Quantum Information Theory (Preskill ch. 5)

A very brief summary

Key Elements of Information Theory.

Classical Information Theory is mostly about two things.

(1) How much redundancy is present in a typical message?
Shannon showed that an n-letter message composed of letters a_x drawn from an alphabet ($a_1, a_2, a_3,\ldots a_k$), with a priori probability of occurrence $p(a_x)$. can be compressed to $n\mathrm{H}(p)$, where $\mathrm{H}(p)$ is the Shannon Entropy.

(2) How much redundancy must be added to a message in order to communicate reliably over a noisy channel. Shannon's Noisy Channel Coding Theorem tells us that the code rate R must be less than or equal to the Channel Capacity

$$R \leq 1 - H(\mathrm{p}) = C(p)$$

(3) There are other entropic measures such as Joint Entropy, Conditional Entropy (page 8), Mutual Information (page 9), etc. All of these ideas and concepts translate to Quantum Information.

Quantum Information Theory is largely based on further developments of the ideas and concepts above.

Key Elements of Quantum Information Theory.

(1) A quantum message consists of letters drawn from a quantum alphabet, $\{\rho_x, p(\mathrm{x})\}$. The message is thus of the form $\rho = \Sigma_x\, p(\mathrm{x})\rho_x$
Note that the message is a tensor product $\rho = \rho_1 \otimes \rho_2 \otimes \rho_3 \otimes \ldots \otimes \rho_k$, where the individual letters can be pure or mixed states.

We define the Von Neumann Entropy $S(\rho) = -Tr(\rho \mathrm{Log}(\rho))$. It can be shown that $S(\rho)$ is the number of incompressible quantum bits per letter in the message.

In the eigenbasis of ρ, the letters can be expressed as $\rho = \Sigma_\lambda\, \rho\, \lambda|\lambda\rangle\langle\lambda|$. Also,

$$S(\rho) = \mathrm{Tr}\ (\ \Sigma_\lambda\, \lambda \log(\lambda)\ |\lambda\rangle\langle\lambda|) = \mathrm{H}(\Lambda) \qquad \text{(page 11).}$$

Thus, if the alphabet consists of mutually orthogonal pure states then the quantum source reduces to a classical source, i.e, $S(\rho) = \mathrm{H}(\Lambda)$, and most everything we have learned about Shannon Information carries over.

(2) The Von Neuman entropy has a number of information theoretical properties, the first 3 of which are straightforward, while the remaining 9 listed by Preskill are increasingly obscure when working one's way down the list. (page 13).

Quantum Data Compression

(1) This is the quantum analogy to Shannons Noiseless Coding Theorem. Preskill offers an example of quantum data compression, but needs to do an awfull lot of work to achieve a very modest gain. Thus, we should probably conclude that this is a proof of principle, but not necessarily something that we might do to improve quantum data storage or tyhroughput. To summarize:

(2) Preskills first step is to use an alphabet of non-orthogonal letters, $|\uparrow_Z\rangle$, $|\uparrow_X\rangle$, each occurring with probability $p = 1/2$. The letters are then of the form

$$\rho\ = 1/2\ |\uparrow_z\rangle\langle\downarrow_z| + 1/2\ |\uparrow_x\rangle\langle\downarrow_x|$$

(3) Looking at the message in the diagonal basis, $|0'\rangle$ and $|1'\rangle$ (page 160), it is clear that both letters overlap strongly with $|0'\rangle$ and weakly with $|1'\rangle$. In this situation it is possible for Alice to significantly compress her message by projecting onto Likely and unlikely Subspaces (page 17). The non-obvious choices of letters, $|\uparrow_Z\rangle$, $|\uparrow_X\rangle$, and thus the diagonal representation $|0'\rangle$, $|1'\rangle$ for the compressed state, is chosen to ensure that "compression" onto the likely and unlikely subspaces is likely to succeed (page 20/21). Preskill adds a few more bells and whistles to recover as much as possible of the information projected onto the unlikely subspace, page (21/22). When all is said and done, the fidelity of the transmitted state is only improved by ~2% relative to the simple strategy of sending two bits and having Bob guess the third.

Quantum Information Theory (Preskill ch. 5)

Main Topics of CIT:

(1) Transmission of classical info over quantum channels

(2) Information/disturbance tradeoff in QM

(3) Quantifying entanglement

(3) Transmission of quantum info over quantum channels

Our Program: (1) & (4) ~ 2 Lectures

Key Concept – Incompressible information content

Classical Measure:	Shannon Entropy
Quantum Measure:	von Neumann Entropy

Review of Classical Information Theory

(Shannon for Dummies, Preskill 5.1)

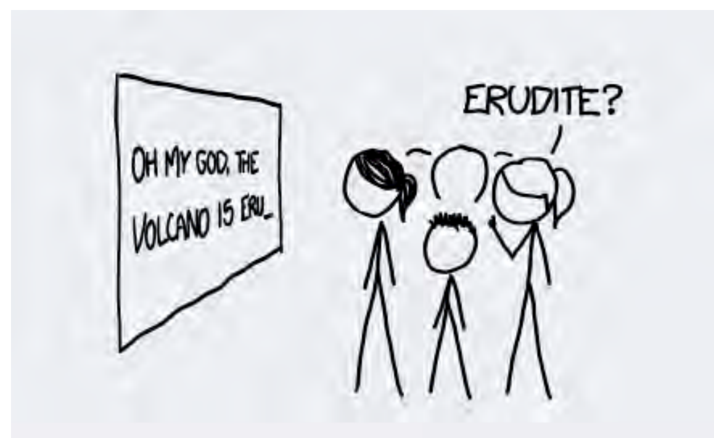
Shannon, 1948: Core findings of classical info theory

(1) How much data can be compressed (redundancy)

