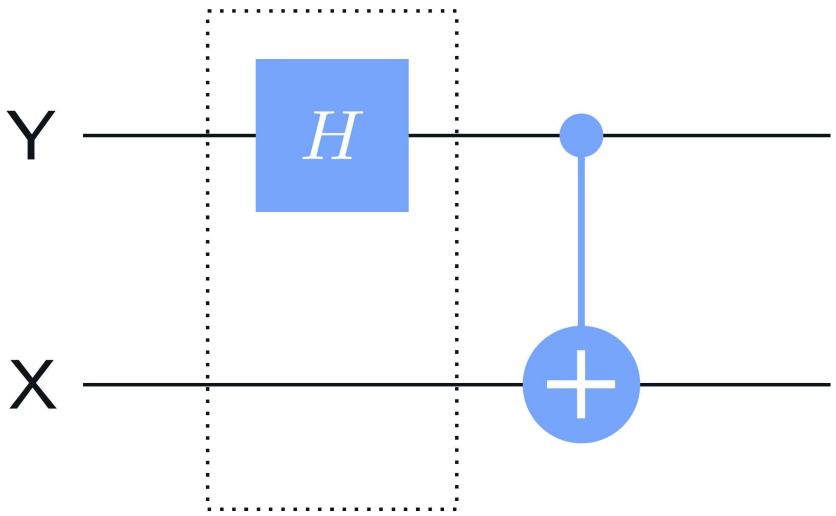
> *second-from-top* qubit has index 1, and corresponds to the position *second-from-right* in a tuple, and so on, down to the *bottommost* qubit, which has the highest index, and corresponds to the *leftmost* position in a tuple.
>
> In particular, Qiskit's default names for the qubits in an n-qubit circuit are represented by the n-tuple $(\mathsf{q}_{n-1}, \ldots, \mathsf{q}_0)$, with q_0 being the qubit on the top and q_{n-1} on the bottom in quantum circuit diagrams.

Please be aware that this is a reversal of a more common convention for ordering qubits in circuits, and is a frequent source of confusion. Further information on this ordering convention can be found on the Bit-ordering in Qiskit documentation page.

Although we sometimes deviate from the specific default names $\mathsf{q}_0, \ldots, \mathsf{q}_{n-1}$ used for qubits by Qiskit, we will always follow the ordering convention described above when interpreting circuit diagrams throughout this course. Thus, our interpretation of the circuit above is that it describes an operation on a pair of qubits (X, Y). If the input to the circuit is a quantum state $|\psi\rangle \otimes |\phi\rangle$, for instance, then this means that the lower qubit X starts in the state $|\psi\rangle$ and the upper qubit Y starts in the state $|\phi\rangle$.

To understand what the circuit does, we can go from left to right through its operations.

1. The first operation is a Hadamard operation on Y:

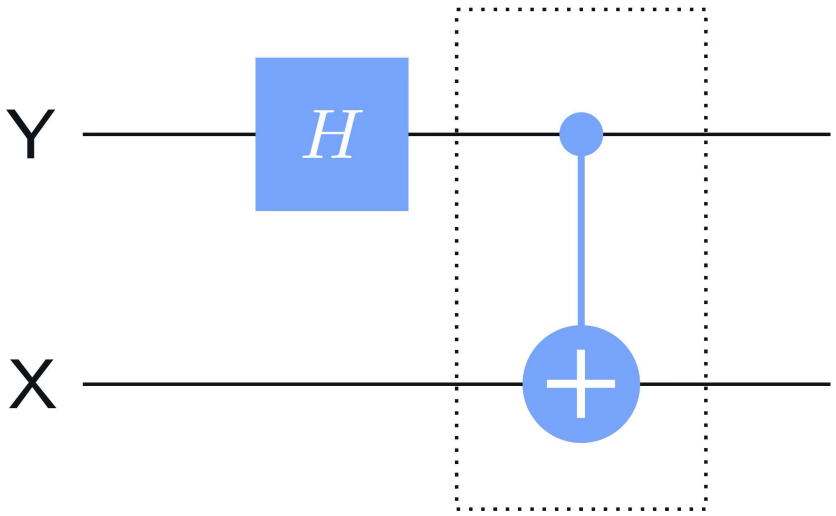


When applying a gate to a single qubit like this, nothing happens to the other qubits (which is just one other qubit in this case). Nothing happening is equivalent to the identity operation being performed. The dotted rectangle in the figure above therefore represents this operation:

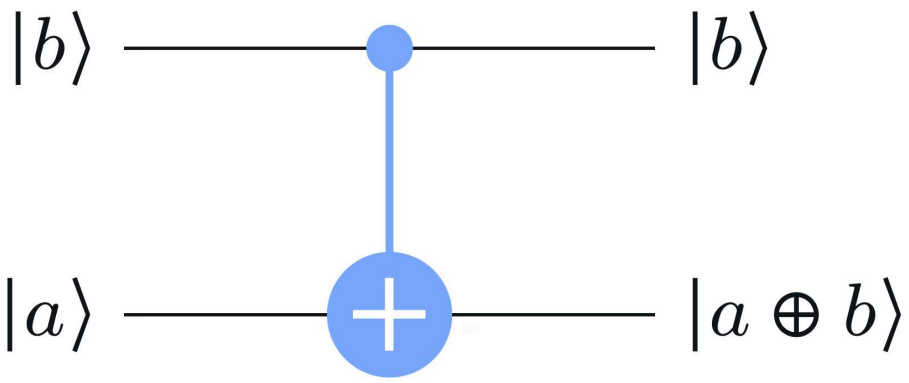
$$\mathbb{I} \otimes H = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

Note that the identity matrix is on the left of the tensor product and H is on the right, which is consistent with Qiskit's ordering convention.

2. The second operation is the controlled-NOT operation, where Y is the control and X is the target:



The controlled-NOT gate's action on standard basis states is as follows:



Given that we order the qubits as (X, Y), with X being on the bottom and Y being on the top of our circuit, the matrix representation of the controlled-NOT gate is this:

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

The unitary operation implemented by the entire circuit, which we'll give the name U, is the composition of the operations:

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix}$$

In particular, recalling our notation for the Bell states,

$$
\begin{aligned}
|\phi^+\rangle &= \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle \\
|\phi^-\rangle &= \frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle \\
|\psi^+\rangle &= \frac{1}{\sqrt{2}}|01\rangle + \frac{1}{\sqrt{2}}|10\rangle \\
|\psi^-\rangle &= \frac{1}{\sqrt{2}}|01\rangle - \frac{1}{\sqrt{2}}|10\rangle,
\end{aligned}
$$

we find that

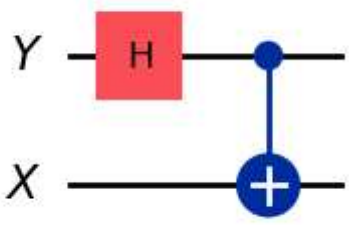
$$
\begin{aligned}
U|00\rangle &= |\phi^+\rangle \\
U|01\rangle &= |\phi^-\rangle \\
U|10\rangle &= |\psi^+\rangle \\
U|11\rangle &= -|\psi^-\rangle.
\end{aligned}
$$

This circuit therefore gives us a way to create the state $|\phi^+\rangle$ if we run it on two qubits initialized to $|00\rangle$. More generally, it provides us with a way to convert the standard basis to the Bell basis. (Note that, while it is not important for this example, the -1 phase factor on the last state, $-|\psi^-\rangle$, could be eliminated if we wanted by making a small addition to the circuit. For instance, we could add a controlled-Z gate at the beginning, which is similar to a controlled-NOT gate except that a Z operation is applied to the target qubit rather than a NOT operation when the control is set to 1. Alternatively, we could add a swap gate at the end. Either choice eliminates the minus sign without affecting the circuit's action on the other three standard basis states.)

