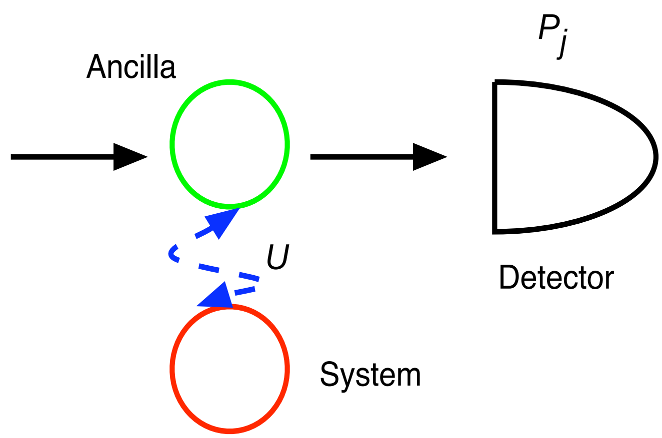
Indirect measurement

Because in practice it is often difficult to directly measure a quantum system without destroying it, most actual measurement are done indirectly.



1. Prepare an *extra* system (an *ancilla*) in a *known* initial state.

2. Have the system and the ancilla interact by carrying out some circuit (performing a unitary transformation on the two of them).

3. Measure the ancilla and discard it.

What does this look like mathematically? If the system is in state $|\psi\rangle$, and the ancilla is in the standard state (say $|0\rangle$), then the *joint* state of the system-plus-ancilla is

$$|\psi\rangle \otimes |0\rangle.$$

These then undergo a joint unitary transformation

$$\hat{U}(|\psi\rangle \otimes |0\rangle).$$

The ancilla is then measured. This measurement is given by some set of orthogonal projectors $\hat{\mathcal{P}}_j = \hat{\mathcal{P}}_j^\dagger$, where j labels the possible outcomes.

$$\hat{\mathcal{P}}_j \hat{\mathcal{P}}_k = \delta_{jk} \hat{\mathcal{P}}_j, \quad \sum_j \hat{\mathcal{P}}_j = \hat{I}.$$

For the present, we assume that this must be a *complete local measurement* which leaves the state and ancilla unentangled.

If the measurement is local, then the projectors have the form

$$\hat{\mathcal{P}}_j = \hat{I} \otimes \hat{Q}_j$$

where the $\{\hat{Q}_j\}$ are a decomposition of the identity on the ancilla space. In order for them to represent a complete local measurement, they must all be one-dimensional:

$$\hat{Q}_j = |\phi_j\rangle\langle\phi_j|,$$

where $\{|\phi_j\rangle\}$ is an orthonormal basis on the ancilla space.

The probability of outcome j is

$$p_j = (\langle\psi| \otimes \langle 0|)\hat{U}^\dagger \hat{\mathcal{P}}_j \hat{U}(|\psi\rangle \otimes |0\rangle).$$

Note that $\{\hat{U}^\dagger \hat{\mathcal{P}}_j \hat{U}\}$ is still a set of orthogonal projectors! So we can also think of this as a joint measurement on the system and ancilla.

Note that the above expression for p_j is bilinear in $|\psi\rangle$. This means that we can find a set of positive operators $\hat{E}_j$ on the space of the system alone such that

$$p_j = \langle\psi|\hat{E}_j|\psi\rangle, \quad \sum_j \hat{E}_j = \hat{I}.$$

Here is one way of seeing that. Let $\{|i\rangle_S\}$ be a basis for the system and $\{|j\rangle_A\}$ be a basis for the ancilla (which includes the state $|0\rangle$). Then $\{|i\rangle_S \otimes |j\rangle_A\}$ is a basis for the joint system and ancilla. Any operator can be written in terms of this basis

$$\hat{O} = \sum_{i,i',j,j'} O_{(ij)(i'j')}|i\rangle\langle i'| \otimes |j\rangle\langle j'|.$$