(2) Reliable communication rate over noisy channel (Redundancy needed to protect against errors)

Shannon Entropy and **Data Compression**

(Shannon's noiseless coding theorem)



Quantum Information Theory (Preskill ch. 5)

Main Topics of CIT:

(1) Transmission of classical info over quantum channels

(2) Information/disturbance tradeoff in QM

(3) Quantifying entanglement

(3) Transmission of quantum info over quantum channels

Our Program: (1) & (4) ~ 2 Lectures

Key Concept – Incompressible information content

Classical Measure:	Shannon Entropy
Quantum Measure:	von Neumann Entropy

Review of Classical Information Theory

(Shannon for Dummies, Preskill 5.1)

Shannon, 1948: Core findings of classical info theory

(1) How much data can be compressed (redundancy)

(2) Reliable communication rate over noisy channel (Redundancy needed to protect against errors)

Shannon Entropy and **Data Compression**

(Shannon's noiseless coding theorem)

Message = String of letters chosen from $\{a_1, a_2, \ldots, a_k\}$

A priori probability of occurrence: $p(a_x)$, $\sum p(a_x) = 1$

Basic Question: given message w/ $n \gg 1$ letters

Can we compress to length $< n$?

Quantum Information Theory (Preskill ch. 5)

But: $p(x_1 \ldots x_n)$ is just one of many typical strings with the same number of occurrences of each letter and thus identical a priori probabilities $p(\text{typical})$. Then for n large enough, we also have

$$1-\varepsilon \leq \underbrace{\sum p(\text{typical})}_{= N(\varepsilon,\delta) \times p(\text{typical})} \leq 1 \quad (2)$$

$N(\varepsilon,\delta)$: # of typical strings

Taking the ratio $\frac{1}{(1)} \times (2)$ gives us the final result

$$\boxed{(1-\varepsilon)\, 2^{n(H-\delta)} \leq N(\varepsilon,\delta) \leq 2^{n(H+\delta)}}$$

Conclusion (Shannons Noiseless Coding theorem)

* We can encode all typical strings w/blocks of $n(H+\delta)$ bits
* Atypical strings occur w/prob. $< \varepsilon$, where $\varepsilon, \delta \to 0$ for $n \to \infty$
* An optimal code thus compresses each letter to $H(X)$ bits

Quantum Information Theory (Preskill ch. 5)

Joint and Conditional Entropy, Mutual Information

Consider the following scenario:

Alice —— noisy channel ——→ **Bob**

$X = \{x, p(x)\}$ errors specified by $p(y|x)$ $Y = \{y, p(y)\}$

known about channel

Bayes Rule: $$p(x|y) = \frac{p(y|x)\,p(x)}{p(y)}$$

known about Alices alphabet

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

Bob uses this to estimate the prob. that Alice sent x given he received y. The "width" of the distribution $p(x|y)$ is thus a measure of Bob's information gain per letter.

Think about this in terms of joint events

$$\{X,Y\} = \{(x,y),\ p(x,y)\}$$

- **Joint entropy**

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

This is a measure of information content per letter in the combined strings

* Assume Bob measures the value of a letter y in the message

* He gets $H(Y)$ bits of info about the letter pair x, y

* Bob's remaining uncertainty about the letter x is then tied to his lack of knowledge about x given that he knows y.

Quantum Information Theory (Preskill ch. 5)

Think about this in terms of joint events

$$\{X,Y\} = \{(x,y),\ p(x,y)\}$$

- **Joint entropy**

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

This is a measure of information content per letter in the combined strings

* Assume Bob measures the value of a letter y in the message
* He gets $H(Y)$ bits of info about the letter pair x,y
* Bob's remaining uncertainty about the letter x is then tied to his lack of knowledge about x given that he knows y.

The entropy of X conditioned on Y is therefore

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

$$\Rightarrow H(X|Y) = H(X,Y) - H(Y)$$

The Conditional Entropy $H(X|Y)$

is the number of bits of info per letter in Alice's message that Bob is missing due to channel errors

– measure of information loss due to errors –

Equivalently, it is the # of extra bits Alice must send to ensure Bob gets the complete message in the presence of channel errors.

Quantum Information Theory (Preskill ch. 5)

The entropy of X conditioned on Y is therefore

$$H(X,Y) \equiv H(Y) + H(X|Y)$$

$$\Rightarrow H(X|Y) \equiv H(X,Y) - H(Y)$$

The **Conditional Entropy** $H(X|Y)$

is the number of bits of info per letter in Alice's message that Bob is missing due to channel errors

– measure of information loss due to errors –

Equivalently, it is the # of extra bits Alice must send to ensure Bob gets the complete message in the presence of channel errors.

Note: From the above,

$$H(X|Y) = H(X,Y) - H(Y)$$

$$= -\sum_{(x,y)} p(x,y) \, \text{Log}\, p(x|y) = \sum_{(x,y)} -p(x,y) \, \text{Log} \frac{p(x,y)}{p(y)}$$

$$= -\sum_{(x,y)} p(x,y) \, \text{Log}(x,y) + \sum_{y} p(y) \, \text{Log}\, p(y)$$

Note: For this derivation we use Bayes rule

$$p(x|y) = \frac{p(y|x)\, p(x)}{p(y)}$$

where $p(y|x)$ is a known property of the channel

and $p(x)$ is a known property of the alphabet X.

Logs are averaged over letter pairs x,y or letters y

Quantum Information Theory (Preskill ch. 5)

The entropy of X conditioned on Y is therefore as

$$H(X,Y) = H(Y) + H(X|Y)$$
$$\Rightarrow H(X|Y) = H(X,Y) - H(Y)$$

The **Conditional Entropy** $H(X|Y)$

is the number of bits of info per letter in Alice's message that Bob is missing due to channel errors

– measure of information loss due to errors –

Equivalently, it is the # of extra bits Alice must send to ensure Bob gets the complete message in the presence of channel errors.

