Quantum Information Theory

oliverobrien111

July 2021

1 Shannon entropy

Alternative definition is using the surprisal $\gamma(x) = -\log p(x)$ which measures how likely an event is. The Shannon entropy is the expected value of the surprisal. A good way of thinking about it is the information gained on average when you learn the value of X.

For a string chosen from the binary alphabet the number of distinct strings of length n is the binomial coefficient $\binom{n}{np}$ which is approximated by Strilings approximation $\log n! = n \log n - n$ to give

$$\log \binom{n}{np} \approx nH(p)$$

for the entropy function:

$$H(p) = -p \log p - (1-p) \log(1-p) \tag{1}$$

Therefore, the strings can be specified by a block code of $2^{nH(p)}$ letters. As $H(p) \leq 1$, the block code is shorter than the message. This reflects the fact that for large n the chance of an unlikely string like 1111111111... becomes very small, so can be left out of the block code. This generalises to n letters to give Shannon entropy:

$$H(X) = \sum_x -p(x) \log p(x) \tag{2}$$

1.1 Properties of H

$$H(X) = H(\{p(x)\}) = H(p_x)$$

Permutation invariant. If you have a π permutation of X then:

$$H(p_X \cdot \pi) = H(p_X) = -\sum_{x \in J} p(\pi(x)) \log p(\pi(x)) = -\sum_{x \in J} p(x) \log p(x)$$

$H(X) \geq 0$

1.2 Formal Proof

A typical sequence $\boldsymbol{u}$ is defined as one which for which the probability of occurrence satisfies:

$$2^{-n(H(X)+\epsilon} \leq p(\boldsymbol{u}) \leq 2^{-n(H(X)-\epsilon} \tag{3}$$

Typical sequence theorem states that the probability of getting any typical sequence can be made arbitrarily close to 1 for large enough n. Let $T_\epsilon^{(n)}$ be the set of typical sequences, then the probability of getting any typical sequence is bounded by:

$$2^{-n(H(X)+\epsilon}|T_\epsilon^{(n)}| \leq \sum_{\boldsymbol{u}\in T_\epsilon^{(n)}} p(\boldsymbol{u}) \leq 1$$

using the left hand inequality of 3. Therefore, $|T_\epsilon^{(n)}| \to 2^{nH(X)}$ as $\epsilon \to 0$.

Pick ϵ such that $R > H(X) + \epsilon$. Break set of possible sequences into typical set and its complement A_ϵ^n. Encode any value in A_ϵ^n to flag bit 0 and assign a codeword of length nR to elements in $T_\epsilon^{(n)}$ (which is possible as $|T_\epsilon^{(n)}| \leq 2^{n(H(X)+\epsilon} < 2^{nR}$). Probability of failure is:

$$\sum_{\boldsymbol{u}\in A_\epsilon^n} p(\boldsymbol{u}) \tag{4}$$

which by typical sequence theorem can be made arbitrarily small (as the probability of being in $T_\epsilon^{(n)}$ tends to 1).

1.2.1 Converse

Pick a subset of the typical set S^n with $|S^n| = 2^{nR}$ with $R < H$. The probability of a sequence being in S^n is:

$$P(S^n) = \sum_{\boldsymbol{u}} p(\boldsymbol{u}) = 2^{nR}2^{-nH(X)}$$

($p(\boldsymbol{u}) = 2^{-nH(X)}$ in limit as $n \to \infty$). As $H(X) - R > 0$, this tends to 0 as n tends to infinity.

Intuitively it doesn't work as the number of sequences we missed grows exponentially as n heads to infinity.

Joint Entropy:

$$H(X,Y) = -\sum_{\boldsymbol{x},\boldsymbol{y}} p(x,y)\log p(x,y) \tag{5}$$

$$H(X,Y) = H(X) + H(Y)$$

Conditional Entropy (imagine X are the bits received over a noisy channel so this is the entropy of the source Y given the data received):

$$H(Y|X) = \sum_{\boldsymbol{x}} p(x)H(Y|X=x) = \sum_{\boldsymbol{xy}} p(x)p(y|x)\log p(y|x) \tag{6}$$

After every letter is received a new optimal code can be found that will specify the remaining string with $H(Y|X)$ bits per letter. Therefore, **chain rule**

$$H(X,Y) = H(Y|X) + H(X) \tag{7}$$

Relative Entropy (defines a sort of distance between two probability distributions but is not a metric as not symmetric and does not satisfy triangle inequality):

$$D(p||q) = \sum_{\boldsymbol{x}} p(x) \frac{p(x)}{q(x)} \tag{8}$$

only 0 for $p = q$. The above is only well-defined if $p << q$ (meaning that $q(x) = 0 \Rightarrow p(x) = 0$).

Proof of Gibbs inequality

$$A = \{x \in J, p(x) > 0\}$$

$$D(p||q) = \sum_{x \in A} p(x) \log \frac{p(x)}{q(x)} = -\sum_{x \in A} p(x) \log \frac{q(x)}{p(x)}$$

Define a random variable to take values $\log \frac{p(x)}{q(x)}$ with probablity $p(x)$ then.

$$-D(p||q) = \mathbb{E}_p(\log \frac{q(X)}{p(X)})$$

using jensens inequality

$$-D(p||q) \leq \log \mathbb{E}_p(\frac{q(X)}{p(X)}) = \log \sum_A p(x) \frac{q(X)}{p(X)} = \log \sum_A q(x) \leq \log \sum_J q(x) = 0$$

so

$$D(p||q) \geq 0$$

Mutual information:

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{9}$$

If you were to fill out a venn diagram with overlapping circles $H(X)$ and $H(Y)$ the union would be $H(X,Y)$, the intersection would be $I(X,Y)$ and the remainder of the circle $H(X)$ would be $H(X|Y)$ as once you are given one of the values there is no longer any entropy coming from its circle.

Neat result. As $D(p||q) \geq 0$ and given $q(x) = \frac{1}{|J|}$ for alphabet J,

$$D(p||q) = \sum p(x) \log \frac{p(x)}{\frac{1}{|J|}} = -H(X) + \sum p(x) \log |J|$$

$$H(X) \leq \log |J|$$

Jensen's Inequality: (for concave functions)

$$\mathbb{E}(f(X)) \leq f(\mathbb{E}(X)) \tag{10}$$

Need to learn what concavity really means as used lots in quantum part of course. Important to know that $H(X)$ is concave.

Subadditivity Prove this with $D(p||q) \geq 0$ and jensens inequality. Do it on example sheet 1.

$$H(X,Y) \geq H(X) + H(Y) \tag{11}$$

5.28 of chapter 5 of Caltech I don't understand

Can use relative entropy as a parent quantity to get all the types of entropy from above. But only if we lift the restriction of the distributions having total probablity 1. For example if we take $q(x) = 1 \forall x$ then $D(P||Q) = -H(X)$. Also, $I(X:Y) = D(p(x,y)||p(x)p(y))$ and $H(Y|X) = D(p(x,y)||p(x))$. Prove and check these on example sheet.

1.3 Shannon's Noisy Channel Theorem

For certain codewords, their images after applying the channel map will represent disjoint subsets in the asymptotic limit. The typical number of sequences that will be received is $|T_n| \approx 2^{nH(Y|X)}$, whereas the size of the range is $2^{nH(Y)}$, so the maximum achievable rate (number of bits communicated per use of channel) is $\frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{nI(Y;X)}$.

If a message of length n is sent along a channel with an error rate of p. Then roughly np bits will flip, leading to $2^{nH(p)}$ typical output strings (it would be 2^{np} strings but we have to account for the encoding reducing the number of strings). In order for this input to be accurately distinguished from any other this "sphere" of possibilities must be distinct from the rest. Therefore, there must be at least $2^{nR}2^{nH(p)}$ possible output strings. Therefore, $R \leq 1 - H(p) = C$.

Can be shown that even picking random codewords gives the optimal rate in the asymptotic limit. If we adopt the decoding method of drawing a "Hamming sphere" of radius $2^{n(H(p)+\delta}$ around the received string and looking for a codeword within this radius. We would typically expect their to be at least one or our assumption about the error in the channel is wrong/we need a bigger delta. The chance of there being two can be calculated as the fraction of space occupied by the s"sphere" is:

$$\frac{2^{n(H(p)+\delta)}}{2^n} = 2^{-n(C-\delta)}$$

so the chance of one of the $2nR$ codewords lieing there is:

$$2^{-n(C-R-\delta)}$$

As δ can be as small as we like, we can pick R as close as we want to C and this will still vanish asymptotically. APPARENTLY THE AVERAGE IS TAKEN HERE BUT I DON'T SEE WHERE.

1.4 Shannon's Noisy Channel Coding Theorem - lectures

1.4.1 Discrete Memoryless Channel (DMC)

Action of each successive uses of $\mathfrak{N}$ is identical and independent to the previous use/the noise affecting each successive inputs in uncorrelated.

$$p(u^{(n)}|x^{(n)}) = \sum_{i=1}^{n} p(u_i|x_i)$$

Might as well restrict to only considering a single use of the channel. Can write channel matrix as $p_{ij} = p(y_i|x_i)$. The channel matrix is symmetric if the rows are permutations of each other.

1.4.2 Example - Memoryless Binary Symmetric Channel (m.b.s.c)

$$J_x = \{0, 1\} = J_y$$

Flips the bit with probability p. So we need an error-correcting code, e.g. the repetition code using three bits at once.

Rate: The encoding decoding pair is said to have a rate R if $|M|$ (number of possible messages) $= 2^{nR}$ for a given number of channel uses n.

Maximum probability of error corresponding to C_n:

$$p_{err}^{(n)}(C_n) = P(\mathfrak{D}_n(Y^{(n)}) \neq m|X^{(n)} = \epsilon_n(m))$$

Achievable rate: A $R \in \mathbb{R}$ is said to be an achievable rate if there exists a sequence of codes $((C_n)_n)$ of rate R s.t. $p_{err}^{(n)} \to 0$ as $n \to \infty$.
Channel Capacity: Maximum rate of reliable transmission of information $C(\mathfrak{N}) = \sup\{R : \text{R is an achievable rate}\}$

Shannon's Theorem says that $C(\mathfrak{N}) = max_{\{p(x)\}} I(X : Y)$. Not going to do this proof as it is not very connected to the quantum proof.