Let's create the circuit in Qiskit and check that our calculations are correct.

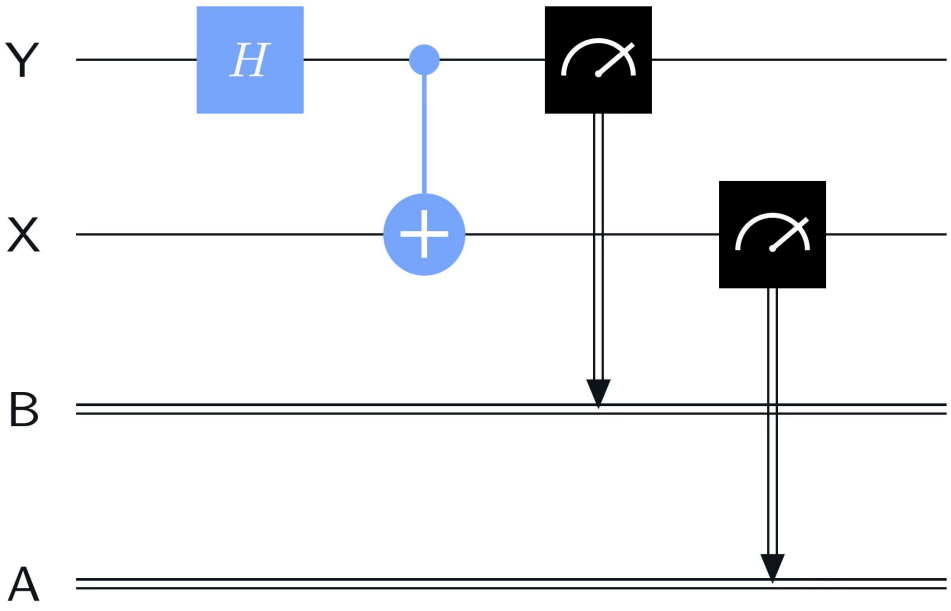
```
X = QuantumRegister(1, "X")
Y = QuantumRegister(1, "Y")
circuit = QuantumCircuit(Y,X)
circuit.h(Y)
circuit.cx(Y, X)

display(circuit.draw(output="mpl"))
display(Operator.from_circuit(circuit).draw("latex"))
```

Output:



$$\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 & 0 \end{bmatrix}$$

In general, quantum circuits can contain any number of qubit wires. We may also include *classical bit* wires, which are indicated by double lines, like in this example:

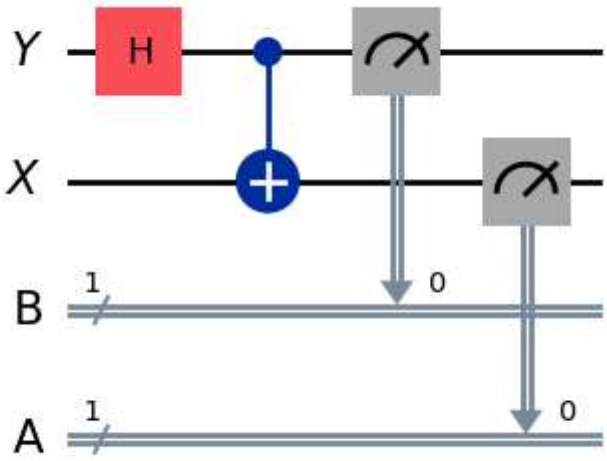


Here we have a Hadamard gate and a controlled-NOT gate on two qubits X and Y, just like in the previous example. We also have two *classical* bits, A and B, as well as two measurement gates. The measurement gates represent standard basis measurements: the qubits are changed into their post-measurement states, while the measurement outcomes are *overwritten* onto the classical bits to which the arrows point.

Here's an implementation of this circuit using Qiskit:

```
X = QuantumRegister(1, "X")
Y = QuantumRegister(1, "Y")
A = ClassicalRegister(1, "A")
B = ClassicalRegister(1, "B")

circuit = QuantumCircuit(Y, X, B, A)
circuit.h(Y)
circuit.cx(Y, X)
circuit.measure(Y, B)
circuit.measure(X, A)

display(circuit.draw(output="mpl"))
```

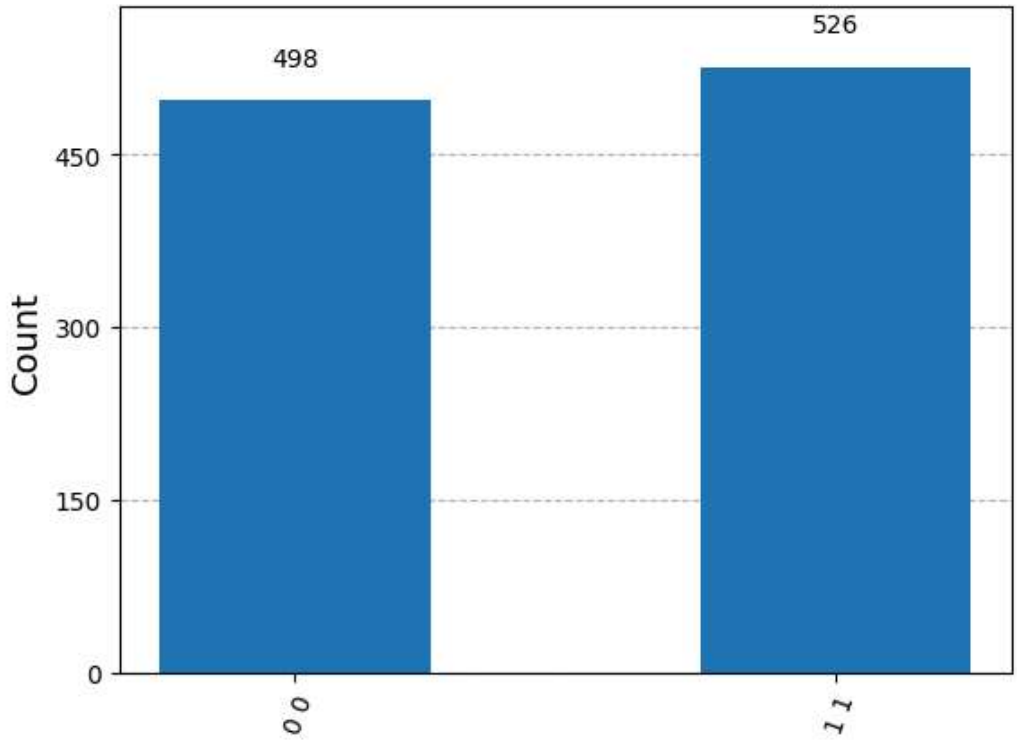
Output:



The circuit can be simulated using the Aer simulator like this:

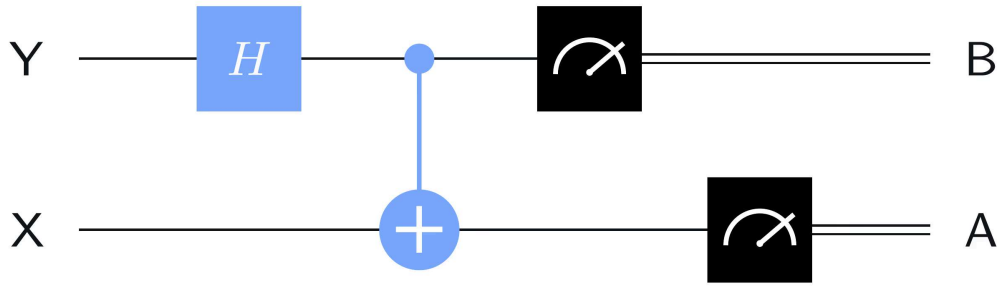
```
result = AerSimulator().run(circuit).result()
statistics = result.get_counts()
display(plot_histogram(statistics))
```

