LECTURE 12 (February 26th)

TODAY Near-term Quantum Advantage

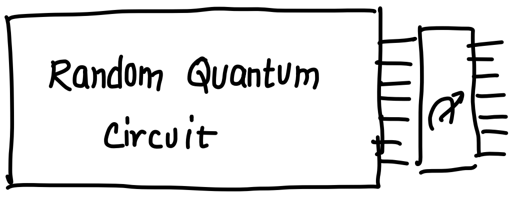
YZ-search problem is in BQP^O not in BPP^O } provable quantum advantage & we can instantiate with a cryptographic hash function

& in NP^O } verifiable in poly-time by classical algorithms

But the problem is that the quantum circuit to solve it can't be implemented on current quantum devices which are noisy & limited to small-depth computation

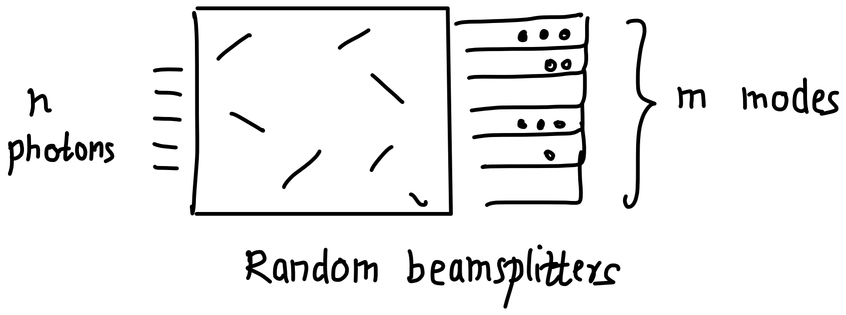
Near-term experiments are based on random circuit or boson sampling

Random Circuit Sampling



Given a random quantum circuit obtained from a "simple" family, sample from the output distribution

Boson sampling



Sample from output distribution of boson sampling experiment

These are near-term, we have some evidence of quantum advantage, although there is still a lot we don't know but not verifiable easily

Holy-grail = Provable quantum advantage + Near-term + Verifiable

Now we are going to focus on Random Circuit Sampling & consider what evidence of quantum advantage do we have. We won't cover boson sampling here

Note: Both these tasks are practically useless (except for maybe generating randomness) but for now we want to demonstrate quantum advantage experimentally

Warning: This is a rapidly evolving field and we are only going to talk about some initial results.

Practically, it is not clear whether the evidence is robust in the presence of noise and whether we have effectively demonstrated quantum advantage since these experiments are hard to scale & verify

Sampling from the output distribution of a quantum circuit is #P-hard

↳ count # solutions to a SAT formula

$$|\psi\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle \underbrace{|\varphi(x)\rangle}_{\text{output qubit}}$$

φ → SAT-formula

$$\mathbb{P}[\text{ output is } 1] = \frac{\#\text{ satisfying assignments}}{2^n}$$

#P- is a counting complexity class not a decision one

The closest decision class is PP which we recall is the class of languages where poly-time randomized algorithms can do better than random guessing

It can also be described as

PP = output the highest order bit of # solutions to a #P-problem

Also, $P^{PP} = P^{\#P}$

Solving a #P-hard problem is as hard as solving an NP problem but a very well-known theorem of Toda says that in fact

$PH \subseteq P^{\#P}$, so it is even more difficult than the entire polynomial hierarchy

Above we encoded a #P-complete problem (#SAT) in the guise of computing the acceptance probability of a quantum circuit exactly

