Title: The power of random quantum circuits

Speakers: Bill Fefferman

Series: Perimeter Institute Quantum Discussions

Date: May 12, 2021 - 4:00 PM

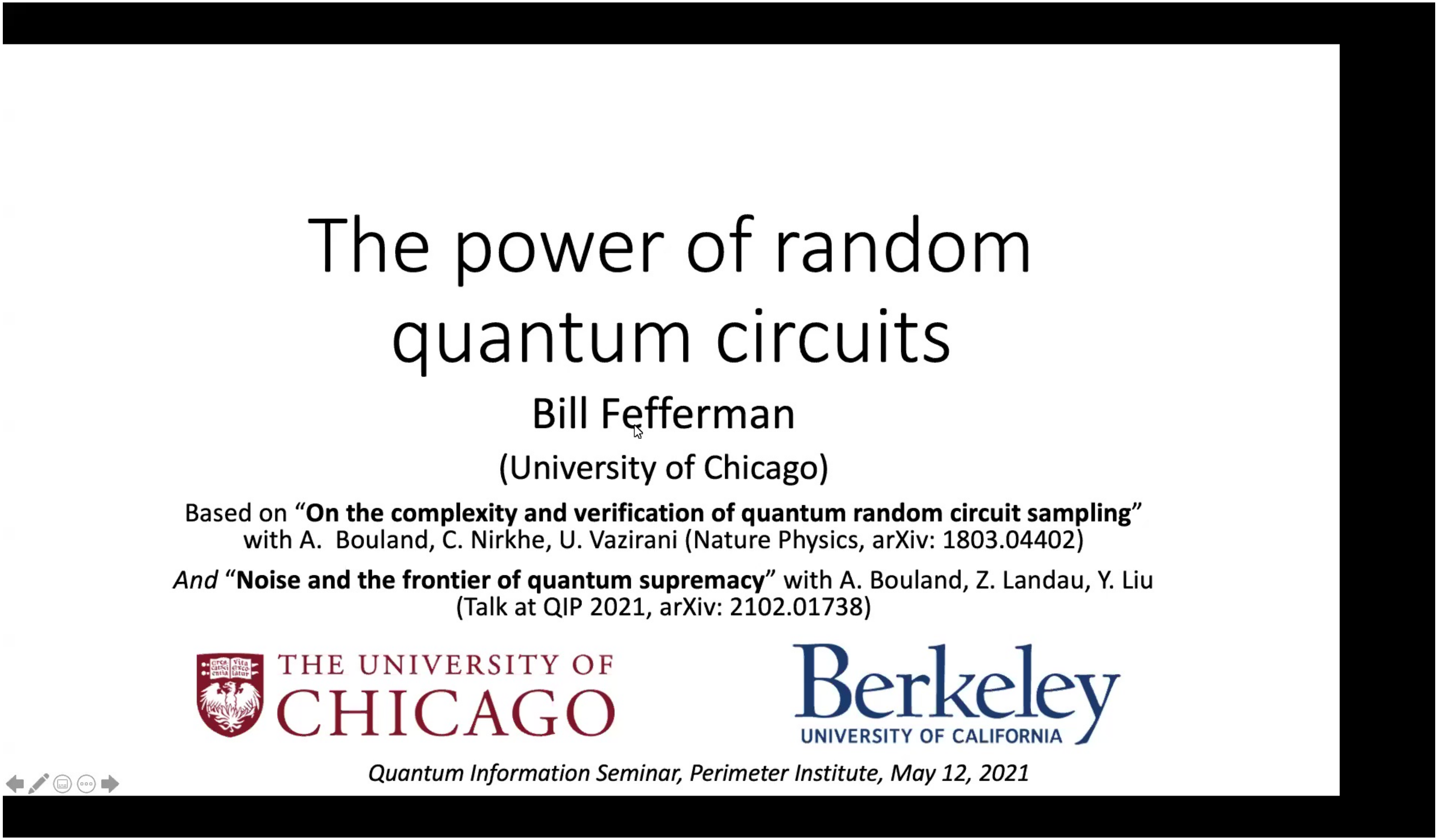
URL: http://pirsa.org/21050015

Abstract: In recent years, random quantum circuits have played a central role in the theory of quantum computation. Much of this prominence is due to recent random quantum circuit sampling experiments which have constituted the first claims of "quantum supremacy". &nbsp; While random quantum circuits enjoy certain advantages that make them ideal for implementation by near-term quantum experiments, it is unclear a priori why they should be difficult to simulate classically. While we know several examples of quantum algorithms which attain exponential speedups over classical computation, they all seem to rely on highly structured circuits (such as quantum Fourier transforms) which are far from typical. Why then should we expect a generic quantum circuit to realize a large computational advantage?

In this talk we will explain the complexity theoretic basis for the classical hardness of random circuit sampling.

This talk will be based on joint work with Adam Bouland, Chinmay Nirkhe, and Umesh Vazirani (https://arxiv.org/abs/1803.04402), as well as Adam Bouland, Yunchao Liu and Zeph Landau (https://arxiv.org/abs/2102.01738).

Zoom Link: https://pitp.zoom.us/j/92408058752?pwd=b1JoeTVISlpjaXk3ZjBoSi9pYjNUZz09



The power of random quantum circuits

Bill Fefferman

(University of Chicago)

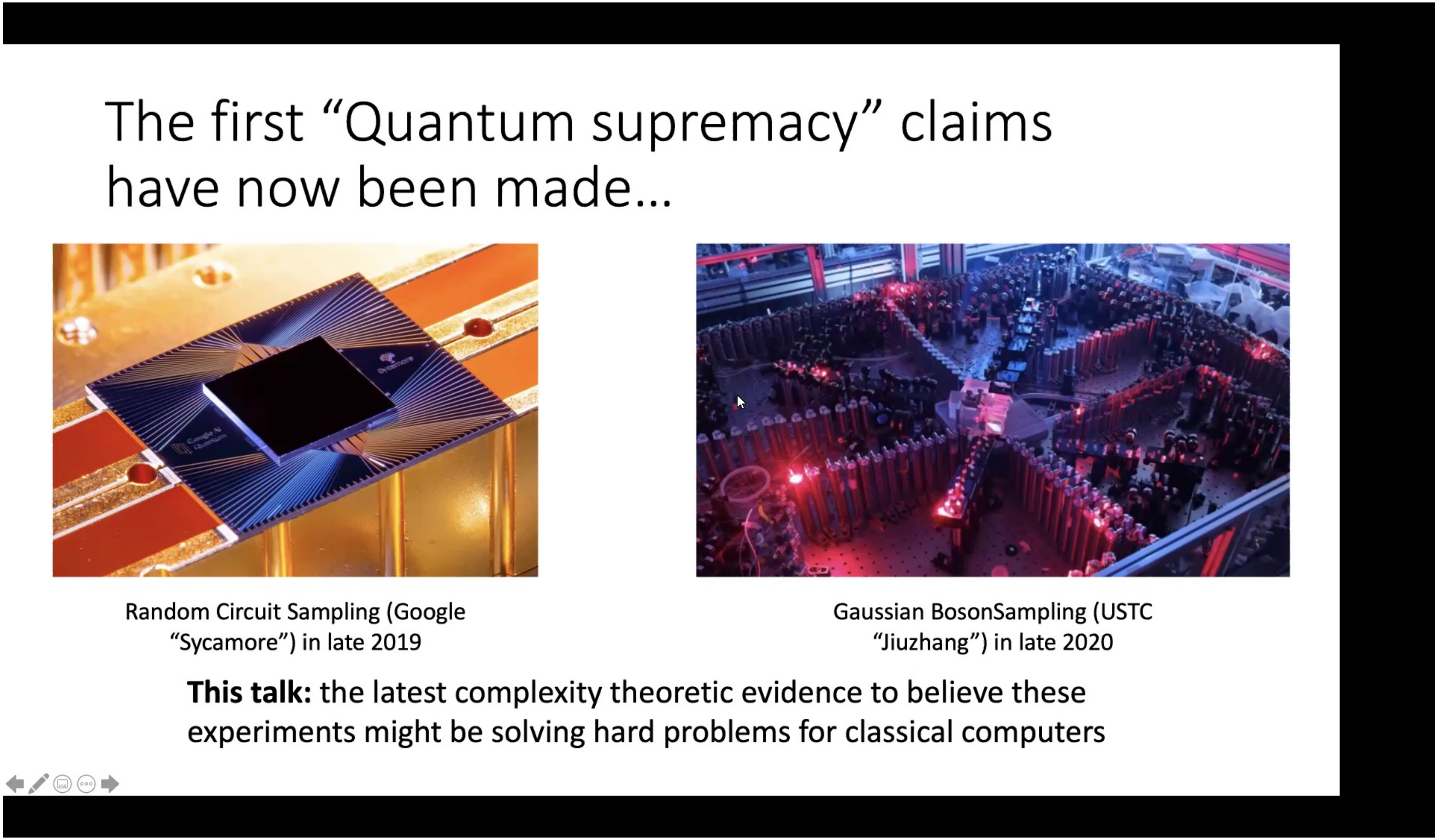
Based on “**On the complexity and verification of quantum random circuit sampling**”
with A. Bouland, C. Nirkhe, U. Vazirani (Nature Physics, arXiv: 1803.04402)

And “**Noise and the frontier of quantum supremacy**” with A. Bouland, Z. Landau, Y. Liu
(Talk at QIP 2021, arXiv: 2102.01738)

