Quantum Information

So far we have learned about

Qubits

Mixed states

Classical information theory (Shannon entropy)

Today we will begin to mash these together into a quantum theory of information

For that we'll need

Partial trace

Generalized measurement

Schumacher Compression
(Von Neumann entropy)

Schumacher Compression

(N+C:12.2, P: 5.3)

Classical: $X : x \in \{x_1, \ldots, x_d\}$ with probs. $\{p_1, \ldots, p_d\}$

$K_N[X]$ bits required to store N outcomes

Compression rate $H[X] = \lim_{N\to\infty} \frac{K_N[X]}{N}$

Quantum: $X : |x\rangle \in \{|x_1\rangle, \ldots, |x_d\rangle\}$ with probs. $\{p_1, \ldots, p_d\}$

How many qubits K_N^Q required to store N outcomes?

What is the best possible compression rate?

$$H^Q[X] = S[X] = \lim_{N\to\infty} \frac{K_N[X]}{N}$$

Can be easily found using the classical results:

We first store the possible states (this is just some meta data)

Suppose the possible states are orthogonal $\langle x_i | x_j \rangle = \delta_{ij}$

This lets us measure the outcomes to determine the states

Then we store the x_i as a classical variable, using qubits as bits $0 \to |0\rangle, \; 1 \to |1\rangle$

Total number of qubits required

$$K_N^Q[X] = d + K_N^C[X] \quad \therefore \quad H^Q[X] = H^C[X]$$

What if the states are not orthogonal?

Recall that everything is calculated from the density matrix

So getting states $|x_j\rangle$ with probs p_j is equivalent to getting states $|\tilde{j}\rangle$ with probs q_j if

$$\rho = \sum_{j=1}^{d} p_j |x_j\rangle\langle x_j| = \sum_{j=0}^{D-1} q_j |\tilde{j}\rangle\langle\tilde{j}|$$

So then we can find an equivalent distribution that has orthogonal states by diagonalizing $\langle\tilde{j}|\tilde{k}\rangle = \delta_{jk}$

The compression rate then follows from before

$$H^Q[X] = H[\{q_j\}] = -\sum_j q_j \log q_j$$

Which can be more easily expressed

$$H^Q[X] = S(\rho) = -\mathrm{tr}(\rho \log \rho)$$

The Von Neumann entropy

We can also do it from the very begining:
Schumacher Compression (generalization of Shannon's source coding theorem to density matrices)

A simple preliminary example

$|0\rangle$ WITH PROB. $1/2$
$|+\rangle$ WITH PROB. $1/2$
$\Rightarrow \quad \rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +| = \frac{1}{4}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$

Diagonalizing this $\quad \rho = q_0 |\tilde{0}\rangle\langle\tilde{0}| + q_1 |\tilde{1}\rangle\langle\tilde{1}|$

$|\tilde{0}\rangle = \begin{pmatrix} \cos\pi/8 \\ \sin\pi/8 \end{pmatrix} \quad q_0 = \cos^2\pi/8 \simeq 0.85 \quad , \quad |\tilde{1}\rangle = \begin{pmatrix} \sin\pi/8 \\ -\cos\pi/8 \end{pmatrix} \quad q_1 = \sin^2\pi/8 \simeq 0.15 \quad , \quad \rho \approx |\tilde{0}\rangle\langle\tilde{0}|$

The latter state is quite unlikely, so we could store N outcomes in a single $|\tilde{0}\rangle$ and be right most of the time

This approximation isn't awful, but it isn't good enough for good compression

Consider three outcomes

$$\rho^{\otimes 3} = \rho\otimes\rho\otimes\rho = \sum_{i\in\{0,1\}}\sum_{j\in\{0,1\}}\sum_{k\in\{0,1\}} q_i q_j q_k |\tilde{i}\tilde{j}\tilde{k}\rangle\langle\tilde{i}\tilde{j}\tilde{k}|$$

$|\tilde{0}\tilde{0}\tilde{0}\rangle$ WITH PROB. $\cos^6\left(\frac{\pi}{8}\right) \approx 0.62$

$|\tilde{0}\tilde{0}\tilde{1}\rangle, |\tilde{0}\tilde{1}\tilde{0}\rangle, |\tilde{1}\tilde{0}\tilde{0}\rangle$ WITH PROB. $\cos^4\left(\frac{\pi}{8}\right)\sin^2\left(\frac{\pi}{8}\right) \approx 0.11$

