Lecture 5

Oct 8, 2025

Previously,

- Density matrices are mathematical descriptions of prob. distributions over pure quantum states.

- If ρ_{AB} is the d.m. describing the state of a system between two players/halves of the system called Alice and Bob

- Def (Partial trace) Let $|b_1\rangle \dots |b_{d_B}\rangle$ be a basis for Bob's qubits/system, then for any matrix ρ_{AB}

$$tr_B(\rho_{AB}) := \sum_{j=1}^{d_B} (\mathbb{1}_A \otimes \langle b_j|)\rho_{AB}(\mathbb{1}_A \otimes |b_j\rangle).$$

- We define ρ_A to be the unique matrix s.t.

$\forall\ |\psi\rangle \in \mathbb{C}^{d_A}$,

if we measure Alice's system in a basis including $|\psi\rangle$,

the prob. of measuring $|\psi\rangle$ w.r.t. ρ_A equals that

w.r.t. ρ_{AB}.

$$\Pr\left[\chi = |\psi\rangle \text{ w.r.t. } \rho_A\right] = \mathrm{tr}\left(|\psi\rangle\langle\psi|\,\rho_A\right)$$

$$= \langle\psi|\rho_A|\psi\rangle$$

$$\Pr\left[\chi = |\psi\rangle \text{ w.r.t. } \rho_{AB}\right] = \mathrm{tr}\left(\left(|\psi\rangle\langle\psi| \otimes \mathbb{1}\right)\rho_{AB}\right)$$

$\underline{\text{Thm}}$ $\rho_A = \mathrm{tr}_B(\rho_{AB})$ iff $\forall\, |\psi\rangle$,

$$\mathrm{tr}(|\psi\rangle\langle\psi|\,\rho_A) \qquad = \mathrm{tr}\Big(\big(|\psi\rangle\langle\psi| \otimes \mathbb{1}_B\big)\,\rho_{AB}\Big).$$

$\underline{\text{Pf}}$.

$$\mathrm{tr}\Big(\big(M_A \otimes \mathbb{1}_B\big)\,\rho_{AB}\Big)$$

$$= \sum_{i=1}^{d_A}\sum_{j=1}^{d_B} \langle a_i| \otimes \langle b_j|\, \Big(M_A \otimes \mathbb{1}_B\Big)\, \rho_{AB}\, |a_i\rangle \otimes |b_j\rangle$$

$$= \sum_{i=1}^{d_A}\sum_{j=1}^{d_B} \Big(\langle a_i| M_A\Big)\Big(\mathbb{1}_A \otimes \langle b_j|\Big)\, \rho_{AB}\, \Big(\mathbb{1}_A \otimes |b_j\rangle\Big)\,\Big(|a_i\rangle\Big)$$

$$= \sum_{i=1}^{d_A} \langle a_i| M_A\, \mathrm{tr}_B(\rho_{AB})\, |a_i\rangle$$

$$= \mathrm{tr}\Big(M_A\, \mathrm{tr}_B(\rho_{AB})\Big).$$

If $\rho_A = tr_B(\rho_{AB})$, then $\Rightarrow$ is obvious.

For other direction, if $\rho_A \neq tr_B(\rho_{AB})$, then $\exists$ coordinate (i,j) where they differ. i.e.

$$\langle i|\rho_A|j\rangle \neq \langle i|\, tr_B(\rho_{AB})\,|j\rangle$$

$$\| \qquad\qquad\qquad\qquad \|$$

$$tr\left(|j\rangle\langle i|\,\rho_A\right) \neq tr\left(|j\rangle\langle i|\, tr_B(\rho_{AB})\right). \qquad \square$$

Corollaries

- Using $M_A = \mathbb{1}_A$, this proves that $tr\,\rho_A = 1$.
- Using $M_A = |\psi\rangle\langle\psi|$, this proves $\rho_A \geq 0$.

Notice if $\rho_{AB} = |\psi\rangle\langle\psi|$ pure and $M_a = |x\rangle\langle x|$ for $x \in \{0, \dots, d_A - 1\}$ then

$$\mathrm{tr}\left(\left(M_A \otimes \mathbb{1}\right) \rho_{AB}\right) = \mathrm{tr}\left(\left(|x\rangle\langle x| \otimes \mathbb{1}\right) |\psi\rangle\langle\psi|\right)$$

$$= \left| \langle x| \otimes \mathbb{1} \, |\psi\rangle \right|^2$$

$$= \text{prob. Alice measures } x \text{ on } |\psi\rangle.$$

By thm, we can compute this directly from ρ_A.

More intuition:

Consider the d.m. $\rho \in \mathcal{L}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$

that is diagonal i.e. $\langle i,j|\rho|i,j\rangle \geq 0$

$\langle i,j|\rho|i',j'\rangle = 0$ unless $i = i'$, $j = j'$.

