Lecture 2: two-qubit system

Xiongfeng Ma[1,*]

[1] *Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China*

In the last chapter, we have learned about one qubit system. Here, we shall learn about a two-qubit system. It turns out two-qubit system has much more fun. It will take three weeks for us to cover these materials. Guess how many qubits we can cover by the end of the course?

Two-qubit state, mixed state, Bloch "ball", positive-operator valued measure (POVM); Super operator, purification of mixed state and POVM; Using teleportation for operation (Gottesman-Chuang'99), remote state preparation; Bell's inequality, CHSH/CH/Eberhard inequality; experiment development and loopholes, entanglement; Quantum dense coding, teleportation (experiment development);

* xma@tsinghua.edu.cn

I. REVIEW: ONE-QUBIT SYSTEM

1. state: ray.
2. normalize state: vector.
3. representation way: $|u\rangle = \cos\frac{\theta}{2} + e^{i\varphi}\sin\frac{\theta}{2}$. (Bloch Sphere[1])
4. density matrix: $\rho = \sum_i \lambda_i |\phi_i\rangle\langle\phi_i|$.[2]
5. $\rho^+ = \rho, Tr(\rho) = 1$.
6. $\forall|\phi\rangle, \langle\phi|\rho|\phi\rangle \geq 0$.[3]
7. For pure qubit, Assume $\rho = |\phi\rangle\langle\phi|$, we have $\rho^2 = \rho$. Also, we can get the eigenvalues and eigenvectors of ρ (Assume $|\Phi\rangle$ is orthogonal of $|\phi\rangle$):
$$\begin{aligned} \rho|\phi\rangle &= |\phi\rangle\langle\phi|\phi\rangle = |\phi\rangle. \\ \rho|\Phi\rangle &= |\phi\rangle\langle\phi|\Phi\rangle = 0. \end{aligned} \tag{1}$$
8. $\langle\psi|M|\psi\rangle = Tr(\langle\psi|M|\psi\rangle) = Tr(M|\psi\rangle\langle\psi|) = Tr(M\rho)$.[4]

A. Bloch sphere

A useful representation of the state of a single qubit is the Bloch sphere representation. Since the overall phase is irrelevant, a pure state of a qubit can be written as

$$|u\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle. \tag{2}$$

Therefore, it is convenient to represent it as a vector living the surface of a unit sphere with the spherical coordinate $(r = 1, \theta, \varphi)$.

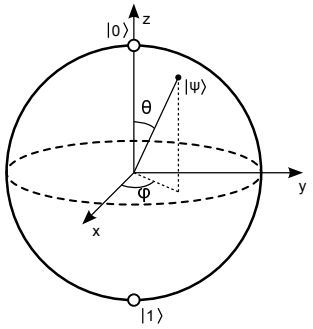

FIG. 1. Bloch sphere.

In general, a qubit might be in a mixed state, and then it will be within the Bloch sphere instead on the surface. In general, the density matrix of a qubit can be written as

$$\begin{aligned} \rho &= \frac{1}{2}(\sigma_0 + \vec{P}\cdot\vec{\sigma}) \\ &= \frac{1}{2}\begin{pmatrix} 1+P_z & P_x - iP_y \\ P_x + iP_y & 1-P_z \end{pmatrix} \end{aligned} \tag{3}$$

where $\vec{P} = (P_x, P_y, P_z)$ is a vector and $|\vec{P}| \leq 1$. When $|\vec{P}| = 1$, Eq. (3) represents a pure qubit.

---

[1] No standard way to visualize two or more qubits system until now.
[2] why we can't distinguish $|u\rangle$ and $e^{i\theta}|u\rangle$? because $|u\rangle\langle u| = e^{i\theta}|u\rangle\langle u|e^{-i\theta}$. Their density matrices are always the same.
[3] By physical meaning of the measurement
[4] $Tr(AB) = \sum_i\sum_j a_{ij}b_{ji} = \sum_i\sum_j b_{ij}a_{ji} = Tr(BA)$.

II. NOTATIONS

Matrix tensor product $\otimes$, the tensor product of two matrices is

$$\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \otimes \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} = \begin{bmatrix} a_{1,1} \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} & a_{1,2} \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} \\ a_{2,1} \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} & a_{2,2} \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a_{1,1}b_{1,1} & a_{1,1}b_{1,2} & a_{1,2}b_{1,1} & a_{1,2}b_{1,2} \\ a_{1,1}b_{2,1} & a_{1,1}b_{2,2} & a_{1,2}b_{2,1} & a_{1,2}b_{2,2} \\ a_{2,1}b_{1,1} & a_{2,1}b_{1,2} & a_{2,2}b_{1,1} & a_{2,2}b_{1,2} \\ a_{2,1}b_{2,1} & a_{2,1}b_{2,2} & a_{2,2}b_{2,1} & a_{2,2}b_{2,2} \end{bmatrix}. \tag{4}$$

Here, we simplify denote it as

$$|\phi\rangle \otimes |\psi\rangle = |\phi\rangle|\psi\rangle \tag{5}$$

A. Trace

Trace $Tr(\rho)$: summation of the diagonal terms.