For m.b.s.c

$I(X : Y) = H(Y) - H(Y|X) = H(Y) - (-(1-p)\log(1-p) - p\log p) = H(Y) - h(p) \leq \log|J_Y| - h(p) = 1 - h(p)$

Does there exist some distribution $\{p(x)\}$ for which $H(Y) = 1$ because if there is then this bound is saturated.

$$H(Y) = -\sum p(y)\log p(y)$$

$$p(y) = \sum_x p(x,y) = \sum_x p(x)p(y|x)$$

So look for $p(x)$ that makes $p(y)$ equiprobable. Try $p(x) = \frac{1}{2}$, and this indeed gives $H(Y) = 1$. Therefore $C(\mathfrak{N}) = 1 - h(p)$. (I am confused about this as the H(Y) only reaches its maximum for $p = 1/2$ when $h(p)$ is also 1 so surely this doesn't work? ask in office hours).

2 Quantum

Can associate many ensembles with the same state. You can use any $\{p_i, |\psi_i\rangle\}$ as long as $p_i \geq 0$, $\sum p_i = 1$ and $\langle\psi_i|\psi_i\rangle = 1$. You can find an infinite number of these ensembles that give the same spectral decomposition (are identical) as the $|\psi_i\rangle$ need not even by orthogonal and can basically be anything as long as they have norm 1. e.g. $\rho = \frac{I}{d} = \sum_j \frac{1}{d} |e_j\rangle\langle e_j| = \sum_k \frac{1}{d} |\psi_k\rangle\langle\psi_k|$ so ensembles $\{\frac{1}{d}, |e_j\rangle\}$ and $\{\frac{1}{d}, |\psi_k\rangle\}$ are both equally valid.

Expectation of observable A in state ρ: $< A >=< A >_\rho = Tr(A\rho)$ which is a positive linear functional.

System of interest to us in the course is often a subsystem S of a composite system SE. The density matrix formalism provides a description of states of subsystems. Consider a comp. system AB then the underlying hilbert space is $\mathcal{H}_A \otimes \mathcal{H}_B$. If AB is in the state $\rho_{AB} \in \mathfrak{D}(\mathcal{H}_A \otimes \mathcal{H}_B) then state A is given by the reduce state : \rho_A = Tr_B \rho_{AB}. Consider orthonomal basis \{ |i_A\rangle\}$ in $\mathcal{H}_A$ and $\{|\alpha_B\rangle\}$ in $\mathcal{H}_B$ then we have $\{|i_A\rangle \otimes |\alpha_B\rangle\}$ in $\mathcal{H}_A \otimes \mathcal{H}_B$. Can always write $A = \sum a_{ij} |i\rangle\langle j|$ with $a_{ij} = \langle i| A |j\rangle$.

$$\rho_{AB} = \sum_{i,j=1}^{d_A} \sum_{\alpha,\beta=1}^{d_B} r_{i\alpha,j\beta} |i_A\rangle |\alpha_B\rangle \langle j_A| \langle\beta_B|$$

$$\rho_A = Tr_B \rho_A B = Tr_B(\sum_{i,j=1}^{d_A} \sum_{\alpha,\beta=1}^{d_B} r_{i\alpha,j\beta} |i_A\rangle\langle j_A| \otimes |\alpha_B\rangle\langle\beta_B|) = \sum_{i,j=1}^{d_A} \sum_{\alpha=1}^{d_B} r_{i\alpha,j\alpha} |i_A\rangle\langle j_A|$$

Consider an observable $M_{AB} = M_A \otimes I_B$ then we claim that:

$$< M_{AB} >_{\rho_{AB}} = Tr(M_{AB}\rho_{AB})$$

Proof:

$$< M_{AB} = Tr(M_{AB} \sum_{i,j=1}^{d_A} \sum_{\alpha,\beta=1}^{d_B} r_{i\alpha,j\beta} |i_A\rangle\langle j_A| \otimes |\alpha_B\rangle\langle\beta_B|) = \sum_{i,j=1}^{d_A} \sum_{\alpha=1}^{d_B} r_{i\alpha,j\alpha} Tr(M_A |i\rangle\langle j| = Tr(M_A\rho_A)$$

Example: Consider state of 2 qubits $\rho_{AB} = |phi^+_{AB}\rangle\langle\phi^+_{AB}|$ with $|phi^+_{AB}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Therefore, $\rho_A = Tr_B\rho_{AB} = \frac{I}{2}$ which is an example of when

we know everything about the system but still have no information about A or B.

A pure state is known as a product state if it can be written: $\psi_{AB} = |\psi_A\rangle \otimes |\psi_B\rangle$ and if it cannot be written like this then it is called entangled.

A mixed state is a seperable state if it can be written as an ensemble with pure states: $\rho_{AB} = \sum_i p_i \omega_A^i \otimes \tau_B^i . Otherwise it is called an entangled state.$

$The four pure states in \mathrm{C}^2 \otimes \mathbb{C}^2$ are called Bell states and EPR states.

2.1 Schmidt decomposition

Any state in space $H_A \otimes H_B$ can be described using coefficients of basis states of the form $|i_A\rangle \otimes |i_B\rangle$ where $|i_A\rangle$ for some set of basis eigenstates $|i_A\rangle$, $|i_B\rangle$. The schmidt rank is the number of positive Schmidt coefficents.

There always exists a set of orthonormal vectors $\{|i_A\rangle\}$ and $\{|i_B\rangle\}$ such that:

$$|\psi_{AB}\rangle = \sum_{i=1}^{\kappa = min(d_a, d_b)} \lambda_i |i_A\rangle |i_B\rangle$$

Proof:

$$\psi_{AB} = \sum_{r,\alpha} a_{r\alpha} |\gamma_A\rangle |\alpha_B\rangle$$

Singular Value Decompisition (SVD): Need to know proof look up in notes There exists unitaries U $d_A \times d_A$ and V $d_B \times d_B$ s.t $A = UDV$ with D a diagonal matrix $d_A \times d_B$. Therefore,

$$\alpha_{r,\alpha} = \sum_{i=1}^{d_A} \sum_{\beta=1}^{d_B} u_{\gamma i} d_{i\beta} u_{\beta\alpha}$$

$$|\psi_{AB}\rangle = \sum \sum d_{i\beta} (\sum u_{ri} |r_A\rangle)(\sum v_{\beta\alpha} |\alpha_B\rangle)$$

as $d_{i\beta} =_{i\beta} d_{ii}$

$$|\psi_{AB}\rangle = \sum_i d_{ii} (|i_A\rangle)(|i_B\rangle) = \sum_i \lambda_i (|i_A\rangle)(|i_B\rangle)$$

Now need to check these $\{|i_A\rangle\}$, $\{|i_B\rangle\}$ are an orthonormal basis. So need to check they are orthonormal and complete ($\sum_i^{min(d_A, d_B)} |i_A\rangle \langle i_B| = 1$). To check completeness consider:

$$\rho_{AB} = |\psi_{AB}\rangle \langle \psi_{AB}|$$

$$\rho_A == Tr_B |\psi_{AB}\rangle \langle \psi_{AB}| = Tr_B (\sum_i \lambda_i |i_A\rangle |i_B\rangle)(\sum_j \lambda_j |j_A\rangle |j_B\rangle) = \sum_{ij} \lambda_{ij} |i_A\rangle \langle j_A| \otimes Tr_B |j_A\rangle \langle j_B| = \sum_i \lambda_i^2$$

This means that if AB is in a pure state $|\psi_{AB}\rangle$ then ρ_A and ρ_B have identical sets of non-zero eigenvalues. If $d_A > d_B$ then ρ_A has set of eigenvalues $\{\frac{2}{i}\}^{min(d_A,d_B)}$ and the rest are zero.

Is the schmidt deompistion unique? No if the eginvalues are degenerate. As you can generate the schmidt decompistion by diagonlising ρ_A and ρ_B then matching eigenvectors of the same eigenvalues (which is non-unique if there are degenerate eigenstates.

The set of $\{\lambda_i\}$ are called Schmidt coefficents, and the set $\{|i_A\rangle\}$, $\{|i_B\rangle\}$ are the schmidt baseses, and $n(\psi_{AB})$ (the number of non-zero schmidt coefficents) is called the Schmidt rank. The schmidt rank is the simplest signature of entanglement as $|\psi_{AB}\rangle$ is entangled iff $n(\psi_{AB}) > 1$. Easy to see that $n(\psi_{AB}) = 1 \implies$ product state as then can write $\psi_{AB} = |i_A\rangle \otimes |j_B\rangle$ as only one non-zero schmidt coefficent.

2.2 Purification

It is possible to convert a mixed state into a pure state by adding a purifying reference system R with Hilbert space H_R, and defining a pure state $|\psi_{AR}\rangle \in H_A \otimes H_B$ such that:

$$\rho_A = Tr_R |\psi_{AR}\rangle \langle\psi_{AR}| = \sum_{i=1} \lambda_i^2 |i_A\rangle \langle i_A|$$

Given a ρ_A we write its spectral decompistion (diagonlise it):

$$\rho_A = \sum_{i=1}^{d_A} p_i |i_a\rangle \langle i_A|$$

Define

$$|\psi_{AR}\rangle = \sum_{i=1}^{d_A} \sqrt{p_i} |i_A\rangle |i_R\rangle$$

then it is easy to check this statifies the relation above. More generally there is a canonical way of writing a purification:

$$|\psi_{AR}\rangle = (\sqrt{\rho_A} \otimes I) |\tilde{\Omega}\rangle$$

for $|\tilde{\Omega}\rangle = \sqrt{d} |\Omega\rangle$ and $|\Omega\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i_A\rangle |i_R\rangle$. Need to check that $\rho_A = Tr_R |\psi_{AR}\rangle \langle\psi_{AR}|$ Usefulness of pure states:

$$\rho = \sum_i |i\rangle \langle i| \implies f(\rho) = \sum f(\lambda_i) |i\rangle \langle i| \implies \sqrt{\rho} = \sum \sqrt{\lambda_i} |i\rangle \langle i|$$

2.3 No-cloning theorem (N.C thm)

There does not exist a universal quantum copier. You cannot make perfect copies of arbitary unknown quantum states. You can copy some states under-some conditions.

Proof by contradiction:

Assume there exists a universal quantum copier which takes arbitary states $|\psi\rangle$ or $|\phi\rangle$ and blank slate $|s\rangle$. Assume there exists U s.t.

$$U(|\psi\rangle|s\rangle) = |\psi\rangle|\psi\rangle, U(|\phi\rangle|s\rangle) = |\phi\rangle|\phi\rangle$$

Take inner product of LHS

$$\langle\psi|\langle s|U^\dagger U|\psi\rangle|s\rangle = \langle\psi|\langle\psi|\phi\rangle|\phi\rangle$$

$$\langle\psi|\phi\rangle\langle s|s\rangle = \langle\psi|\phi\rangle^2$$

so cannot copy arbitary states can only copy orthogonal states with the same copier as must have $\langle\psi|\phi\rangle = 0$ (orthogonal) or $\langle\psi|\phi\rangle = 1$ (identical states).