This is a prob. dist over states $|i,j\rangle$

given by $pr(i,j) = \langle i,j|\rho|i,j\rangle$.

Then $\langle i|\rho_A|i\rangle = \langle i|\, tr_B(\rho_{AB})|i\rangle$.

$$= \sum_j \langle i,j|\, \rho_{AB}\, |i,j\rangle$$

$$= \sum_j pr(i,j)$$

$$= pr(i).$$

So the partial trace is computing the marginal

(aka reduced) distribution.

Next let us prove that the def of partial trace does not depend on choice of $\{|b_i\rangle\}$.

Ex. If $\langle\psi|\rho|\psi\rangle = \langle\psi|\rho'|\psi\rangle$ for all unit vectors $|\psi\rangle$ then $\rho = \rho'$.

For any $|\psi\rangle$, let $\rho_B^{\psi} := (\langle\psi_A| \otimes \mathbb{1}_B)\, \rho_{AB} (|\psi_A\rangle \otimes \mathbb{1}_B)$.

Now,

$$\langle\psi|\, tr_B(\rho_{AB})\, |\psi\rangle$$

$$= \langle\psi_A| \left(\sum_{i=1}^{d_B} (\mathbb{1}_A \otimes \langle b_i|)\, \rho_{AB}\, (\mathbb{1}_A \otimes |b_i\rangle) \right) |\psi_A\rangle$$

$$= \sum_{i=1}^{d_B} \langle b_i|\, \rho_B^{\psi}\, |b_i\rangle = \sum_{i=1}^{d_B} tr(\rho_B^{\psi})$$

which is independent of $\{|b_i\rangle\}$ as trace is basis indep.

1 Let V_B be any unitary. Let $|1\rangle \dots |d_B\rangle$ be basis for B.

$$\mathrm{tr}_B\left((\mathbb{1}_A \otimes V_B)\, \rho_{AB}\, (\mathbb{1}_A \otimes V_B^\dagger) \right)$$

$$= \underbrace{\sum_{i=1}^{d_B} \left(\mathbb{1}_A \otimes \langle i| V_B \right) \rho_{AB} \left(\mathbb{1} \otimes V_B^\dagger |i\rangle \right)}_{\text{equivalent to taking partial trace w.r.t. basis } \{V|i\rangle\}.}$$

$$= \mathrm{tr}_B(\rho_{AB})$$ ← Bob's actions do not change ρ_A.

2 Let U_A be any unitary.

$$U_A \otimes \mathbb{1}_B\ \mathrm{tr}_B(\rho_{AB})\ U_A^\dagger \otimes \mathbb{1}_B$$

$$= \sum_i \left(U_A \otimes \langle i|_B \right) \rho_{AB} \left(U_A \otimes |i\rangle_B \right)$$

$$= \sum_i \left(\mathbb{1}_A \otimes \langle i|_B \right) U_A\, \rho_{AB}\, U_A^\dagger \left(\mathbb{1}_A \otimes |i\rangle_B \right)$$

$$= \mathrm{tr}_B\left(U_A\, \rho_{AB}\, U_A^\dagger \right).$$

Def. The maximally mixed state in d-dimensions is $\frac{1}{d} \mathbb{1}_d$. For qubits, $\frac{1}{2^n} \mathbb{1}_{2^n}$.

Def. A state ρ_{AB} betwen Alice and Bob each of n qubits is maximally entangled if $\rho_A = \frac{1}{2^n} \mathbb{1}_{2^n}$.

Ex. The 4 states we saw during teleportation:

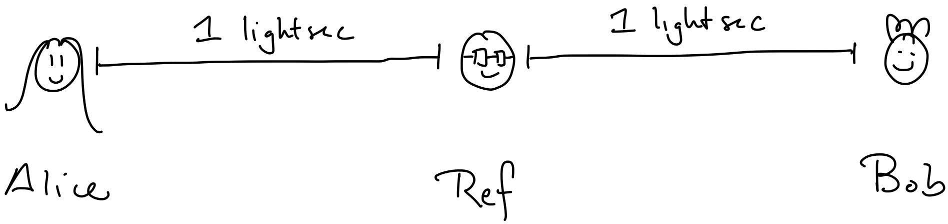
$$|\Phi_{ab}\rangle = \frac{1}{\sqrt{2}} \left(|0,a\rangle + (-1)^b \, |1, 1 \oplus a\rangle \right)$$

are maximally entangled.

Not. ρ_A is sometimes referred to as the reduced density matrix of ρ_{AB}.

Today: The CHSH game and "spooky-action" at a distance.