We can express the joint state $|\psi\rangle \otimes |0\rangle$ as

$$|\psi\rangle \otimes |0\rangle = \sum_i \alpha_i |i\rangle \otimes |0\rangle.$$

The expectation value then becomes

$$(\langle\psi|\otimes\langle 0|)\hat{O}(|\psi\rangle\otimes|0\rangle) = \sum_{i,i'} \alpha_i^* O_{(i0)(i'0)} \alpha_{i'},$$

which is the same as $\langle\psi|\hat{E}|\psi\rangle$ for an operator on the system space alone defined by

$$\hat{E} = \sum_{i,i'} O_{(i0)(i'0)} |i\rangle\langle i'|.$$

In the case of our indirect measurement, the operators $\hat{E}_j$ would be defined by substituting $\hat{O} = \hat{U}^\dagger \hat{\mathcal{P}}_j \hat{U}$. The operators $\hat{E}_j$ must be positive because the probabilities p_j are non-negative for every $|\psi\rangle$, and they must sum to the identity because $\sum_j p_j = 1$ for every $|\psi\rangle$. But note that unlike projective measurements, the $\hat{E}_j$ need not be projectors! This is therefore a generalization, called *Positive Operator Valued Measurement* (or POVM), which includes projective measurements as a special case.

This derivation in terms of indices with respect to a particular basis is perfectly adequate. But it can be made much more elegantly by a useful mathematical tool: the *partial trace.*

Recall that the usual trace is the sum of the diagonal elements of an operator:

$$\mathrm{Tr}\{\hat{O}\} = \sum_j \langle j|\hat{O}|j\rangle.$$

This is a linear map from operators to complex numbers: $\mathrm{Tr}\{a\hat{A} + b\hat{B}\} = a\mathrm{Tr}\{\hat{A}\} + b\mathrm{Tr}\{\hat{B}\}$.

Suppose now that our space is a tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$. For a product operator $\hat{A} \otimes \hat{B}$ on this space, the *partial trace of* $\hat{A} \otimes \hat{B}$ *over* B is

$$\mathrm{Tr}_B\{\hat{A} \otimes \hat{B}\} = \mathrm{Tr}\{\hat{B}\}\hat{A},$$

where the trace on the right-hand side is on $\mathcal{H}_B$, and the result is an operator on $\mathcal{H}_A$.

Naturally, we could also trace over A:

$$\mathrm{Tr}_A\{\hat{A} \otimes \hat{B}\} = \mathrm{Tr}\{\hat{A}\}\hat{B}.$$

The definition extends to general operators $\hat{O}$ on $\mathcal{H}_A \otimes \mathcal{H}_B$ by linearity:

$$\hat{O} = \sum_\ell \hat{A}_\ell \otimes \hat{B}_\ell, \quad \mathrm{Tr}_B\{\hat{O}\} = \sum_\ell \mathrm{Tr}\{\hat{B}_\ell\}\hat{A}_\ell.$$

If we represent $\hat{O}$ in terms of bases $\{|i\rangle\}$ and $\{|j\rangle\}$ on $\mathcal{H}_A$ and $\mathcal{H}_B$

$$\hat{O} = \sum_{i,i',j,j'} O_{(ij)(i'j')} |i\rangle\langle i'| \otimes |j\rangle\langle j'|,$$

then

$$\mathrm{Tr}_B\{\hat{O}\} = \sum_{i,i',j} O_{(ij)(i'j)} |i\rangle\langle i'|,$$

$$\mathrm{Tr}_A\{\hat{O}\} = \sum_{i,j,j'} O_{(ij)(ij')} |j\rangle\langle j'|.$$

We can then define the operators $\hat{E}_j$ in terms of the partial trace in a very natural (and basis-independent) way. Note that

$$p_j = \mathrm{Tr}\{(|\psi\rangle\langle\psi| \otimes |0\rangle\langle 0|)(\hat{U}^\dagger \hat{\mathcal{P}}_j \hat{U})\}.$$

We define the positive operators

$$\hat{E}_j = \mathrm{Tr}_{\mathrm{anc}}\{(\hat{I}_{\mathrm{sys}} \otimes |0\rangle\langle 0|)(\hat{U}^\dagger \hat{\mathcal{P}}_j \hat{U})\}.$$

This gives us

$$p_j = \mathrm{Tr}\{|\psi\rangle\langle\psi|\hat{E}_j\} = \langle\psi|\hat{E}_j|\psi\rangle,$$

as desired.