$|\tilde{0}\tilde{1}\tilde{1}\rangle, |\tilde{1}\tilde{0}\tilde{1}\rangle, |\tilde{1}\tilde{1}\tilde{0}\rangle$ WITH PROB. $\cos^2\left(\frac{\pi}{8}\right)\sin^4\left(\frac{\pi}{8}\right) \approx 0.02$

$|\tilde{1}\tilde{1}\tilde{1}\rangle$ WITH PROB. $\sin^6(\pi/8) \approx 0.003$

Most of the state is within a 4D subspace

$\{|\tilde{0}\tilde{0}\tilde{0}\rangle, |\tilde{0}\tilde{0}\tilde{1}\rangle, |\tilde{0}\tilde{1}\tilde{0}\rangle, |\tilde{1}\tilde{0}\tilde{0}\rangle\}$ WITH PROB. ≈ 0.94

Using the approximation

$$\rho^{\otimes 3} \approx \frac{\mathbb{P}\rho\mathbb{P}}{\mathrm{tr}(\mathbb{P}\rho)}$$

$\mathbb{P}$ IS PROJECTOR ONTO

We lose little, but can encode in 2 qubits instead of 3

In general, a random source generates states of a quantum system such that

$$\rho = \sum_{j=0}^{D-1} q_j |\tilde{j}\rangle\langle\tilde{j}| , \quad \langle \tilde{i} | \tilde{j} \rangle = \delta_{ij}$$

So for N outcomes

$$\rho^{\otimes N} = \sum_{j_1 \ldots j_N} q_{j_1 \ldots j_N} |\tilde{j}_1 \ldots \tilde{j}_N\rangle\langle\tilde{j}_1 \ldots \tilde{j}_N| , \qquad q_{j_1 \ldots j_N} = q_{j_1} \times \ldots \times q_{j_N}$$

Consider the projector $\mathbb{P}_k$ onto the k most likely eigenstates

$$\mathrm{tr}(\mathbb{P}_k \rho^{\otimes N}) = 1 - \varepsilon$$

We consider k for which $\varepsilon \ll 1$

This is the subspace on which most of $\rho^{\otimes N}$ has support

N outcomes can be compressed to $K_N^Q = \log_2 k$ qubits

$$H^Q(\rho) = \frac{K_N^Q}{N} = \frac{\log_2 k}{N}$$

So what k do we need?

Law of large numbers:

The most likely strings have $q_0 N$ $|\tilde{0}\rangle$'s AND $q_1 N$ $|\hat{1}\rangle$'s

Number of such states:

$$\binom{N}{Nq_1} = \frac{N!}{(Nq_1)!\,(N-Nq_1)!} = \frac{N!}{(Nq_0)!\,(Nq_1)!}$$

Using this for k, the number of qubits required is

$$K_N^Q = \log_2 \binom{N}{Nq_1} = N\log_2 N - Nq_0\log_2 Nq_0 - Nq_1\log_2 Nq_1 + O(\log N)$$

$$= N H(q_1) + O(\log N)$$

Using Stirling's approximation $\log a! = a\log a - a + O(\log a)$

This k leads to quite a large ε, since most likely the string is not one of the most likely strings

It can be shown (exercises) that ε can be made arbitrarily small when the number of qubits used is

$$K_N^Q = N(H(q_1) + \delta) \qquad H(q_1) = -tr(\rho \log \rho)$$

For arbitrarily small δ and $N \to \infty$

So our compression rate is

$$\frac{K_N^Q}{N} = H[X] = H(q_1) = -\rho \log \rho$$

This holds for any dimension, D, of the quantum system

The quantity $S(\rho) = -\rho \log \rho$

Is the Von Neumann entropy: quantum version of Shannon

Generalized measurement

(POVM: N+C 2.2.6, P 3.1.2)

So far we've just considered projective measurements

single qubit : $\{|0\rangle, |1\rangle\}$, $\{|\varphi_0\rangle, |\varphi_1\rangle\}$

This is not the most general form

For example, suppose you want to measure both σ_x and σ_z

This is impossible due to the uncertainty principle

What if you need to be pretty sure about σ_x , but not certain.
And only have a little hint about σ_z ?