Partial trace $Tr_B(\rho_{AB})$: Assume A is n-dimension state, B is m-dimension state, then ρ_{AB} can be regarded as a $nm \times nm$ matrix. $Tr_B(\rho_{AB})$ is a $n \times n$ matrix, where

$$(Tr_B(\rho_{AB}))_{ij} = \sum_{k=1}^{m} (\rho_{AB})_{ik,jk} \tag{6}$$

Particularly, $Tr_B(A \otimes B) = Tr(B)A$.

III. TWO-QUBIT SYSTEM

A. an interesting "paradox"

When considering a two qubits state which is written as

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B). \tag{7}$$

When measure the system A in the Z basis, with probability 1/2, the measurement result is $|0\rangle$ and the prepared state is $|0\rangle_A|0\rangle_B$. With probability 1/2, the measurement result is $|1\rangle$ and the prepared state is $|1\rangle_A|1\rangle_B$. Naively, one might express system A in the state of

$$|\psi\rangle_A = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |+\rangle. \tag{8}$$

Consider another two-qubit state,

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|+\rangle_A|+\rangle_B + |-\rangle_A|-\rangle_B). \tag{9}$$

Following the same naive argument, we should have

$$|\psi\rangle_A = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) = |0\rangle. \tag{10}$$

But, on the other hand,

$$\begin{aligned} |\psi\rangle_{AB} &= \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B) \\ &= \frac{1}{2\sqrt{2}}(|+\rangle + |-\rangle)(|+\rangle + |-\rangle) + \frac{1}{2\sqrt{2}}(|+\rangle - |-\rangle)(|+\rangle - |-\rangle) \\ &= \frac{1}{\sqrt{2}}(|+\rangle_A|+\rangle_B + |-\rangle_A|-\rangle_B). \end{aligned} \tag{11}$$

So now we have: $|0\rangle = |+\rangle$. What is wrong?

Explanation: The key of the above paradox is that $|\psi\rangle_A$ is not a pure state any more, we only can write it as a density matrix. The density operator of subsystem A is given by

$$\rho_A = tr_B(|\psi\rangle_{AB}\langle\psi|_{AB}) = \frac{1}{\sqrt{2}}(|0\rangle\langle 0| + |1\rangle\langle 1|). \tag{12}$$

B. EPR paradox

Alice and Bob share a two qubits system:

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B).$$

Alice gets qubit A, Bob gets qubit B. Then Alice goes to a planet which far from Bob. Then if Alice wants to tell 0 to B, she measures A in $|+/-\rangle$ basis; if she wants to tell 1 to B, she measures A with $|0/1\rangle$ basis.

We assume that after measurement for A, Alice gets $|0\rangle$. Then for Bob, he measures B after a while with basis $|0/1\rangle$, if he gets $|0\rangle$ with probability 1, he can say Alice wants to tell him 0.

So by this process, we transform information which is faster than light. What's the problem?

Explanation: For Bob, if he gets $|0\rangle$, he doesn't know which basis Alice uses. Because if Alice chooses $|0/1\rangle$ basis, Bob has 0.5 probability to get $|0\rangle$; if Alice chooses $|+/-\rangle$ basis, Bob also has 0.5 probability to get $|0\rangle$.

This experiments tells us, quantum has the **non-locality** and also **no-signalling** properties.

Something are changed, but we do not know. We simply use the same state to describe the system.

C. Density matrix of subsystem

For a bipartite system ρ_{AB}, the density matrix of subsystem A can be denoted as $Tr_B(\rho_{AB})$. For example, $|\psi\rangle_{AB} = a|0\rangle_A|0\rangle_B + b|1\rangle_A|1\rangle_B$. Then $\rho_A = \mathrm{Tr}_B(|\psi\rangle_{AB}\langle\psi|_{AB}) = aa^*|0\rangle\langle 0| + bb^*|1\rangle\langle 1|$. In general, there are some properties of ρ_A:

- $\rho_A^\dagger = \rho_A$.
- $\rho_A \geq 0$.
- $\mathrm{Tr}(\rho_A) = 1$.

Comment: If ρ_A is pure state, then $\rho_A^2 = \rho_A$, the purity $\mathrm{Tr}(\rho_A^2) \leq 1 \Rightarrow \mathrm{Tr}(\rho_A^2) = 1$.

An example $|\psi\rangle_{AB} = |0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B = |+\rangle_A|+\rangle_B + |-\rangle_A|-\rangle_B$ has two representations. There are two cases to measure $\rho_A = \frac{1}{2}\mathbf{I}$.

Case 1: How to get $\rho_A = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$?

1. Prepare $|0\rangle|0\rangle + |1\rangle|1\rangle$
2. Measure B in $z = 0$ and $z = 1$
3. Given z-measure result. $A = |0\rangle$ or $|1\rangle$.
4. $\rho_{A1} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$.

Case 2: How to get $\rho_A = \frac{1}{2}(|+\rangle\langle +| + |-\rangle\langle -|)$?