Uses of POVMs

Since POVMs are more general than projective measurements, in some cases they make possible new and different tricks.

Remember the problem of distinguishing two nonorthogonal states $|\psi\rangle$ and $|\phi\rangle$, $\langle\psi|\phi\rangle \neq 0$. The best we can do with a projective measurement is use $\hat{\mathcal{P}}_0 = |\psi\rangle\langle\psi|$, $\hat{\mathcal{P}}_1 = \hat{I} - \hat{\mathcal{P}}_0$. In this case, if we get result 1 we know that the system *cannot* have been in state ψ, since $\langle\psi|\hat{\mathcal{P}}_1|\psi\rangle = 0$. So it must have been in state ϕ. In the case of result 0, however, we cannot tell.

We can design a similar projective measurement based on $|\phi\rangle$. But we cannot treat the two states symmetrically and still have results that tell us which state was prepared with certainty.

What if we are allowed to do a POVM? Consider the following three operators:

$$\hat{E}_0 = a(\hat{I} - |\psi\rangle\langle\psi|),$$
$$\hat{E}_1 = a(\hat{I} - |\phi\rangle\langle\phi|),$$
$$\hat{E}_2 = \hat{I} - \hat{E}_0 - \hat{E}_1.$$

Clearly $\sum_j \hat{E}_j = \hat{I}$. The number $1/2 \leq a \leq 1$ must be chosen small enough that $\hat{E}_2$ is a *positive* operator. What are the probabilities of these outputs for the states $|\psi\rangle$ and $|\phi\rangle$?

In state $|\psi\rangle$,

$$p_0 = 0, p_1 = a(1 - |\langle\psi|\phi\rangle|^2), p_2 = 1 - p_1.$$

In $|\phi\rangle$,

$$p_0 = a(1 - |\langle\psi|\phi\rangle|^2), p_1 = 0, p_2 = 1 - p_0.$$

So if we get result 1 we know the system *must* have been in state ψ, and if we get result 0 we know it *must* have been in state ϕ. Only if we get result 2 are we unable to tell.

Generalized measurements

POVMs tell us the *probabilities* of outcomes for an indirect measurement, but they don't say what state the system is left in. It would be useful to generalize this aspect of projective measurements as well.

We prepare the ancilla in state $|0\rangle$ and do a joint unitary $\hat{U}$. We then make a complete measurement of the ancilla. Let the eigenbasis for this measurement be $|j\rangle$. Since we have control over both the initial state of the ancilla and the unitary transformation $\hat{U}$ (i.e., which circuit is performed), without loss of generality we can assume that the initial ancilla state $|0\rangle$ is also a member of this basis. After measurement result j, the (unrenormalized) system and ancilla are in the state

$$(\hat{I} \otimes |j\rangle\langle j|)\hat{U}(|\psi\rangle \otimes |0\rangle) = |\psi_j\rangle \otimes |j\rangle.$$

What is $|\psi_j\rangle$?