Second proof by contradiction Use qubits: $\psi = \alpha|0\rangle + \beta|1\rangle$, and $|s\rangle = |0\rangle$. Therefore assume that U exists such that:

$$U|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle = \alpha^2|0\rangle|0\rangle + \alpha\beta|0\rangle|1\rangle + \alpha\beta|1\rangle|0\rangle + \beta^2|1\rangle|1\rangle$$

$$U|0\rangle|0\rangle = |0\rangle|0\rangle, U|1\rangle|0\rangle = |1\rangle|1\rangle$$

So,

$$U(|\psi\rangle|0\rangle) = U(\alpha|0\rangle + \beta|1\rangle)|0\rangle = \alpha|0\rangle|0\rangle + \beta|1\rangle|1\rangle$$

So this is a contradiction unless $\alpha = 1, \beta = 0$ or $\alpha = 0, \beta = 1$.

An implication of the no-cloning theorem is that superluminal communication is impossible.

2.4 Maximally Entangled States

$$|\psi_{AB}\rangle, \lambda_i = \frac{1}{\sqrt{d}}, d = min(d_A, d_B)$$

$$|\psi_{AB}\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^{d}|i_A\rangle|i_B\rangle$$

Its reduced states are completely mixed states if $d_A = d_B = d$:

$$\rho_A = \frac{1}{d}\sum|i_A\rangle\langle i_A|, \rho_B = \frac{1}{d}\sum|i_B\rangle\langle i_B|$$

If $d_A < d_B$ then ρ_A is a completely mixed state and ρ_B is a completely mixed state on their supports:

$$\rho_A = \frac{1}{d_A}\sum|i_A\rangle\langle i_A|, \rho_B = \frac{1}{d_A}\sum^{d_A}|i_B\rangle\langle i_B|$$

Properties
$|\psi_{AB}\rangle = \mathrm{H}_A \otimes \mathrm{H}_B \approx \mathbb{C}^d \otimes \mathbb{C}^d$ These are two qudits. Fix the orthonormal basis on $\mathbb{C}^2$ of $\{|i\rangle\}_{i=1}^d$ and let:

$$|\Omega\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle |i\rangle$$

Lemma 1: $\forall A, B \in \mathfrak{B}(\mathbb{C}^d)$ we have $\langle\Omega| A \otimes B |\Omega\rangle = Tr(A^T B)$ where T is the transposition in the chosen basis (schmidt basis of $|\Omega\rangle$

Lemma 2: Ricoche trick $(A \otimes I)|\Omega\rangle = (I \otimes A^T)|\Omega\rangle$
Proof:

$$A = \sum_{i=1}^{d} a_{jk} |j\rangle \langle k|$$

LHS of lemma 2:

$$= \frac{1}{\sqrt{d}} \sum_i [(\sum_{jk} a_{jk} |j\rangle \langle k|) \otimes I] |i\rangle \otimes = \frac{1}{\sqrt{d}} \sum_{ij} a_{ji} |j\rangle |i\rangle$$

$$A^T = \sum_{kj} |j\rangle \langle k|$$

RHS of lemma 2:

$$= \frac{1}{\sqrt{d}} \sum_{ij} |i\rangle \otimes a_{ij} |j\rangle = \frac{1}{\sqrt{d}} \sum_{ij} a_{ij} |j\rangle |i\rangle$$

Remember these lemmas we will use it over and over again.
Implications of Lemma 1
Given a state ρ_A then the canonical purification is$|\psi_{AB}\rangle = \sqrt{d}(\sqrt{\rho_A} \otimes I)|\Omega\rangle$ as:

$$\rho_A = Tr_B |\psi_{AB}\rangle \langle\psi_{AB}| = Tr_B d(\sqrt{\rho_A} \otimes I)|\Omega\rangle \langle\Omega| (\sqrt{\rho_A} \otimes I) = Tr_B \sum_{ij} \sqrt{\rho_A} |i\rangle \langle j| \sqrt{\rho_A} \otimes |i\rangle \langle j| = \sum_i \sqrt{\rho_A} |i\rangle \langle$$

Every biparitate state can be written:

$$|\psi_{AB}\rangle = (I \otimes R)|\Omega\rangle$$

for some operator R. Proof by construction:

$$|\psi_{AB}\rangle = \sum \lambda_i |i_A\rangle |i_B\rangle \in \mathrm{H}_A \otimes \mathrm{H}_B$$

Choose two isometries U, V s.t.

$$U^\dagger U = V^\dagger V = I, U|i\rangle = |i_A\rangle, V|i\rangle = |i_B\rangle, D = \sum \lambda_k |k\rangle \langle k|$$

Let $R = \sqrt{d} V D U^T$ therefore:

$$(I \otimes R)|\Omega\rangle = (I \otimes VDU^T)|\Omega\rangle = (I \otimes VD)(I \otimes U^T)|\Omega\rangle =^{\text{lemma 2}} (U \otimes VD)|\Omega\rangle$$

$$(I\otimes R)\left|\Omega\right\rangle = \sum_i (U\otimes VD)\left|i\right\rangle\otimes\left|i\right\rangle = \sum_i \left|i_A\right\rangle\otimes V\sum_k \lambda_k \left|k\right\rangle\left\langle k\right|\left|i\right\rangle = \sum_i \left|i_A\right\rangle\otimes V\lambda_i\left|i\right\rangle = \sum_i \lambda_i \left|i_A\right\rangle\otimes\left|i_B\right\rangle$$

Aside about Schmidt
$|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\xi_B\rangle \neq \sum_{i>1} \lambda_i |i_A\rangle |i_B\rangle$ as the schmidt coefficents are unique as they are the eigenvalues of ρ_A and ρ_B, so the schmidt rank is unique.

3 Time evolution of open quantum system

Dynamics is given by a quantum operation (also called a quantum channel).
Quantum Operation: $\Lambda : \mathrm{D(H)} \to \mathrm{D(K)}$ and is a completely positive, trace-preserving map (CPTP) map
- It enables a description of discrete state changes

$$\Lambda\rho_{t=0} \to \rho'_{t>0} = \Lambda(\rho)$$

Λ is a superoperator and can map from operators in general H to other operators on H: $\Lambda : \mathrm{B(H)} \to \mathrm{B(K)}$.
Properties of Λ
- Linearity $\Lambda(p_1\rho_1 + p_2\rho_2) = p_1\Lambda(\rho_1) + p_2\Lambda(\rho_2)$
- Trace preserving: $Tr\Lambda(\rho) = Tr\rho = 1$, $\rho \in \mathrm{D(H)}$ this corresponds to the conservation of probablity.
- Positivity [positive(-semidefiniteness) preserving] $\rho \geq 0 \implies \Lambda(\rho) \geq 0$
- Completely positive (CP):

$$(\Lambda \otimes id_B) : \mathrm{D(H}_A) \otimes \mathrm{B(H}_B) \to \mathrm{D(K)} \otimes \mathrm{B(H}_B)$$

is completely positive if $(\Lambda \otimes id_B)$ is positive for all B:

$$(\Lambda \otimes id_B)\rho_{AB} \geq 0$$

Further consideration of completely positiveness (CP)

$$\mathrm{H} = \mathbb{C}^m, \mathrm{K} = \mathbb{C}^n$$

use notation: M_m is set of complex $m \times m$ matrices, M_m^+ is set of positive semi definite complex $m \times m$ matrices, $\mathrm{B}(\mathbb{C}^n) = M_m$, $\mathrm{D(H)} = \{\rho \in M_m^+, Tr\rho = 1\}$.

A linear map $\Lambda : M_m \to M_n$ is positive if $\Lambda(A) \in M_n^+$ if $A \in M_m^+$.

k-positive for any $k \geq 1$ if $(\Lambda \otimes id_k)$ is positive

Is completely positive if it is k-positive for all positive integres k

3.0.1 Necessary and sufficent condtion for complete positivity

A linear map $\Lambda : \mathrm{B(H)} \to \mathrm{B(K)}$ with $\mathrm{H} = \mathbb{C}^d$ and $\mathrm{K} = \mathbb{C}^{d'}$ is completely positive iff

$$(\Lambda \otimes id_d)\left|\Omega\right\rangle\left\langle\Omega\right| \geq 0$$

Proof: If it is completely positive then it is obvious as it is a completely positive operator acting on a positive state so it will be positive. Now proof opposite direction:

Consider an arbitary $k \geq 1$ and a bipartite state $\rho \in M_d \otimes M_k$ with spectral decomposition: $\rho = \sum_{i=1}^{d} \lambda_i |\psi_i\rangle \langle\psi_i|$.

$$(\Lambda \otimes id_k)\rho \geq 0 \iff \sum \lambda_i (\Lambda \otimes id_k) |\psi_i\rangle \langle\psi_i| \geq 0 \iff (\Lambda \otimes id_k) |\psi_i\rangle \langle\psi_i| \geq 0 \forall i$$

second iff holds as $\lambda_i \geq 0$. As any bipartite state can be written $|\psi\rangle = (I \otimes R) |\Omega\rangle$. Since $|\Omega_i\rangle \in \mathbb{C}^d \otimes \mathbb{C}^k$ there exists R_i s.t. $|\psi_i\rangle = (I \otimes R_i) |\Omega\rangle$. So the above can be written as:

$$(\Lambda \otimes id_k)(I \otimes R_i) |\Omega\rangle \langle\Omega| (I_d \otimes R_i^\dagger) \geq 0$$

Note that $R_i \in \mathcal{B}(\mathbb{C}^d, \mathbb{C}^d)$. Define a superoperator $Q_i : \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^k)$ such that $Q_i(.)R_i(.)R_i^\dagger$. So:

$$(\Lambda \otimes id_k)(I \otimes R_i) |\Omega\rangle \langle\Omega| (I_d \otimes R_i^\dagger) = (\Lambda \otimes id_k)(id_d \otimes Q_i) |\Omega\rangle \langle\Omega| = (id_{d'} \otimes Q_i)(\Lambda \otimes id_d)(|\Omega\rangle \langle\Omega| \geq 0$$

$$(id_{d'} \otimes R_i)(\Lambda \otimes id_d)(|\Omega\rangle \langle\Omega| (id_{d'} \otimes R_i^\dagger) \geq 0$$

let $A = (id_{d'} \otimes R_i)$ and $B = (\Lambda \otimes id_d) |\Omega\rangle \langle\Omega|$ as

$$B \geq 0 \implies ABA^\dagger \geq 0$$

$$J(\Lambda) = (\Lambda \otimes id_d) |\Omega\rangle \langle\Omega| \geq 0 \implies (\Lambda \otimes id_d)\rho \geq 0$$

Above operator is called the Choi operator/matrix/state

4 Lecture 10

Operator theory gives us a theorem called the Strinespung's Dilation Theorem ("going to the church of the larger hilbert space"). We won't prove this but we will consider lots of equivalent theorems:
Let $\Lambda : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ denote a linear CPTP map. Then there exists a hilbert space $\mathcal{H}'$ and a unitary operator $U \in \mathcal{B}(\mathcal{H} \otimes \mathcal{H}')$ such that for all $\rho \in \mathcal{D}(\mathcal{H})$:

$$\Lambda(\rho) = Tr_{\mathcal{H}'}(U(\rho \otimes \phi)U^\dagger)$$

where ϕ is some fixed state in $\mathcal{B}(\mathcal{H}')$.
WLOG we can take $\phi = |\phi\rangle \langle\phi|$ (a pure state).