You could take your qubit and an additional 'ancilla' qubit, and interact them a little bit

$$|\psi\rangle \otimes |0\rangle = a|0\rangle\otimes|0\rangle + b|1\rangle\otimes|0\rangle \longrightarrow a|0\rangle\otimes|0\rangle + b|1\rangle\otimes\left(\sqrt{1-\varepsilon^2}|0\rangle + \varepsilon|1\rangle\right)$$

The ancilla now knows a little bit about the z basis state of our qubit

$$a|0\rangle \otimes |0\rangle + b|1\rangle \otimes \left(\sqrt{1-\varepsilon^2}|0\rangle + \varepsilon|1\rangle\right)$$

If we measure the ancilla in the Z basis and get the result $|0\rangle$, it tells us very little. If we get $|1\rangle$ we know that our qubit is in state $|1\rangle$ too. But this is very rare $(\varepsilon \ll 1)$. So, on average, we get very little Z basis information.

What about the X basis?

$$a|0\rangle \otimes |0\rangle + b|1\rangle \otimes \left(\sqrt{1-\varepsilon^2}|0\rangle + \varepsilon|1\rangle\right) = \left(a|0\rangle + b\sqrt{1-\varepsilon^2}|1\rangle\right) \otimes |0\rangle + \varepsilon b|1\rangle \otimes |1\rangle$$

$$\therefore \rho_A = \left(a|0\rangle + b\sqrt{1-\varepsilon^2}|1\rangle\right)\left(a^*\langle 0| + b^*\sqrt{1-\varepsilon^2}\langle 1|\right) + \varepsilon b|1\rangle\langle 1| \approx |\psi\rangle\langle\psi|$$

The state of our qubit has hardly changed, so the X basis measurement will give almost the same result as it would've without the interaction. It will be a bit messed up, though. So we can not entirely trust the result.

Here we have done two measurements. Equivalently, it's a measurement with four outcomes. You can't do that on a qubit using only projectors!

Consider a system, S, whose state want to measure

$$\rho_S \in \mathcal{H}_S$$

And another system, M, whose state we know.

$$\rho_B = |\varphi\rangle\langle\varphi| \in \mathcal{H}_M$$

We can then perform projective measurements on the combined system

$$\mathbb{P}_j = |\mu_j\rangle\langle\mu_j| \in \mathcal{H}_S \otimes \mathcal{H}_M \qquad p_j = tr(\mathbb{P}_j \; \rho_S \otimes \rho_B)$$

These could be entangling measurements, or they could measure a product basis after interaction.

They could be one big measurement with many outcomes, or the combination of multiple ones.

If multiple, the basis choice for later measurements could depend on the outcome of earlier ones

The bigger system means more outcomes are possible. But because the state of the ancilla system is known, all information extracted comes from the our system.

Can we calculate the probabilties using the state of S alone?

$$tr(\rho_A \otimes \rho_B \; \mathbb{P}_j) = \sum_{k_A, k_B} \langle k_A| \otimes \langle k_B| \left([\rho_A \otimes |\varphi\rangle\langle\varphi|] \, \mathbb{P}_j \right) |k_A\rangle \otimes |k_B\rangle$$

For simplicity, consider the basis for B that includes $|\varphi\rangle$

$$tr(\rho_A \otimes \rho_B \; \mathbb{P}_j) = \sum_{k_A} \langle k_A| \otimes \langle \varphi| \left([\rho_A \otimes |\varphi\rangle\langle\varphi|] \, \mathbb{P}_j \right) |k_A\rangle \otimes |\varphi\rangle$$

$$= \sum_{k_A} \left[\left(\langle k_A| \rho_A \right) \otimes \langle \varphi| \right] \mathbb{P}_j \left[|k_A\rangle \otimes |\varphi\rangle \right]$$

$$= tr(\rho_A E_j) \, ,$$

$$E_j = \langle \varphi | \mathbb{P}_j | \varphi \rangle = tr_M(\mathbb{P}_j |\varphi\rangle\langle\varphi|)$$

These operators clearly satisfy the following conditions

Since the probabilities must sum to 1 for any state of A

$$\sum_j E_j = \mathbb{1}$$

Since probabilities are non-negative, the operators must be positive

$$\mathrm{tr}\left(|\psi\rangle\langle\psi| E_j\right) \geq 0 \quad \forall\, |\psi\rangle \in \mathcal{H}_A$$

