4. Quantum information and quantum error correction

4.1 Measure of quantum information

A classical random variable $X$ that takes on values $x_i$ with probabilities $p_i \Rightarrow$ Shannon entropy

\[ S_{\text{Sh}} = -\sum_j p_j \log p_j \]

as a measure of information.

Consider orthogonal states with probability $p_j$:

\[ x_j \leftrightarrow |x_j\rangle \quad \langle x_j| x_k \rangle = \delta_{j,k} \]

\[ g = \sum_j p_j \langle x_j| x_j \rangle \]

Eq. (1) can be written as

\[ S_{\text{Sh}} = -\text{Tr} \left\{ g \log g \right\} = S_N \]

i.e. this is the von Neumann entropy. We now use $S_N$ also as a measure of information contained in a quantum state $g$.

\[ S_N = -\text{Tr} \left\{ g \log g \right\} \]

rem.: the basis of the log is 2
some properties (some known and some new)

(i) pure state \( \rho = |\psi\rangle \langle \psi| \quad S_N = 0 \)

(ii) maximally mixed state \( \rho = \frac{1}{N} I_N \quad S_N = \log N \)

(iii) \( 0 \leq S_N \leq \log (\text{dim} \mathcal{H}) \) \hspace{1cm} (6)

(iv) \( S_N \) is concave \( \lambda_1, \lambda_2, \ldots, \lambda_n \geq 0 \quad \sum \lambda_i = 1 \)

\[
S_N(\rho) = \sum_{i} \lambda_i S_N(\rho_i) \geq \lambda_1 S_N(\rho_1) + \ldots + \lambda_n S_N(\rho_n)
\]

This reflects that if we know the decomposition \( \rho = \sum \lambda_i \rho_i \) of a state we have more a-priori knowledge, i.e. "measurement" of \( \rho \) would yield less information.

(v) \( S_N(U \rho U^{-1}) = S_N(\rho) \) if \( U^+ = U^{-1} \) \hspace{1cm} (8)

(vi) subadditivity

\[
S_N(\rho_{AB}) \leq \frac{1}{N} S_N(\rho_A) + S_N(\rho_B) \hspace{1cm} (9)
\]

where \( \rho_A = \text{Tr}_B \{ \rho_{AB} \} ; \quad \rho_B = \text{Tr}_A \{ \rho_{AB} \} \)

if \( \rho \) factorizes, i.e. if \( \rho_{AB} = \rho_A \cdot \rho_B \)

\[
S_N(\rho_A \cdot \rho_B) = S_N(\rho_A) + S_N(\rho_B)
\]

\( (\rho_A, \rho_B \text{ do not contain information about entanglement}) \)
(vii) Strong subadditivity

\[
S_N(S_{ABC}) + S_N(S_A) \leq S_N(S_{AB}) + S_N(S_{AC})
\]

This is non-trivial! It is basis for many formal proofs in quantum information theory.

(1x) Araki-Lieb inequality

\[
S_N(S_{AB}) \geq |S_N(S_A) - S_N(S_B)|
\]

This is in contrast to the classical case!!

where

\[
S_{sh}(x,y) \geq S_{sh}(x), S_{sh}(y)
\]

in q-mechanics it is however possible that

\[
S_P(S_{AB}) < S_N(S_A), S_N(S_B)
\]

even \( S_N(S_{AB}) = 0 \) if \( S_{AB} \) is pure but \( A \) and \( B \) are entangled and thus \( S_N(S_A) = S_N(S_B) = 0 \)

**Example** \( |\psi_+\rangle = \frac{1}{\sqrt{2}} \left( |00\rangle + |11\rangle \right) \)

\[
S_A = Tr_B \{|\psi_+\rangle \langle \psi_+| S = \frac{1}{2} \ \mathbb{I} = S_B \ \ (S_N \rightarrow S)
\]

\[
S(S_A) = S(S_B) = 1 \ \ \text{but} \ \ S(S_{AB}) = 0
\]

- 56 -
smallest Hilbert space with non-zero $S_N \neq \mathbb{C}^2$
(spin $1/2$)
(12) $\mathcal{H} = \{ |0\rangle, |1\rangle \}$

a general qubit state has the form

(13) $S = \frac{1}{2} ( 1 + \vec{s} \cdot \vec{s} )$

and $\vec{s} = \mathbb{R}^3$ with $| \vec{s} | \leq 1$ called polarization vector

obviously $S_z = \text{Tr} \{ S \vec{s}_z \}$, diagonalizing $\mathcal{H}_z$ yields