The key insight is that this expression is linear is $|\psi\rangle$; therefore, there must be some operator $\hat{M}_j$ such that

$$|\psi_j\rangle = \hat{M}_j|\psi\rangle.$$

The probabilities are $p_j = \langle\psi_j|\psi_j\rangle = \langle\psi|\hat{M}_j^\dagger\hat{M}_j|\psi\rangle$, so we must have

$$\hat{E}_j = \hat{M}_j^\dagger\hat{M}_j, \quad \sum_j \hat{M}_j^\dagger\hat{M}_j = \hat{I}.$$

We call the set of operators $\{\hat{M}_j\}$ the *measurement operators* corresponding to the generalized measurement; $\hat{E}_j$ is the *POVM element* corresponding to outcome j. Note that $\hat{M}_j^\dagger\hat{M}_j$ is automatically a positive operator. This description in terms of operators $\{\hat{M}_j\}$ is called a *Kraus representation*.

How do we find the operators $\{\hat{M}_j\}$, given the initial state $|0\rangle$, the unitary transformation $\hat{U}$, and the eigenbasis $\{|j\rangle\}$? We could write $\hat{U}$ in terms of its matrix elements $U_{(ij)(i'j')}$, but we will instead *partially* decompose $\hat{U}$ just in terms of the ancilla basis:

$$\hat{U} = \sum_{j,j'} \hat{A}_{jj'} \otimes |j\rangle\langle j'|,$$

where the operators $\hat{A}_{jj'}$ act on the system Hilbert space. (In effect, we have chopped up the matrix $\hat{U}$ into $D \times D$ blocks.)

We then get

$$\hat{U}(|\psi\rangle \otimes |0\rangle) = \sum_j \hat{A}_{j0}|\psi\rangle \otimes |j\rangle,$$

and our expression for $|\psi_j\rangle$ becomes

$$|\psi_j\rangle = \hat{A}_{j0}|\psi\rangle \equiv \hat{M}_j|\psi\rangle.$$

We have seen that for an ancilla in state $|0\rangle$, a unitary transformation $\hat{U}$ and a measurement basis $|j\rangle$, we can find measurement operators $\{\hat{M}_j\}$ which describe the measurement outcome

$$|\psi\rangle \rightarrow \hat{M}_j|\psi\rangle/\sqrt{p_j}, \quad p_j = \langle \hat{M}_j^\dagger \hat{M}_j \rangle.$$

But is the converse true? Given any set of m operators $\{\hat{M}_j\}$, can we find an ancilla, unitary transformation and measurement which will enact this generalized measurement? The answer is *yes*, given a couple of important provisos.

First (and most obvious): if there are m measurement operators, the ancilla must have dimension $\geq m$. Next, we need to find a unitary transformation $\hat{U}$ such that

$$\hat{U}(|\psi\rangle \otimes |0\rangle) = \sum_j \hat{M}_j|\psi\rangle \otimes |j\rangle.$$

Writing $\hat{U}$ in terms of a partial decomposition

$$\hat{U} = \sum_{j,j'} \hat{A}_{jj'} \otimes |j\rangle\langle j'|,$$

we identify $\hat{A}_{j0} = \hat{M}_j$. Unitarity requires that

$$\hat{U}^\dagger \hat{U} = \hat{I} \otimes \hat{I}$$

which means that

$$\sum_k \hat{A}^\dagger_{jk} \hat{A}_{kj'} = \delta_{jj'} \hat{I}.$$

In particular, for $j = j' = 0$ we must have

$$\sum_j \hat{M}^\dagger_j \hat{M}_j = \hat{I}.$$

Given that, it is always possible to find a set of other operators $\hat{A}_{jj'}$ for $j' > 0$ which make $\hat{U}$ unitary. So for any set of measurement operators that satisfy this sum, we can do the generalized measurement.