Any set of operators satisfying these conditions corresponds to a valid measurement, and is knows as a POVM

The more general form of measurement in QM is then

Measurements correspond to a POVM $\{E_j\}$

and yield outcome j with probability $p(j) = \mathrm{tr}(E_j \rho)$

Note that we have lost the ability to calculate final states, but we usually don't need this in situations where we need POVMs

Partial trace

The state of two systems A and B can be described by a density matrix (or wave function if pure) acting on the joint Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$

We don't care about B, we only want to make calculations regarding A. We want a density matrix that acts on $\mathcal{H}_A$ alone and gives us all information about A

So given the state of A and B: ρ_{AB}

We want to find a ρ_A such that

$$\langle O \rangle = tr(O \rho_A)$$

For some observable O acting on A

If both A and B are qubits, then for ρ_A

$$\langle O \rangle = tr(O\rho_A) = \frac{1}{2}\sum_\alpha tr(O\sigma_\alpha)\langle\sigma_\alpha\rangle \qquad \alpha \in \{0,x,y,z\}$$

The state ρ_{AB} can be written

$$\rho_{AB} = \frac{1}{4}\sum_{\alpha,\beta}\langle\sigma_\alpha\sigma_\beta\rangle\sigma_\alpha\otimes\sigma_\beta = \frac{1}{4}\sum_\alpha \sigma_\alpha \otimes\left[\sum_\beta \langle\sigma_\alpha\sigma_\beta\rangle\sigma_\beta\right] \qquad \alpha,\beta\in\{0,x,y,z\}$$

The expectation value for an observable O, acting only on A is then

$$\langle O \rangle = tr([O\otimes \mathbb{1}]\rho_{AB}) = tr\Big(\underbrace{\frac{1}{2}\sum_\alpha (O\sigma_\alpha)}_{A\ part}\otimes\underbrace{\left[\frac{1}{2}\sum_\beta\langle\sigma_\alpha\sigma_\beta\rangle\sigma_\beta\right]}_{B\ part}\Big)$$

$$= \frac{1}{2}\sum_\alpha tr(O\sigma_\alpha)\, tr\left[\frac{1}{2}\sum_\beta\langle\sigma_\alpha\sigma_\beta\rangle\sigma_\beta\right]$$

If the observables are Paulis, we find

$$\rho_A = \frac{1}{2}\sum_\alpha \sigma_\alpha\, tr\left[\frac{1}{2}\sum_\beta\langle\sigma_\alpha\sigma_\beta\rangle\sigma_\beta\right], \quad \langle\sigma_\alpha\rangle = tr\left[\frac{1}{2}\sum_\beta\langle\sigma_\alpha\sigma_\beta\rangle\sigma_\beta\right] = \langle\sigma_\alpha\sigma_0\rangle$$

The B part of ρ_{AB} is then simply replaced with a number, that comes from tracing over H_B

This gives intuition for the general case

To find the density matrix for a A, given that for AB, we need to perform a trace that only traces over B

Rather than outputting a number, like a normal trace, it results in a smaller matrix, that acts only on A

This is called the partial trace

$$\rho_A = tr_B(\rho_{AB})$$

First, what is the normal trace?

$$M = m_{00}|0\rangle\langle 0| + m_{01}|0\rangle\langle 1| + \ldots + m_{10}|1\rangle\langle 0| + m_{11}|1\rangle\langle 1| + \ldots$$

$$tr\,M = m_{00}\cancel{|0\rangle\langle 0|} + \cancel{m_{01}|0\rangle\langle 1|} + \ldots + \cancel{m_{10}|1\rangle\langle 0|} + m_{11}\cancel{|1\rangle\langle 1|} + \ldots = \sum_j m_{jj}$$

It takes a matrix and removes the bits we don't want (off diagonal elements and outer products)

Note: any orthonormal basis can be used. The result is basis independent

Partial trace can be thought of in the same way

First express the matrix in an orthonormal product basis

$$\rho_{AB} = \sum_{ijkl} C_{ijkl} |\alpha_i\rangle\langle\alpha_k| \otimes |\beta_j\rangle\langle\beta_l|$$

$$|\alpha_i\rangle \in \mathcal{H}_A, \quad |\beta_i\rangle \in \mathcal{H}_B$$

$$\langle\alpha_i|\alpha_j\rangle = \langle\beta_i|\beta_j\rangle = \delta_{ij}$$