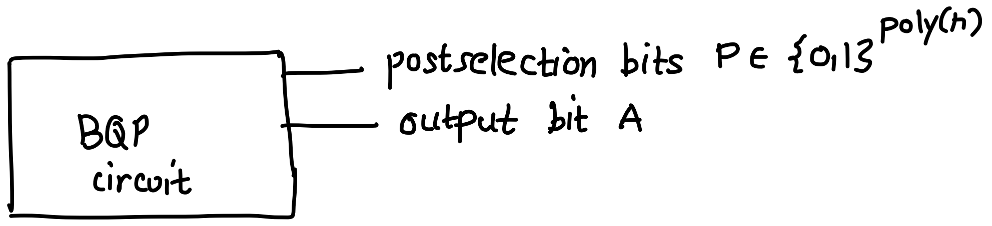
We don't believe quantum circuits can solve NP or #P-complete problems in poly-time

But this is different from problems in BQP where we don't know the exact acceptance probabilities

This seems promising but we need simpler classes of quantum circuits and need that this is robust to errors & noise which exact sampling is not

In order to do this, we need the notion of postselection & the complexity class postBQP

PostBQP



A problem is in postBQP if

(1) $\mathbb{P}[\underset{P}{\text{postselection}} \text{ bits} = 0\cdots 0] \geq 2^{-\text{poly}(n)}$

(2) $\mathbb{P}[A \text{ is correct} \mid P = 0\cdots 0] \geq \frac{2}{3}$ ⇒ This can be amplified

We are conditioning on an event of exponentially small probability

This is not physically viable but is a very powerful theoretical tool

Postselection gives a lot of computational power

$\boxed{NP \subseteq \text{post BQP}}$

$$|\psi\rangle = \sqrt{1-\varepsilon}\left(\frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle|\overset{\text{SAT formula}}{\varphi}(x)\rangle\right) + \sqrt{\underset{=\frac{1}{16^n}}{\varepsilon}}\,|\text{abort}\rangle\overset{\text{special state}}{|1\rangle}$$

If we postselect on 2nd qubit being 1,

$$\text{unnormalized } |\psi\rangle = \frac{\sqrt{1-\varepsilon}}{\sqrt{2^n}}\sum_{x:\,\varphi(x)=1}|x\rangle + \sqrt{\varepsilon}\,|\text{abort}\rangle$$

If φ is unsatisfiable measuring first register gives abort

Otherwise in the worst-case φ has a single satisfying assignment x^*
so the unnormalized state is

$$\frac{\sqrt{1-\varepsilon}}{\sqrt{2^n}}|x^*\rangle + \sqrt{\varepsilon}\,|\text{abort}\rangle$$

$$\approx \frac{1}{\sqrt{2^n}} |x^*\rangle + \frac{1}{4^n} |abort\rangle$$

$$\mathbb{P}[\text{we measure } x^*] \geq 1 - 2^{-\theta(n)}$$

We can define the classical version post BPP similarly and the above also works for post BPP

Theorem post BQP = PP

postBQP ⊆ PP follows from just minor modifications to the BQP ⊆ PP proof we saw earlier

The other direction is non-trivial and was shown by Aaronson

We are not going to cover the proof in the lecture but I might try to make an exercise out of it

Theorem post BQP = post BPP ⟹ PH collapses to the third level

Proof Known results say postBPP $\subseteq NP^{NP^{NP}} = \Sigma_3^P$ and $PH \subseteq P^{PP} = P^{\#P}$

Thus, postBQP = postBPP implies

$$PH \subseteq P^{PP} = P^{PostBQP} = P^{postBPP} = P^{\Sigma_3^P} \subseteq \Sigma_3^P \qquad \square$$

Now the punchline is, you can take a simple quantum circuit class C

for example, IQP circuits which look like $H^{\otimes n} D H^{\otimes n}$ where D is a diagonal unitary in the computational basis

These circuits are way less powerful than BQP but if we give them the power of postselection, they become as powerful as postBQP

Theorem post IQP = postBQP

Now, if there was an exact classical sampler to sample from output distribution for an IQP circuit, then we could classically postselect and

$$\text{post BPP} = \text{postIQP} = \text{post BQP} \implies \text{PH collapses}$$