Note: From the above,

$$H(X|Y) = H(X,Y) - H(Y)$$
$$= -\sum_{(x,y)} p(x,y) \operatorname{Log} p(x|y) = \sum_{(x,y)} -p(x,y) \operatorname{Log} \frac{p(x,y)}{p(y)}$$
$$= -\sum_{(x,y)} p(x,y) \operatorname{Log}(x,y) + \sum_{y} p(y) \operatorname{Log} p(y)$$

We can similarly quantify the # of bits of info about X that Bob has gained by measuring Y .

This is the Mutual Information:

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$
$$= H(X) - H(X|Y) = H(Y) - H(Y|X)$$

Note: When we added the info content of X to the info content of Y we overcounted the total info because some info is common to X and Y, and must be subtracted to get the proper measure for the Mutual Information

Quantum Information Theory (Preskill ch. 5)

Shannon's Noisy Channel Coding Theorem

noisy channel

Alice $\longrightarrow$ Bob

$X = \{x, p(x)\}$ errors specified by $p(y|x)$ $Y = \{y, p(y)\}$

Alice & Bob need redundancy to communicate reliably over a noisy channel. How much?

Key Question: Can we always find a reliable code when the message length $n \to \infty$?

Binary alphabet: $\{0, 1\}$

p(no flip) $= 1 - p$

p(flip) $= p$

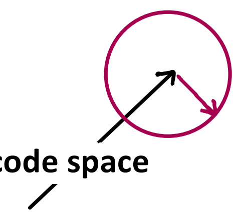
$$\Rightarrow \begin{cases} p(0|0) = p(1|1) = 1-p \\ p(0|1) = p(1|0) = p \end{cases}$$

Basic Idea: Encode k bits of info in block of size n

We define the Code Rate $R = k/n$

Optimal Code: max # of bits must flip to interchange code words

n bits $\rightarrow$ np flips



diffuses into error bubble of $2^{nH(p)}$ words

words in code space

Hamming radius np

Reliable Decoding: Error bubbles must not overlap

$$\Rightarrow \underbrace{2^k \, 2^{nH(p)}}_{\text{\# of words required}} \leq 2^n \leftarrow \text{total \# of words in code space}$$

Setting $k = nR$ and solving for the Code Rate we get

$$\boxed{R \leq 1 - H(p) \equiv C(p)} \leftarrow \text{Channel Capacity}$$

Quantum Information Theory (Preskill ch. 5)

Quantum Information Theory:

Key Results from Classical Information Theory

* Message of letters drawn from ensemble $\{x, p(x)\}$

  Shannon Info $H(X) =$ # of incompressible bits per letter in limit $n \to \infty$

* Correlation between sent (X) and received (Y) messages

  Mutual Info
  $$I(X,Y) = H(X) - H(X|Y)$$
  $$= H(Y) - H(Y|X)$$

  # of bits of info about (X) learned from (Y)

Quantum Information Theory

⇨ Need to generalize these concepts

Basic Scenario:

Alice sends letters drawn from the ensemble

$$\{\rho_x, p_x\} \Rightarrow \rho = \sum_x p_x \rho_x$$

Bob reads message by measuring the POVM

$$\{F_a\} \Rightarrow P(a) = \mathrm{Tr}(F_a \rho)$$

We define the von Neumann Entropy

$$\boxed{S(\rho) = -\mathrm{Tr}(\rho \log \rho)}$$

In the eigenbasis of ρ we have

$$\rho = \sum_\lambda \lambda |\lambda\rangle\langle\lambda| \Rightarrow$$
$$S(\rho) = \mathrm{Tr} \sum_\lambda \lambda \log \lambda |\lambda\rangle\langle\lambda| = H(\Lambda)$$

Shannon Entropy of the ensemble $\Lambda = \{|\lambda\rangle, \lambda\}$

Quantum Information Theory (Preskill ch. 5)

Basic Scenario:

Alice sends letters drawn from the ensemble

$$\{\rho_x, p_x\} \Rightarrow \rho = \sum_x p_x \rho_x$$

Bob reads message by measuring the POVM

$$\{F_a\} \Rightarrow P(a) = \mathrm{Tr}(F_a \rho)$$

We define the von Neumann Entropy

$$\boxed{S(\rho) = -\mathrm{Tr}(\rho \log \rho)}$$

In the eigenbasis of ρ we have

$$\rho = \sum_\lambda \lambda |\lambda\rangle\langle\lambda| \Rightarrow$$

$$S(\rho) = \mathrm{Tr} \sum_\lambda \lambda \log \lambda |\lambda\rangle\langle\lambda| = H(\Lambda)$$

Shannon Entropy of the ensemble $\Lambda = \{|\lambda\rangle, \lambda\}$

Conclusion:

* If the alphabet consists of mutually orthogonal pure states then the quantum source reduces to a classical source
* In that case all signal states are perfectly distinguishable
* $S(\rho) = H(\Lambda)$

We can show the Von Neumann entropy quantifies

* The incompressible information content of a quantum source
* The quantum information content per quantum letter
* The classical information content per quantum letter (extractable by POVM)
* Entanglement of a bipartite pure state

Quantum Information Theory (Preskill ch. 5)

Conclusion:

* If the alphabet consists of mutually orthogonal pure states then the quantum source reduces to a classical source
* In that case all signal states are perfectly distinguishable
* $S(\rho) = H(\Lambda)$

We can show the Von Neumann entropy quantifies

* The incompressible information content of a quantum source
* The quantum information content per quantum letter
* The classical information content per quantum letter (extractable by POVM)
* Entanglement of a bipartite pure state

Mathematical properties of $S(\rho)$ (No proofs, see Preskill)