Then remove terms than are off diagonal on subsystem B (the one we are tracing over)

$$\rightarrow \sum_{ijk\cancel{l}} C_{ijkj} |\alpha_i\rangle\langle\alpha_k| \otimes |\beta_j\rangle\langle\beta_j|$$

Then remove the outer products that act on subsystem B

$$\rightarrow \sum_{ijk} C_{ijkj} |\alpha_i\rangle\langle\alpha_k| = \rho_A$$

Example: General case of two qubits

$$
\begin{aligned}
\rho_{AB} = {} & C_{0000} |\alpha_0\rangle\langle\alpha_0| \otimes \cancel{|\beta_0\rangle\langle\beta_0|} \\
& \cancel{+ C_{0001} |\alpha_0\rangle\langle\alpha_0| \otimes |\beta_0\rangle\langle\beta_1|} \\
& \cancel{+ C_{0100} |\alpha_0\rangle\langle\alpha_0| \otimes |\beta_1\rangle\langle\beta_0|} \\
& + C_{0101} |\alpha_0\rangle\langle\alpha_0| \otimes \cancel{|\beta_1\rangle\langle\beta_1|} \\
& + C_{0100} |\alpha_0\rangle\langle\alpha_1| \otimes \cancel{|\beta_0\rangle\langle\beta_0|} \\
& \cancel{+ C_{0011} |\alpha_0\rangle\langle\alpha_1| \otimes |\beta_0\rangle\langle\beta_1|} \\
& \cancel{+ C_{0110} |\alpha_0\rangle\langle\alpha_1| \otimes |\beta_1\rangle\langle\beta_0|} \\
& + C_{0111} |\alpha_0\rangle\langle\alpha_1| \otimes \cancel{|\beta_1\rangle\langle\beta_1|} \\
& + C_{1000} |\alpha_1\rangle\langle\alpha_0| \otimes \cancel{|\beta_0\rangle\langle\beta_0|} \\
& \cancel{+ C_{1001} |\alpha_1\rangle\langle\alpha_0| \otimes |\beta_0\rangle\langle\beta_1|} \\
& \cancel{+ C_{1100} |\alpha_1\rangle\langle\alpha_0| \otimes |\beta_1\rangle\langle\beta_0|} \\
& + C_{1101} |\alpha_1\rangle\langle\alpha_0| \otimes \cancel{|\beta_1\rangle\langle\beta_1|} \\
& + C_{1010} |\alpha_1\rangle\langle\alpha_1| \otimes \cancel{|\beta_0\rangle\langle\beta_0|} \\
& \cancel{+ C_{1011} |\alpha_1\rangle\langle\alpha_1| \otimes |\beta_0\rangle\langle\beta_1|} \\
& \cancel{+ C_{1110} |\alpha_1\rangle\langle\alpha_1| \otimes |\beta_1\rangle\langle\beta_0|} \\
& + C_{1111} |\alpha_1\rangle\langle\alpha_1| \otimes \cancel{|\beta_1\rangle\langle\beta_1|}
\end{aligned}
$$

$\therefore$

$$
\begin{aligned}
\rho_A = {} & \left(C_{0000} + C_{0101}\right) |\alpha_0\rangle\langle\alpha_0| \\
& + \left(C_{0010} + C_{0111}\right) |\alpha_0\rangle\langle\alpha_1| \\
& + \left(C_{1000} + C_{1111}\right) |\alpha_1\rangle\langle\alpha_0| \\
& + \left(C_{1010} + C_{1111}\right) |\alpha_1\rangle\langle\alpha_1|
\end{aligned}
$$

More formally, the trace is defined on a Hilbert space $\mathcal{H}$ for a general matrix

$$M = \sum_{i,j} m_{ij} |i\rangle\langle j|$$

as

$$tr(M) = \sum_{i} m_{ii}$$

The partial trace over system B for a composite space $\mathcal{H}_A \otimes \mathcal{H}_B$ is defined for a general matrix

$$M = \sum_{\substack{i_A i_B \\ j_A j_B}} m_{i_A i_B j_A j_B} |i_A\rangle \otimes |i_B\rangle \langle j_A| \otimes \langle j_B|$$

as

$$tr_B(M) = \sum_{i_B} m_{i_A i_B j_A i_B} |i_A\rangle\langle j_A|$$