THE UNIVERSITY OF CHICAGO

Berkeley
UNIVERSITY OF CALIFORNIA

Quantum Information Seminar, Perimeter Institute, May 12, 2021

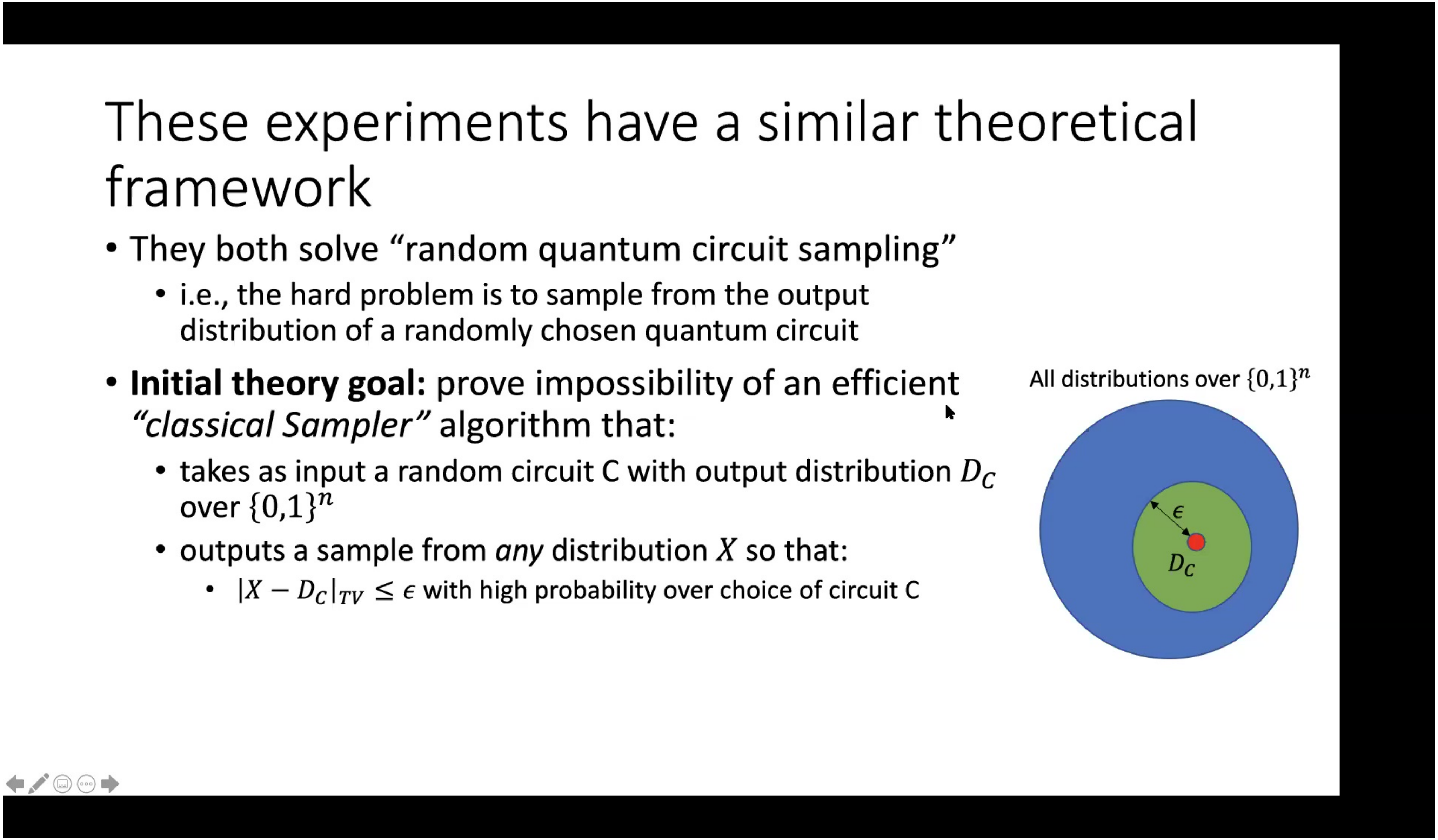


The first “Quantum supremacy” claims have now been made…

Random Circuit Sampling (Google “Sycamore”) in late 2019

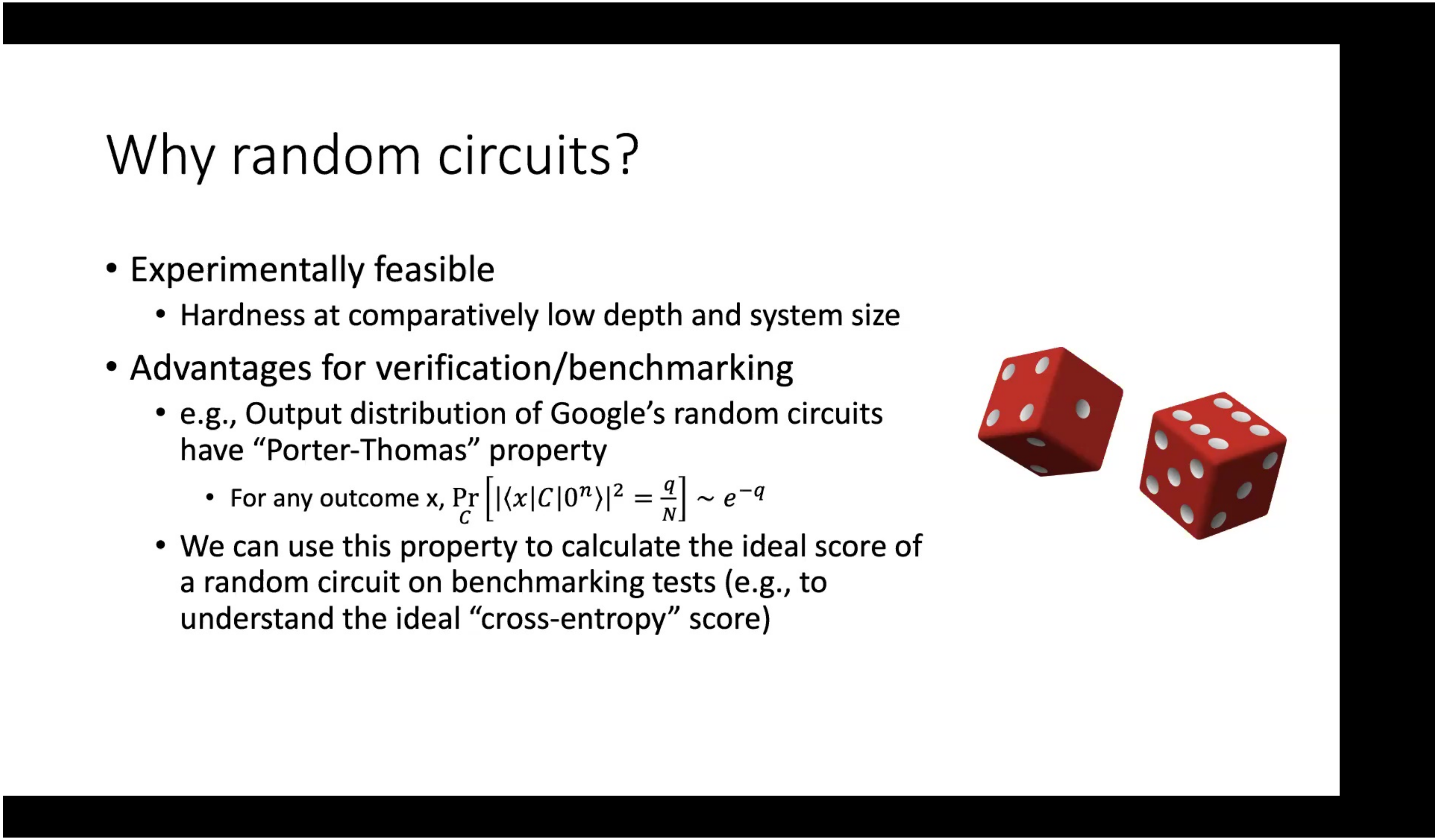
Gaussian BosonSampling (USTC “Jiuzhang”) in late 2020

This talk: the latest complexity theoretic evidence to believe these experiments might be solving hard problems for classical computers



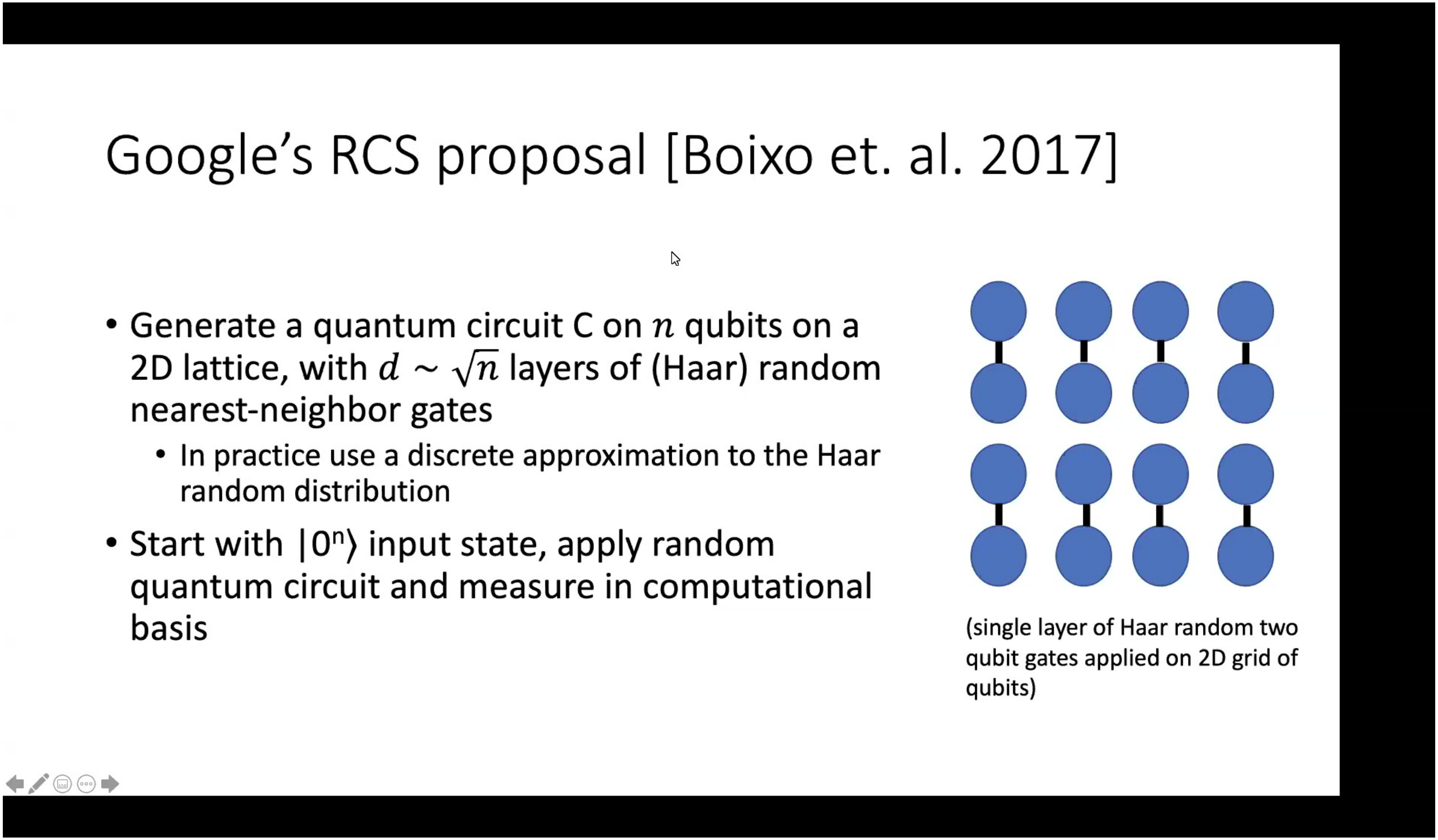
These experiments have a similar theoretical framework

- They both solve “random quantum circuit sampling”
  - i.e., the hard problem is to sample from the output distribution of a randomly chosen quantum circuit
- **Initial theory goal:** prove impossibility of an efficient *“classical Sampler”* algorithm that:
  - takes as input a random circuit C with output distribution D_C over $\{0,1\}^n$
  - outputs a sample from *any* distribution X so that:
    - $|X - D_C|_{TV} \leq \epsilon$ with high probability over choice of circuit C



Why random circuits?