(14) $S = \frac{1}{2} (1 + |s_z|^2) |S_z\rangle\langle S_z| + \frac{1}{2} (1 - |s_z|^2) |1\rangle\langle 1| + |0\rangle\langle 0|

representation on the Bloch sphere

- pure states $|s| = 1$
- max. mixed state $|s| = 0$
4.2 Quantum information transfer

There is a generalization of Shannon's noiseless coding theorem to qubit

**Pure states:** Schumacher theorem

In order to encode (store) the quantum information of \( n \) copies of a pure state \( \rho = |\psi\rangle \langle \psi| \) a Hilbert space \( \mathcal{H}_0 \) is required in the limit \( n \to \infty \) with

\[
\log (\dim \mathcal{H}_0) = n S_N(\rho)
\]

Thus to transfer \( n \) qubits each in state \( \rho \) over a lossless channel \( n S_N(\rho) \) qubits need to be transferred.

**Example** Alice wants to send to Bob a qubit word of length 3

\[
|2\psi\rangle = |2x\rangle |2y\rangle |2z\rangle \quad |2\psi_1\rangle = \begin{cases} 11x] \vspace{0.25cm} \\
11z\end{cases}
\]

Can be done most efficiently using only 4 states

\[
\{1000\rangle, 1100\rangle, 1010\rangle, 1001\rangle\}
\]

Instead of the \( 2^3 = 8 \) possible combinations of \( 11x\rangle \) and \( 11z\rangle \)
one can easily calculate

\[ \left| \langle 000 | \psi \rangle \right|^2 = \left| \cos^2 \left( \frac{\pi}{3} \right) \right|^2 = 0.6219 \]

\[ \left| \langle 001 | \psi \rangle \right|^2 = \left| \langle 010 | \psi \rangle \right|^2 = \left| \langle 001 | \psi \rangle \right|^2 = \left[ \cos^2 \left( \frac{\pi}{3} \right) \right]^2 \sin^2 \frac{2\pi}{3} \]

\[ = 0.1067 \]

i.e. for a random qubit word the probability to find it in the subspace spanned by (17) is

\[ P_{\text{success}} = 0.6219 + 3 \cdot 0.1067 = 0.8419 \]

for \( n \to \infty \), \( P_{\text{success}} \to 1 \)

For mixed states, the generalization is less trivial. Here we introduce an alphabet of orthogonal mixed states \( \hat{\Sigma}_j \):

\[ \text{Tr} \{ \hat{\Sigma}; \hat{\Sigma}_j \} \approx \delta_j \]

If then a general state is decomposed into \( \Sigma_j \):

\[ \hat{\Sigma} = \sum_j p_j \hat{\Sigma}_j \]

\[ \mathcal{S}(\hat{\Sigma}) = -\text{Tr} \{ \hat{\Sigma} \log \hat{\Sigma} \} = -\sum_j \text{Tr} \left\{ p_j \hat{\Sigma}_j \log (p_j \hat{\Sigma}_j) \right\} \]

\[ = -\sum_j p_j \log p_j \]

\[ \text{classical information in } p_j \]

\[ \text{from mixedness of alphabet} \]
Holevo information

\[
X(\rho) = S(\rho) + \sum_j p_j \text{Tr} \left\{ \rho_j \log \rho_j \right\} \\
= S(\rho) - \sum_j p_j S(\rho_j)
\]

Holevo conjecture

To transfer a \( n \) qubit-word of mixed states \( \rho_j \)
which appear with probability \( p_j \),

\[ n X(\rho_j; \hat{\rho}_j) \leq n S(\rho) \quad n \to \infty \]

qubits are needed.