Theorem implies that any quantum operator can be composed of the following building blocks:
- Add an ancilla with Hilbert space $\mathcal{H}'$ and consider it to be in a fixed state ϕ.
- A unitary transformation of composite system: $\mathcal{H} \otimes \mathcal{H}'$
- Discard the ancilla, $Tr_{\mathcal{H}'}$

4.1 Kraus representation/ form of operator sum representation

Involves linear operators acting only on the system itself. For $\Lambda : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ there exists a finite set of linear operators $\{A_k\}_{k=1}$ such that $A_k \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ such that:

$$\Lambda(\rho) = \sum_k A_k \rho A_k^\dagger, \sum_k A_k^\dagger A_k = I_{\mathcal{H}}$$

RHS is implied by LHS and the fact Λ is trace preserving:

$$1 = Tr\rho = Tr\Lambda(\rho) = Tr(\sum_k A_k \rho A_k^\dagger) = \sum_k Tr(A_k \rho A_k^\dagger) = \sum_k Tr(A_k A_k^\dagger \rho) = Tr(\sum_k A_k A_k^\dagger \rho) = Tr\rho \forall \rho$$

Kraus form is just restatement of Stinesprings Thm

$$\Lambda(\rho) = Tr_{\mathcal{H}'}(U(\rho \otimes \phi)U^\dagger) = \sum_k \langle e_k| U(\rho \otimes \phi)U^\dagger |e_k\rangle = \sum \langle e_k| U(\rho \otimes |\phi\rangle \langle\phi|)U^\dagger |e_k\rangle = \sum_k A_k \rho A_k^\dagger$$

with $A_k = \langle e_k| U |\phi\rangle \in \mathcal{B}(\mathcal{H}', \mathcal{H})$

$$U = \sum_{\alpha\beta ij} U_{\alpha i, \beta j} |f_\alpha\rangle |e_i\rangle \langle f_\beta| \langle e_j|, |\phi\rangle = \sum c_l |e_l\rangle$$

$$A_k = \langle e_k| U |\phi\rangle \langle e_k| \sum_{\alpha\beta ij} |f_\alpha\rangle |e_i\rangle \langle f_\beta| \langle e_j| \sum c_l |e_l\rangle = \sum_{\alpha\beta l} c_l U_{\alpha k, \beta l} |f_\alpha\rangle \langle f_\beta| \in \mathcal{B}(\mathcal{H})$$

Above assume that $\mathcal{H} = \mathcal{K}$.

Prove that any map of Kraus form is a linear CPTP map.

$$\Lambda(\rho) = \sum_k A_k \rho A_k^\dagger$$

Linearity is checked easily on $\mathcal{B}(\mathcal{H})$.
Show completely positive by considering $(\Lambda \otimes id_d) |\Omega\rangle \langle\Omega| \geq 0$
Need to show that $\langle\psi| (\Lambda \otimes id_d) |\Omega\rangle \langle\Omega| |\psi\rangle \geq 0 \forall |\psi\rangle \in \mathbb{C}^{d'} \otimes \mathbb{C}^d$. Assume has kraus form and show that this is therefore satisfied:

$$\langle\psi| (\Lambda \otimes id_d) |\Omega\rangle \langle\Omega| = \langle\psi| (A_k I) |\Omega\rangle \langle\Omega| (A_k^\dagger \otimes I) |\psi\rangle$$

Define $|\phi_k\rangle = (A_k^\dagger \otimes I) |\psi\rangle$.

$$= \sum_k \langle\psi_k| \Omega |\psi_k\rangle \geq 0$$

The above is true as $\Omega = |\Omega\rangle \langle\Omega|$ is a positive operator.

All properties of Λ can be transcribed into the properties of $J(\Lambda)$. There exists an isomorphism between linear maps and operators. This reduces the study

of linear maps to the study of linear operators. THis is the Choi-Jamilkowski (C-J) isomorphism. Very important part of the lectures is learning the following proof - will be in lecture notes.

Choi-Jamilkowski (C-J) isomorphism
The follwing equations provide a 1-1 correspondance between linear maps $\Lambda : M_d \to M_{d'}$ and operators $J \in \mathcal{B}(\mathbb{C}^{d'} \otimes \mathbb{C}^d)$.

$$(a) : J = (\Lambda \otimes id_d) |\Omega\rangle \langle\Omega|, (b) : Tr(A\Lambda(B)) = dTr(J(A \otimes B^T))$$

there are more correpsondences than listed here and THESE CORRESPONDANCES WILL BE IN THE EXAM
The maps $\Lambda \to J$ and $J \to \Lambda$ defined by the above are mutual inverses and lead to the following correspondances:
1. Λ is completely positive $\iff J(\Lambda) \geq 0$
2. Λ is trace preserving $\iff Tr_{\mathbb{C}^{d'}} J = \frac{I_d}{d}$
Proof: (first prove from LHS to RHS of the above set of equations, then we will prove 2)
Adjoint: Adjoint of a linear map $\Lambda : M_d \to M_{d'}$ is $\Lambda^* : M_{d'} \to M_d$ defined through:

$$Tr(A\Lambda(B)) = Tr(\Lambda^*(A)B)$$

for $B \in M_d, A \in M_{d'}$

5 Example Class 1

Definition of complex inner product:

$$(A, B) = (\bar{B, A}), (A, \lambda B + C) = \lambda(A, B) + (A, C), (A, A) \geq 0$$

Definition of vector space:
Closed under $A + B$, and closed under λA for $\lambda \in F$ with F either $\mathbb{R}$ or $\mathbb{C}$.

Useful trick is to always write the entropy with the maximal number of variables so you can just multiply together the logarithms and then express it all in terms of a relative entropy.

$$\sum_{xy} p(x, y) \log p(x) + \sum_{x,y} p(x, y) \log p(x, y)$$

6 Lecture 10

Going to show that C-J isomorphism proves Stinesprungs Dilation Thereom and Kraus Representation. Proof of $(a) \implies (b)$:

$$dTr(J(A \otimes B^T)) = dTr(((\Lambda \otimes id_d) |\Omega\rangle \langle\Omega|)(A \otimes B^T)) = dTr(|\Omega\rangle \langle\Omega| (\Lambda^* \otimes I)(A \otimes B^T))$$

$dTr(J(A\otimes B^T)) = dTr(|\Omega\rangle\langle\Omega|(\Lambda^*(A)\otimes B^T)) = dTr((\Lambda^*(A)\otimes B^T)|\Omega\rangle\langle\Omega|) = dTr((\Lambda^*(A)\otimes I)(I\otimes B^T)|\Omega\rangle\langle\Omega$

Use ricochet:

$$(I\otimes B^T)|\Omega\rangle = (B\otimes I)|\Omega\rangle$$

so

$$dTr(J(A\otimes B^T)) = dTr((\Lambda^*(A)B\otimes I)|\Omega\rangle\langle\Omega|) = Tr(\sum_{ij}(\Lambda^*(A)B|i\rangle\langle j|\otimes|i\rangle\langle j|))$$

$$dTr(J(A\otimes B^T)) = Tr\sum_i(\Lambda^*(A)B)|i\rangle\langle i| = Tr(\Lambda^*(A)B) = Tr(A\Lambda(B))$$

Important exercise: Prove that $\Lambda(\rho) = Tr_{\mathbb{C}^d}J(I\otimes\rho^T)$.

In order to prove that this correspondance is a bijection we need to prove injective and surjective. To every Λ there is a unique J. We have just proved injective as for every Λ there is a unique J found using (a). Now need to prove that eeach J is the image of at least one :

Take $J\in\mathcal{B}(\mathbb{C}^{d'}\otimes\mathbb{C}^d$, consider orthonormal basis $\{|\psi_i\rangle\}$ of $\mathbb{C}^{d'}\otimes\mathbb{C}^d$. Then expand J in basis:

$$J = \sum\alpha_{ij}|\psi_i\rangle\langle\psi_j|$$

As $|\psi_i\rangle$ are bipartite states:

$$|\psi_i\rangle = (R_i\otimes I)|\Omega\rangle, R_i\in\mathcal{B}(\mathbb{C}^d,\mathbb{C}^{d'})$$

(we are always trying to bring in the MES

$$J = \sum\alpha_{ij}(R_i\otimes I)|\Omega\rangle\langle\Omega|(R_j^\dagger\otimes I)$$

Define Λ to be the map which acts on any $A\in M_d$ as follows:

$$\Lambda(A) = \sum\alpha_{ij}R_iAR_j^\dagger$$

Obviously Λ maps $M_d\to M_{d'}$. Consider the action of this Λ:

$$(\Lambda\otimes I)|\Omega\rangle\langle\Omega| = \sum\alpha_{ij}(R_i\otimes I)|\Omega\rangle\langle\Omega|(R_j^\dagger\otimes I)$$

This is clearly J so for every J there is a Λ that maps to it by construction. Check $\langle\psi_i||\psi_j\rangle = \delta_{ij} \implies R_iR_j^\dagger = \delta_{ij}I$. Now need to prove correspondences, Λ is Trace Preserving $\iff Tr_{\mathbb{C}^{d'}}J = \frac{I_d}{d}$.