Output:



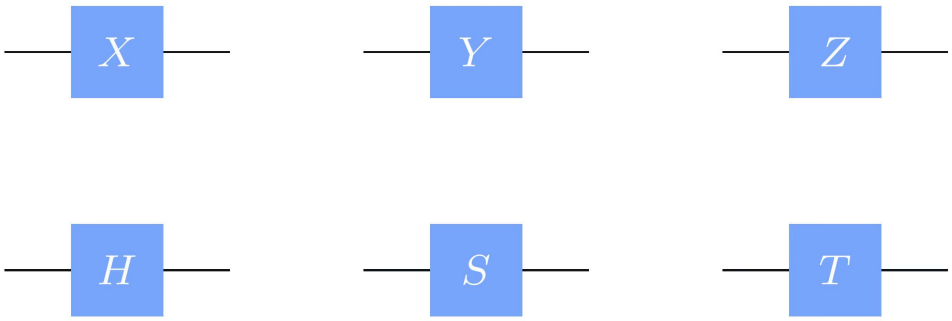
It's often convenient to depict a measurement as a gate that takes a qubit as input and outputs a classical bit (as opposed to outputting the qubit in its post-measurement state and writing the result to a separate classical bit). This means the measured qubit has been discarded and can safely be ignored thereafter, its state having changed into $|0\rangle$ or $|1\rangle$ depending upon the measurement outcome.

For example, the following circuit diagram represents the same process as the one in the previous diagram, but where we disregard X and Y after measuring them:

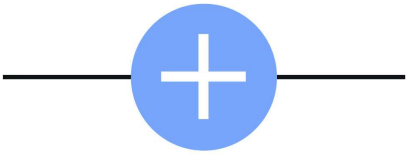


As the course continues, we'll see more examples of quantum circuits, which are usually more complicated than the simple examples above. Here are some examples of symbols used to denote gates that commonly appear in circuit diagrams:

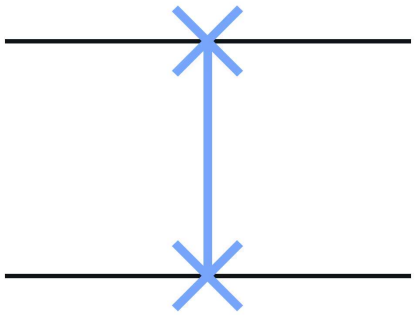
– Single-qubit gates are generally shown as squares with a letter indicating which operation it is, like this:



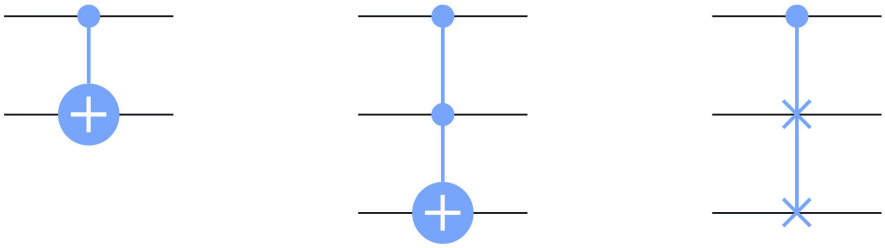
Not gates (or, equivlently, X gates) are also sometimes denoted by a circle around a plus sign:



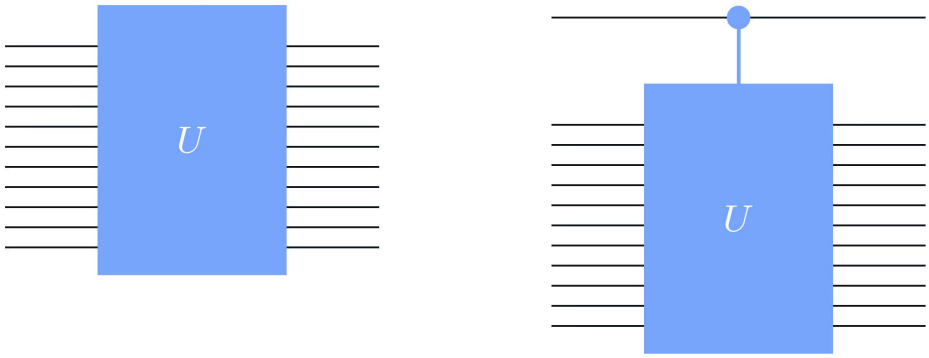
– Swap gates are denoted as follows:



– Controlled-gates, meaning gates that describe controlled-unitary operations, are denoted by a filled-in circle (indicating the control) connected by a vertical line to whatever operation is being controlled. For instance, controlled-NOT gates, controlled-controlled-NOT (or Toffoli) gates, and controlled-swap (Fredkin) gates are denoted like this:



– Arbitrary unitary operations on multiple qubits may be viewed as gates. They are depicted by rectangles labeled by the name of the unitary operation. For instance, here is a depiction of an (unspecified) unitary operation U as a gate, along with a controlled version of this gate:



Inner products, orthonormality, and projections

To better prepare ourselves to explore the capabilities and limitations of quantum circuits, we now introduce some additional mathematical concepts — namely the *inner product* between vectors (and its connection to the Euclidean norm), the notions of *orthogonality* and *orthonormality* for sets of vectors, and *projection* matrices, which will allow us to introduce a handy generalization of standard basis measurements.

Inner products

Recall from the Single systems lesson that, when we use the Dirac notation to refer to an arbitrary column vector as a ket, such as

$$|\psi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix},$$

the corresponding bra vector is the *conjugate transpose* of this vector:

$$\langle\psi| = \big(|\psi\rangle\big)^\dagger = \begin{pmatrix} \overline{\alpha_1} & \overline{\alpha_2} & \cdots & \overline{\alpha_n} \end{pmatrix}. \tag{1}$$

Alternatively, if we have some classical state set Σ in mind, and we express a column vector as a ket, such as

$$|\psi\rangle = \sum_{a\in\Sigma} \alpha_a |a\rangle,$$

then the corresponding row (or bra) vector is the conjugate transpose

$$\langle\psi| = \sum_{a\in\Sigma} \overline{\alpha_a} \langle a|. \tag{2}$$

We also observed that the product of a bra vector and a ket vector, viewed as matrices either having a single row or a single column, results in a scalar. Specifically, if we have two (column) vectors

$$|\psi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \quad \text{and} \quad |\phi\rangle = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix},$$

so that the row vector $\langle\psi|$ is as in equation (1), then

$$\langle\psi|\phi\rangle = \langle\psi||\phi\rangle = \begin{pmatrix} \overline{\alpha_1} & \overline{\alpha_2} & \cdots & \overline{\alpha_n} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \overline{\alpha_1}\beta_1 + \cdots + \overline{\alpha_n}\beta_n.$$

Alternatively, if we have two column vectors that we have written as

$$|\psi\rangle = \sum_{a\in\Sigma} \alpha_a |a\rangle \quad \text{and} \quad |\phi\rangle = \sum_{b\in\Sigma} \beta_b |b\rangle,$$

so that $\langle\psi|$ is the row vector (2), we find that

$$\begin{aligned} \langle\psi|\phi\rangle &= \langle\psi||\phi\rangle \\ &= \left(\sum_{a\in\Sigma} \overline{\alpha_a}\langle a|\right)\left(\sum_{b\in\Sigma} \beta_b |b\rangle\right) \\ &= \sum_{a\in\Sigma}\sum_{b\in\Sigma} \overline{\alpha_a}\beta_b \langle a|b\rangle \\ &= \sum_{a\in\Sigma} \overline{\alpha_a}\beta_a, \end{aligned}$$

where the last equality follows from the observation that $\langle a|a\rangle = 1$ and $\langle a|b\rangle = 0$ for classical states a and b satisfying $a \neq b$.