Consider the following game:



Ref samples $x, y \in \{0, 1\}$ and sends x to Alice and y to Bob. He asks for answers $a, b \in \{0, 1\}$ from Alice and Bob, respectively.

Alice and Bob must answer in $2 + \epsilon$ lightsecs.

They win if $a \oplus b = x \cdot y$.

What is their opt. success prob.?

Space-seperation is just to force no communication.

If Alice and Bob employ classical strategies, the best strategy only wins w pr $= 3/4$.

But if they use quantum strategies, they can win with probabilities up to $\cos^2 \pi/8 \sim 0.878$.

Next two lectures: understand the quantum strategy, prove it is optimal, understand how it characterizes the state of Alice and Bob's q.c.'s.

<u>Classical</u> Best deterministic strategy $a = 0$, $b = 0$ independent of input.

What if Alice & Bob share classical randomness?

Let $A_r, B_r : \{0,1\} \to \{0,1\}$ be the strategies for rand. r.

Then

$$\Pr[\text{win}] = \Pr_{x,y,r}\left[A_r(x) \oplus B_r(y) = xy\right]$$

$$= \sum_r \Pr[r] \cdot \Pr_{x,y}\left[A_r(x) \oplus B_r(y) = xy\right]$$

$$\leq \max_r \Pr_{x,y}\left[A_r(x) \oplus B_r(y) = xy \right]$$

$$\leq 3/4.$$

<u>Quantum</u> What if Alice and Bob share an EPR pair?

<u>Idea</u> Depending on the input $x \in \{0,1\}$, Alice makes a measurement of her half of the EPR pair.

Strategy: $\{ |\alpha_x^a\rangle \in \mathbb{C}^2 \}$ for Alice

where $\langle \alpha_x^0 | \alpha_x^1 \rangle = 0$.

likewise $\{ |\beta_y^b\rangle \in \mathbb{C}^2 \}$ for Bob,

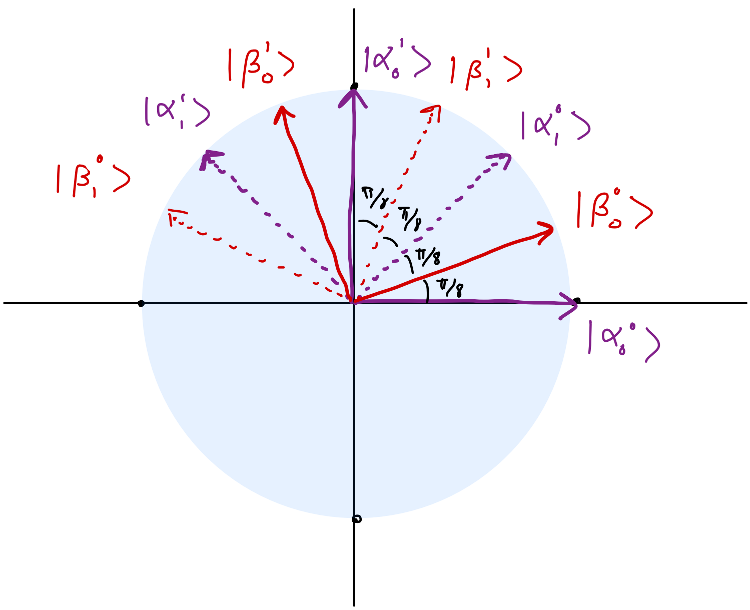
where $\langle \beta_x^0 | \beta_x^1 \rangle = 0$.

When Alice gets input x, she measures her half of the EPR pair with basis $|\alpha_x^0\rangle, |\alpha_x^1\rangle$ and answers accordingly. Likewise for Bob. So,

$$\Pr[a,b \mid x,y] = \left| \langle \alpha_x^a, \beta_y^b \mid EPR \rangle \right|^2.$$

Let us first write out the optimal strategy and then return to the more interesting part of proving it's optimality.

Easiest to draw a picture.



How do we analyze this?

① Recall from earlier lectures our def

$$|\theta\rangle := \cos\theta\,|0\rangle + \sin\theta\,|1\rangle$$

and $\langle \theta | \theta + \gamma \rangle = \cos \gamma.$

(2) For any basis $|v\rangle, |w\rangle$ of $\mathbb{C}^2$,

$$|EPR\rangle = \frac{1}{\sqrt{2}} \Big(|v\rangle |v^*\rangle + |w\rangle |w^*\rangle \Big) \qquad \text{(on hw 2)}$$

Using these 2 facts and the angles between vectors we can analyze the success of this game.