1. Prepare $|+\rangle|+\rangle + |-\rangle|-\rangle$
2. measure B in $x = |+\rangle$ and $x = |-\rangle$.
3. Given x-measure results, $A = |+\rangle$ or $|-\rangle$
4. $\rho_{A2} = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|$.

$\rho_{A1} = \rho_{A2}$. From Alice's point of view, they are the same, although Bob knows which of the two pure states it is. When Bob tells Alice the information of the state, the states collapses. (Information is physical; physics is informational.)

Coherence (classical mixture or superposition). $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |\psi\rangle\langle\psi| = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. If we measure in x we always get the same. But it is different for $\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$ (which can be distinguished by measuring in z).

We prefer the pure state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ to generated randomness, because additional information can not change the state, while the other one can be attacked by measuring the entangled qubit.

Comment: Given a pure state, we know every information in the system and can not change it by getting more information, since the entropy is already zero (not the case for general density matrix).

D. The properties for general density matrix

In general, the state is represented by a density operator. In the case where the state of the subsystem is a ray, and we say that the state is pure. Otherwise the state is mixed. If $\rho_A = \rho_A^2$, then $|\psi_A\rangle$ is a pure state. Otherwise, the density matrix of A is $\rho_A = \sum_a p_a |a\rangle\langle a|$, where $\sum_a p_a = 1$, $0 < p_a < 1$. The trace distance $tr\rho_A^2 = \sum_a p_a^2 < \sum_a p_a = 1$.

Properties of a general density matrix

1. self-adjoint: $\rho_A = \rho_A^\dagger$
2. positivity: $\forall|\psi\rangle$, $\langle\psi|\rho_A|\psi\rangle \geq 0$
3. completeness: $Tr(\rho_A) = 1$

E. Schmidt decomposition

For any pure state $|\psi\rangle_{AB}$ of a bipartite system, there are orthonormal bases $\{|i\rangle_A\}$ and $\{|i\rangle_B\}$ such that:

$$|\psi\rangle_{AB} = \sum_i \sqrt{p_i}|i\rangle_A \otimes |i'\rangle_B. \tag{13}$$

The subsystems A and B have the same eigenvalues, p_is. The number of p_is is called the Schmidt number of $|\psi\rangle_{AB}$. We denote that the pure state is a entangled state when the the Schmidt number is greater than one. It is easy to see that the Bell states are entangled states.

For pure $|\psi\rangle_{AB}x$, you can always find $\rho_A = \rho_B$ in any basis of ρ_A.

F. Bell basis

The dimension of the 2-qubit Hilbert space is 4. Thus, there are 4 basis states. Bell state basis is widely used, especially for the case involving entanglement. The 4 Bell states in the Z basis are,

$$\begin{aligned} \Phi^+ &= |00\rangle + |11\rangle \\ \Phi^- &= |00\rangle - |11\rangle \\ \Psi^+ &= |01\rangle + |10\rangle \\ \Psi^- &= |01\rangle - |10\rangle \end{aligned} \tag{14}$$

in the X basis are

$$\begin{aligned} \Phi^+ &= |++\rangle + |--\rangle \\ \Phi^- &= |-+\rangle + |+-\rangle \\ \Psi^+ &= |++\rangle - |--\rangle \\ \Psi^- &= |-+\rangle - |+-\rangle \end{aligned} \tag{15}$$

in the Y basis are

$$\begin{aligned}\Phi^+ &= |+i-i\rangle + |-i+i\rangle \\ \Phi^- &= |+i+i\rangle + |-i-i\rangle \\ \Psi^+ &= -i(|+i+i\rangle - |-i-i\rangle) \\ \Psi^- &= i(|+i-i\rangle - |-i+i\rangle)\end{aligned} \tag{16}$$

Many interesting simple quantum information phenomenons come with Bell states, such as Bell's inequality, Teleportation, super dense coding, quantum key distribution, and Deutsch's algorithm.

G. Qubit tomography

Quantum state tomography is the process of reconstructing the quantum state for a quantum system by proper measurements. The quantum state can be pure (vector) or in general mixed (density matrix). A set of measurements is called tomographically complete if it can uniquely identify the state. That is, the measurement outcomes are able to provide all the information about the state. In the classical physics, it corresponds to system calibration.

Let us take qubit tomography for example. As given in Eq. (3), the Z basis measure provides the information on P_z. By changing the Z and X basis, by the Hadamard transformation, we can conclude that the X basis measure provides the information on P_x. Similarly, the Y basis measure provides the information on P_y. Thus, X, Y, and Z measurements is tomographically complete for a qubit. Another way to put this is,

$$\rho = \frac{1}{2}[tr(\rho)\sigma_0 + tr(\sigma_x\rho)\sigma_x + tr(\sigma_y\rho)\sigma_y + tr(\sigma_z\rho)\sigma_z]. \tag{17}$$

Now, let us move a bit further, tomography for n qubits,

$$\rho = 2^{-n} \sum_{v_1,v_2,\ldots,v_n \in \{0,x,y,z\}} tr(\sigma_{v1} \otimes \sigma_{v2} \otimes \cdots \otimes \sigma_{vn}\rho)\sigma_{v1} \otimes \sigma_{v2} \otimes \cdots \otimes \sigma_{vn}, \tag{18}$$

where the sum is over all possible the identity and Pauli matrices.