The value $\langle\psi|\phi\rangle$ is called the *inner product* between the vectors $|\psi\rangle$ and $|\phi\rangle$. Inner products are critically important in quantum information and computation — we would not get far in understanding quantum information at a mathematical level without them.

Let us now collect together some basic facts about inner products of vectors.

1. **Relationship to the Euclidean norm.** The inner product of any vector

$$|\psi\rangle = \sum_{a\in\Sigma} \alpha_a |a\rangle$$

with itself is

$$\langle\psi|\psi\rangle = \sum_{a\in\Sigma} \overline{\alpha_a}\alpha_a = \sum_{a\in\Sigma} |\alpha_a|^2 = \bigl\||\psi\rangle\bigr\|^2.$$

Thus, the Euclidean norm of a vector may alternatively be expressed as

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

Notice that the Euclidean norm of a vector must always be a nonnegative real number. Moreover, the only way the Euclidean norm of a vector can be equal to zero is if every one of the entries is equal to zero, which is to say that the vector is the zero vector.
We can summarize these observations like this: for every vector $|\psi\rangle$ we have

$$\langle\psi|\psi\rangle \geq 0,$$

with $\langle\psi|\psi\rangle = 0$ if and only if $|\psi\rangle = 0$. This property of the inner product is sometimes referred to as *positive definiteness*.

2. **Conjugate symmetry.** For any two vectors

$$|\psi\rangle = \sum_{a\in\Sigma} \alpha_a |a\rangle \quad \text{and} \quad |\phi\rangle = \sum_{b\in\Sigma} \beta_b |b\rangle,$$

we have

$$\langle\psi|\phi\rangle = \sum_{a\in\Sigma} \overline{\alpha_a}\beta_a \quad \text{and} \quad \langle\phi|\psi\rangle = \sum_{a\in\Sigma} \overline{\beta_a}\alpha_a,$$

and therefore

$$\overline{\langle \psi | \phi \rangle} = \langle \phi | \psi \rangle.$$

3. **Linearity in the second argument (and conjugate linearity in the first).** Let us suppose that $|\psi\rangle$, $|\phi_1\rangle$, and $|\phi_2\rangle$ are vectors and α_1 and α_2 are complex numbers. If we define a new vector

   $$|\phi\rangle = \alpha_1 |\phi_1\rangle + \alpha_2 |\phi_2\rangle,$$

   then

   $$\langle \psi | \phi \rangle = \langle \psi | \bigl(\alpha_1 |\phi_1\rangle + \alpha_2 |\phi_2\rangle\bigr) = \alpha_1 \langle \psi | \phi_1 \rangle + \alpha_2 \langle \psi | \phi_2 \rangle.$$

   That is to say, the inner product is *linear* in the second argument. This can be verified either through the formulas above or simply by noting that matrix multiplication is linear in each argument (and specifically in the second argument).
   Combining this fact with conjugate symmetry reveals that the inner product is *conjugate linear* in the first argument. That is, if $|\psi_1\rangle$, $|\psi_2\rangle$, and $|\phi\rangle$ are vectors and α_1 and α_2 are complex numbers, and we define

   $$|\psi\rangle = \alpha_1 |\psi_1\rangle + \alpha_2 |\psi_2\rangle,$$

   then

   $$\langle \psi | \phi \rangle = \bigl(\overline{\alpha_1} \langle \psi_1 | + \overline{\alpha_2} \langle \psi_2 |\bigr) |\phi\rangle = \overline{\alpha_1} \langle \psi_1 | \phi \rangle + \overline{\alpha_2} \langle \psi_2 | \phi \rangle.$$

4. **The Cauchy–Schwarz inequality.** For every choice of vectors $|\phi\rangle$ and $|\psi\rangle$ having the same number of entries, we have

   $$\bigl|\langle \psi | \phi \rangle\bigr| \leq \bigl\| |\psi\rangle \bigr\| \bigl\| |\phi\rangle \bigr\|.$$

   This is an incredibly handy inequality that gets used quite extensively in quantum information (and in many other fields of study).

Orthogonal and orthonormal sets

Two vectors $|\phi\rangle$ and $|\psi\rangle$ are said to be *orthogonal* if their inner product is zero:

$$\langle \psi | \phi \rangle = 0.$$

Geometrically, we can think about orthogonal vectors as vectors at right angles to each other.

A set of vectors $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ is called an *orthogonal set* if every vector in the set is orthogonal to every other vector in the set. That is, this set is orthogonal if

$$\langle \psi_j | \psi_k \rangle = 0$$

for all choices of $j, k \in \{1, \ldots, m\}$ for which $j \neq k$.

A set of vectors $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ is called an *orthonormal* set if it is an orthogonal set and, in addition, every vector in the set is a unit vector. Alternatively, this set is an orthonormal set if we have

$$\langle \psi_j | \psi_k \rangle = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \tag{3}$$

for all choices of $j, k \in \{1, \ldots, m\}$.

Finally, a set $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ is an *orthonormal basis* if, in addition to being an orthonormal set, it forms a basis. This is equivalent to $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ being an orthonormal set and m being equal to the dimension of the space from which $|\psi_1\rangle, \ldots, |\psi_m\rangle$ are drawn.

For example, for any classical state set Σ, the set of all standard basis vectors

$$\big\{|a\rangle \,:\, a \in \Sigma\big\}$$

is an orthonormal basis. The set $\{|+\rangle, |-\rangle\}$ is an orthonormal basis for the 2-dimensional space corresponding to a single qubit, and the Bell basis $\{|\phi^+\rangle, |\phi^-\rangle, |\psi^+\rangle, |\psi^-\rangle\}$ is an orthonormal basis for the 4-dimensional space corresponding to two qubits.

Extending orthonormal sets to orthonormal bases

Suppose that $|\psi_1\rangle, \ldots, |\psi_m\rangle$ are vectors that live in an n-dimensional space, and assume moreover that $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ is an orthonormal set. Orthonormal sets are always linearly independent sets, so these vectors necessarily span a subspace of dimension m. From this we conclude that $m \leq n$ because the dimension of the subspace spanned by these vectors

cannot be larger than the dimension of the entire space from which they're drawn.

If it is the case that $m < n$, then it is always possible to choose an additional $n - m$ vectors $|\psi_{m+1}\rangle, \ldots, |\psi_n\rangle$ so that $\{|\psi_1\rangle, \ldots, |\psi_n\rangle\}$ forms an orthonormal basis. A procedure known as the *Gram–Schmidt orthogonalization process* can be used to construct these vectors.

Orthonormal sets and unitary matrices

Orthonormal sets of vectors are closely connected with unitary matrices. One way to express this connection is to say that the following three statements are logically equivalent (meaning that they are all true or all false) for any choice of a square matrix U:

1. The matrix U is unitary (i.e., $U^\dagger U = \mathbb{I} = UU^\dagger$).
2. The rows of U form an orthonormal set.
3. The columns of U form an orthonormal set.