4.3 Quantum error correction

Many ideas of classical error correction can be transformed, however there are two specialities

(i) qubits have more than bit flip errors

(ii) the state of a qubit should not be measured during transfer
**qubit errors**

result from coupling of qubit $|0\rangle, |1\rangle$ to environment $|E\rangle$

- **Spin flip**
  \[
  (24) \quad
  \begin{align*}
  |0\rangle |E\rangle & \rightarrow |1\rangle |E_{01}\rangle, \\
  |1\rangle |E\rangle & \rightarrow |0\rangle |E_{10}\rangle.
  \end{align*}
  \]

- **Dephasing**
  \[
  (25) \quad
  \begin{align*}
  |0\rangle |E\rangle & \rightarrow |0\rangle |E_{00}\rangle, \\
  |1\rangle |E\rangle & \rightarrow |1\rangle |E_{11}\rangle.
  \end{align*}
  \]

In the course of interaction the environment states become more and more orthogonal:

- **Dephasing**
  \[
  (26) \quad
  \begin{align*}
  \langle E_{01}(t) | E_{10}(t) \rangle & \rightarrow 0, \\
  \langle E_{00}(t) | E_{11}(t) \rangle & \rightarrow 0.
  \end{align*}
  \]

This gives to the following evolution of the reduced density matrix starting from $|\psi\rangle = |\alpha\rangle + \beta |1\rangle$

\[
S_0 = \begin{pmatrix}
|\alpha|^2 & \alpha \beta^* \\
\alpha^* \beta & |\beta|^2
\end{pmatrix}
\quad
\rightarrow
\quad
\begin{pmatrix}
|\beta|^2 & \alpha \beta^* \langle E_{01}(t) | E_{10}(t) \rangle \\
\alpha^* \beta \langle E_{00}(t) | E_{11}(t) \rangle & |\alpha|^2
\end{pmatrix}
\]

\[
(27) \quad
\begin{align*}
\text{spin flip} \quad & \rightarrow \\
\text{dephasing} & \rightarrow
\end{align*}
\]

in both cases non-diagonal terms decay

-61- "coherence"
in general

\( 10 \rangle |E^\rangle \xrightarrow{U(t)} 10 \rangle |E_{00}\rangle + 11 \rangle |E_{01}\rangle \)

\( 10 \rangle |E_{10}\rangle + 11 \rangle |E_{11}\rangle \)

this can be written as:

\( (\alpha |10\rangle + \beta |11\rangle) |E_0\rangle \)

\( + (\alpha |11\rangle + \beta |10\rangle) |E_x\rangle \)

\( + (\alpha |10\rangle - \beta |11\rangle) |E_z\rangle \)

\( + (-\alpha |11\rangle + \beta |10\rangle) |E_y\rangle \)

\( = (\hat{1} |12_0\rangle |E_0\rangle + (\hat{X} |12_0\rangle |E_x\rangle + (\hat{Y} |12_0\rangle |E_y\rangle + (\hat{Z} |12_0\rangle |E_z\rangle \)

where

\( \hat{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \hat{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \hat{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \)

\( \hat{X} = \text{spin flip} \quad \hat{Y} = \text{both} \quad \hat{Z} = \text{dephasing} \)

\( \hat{X} = \hat{Z} \hat{X} \)

\( \Rightarrow \) evov correcting code should correct in addition to spin flip X also dephasing Z and combination ZX

3-qubit code corrects spin flip X

evov operation

\( \hat{M} = (1-\rho) \hat{I} + \rho \hat{X} \)
this is just the classical repetition code \( k=1, n=3, d=3 \)

\[(34) \quad |0\rangle \rightarrow |0,0,0\rangle \quad |1\rangle \rightarrow |1,1,1\rangle\]

*how to create code without measurement of state \( |z\rangle = \alpha |0\rangle + \beta |1\rangle \), i.e. without knowing \( \alpha, \beta \)?

**CNOT operation (unitary)**

\[(35) \quad |z\rangle|\Phi\rangle \quad \rightarrow \quad \text{CNOT} |z\rangle|\tilde{\Phi}\rangle \left\{ \begin{array}{l}
\text{leaves } |z\rangle \\
\text{unchanged}
\end{array} \right.\]

\[(36) \quad |0\rangle|0\rangle \otimes I + |1\rangle|1\rangle \otimes X \quad \Rightarrow \quad \text{operation on target bit}
\]

\[\text{projector to basis of control bit}\]

\[\text{if } |x\rangle, |y\rangle \in \{0,1\}^3\]

\[(37) \quad \text{CNOT } |x\rangle|y\rangle = |x\rangle (y+ \text{X mod} 2) \]

entanglement of control+target

\[(38) \quad |00\rangle \rightarrow |100\rangle \\
|01\rangle \rightarrow |101\rangle \\
|10\rangle \rightarrow |111\rangle \\
|11\rangle \rightarrow |1\rangle \]