$$\Pr[\text{win}] = \sum_{x,y} \Pr[x,y] \cdot \sum_{a,b} \Pr[a,b \mid x,y] \cdot \underbrace{\mathbb{1}\{a \oplus b = xy\}}_{\text{win condition}}$$

$$\begin{aligned}
= \frac{1}{4} \Big[& |\langle \alpha_0^0, \beta_0^0 | EPR\rangle|^2 + |\langle \alpha_0^1, \beta_0^1 | EPR\rangle|^2 && (x,y) = (0,0) \\
+ & |\langle \alpha_0^0, \beta_1^0 | EPR\rangle|^2 + |\langle \alpha_0^1, \beta_1^1 | EPR\rangle|^2 && (x,y) = (0,1) \\
+ & |\langle \alpha_1^0, \beta_0^0 | EPR\rangle|^2 + |\langle \alpha_1^1, \beta_0^1 | EPR\rangle|^2 && (x,y) = (1,0) \\
+ & |\langle \alpha_1^0, \beta_1^1 | EPR\rangle|^2 + |\langle \alpha_1^1, \beta_1^0 | EPR\rangle|^2 \Big] && (x,y) = (1,1)
\end{aligned}$$

Recognize

$$\langle \alpha_x^a, \beta_y^b \,|\, EPR\rangle = \langle \alpha_x^a, \beta_y^b | \cdot \frac{|\beta_y^0\rangle|\beta_y^0\rangle + |\beta_y^1\rangle|\beta_y^1\rangle}{\sqrt{2}}$$

$$= \frac{1}{\sqrt{2}} \langle \alpha_x^a | \beta_y^b \rangle$$

$$= \frac{1}{\sqrt{2}} \cos\theta \leftarrow \text{angle between } |\alpha_x^a\rangle, |\beta_y^b\rangle$$

if Alice & Bob measure in basis differing in θ, their measurements then

$$\Pr[a = b] = \cos^2\theta \quad \text{and} \quad \Pr[a = 0] = \frac{1}{2}.$$

Notice by design, all terms in $\Pr[\text{win}]$ evaluate to $\pm \frac{1}{\sqrt{2}} \cos \pi/8$.

$$\text{So } \Pr[\text{win}] = \frac{1}{4}\left[8 \cdot \frac{1}{2} \cos^2 \pi/8 \right] = \cos^2 \pi/8$$

$$= 0.854.$$

So Alice and Bob can win game with quantum strategies with higher prob than classical.

"Clauser, Hauser, Shimony, Holt" (CHSH)

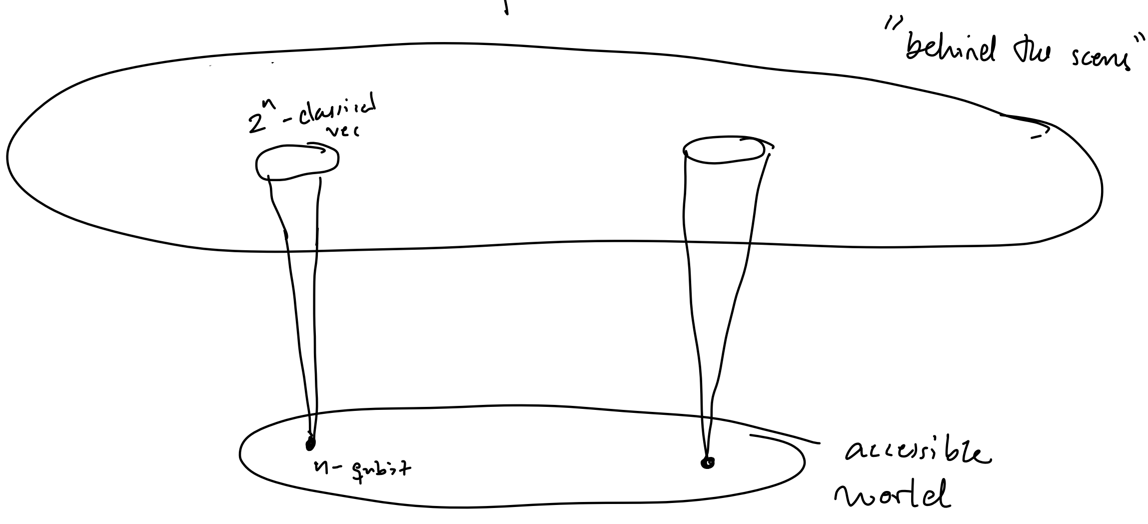
experiment.

aka Bell inequality experiment.

aka a non-local game.

We know that q. mechanics says that n-qubit states can be described by a 2^n-dim vector. Meaning, that if nature is doing a lot of computation "behind the scenes"

But where is that info stored?



CHSH game proves that ↑ picture is incorrect!

The information about the state of the system cannot be held locally assuming a speed limit on information.