This equivalence is actually pretty straightforward when we think about how matrix multiplication and the conjugate transpose work. Suppose, for instance, that we have a 3×3 matrix like this:

$$U = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,1} & \alpha_{3,2} & \alpha_{3,3} \end{pmatrix}$$

The conjugate transpose of U looks like this:

$$U^\dagger = \begin{pmatrix} \overline{\alpha_{1,1}} & \overline{\alpha_{2,1}} & \overline{\alpha_{3,1}} \\ \overline{\alpha_{1,2}} & \overline{\alpha_{2,2}} & \overline{\alpha_{3,2}} \\ \overline{\alpha_{1,3}} & \overline{\alpha_{2,3}} & \overline{\alpha_{3,3}} \end{pmatrix}$$

Multiplying the two matrices, with the conjugate transpose on the left-hand side, gives us this matrix:

$$\begin{pmatrix} \overline{\alpha_{1,1}} & \overline{\alpha_{2,1}} & \overline{\alpha_{3,1}} \\ \overline{\alpha_{1,2}} & \overline{\alpha_{2,2}} & \overline{\alpha_{3,2}} \\ \overline{\alpha_{1,3}} & \overline{\alpha_{2,3}} & \overline{\alpha_{3,3}} \end{pmatrix} \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,1} & \alpha_{3,2} & \alpha_{3,3} \end{pmatrix}$$
$$= \begin{pmatrix} \overline{\alpha_{1,1}}\alpha_{1,1} + \overline{\alpha_{2,1}}\alpha_{2,1} + \overline{\alpha_{3,1}}\alpha_{3,1} & \overline{\alpha_{1,1}}\alpha_{1,2} + \overline{\alpha_{2,1}}\alpha_{2,2} + \overline{\alpha_{3,1}}\alpha_{3,2} \\ \overline{\alpha_{1,2}}\alpha_{1,1} + \overline{\alpha_{2,2}}\alpha_{2,1} + \overline{\alpha_{3,2}}\alpha_{3,1} & \overline{\alpha_{1,2}}\alpha_{1,2} + \overline{\alpha_{2,2}}\alpha_{2,2} + \overline{\alpha_{3,2}}\alpha_{3,2} \\ \overline{\alpha_{1,3}}\alpha_{1,1} + \overline{\alpha_{2,3}}\alpha_{2,1} + \overline{\alpha_{3,3}}\alpha_{3,1} & \overline{\alpha_{1,3}}\alpha_{1,2} + \overline{\alpha_{2,3}}\alpha_{2,2} + \overline{\alpha_{3,3}}\alpha_{3,2} \end{pmatrix}$$

If we form three vectors from the columns of U,

$$|\psi_1\rangle = \begin{pmatrix} \alpha_{1,1} \\ \alpha_{2,1} \\ \alpha_{3,1} \end{pmatrix}, \quad |\psi_2\rangle = \begin{pmatrix} \alpha_{1,2} \\ \alpha_{2,2} \\ \alpha_{3,2} \end{pmatrix}, \quad |\psi_3\rangle = \begin{pmatrix} \alpha_{1,3} \\ \alpha_{2,3} \\ \alpha_{3,3} \end{pmatrix},$$

then we can alternatively express the product above as follows:

$$U^\dagger U = \begin{pmatrix} \langle\psi_1|\psi_1\rangle & \langle\psi_1|\psi_2\rangle & \langle\psi_1|\psi_3\rangle \\ \langle\psi_2|\psi_1\rangle & \langle\psi_2|\psi_2\rangle & \langle\psi_2|\psi_3\rangle \\ \langle\psi_3|\psi_1\rangle & \langle\psi_3|\psi_2\rangle & \langle\psi_3|\psi_3\rangle \end{pmatrix}$$

Referring to the equation (3), we now see that the condition that this matrix is equal to the identity matrix is equivalent to the orthonormality of the set $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$. This argument generalizes to unitary matrices of any size. The fact that the rows of a matrix form an orthonormal basis if and only if the matrix is unitary then follows from the fact that a matrix is unitary if and only if its transpose is unitary.

Given the equivalence described above, together with the fact that every orthonormal set can be extended to form an orthonormal basis, we conclude the following useful fact: Given any orthonormal set of vectors $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ drawn from an n-dimensional space, there exists a unitary matrix U whose first m columns are the vectors $|\psi_1\rangle, \ldots, |\psi_m\rangle$. Pictorially, we can always find a unitary matrix having this form:

$$U = \begin{pmatrix} | & | & & | & | & & | \\ |\psi_1\rangle & |\psi_2\rangle & \cdots & |\psi_m\rangle & |\psi_{m+1}\rangle & \cdots & |\psi_n\rangle \\ | & | & & | & | & & | \end{pmatrix}.$$

Here, the last $n - m$ columns are filled in with any choice of vectors $|\psi_{m+1}\rangle, \ldots, |\psi_n\rangle$ that make $\{|\psi_1\rangle, \ldots, |\psi_n\rangle\}$ an orthonormal basis.

Projections and projective measurements

Projection matrices

A square matrix Π is called a *projection* if it satisfies two properties:

1. $\Pi = \Pi^\dagger$.
2. $\Pi^2 = \Pi$.

Matrices that satisfy the first condition — that they are equal to their own conjugate transpose — are called *Hermitian matrices*, and matrices that satisfy the second condition — that squaring them leaves them unchanged — are called *idempotent* matrices.

As a word of caution, the word *projection* is sometimes used to refer to any matrix that satisfies just the second condition but not necessarily the first, and when this is done the term *orthogonal projection* is typically used to refer to matrices satisfying both properties. In the context of quantum information and computation, however, terms *projection* and *projection matrix* more typically refer to matrices satisfying both conditions.

An example of a projection is the matrix

$$\Pi = |\psi\rangle\langle\psi| \tag{4}$$

for any unit vector $|\psi\rangle$. We can see that this matrix is Hermitian as follows:

$$\Pi^\dagger = \big(|\psi\rangle\langle\psi|\big)^\dagger = \big(\langle\psi|\big)^\dagger\big(|\psi\rangle\big)^\dagger = |\psi\rangle\langle\psi| = \Pi.$$

Here, to obtain the second equality, we have used the formula

$$(AB)^\dagger = B^\dagger A^\dagger,$$

which is always true, for any two matrices A and B for which the product AB makes sense.

To see that the matrix Π in (4) is idempotent, we can use the assumption that $|\psi\rangle$ is a unit vector, so that it satisfies $\langle\psi|\psi\rangle = 1$. Thus, we have

$$\Pi^2 = \big(|\psi\rangle\langle\psi|\big)^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \Pi.$$

More generally, if $\big\{|\psi_1\rangle, \ldots, |\psi_m\rangle\big\}$ is any orthonormal set of vectors, then the matrix

$$\Pi = \sum_{k=1}^{m} |\psi_k\rangle\langle\psi_k| \tag{5}$$

is a projection. Specifically, we have

$$\begin{aligned}
\Pi^\dagger &= \left(\sum_{k=1}^{m} |\psi_k\rangle\langle\psi_k|\right)^\dagger \\
&= \sum_{k=1}^{m} \left(|\psi_k\rangle\langle\psi_k|\right)^\dagger \\
&= \sum_{k=1}^{m} |\psi_k\rangle\langle\psi_k| \\
&= \Pi,
\end{aligned}$$

and

$$\begin{aligned}
\Pi^2 &= \left(\sum_{j=1}^{m} |\psi_j\rangle\langle\psi_j|\right)\left(\sum_{k=1}^{m} |\psi_k\rangle\langle\psi_k|\right) \\
&= \sum_{j=1}^{m}\sum_{k=1}^{m} |\psi_j\rangle\langle\psi_j|\psi_k\rangle\langle\psi_k| \\
&= \sum_{k=1}^{m} |\psi_k\rangle\langle\psi_k| \\
&= \Pi,
\end{aligned}$$

where the orthonormality of $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ implies the second-to-last equality.

In fact, this exhausts all of the possibilities; *every* projection Π can be written in the form (5) for some choice of an orthonormal set $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$. (Technically speaking, the zero matrix $\Pi = 0$, which is a projection, is a special case. To fit it into the general form (5) we must allow the possibility that the sum is empty, resulting in the zero matrix.)