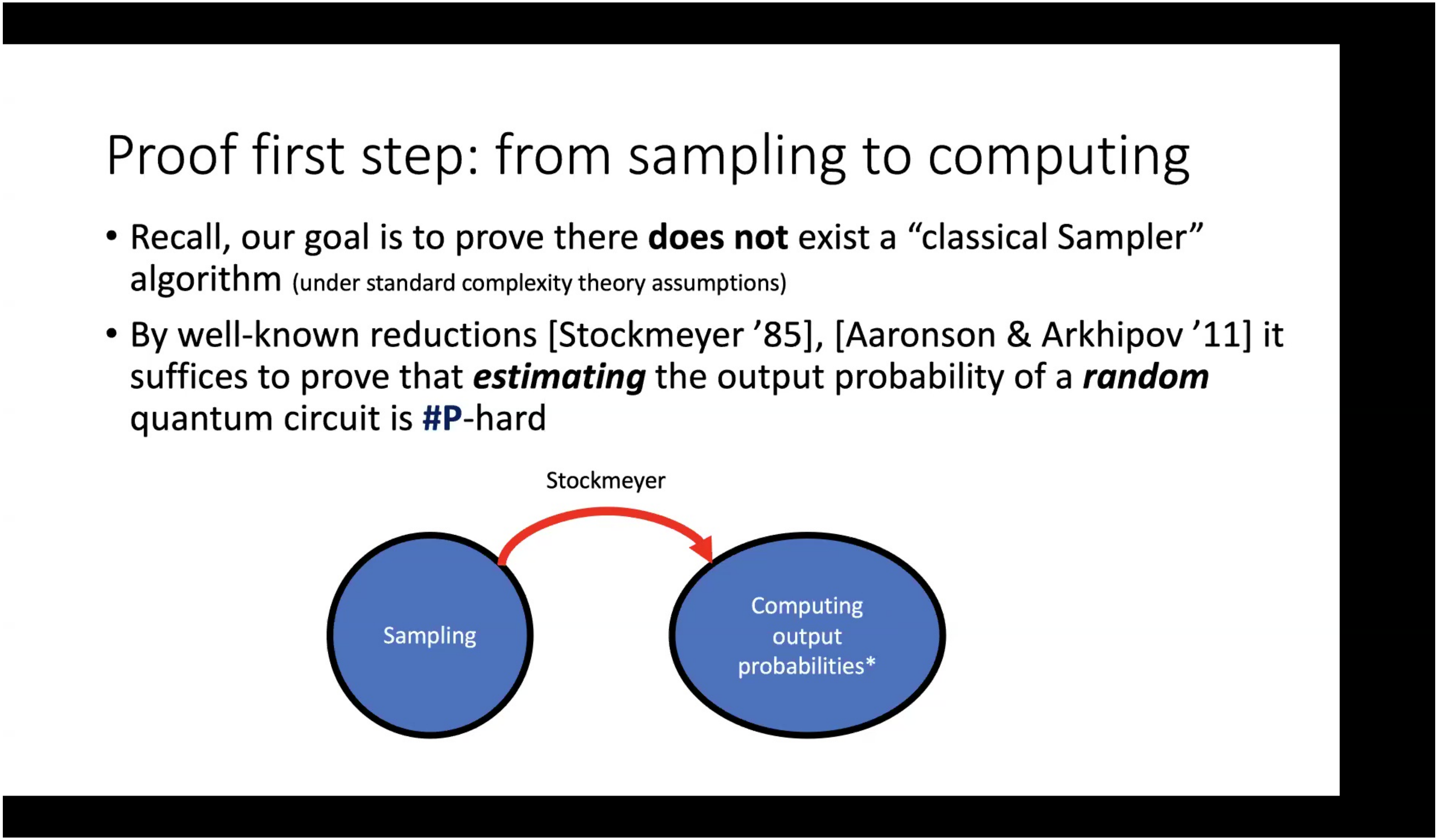
- Experimentally feasible
  - Hardness at comparatively low depth and system size
- Advantages for verification/benchmarking
  - e.g., Output distribution of Google's random circuits have "Porter-Thomas" property
    - For any outcome x, $\Pr_{C}\left[|\langle x|C|0^n\rangle|^2 = \frac{q}{N}\right] \sim e^{-q}$
  - We can use this property to calculate the ideal score of a random circuit on benchmarking tests (e.g., to understand the ideal "cross-entropy" score)



Google's RCS proposal [Boixo et. al. 2017]

- Generate a quantum circuit C on n qubits on a 2D lattice, with $d \sim \sqrt{n}$ layers of (Haar) random nearest-neighbor gates
  - In practice use a discrete approximation to the Haar random distribution
- Start with $|0^n\rangle$ input state, apply random quantum circuit and measure in computational basis

(single layer of Haar random two qubit gates applied on 2D grid of qubits)



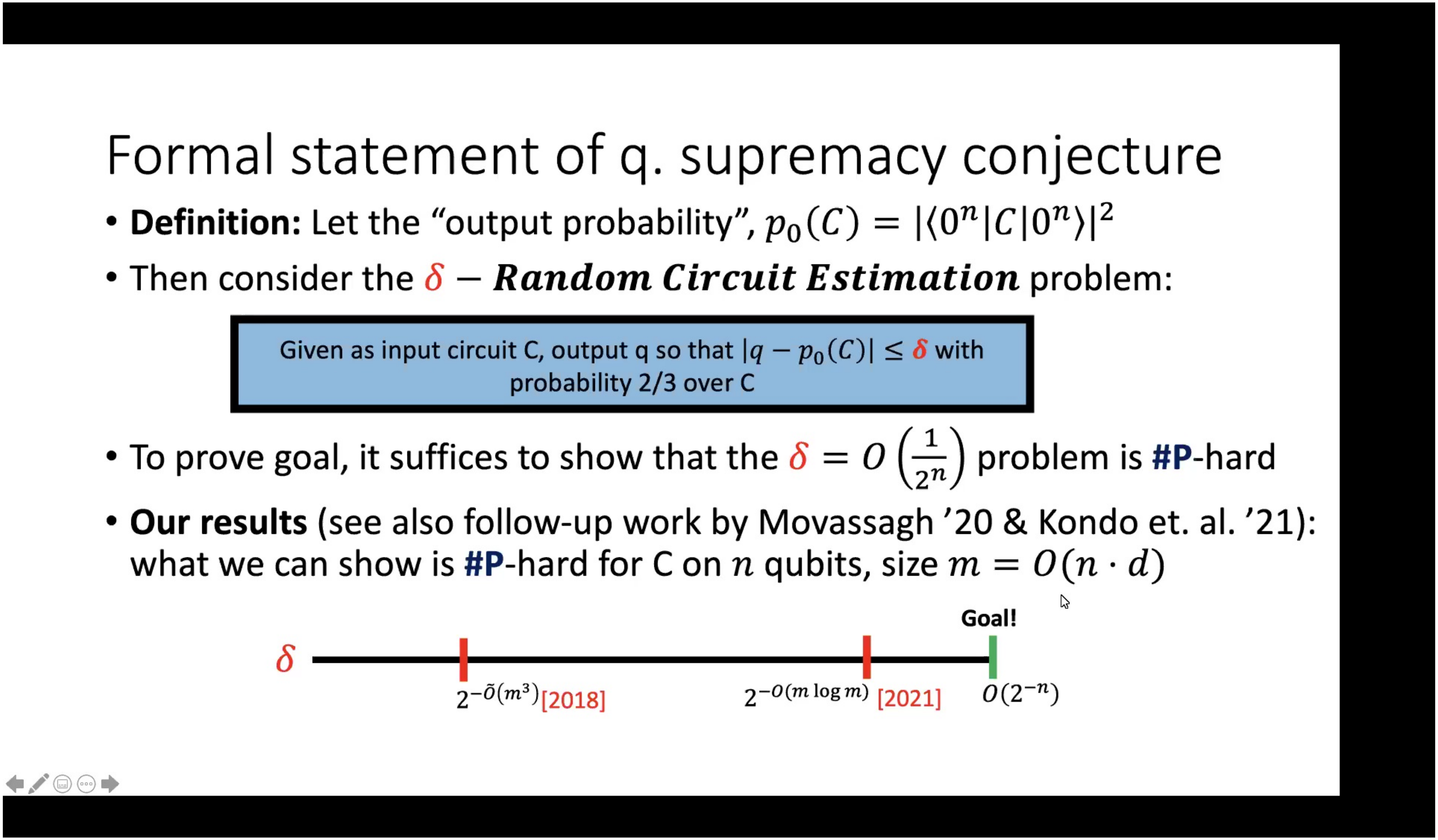
Proof first step: from sampling to computing

- Recall, our goal is to prove there **does not** exist a “classical Sampler” algorithm (under standard complexity theory assumptions)
- By well-known reductions [Stockmeyer ’85], [Aaronson & Arkhipov ’11] it suffices to prove that ***estimating*** the output probability of a ***random*** quantum circuit is **#P**-hard