This is because local info could not produce the statistics needed to win the game with prob. > $3/4$.

To describe the state of the system, we need to have a global description.

Indeed $\rho = |EPR\rangle\langle EPR|_{AB}$, then $\rho_A = \rho_B = \frac{1}{2}\mathbb{1}_2$.

The key is that the coins are correlated!

Can we experimentally verify CHSH?

Can actually conduct this and has been done!

Does it prove that quantum mechanics is <u>correct</u>?

<u>No</u>. Since the q. advantage is only probabilistic

it is <u>evidence</u> that our world isn't classical.

Suppose a ref observes Alice & Bob winning the game 80% of the time. How many games would they have to play before he is convinced with $1-10^{-9}$ confidence that the world isn't classical?

Chernoff bounds: $X = \sum_i^n X_i \leftarrow$ per game success

$$\Pr\left[X \geq (1+\delta)\mu\right] \leq \exp\left(-\frac{\delta^2 \mu}{2+\delta}\right)$$

here $\mu = 0.75n$ $\quad \delta = \frac{0.8}{0.75} < 1.07$

$$\Pr\left[X \geq 0.8n\right] \leq \exp\left(-0.27n\right)$$

want $\exp\left(-0.27n\right) \leq 10^{-9}$

$$0.27n \geq 9 \ln 10$$

$$n \geq 77.$$

Does it prove the world is <u>quantum</u>? <u>No</u>.

For all we know, there is a stronger <u>theory</u> consistent with quantum mechanics.

<u>Ex</u>. Noisy telepathy! Alice and Bob are telepathic with a success rate of 80%. Perfect telepathy and they win with probability 1.

Is $\cos^2 \pi/8$, the optimal quantum strategy?

Can we do better if Alice and Bob share multiple EPR pairs?

Answer: YES!

But to prove it, we will need a way to characterize all quantum strategies Alice & Bob could apply.

General quantum strategy of Alice

Input: ρ_{AB} ← some state over Alice and Bob.
Could include ancillas.

Her algorithm:

- Apply unitary U_1 to her qubits.
- Measure qubit 1.
- Apply U_2 to her qubits.
- Measure qubit 1.

⋮

- Apply U_T to her qubits.
- Measure qubit 1 and output value.

Note: this is general enough to include any classical processing of measurements by hw 1 result.

We need to simplify her algorithm.

Step 1: Principle of deferred measurement

Alice's strategy is equivalent to a strategy with only 1 measurement (the final measurement). The new strategy uses T additional ancilla qubits.

Pf.

Recall $CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ or $CNOT|x\rangle|y\rangle = |x\rangle|y\oplus x\rangle$.

Let σ_{AB} be the state of the computation before the measument of the 1^{st} qubit. After the measurement the state will be:

$$\sigma_{AB}' = \sum_{k\in\{0,1\}} (|k\rangle\langle k| \otimes \mathbb{1})\, \sigma_{AB}\, (|k\rangle\langle k| \otimes \mathbb{1}).$$

Consider instead initializing an additional ancilla so the state is $\sigma_{AB} \otimes |0\rangle\langle 0|_{Anc}$. Apply CNOT between qubit 1 and ancilla. Then ignore the ancilla for the remainder of the algorithm.

Since the remainder of the alg. ignores the ancilla, we know the measurements are equivalent to that of the state

$$\mathrm{tr}_{anc}\left(CNOT_{1,anc} \left(\sigma_{AB} \otimes |0\rangle\langle 0|_{anc} \right) CNOT_{1,anc} \right)$$

Suffices to show that this equals σ'_{AB}.

Let $\sigma_{AB} = \sum_{i,j \in \{0,1\}} |i\rangle\langle j|_1 \otimes \sigma_{ij}$ $\leftarrow$ generic decomp.

$$\left(\sigma_{ij} = (\langle i| \otimes \mathbb{1}) \sigma (|j\rangle \otimes \mathbb{1}) \right) \quad \sigma = \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{pmatrix}.$$

$$\mathrm{tr}_{anc}\left(CNOT_{1,anc} \left(\sigma_{AB} \otimes |0\rangle\langle 0|_{anc} \right) CNOT_{1,anc} \right)$$

$$= \mathrm{tr}_{anc}\left(\sum_{i,j \in \{0,1\}} |i,i\rangle\langle j,j|_{1,anc} \otimes \sigma_{ij} \right)$$

$$= \sum_{k \in \{0,1\}} \langle k|_{anc} \otimes \mathbb{1}_{AB} \left(\sum_{i,j \in \{0,1\}} |i,i\rangle\langle j,j|_{1,anc} \otimes \sigma_{ij} \right) |k\rangle_{anc} \otimes \mathbb{1}_{AB}$$