In another related concept, quantum process tomography, known quantum states are used to probe a quantum process to find out how the process can be described. Similarly, quantum measurement tomography works to find out what measurement is being performed.

The general principle behind quantum state tomography is that by repeatedly performing many different measurements on quantum systems described by identical density matrices, frequency counts can be used to infer probabilities, and these probabilities are combined with Born's rule to determine a density matrix which fits the best with the observations.

IV. POSITIVE OPERATOR-VALUED MEASURE

For the bipartite state $|\psi\rangle_{AB}$,

$$|\psi\rangle_{AB} = a|0\rangle_A|0\rangle_B + b|1\rangle_A|1\rangle_B, \tag{19}$$

if we only measure system A, the observable can be expressed as

$$M_A \otimes I_B \tag{20}$$

where M_A is a self-adjoint operator acting on system A. Then the expectation value of the measurement outcome is,

$$\begin{aligned}\langle\psi|M_A \otimes I_B|\psi\rangle &= (a^*\langle 0|_A\langle 0|_B + b^*\langle 1|_A\langle 1|_B)(M_A \otimes I_B)(a|0\rangle_A|0\rangle_B + b|1\rangle_A|1\rangle_B) \\ &= |a|^2\langle 0|M_A|0\rangle + |b|^2\langle 1|M_A|1\rangle.\end{aligned} \tag{21}$$

Since we know that the density matrix of $\rho_A = |a|^2|0\rangle\langle 0| + |b|^2|1\rangle\langle 1|$, Eq. (21) can be written as,

$$\langle M_A\rangle = tr(M_A\rho_A) \tag{22}$$

We can understands the difference between the Positive Operator-Valued Measure (POVM) and PVM with the same sense that a density matrix is to a pure state. The POVM can be used to describe the effect of PVM acts on a large system. The operators $\{E_a\}$ form a complete set of Hermitian nonnegative operators and satisfy that:

1. Hermiticity: $E_a = E_a^\dagger$.

2. Positivity :$\langle\psi|E_a|\psi\rangle \geq 0$.

3. Completeness: $\sum_a E_a = I$.

A. State tomography

How to measure $\rho = \frac{1}{2} = (I + \vec{P}\cdot\vec{\sigma})$?

Measure states by PVM: $|\psi\rangle_{AB} = a|00\rangle + b|11\rangle$. $M_A \otimes I$. The expectation

$$\begin{aligned}\langle M_A \otimes I\rangle &= \mathrm{Tr}(M_A \otimes I|\psi\rangle\langle\psi|) \\ &= {}_{AB}\langle\psi|M_A \otimes I_B|\psi\rangle_{AB} \\ &= |a|^2\langle 0|M_A|0\rangle + |b|^2\langle 1|M_A|1\rangle = \mathrm{Tr}(\rho_A M_A)\end{aligned}$$

Measure density matrix by POVM: Expectation $\langle M\rangle = \mathrm{Tr}(M\rho)$.
$\{E_a\}$, (1) $E_a = E_a^\dagger$; (2) $\forall|\psi\rangle, \langle\psi|E_a|\psi\rangle \geq 0$;(3) $\sum_a E_a = 1$.
The probability to get a is $\mathrm{Tr}(E_a\rho)$. e.g. PVM, qubit, z-basis, $a = 0, 1$. $E_0 = |0\rangle\langle 0|, E_1 = |1\rangle\langle 1|$.
PVM is the special case of POVM.

B. Measurement

Many copies (the system reproduces the same state).
Measure in x-basis: $+, -$, $\Pr(\pm) = \mathrm{Tr}(\rho E_\pm) = \langle\pm|\rho|\pm\rangle = \frac{1}{2}(1 \pm P_x)$ (since $|\pm\rangle$ is the eigen state of σ_x and σ_x is orthogonal to both σ_y and σ_z).
$\Pr(\pm i) = \frac{1}{2}(1 \pm P_y), \Pr(0/1) = \frac{1}{2}(1 \pm P_z)$.
How to measure arbitrary many qubit?
$\rho = \sum_x p(x)|x\rangle\langle\otimes|\ f(x)$, e.g. $\frac{1}{2}[|0\rangle\langle\otimes|\ \rho_0 + |1\rangle\langle\otimes|\ \rho_1]$.
Measure x, y, z for $\rho = \frac{1}{2}(I + \vec{P}\cdot\vec{\sigma})$. $P_i = \mathrm{Tr}(\rho\sigma_i)(i = x, y, z)$.

$$\rho = \frac{1}{2}\left(\mathrm{Tr}(\rho I)I + \sum_{i=x,y,z}\mathrm{Tr}(\sigma_i\rho)\sigma_i\right)$$

two-qubit system. Use $\sigma_i \otimes \sigma$ to measure ρ_{AB}.
For $\frac{1}{\sqrt{2}}(|00\rangle_{AB} \pm |11\rangle_{AB})$, $\frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$ (can not be written as tensor product),

$\rho_A = \frac{1}{2}[10; 01] = \rho_B, \rho_{AB} = \rho_A \otimes \rho_B = \begin{pmatrix} & & & \\ & & & \\ & & & \\ & & & \end{pmatrix}_{4\times 4} = \sum_{i,j\in\{I,x,y,z\}} P_{ij}\sigma_i \otimes \sigma_j$.

