PHY682 Special Topics in Solid-State Physics: Quantum Information Science

Lecture time: 2:40-4:00PM Monday & Wednesday

Today 11/2:

1. Homework 6 and Final presentation
2. Today: quantum data compression
3. Week 10's topics: **No clones in quantum**

Student presentation

There will be seven groups each having 3 student members. Each presentation will be 20mins (presentation) + 5 mins (Q&A).

[11/30] Group 1: Bak, Gokhale & Nghiem Vu; Group 2: Bashir, Gordon & Yu; Group **7**: Wallace, Wu & Zhao;

[12/1] Group 4: Gregory, Lee & Xu; Group 5: Chheta, Sukeno & Zou; Group 6: Thotakura, Zhang, Zhu

[12/7 (last day of class)] (Student presentation) Group **3**: Farno, Guo & Singletary

Next, we turn to Quantum Data Compression (a.k.a. Noiseless Quantum Shannon Channel Coding Theorem)

Shannon entropy (Classical)

- We have seen the von Neumann entropy of a density matrix (log is base-2):

$$S_V(\rho) \equiv -Tr\rho \log \rho \quad = -\sum_k \lambda_k \log(\lambda_k) \qquad \lambda_k\text{'s are eigenvalues of } \rho \text{ and } \sum_k \lambda_k = 1$$

- Shannon entropy (already used above) for a probability distribution:

$$H(X) = H(\{p_x\}) = -\sum_x p_x \log(p_x)$$

➔ Average number of bits needed to represent a symbol ***x*** (selected from a set with distribution {$\boldsymbol{p_x}$}

- Example: four symbols a,b,c,d with {p_a=1/2, p_b=1/4, p_c=p_d=1/8}

  - Use the encoding: a→0, b→10, c→110, d→111 Would I encode using ~~a → 0, b → 01, c → 10, d → 11~~ ?

  - Average bits = 1 x 1/2 + 2 x ¼ + 3 x 1/8 x 2 = 7/4 ≥ H(X) [cannot be less than Shannon entropy]

    → d a a b a b c a

    - Question: 11100100101100 encodes what message? Ans: daababca

Shannon's noiseless channel coding theorem

source → channel → channel has no noise

- Consider a source consists of a sequence of random variable $X_1, X_2, \ldots$, whose values ($x_1, x_2, \ldots$ are drawn from alphabet of e.g. {a,b,c,d}) represent output of the source

  $X_1 X_2 \cdots X_n$ a b (a b) not correlated

- Assume different uses of the source are independent and identically distributed (iid)

- Example: binary alphabet--- source emitting each **X** with 0 with probability p; 1 with probability 1-p

  prob p 0 1 0 1 1 … 0 (1-p)

→ Divide the sequence $x_1, x_2, \ldots x_n$ to two types:
(i) typical: those that occur with high probability ----- [for large n, a fraction of p is 0, and a fraction of 1-p is 1]
(ii) atypical: those that rarely occur

# of 0's # of 1's Shannon entropy

$$p(x_1, x_2, \ldots, x_n) = p(x_1)p(x_2)\ldots p(x_n)$$

for typical squence, each appears with prob. $\approx p^{np}(1-p)^{n(1-p)} = 2^{-nH(X)}$

$H(X) = -p\log_2 p - (1-p)\log_2(1-p)$

$2^{nR} > \# 2^{nH} = \#$ typical sequence

→ For atypical sequence: do nothing (cannot compress); for typical sequence, can compress using $nH(X)$ bits

Shannon's noiseless channel coding theorem

average # of bits

R =

Suppose $\{X_i\}$ is an i.i.d. information source with entropy rate $H(X)$.
If $R > H(X)$, then there exists a reliable compression scheme of rate R for the source.
if $R < H(X)$, then any compression scheme will not be reliable.

[See N&C Thm 12.4]

- How to generalize to quantum regime?
  (1) Alphabet is drawn from a set of quantum states $\{|\phi_x\rangle\}$
  (2) $\{X_i\}$ → an ensemble $\rho = \sum_x q_x |\phi_x\rangle\langle\phi_x|$
  (3) Typical sequence → typical subspace; atypical sequence → atypical subspace

(4) Shannon entropy $H(X)$ → von Neumann entropy $S_V(\rho)$

➔ Expect: Schumacher's quantum data compression or q. Shannon noiseless channel coding

Suppose $\{\rho\}$ is an i.i.d. quantum information source with entropy rate $S(\rho)$.
If $R > S(\rho)$, then there exists a reliable compression scheme of rate R for the source.
if $R < S(\rho)$, then any compression scheme will not be reliable.