equals 0 unless $k = i$ and $j = k$

$$= \sum_{k \in \{0,1\}} |k\rangle\langle k|_1 \otimes \sigma_{kk}$$

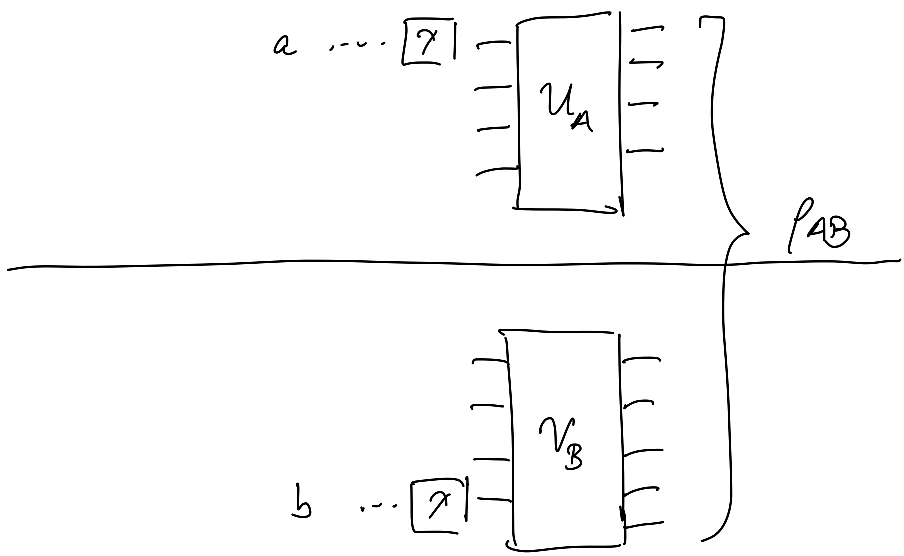
Notice $\sigma' = \sum_{k\in\{0,1\}} (|k\rangle\langle k| \otimes \mathbb{1})\, \sigma_{AB}\, (|k\rangle\langle k| \otimes \mathbb{1})$.

$= \sum_k |k\rangle\langle k| \otimes \sigma_{kk}$.

Therefore, entangling with ancilla and "discarding" ancilla is equivalent to measurement.

So, Alice strategy is to apply a unitary U_A and then measure the first qubit, without loss of generality.

Same with Bob.



What is Alice's prob of outputting 1 ?

$$\Pr(\text{outputting } i) = tr\left(\left(|i\rangle\langle i| \otimes \mathbb{1}\right) U_A \rho_A U_A^\dagger\right)$$

$$= tr\left(U_A^\dagger \left(|i\rangle\langle i| \otimes \mathbb{1}\right) U_A \rho_A\right)$$

Let $M^i = U_A^\dagger |i\rangle\langle i| \otimes \mathbb{1}\, U_A$.

Then $\Pr(\text{outputting } i) = tr(M_i \rho)$.

Notice $M^0 + M^1 = U_A^\dagger \left(\sum |i\rangle\langle i| \otimes \mathbb{1}\right) U_A$

$$= U_A^\dagger \mathbb{1} U_A = \mathbb{1}$$

and $M^i \geq 0$, and $(M^i)^2 = M^i$ projector.

This is a special case of a

pos. valued operator valued measurement.

will show up in hw 2.

For now, suffices to observe that Alice's strategy can be described by M^0, M^1 proj. for $M^0 + M^1 = \mathbb{1}_A$.

In general, for the CHSH game, Alice's action depends on input x, so her strategy can be expressed by pairs

$\{A_x^0, A_x^1\}_{x \in \{0,1\}}$ s.t. A_x^a proj. and $A_x^0 + A_x^1 = \mathbb{1}_A$.

Same with Bob: $\{B_y^0, B_y^1\}$

Today:

Complete analysis of CHSH game.

Recall we defined the game as Alice and Bob recieve $x, y \in \{0,1\}$ and answer with $a, b \in \{0,1\}$.

Win condition: $a \oplus b = x \cdot y$.

Switch to $\{a, b\} \in \{-1, +1\}$ and

win condition: $a \cdot b = (-1)^{x \cdot y}$.

So last time we showed that a general strategy for Alice and Bob can be mathematically expressed as

$\{A_x^{(1)}, A_x^{(-1)}\}_{x \in \{0,1\}}$ s.t. $A_x^{(a)}$ proj. and $A_x^{(1)} + A_x^{(-1)} = \mathbb{1}_A$.

Same with Bob: $\{B_y^{(1)}, B_y^{(-1)}\}$

Assume Alice and Bob share a state ρ_{AB}.