Stockmeyer

Sampling

Computing output probabilities*

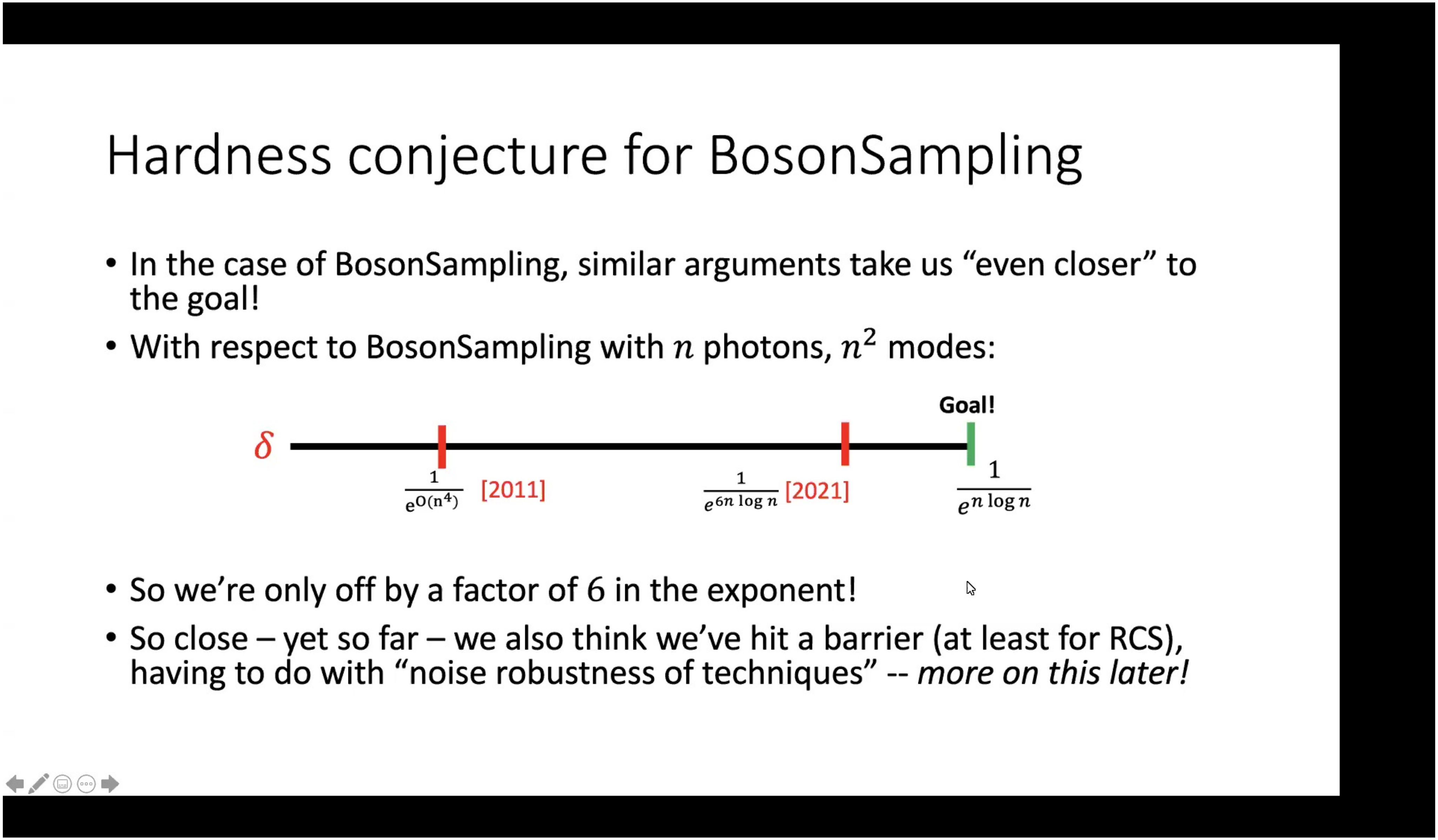


Formal statement of q. supremacy conjecture

- **Definition:** Let the “output probability”, $p_0(C) = |\langle 0^n | C | 0^n \rangle|^2$
- Then consider the $\delta - \boldsymbol{Random\ Circuit\ Estimation}$ problem:

Given as input circuit C, output q so that $|q - p_0(C)| \leq \boldsymbol{\delta}$ with probability 2/3 over C

- To prove goal, it suffices to show that the $\delta = O\left(\frac{1}{2^n}\right)$ problem is **#P**-hard
- **Our results** (see also follow-up work by Movassagh ’20 & Kondo et. al. ’21): what we can show is **#P**-hard for C on n qubits, size $m = O(n \cdot d)$

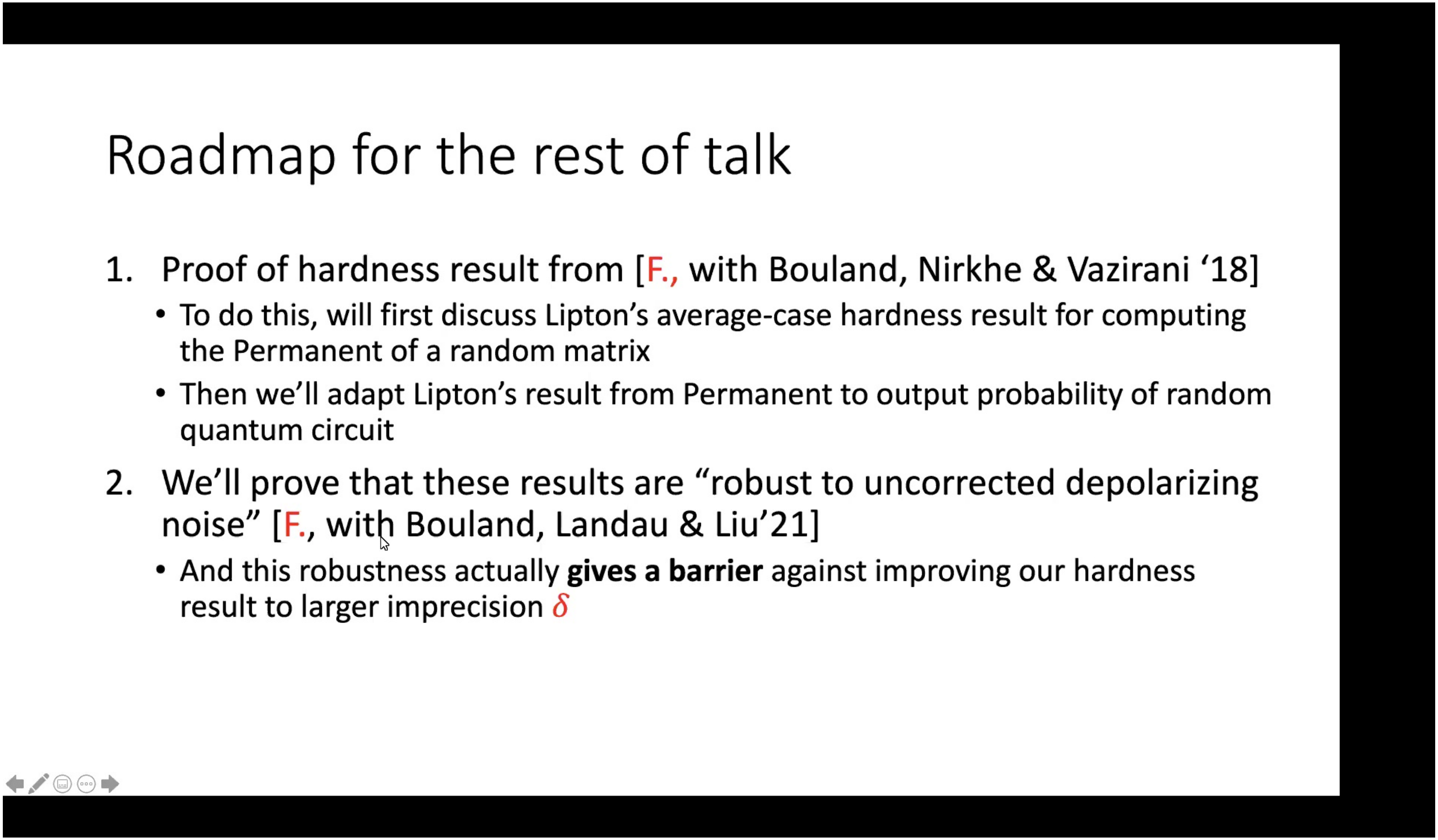


Hardness conjecture for BosonSampling

- In the case of BosonSampling, similar arguments take us “even closer” to the goal!
- With respect to BosonSampling with n photons, n^2 modes:

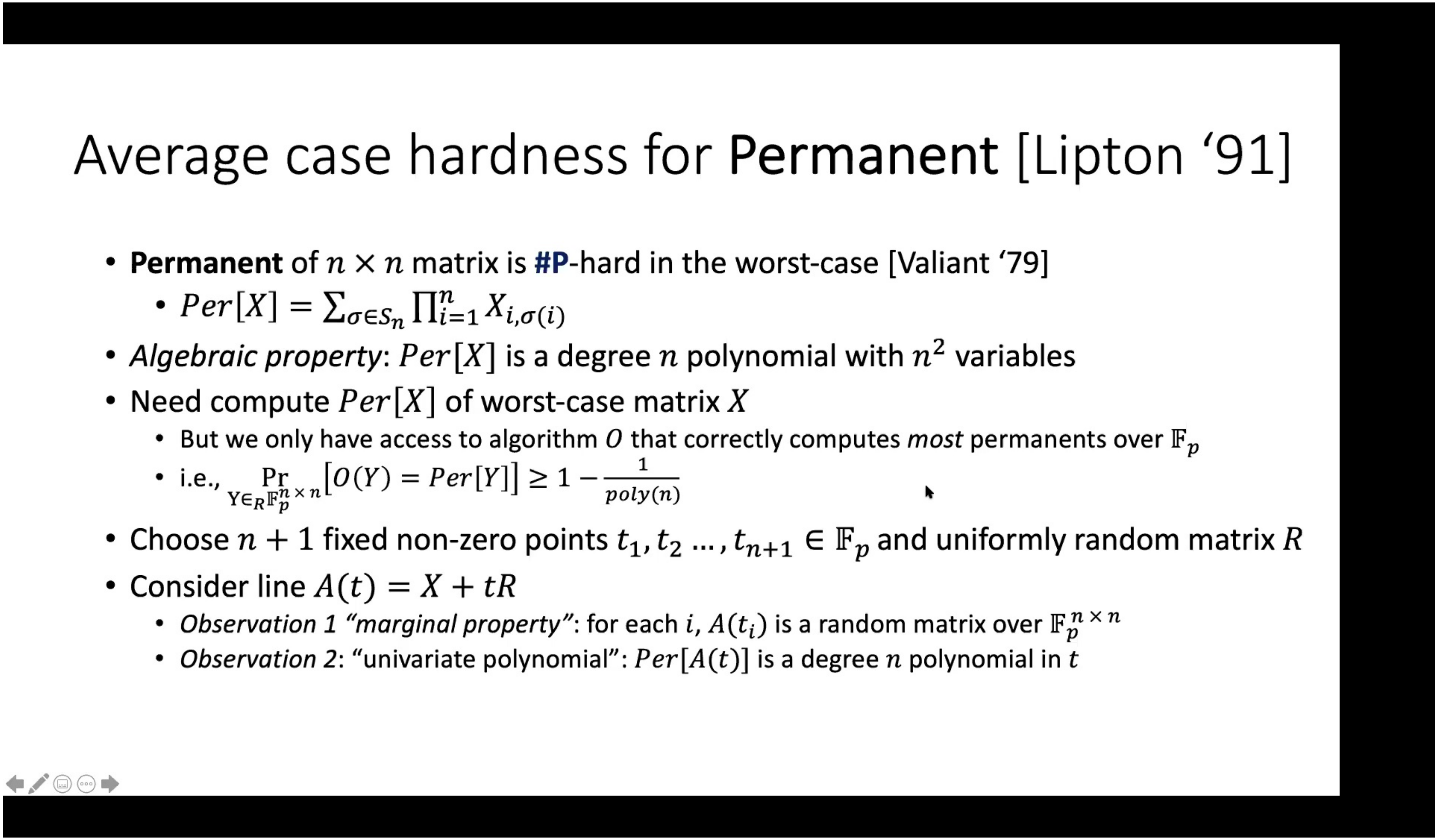
Goal!
δ
$\frac{1}{e^{O(n^4)}}$ [2011]
$\frac{1}{e^{6n \log n}}$ [2021]
$\frac{1}{e^{n \log n}}$

- So we’re only off by a factor of 6 in the exponent!
- So close – yet so far – we also think we’ve hit a barrier (at least for RCS), having to do with “noise robustness of techniques” -- *more on this later!*



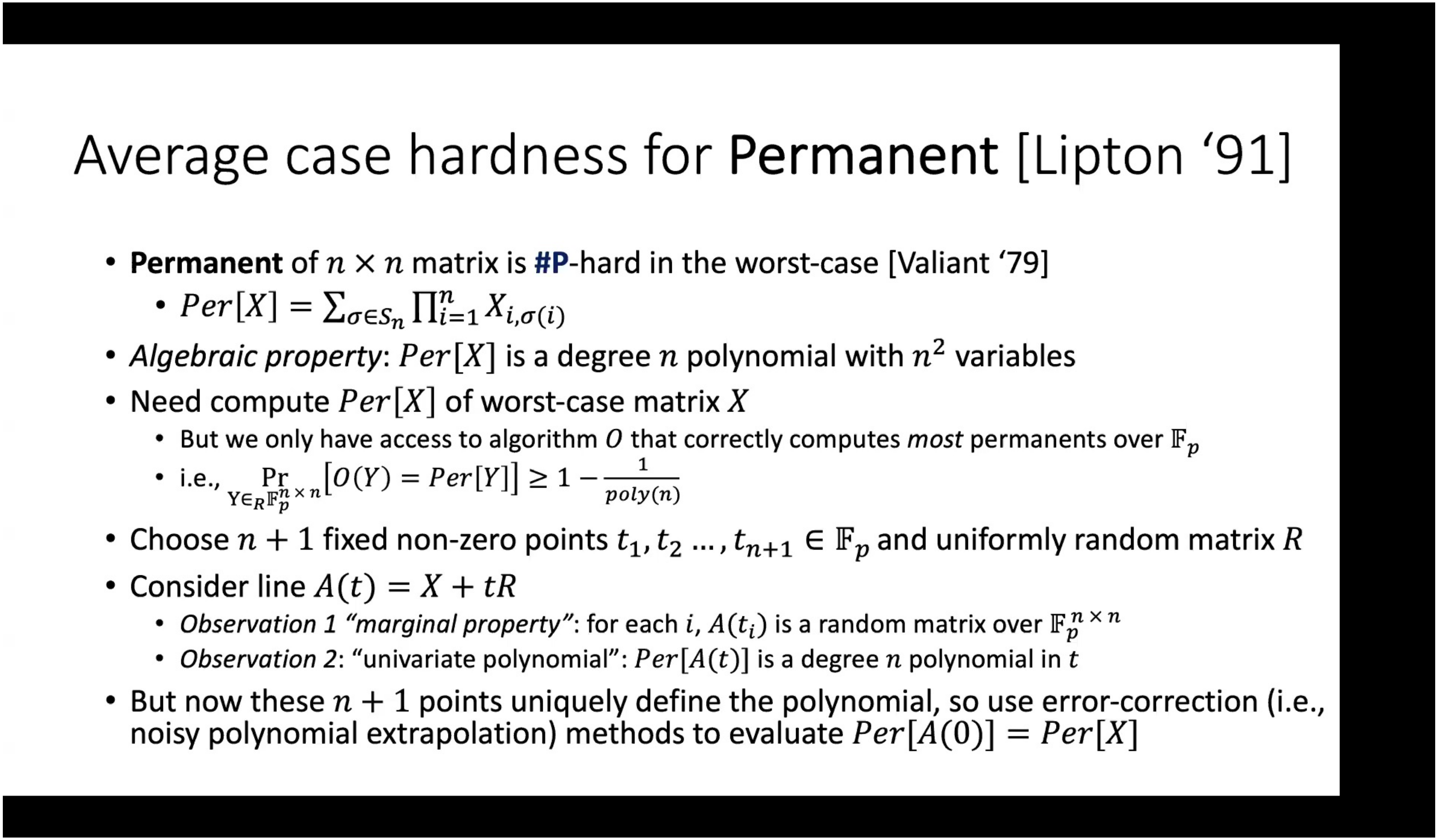
Roadmap for the rest of talk

1. Proof of hardness result from [F., with Bouland, Nirkhe & Vazirani ‘18]
   - To do this, will first discuss Lipton’s average-case hardness result for computing the Permanent of a random matrix
   - Then we’ll adapt Lipton’s result from Permanent to output probability of random quantum circuit
2. We’ll prove that these results are “robust to uncorrected depolarizing noise” [F., with Bouland, Landau & Liu’21]
   - And this robustness actually **gives a barrier** against improving our hardness result to larger imprecision δ



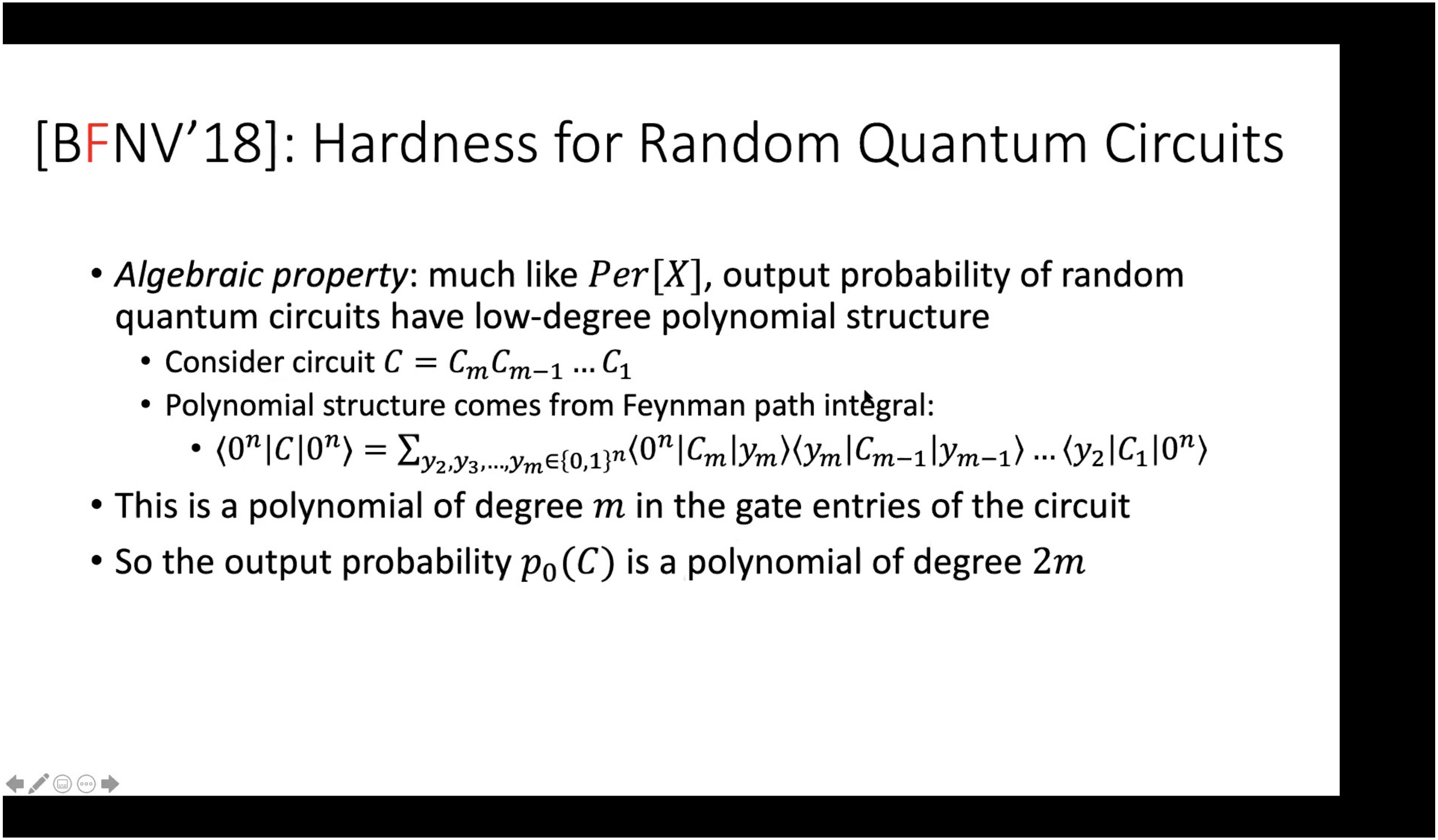
Average case hardness for **Permanent** [Lipton '91]