First prove Lemma 1: Λ is Trace Preserving $\iff \Lambda^*(I_{d'}) = I_d$:

$$Tr(\Lambda(A)) = Tr(I_{d'}\Lambda(A)) = Tr(\Lambda^*(I_{d'}A) = Tr(A) \implies \Lambda^*(I_{d'} = I_d$$

inverse:

$$Tr\Lambda(A) = Tr(\Lambda^*(I_{d'})A) = Tr(I_dA) = Tr(A)$$

Prove correspondence using Lemma:

$$Tr_{\mathbb{C}^{d'}} J = T_{\mathbb{C}^{d'}} (\Lambda \otimes I) |\Omega\rangle \langle\Omega|$$

Write out the above in terms of i, j and take the trace over the first element but remember we are assuming it is trace preserving so $Tr\Lambda |i\rangle \langle j| = Tr |i\rangle \langle j| = \delta_{ij}$. Need Lemma and the C-J $TrA\Lambda(B) = dTr(J(A \otimes B^T))$ to prove the converse:

$$Tr\Lambda(B) = dTr_{\mathbb{C}^{d'}} Tr_{\mathbb{C}^d}(J(I \otimes B^T)) = dTr_{\mathbb{C}^d} Tr_{\mathbb{C}^{d'}}(J(I_{d'} \otimes B^T))$$

As

$$Tr_A(M_{AB}(I_A \otimes N_B)) = Tr_A(M_{AB})N_B$$

so

$$Tr\Lambda(B) = dTr_{\mathbb{C}^d}((T_{\mathbb{C}^{d'}} J)B^T) = Tr_{\mathbb{C}^d}(I_d B^T) = Tr(B^T) = TrB$$

so Λ is trace preserving if $Tr_{\mathbb{C}^{d'}} J = \frac{I_d}{d}$.

7 Lecture 12

Prove $\Lambda(\rho) = dTr_{\mathbb{C}^d}(J(I \otimes \rho^T)) = dTr_{\mathbb{C}^d}((\Lambda I) |\Omega\rangle \langle\Omega| (I \otimes \rho^T)) = Tr \sum_{ij} (\Lambda(|i\rangle j) \otimes |i\rangle braj\rho^T = \sum_{ij} \rho_{ij} \Lambda(|i\rangle \langle j|) = \Lambda(\rho)$.

Important property is unital. Iff Λ is trace preserving then $\Lambda^*(I_{d'}) = I_d$.

Proof that Λ is a linear CPTP map implies Kraus representation. We know that $J(\Lambda) \geq 0$ and $TrJ(\Lambda) = 1$ which implies that $J(\Lambda)$ is a density matrix. So you can associate to it an ensemble of states $\{p_i, |\psi_i\rangle\}$. Therefore can write $J(\Lambda) = \sum_i p_i |\psi_i\rangle \langle\psi_i|$. There must exist R_i s.t. $|\psi_i\rangle = (R_i \otimes I) |\Omega\rangle$. Subsitute this into defintion of $J(\Lambda)$:

$$J(\Lambda) = \sum p_i (R_i \otimes I) |\Omega\rangle \langle\Omega| (R_i^\dagger \otimes I)$$

Define $\mathrm{A}_i = \sqrt{p_i} R_i$ so

$$J(\Lambda) = \sum (A_i \otimes I) |\Omega\rangle \langle\Omega| (A_i^\dagger$$

Therefore, $\Lambda(\rho) = \sum_i A_i \rho A_i^\dagger$. Check that this ensemble has identity property:

$$Tr\Lambda(\rho) = Tr \sum A_i \rho A_i^\dagger = Tr(\sum A_i A_i^\dagger)\rho \implies \sum A_i A_i^\dagger = I$$

This tell us how many kraus operators there are as we get it from the ensemble needed to represent $J(\Lambda)$. Therefore, the number of kraus operators must be at least the rank of $J(\Lambda)$. If you consider $\{p_i, |\psi_i\rangle\}$ s.t that $J(\Lambda) = \sum p_i |\psi_i\rangle \langle\psi_i|$ is the spectral decomposition (that $|\psi_i\rangle$ are orthgonal) then this inequality holds with equality (number of kraus operators = rank $J(\Lambda)$).

Lastly we want to connect Kraus to Stinesprings Dilation Thereom. Consider $\Lambda : \mathcal{B}(\mathcal{H}_\mathcal{A}) \to \mathcal{B}(\mathcal{H}_\mathcal{B})$, and $U \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_E, \mathcal{H}_B \otimes \mathcal{E})$ with $|\phi_E\rangle besome fixed state in \mathrm{H}_E$, and consider an orthonormal basis $\{|k\rangle\}_{k=1}^r$ in $\mathcal{H}_E$ (with $\mathcal{H}_E$ selected to have the same dimensions as the kraus operators. Define U through the following relation:

$$U(|\psi_A\rangle \otimes |\psi_E\rangle) = \sum_{k=1}^{r} A_k |\psi_A\rangle \otimes |k\rangle$$

CHECK that U defined through this is an isometry (preserves inner products).

Example of a map that is positive but not completely positive. Example of this is Transposition.

$$T : \rho \to \rho^T$$

must be positive as the eigenvalues are unchanged so $T \geq 0$. Now consider:

$$(T \otimes I) \ngeq 0$$

Whenever looking for counter examples it often comes from entanglement like here. Consdier $\rho = |\phi\rangle \langle\phi|$ for $\phi = \frac{1}{\sqrt{d}} \sum_{k=1}^{d} |i\rangle |i\rangle$.

$$\tilde{\rho} = (T \otimes I)\rho = \frac{1}{d} \sum_{ij} (T \otimes I)(|i\rangle \langle j| \otimes |i\rangle \langle j| = \frac{1}{d} \sum_{ij} |j\rangle \langle j| \otimes |i\rangle \langle j|$$

Therefore this is just the swap operator so $\tilde{\rho} = \frac{F}{d}$ so $\tilde{r}ho^2 = I$ so eigenvalues are ± 1 so $\tilde{\rho}$ has a negative eigenvalue so the transposition operator is not completely positive.

7.1 Distance between quantum states

Given 2 states $\rho, \sigma \in D(\mathcal{H})$, how well can you distinguish between them? So we need a concept of distance/distinguishability between states.

7.1.1 Trace distance

$D(\rho, \sigma) = \frac{1}{2}||\rho - \sigma||_1$ with $||A||_1 = Tr|A|$ with $|A| = \sqrt{A^\dagger A}$.

Therefore here $A = \rho - \sigma$ and has eigenvalue decomposition $A = \sum a_i |\psi_i\rangle \langle\psi_i| = \sum_{a_i \geq 0} a_i |\psi + i\rangle \langle\psi_i| + \sum_{a_i < 0} a_i |\psi + i\rangle \langle\psi_i|$. Let $Q = \sum_{a_i \geq 0} a_i |\psi + i\rangle \langle\psi_i|$ and $R = \sum_{a_i < 0} (-a_i) |\psi + i\rangle \langle\psi_i|$ so $A = Q - R$. We have Q and R have mutually orthonogonal supports. Consider distance:

$$D(\rho, \sigma) = Tr|A| = \sum |a_i| = \sum_{a_i \geq 0} a_i + \sum_{a_i < 0} (-a_i) = \frac{1}{2} TrQ + \frac{1}{2} TrR$$

We know that $Tr(Q) = Tr(R)$ as $A = \rho - \sigma = Q - R$ but $Tr(A) = 0$ as ρ and σ are both density matrices so have trace 1. so:

$$D(\rho, \sigma) = TrQ$$

Lemma 1: $D(\rho,\sigma) = max_P TrP(\rho-\sigma)$ for $P^\dagger = P$ and $P^2 = P$ and $0 \geq P \geq I$. Proof: Choose P projection on to support of Q.

$$TrP(\rho-\sigma) = TrP(Q-R) = Tr(PQ) = Tr(PQP) = Tr(Q)$$

But $D(\rho,\sigma) = \frac{1}{2}TrQ + \frac{1}{2}TrR = TrQ = TrR$. Conversely, for all projection operator P:

$$Tr(P(\rho-\sigma)) = Tr(P(Q-R)) \leq Tr(PQ) \leq Tr(Q) = D(\rho,\sigma)$$

first inequality as $TrPR = TrP^{1/2}RP^{1/2} \geq 0$ as $R \geq 0$ and second inequality is because $TrPQ = TrQ^{1/2}PQ^{1/2} \leq TrQ^{1/2}Q^{1/2} = Tr(Q)$. So this gives $D(\rho,\sigma) \geq max_P Tr(P(\rho-\sigma))$ and as we have showed this is achieved $D(\rho,\sigma) = max_P Tr(P(\rho-\sigma))$.

7.1.2 Properties of $D(\rho,\sigma)$

It is a metric as: 1. Symmetric $D(\rho,\sigma) = D(\sigma,\rho)$
2. $D(\rho,\sigma) \geq 0 = 0 \iff \rho =$
3. Triangle inequality holds

2. If $\rho = \sigma$ then $D(\rho,\sigma) = 0$. Conversly ff $D(\rho,\sigma) = 0$then as $D(\rho,\sigma) = TrQ = TrR \implies a_i = 0 \forall i$ with $A = \rho - \sigma$ so must have $\rho_i - \sigma_i = 0$ so $\rho = \sigma$.
3. $D(\rho,\sigma) \leq D(\rho,\omega) + D(\omega,\sigma)$. Proof: From lemma 1 we must have a projection: $D(\rho,\sigma) = Tr(P(\rho-\sigma)) = Tr(P(\rho-\omega))+Tr(P(\omega-\sigma)) \leq D(\rho,\omega)+D(\omega,\sigma)$ with the last inequality because D is the maximisation of the trace quantity over P.

Lemma 2: Monotonicity under quantum operations (Data processing inequality)

$$\Lambda : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$$

then

$$D(\Lambda(\rho),\Lambda(\sigma)) \leq D(\rho,\sigma)$$

Proof: Let P be a projection such that $D(\Lambda(\rho),\Lambda(\sigma)) = Tr(P(\Lambda(\rho)-\Lambda(\sigma)))$. We saw that $Tr(Q) = Tr(R)$ so $Tr(\Lambda(Q)) = Tr(\Lambda(R))$ as trace preserving. Prove that $D(\rho,\sigma) = \frac{1}{2}TrQ + \frac{1}{2}TrR = \frac{1}{2}Tr(\Lambda(Q)) + \frac{1}{2}Tr((R)) = Tr(\Lambda(Q)) = Tr(\Lambda(Q)^{1/2}I\Lambda(Q)^{1/2}) \geq Tr(P\Lambda(Q)) \geq Tr(P\Lambda(Q) - P\Lambda(R)) = Tr(P(\Lambda(Q)) - (R)) = D(\Lambda(\rho) - \Lambda(\sigma))$.

7.2 Fidelity

$$F(\rho,\sigma) = Tr\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}||\sqrt{\rho}\sqrt{\sigma}||_1$$

Want to prove that $F(\rho,\sigma) = F(\sigma,\rho)$ and $0 \leq F(\rho,\sigma) \leq 1$.

If $\rho,\sigma] = 0$ then they have common eigenbasis $\{|i\rangle\}$ so can write $\rho = \sum \lambda_i |e_i\rangle \langle e_i|$

and $\sigma = \sum \mu_i |e_i\rangle \langle e_i|$ so $F(\rho,\sigma) = Tr(\sum \sqrt{\lambda_i}\sqrt{\mu_i}|e_i\rangle\langle e_i|) = \sum \sqrt{\lambda_i}\sqrt{\mu_i} = F_u(,\mu)$ with F_u the classical fidelity of classical probability distributions.