Beccause $\mathrm{Tr}\rho_{AB} = \sum_{i,j} P_{ij}\mathrm{Tr}\sigma_i \otimes \sigma_j = P_{II}\mathrm{Tr}(\sigma_I)\mathrm{Tr}(\sigma_I) = 4P_{II} = 1$, $P_{II} = \frac{1}{4}$.
$P_{ij} = \mathrm{Tr}(\sigma_{i'j'}P_{i'j'}(\sigma_{i'} \otimes \sigma_{j'})(\sigma_i \otimes \sigma_j)) = \mathrm{Tr}(\rho\sigma_i \otimes \sigma_j)$.

n-qubit system. $\bigotimes_{v_i=\in\{I,x,y,z\}}\sigma_{v_i}, \rho = \sigma_{v_i}\mathrm{Tr}(\rho\bigotimes_{v_i}\sigma_{v_i})\bigotimes(\sigma_{v_i})$. 3^n measurement should be performed, but the DOF is less, so this tomography is not optimal.

Individual measurement: 4^n terms (nothing for I).

Joint measurement (BSM: Bell-state measurement): $\phi^{\pm} = |00\rangle \pm |11\rangle, \psi^{\pm} = |01\rangle \pm |10\rangle$

Hamard $H = \frac{1}{\sqrt{2}}[1, 1; 1, -1]$ operation, C-NOT operation (control-NOT, flip target qubit when input control qubit $|1\rangle$, e.g. $|1\rangle|0\rangle \mapsto |1\rangle|1\rangle$)) is 4×4 unitary matrix:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

$|00\rangle + |11\rangle \overrightarrow{C-NOT} |+\rangle|0\rangle \overrightarrow{H \times I} |0\rangle|0\rangle$,
$|00\rangle - |11\rangle \overrightarrow{C-NOT, H \otimes I} |1\rangle|0\rangle$,
$|01\rangle + |10\rangle \mapsto |0\rangle|1\rangle, |01\rangle - |10\rangle \mapsto |1\rangle|1\rangle$,
$|00\rangle \mapsto |+\rangle|0\rangle$.
Then measure in z-basis.

V. QUANTUM CHANNEL

The quantum channel is also called "super operator". The super means that the map takes operators to operators, rather than vectors to vectors. Unitary evolution on $\mathcal{H}_A \otimes \mathcal{H}_B$ will not in general appear to be unitary if we restrict our attention to $\mathcal{H}_A$ alone. Rather, evolution in HA will be described by a quantum channel, (which can be inverted by another channel only if unitary). A general channel $\mathcal{E}$ has an operator-sum representation:

$$\begin{aligned} \mathcal{E}(\rho) &= \sum_a M_a \rho M_a^{\dagger}, \\ \sum_a M_a^{\dagger} M_a &= I. \end{aligned} \tag{23}$$

VI. ENSEMBLES AND PURIFICATION

A mixed state of a system A can be prepared as an ensemble of pure states in many different ways, all of which are experimentally indistinguishable if we observe system A alone.

$$\rho_A = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{24}$$

with $\sum p_i = 1$.

For any such ρ_A, we can construct a "purification" of ρ_A, $|\Psi_1\rangle_{AB}$,

$$|\Psi_1\rangle_{AB} = \sum_i \sqrt{p_i} |\psi_i\rangle_A |\alpha_i\rangle_B, \tag{25}$$

where $\{|\alpha_i\rangle_B\}$ are mutually orthogonal and normalized.

The relation between two purifications, $|\Psi_1\rangle_{AB}$ and $|\Psi_2\rangle_{AB}$, is given by,

$$|\Psi_1\rangle_{AB} = (I_A \otimes U_B)|\Psi_2\rangle_{AB}, \tag{26}$$

the two states differ by an unitary change of basis acting in $\mathcal{H}_B$ alone.

Given $\rho_A = \frac{1}{2}(|0\rangle\langle +| \ |1\rangle\langle|)$, find Ψ_{AB} s.t. $\mathrm{Tr}_{AB} = \rho_A$. Answer: $|00\rangle + |11\rangle$, not unique, but all purification can be linked by linear operation. e.g. $\begin{pmatrix} \frac{4}{9} & \frac{1}{100} \\ \frac{1}{100} & \frac{5}{9} \end{pmatrix}$

For diagonal terms, $\sqrt{\frac{4}{9}}|0\rangle|0\rangle + \sqrt{\frac{5}{9}}|1\rangle$. For general matrix, extends ρ_A to pure ρ_{AB}. Choose a basis such that $|\psi_{AB}\rangle = \sum \sqrt{x_i}|i\rangle_A |i\rangle_B$ is diagonalized.