These circuits cannot solve many problems but for the specific problems they solve, a classical computer could not solve them unless PH collapses

The same argument also works if we have a multiplicative approximation with a classical sampler, i.e.

$$\forall \text{ outcomes } y, \quad \frac{\mathbb{P}[\text{classical sampler outputs } y]}{\mathbb{P}[\text{quantum circuit outputs } y]} \in \left(\frac{1}{1+\varepsilon}, 1+\varepsilon\right)$$

Caveats ① We have shown existence which says in the worst-case sampling from an IQP circuit is hard classically but if it was a single pathological case, it may not be useful experimentally

Can we say that on average this task is hard ?

② Again the above assumes exact or multiplicative error which is not experimentally feasible

Can we say that this is still hard if the classical sampler samples from a distribution that is ε-close in total variation distance ?

TV distance b/w distributions p & q

→ classical sampler

→ quantum sampler

$$= \frac{1}{2} \sum_{y \in \{0,1\}^n} |p_c(y) - q_c(y)| \quad \text{where} \quad q_c(y) = |\langle y | C | 0 \rangle|^2$$

Let us see how to handle caveat ② first

Suppose a classical sampler outputs from a distribution that is ε-close in TV-dist.

$$\text{Then,} \quad \mathbb{E}_y \left[|p_c(y) - q_c(y)| \right] \leq \frac{2\varepsilon}{2^n}$$

$$\implies \text{For } 99\% \text{ of } y\text{'s}, \quad |p_c(y) - q_c(y)| \leq \frac{200\,\varepsilon}{2^n}$$

From sampling to estimating probabilities

For a randomized poly-time sampler,

$$\mathbb{P}[\,A \text{ outputs } y\,] = \frac{\#\text{random choices that lead to } y}{2^{\mathrm{poly}(n)}} \Big\} \Rightarrow \text{This is a problem in } \#P$$

A classic result of Stockmeyer says that # solutions to a #P-problem can be multiplicatively approximated with a randomized poly-time algorithm that has an NP-oracle

$\boxed{\text{Theorem}}$ (Stockmeyer) Let $f : \{0,1\}^m \longrightarrow \{0,1\}^n$ be the classical sampler that takes as input description of circuit and some random bits, and let $y \in \{0,1\}^n$.

There is a $\mathrm{FBPP}^{\mathrm{NP}}$ algorithm that runs in $\mathrm{poly}(n, 1/\delta)$ time and outputs $\hat{p}_y$ satisfying

$$\hat{p}_y \in [1 \pm \delta] \cdot \mathbb{P}_{x \in \{0,1\}^m}[f(x) = y]$$

Applying Stockmeyer's result with $\delta = 1/\mathrm{poly}(n)$, we get an estimate $\hat{p}_y$ in $\mathrm{poly}(n)$ time

where $\hat{p}_y \in \left(1 \pm \frac{1}{\mathrm{poly}(n)}\right) p_y$

This means for most y, $\hat{p}_y \in \left(1 \pm \frac{1}{\mathrm{poly}(n)}\right) p_y \in \left(1 \pm \frac{1}{\mathrm{poly}(n)}\right)\left(q_y \pm O\left(\frac{\varepsilon}{2^n}\right)\right)$

$$\Rightarrow \quad \hat{p}_y \in \left(1 \pm \frac{1}{\mathrm{poly}(n)}\right) q_y \pm O\left(\frac{\varepsilon}{2^n}\right) \quad \text{for most } y\text{'s}.$$

] This is true for any circuit C

Now suppose we sample a random circuit $C \in \mathcal{C}$, then

$$\mathbb{P}_{C,y}\left[\ \hat{p}_y \in \left(1 \pm \frac{1}{\mathrm{poly}(n)}\right) q_y \pm O\left(\frac{\varepsilon}{2^n}\right)\right] \geqslant 0.99$$

NEXT TIME Starting from the above, what other assumptions do we need on C to conclude that no TV-distance error sampler exists?