Then $\Pr[a, b \mid x, y] = tr\left(A_x^{(a)} \otimes B_y^{(b)} \rho_{AB}\right)$.

$$\Pr[\text{win}] = \sum_{x,y} \Pr[x,y] \cdot \sum_{a,b} \Pr[a,b \mid x,y] \cdot \mathbb{1}\{a \cdot b = (-1)^{xy}\}$$

$$= \frac{1}{4} \sum_{a,b,x,y}^{1} \mathbb{1}\{a \cdot b = (-1)^{xy}\} \operatorname{tr}\left(A_x^{(a)} \otimes B_y^{(b)} \rho_{AB}\right)$$

$$= \frac{1}{4} \sum_{a,b,x,y} \left(\frac{1}{2} + \frac{a \cdot b \cdot (-1)^{xy}}{2}\right) \operatorname{tr}\left(A_x^{(a)} \otimes B_y^{(b)} \rho_{AB}\right)$$

$$= \frac{1}{2} + \frac{1}{8} \sum_{a,b,x,y} ab(-1)^{xy} \operatorname{tr}\left(A_x^{(a)} \otimes B_y^{(b)} \rho_{AB}\right)$$

$$(*) \qquad = \frac{1}{2} + \frac{1}{8} \operatorname{tr}\left(\left(\sum_{a,b,x,y} (-1)^{xy} \left(a A_x^{(a)} \otimes b B_y^{(b)}\right)\right) \rho_{AB}\right).$$

To simplify, lets define $A_x := A_x^{(1)} - A_x^{(-1)}$.

Since $A_x^{(1)} + A_x^{(-1)} = \mathbb{1}_A$ and $A_x^{(a)}$ are projectors,

$$A_x = \mathbb{1}_A - 2A_x^{(1)} \quad \text{and}$$

$$A_x^2 = \mathbb{1}_A + 4 A_x^{(1)\,2} - 4 A_x^{(1)} = \mathbb{1}_A \quad \text{since } A_x^{(1)} \text{ proj.}$$

so A_x^2 has eigenvalues $(-1, 1)$.

Such matrices are called "<u>binary</u> <u>observables</u>".

Notice $\mathbb{E}(a) = \mathrm{tr}(A_x \rho_A)$.

Back to solving $(*)$. Notice that

$$A_x \otimes B_y = \left(\sum_a a A_x^{(a)}\right) \otimes \left(\sum_b b B_y^{(b)}\right)$$

$$= \sum_{a,b} a A_x^{(a)} \otimes b B_y^{(b)}.$$

So $$(*) = \frac{1}{2} + \frac{1}{8} \mathrm{tr}\left(\underbrace{\left(\sum_{x,y} (-1)^{xy} A_x \otimes B_y\right)}_{:= CHSH} \rho_{AB} \right)$$

CHSH is an operator $\in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$. So

Claim: $$\mathrm{tr}(CHSH\ \rho_{AB}) \leq \underbrace{\|CHSH\|_{op}}_{\text{max singular value.}}$$

Pf. $\rho_{AB} = \sum p_i |\psi_i\rangle\langle\psi_i|$.

unit norm

Then $$\mathrm{tr}(CHSH\ \rho_{AB}) = \sum_i p_i \langle\psi_i|CHSH|\psi_i\rangle$$

$$\leq \sum_i p_i \|CHSH\| = \|CHSH\|. \quad \square$$

So, $\Pr[\text{win}] \leq \frac{1}{2} + \frac{1}{8} \| CHSH \|$.

Note: $\cos^2 \pi/8 = \frac{1}{2} + \frac{2\sqrt{2}}{8}$.

So suffices to prove, $\|CHSH\| \leq 2\sqrt{2}$, or $\|CHSH^2\| \leq 8$.

Pf.

$$CHSH = A_0 \otimes B_0 + A_1 \otimes B_0 + A_0 \otimes B_1 - A_1 \otimes B_1 .$$

$$= (A_0 + A_1) \otimes B_0 + (A_0 - A_1) \otimes B_1 ,$$

Use $A_0^2 = A_1^2 = B_0^2 = B_1^2 = \mathbb{1}$ to get

$$CHSH^2 = (A_0 + A_1)^2 \otimes \mathbb{1} + (A_0 - A_1)^2 \otimes \mathbb{1}$$

(no commutation!)

$$+ (A_0 + A_1)(A_0 - A_1) \otimes B_0 B_1$$

$$+ (A_0 - A_1)(A_0 + A_1) \otimes B_1 B_0 .$$

$$= 4\mathbb{1} + (\underbrace{A_0 A_1 + A_1 A_0 - A_0 A_1 - A_1 A_0}_{0}) \otimes \mathbb{1}$$

$$+ (A_0 A_1 - A_1 A_0) \otimes (-B_0 B_1)$$

$$+ (A_0 A_1 - A_1 A_0) \otimes (B_1 B_0)$$

$$= 4\mathbb{1} + [A_0, A_1] \otimes [B_1, B_0]$$

where $[A_0, A_1]$ = commutator $:= A_0 A_1 - A_1 A_0$.