If $\rho = |\psi\rangle\langle\psi|$ and $\sqrt{\rho} = |\psi\rangle\langle\psi|$ then

$$F(\rho,\sigma) = F(|\psi\rangle\langle\psi|,\sigma) = Tr(\sqrt{\langle\psi|\sigma|\psi\rangle}|\psi\rangle bra\psi) = \sqrt{\langle\psi|\sigma|\psi\rangle}$$

If $\rho = |\phi\rangle\langle\phi|$ and $\sigma = |\psi\rangle\langle\psi|$ then:

$$F(\rho,\sigma) = |\langle\phi|\psi\rangle|$$

Also can be shown that

$$F(U\rho U^\dagger, U\sigma U^\dagger) = F(\rho,\sigma)$$

7.2.1 Uhlmann's Theorem

Let ρ_A, $\sigma_A \in \mathcal{D}(\mathcal{H}_A)$ then

$$F(\rho_A,\sigma_A) = max_{|\psi_{AR}\rangle,|\sigma_{AR}\rangle}|\langle\psi_{AR}|\psi_{AR}\rangle|$$

Lemma 3: $\forall A \in \mathcal{B}(\mathcal{H})$ such that $||A||_1 = Tr|A| = max_{U\text{unitaries}}|Tr(UA)|$:

$$Tr|A| \geq |Tr(UA)| \forall U \text{ unitary}$$

Note that all purifications are equivalent up to unitaries on the reference system.
Proof of Uhlmann's thereom: Choose cannonical purification:

$$|\psi^\rho_{AR}\rangle = \sqrt{d}(\sqrt{\rho_A}\otimes I_R)|\Omega\rangle = (\sqrt{\rho_A}\otimes I_R)|\tilde{\Omega}\rangle$$

$$|\psi^\sigma_{AR}\rangle = \sqrt{d}(\sqrt{\sigma_A}\otimes I_R)|\Omega\rangle = (\sqrt{\sigma_A}\otimes I_R)|\tilde{\Omega}\rangle$$

So Uhlmann's theorem can be written as:

$$F(\rho,\sigma) = max_{U^\rho_R,U^\sigma_R}|\langle\psi^\rho_{AR}|(I\otimes U^\dagger_R)(I_A\otimes U^\sigma_R)|\psi^\sigma_{AR}\rangle| = max_{U^\rho_R,U^\sigma_R}|\langle\psi^\rho_{AR}|(I\otimes U^\dagger_R U^\sigma_R)|\psi^\sigma_{AR}\rangle|$$

Consider $\langle\psi^\rho_{AR}|(I\otimes U_R)|\psi^\sigma_{AR}\rangle$ for $U_R = U^\dagger_R U^\sigma_R$:

$$|\langle\psi^\rho_{AR}|(I\otimes U_R)|\psi^\sigma_{AR}\rangle| = |\langle\tilde{\Omega}|(\sqrt{\rho_A}\otimes I)(I\otimes U_R)(\sqrt{\sigma_A}\otimes I_R)|\tilde{\Omega}\rangle| = |\langle\tilde{\Omega}|(\sqrt{\rho_A}\sqrt{\sigma_A}\otimes I)(I\otimes U_R)|\tilde{\Omega}\rangle|$$

$$|\langle\psi^\rho_{AR}|(I\otimes U_R)|\psi^\sigma_{AR}\rangle| = |\langle\tilde{\Omega}|(\sqrt{\rho_A}\sqrt{\sigma_A}U^T_A\otimes I)|\tilde{\Omega}\rangle| = \sum_i \langle i|\sqrt{\rho_A}\sqrt{\sigma_A}U^T|i\rangle = Tr|\sqrt{\rho_A}\sqrt{\sigma_A}U^T|$$

Use lemma to show that

$$|\langle\psi^\rho_{AR}|(I\otimes U_R)|\psi^\sigma_{AR}\rangle| = Tr|\sqrt{\rho_A}\sqrt{\sigma_A}U^T| \leq Tr|\sqrt{\rho_A}\sqrt{\sigma_A}| = F(\rho_A,\sigma_A)$$

So: $F(\rho_A,\sigma_A) \geq max_{|\psi_{AR}\rangle,|\sigma_{AR}\rangle}|\langle\psi_{AR}|\psi_{AR}\rangle|$. Consider $X = \sqrt{\rho_A}\sqrt{\sigma_A}$ has polar decomposition $X = V|X|$, if we choose $U = V^\dagger$ then we get that equality holds.

7.2.2 Implications of Uhlmann's Thereom

1) $F(\rho,\sigma) = F(\sigma,\rho)$ as $|\langle\psi^\rho|\,|\psi^\sigma\rangle| = |\langle\psi^\sigma|\,|\psi^\rho\rangle$.
2) $0 \leq F(\rho,\sigma) \leq 1$. Proof: $F(\rho,\sigma) = 1 \iff \rho = \sigma$ as if $\rho = \sigma$ then can choose purification s.t. $|\langle\psi^\rho|\,|\psi^\sigma\rangle| = 1$. If $F(\rho,\sigma) = 1$ then there exists a purification start s.t. $|\langle\psi^\rho|\,|\psi^\sigma\rangle| = 1$ so $|\psi^\rho\rangle = |\psi^\sigma\rangle \implies \rho = \sigma$.

Lemma 4: Monotoncity under partial trace

Start with bipartite states: ρ_{AB} and $\sigma_{AB} \in \mathcal{D}(\mathcal{H}_A, \mathcal{H}_B)$. Then:

$$F(\rho_{AB}, \sigma_{AB}) \leq F(\rho_A, \sigma_A)$$

Proof: Follows from Uhlmanns theorem as it implies there must exist purifications: $|\psi^\rho_{ABC}\rangle$ and $|\psi^\sigma_{ABC}\rangle$ s.t.

$$F(\rho_{AB}, \sigma_{AB}) = |\langle\psi^\rho_{ABC}|\,|\psi^\sigma_{ABC}\rangle|$$

These states are also purfications of ρ_A and σ_A as $\rho_A = Tr_{BC}\,\langle\psi^\rho_{ABC}|\,|\psi^\rho_{ABC}\rangle$. So by Uhlmann's thereom:

$$F(\rho_A, \sigma_A) = max_{|\psi_\rho\rangle|\psi_\sigma\rangle}|\langle\psi_\rho|\,|\psi_\sigma\rangle| \geq |\langle\psi^\rho_{ABC}|\,|\psi^\sigma_{ABC}\rangle| = F(\rho_{AB}, \sigma_{AB})$$

8 Measurement Postulate

8.1 Von Neumann/ Projective Measurement Postulate

System is in a state $\rho \in \mathcal{D}(\mathcal{H})$. Measure an observable $A \in \mathcal{B}(\mathcal{H})$, the outcome is an eigenvalue of A (given by $A\,|\phi_A\rangle = a\,|\phi_A\rangle$). So can generally write the spectral decomposition as: $A = \sum_a aP_a$. Where P_a is an orthongonal projector onto eigenspace corresponding to eigenvalue a. The probability of the outcome being a is given by $Tr(P_a\rho)$ If the outcome is a then $\rho \to \rho' = \frac{P_a\rho P_a}{Tr(P_a\rho)}$.

Need to generalise this postulate as:
1) Suppose you have a spin-1/2 particle preprared in state $|\psi\rangle$ s.t. $\boldsymbol{\sigma}\cdot\boldsymbol{n}\,|\psi\rangle = |\psi\rangle$. E.g. if $\boldsymbol{n} = (0,0,1)$ then this would reduce to $\sigma_z\,|\psi\rangle = |\psi\rangle$ so $|\psi\rangle = 0$. Can we do a measurement to find the direction of $\boldsymbol{n}$? There is no projective measurement we can use to determine this measurement as the direction $\boldsymbol{n}$ is not an observable (an $\mathbb{R}^3$ is not a self-adjoint operator), however it is a valid question so we want a different measurement postulate that we can use to solve this.
ii) If we want to do a measurement where we only care about the result of the measurment on $A < B$ but this is not possible to write purely as projection operators.

Intution: All that matters ito determine the measurement are two things:
1) Probability of an outcome $p(a) = Tr(P_a\rho)$

2) Post-measuremnt state $\rho \to \rho' = \frac{P_a \rho P_a}{Tr(P_a \rho)}$
All that you require of P_a is that:

$$p(a) \geq 0 \implies Tr(P_a \rho) \geq 0 \implies P_a \geq 0$$

$$1 = \sum_a p(a) = Tr(\sum_a P_a \rho) \implies \sum_a P_a = 1$$

Because we stipulate that we are measuring an observable A we get an additional condition:

$$P_a P_b = \delta_{ab} P_b$$

It is this last condition that is relaxed to get the generalised measurement postulate.

8.1.1 Generalised Measurement Postulate

System is in a state $|\psi\rangle \in \mathcal{H}$ and a complete quantum measurement is described by a finite collection of operators $\{M_a\}, M_a \in \mathcal{B}(\mathcal{H})$.
The subscript a labels the possible outcomes.
They satisfy a completeness relation:$\sum_a M_a^\dagger M_a = I$
If initial state of the system is $|\psi\rangle$ then the probability of getting a:

$$p(a) = |\psi\rangle M_a^\dagger M_a |\psi\rangle$$

If a is the outcome then the state $|\psi\rangle$ collapses to:

$$|\psi'\rangle = \frac{M_a |\psi\rangle}{\sqrt{\langle\psi| M_a^\dagger M_a |\psi\rangle}}$$

Completeness relation gives us:

$$\sum_a p(a) = \langle\psi| \sum_a M_a^\dagger M_a |\psi\rangle = \langle\psi| |\psi\rangle = 1$$

Sanity check:
Projective measurement is a special case of this generalised measurement. Just pick M_a to be projectors we get back $M_a^\dagger M_a = P_a^2 = P_a$.

8.1.2 POVM

Very often we don't care about the post-measurement state, so there is a parrticular formulism form this called POVM (positive operator valued measure doesn't really mean anything) which simiplifies the generalised measurement postulate.

Define $E_a = M_a^\dagger M_a \geq 0$ (obvious apparently?).

$$\sum_a E_a = \sum_a M_a^\dagger M_a = I \text{ (completeness relation)}$$

$\{E_a\}$ gives a positive definite partition of unity (defined as $\sum_a E_a = I$). Look at measurement and consider the probability of the outcome then:

$$p(a) = Tr(\rho E_a) \geq 0$$

$$\sum_a p(a) = Tr(\rho \sum_a E_a) = 1$$

But this is not sufficent to infer the post-measurement state. Could choose $M_a = \sqrt{E_a}$ but generally not given M_a just E_a. So POVM can be defined independantly it means the collection of operators $\{E_a\}$ doesn't need any reference to the generalised measurement. These E_a are called POVM elements.