(1) **Purity** $\rho = |\psi\rangle\langle\psi| \rightarrow S(\rho) = 0$

(2) **Invariance** $S(U\rho U^{+}) = S(\rho)$

(3) **Maximum** ρ has d eigenvalues $\neq 0 \rightarrow S(\rho) \leq \log d$

(4) **Concavity** For $\lambda_1, \lambda_2, \ldots \lambda_n \geq 0$, $\sum_i \lambda_i = 1 \rightarrow$

$$S(\lambda_1\rho_1 + \ldots + \lambda_n\rho_n) \geq \lambda_1 S(\rho_1) + \ldots \lambda_n S(\rho_n)$$

(vNE grows when ignorant of how the state was prepared)

(6) **Entropy of Measurement** (Q Meas adds randomness)

Measure $A = \sum_y a_y |a_y\rangle\langle a_y| \rightarrow$ outcomes $Y = \{a_y, p(a_y)\}$

S. E. of outcomes $H(Y) \geq S(\rho)$, w/ "=" for $[A, \rho] = 0$

(7) **Entropy of Preparation** (Mix N. O. states cannot recover full info)

Draw from $\{|\varphi_x\rangle, p_x\}$, $\rho = \sum_x p_x |\varphi_x\rangle\langle\varphi_x| \rightarrow H(X) \geq S(\rho)$

Quantum Information Theory (Preskill ch. 5)

Mathematical properties of $S(\rho)$ (No proofs, see Preskill)

(1) Purity $\rho = |\psi\rangle\langle\psi| \rightarrow S(\rho) = 0$

(2) Invariance $S(U\rho U^{+}) = S(\rho)$

(3) Maximum ρ has d eigenvalues $\neq 0 \rightarrow S(\rho) \leq \log d$

(4) Concavity For $\lambda_1, \lambda_2, \dots \lambda_n \geq 0$, $\sum_i \lambda_i = 1 \rightarrow$

$$S(\lambda_1\rho_1 + \dots + \lambda_n\rho_n) \geq \lambda_1 S(\rho_1) + \dots \lambda_n S(\rho_n)$$

(vNE grows when ignorant of how the state was prepared)

(6) Entropy of Measurement (Q Meas adds randomness)

Measure $A = \sum_y a_y |a_y\rangle\langle a_y| \rightarrow$ outcomes $Y = \{a_y, p(a_y)\}$

S. E. of outcomes $H(Y) \geq S(\rho)$, w/ "=" for $[A, \rho] = 0$

(7) Entropy of Preparation (Mix N. O. states cannot recover full info)

Draw from $\{|\varphi_x\rangle, p_x\}$, $\rho = \sum_x p_x |\varphi_x\rangle\langle\varphi_x| \rightarrow H(X) \geq S(\rho)$

Mathematical properties of $S(\rho)$ (No proofs, see Preskill)

(8) Subadditivity (info in whole ≤ sum of info in parts)

$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$ (classical $H(X,Y) \leq H(X) + H(Y)$ with "=" when X, Y uncorrelated)

(9) Triangle inequality (uncertainty about whole can be less than uncertainty about parts)

$S(\rho_{AB}) \geq |S(\rho_A) - S(\rho_B)|$ (classical $H(X,Y) \geq H(X), H(Y)$)

Quantum Information Theory (Preskill ch. 5)

Quantum Information Theory (Preskill ch. 5)

Quantum Data Compression

(Quantum analog of Shannons Noiseless Coding Theorem)

Starting Point: n – letter message drawn from $\{|\varphi_x\rangle, p_x\}$

need not be orthogonal

Each letter described by $\rho = \sum_x p_x |\varphi_x\rangle\langle\varphi_x|$

Message described by $\rho^n = \rho \otimes \rho \otimes \ldots \otimes \rho$

Basic Question: How redundant is this information?

– is there a "quantum code" which can compress to a smaller Hilbert space while retaining the fidelity of the encoded quantum information ?

Answer: Optimal Compression requires

$\log(\dim \mathcal{H}) = n S(\rho)$ qubits

(Schumacher's Theorem)

Corrollary: The von Neumann entropy is the # of qubits carried per letter in a message. We can always compress unless $\rho = \frac{1}{2}\mathbb{1}$

Example of how we might do this:

Alice sends a message using the alphabet $|\uparrow_z\rangle,\ p = 1/2$; $|\uparrow_x\rangle,\ p = 1/2$

$$\Rightarrow \rho = \frac{1}{2}|\uparrow_z\rangle\langle\uparrow_z| + \frac{1}{2}|\uparrow_x\rangle\langle\uparrow_x| = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$$

(z or x basis)

Symmetry → eigenvectors are $\uparrow, \downarrow$ along $\vec{n} = \frac{1}{\sqrt{2}}(\vec{x} + \vec{z})$

eigenvectors

$$\begin{cases} |0'\rangle = |\uparrow_{\vec{n}}\rangle = \begin{pmatrix} \cos \pi/8 \\ \sin \pi/8 \end{pmatrix} \\ |1'\rangle = |\downarrow_{\vec{n}}\rangle = \begin{pmatrix} \sin \pi/8 \\ -\cos \pi/8 \end{pmatrix} \end{cases}$$

eigenvalues

$$\begin{cases} \lambda(0') = \frac{1}{2} + \frac{1}{2\sqrt{2}} = \cos^2 \pi/8 \\ \lambda(1') = \frac{1}{2} - \frac{1}{2\sqrt{2}} = \sin^2 \pi/8 \end{cases}$$

Quantum Information Theory (Preskill ch. 5)

Corrollary: The von Neumann entropy is the # of qubits carried per letter in a message. We can always compress unless $\rho = \frac{1}{2}\mathbb{1}$

Example of how we might do this:

Alice sends a message using the alphabet $|\uparrow_z\rangle,\ p=1/2$; $|\uparrow_x\rangle,\ p=1/2$

$$\Rightarrow\ \rho = \tfrac{1}{2}|\uparrow_z\rangle\langle\uparrow_z| + \tfrac{1}{2}|\uparrow_x\rangle\langle\uparrow_x| = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$$

(z or x basis)

Symmetry → eigenvectors are $\uparrow, \downarrow$ along $\vec{n} = \frac{1}{\sqrt{2}}(\vec{x}+\vec{z})$

eigenvectors

$$|0'\rangle = |\uparrow_{\vec{n}}\rangle = \begin{pmatrix} \cos\pi/8 \\ \sin\pi/8 \end{pmatrix}$$

$$|1'\rangle = |\downarrow_{\vec{n}}\rangle = \begin{pmatrix} \sin\pi/8 \\ -\cos\pi/8 \end{pmatrix}$$

eigenvalues

$$\lambda(0') = \frac{1}{2} + \frac{1}{2\sqrt{2}} = \cos^2\pi/8$$

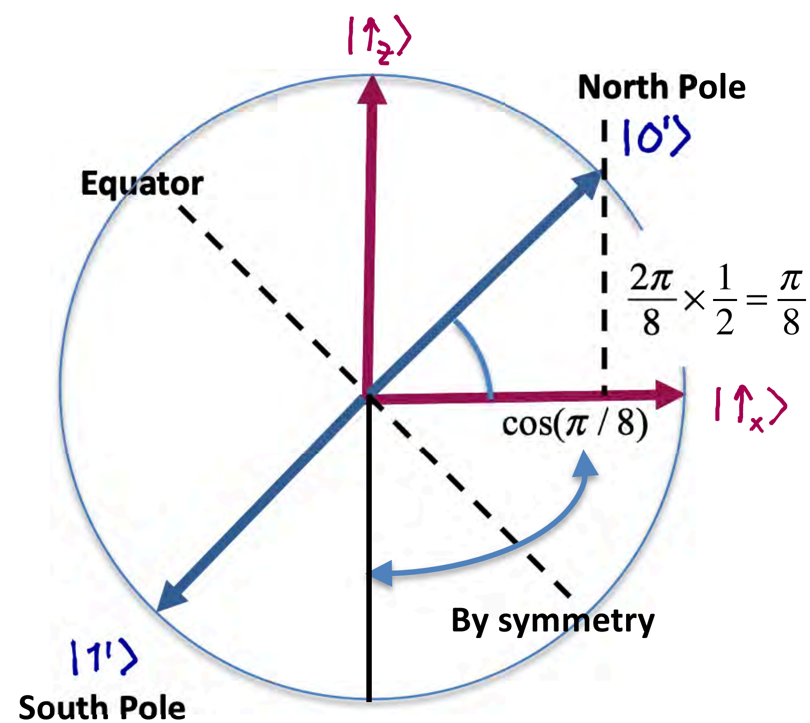
$$\lambda(1') = \frac{1}{2} - \frac{1}{2\sqrt{2}} = \sin^2\pi/8$$

Overlap

$$\langle 0'|\uparrow_z\rangle^2 = \langle 0'|\uparrow_x\rangle^2 = \cos^2\pi/8 = 0.8535$$

$$\langle 1'|\uparrow_z\rangle^2 = \langle 1'|\uparrow_x\rangle^2 = \sin^2\pi/8 = 0.1465$$

Visualization on the Bloch Sphere



Quantum Information Theory (Preskill ch. 5)

Corrollary: The von Neumann entropy is the # of qubits carried per letter in a message. We can always compress unless $\rho = \frac{1}{2}\mathbb{1}$

Example of how we might do this:

Alice sends a message using the alphabet $\begin{array}{ll} |\uparrow_z\rangle, & p = 1/2 \\ |\uparrow_x\rangle, & p = 1/2 \end{array}$

$$\Rightarrow \quad \rho = \tfrac{1}{2}|\uparrow_z\rangle\langle\uparrow_z| + \tfrac{1}{2}|\uparrow_x\rangle\langle\uparrow_x| = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$$

(z or x basis)

Symmetry → eigenvectors are $\uparrow, \downarrow$ along $\vec{n} = \frac{1}{\sqrt{2}}(\vec{x} + \vec{z})$

eigenvectors

$$\begin{cases} |0'\rangle = |\uparrow_{\vec{n}}\rangle = \begin{pmatrix} \cos \pi/8 \\ \sin \pi/8 \end{pmatrix} \\ |1'\rangle = |\downarrow_{\vec{n}}\rangle = \begin{pmatrix} \sin \pi/8 \\ -\cos \pi/8 \end{pmatrix} \end{cases}$$