Schumacher's quantum data compression

= Quantum Shannon noiseless channel coding

e.g. $|\phi_{x=0}\rangle = |0\rangle$, $p_0 = \frac{1}{2}$; $|\phi_{x=1}\rangle = |+\rangle$, $p_1 = \frac{1}{2}$

❖ How to generalize to quantum regime?

(1) Alphabet is drawn from a set of quantum states $\{|\phi_x\rangle\}$

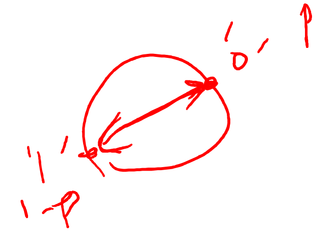
(2) $\{X_i\}$ → an ensemble $\rho = \sum_x q_x |\phi_x\rangle\langle\phi_x|$

$|\phi_x\rangle$ not eigenstates & not eigenvalues q_x

(3) Typical sequence → typical subspace; atypical sequence → atypical subspace

(4) $H(x) \longrightarrow S(\rho)$

$$\text{Diagonalize } \rho = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i| \qquad \text{For a qubit: } \rho = p|'0'\rangle\langle '0'| + (1-p)|'1'\rangle\langle '1'|$$



- Source is i.i.d. so (on average) is emitting a message of length n: $\boldsymbol{\rho}^{\otimes n} \equiv \underbrace{\boldsymbol{\rho} \otimes \cdots \otimes \boldsymbol{\rho}}_{n\ \rho\text{'s}}$

➢ We focus on qubit case. Typical subspace $\boldsymbol{\Lambda}$ is spanned by those 'sequences' that have a fraction of p is '0', and a fraction of 1-p is '1'

→ can represented by a projector: (about $2^{n\,S(\rho)} = 2^{n\,H(p)}$ such sequences)

↳ dimension of the subspace

$$P_\Lambda = \sum_{\text{typical } z's} |z_1\rangle\langle z_1| \otimes |z_2\rangle\langle z_2| \otimes \cdots \otimes |z_n\rangle\langle z_n|$$

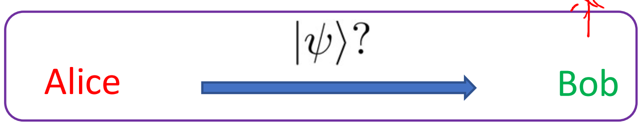
Quantum Data Compression (and Transmission)

➢ There exists a unitary transformation U which takes any n-qubit state $|\phi_\Lambda\rangle$ in $\boldsymbol{\Lambda}$ to

$$\boldsymbol{U}|\phi_\Lambda\rangle = |\phi_{\text{compressed}}\rangle|0_{\text{rest}}\rangle$$ [Expect decompression is via U^{-1}]

where $|\phi_{\text{compressed}}\rangle$ is $n(S(\rho)+\delta)$-qubit, and $|0_{\text{rest}}\rangle$ is $|0\rangle \otimes \cdots \otimes |0\rangle$

$n - \quad n S(\rho)$



For an input state $|\psi\rangle = |\phi_1\rangle \cdots |\phi_n\rangle$, apply $\boldsymbol{U}$ on $|\psi\rangle$, and measure the state of the last $(n - n(S+\delta))$ qubits.

1. If result is $|0_{\text{rest}}\rangle$, Alice successfully compresses $|\psi\rangle$ onto $|\psi_{\text{compressed}}\rangle|0_{\text{rest}}\rangle$, and sends $|\psi_{\text{compressed}}\rangle$ to Bob, who can decompress it. $U^{-1}|\phi_{\text{compressed}}\,0\rangle = |\phi_\Lambda\rangle$

2. If result other than $|0_{\text{rest}}\rangle$, she fails to compress her message. The best she can do is send a state $|0'_{\text{compressed}}\rangle$ where $\boldsymbol{U}|\lambda_{n,1}\rangle = |0'_{\text{compressed}}\rangle|0_{\text{rest}}\rangle$ [with $|\lambda_{n,1}\rangle$ having the largest probability in Λ]

Example

$|0\rangle$ $H = \leftrightarrow$ $|+\rangle$ $D = \nearrow$ $\langle D|H\rangle = \frac{1}{\sqrt{2}}$