POVM is compeltely defined by a set of $\{E_a\}$ s.t. $E_a \geq 0$ and $\sum_a E_a = I$ and the probability of outcome a is $p(a) = Tr(\rho E_a), \sum_a p(a) = 1$.

Interesting takeway from last lecture:

$$TrQ = Tr((P + P^\perp)Q) = Tr(PQ) + Tr(P^\perp Q) \geq Tr(PQ)$$

9 Lecture 15

Pure POVM: $\{E_i\}$ is pure if E_i are all rank 1 projections. This means there must exist some $|\phi_i\rangle$ s.t. $E_a = |\phi_i\rangle \langle\phi_i|$.

9.1 Implementing/realizing a generalised measurement

This is very similar to the stinespring realisatiion ("church of the higher hilbert space")
Let the system to be measured have a hilbert spce $\mathcal{H}_A$
Measurement is described by $\{M_a\}$
1. Add an anncilla B of hilbert space $\mathcal{H}_B$ s.t. $dim\mathcal{H}_B =$ number of measurement operators $\{M_a\}$. So can make a one to one correspondance from an orthonormal basis of B to the measurement operators $\{|e_a\rangle\} \to \{M_a\}$
2. Consider B initially in a pure state $|\phi\rangle$ which is uncorrelated with A. So initial state A:$|\psi\rangle$, B :$|\phi\rangle$ so AB: $|\psi\rangle \otimes |\phi\rangle$.
3. Unitary transformation/evolution U of AB. U acts on AB defined through:

$$|\Psi\rangle = U(|\psi\rangle \otimes |\phi\rangle) = \sum_a M_a |\psi\rangle \otimes e_a$$

(In the stinespring/Kraus situtaiton we had $U(|\psi\rangle \otimes |\phi\rangle = \sum A_k |\psi\rangle \otimes |k\rangle$.)
Define $|\tilde{\Psi}\rangle = U(|\tilde{\psi}\rangle \otimes |\phi\rangle) = \sum M_a |\tilde{\psi}\rangle \otimes |e_a\rangle$. Therefore:

$$\langle\Psi| |\tilde{\Psi}\rangle = \langle\psi| |\tilde{\psi}\rangle$$

So this preserves inner products and is therefore a unitary transformation. Here we had some weird stuff about unitaries on subspaces there is an exercise it is worht doing. minute 15 lecture 15.
4. Make a projective measurement $\{P_a\} = I_a \otimes |e_a\rangle \langle e_a|$ of $|\Psi\rangle_A B$. Obviously, $P_a^2 = P_a$, $P_a = P_a^\dagger$ and $P_a P_b = \delta_{ab} P_a$. Probability of 'a' $= p(a) = \langle\Psi| P_a |\Psi\rangle$. If you substitute in the values of this and using the fact $\langle e_a| |e_b\rangle = \delta_{ab}$ and $\sum M_a^\dagger M_a = I$ this reduces to $p(a) = \langle\psi| M_a^\dagger M_a |\psi\rangle$.
Consider post-measurement state. If outcome is 'a' then: $|\Psi\rangle = U(|\psi\rangle \otimes |\phi\rangle) \to |\Psi'\rangle = P_a |\Psi\rangle$. Therefore, $|\Psi'\rangle = (I_A \otimes |e_a\rangle \langle e_a|) \sum_{a'} M_{a'} |\psi\rangle \otimes |e_a\rangle = M_a |\psi\rangle \otimes |e_a\rangle$. So If we take the partial trace over B to give the postmeasurement state of A:

$$Tr_B |\Psi'\rangle \langle\Psi'| = M_a |\psi\rangle \langle\psi| M_a^\dagger$$

Go back to example $\boldsymbol{\sigma} \cdot \boldsymbol{n} |\psi\rangle = |\psi\rangle$ where $\boldsymbol{n}$ is one of a finite set of linearly dependant vectors s.t. $\sum \lambda_a n_a = 0$ with $\lambda + a \geq 0, \sum \lambda_a = 1$.
Let $E_a = \lambda_a (I + \boldsymbol{n} \cdot \sigma)$. Claim $E_a = 2\lambda_a P_{\boldsymbol{n}_a}, P_{\boldsymbol{n}_a} = |\uparrow_{n_a}\rangle \langle\uparrow_{n_a}|$ is a valid POVM (check this as exercise).
Case 1: $\boldsymbol{n} \in \{\boldsymbol{n}_1, \boldsymbol{n}_2\}, \sum \lambda_a n_a = 0$ so $\boldsymbol{n}_1 = -\boldsymbol{n}_2$ (another excercise very good to do) check that $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ become orthogonal projectors (POVM -¿ projective measure). $E_a = 2\lambda P_{n_1}$
Case 2: $\boldsymbol{n} \in \{\boldsymbol{n}_1, \boldsymbol{n}_2, \boldsymbol{n}_3\}, \lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{3}$ then check that $E_a = \frac{2}{3} P_{\boldsymbol{n}_a}$.

9.2 Drop in session

Richoet relation between two different spaces:

$$A : \mathbb{C}^d \to '$$

$$d(I_d \otimes A) |\Omega_d\rangle = d'(A^T I_{d'}) |\Omega_{d'}\rangle$$

Defintion of fidelity:

$$F(\rho, \sigma) = Tr(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}})$$

$$BAB^\dagger \geq 0 \iff A \geq 0$$

9.3 Entanglement

Bipartite entanglement: Pure state either product state like below or it is entangled

$$|\psi_{AB}\rangle = |\phi_A\rangle \otimes |\xi_B\rangle$$

Densitiy state is either separable like below or it is entangled

$$\rho_{AB} = \sum p_u \omega_i^A \otimes \nu_i^B$$

Bell States

$$|\phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), |\psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$$

Claim they can be characterised by 2 bits:
1. Parity bit (0 if spins are parallel, 1 if spins are antiparallel)
2. Phase bit (0 if superposition has + sign, 1 if superposition has - sign)
e.g. $10 \to |\psi^{+}\rangle$. How to recover this information. Can only be done with a joint measurement so need bits in same location. The projection operators:

$$P_{00} = |\phi^{+}\rangle \langle\phi^{+}|, P_{11} = |\psi^{-}\rangle \langle\psi^{-}|, ...$$

What if they are in different locations??

We have:

$$(\sigma_z^{(A)} \otimes I) : |\phi^{\pm}\rangle \to |\phi^{\mp}\rangle$$
$$(\sigma_x^{(A)} \otimes I) : |\phi^{\pm}\rangle \to |\psi^{\pm}\rangle$$

When you act locally with σ_x or σ_z you just change one bell state into another and with both local operations and classical communication you can only determine the parity bit or the phase bit but not both.

10 Lecture 15

Look at provided example handout to see who you can convert from one Bell state to another bipartite state that is not maximially entangled. Easy way to check if a state is maximally entangeled to check if $Tr_A |\psi_{AB}\rangle \langle\psi_{AB}| = \frac{I}{2}$.

10.1 Marjorization

Let $\boldsymbol{x} = (x_1, .., x_n) \in \mathbb{R}^n$ and $\boldsymbol{y} = (y_1, ..., y_n) \in \mathbb{R}^n$ and take $\boldsymbol{x}$, $\boldsymbol{y}$ and reorder into decreasing order:

$$\boldsymbol{x}^{\downarrow} = (x_1^{\downarrow}, ..., x_n^{\downarrow}), x_1^{\downarrow} \geq x_2^{\downarrow}$$
$$\boldsymbol{y}^{\downarrow} = (y_1^{\downarrow}, ..., y_n^{\downarrow}), y_1^{\downarrow} \geq y_2^{\downarrow}$$

$\boldsymbol{x}$ is majorised by $\boldsymbol{y}$ ($\boldsymbol{x} \prec \boldsymbol{y}$) if:

$$\sum_{n=1}^{k} x_l^{\downarrow} \leq \sum_{i=1}^{k} y_i^{\downarrow} \forall k = 1, 2, .., n$$

with equality holding when $k = n$. Lots of extra material in online lecture notes that is not examinable just the lecture content.

This concept of majorization can be extended to density matrices $\rho, \sigma \in D(\mathcal{H})$. As $\rho \prec \sigma$ if $\lambda(\rho) \prec \lambda(\sigma)$ where $(\rho), \lambda(\sigma)$ are vectors of e-values of ρ and σ.

10.1.1 Nielsen's Majorization Thereom

A bipartite pure state $|\psi\rangle$ shared between ALice and Bob can be transformed to $|\phi\rangle$ by LOCC $\iff \lambda_\psi \prec \lambda_\phi$. This implies that entanglement cannot be increased or created by LOCC. Equivalent to saying the schmidt rank cannot be increased by LOCC. Claim that for schmidt ranks $_\phi$ and n_ψ:

$$\lambda_\psi \prec \lambda_\phi \implies n_\phi \not> n_\psi$$

Proof by contradiction. Assume $\lambda_\psi \prec \lambda_\phi, n_\phi > n_\psi$. $\lambda_\phi = (\mu_1, ..., \mu_m, 0, 0..0)$ has a greater number of non-zero entries than $\lambda_\psi = (\nu_1, ..., \nu_j, 0, 0, 0)$. So there exists an integer m s.t. $\mu_n \neq 0$ but $\nu_m = 0$. Therefore, $\sum_{i=1}^{m-1} = 1$ but $\sum_1^{m-1} \mu_i \neq 1$ so this violates $\sum_{i=1}^k \nu_i \leq \sum_{i=1}^k \mu_i$ for $k = m - 1$.

10.2 Superdense coding

Alice has 2 classical bits and wants to send them to Bob, but no classical communication is possible and can only send one qubit. This is possible if Alice and Bob share a maximally entangled state to start with. She acts with nothing, σ_z, $_x$ and $\sigma_z\sigma_x$ to encode 00,01,10 and 11. This transforms the bell state into one of the four bell states. She then sends her qubit to bob and he performs a bell measurment. If Eve intercepts the transfered qubit she won't be able to get any information.

10.3 Quantum Teleportation

Alice has a single unknown qubit state $|\psi_C\rangle = \alpha|0\rangle + \beta 1$ but no quantum communication is possible. This is possible if Alice and Bob share a maximally entangled state $|\phi^+\rangle$. Initial state is therefore $|\psi_C\rangle \otimes |\phi^+_{AB}\rangle = \frac{1}{2}|\phi^+_{CA}\rangle \otimes |\psi_B\rangle + \frac{1}{2}|\phi^-_{CA}\rangle \otimes \sigma_z|\psi_B\rangle + \frac{1}{2}|\psi^+_{CA}\rangle \otimes \sigma_X|\psi_B\rangle + \frac{1}{2}|\psi^-_{CA}\rangle \otimes (-i\sigma_y)|\psi_B\rangle$. Alice measures CA in the bell measurement and then communicates which postmeasurement state the system is in to bob. He can then perform the inverse of the σ_i operations to extract the $|\psi_B\rangle$ state.