Denote eigenvalues of ρ as λ_0, λ_1, eigen states $|\psi_0\rangle, |\psi_1\rangle$. $\rho' = \mathrm{diag}(\lambda_0, \lambda_1)$. $\rho = \lambda_0|\psi_0\rangle\langle\psi_0| + \lambda_1|\psi_0\rangle\langle\psi_0|$, then $\phi_{AB} = \lambda_0|\psi_0\rangle_A|0\rangle_B +$
$lambda_1|\psi_1\rangle_A|1\rangle_B$. Comment: Given any mixed states, you can always find a larger Hilbert space such that the density matrix is purified (not unique, but there are relations between such purification).

VII. BELL'S INEQUALITIES

Bell's theorem [**?**]: no physical theory of local hidden variables can ever reproduce all of the predictions of quantum mechanics. Refer to Jhon Preskill notes, Hoi-Kwong Lo notes. Local hidden variables can not express any quantum outcomes. Randomness is in the nature of quantum mechanics.

A. Clauser-Horne-Shimony-Holt inequality

One of the most well known Bell's inequality is the Clauser-Horne-Shimony-Holt (CHSH) inequality [**?**]. There are many ways to express it. We study it from a quantum game point of view.

As shown in Fig. 2, two space-like separated parties, Alice and Bob, randomly choose input bit settings x and y and generate outputs bits a and b based on their inputs and pre-shared quantum (ρ) and classical (λ) resources, respectively. The probability distribution $p(a,b|x,y)$, obtaining outputs a and b conditioned on inputs x and y, are determined by specific strategies of Alice and Bob. By assuming that the input settings x and y are chosen fully randomly and equally likely, the CHSH inequality is defined by a linear combination of the probability distribution $p(a,b|x,y)$ according to

$$S = \sum_{a,b,x,y} (-1)^{a\oplus b+x\cdot y} p(a,b|x,y) \leq S_C = 2, \tag{27}$$

where the plus operation $\oplus$ is modulo 2, $\cdot$ is numerical multiplication, and S_C is the (classical) bound of Bell value S for all LHVMs.

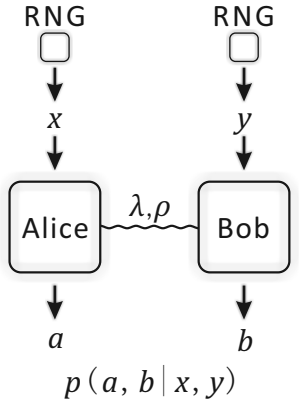

FIG. 2. Bipartite Bell inequality. The inputs of Alice and Bob, x and y, are decided by perfect random number generators (RNGs), which produce uniformly distributed random numbers.

Similarly, there is an achievable bound $S_Q = 2\sqrt{2}$ for the quantum theory [**?**]. In this case, a violation of the classical bound S_C indicates the need for alternative theories other than LHVMs, such as quantum theory. For general no signalling (NS) theories [**?**], denote the corresponded upper bound as $S_{NS} = 4$. It is straightforward to see that $S_{NS} \geq S_Q \geq S_C$.

Different strategies impose different constraints on the probability distribution.

- Classical: $p(a,b|x,y) = \sum_\lambda q(\lambda)p(a|x,\lambda)p(b|y,\lambda)$
- Quantum: $p(a,b|x,y) = \mathrm{Tr}[\rho_{AB} M_a^x \otimes M_b^y]$
- No-signaling: $\sum_a p(a,b|x,y) = \sum_a p(a,b|x',y), \sum_b p(a,b|x,y) = \sum_b p(a,b|x,y')$

B. Experiment loopholes

Since the first experiment in the early 1980s [**?**], lots of lab demonstrations of the CHSH inequality have been presented. In practice, the conclusion of the violation of a Bell test is conditioned on several assumptions. Experimental demonstrations suffer from three major loopholes.

Locality loophole: The measurement events of Alice and Bob should be space-like separated. If this condition is not satisfied, Bell's inequality can be violated even with LHVMs by signaling. This loophole can be closed by separating Alice and Bob sufficient apart such that the measurement events are space-like separated. In experiment, this loophole is closed in optical systems [**?**] and shown to be promising to be closed in atomic systems [**?**].

Efficiency loophole: The detection efficiency must be higher than a threshold to ensure the violation without assuming fair sampling. In the famous CH or Eberhard [**?**] test, it is shown that the efficiency should be at least 2/3 for each party, which is also proved to be a tight bound [**? ?**] for all bipartite Bell test with two inputs. The efficiency loophole can be closed in different realizations [**? ? ?**].

Randomness loophole: The inputs x and y should be random and thus cannot be predetermined. Also, we require x and y to be uncorrelated to each other and also from different runs [**? ?**]. In experiment, this loophole cannot be closed perfectly, as we can never unconditionally certify the randomness without a faithful Bell test, which in turn requires faithful randomness. Thus, we have to assume the existence of a true random seed. In practice, we can make use of independent random number generators, such as causally disconnected cosmic photons [**?**]. On the other hand, if we can characterize the randomness well, we can also check whether the input randomness satisfy the requirement [**? ? ? ? ?**] that guarantee the conclusion even with imperfect randomness input.