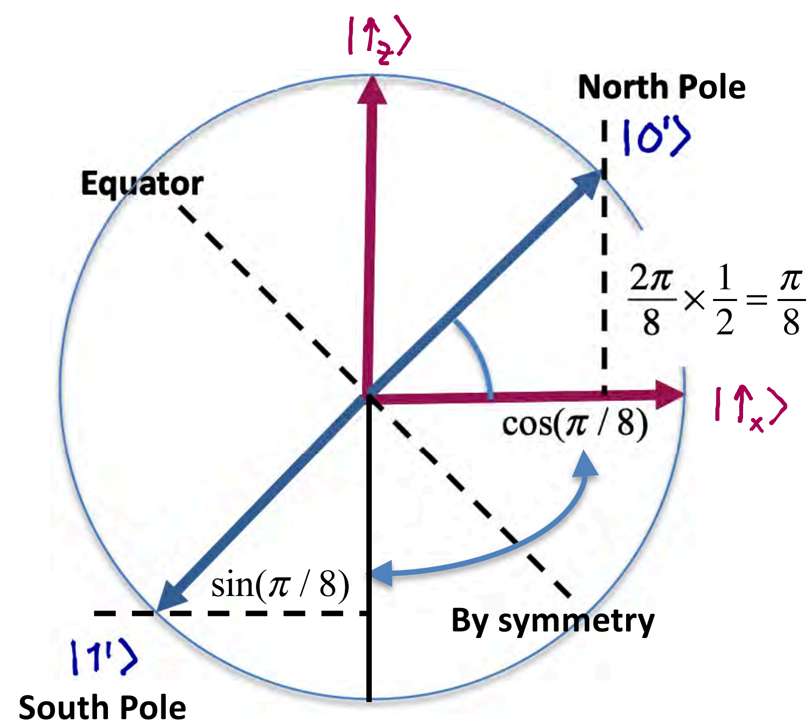
eigenvalues

$$\begin{cases} \lambda(0') = \frac{1}{2} + \frac{1}{2\sqrt{2}} = \cos^2 \pi/8 \\ \lambda(1') = \frac{1}{2} - \frac{1}{2\sqrt{2}} = \sin^2 \pi/8 \end{cases}$$

Overlap

$$\begin{cases} \langle 0'|\uparrow_z\rangle^2 = \langle 0'|\uparrow_x\rangle^2 = \cos^2 \pi/8 = 0.8535 \\ \langle 1'|\uparrow_z\rangle^2 = \langle 1'|\uparrow_x\rangle^2 = \sin^2 \pi/8 = 0.1465 \end{cases}$$

Visualization on the Bloch Sphere



Quantum Information Theory (Preskill ch. 5)

Corrollary: The von Neumann entropy is the # of qubits carried per letter in a message. We can always compress unless $\rho = \frac{1}{2}\mathbb{1}$

Example of how we might do this:

Alice sends a message using the alphabet $|\uparrow_z\rangle,\ p = 1/2$; $|\uparrow_x\rangle,\ p = 1/2$

$$\Rightarrow \rho = \frac{1}{2}|\uparrow_z\rangle\langle\uparrow_z| + \frac{1}{2}|\uparrow_x\rangle\langle\uparrow_x| = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$$

(z or x basis)

Symmetry → eigenvectors are $\uparrow, \downarrow$ along $\vec{n} = \frac{1}{\sqrt{2}}(\vec{x}+\vec{z})$

eigenvectors
$$|0'\rangle = |\uparrow_{\vec{n}}\rangle = \begin{pmatrix} \cos\pi/8 \\ \sin\pi/8 \end{pmatrix}$$
$$|1'\rangle = |\downarrow_{\vec{n}}\rangle = \begin{pmatrix} \sin\pi/8 \\ -\cos\pi/8 \end{pmatrix}$$

eigenvalues
$$\lambda(0') = \frac{1}{2} + \frac{1}{2\sqrt{2}} = \cos^2\pi/8$$
$$\lambda(1') = \frac{1}{2} - \frac{1}{2\sqrt{2}} = \sin^2\pi/8$$

Overlap
$$\langle 0'|\uparrow_z\rangle^2 = \langle 0'|\uparrow_x\rangle^2 = \cos^2\pi/8 = 0.8535$$
$$\langle 1'|\uparrow_z\rangle^2 = \langle 1'|\uparrow_x\rangle^2 = \sin^2\pi/8 = 0.1465$$

Bob does not know what was sent → **best guess** $|\psi\rangle = |0'\rangle$

$$\Rightarrow F = \langle\psi|\rho|\psi\rangle = \frac{1}{2}|\langle\uparrow_z|\psi\rangle|^2 + \frac{1}{2}|\langle\uparrow_x|\psi\rangle|^2 = 0.8535$$

Scenario:

Alice wants to send a 3-qubit message, but can transmit only 2 qubits. Could send Bob those qubits $(F=1)$ and have Bob guess $|0'\rangle$ for the third

⇒ **Baseline Fidelity** (no tricks) $F = 0.8535$

How to improve:

Diagonalize ρ — Likely subspace $|0'\rangle$; Unlikely subspace $|1'\rangle$ — for 1 bit

Let $|\Psi\rangle = \underbrace{|\psi_1\rangle|\psi_2\rangle|\psi_3\rangle}_{\text{either } |\uparrow_z\rangle \text{ or } |\uparrow_x\rangle}$

Note: all possible $|\Psi\rangle$ have the same overlap with states of the type $|i\rangle|j\rangle|k\rangle$ where $i,j,k \in \{0',1'\}$

Quantum Information Theory (Preskill ch. 5)

Corrollary: The von Neumann entropy is the # of qubits carried per letter in a message. We can always compress unless $\rho = \frac{1}{2}\mathbb{1}$

Example of how we might do this:

Alice sends a message using the alphabet $|\uparrow_z\rangle,\ p = 1/2$; $|\uparrow_x\rangle,\ p = 1/2$

$$\rho = \frac{1}{2}|\uparrow_z\rangle\langle\uparrow_z| + \frac{1}{2}|\uparrow_x\rangle\langle\uparrow_x| = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$$

(z or x basis)

Symmetry → eigenvectors are $\uparrow, \downarrow$ along $\vec{n} = \frac{1}{\sqrt{2}}(\vec{x}+\vec{z})$

eigenvectors

$$|0'\rangle = |\uparrow_{\vec{n}}\rangle = \begin{pmatrix} \cos\pi/8 \\ \sin\pi/8 \end{pmatrix}$$

$$|1'\rangle = |\downarrow_{\vec{n}}\rangle = \begin{pmatrix} \sin\pi/8 \\ -\cos\pi/8 \end{pmatrix}$$

eigenvalues

$$\lambda(0') = \frac{1}{2} + \frac{1}{2\sqrt{2}} = \cos^2\pi/8$$

$$\lambda(1') = \frac{1}{2} - \frac{1}{2\sqrt{2}} = \sin^2\pi/8$$

Overlap

$$\langle 0'|\uparrow_z\rangle^2 = \langle 0'|\uparrow_x\rangle^2 = \cos^2\pi/8 = 0.8535$$

$$\langle 1'|\uparrow_z\rangle^2 = \langle 1'|\uparrow_x\rangle^2 = \sin^2\pi/8 = 0.1465$$