Suppose the ensemble consists of $\{(|H\rangle, p_H = \frac{1}{2}), (|D\rangle, p_D = \frac{1}{2})\}$, where $|H\rangle$ is the state of horizontal polarization while $|D\rangle$ is $45°$,

$$|H\rangle = \begin{pmatrix}1\\0\end{pmatrix},\ |D\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \implies \rho = \frac{1}{2}|H\rangle\langle H| + \frac{1}{2}|D\rangle\langle D| = \begin{pmatrix}\frac{3}{4} & \frac{1}{4}\\ \frac{1}{4} & \frac{1}{4}\end{pmatrix}$$

- Diagonalize ρ:

$$|Q\rangle = |22.5°\rangle = \begin{pmatrix}\cos\frac{\pi}{8}\\ \sin\frac{\pi}{8}\end{pmatrix},\ \lambda_Q = \cos^2\frac{\pi}{8}$$

$$|\overline{Q}\rangle = |112.5°\rangle = \begin{pmatrix}-\sin\frac{\pi}{8}\\ \cos\frac{\pi}{8}\end{pmatrix},\ \lambda_{\overline{Q}} = \sin^2\frac{\pi}{8}$$

$$S(\rho) = -\lambda_Q \log_2 \lambda_Q - \lambda_{\overline{Q}} \log_2 \lambda_{\overline{Q}} \approx 0.60088$$

$|QQQ\rangle$ has eigenvalue λ_Q^3

$|QQ\bar{Q}\rangle, |Q\bar{Q}Q\rangle, |\bar{Q}QQ\rangle$ $\lambda_Q^2 \lambda_{\bar{Q}}$

$2^2 = 4$

- 3 x S(ρ)≈1.8 ➔ can use two qubits to encode a three-qubit message

The typical subspace Λ is spanned by $\{|1\rangle, |2\rangle, |3\rangle, |4\rangle\}$, and its dual subspace $\Lambda^{\perp}$ by $\{|5\rangle, |6\rangle, |7\rangle, |8\rangle\}$

$$|1\rangle = |QQQ\rangle,\ |2\rangle = |QQ\overline{Q}\rangle,\ |3\rangle = |Q\overline{Q}Q\rangle,\ |4\rangle = |\overline{Q}QQ\rangle,$$
$$|5\rangle = |Q\overline{Q}\overline{Q}\rangle,\ |6\rangle = |\overline{Q}Q\overline{Q}\rangle,\ |7\rangle = |\overline{Q}\overline{Q}Q\rangle,\ |8\rangle = |\overline{Q}\overline{Q}\overline{Q}\rangle$$

$$\lambda_1 = \cos^6\tfrac{\pi}{8},\ \lambda_2 = \lambda_3 = \lambda_4 = \cos^4\tfrac{\pi}{8}\sin^2\tfrac{\pi}{8},$$
$$\lambda_5 = \lambda_6 = \lambda_7 = \cos^2\tfrac{\pi}{8}\sin^4\tfrac{\pi}{8},\ \lambda_8 = \sin^6\tfrac{\pi}{8}.$$

Probabilities: $p_\Lambda = \mathrm{tr}(\rho^{\otimes 3}\boldsymbol{P}_{\boldsymbol{\Lambda}}) = \sum_{i=1}^{4}\lambda_i \approx \boxed{0.9419}$ vs. $p_{\Lambda^\perp} = \mathrm{tr}(\rho^{\otimes 3}(\mathbf{1} - \boldsymbol{P}_{\boldsymbol{\Lambda}})) \approx 0.0581.$

- Alice and Bob both agree on the unitary transformation U to compress; communicate using first two qubits

always H

indicator of typical subspace Λ

V

indicator of atypical subspace $\Lambda^\perp$

$$\mathrm{U}\begin{pmatrix}|1\rangle\\|2\rangle\\|3\rangle\\|4\rangle\end{pmatrix} \rightarrow \begin{pmatrix}|HHH\rangle\\|HVH\rangle\\|VHH\rangle\\|VVH\rangle\end{pmatrix} \quad \mathrm{U}\begin{pmatrix}|5\rangle\\|6\rangle\\|7\rangle\\|8\rangle\end{pmatrix} \rightarrow \begin{pmatrix}|HHV\rangle\\|HVV\rangle\\|VHV\rangle\\|VVV\rangle\end{pmatrix}$$

Alice's message $|\psi\rangle$ can be expanded as $|\psi\rangle = \sum_{i=1}^{8} a_i|i\rangle$ (We know $\sum_{i=1}^{4} |a_i|^2 \gg \sum_{i=5}^{8} |a_i|^2$)