C. CH and Eberhard's inequality

In the bipartite scenario, a general Bell test can be defined as

$$J = \sum_{a,b,x,y} c_{a,b,x,y} p(a,b|x,y) - J_C \geq 0. \tag{28}$$

Here, J_C is the classical upper bound and we assume that the probability of choosing x and y is uniform. In practice, a Bell inequality contains both valid outputs such as o e and undetected events u, as shown in Fig. 3. If we take the undetected events into account, the CHSH inequality is defined by the CH or Eberhard's inequality.

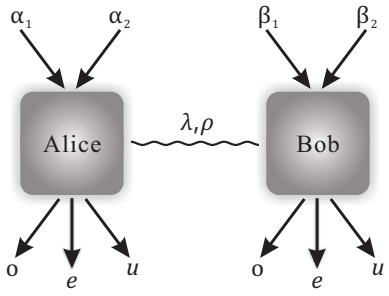

FIG. 3. The CH inequality with two input settings and three possible outputs.

a. CH inequality: The CH inequality is a bipartite Bell inequality with measurement settings $x \in \{\alpha_1, \alpha_2\}$ and $y \in \{\beta_1, \beta_2\}$, and outputs $a, b \in \{o, e, u\}$, corresponding to ordinary, extraordinary, and undetected events, respectively. The CH inequality is defined by a linear combination of the probability distribution $p(a,b|x,y)$ by

$$\begin{aligned} \bar{J} = &-p_{oo}(\alpha_1,\beta_1) - p_{oo}(\alpha_1,\beta_2) - p_{oo}(\alpha_2,\beta_1) \\ &+ p_{oo}(\alpha_2,\beta_2) + p_o^A(\alpha_1) + p_o^B(\beta_1) \geq 0. \end{aligned} \tag{29}$$

Here, we take a more convenient notation of $p(a,b|x,y)$ as $p_{ab}(x,y)$, and $p_o^A(\alpha_1)$ ($p_o^B(\beta_1)$) refers to the probability of obtaining o when Alice's (Bob's) input is α_1 (β_1). It is easy to prove that the CH inequality is satisfied with all LVHMs.

In practice, we have to run many times to sample the probability distribution in Eq. (29). Suppose that the input setting is uniformly random and the number of each input setting is N, then the CH inequality can also be written as

$$\begin{aligned} N\bar{J} \approx &-n_{oo}(\alpha_1,\beta_1) - n_{oo}(\alpha_1,\beta_2) - n_{oo}(\alpha_2,\beta_1) \\ &+ n_{oo}(\alpha_2,\beta_2) + n_o^A(\alpha_1)/2 + n_o^B(\beta_1)/2 \geq 0. \end{aligned} \tag{30}$$

Here, $n_{oo}(x,y)$ denotes coincidence detection of oo with inputs x and y, and $n_o^A(x)$ ($n_o^B(y)$) denotes the single detection o of Alice (Bob) with input x (y). The factor of 2 comes from the change from conditional probability to joint probability. Notice that, Eq. (30) is approximately satisfied for finite samples of N. It becomes the equal sign when N goes to infinity.

b. Eberhard's inequality: In experiment, we can specify the single counts $n_o^A(\alpha_1)$ and $n_o^B(\beta_1)$ to be

$$\begin{aligned} n_o^A(\alpha_1)/2 &= n_{oo}(\alpha_1,\beta_2) + n_{oe}(\alpha_1,\beta_2) + n_{ou}(\alpha_1,\beta_2) \\ n_o^B(\beta_1)/2 &= n_{eo}(\alpha_2,\beta_1) + n_{uo}(\alpha_2,\beta_1) + n_{oo}(\alpha_2,\beta_1) \end{aligned} \tag{31}$$

Then, the CH inequality in finite runs can be expressed by

$$\begin{aligned} J_N = &-n_{oo}(\alpha_1,\beta_1) + n_{oe}(\alpha_1,\beta_2) + n_{ou}(\alpha_1,\beta_2) \\ &+ n_{eo}(\alpha_2,\beta_1) + n_{uo}(\alpha_2,\beta_1) + n_{oo}(\alpha_2,\beta_2) \geq 0, \end{aligned} \tag{32}$$

which is also known as the Eberhard's inequality. In the asymptotical limit, the relation between the CH inequality in Eq. (29) and the Eberhard's inequality in Eq. (32) is

$$\bar{J} = J_N/N. \tag{33}$$

With the no-signaling assumption, we can prove that the CHSH, CH, and Eberhard's inequalities are equivalent.

VIII. QUANTUM TELEPORTATION

The seminal work by C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. K. Wootters is published in 1993 [**?**]. Quantum teleportation is a process by which quantum information can be transmitted (exactly, in principle) from one location to another, with the help of classical communication and previously shared quantum entanglement between the sending and receiving location. Because it depends on classical communication, which can proceed no faster than the speed of light, it cannot be used for superluminal transport or communication of classical bits. It also cannot be used to make copies of a system, as this violates the no-cloning theorem.