How to improve:

Diagonalize ρ — Likely subspace $|0'\rangle$; Unlikely subspace $|1'\rangle$ — for 1 bit

Let $|\Psi\rangle = \underbrace{|\psi_1\rangle|\psi_2\rangle|\psi_3\rangle}_{\text{either } |\uparrow_z\rangle \text{ or } |\uparrow_x\rangle}$

Note: all possible $|\Psi\rangle$ have the same overlap with states of the type $|i\rangle|j\rangle|k\rangle$ where $i,j,k \in \{0',1'\}$

Note: for any $|\Psi\rangle$ drawn from Alice and Bobs alphabet, we have

Likely subspace Λ

$$|\langle 0'0'0'|\Psi\rangle|^2 = \cos^6\pi/8 = \underline{0.6219}$$

$$|\langle 0'0'1'|\Psi\rangle|^2 = |\langle 0'1'0'|\Psi\rangle|^2 = |\langle 1'0'0'|\Psi\rangle|^2 = \cos^4\frac{\pi}{8}\sin^2\frac{\pi}{8} = \underline{0.1067}$$

Unlikely subspace $\Lambda^\perp$

$$|\langle 0'1'1'|\Psi\rangle|^2 = |\langle 1'0'1'|\Psi\rangle|^2 = |\langle 1'1'0'|\Psi\rangle|^2 = \cos^2\frac{\pi}{8}\sin^4\frac{\pi}{8} = \underline{0.0183}$$

$$|\langle 1'1'1'|\Psi\rangle|^2 = \sin^6\frac{\pi}{8} = \underline{0.0031}$$

Quantum Information Theory (Preskill ch. 5)

Overlap $\begin{cases} \langle 0'|\uparrow_z\rangle^2 = \langle 0'|\uparrow_x\rangle^2 = \cos^2\pi/8 = 0.8535 \\ \langle 1'|\uparrow_z\rangle^2 = \langle 1'|\uparrow_x\rangle^2 = \sin^2\pi/8 = 0.1465 \end{cases}$

How to improve:

Diagonalize ρ $\begin{cases} \text{Likely subspace } |0'\rangle \\ \text{Unlikely subspace } |1'\rangle \end{cases}$ for 1 bit

Let $|\psi\rangle = \overbrace{|\psi_1\rangle|\psi_2\rangle|\psi_3\rangle}^{\text{either } |\uparrow_z\rangle \text{ or } |\uparrow_x\rangle}$

Note: all possible $|\psi\rangle$ have the same overlap with states of the type $|i\rangle|j\rangle|k\rangle$ where $i,j,k \in \{0',1'\}$

Note: for any $|\psi\rangle$ drawn from Alice and Bobs alphabet, we have

Likely subspace Λ

$|\langle 0'0'0'|\psi\rangle|^2 = \cos^6\pi/8 = \underline{0.6219}$

$|\langle 0'0'1'|\psi\rangle|^2 = |\langle 0'1'0'|\psi\rangle|^2 = |\langle 1'0'0'|\psi\rangle|^2 = \cos^4\frac{\pi}{8}\sin^2\frac{\pi}{8} = \underline{0.1067}$

Unlikely subspace $\Lambda^\perp$

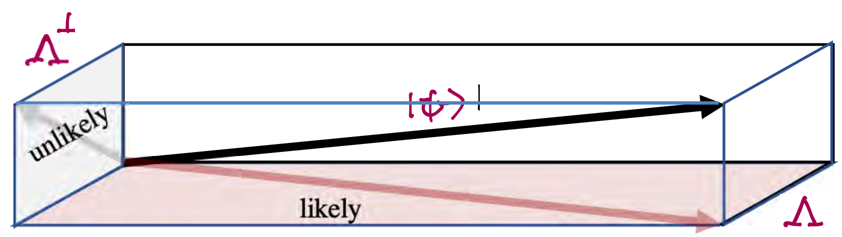
$|\langle 0'1'1'|\psi\rangle|^2 = |\langle 1'0'1'|\psi\rangle|^2 = |\langle 1'1'0'|\psi\rangle|^2 = \cos^2\frac{\pi}{8}\sin^4\frac{\pi}{8} = \underline{0.0183}$

$|\langle 1'1'1'|\psi\rangle|^2 = \sin^6\frac{\pi}{8} = \underline{0.0031}$

This structure of message space suggests Alice should send only the part $\in \Lambda$, which fits in 2 qubits ($\mathrm{Dim}\,\Lambda = 4$). She can do this by performing a measurement that projects her 3–qubit state onto either Λ or $\Lambda^\perp$

$\Rightarrow \begin{cases} P_{likely} = 0.6219 + 3\times 0.1067 = 0.9419 \\ P_{unlikely} = 3\times 0.0183 + 0.0031 = 0.0581 \end{cases}$

Geometric illustration of a space with likely and unlikely subspaces.



The state vector (black) has a large projection on the "likely subspace" (pink), and a much smaller projection on the "unlikely subspace" (grey). Accordingly, when we project onto the likely subspace and ignore the unlikely subspace, we don't loose much information.