Projective measurements

The notion of a measurement of a quantum system is more general than just standard basis measurements. *Projective measurements* are measurements that are described by a collection of projections whose sum is equal to the

identity matrix. In symbols, a collection $\{\Pi_0, \ldots, \Pi_{m-1}\}$ of projection matrices describes a projective measurement if

$$\Pi_0 + \cdots + \Pi_{m-1} = \mathbb{I}.$$

When such a measurement is performed on a system X while it is in some state $|\psi\rangle$, two things happen:

1. For each $k \in \{0, \ldots, m-1\}$, the outcome of the measurement is k with probability equal to

$$\Pr\big(\text{outcome is } k\big) = \big\|\Pi_k|\psi\rangle\big\|^2.$$

2. For whichever outcome k the measurement produces, the state of X becomes

$$\frac{\Pi_k|\psi\rangle}{\big\|\Pi_k|\psi\rangle\big\|}.$$

We can also choose outcomes other than $\{0, \ldots, m-1\}$ for projective measurements if we wish. More generally, for any finite and nonempty set Σ, if we have a collection of projection matrices

$$\{\Pi_a : a \in \Sigma\}$$

that satisfies the condition

$$\sum_{a \in \Sigma} \Pi_a = \mathbb{I},$$

then this collection describes a projective measurement whose possible outcomes coincide with the set Σ, where the rules are the same as before:

1. For each $a \in \Sigma$, the outcome of the measurement is a with probability equal to

$$\Pr\big(\text{outcome is } a\big) = \big\|\Pi_a|\psi\rangle\big\|^2.$$

2. For whichever outcome a the measurement produces, the state of X becomes

$$\frac{\Pi_a|\psi\rangle}{\big\|\Pi_a|\psi\rangle\big\|}.$$

For example, standard basis measurements are equivalent to projective measurements, where Σ is the set of classical states of whatever system X we're talking about and our set of projection matrices is $\{|a\rangle\langle a| : a \in \Sigma\}$.

Another example of a projective measurement, this time on two qubits (X, Y), is given by the set $\{\Pi_0, \Pi_1\}$, where

$$\Pi_0 = |\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+| \quad \text{and} \quad \Pi_1 = |\psi^-\rangle\langle\psi^-|.$$

If we have multiple systems that are jointly in some quantum state and a projective measurement is performed on just one of the systems, the action is similar to what we had for standard basis measurements — and in fact we can now describe this action in much simpler terms than we could before. To be precise, let us suppose that we have two systems (X, Y) in a quantum state $|\psi\rangle$, and a projective measurement described by a collection $\{\Pi_a : a \in \Sigma\}$ is performed on the system X, while nothing is done to Y. Doing this is then equivalent to performing the projective measurement described by the collection

$$\{\Pi_a \otimes \mathbb{I} : a \in \Sigma\}$$

on the joint system (X, Y). Each measurement outcome a results with probability

$$\big\|(\Pi_a \otimes \mathbb{I})|\psi\rangle\big\|^2,$$

and conditioned on the result a appearing, the state of the joint system (X, Y) becomes

$$\frac{(\Pi_a \otimes \mathbb{I})|\psi\rangle}{\big\|(\Pi_a \otimes \mathbb{I})|\psi\rangle\big\|}.$$

Implementing projective measurements using standard basis measurements

Arbitrary projective measurements can be implemented using unitary operations, standard basis measurements, and an extra workspace system, as will now be explained.

Let us suppose that X is a system and $\{\Pi_0, \ldots, \Pi_{m-1}\}$ is a projective measurement on X. We can easily generalize this discussion to projective measurements having different sets of outcomes, but in the interest of convenience and simplicity we will assume the set of possible outcomes for

our measurement is $\{0, \ldots, m-1\}$. Let us note explicitly that m is not necessarily equal to the number of classical states of X — we'll let n be the number of classical states of X, which means that each matrix Π_k is an $n \times n$ projection matrix. Because we assume that $\{\Pi_0 \ldots, \Pi_{m-1}\}$ represents a projective measurement, it is necessarily the case that

$$\sum_{k=0}^{m-1} \Pi_k = \mathbb{I}_n.$$

Our goal is to perform a process that has the same effect as performing this projective measurement on X, but to do this using only unitary operations and standard basis measurements.

We will make use of an extra workspace system Y to do this, and specifically we'll take the classical state set of Y to be $\{0, \ldots, m-1\}$, which is the same as the set of outcomes of the projective measurement. The idea is that we will perform a standard basis measurement on Y, and interpret the outcome of this measurement as being equivalent to the outcome of the projective measurement on X. We'll need to assume that Y is initialized to some fixed state, which we'll choose to be $|0\rangle$. (Any other choice of fixed quantum state vector could be made to work, but choosing $|0\rangle$ makes the explanation to follow much simpler.)

Of course, in order for a standard basis measurement of Y to tell us anything about X, we will need to allow X and Y to interact somehow before measuring Y, by performing a unitary operation on the system (Y, X). First consider this matrix:

$$M = \sum_{k=0}^{m-1} |k\rangle\langle 0| \otimes \Pi_k.$$

Expressed explicitly as a so-called *block matrix*, which is essentially a matrix of matrices that we interpret as a single, larger matrix, M looks like this:

$$M = \begin{pmatrix} \Pi_0 & 0 & \cdots & 0 \\ \Pi_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \Pi_{m-1} & 0 & \cdots & 0 \end{pmatrix}.$$

Here, each 0 represents an $n \times n$ matrix filled entirely with zeros, so that the entire matrix M is an $nm \times nm$ matrix.

Now, M is certainly not a unitary matrix (unless $m = 1$, in which case $\Pi_0 = \mathbb{I}$, giving $M = \mathbb{I}$ in this trivial case) because unitary matrices cannot have any columns (or rows) that are entirely 0; unitary matrices have columns that form orthonormal bases, and the all-zero vector is not a unit vector. However, it is the case that the first n columns of M are orthonormal, and we get this from the assumption that $\{\Pi_0, \ldots, \Pi_{m-1}\}$ is a measurement. To verify this claim, notice that for each $j \in \{0, \ldots, n-1\}$, column number j of M is as follows.

$$|\psi_j\rangle = M|0, j\rangle = \sum_{k=0}^{m-1} |k\rangle \otimes \Pi_k |j\rangle.$$

Note that here we're numbering the columns starting from column 0. Taking the inner product of column i with column j when $i, j \in \{0, \ldots, n-1\}$ gives

$$\begin{aligned}
\langle \psi_i | \psi_j \rangle &= \left(\sum_{k=0}^{m-1} |k\rangle \otimes \Pi_k |i\rangle \right)^{\dagger} \left(\sum_{l=0}^{m-1} |l\rangle \otimes \Pi_l |j\rangle \right) \\
&= \sum_{k=0}^{m-1} \sum_{l=0}^{m-1} \langle k|l\rangle \langle i|\Pi_k \Pi_l |j\rangle \\
&= \sum_{k=0}^{m-1} \langle i|\Pi_k \Pi_k |j\rangle \\
&= \sum_{k=0}^{m-1} \langle i|\Pi_k |j\rangle \\
&= \langle i|\mathbb{I}|j\rangle \\
&= \begin{cases} 1 & i = j \\ 0 & i \neq j, \end{cases}
\end{aligned}$$

which is what we needed to show.

Thus, because the first n columns of the matrix M are orthonormal, we can replace all of the remaining zero entries by some different choice of complex number entries so that the entire matrix is unitary:

$$U = \begin{pmatrix} \Pi_0 & \boxed{?} & \cdots & \boxed{?} \\ \Pi_1 & \boxed{?} & \cdots & \boxed{?} \\ \vdots & \vdots & \ddots & \vdots \\ \Pi_{m-1} & \boxed{?} & \cdots & \boxed{?} \end{pmatrix}$$

If we're given the matrices $\Pi_0, \ldots, \Pi_{m-1}$, we can compute suitable matrices to fill in for the blocks marked $\boxed{?}$ in the equation — using the Gram–Schmidt process — but it does not matter specifically what these matrices are for the sake of this discussion.