$|H\rangle|D\rangle|H\rangle \dots$

$S(\rho)$

Alice applies $\boldsymbol{U}$ on $|\psi\rangle$ followed by measurement on the third qubit

ρ = ensemble description of source

If the result is $|H\rangle$, she sucessfully projects $|\psi\rangle$ into Λ: $a_1|HHH\rangle + a_2|HVH\rangle + a_3|VHH\rangle + a_4|VVH\rangle = |\psi_{\text{compressed}}\rangle|H\rangle$ $\Longrightarrow$ She sends $|\psi_{\text{compressed}}\rangle$ to Bob for uncompressing.

Bob applies $\boldsymbol{U^{-1}}$: $|\psi'\rangle = \boldsymbol{U^{-1}}(|\psi_{\text{compressed}}\rangle|H\rangle) = \sum_{i=1}^{4} a_i|i\rangle$, which has high resemblance to the initial $|\psi\rangle$, $F_1 \equiv |\langle\psi|\psi'\rangle|^2 = p_\Lambda \approx 0.9419$

$|\langle HHH|i'\rangle|^2$

[If result is V, she sends HH to Bob, which will be decomposed to |1⟩ and has fidelity=(.0581)(.6219)≈0.036]

- How good is this? Let us compare it to the case when Alice sends the first two and asks Bob to guess the third letter

H, D, H
D, H, H
D ... V

The best guess he can make is $|Q\rangle$. The fidelity is $F = \frac{1}{2}|\langle H|Q\rangle|^2 + \frac{1}{2}|\langle D|Q\rangle|^2 = 0.8535 < 0.9419$

Recall: Entanglement cost and formation

- How can we achieve the optimal "dilution" process?
- Ans. 'Quantum data compression' (just explained)+ quantum teleportation

$\frac{1}{2} I$

[Schumacher '96] [Bennett et al. '92]

n Φ^+ k pairs of Bell states ψ

$n > k$

1. Alice prepares ***n*** copies of the states locally, and compresses the part that will be shared by Bob

Should make sense now!

k qubits

n copies of ψ

$$|\psi_{12}\rangle = \cos\theta\,|0_1 0_2\rangle + \sin\theta\,|1_1 1_2\rangle$$
$$|\psi_{34}\rangle = \cos\theta\,|0_3 0_4\rangle + \sin\theta\,|1_3 1_4\rangle$$
$$\vdots$$
$$|\psi_{2n-1,2n}\rangle = \cos\theta\,|0_{2n-1} 0_{2n}\rangle + \sin\theta\,|1_{2n-1} 1_{2n}\rangle$$

Compressed into k qubits

$$k \approx n\, S_V(Tr_B|\psi\rangle\langle\psi|)$$

$\rho_A = tr \cdots = \cos^2\theta\, |0\rangle\langle 0| + \sin^2\theta\, |1\rangle\langle 1|$

Teleport

Using k Bell pairs

Bob Decompresses

2. Alice and Bob share n copies of ψ by consuming ***k = n E_c*** copies of Bell states

Noisy channel coding*: classical vs. quantum

[From Nielsen & Chuang]

Theorem 12.7: (**Shannon's noisy channel coding theorem**) For a noisy channel $\mathcal{N}$ the capacity is given by

$$C(\mathcal{N}) = \max_{p(x)} H(X:Y)\,, \qquad (12.67)$$

where the maximum is taken over all input distributions $p(x)$ for X, for one use of the channel, and Y is the corresponding induced random variable at the output of the channel.

$$H(X:Y) = H(Y) - H(Y|X)$$
$$= H(Y) - \sum_x p(x) H(Y|X=x)$$

Theorem 12.8: (**Holevo–Schumacher–Westmoreland (HSW) theorem**) Let $\mathcal{E}$ be a trace-preserving quantum operation. Define

$$\chi(\mathcal{E}) \equiv \max_{\{p_j,\rho_j\}} \left[S\left(\mathcal{E}\left(\sum_j p_j \rho_j\right)\right) - \sum_j p_j S(\mathcal{E}(\rho_j)) \right], \qquad (12.71)$$

where the maximum is over all ensembles $\{p_j, \rho_j\}$ of possible input states ρ_j to the channel. Then $\chi(\mathcal{E})$ is the product state capacity for the channel $\mathcal{E}$, that is, $\chi(\mathcal{E}) = C^{(1)}(\mathcal{E})$.