Quantum Information Theory (Preskill ch. 5)

Overlap
$$\begin{cases} \langle 0'|\uparrow_z\rangle^2 = \langle 0'|\uparrow_x\rangle^2 = \cos^2\pi/8 = 0.8535 \\ \langle 1'|\uparrow_z\rangle^2 = \langle 1'|\uparrow_x\rangle^2 = \sin^2\pi/8 = 0.1465 \end{cases}$$

How to improve:

Diagonalize ρ $\begin{cases} \text{Likely subspace } |0'\rangle \\ \text{Unlikely subspace } |1'\rangle \end{cases}$ for 1 bit

Let $|\psi\rangle = \overbrace{|\psi_1\rangle|\psi_2\rangle|\psi_3\rangle}^{\text{either } |\uparrow_z\rangle \text{ or } |\uparrow_x\rangle}$

Note: all possible $|\psi\rangle$ have the same overlap with states of the type $|i\rangle|j\rangle|k\rangle$ where $i,j,k \in \{0',1'\}$

Note: for any $|\psi\rangle$ drawn from Alice and Bobs alphabet, we have

Likely subspace Λ

$|\langle 0'0'0'|\psi\rangle|^2 = \cos^6\pi/8 = \underline{0.6219}$

$|\langle 0'0'1'|\psi\rangle|^2 = |\langle 0'1'0'|\psi\rangle|^2 = |\langle 1'0'0'|\psi\rangle|^2 = \cos^4\frac{\pi}{8}\sin^2\frac{\pi}{8} = \underline{0.1067}$

Unlikely subspace $\Lambda^\perp$

$|\langle 0'1'1'|\psi\rangle|^2 = |\langle 1'0'1'|\psi\rangle|^2 = |\langle 1'1'0'|\psi\rangle|^2 = \cos^2\frac{\pi}{8}\sin^4\frac{\pi}{8} = \underline{0.0183}$

$|\langle 1'1'1'|\psi\rangle|^2 = \sin^6\frac{\pi}{8} = \underline{0.0031}$

This structure of message space suggests Alice should send only the part $\in \Lambda$, which fits in 2 qubits ($\mathrm{Dim}\,\Lambda = 4$). She can do this by performing a measurement that projects her 3–qubit state onto either Λ or $\Lambda^\perp$

$$\Rightarrow \begin{cases} P_{\text{Likely}} = 0.6219 + 3\times 0.1067 = 0.9419 \\ P_{\text{Unlikely}} = 3\times 0.0183 + 0.0031 = 0.0581 \end{cases}$$

To do this Alice can apply the Unitary Transformation U that maps

$$U: |\psi_{\text{Likely}}\rangle \to |\cdot\rangle|\cdot\rangle|0'\rangle$$

$$U: |\psi_{\text{Unlikely}}\rangle \to |\cdot\rangle|\cdot\rangle|1'\rangle$$

She measures the 3rd qubit

$$\Rightarrow \begin{cases} |0\rangle \to \text{projects onto } \Lambda \\ |1\rangle \to \text{projects onto } \Lambda^\perp \end{cases}$$

If Alice's outcome is $|0'\rangle$ she sends $|\psi_{\text{comp}}\rangle = |\cdot\rangle|\cdot\rangle$

Bob decompresses by appending $|0'\rangle$ and undoing U

$$\Rightarrow |\psi_{\text{Bob}}\rangle = U^{-1}\left(|\psi_{\text{comp}}\rangle|0'\rangle\right)$$

Quantum Information Theory (Preskill ch. 5)

This structure of message space suggests Alice should send only the part $\in \Lambda$, which fits in 2 qubits ($\mathcal{D}im\ \Lambda = 4$). She can do this by performing a measurement that projects her 3–qubit state onto either Λ or $\Lambda^{\perp}$

$$\Rightarrow \begin{cases} P_{Likely} = 0.6219 + 3\times 0.1067 = 0.9419 \\ P_{Unlikely} = 3\times 0.0183 + 0.0031 = 0.0581 \end{cases}$$

To do this Alice can apply the Unitary Transformation U that maps

$$U: |\Psi_{Likely}\rangle \to |\cdot\rangle|\cdot\rangle|0'\rangle$$

$$U: |\Psi_{UnLikely}\rangle \to |\cdot\rangle|\cdot\rangle|1'\rangle$$

She measures the 3rd qubit

$$\Rightarrow \begin{cases} |0\rangle \to \text{projects onto } \Lambda \\ |1\rangle \to \text{projects onto } \Lambda^{\perp} \end{cases}$$

If Alice's outcome is $|0'\rangle$ she sends $|\Psi_{comp}\rangle = |\cdot\rangle|\cdot\rangle$

Bob decompresses by appending $|0'\rangle$ and undoing U

$$\Rightarrow |\Phi_{Bob}\rangle = U^{-1}\left(|\Psi_{comp}\rangle|0'\rangle\right)$$

If Alice's outcome is $|1'\rangle$ she sends $U|0'0'\rangle$

$$\Rightarrow |\tilde{\Phi}_{Bob}\rangle = U^{-1}\left(|0'0'\rangle|0'\rangle\right) = |0'0'0'\rangle$$

This leaves Bob with the state

$$\rho_{Bob} = E|\Psi\rangle\langle\Psi|E + \langle\Psi|1-E|\Psi\rangle|0'0'0'\rangle\langle 0'0'0'|$$

($E =$ projection on Λ)

which has Fidelity

$$\mathcal{F} = \langle\Psi|\rho_{Bob}|\Psi\rangle$$

$$= \langle\Psi|E|\Psi\rangle + \langle\Psi|1-E|\Psi\rangle|\langle\Psi|0'0'0'\rangle|^2$$

$$= 0.9419^2 + 0.0581\times 0.6219 = \underline{0.9234} > \underline{0.8535}$$

As with classical data compression, longer messages allow for more compression or compression without loss.

In Quantum Communication one has the option of choosing an alphabet where the individual letters are mixed states. This makes it much harder to find bounds for compressibility and code rates. See Preskill for more information.