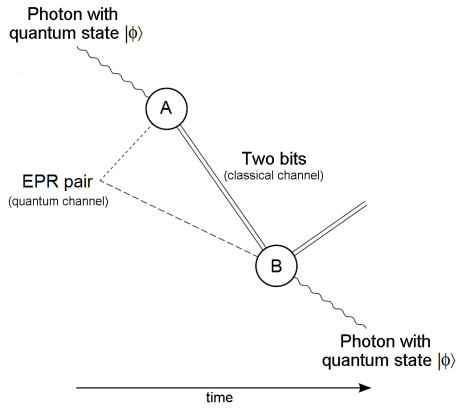

FIG. 4. Schematic diagram for quantum teleportation.

Quantum teleportation can be regarded as a secure way to transfer information [**?**].

IX. QUANTUM SUPER DENSE CODING

The unitary operation $|\psi\rangle_{AB} \to |\phi\rangle_{AB}$ does not necessarily indicate that $\rho_A \to E(\rho_A) = \rho'_A$ is also unitary. The super operation $\mathcal{E}(\rho) = \sum_a M_a \rho M_a^\dagger$, in which $\sum_a M_a M_a^\dagger = I$. (read Preskill lecture notes for more details, related to quantum channel which we will discuss in later lectures).

X. XOR GAMES

Separate boxes (e.g., 16 l.y. away), local referees each give Alice or Bob random variables. $x \in \{0,1\} \to [A] \to a \in \{0,1\}$; $y \in \{0,1\} \to [B] \to b \in \{0,1\}$. In order to win, Alice and Bob need to make sure $a \oplus b = x \cdot y$. No communication in the games, but before the game they can discuss the strategy.

Analysis: $\Pr[x \cdot y = 0] = \frac{3}{4}, \Pr[x \cdot y = 1] = \frac{1}{4}$.

Random strategy can win with probability $\frac{1}{2}$. Always outputs same result, $a \oplus b = 0$, win with probability $\frac{3}{4}$.

Proof: No strategy achieves larger probability than $\frac{3}{4}$. Maximize $S = \sum_{a,b,x,y} (-1)^{a \oplus b + x \cdot y} \Pr[a, b|x, y]$. Note that a can not depend on y and b can not depend on x, so $\Pr[a, b|x, y] = \Pr[a|x] \Pr[b|y]$.

$$\begin{aligned} S &= \sum_{a,b,x,y} (-1)^{a+b+x \cdot y} \Pr[a|x] \Pr[b|y] \\ &= \sum_{x,y} \sum_{a,x} (-1)^a \Pr[a|x] \sum_{b,y} (-1)^b \Pr[b|y] (-1)^{xy} \\ &= \sum_{a,x} (-1)^a \Pr[a|x] \sum_{b} (-1)^b \left[\Pr[b|0] + \Pr[b|1](-1)^x\right] \end{aligned}$$

Case by case, denote the conditional expectations $A(x) = \Pr[a = 0|x] - \Pr[a = 1|x], B(y) = \Pr[b = 0|y] - \Pr[b = 1|y], A, B \in [-1, +1]$. Then $\mathrm{Tr}(\rho \sigma_z) \sim A$

1. $x \cdot y = 0$: $\sum_{a,b} (-1)^{a+b} \Pr[a|x] \Pr[b|y] = \sum_a (-1)^a \Pr[a|x] \sum_b (-1)^b \Pr[b|y]$.
2. $x = y = 1$: $-\sum_a (-1)^a \Pr[a|1] \sum_b (-1)^b \Pr[b|1]$.

$S = A(0)B(0) + A(1)B(0) + A(0)B(1) - A(1)B(1) = A(0)[B(0) + B(1)] + A(1)[B(0) - B(1)] \leq 2$. Use probabilistic explanation, treat A, B as random variables in $\{-1, 1\}$. At least two terms are 0's. Or solve by linear optimization for continuous A, B in $[-1, 1]$.

Relate classical prob to quantum measurement. $x = 0 : A_0, \rho_A$. $\mathrm{Tr}(A_0 \rho_A) = A(0)$. Similar for $x = 1, A_1; y = 0, B_0; y = 1, B_1$.

$\psi_{AB} = |00\rangle + |11\rangle$ $A_0 = \sigma_z, A_1 = \sigma_x, B_0 = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z), B_1 = \frac{1}{\sqrt{2}}(\sigma_x - \sigma_z)$.

The expectation $S = \langle A_0 B_0 + A_0 B_1 + A_1 B_0 - A_1 B_1 \rangle$. $\langle A_0 B_0 \rangle = \mathrm{Tr}(\rho_{AB} \sigma_x \otimes \frac{\sigma_x + \sigma_z}{\sqrt{2}}) = \frac{1}{\sqrt{2}} \langle \sigma_x \otimes \sigma_x \rangle + \frac{1}{\sqrt{2}} \langle \sigma_x \otimes \sigma_z \rangle = \frac{1}{\sqrt{2}} + 0$.

So $S = \frac{\sqrt{2}}{2} \times 3 - (-\frac{\sqrt{2}}{2}) = 2\sqrt{2} > 2$. (Why contradicts the classical results?)