11 Examples class 2

If $|\psi_{AR}\rangle$ and $|\phi_{AR}\rangle$ are two purifications of a state ρ of a composite system AR then there exists a unitary transformation U_R that acts s.t.:

$$|\phi_{AR}\rangle = (I_A \otimes U_R)|\psi_{AR}\rangle$$

Trace does not commute everything it is cyclic so $Tr(AB) = Tr(BA)$ but $Tr(ABC) \neq Tr(BAC)$

$$A \geq 0 \iff BAB^\dagger \geq 0 \forall B$$

as

$$\langle\psi| BAB^\dagger |\psi\rangle = \langle\phi| A |\phi\rangle$$

Remember

$$\sum \Pi_i = \boldsymbol{I}, \Pi_i = \Pi_i^\dagger, \Pi_i\Pi_j = \delta_{ij}\Pi_i$$

You can use ricochet on block diagonal matrices when each block maps a degenerate subspace with non maximally entangled states. As for each block it is mapping a maximally entangeled state as the coefficents are all equal. Below we show this by separating out the eigenspaces and on each space we get the form of maximally entangeld ($\sum |i\rangle |i\rangle$) whereas before we had $\sum \lambda_i |i\rangle |i\rangle$ so it wasn't maxially entangled. If you take the trace of maximally entangled space on one of its spaces you get the identity

$$U_A = _{\lambda \in spec(\rho)} U_A^\lambda$$

$$(U_A \otimes I) |\psi_{AR}\rangle = \sum \lambda \sum_{i|\lambda_i=\lambda} U_A^{(\lambda)} |\tilde{i}_A\rangle \otimes |i_R\rangle = \sum \lambda \sum_{i|\lambda_i=\lambda} |\tilde{i}_A\rangle \otimes (U_A^{(\lambda)})^T |i_R\rangle = (I \otimes U_A^T) |\psi_{AR}\rangle$$

with U_A^λ mapping the span of λ in $\{|\tilde{i}_a\rangle\}$ to the span of in $\{|i_a\rangle\}$.

General expression for density matrix from the defining properties: $\rho \geq 0 \implies \rho = \rho^\dagger, Tr\rho = 1$

$$\rho = \begin{pmatrix} a & b - ic \\ b + ic & 1 - a \end{pmatrix}$$

12 Lecture 17

12.1 Quantum Entropies

Given a state $\rho \in \mathcal{D}(\mathcal{H}$ the von Neumann entropy is:.

$$S(\rho) = Tr(\rho \log \rho) = -\sum_i \lambda_i \log \lambda_i = H(\{\lambda_i\})$$

For spectral decomposition $\rho = \sum \lambda_1 |\psi_i\rangle \langle\psi_i|$ we have $\log \rho = \sum_i (\log \lambda_i) |\psi_i\rangle \langle\psi_i|$.
Properties
1) $S(\rho) \geq 0$ and $S(\rho) = 0 \iff \rho$ is a pure state. If ρ pure then $\lambda_i = 1$ for some i and $\lambda_j = 0$ for the rest so from defintion $S(\rho) = 0$. Other direction if $S(\rho) = 0$ the only way the RHS can be 0 is if only one of the λ_i is 1.
2) $S(U\rho U^\dagger) = S(\rho)$ for all unitary U. As it just depends on the eigenvalues which a unitary operation won't change.
3) Concavity $\{p_i\}$ probability distribution $\{\rho_i\} setof states. S(\sum_i p_i\rho_i) \geq p_i S(\rho_i) 4) \mathrm{S}(\rho) \leq \log d$ and equality only holds if $\rho = \frac{I}{d}$

12.1.1 Quantum Relative Entropy

For $\rho \in \mathcal{D}(\mathcal{H}), \sigma \in \mathcal{B}(\mathcal{H}), \sigma \geq 0$.

$$D(\rho||\sigma) = \begin{cases} Tr(\rho \log \rho) - Tr(\rho \log \sigma) & supp\, \rho < supp\, \sigma \\ \infty & otherwise \end{cases} \tag{12}$$

1) $D(\rho||\sigma) \geq 0$ with equality iff $\rho = \sigma$ (Klein's inequality) this is sort of a "distance measure" but not really.
Proof: Let $\rho = \sum \lambda_i \left|i\right\rangle \left\langle i\right|, \sigma = \sum_\alpha \mu_\alpha \left|\alpha\right\rangle \left\langle\alpha\right|$

$$D(\rho||\sigma) = \sum \lambda_i \log \lambda_i - Tr(\sum \lambda_i \left|i\right\rangle \left\langle i\right| \sum_\alpha (\log \mu_\alpha) \left|\alpha\right\rangle \left\langle\alpha\right| = \sum \lambda_i \log \lambda_i - \sum \lambda_i \log \mu_\alpha |\left\langle i\right| \left|\alpha\right\rangle|^2$$

$$p_{i\alpha} = |\left\langle i\right| \left|\alpha\right\rangle|^2 \geq 0, \sum_i p_{i\alpha} = \sum_i \left\langle\alpha\right| \left|i\right\rangle \left\langle i\right| \left|\alpha\right\rangle = 1, \sum_\alpha p_{i\alpha} = 1$$

So $p_{i\alpha}$ are eleemnts of a doubly stochastic matric so creates probability distributions: $\{p_{i\alpha}\}_i$ and $\{p_{i\alpha}\}_\alpha$. So $\log x$ is concave:

$$\log(\sum_\alpha p_{i\alpha}\mu_\alpha) \geq \sum_\alpha p_{i\alpha} \log \mu_\alpha$$

So:

$$D(\rho||\sigma) = \sum \lambda_i \log \lambda_i - \sum_i \lambda_i \sum_\alpha (\log \mu_\alpha) p_{i\alpha} \geq \sum \lambda_i (\log \lambda_i - \log(\sum_\alpha p_{i\alpha}\mu_\alpha))$$

Consider $r_i = \sum_\alpha p_{i\alpha}\mu_\alpha \geq 0$ which has $\sum_i ri = \sum_{alpha}(\sum_i \rho_{i\alpha})\mu_\alpha - 1$ so r_i is a probability distribution so:

$$D(\rho||\sigma) \geq \sum \lambda_i \log \frac{\lambda_i}{r_i} = D_{KL}(\{\lambda_i\}||\{r_i\}) \geq 0$$

Use this to prove property 4 of von Neuman entropy
For $\sigma = \frac{I}{d}$:

$$0 \leq D(\rho||\sigma) = -S(\rho) - Tr(\rho \log(\frac{I}{d})) = -S(\rho) - Tr(\log(\frac{1}{d})\rho I) = -S(\rho) + \log(d) \implies S(\rho) \leq \log d$$

12.1.2 Data processing inequality

$$D(\Lambda(\rho)||\Lambda(\sigma)) \leq D(\rho||\sigma)$$

Proof uses joint convexity: for $\{p_i\}$, $\{\rho_i\}$ and $\{\sigma\}$ these are jointyl convex if $D(\sum_i p_i\rho_i||\sum_i p_i\sigma_i) \leq \sum_i p_i D(\rho_i||\sigma_i)$. Proof on example sheet.
Also uses: $D(\rho \otimes \omega||\sigma \otimes \nu) = D(\rho||\sigma) + D(\omega||\nu)$ and $= D(\rho||\sigma)$ if $\omega = \nu$
Also uses: $D(U\rho U^\dagger||U\sigma U^\dagger) = D(\rho||\sigma)$ for all unitaries U.

Take two operators $X, Z \in \mathcal{B}(\mathcal{H})$ defined for $k, m. \in \{0, 1, ..., d-1\}$ with:

$$X^k \left|j\right\rangle = \left|j \oplus k\right\rangle, Z^m \left|k\right\rangle = e^{2\pi i m j} d \left|j\right\rangle$$

e.g. for $d = 2$ this would just give the pauli matrices $X = \sigma_x$ and $Z = \sigma_z$.
Define $W_{k,m} = X^k Z^m \in \mathcal{B}(\mathbb{C}^d)$ which are unitary and called the Heisenberg-Weyl operators.

Finally uses $\forall A \in \mathcal{B}(\mathcal{C}^d)$ andd $\frac{1}{d^2}\sum_{k,m} W_{km} A W_{km}^\dagger = Tr(A)\frac{I}{d} \implies \frac{1}{d^2}\sum(I \otimes W_{km})X(I \otimes W_{km}^\dagger = Tr_2 X \otimes \tau$ with $\tau = \frac{I}{d}$.

Proof of Data Processing Inequality:

$$D(\Lambda(\rho)||\Lambda(\sigma)) = D(\Lambda(\rho) \otimes \tau||\Lambda(\sigma) \otimes tau)$$

Using Stinesprings Dilation Theorem: $\Lambda(\rho) = Tr_2 U(\rho \otimes \phi)U^\dagger = Tr_2 X$:

$$\Lambda(\rho) \otimes \tau = Tr_2 X \otimes \tau = \frac{1}{d^2}\sum_{km}(I \otimes W_{km} U(\rho \otimes \phi)U^\dagger(I \otimes W_{km}^\dagger$$

similarly:

$$\Lambda(\sigma) \otimes \tau = Tr_2 X \otimes \tau = \frac{1}{d^2}\sum_{km}(I \otimes W_{km} U(\sigma \otimes \phi)U^\dagger(I \otimes W_{km}^\dagger$$

Define $V_{km} = (I \otimes W_{km})U$ so:

$$D(\Lambda(\rho)||\Lambda(\sigma)) = D(\frac{1}{d^2}\sum_{km} V_{km}(\rho \otimes \phi)V_{km}^\dagger||\frac{1}{d^2}\sum_{km} V_{km}(\sigma \otimes \phi)V_{km}^\dagger)$$

Use joint convexity:

$$D(\Lambda(\rho)||\Lambda(\sigma)) \leq \frac{1}{d^2}\sum_{km} D(V_{km}(\rho \otimes \phi)V_{km}^\dagger||V_{km}(\sigma \otimes \phi)V_{km}^\dagger) Invarianceunderunitary:$$

$$D(\Lambda(\rho)||\Lambda(\sigma)) \leq \frac{1}{d^2}\sum_{km} D(\rho \otimes \phi||\sigma \otimes \phi) = D(\rho||\sigma) + D(\phi||\phi) = D(\rho||\sigma)$$