---

-63-
(i) encoding in 3-qubit code

\[ (\alpha|000\rangle + \beta|111\rangle) 10\rangle 10\rangle \rightarrow \alpha|1000\rangle + \beta|111\rangle \]

2 x CNOT

\[
\begin{array}{c}
|10\rangle \\
|0\rangle \\
|0\rangle
\end{array}
\]

\[
\begin{array}{c}
\text{code}
\end{array}
\]

(ii) interaction with environment

\[ \alpha|1000\rangle + \beta|111\rangle \]

\[
\begin{array}{c}
\alpha|1000\rangle + \beta|111\rangle \\
\alpha|1000\rangle + \beta|101\rangle \\
\alpha|0001\rangle + \beta|1110\rangle \\
\text{other}
\end{array}
\]

\[ \text{probability} \]

\[
\begin{array}{c}
(1-p)^3 \\
p(1-p)^2 \\
p(1-p)^2 \\
\Theta(p^2) \text{ small!}
\end{array}
\]

(iii) error detection

\[ k=7, n=3, d=3 \]

\Rightarrow \text{parity check matrix} \quad n-k=2 \times n=3

\[
H = \begin{pmatrix}
1 & 1 & 0 \\
1 & 0 & 1
\end{pmatrix}
\]

important property

\[ H \begin{pmatrix} 0 \\ 0 \end{pmatrix} = H \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0 \]
This can be used to detect error without measurement of code $\mathcal{C}$.

\[
\begin{align*}
\text{code} \quad \{ &\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\text{measurement of error syndrome}
\end{array}
\end{array}
\end{array}
\end{array}
\end{array}
\end{array}
\end{array}
\end{align*}
\]

\text{measurement of error syndrome}

$X$ error in qubit $j$: $\hat{X}_j$

e.g. $\hat{X}_2 |x, y, z\rangle = |x, \bar{y}, y\rangle = |x, 0 \oplus y, y\rangle$

$\hat{X}_2 |x, 0, y\rangle = |x, 0, y\rangle = |x, \bar{y}, y\rangle$

i.e. in general

\[
\hat{X}_2 |x, y, z\rangle = |x, y \oplus z, z\rangle = |x y z \oplus 010\rangle
\]

$\Rightarrow$ classical error

\[
\begin{align*}
\begin{array}{c}
\text{all errors in the code (40)}
\end{array}
\end{align*}
\]

up to 1 bit flip can be detected

\[
\Rightarrow \quad \begin{pmatrix} \alpha |100\rangle + \beta |111\rangle \end{pmatrix} |100\rangle
\]

$\text{error syndrome}$
(iv) evan cromesion

\[
\begin{align*}
|00\rangle & \Rightarrow \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\
|10\rangle & \Rightarrow \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
|01\rangle & \Rightarrow \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
|11\rangle & \Rightarrow \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}
\end{align*}
\]

**Hamming inequality**

A linear code for \( k \) qubits which can correct 3 independent errors \((X, Y, Z)\) in \( n \) qubits has to have at least \( n \) qubits

\[
2^k \sum_{i=0}^{k} \frac{k}{i} \leq 2^n
\]

The smallest code that can correct all possible errors in one qubit is therefore \((k=1, n=7)\)

\[
2 \cdot 3^7 = 6n \leq 2^n
\]

\[\Rightarrow n=5\text{ or larger}\]

Such a code can be constructed, it is however not very handy.
a very important class of codes are the so-called stabilizer codes

they make use of the fact that the possible errors form a group

\[ \pm \{ I, X, Y, Z \} \]

one can show that "good" stabilizer codes exist when good means

\[ [n, k, d = 2t + 1] \]

\[ R = \frac{k}{n} \quad \rightarrow \quad \text{finite} \]

\[ P = \frac{t}{n} \quad \rightarrow \quad \text{finite} \]

(max. error probability per qubit)

one can show that good codes exist for \[ P < 0.0946 \]

\underline{Remark:} All of the discussion assumes that the coding, error detection and error correction steps themselves do not contribute errors. To include the latter, one has to add additional encodings \( \Rightarrow \) fault tolerant coding and error correction.

one can show that this is possible if \[ P \leq 1 \times 10^{-6} \]

\[ P \leq 1 \times 10^{-6} \]