- **Permanent** of $n \times n$ matrix is **#P**-hard in the worst-case [Valiant '79]
  - $Per[X] = \sum_{\sigma \in S_n} \prod_{i=1}^{n} X_{i,\sigma(i)}$
- *Algebraic property*: $Per[X]$ is a degree n polynomial with n^2 variables
- Need compute $Per[X]$ of worst-case matrix X
  - But we only have access to algorithm O that correctly computes *most* permanents over $\mathbb{F}_p$
  - i.e., $\Pr_{Y \in_R \mathbb{F}_p^{n \times n}}\left[O(Y) = Per[Y]\right] \geq 1 - \frac{1}{poly(n)}$
- Choose $n + 1$ fixed non-zero points $t_1, t_2 \ldots, t_{n+1} \in \mathbb{F}_p$ and uniformly random matrix R
- Consider line $A(t) = X + tR$
  - *Observation 1 "marginal property"*: for each i, $A(t_i)$ is a random matrix over $\mathbb{F}_p^{n \times n}$
  - *Observation 2*: "univariate polynomial": $Per[A(t)]$ is a degree n polynomial in t



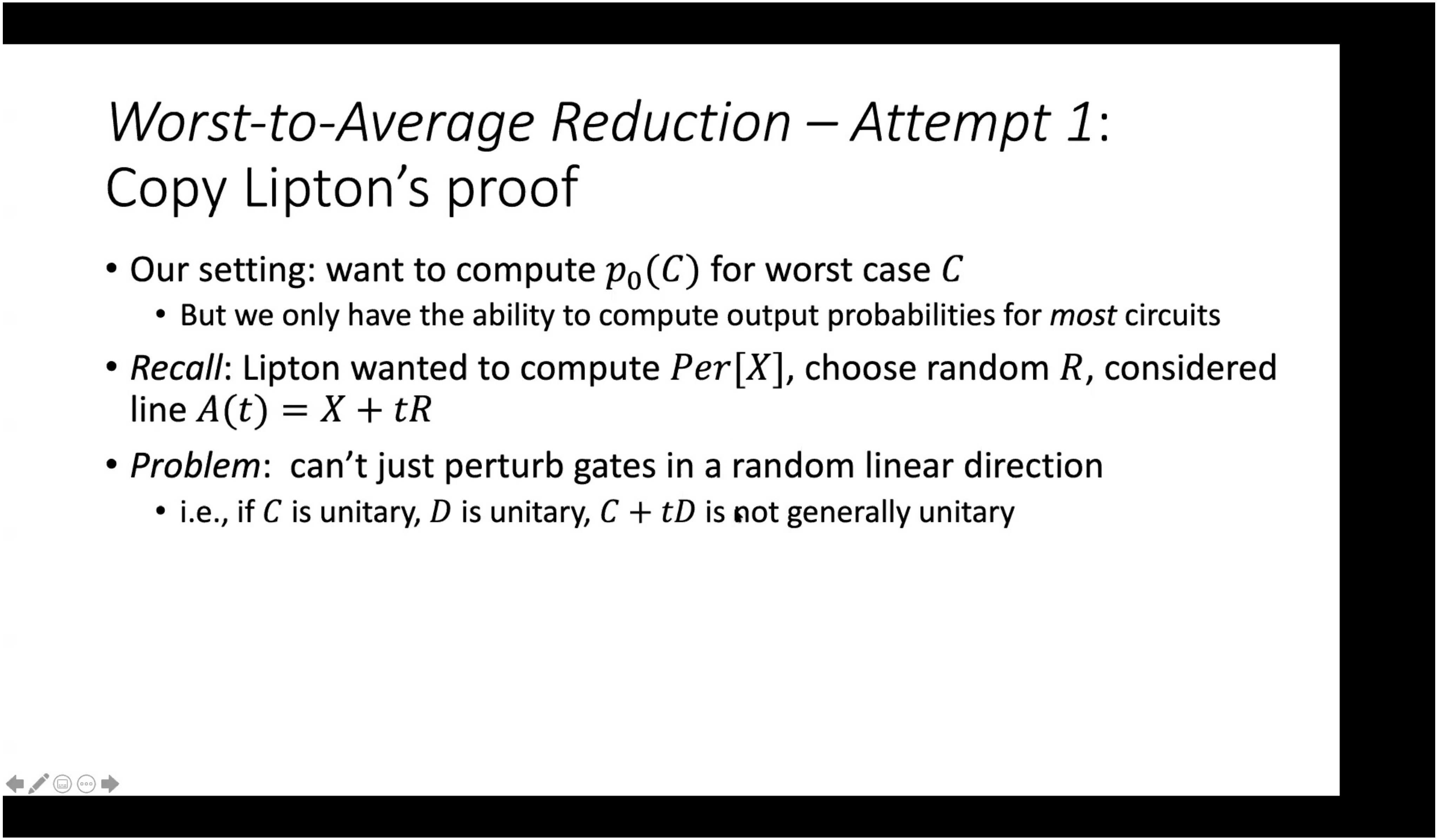
Average case hardness for **Permanent** [Lipton '91]

- **Permanent** of $n \times n$ matrix is **#P**-hard in the worst-case [Valiant '79]
  - $Per[X] = \sum_{\sigma \in S_n} \prod_{i=1}^{n} X_{i,\sigma(i)}$
- *Algebraic property*: $Per[X]$ is a degree n polynomial with n^2 variables
- Need compute $Per[X]$ of worst-case matrix X
  - But we only have access to algorithm O that correctly computes *most* permanents over $\mathbb{F}_p$
  - i.e., $\Pr_{Y \in_R \mathbb{F}_p^{n \times n}}\left[O(Y) = Per[Y]\right] \geq 1 - \frac{1}{poly(n)}$
- Choose $n + 1$ fixed non-zero points $t_1, t_2 \ldots, t_{n+1} \in \mathbb{F}_p$ and uniformly random matrix R
- Consider line $A(t) = X + tR$
  - *Observation 1 "marginal property"*: for each i, $A(t_i)$ is a random matrix over $\mathbb{F}_p^{n \times n}$
  - *Observation 2*: "univariate polynomial": $Per[A(t)]$ is a degree n polynomial in t
- But now these $n + 1$ points uniquely define the polynomial, so use error-correction (i.e., noisy polynomial extrapolation) methods to evaluate $Per[A(0)] = Per[X]$



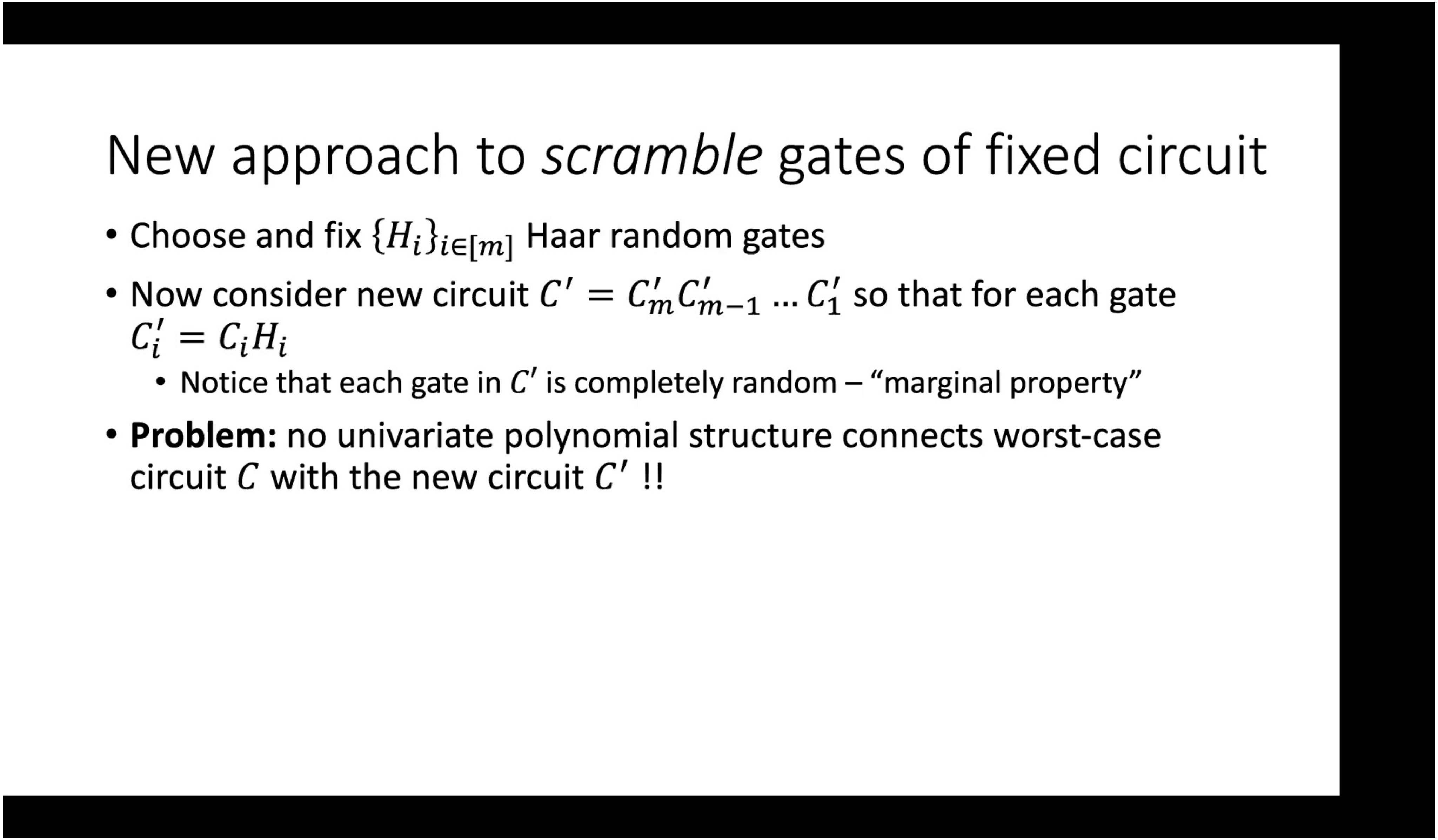
[BFNV’18]: Hardness for Random Quantum Circuits