Finally we can describe the measurement process: we first perform U on the joint system (Y, X) and then measure Y with respect to a standard basis measurement. For an arbitrary state $|\phi\rangle$ of X, we obtain the state

$$U\big(|0\rangle|\phi\rangle\big) = M\big(|0\rangle|\phi\rangle\big) = \sum_{k=0}^{m-1} |k\rangle \otimes \Pi_k|\phi\rangle,$$

where the first equality follows from the fact that U and M agree on their first n columns. When we perform a projective measurement on Y, we obtain each outcome k with probability

$$\big\|\Pi_k|\phi\rangle\big\|^2,$$

in which case the state of (Y, X) becomes

$$|k\rangle \otimes \frac{\Pi_k|\phi\rangle}{\big\|\Pi_k|\phi\rangle\big\|}.$$

Thus, Y stores a copy of the measurement outcome and X changes precisely as it would had the projective measurement described by $\{\Pi_0, \ldots, \Pi_{m-1}\}$ been performed directly on X.

Limitations on quantum information

Despite sharing a common underlying mathematical structure, quantum and classical information have key differences. As a result, there are many examples of tasks that quantum information allows but classical information does not. Before exploring some of these examples, however, we'll take note

of some important limitations on quantum information. Understanding things quantum information *can't* do helps us identify the things it *can* do.

Irrelevance of global phases

The first limitation we'll cover — which is really more of a slight degeneracy in the way that quantum states are represented by quantum state vectors, as opposed to an actual limitation — concerns the notion of a *global phase*.

What we mean by a global phase is this. Let $|\psi\rangle$ and $|\phi\rangle$ be unit vectors representing quantum states of some system, and suppose that there exists a complex number α on the unit circle, meaning that $|\alpha| = 1$, or alternatively $\alpha = e^{i\theta}$ for some real number θ, such that

$$|\phi\rangle = \alpha|\psi\rangle.$$

The vectors $|\psi\rangle$ and $|\phi\rangle$ are then said to *differ by a global phase*. We also sometimes refer to α as a *global phase*, although this is context-dependent; any number on the unit circle can be thought of as a global phase when multiplied to a unit vector.

Consider what happens when a system is in one of the two quantum states $|\psi\rangle$ and $|\phi\rangle$, and the system undergoes a standard basis measurement. In the first case, in which the system is in the state $|\psi\rangle$, the probability of measuring any classical state a is

$$\big|\langle a|\psi\rangle\big|^2.$$

In the second case, in which the system is in the state $|\phi\rangle$, the probability of measuring any classical state a is

$$\big|\langle a|\phi\rangle\big|^2 = \big|\alpha\langle a|\psi\rangle\big|^2 = |\alpha|^2\big|\langle a|\psi\rangle\big|^2 = \big|\langle a|\psi\rangle\big|^2,$$

because $|\alpha| = 1$. That is, the probability of an outcome appearing is the same for both states.

Now consider what happens when we apply an arbitrary unitary operation U to both states. In the first case, in which the initial state is $|\psi\rangle$, the state becomes

$$U|\psi\rangle,$$

and in the second case, in which the initial state is $|\phi\rangle$, it becomes

$$U|\phi\rangle = \alpha U|\psi\rangle.$$

That is, the two resulting states still differ by the same global phase α.

Consequently, two quantum states $|\psi\rangle$ and $|\phi\rangle$ that differ by a global phase are completely indistinguishable; no matter what operation, or sequence of operations, we apply to the two states, they will always differ by a global phase, and performing a standard basis measurement will produce outcomes with precisely the same probabilities as the other. For this reason, two quantum state vectors that differ by a global phase are considered to be equivalent, and are effectively viewed as being the same state.

For example, the quantum states

$$|-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle \quad \text{and} \quad -|-\rangle = -\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$

differ by a global phase (which is -1 in this example), and are therefore considered to be the same state.

On the other hand, the quantum states

$$|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \quad \text{and} \quad |-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$$

do not differ by a global phase. Although the only difference between the two states is that a plus sign turns into a minus sign, this is not a *global* phase difference, it is a *relative* phase difference because it does not affect every vector entry, but only a proper subset of the entries. This is consistent with what we have already observed previously, which is that the states $|+\rangle$ and $|-\rangle$ can be discriminated perfectly. In particular, performing a Hadamard operation and then measuring yields outcome probabilities as follows:

$$\begin{aligned} \big|\langle 0|H|+\rangle\big|^2 &= 1 \qquad & \big|\langle 0|H|-\rangle\big|^2 &= 0 \\ \big|\langle 1|H|+\rangle\big|^2 &= 0 \qquad & \big|\langle 1|H|-\rangle\big|^2 &= 1. \end{aligned}$$

No-cloning theorem

The *no-cloning theorem* shows it is impossible to create a perfect copy of an unknown quantum state.

> **Theorem (No-cloning theorem).** Let Σ be a classical state set having at least two elements, and let X and Y be systems sharing the same classical state set Σ. There does not exist a quantum state $|\phi\rangle$ of Y and a unitary operation U on the pair (X, Y) such that $U\big(|\psi\rangle \otimes |\phi\rangle\big) = |\psi\rangle \otimes |\psi\rangle$ for every state $|\psi\rangle$ of X.

That is, there is no way to initialize the system Y (to any state $|\phi\rangle$ whatsoever) and perform a unitary operation U on the joint system (X, Y) so that the effect is for the state $|\psi\rangle$ of X to be *cloned* — resulting in (X, Y) being in the state $|\psi\rangle \otimes |\psi\rangle$.

The proof of this theorem is actually quite simple: it boils down to the observation that the mapping

$$|\psi\rangle \otimes |\phi\rangle \mapsto |\psi\rangle \otimes |\psi\rangle$$

is not linear in $|\psi\rangle$.

In particular, because Σ has at least two elements, we may choose $a, b \in \Sigma$ with $a \neq b$. If there did exist a quantum state $|\phi\rangle$ of Y and a unitary operation U on the pair (X, Y) for which $U\big(|\psi\rangle \otimes |\phi\rangle\big) = |\psi\rangle \otimes |\psi\rangle$ for every quantum state $|\psi\rangle$ of X, then it would be the case that

$$U\big(|a\rangle \otimes |\phi\rangle\big) = |a\rangle \otimes |a\rangle \quad \text{and} \quad U\big(|b\rangle \otimes |\phi\rangle\big) = |b\rangle \otimes |b\rangle.$$

By linearity, meaning specifically the linearity of the tensor product in the first argument and the linearity of matrix-vector multiplication in the second (vector) argument, we must therefore have

$$U\left(\left(\frac{1}{\sqrt{2}}|a\rangle + \frac{1}{\sqrt{2}}|b\rangle\right) \otimes |\phi\rangle\right) = \frac{1}{\sqrt{2}}|a\rangle \otimes |a\rangle + \frac{1}{\sqrt{2}}|b\rangle \otimes |b\rangle.$$

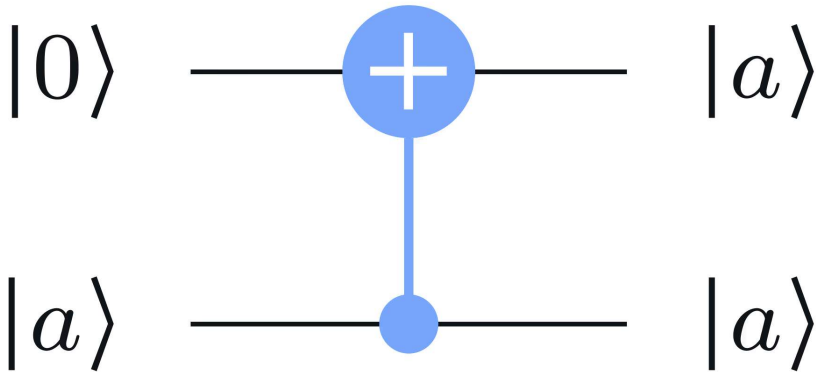
However, the requirement that $U\big(|\psi\rangle \otimes |\phi\rangle\big) = |\psi\rangle \otimes |\psi\rangle$ for every quantum state $|\psi\rangle$ demands that

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

Therefore there cannot exist a state $|\phi\rangle$ and a unitary operation U for which $U\big(|\psi\rangle \otimes |\phi\rangle\big) = |\psi\rangle \otimes |\psi\rangle$ for every quantum state vector $|\psi\rangle$.