If the measurement has only two possible outcomes, and the measurement operators are Hermitian, then we can choose the ancilla to be a single q-bit, and there is a particularly simple form for this unitary. Let

$$\hat{M}_i = \hat{M}_i^\dagger, \quad \hat{M}_0^2 + \hat{M}_1^2 = \hat{I}, \quad [\hat{M}_0, \hat{M}_1] = 0.$$

$$\hat{U} = \hat{M}_0 \otimes \hat{Z} + \hat{M}_1 \otimes \hat{X},$$

It is easy to check that this is unitary:

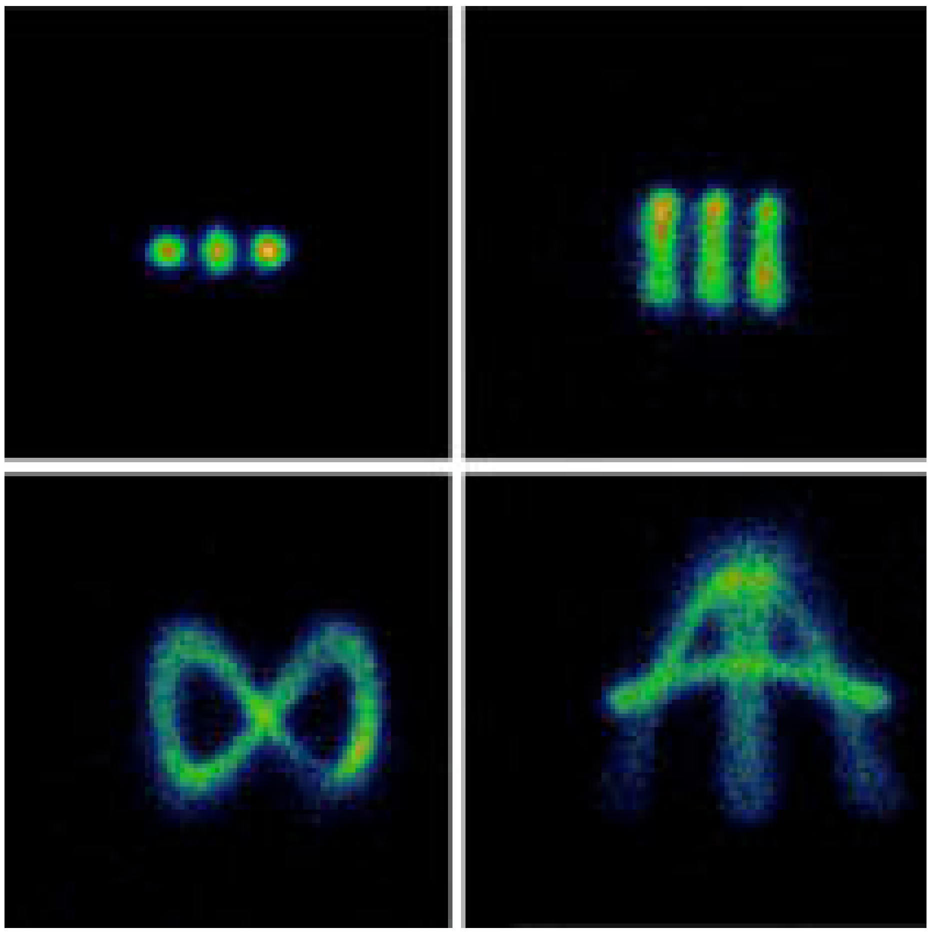
$$\hat{U}^\dagger \hat{U} = (\hat{M}_0^2 + \hat{M}_1^2) \otimes \hat{I} + \hat{M}_0 \hat{M}_1 \otimes (\hat{X}\hat{Z} + \hat{Z}\hat{X})$$
$$= \hat{I} \otimes \hat{I} = \hat{I},$$

and that

$$\hat{U}(|\psi\rangle \otimes |0\rangle) = \hat{M}_0|\psi\rangle \otimes |0\rangle + \hat{M}_1|\psi\rangle \otimes |1\rangle.$$

Other cases are more complicated than this, but the principle is similar. (Actually, any generalized measurement can be broken down into two-outcome measurements like this and unitaries conditioned on the outcomes.)

What does such a generalized measurement look like in practice? Here is a picture of ions in an electromagnetic trap undergoing an indirect measurement by a laser:



The laser is tuned in resonance with a transition from a particular electronic state to an excited, short lived state. If an ion is in that initial state, the laser makes it fluoresce.