- *Algebraic property*: much like $Per[X]$, output probability of random quantum circuits have low-degree polynomial structure
  - Consider circuit $C = C_m C_{m-1} \ldots C_1$
  - Polynomial structure comes from Feynman path integral:
    - $\langle 0^n | C | 0^n \rangle = \sum_{y_2, y_3, \ldots, y_m \in \{0,1\}^n} \langle 0^n | C_m | y_m \rangle \langle y_m | C_{m-1} | y_{m-1} \rangle \ldots \langle y_2 | C_1 | 0^n \rangle$
- This is a polynomial of degree m in the gate entries of the circuit
- So the output probability $p_0(C)$ is a polynomial of degree $2m$



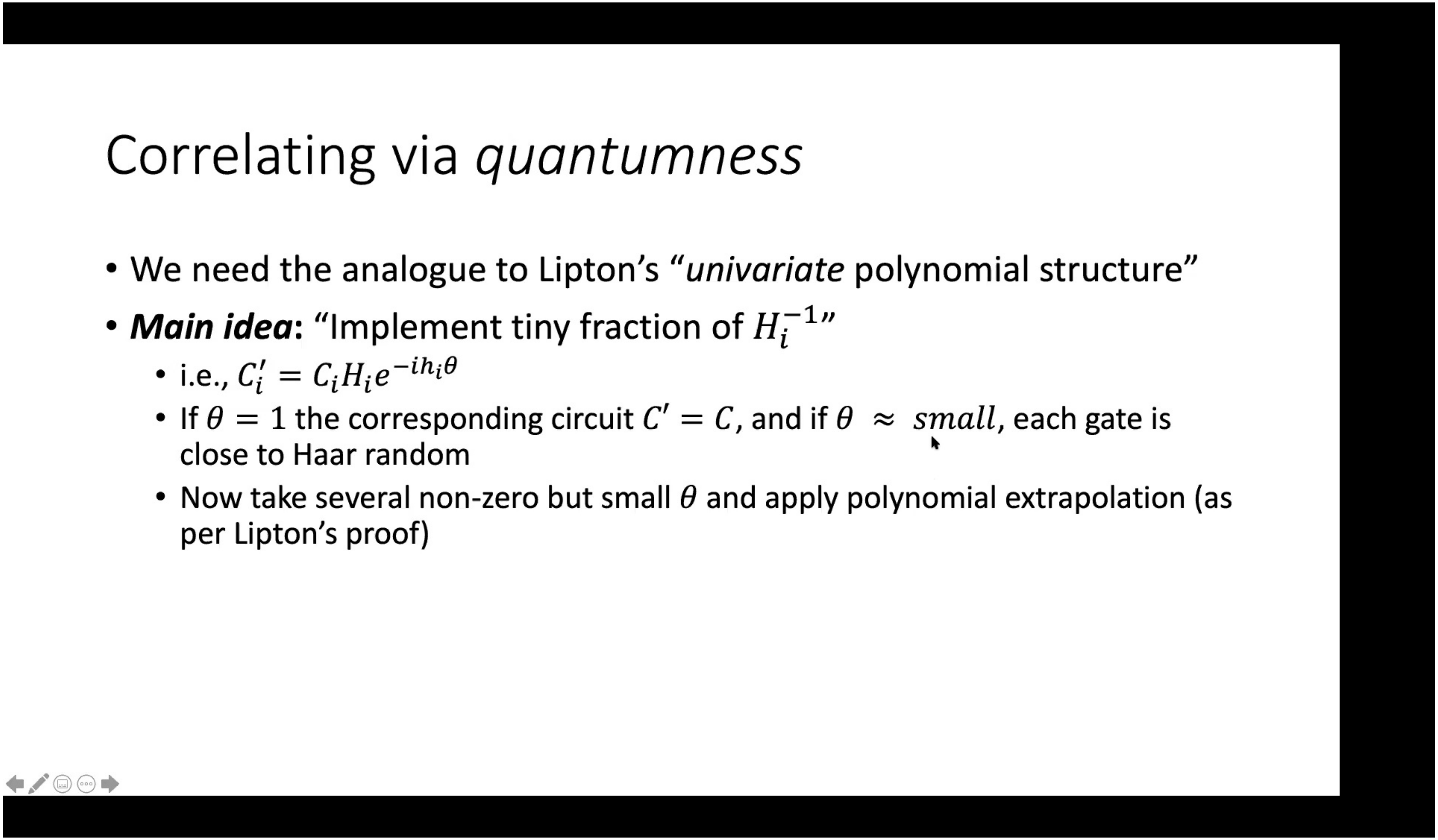
Worst-to-Average Reduction – Attempt 1: Copy Lipton's proof

- Our setting: want to compute $p_0(C)$ for worst case C
  - But we only have the ability to compute output probabilities for *most* circuits
- *Recall*: Lipton wanted to compute $Per[X]$, choose random R, considered line $A(t) = X + tR$
- *Problem*: can't just perturb gates in a random linear direction
  - i.e., if C is unitary, D is unitary, $C + tD$ is not generally unitary



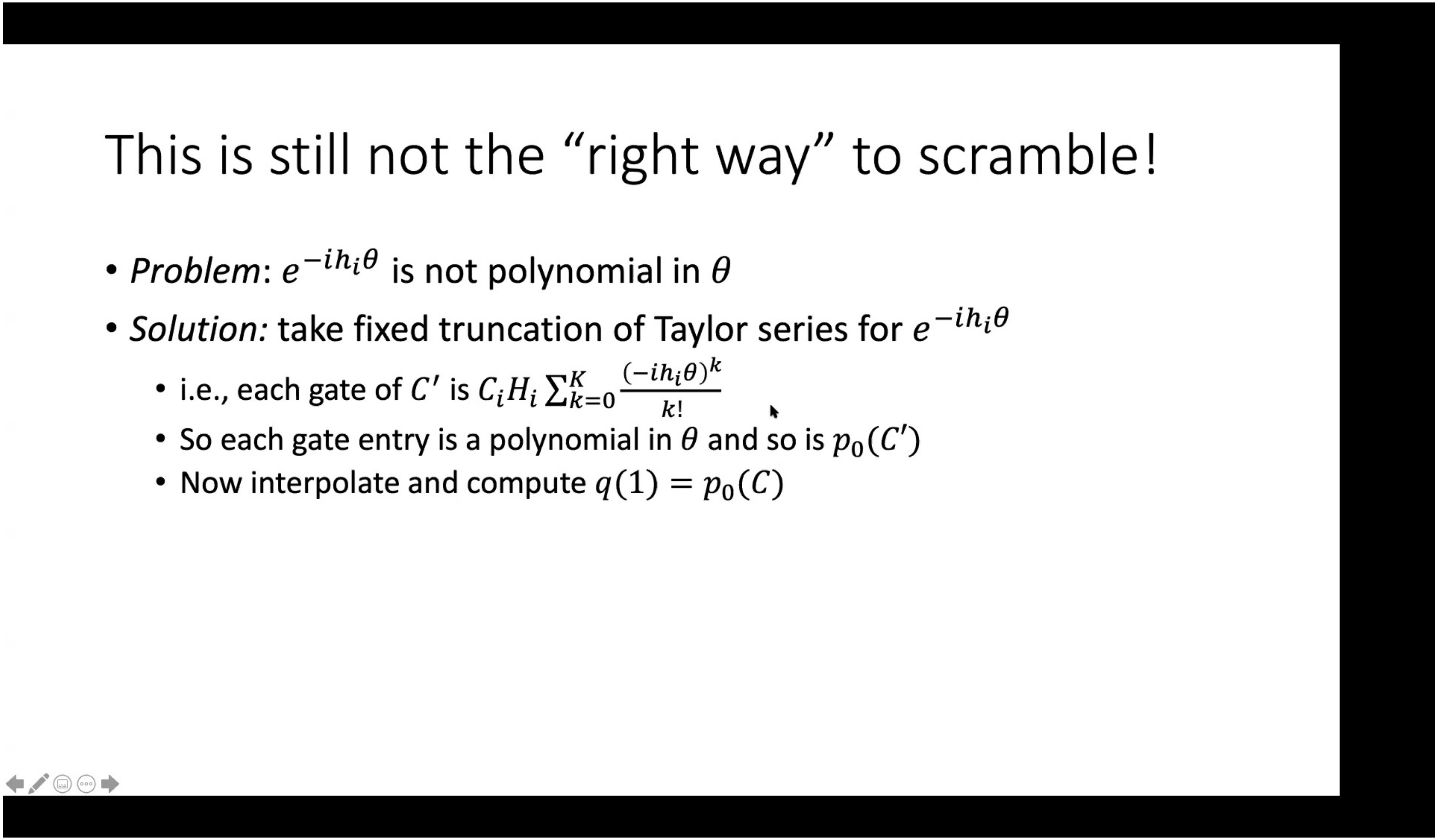
New approach to *scramble* gates of fixed circuit

- Choose and fix $\{H_i\}_{i\in[m]}$ Haar random gates
- Now consider new circuit $C' = C'_m C'_{m-1} \ldots C'_1$ so that for each gate $C'_i = C_i H_i$
  - Notice that each gate in C' is completely random – “marginal property”
- **Problem:** no univariate polynomial structure connects worst-case circuit C with the new circuit C' !!