A few remarks concerning the no-cloning theorem are in order. The first one is that the statement of the no-cloning theorem above is absolute, in the sense that it states that *perfect* cloning is impossible — but it does not say anything about possibly cloning with limited accuracy, where we might succeed in producing an approximate clone (with respect to some way of measuring how similar two different quantum states might be). There are, in fact, statements of the no-cloning theorem that place limitations on approximate cloning, as well as methods to achieve approximate cloning with limited accuracy.

The second remark is that the no-cloning theorem is a statement about the impossibility of cloning an *arbitrary* state $|\psi\rangle$. In contrast, we can easily create a clone of any standard basis state, for instance. For example, we can clone a qubit standard basis state using a controlled-NOT operation:



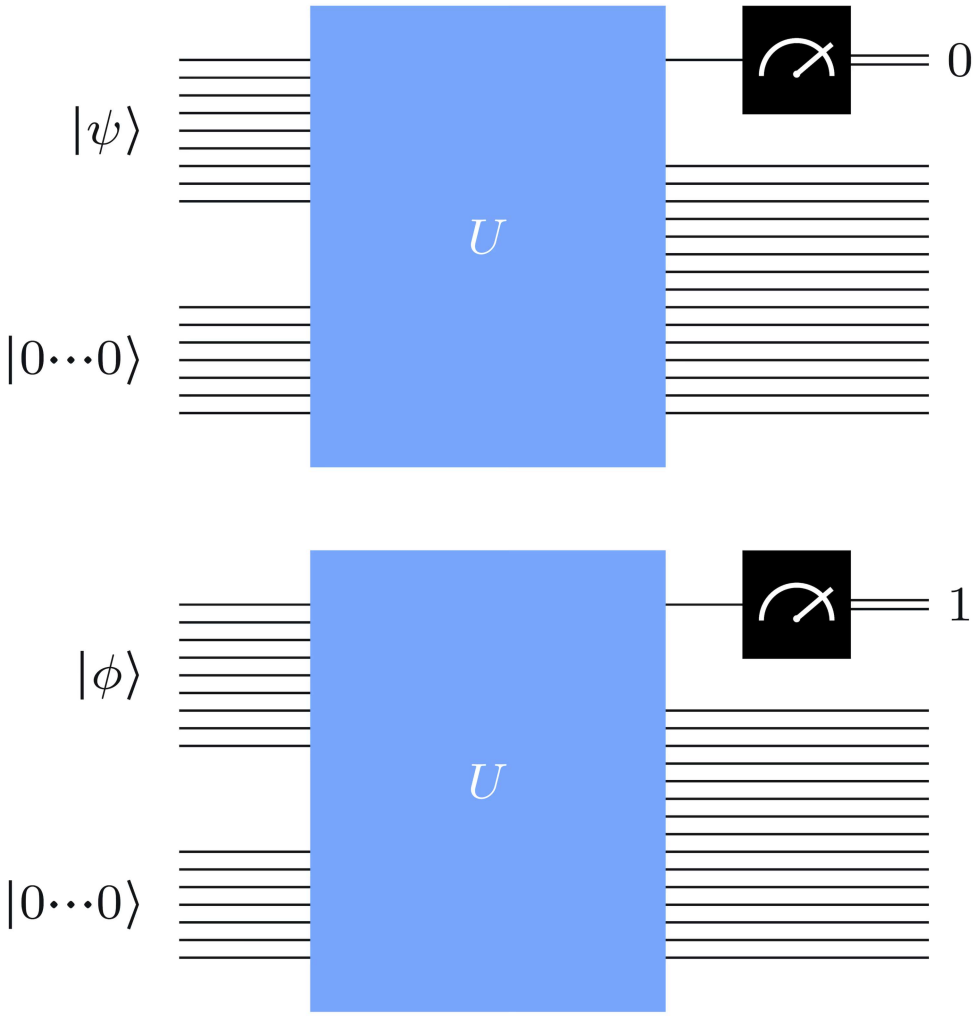
While there is no difficulty in creating a clone of a standard basis state, this does not contradict the no-cloning theorem — this approach of using a controlled-NOT gate would not succeed in creating a clone of the state $|+\rangle$, for instance.

One final remark about the no-cloning theorem is that it really isn't unique to quantum information — it's also impossible to clone an arbitrary probabilistic state using a classical (deterministic or probabilistic) process. This is pretty intuitive. Imagine someone hands you a system in some probabilistic state, but you're not sure what that probabilistic state is. For example, maybe they randomly generated a number between 1 and 10, but they didn't tell you how they generated that number. There's certainly no physical process through which you can obtain two *independent* copies of that same probabilistic state: all you have in your hands is a number between 1 and 10, and there just isn't enough information present for you to somehow reconstruct the probabilities for all of the other outcomes to appear. Mathematically speaking, a version of the no-cloning theorem for probabilistic states can be proved in exactly the same way as the regular no-cloning theorem (for quantum states). That is, cloning an arbitrary probabilistic state is a non-linear process, so it cannot possibly be represented by a stochastic matrix.

Non-orthogonal states cannot be perfectly discriminated

For the final limitation to be covered in this lesson, we'll show that if we have two quantum states $|\psi\rangle$ and $|\phi\rangle$ that are not orthogonal, which means that $\langle\phi|\psi\rangle \neq 0$, then it's impossible to discriminate them (or, in other words, to tell them apart) perfectly. In fact, we'll show something logically equivalent: if we do have a way to discriminate two states perfectly, without any error, then they must be orthogonal.

We will restrict our attention to quantum circuits that consist of any number of unitary gates, followed by a single standard basis measurement of the top qubit. What we require of a quantum circuit, to say that it perfectly discriminates the states $|\psi\rangle$ and $|\phi\rangle$, is that the measurement always yields the value 0 for one of the two states and always yields 1 for the other state. To be precise, we shall assume that we have a quantum circuit that operates as the following diagrams suggest:



The box labeled U denotes the unitary operation representing the combined action of all of the unitary gates in our circuit, but not including the final measurement. There is no loss of generality in assuming that the measurement outputs 0 for $|\psi\rangle$ and 1 for $|\phi\rangle$; the analysis would not differ fundamentally if these output values were reversed.

Notice that, in addition to the qubits that initially store either $|\psi\rangle$ or $|\phi\rangle$, the circuit is free to make use of any number of additional *workspace* qubits. These qubits are initially each set to the $|0\rangle$ state — so their combined state is denoted $|0\cdots0\rangle$ in the figures — and these qubits can be used by the circuit in any way that might be beneficial. It is very common to make use of workspace qubits in quantum circuits like this.

Now, consider what happens when we run our circuit on the state $|\psi\rangle$ (along with the initialized workspace qubits). The resulting state, immediately prior

to the measurement being performed, can be written as

$$U\big(|0\cdots 0\rangle|\psi\rangle\big) = |\gamma_0\rangle|0\rangle + |\gamma_1\rangle|1\rangle$$

for two vectors $|\gamma_0\rangle$ and $|\gamma_1\rangle$ that correspond to all of the qubits except the top qubit. In general, for such a state the probabilities that a measurement of the top qubit yields the outcomes 0 and 1 are as follows:

$$\Pr(\text{outcome is } 0) = \big\||\gamma_0\rangle\big\|^2 \qquad \text{and} \qquad \Pr(\text{outcome is } 1) = \big\||\gamma_1\rangle\big|$$

Multiple systems Entanglement in action →