Now $\| [A_0, A_1] \| \leq \| A_0 A_1 \| + \| A_1 A_0 \|$

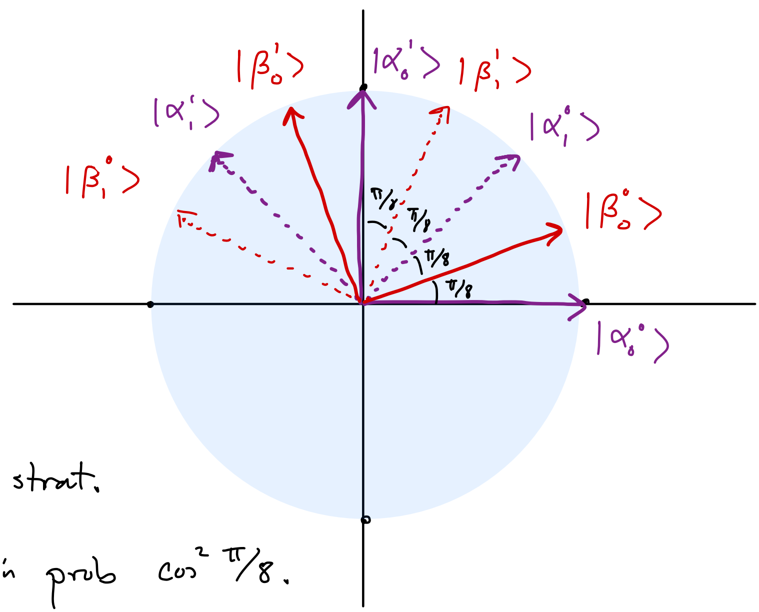
$$\leq \|A_0\| \cdot \|A_1\| + \|A_1\| \cdot \|A_0\| = 2.$$

So $\| CHSH \|^2 = 4 + \| [A_0, A_1] \| \cdot \| [B_0, B_1] \|$

$$\leq 4 + 2 \cdot 2 = 8. \quad \blacksquare$$

What have we solved?

We proved, the following strategy wins with prob. $\cos^2 \pi/8$.



And that every q. strat. is bounded by win prob $\cos^2 \pi/8$.

Def. (Value of game) For game G,

$\omega(G)$ = max prob winning over classical strat

$\omega^*(G)$ = sup prob. winning over q. strat.

$\omega^*(\text{CHSH}) = \cos^2 \pi/8$ and $\omega(\text{CHSH}) = 3/4$.

What was the opt strategy, we found?

$A_0^* = |\alpha_0^0\rangle\langle\alpha_0^0| - |\alpha_0^1\rangle\langle\alpha_0^1| = Z$ (phase-flip)

$A_1^* = X$ (bit-flip)

$B_0^* = H$ (Hadamard)

$B_1^* = \tilde{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}$ ("rotated" Hadamard)

Lets notice that $A_0^* A_1^* = -A_1^* A_0^*$ and $B_0^* B_1^* = -B_1^* B_0^*$.

(anti-commutation).

Furthermore, $|EPR\rangle$ is the unique $2\sqrt{2}$ eigenvector of

$$\text{CHSH}^* = A_0 \otimes B_0 + A_1 \otimes B_0 + A_0 \otimes B_1 - A_1 \otimes B_1$$

(easy to computer verify).

Are there any other optimal strategies?

Thm All optimal strategies are unitarily equivalent to to the strategy $(A_0^*, A_1^*, B_0^*, B_1^*, |EPR\rangle)$ given.

we will formally define unitarily equivalent.

Pf. If a strategy is optimal, then all ineq. must be tight:

$$\mathrm{tr}(CHSH\ \rho_{AB}) = 2\sqrt{2} \quad \text{and} \quad \|CHSH\| = 2\sqrt{2}.$$

so $\rho_{AB} = \sum p_i |\psi_i\rangle\langle\psi_i|$ with $\langle\psi_i|CHSH|\psi_i\rangle = 2\sqrt{2}$.

So, ρ_{AB} is a linear combination of "pure strategies", each of which is a $2\sqrt{2}$ eigenvalued eigenvector of CHSH.

Note: $\|CHSH\| = 2\sqrt{2} \iff \|[A_0, A_1]\| = 2$ and $\|[B_0, B_1]\| = 2$.

Does $\|[A_0, A_1]\| = 2$ imply $A_0 A_1 = -A_1 A_0$?