Quantum operations

Something to reflect on: this prescription for generalized measurements is *very* general indeed. In fact, it includes things which don't look very much like measurements. Consider the following examples:

$$\hat{M}_0 = (1/\sqrt{2})\hat{I}, \quad \hat{M}_1 = (1/\sqrt{2})\hat{I}.$$

This has $p_0 = p_1 = 1/2$ for all states, and leaves the state totally unchanged. In effect, we haven't measured anything; we've just generated a bit of randomness.

$$\hat{M}_0 = \epsilon|0\rangle\langle 0|, \ \hat{M}_1 = \epsilon|1\rangle\langle 1|, \ \hat{M}_2 = \sqrt{1-\epsilon^2}\hat{I},$$

with $\epsilon \ll 1$. Most of the time, this does nothing. But every once in a while, it does a measurement in the Z basis. We can think of this as an *unreliable measurement device.*

Here is a different kind of unreliable device:

$$\hat{M}_0 = (1/\sqrt{2})\hat{U}_0, \quad \hat{M}_1 = (1/\sqrt{2})\hat{U}_1,$$

where $\hat{U}_0$ and $\hat{U}_1$ are unitary. This is a *probabilistic gate*: half the time it performs unitary $\hat{U}_0$, and half the time $\hat{U}_1$. Such gates arise naturally in certain schemes for quantum computation (notably, the *linear optical* scheme of Knill, Laflamme and Milburn). We can also think of the Bell state measurement in quantum teleportation as performing a probabilistic gate on the transmitted q-bit.

All of these cases had random outcomes. Are there any which are deterministic? This would correspond to a set with only a single operator $\hat{M}$. In this case, our sum normalization implies that $\hat{M}^\dagger \hat{M} = \hat{I}$, i.e., $\hat{M}$ must be unitary. So we see that unitaries are the only deterministic operations in quantum mechanics; and that they, too, count as generalized measurements.

Because of this tremendous breadth, generalized measurements are often referred to as *quantum operations*. These include everything we've seen so far: unitaries, projective measurements, POVMs, and a good deal more besides.

In fact, the class of quantum operations can be expanded in several ways. One (which we will see later) corresponds to doing an *incomplete* measurement on the ancilla. In order to treat this, we will need to learn about density matrices and mixed states. Another (which we will not consider in this course) corresponds to operations with *postselection*, i.e., keeping the system only for certain measurement outcomes and discarding it otherwise. This lets one relax the requirement that the POVM elements add up to the identity.

Information and Entropy

How much information do we get about a quantum system from a measurement? This is a subtle question, when applied to quantum mechanics, and we will continue to address it throughout the course. One question, though, has a straightforward classical answer: how much information is *produced* by a measurement?

Classically, information is quantified by the *Shannon entropy*:

$$S = -\sum_{j=1}^{m} p_j \log_2 p_j ,$$

where the $\{p_j\}$ are the probabilities of the m different outcomes and the information is measured in bits. We can calculate the information production by plugging in the expressions $p_j = \langle \hat{E}_j \rangle$.

The Shannon entropy characterizes the *uncertainty* of a measurement; the more uncertain the outcome, the more we learn by making the measurement. It is easy to see that S has a minimum of 0, which is achieved when $p_j = 1$ for some j and all other probabilities are 0. The (unique) maximum occurs when all outcomes are equally likely: $p_j = 1/m$, and $S = \log_2 m$.

It would be a mistake, though, to assume that the Shannon entropy of a measurement is equal to the information we gain *about the system.* In fact, we may gain no information at all, as we saw in the case $\hat{M}_0 = \hat{M}_1 = \hat{I}/\sqrt{2}$. The most we can say is that S gives an upper bound on the information gain. Characterizing the information in quantum systems is a difficult and subtle issue.

Next time: Classical computation