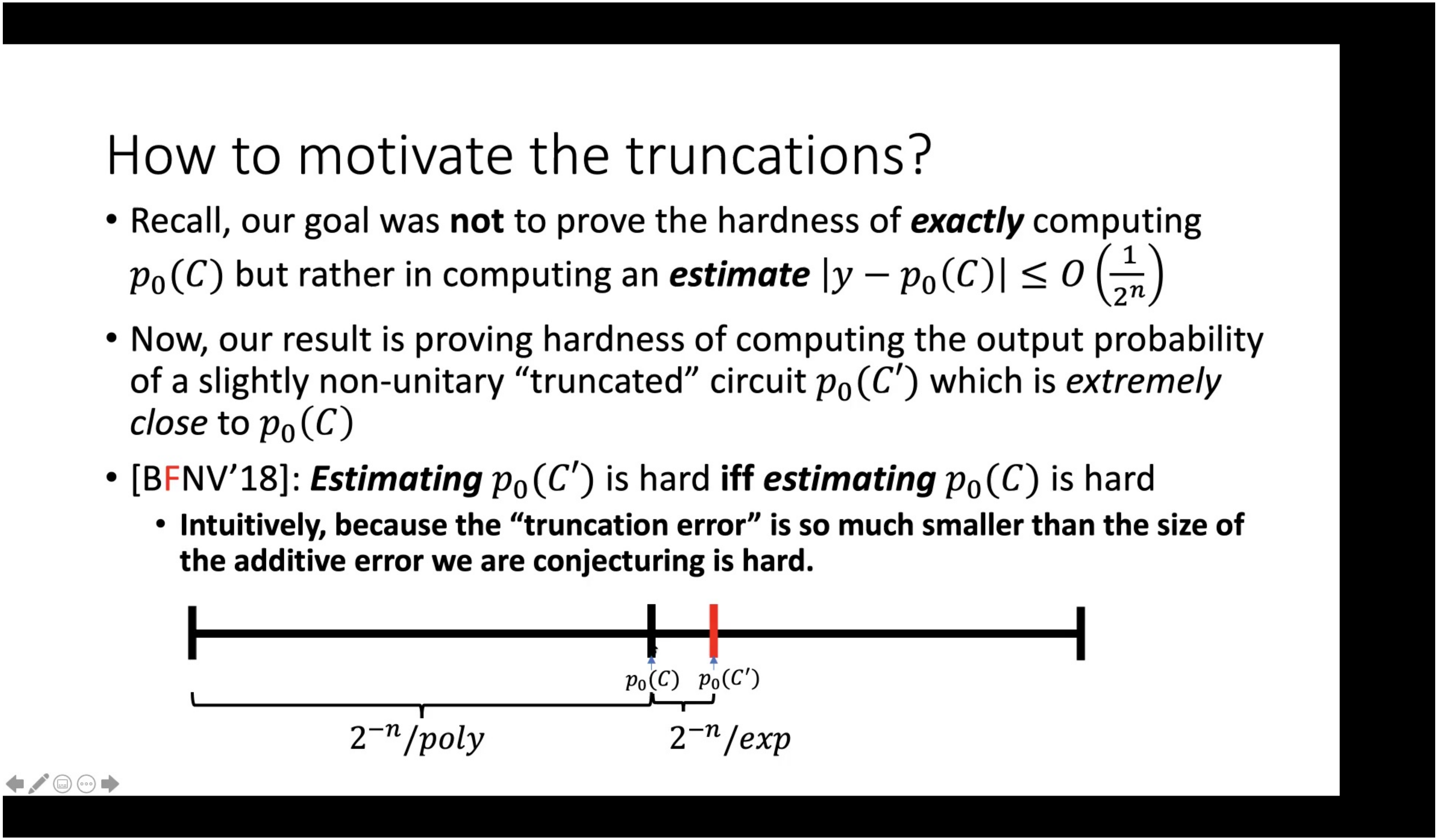
Correlating via *quantumness*

- We need the analogue to Lipton's "*univariate* polynomial structure"
- ***Main idea***: "Implement tiny fraction of H_i^{-1}"
  - i.e., $C_i' = C_i H_i e^{-ih_i\theta}$
  - If $\theta = 1$ the corresponding circuit $C' = C$, and if $\theta \approx small$, each gate is close to Haar random
  - Now take several non-zero but small θ and apply polynomial extrapolation (as per Lipton's proof)



This is still not the “right way” to scramble!

- *Problem*: $e^{-ih_i\theta}$ is not polynomial in θ
- *Solution:* take fixed truncation of Taylor series for $e^{-ih_i\theta}$
    - i.e., each gate of C' is $C_i H_i \sum_{k=0}^{K} \frac{(-ih_i\theta)^k}{k!}$
    - So each gate entry is a polynomial in θ and so is $p_0(C')$
    - Now interpolate and compute $q(1) = p_0(C)$

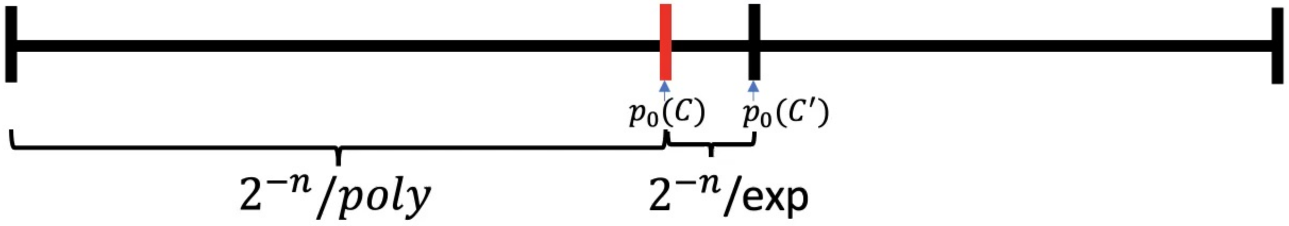


How to motivate the truncations?

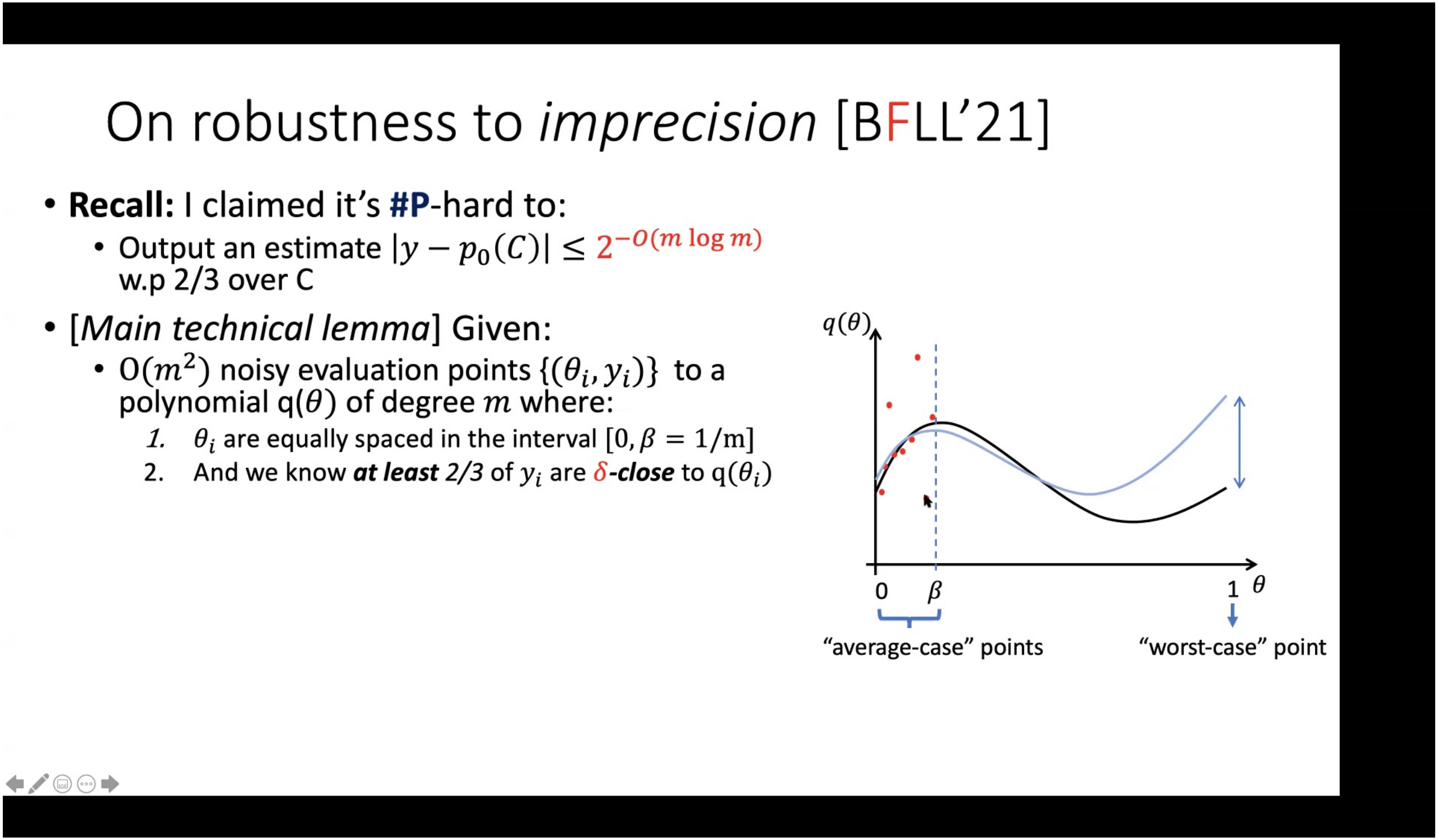
- Recall, our goal was **not** to prove the hardness of ***exactly*** computing $p_0(C)$ but rather in computing an ***estimate*** $|y - p_0(C)| \leq O\left(\frac{1}{2^n}\right)$
- Now, our result is proving hardness of computing the output probability of a slightly non-unitary "truncated" circuit $p_0(C')$ which is *extremely close* to $p_0(C)$
- [BFNV'18]: ***Estimating*** $p_0(C')$ is hard **iff *estimating*** $p_0(C)$ is hard
    - **Intuitively, because the "truncation error" is so much smaller than the size of the additive error we are conjecturing is hard.**

Movassagh’s result

- In recent follow-up work, hardness has been shown around original output probability
  - i.e., computing $p_0(C) \pm 2^{-n^3}$ is **#P**-hard [Movassagh ’19,’20]



- To do this, Movassagh gives a new method to interpolate between the worst-case and random quantum circuit, using the “Cayley path”, which stays unitary throughout the entire path

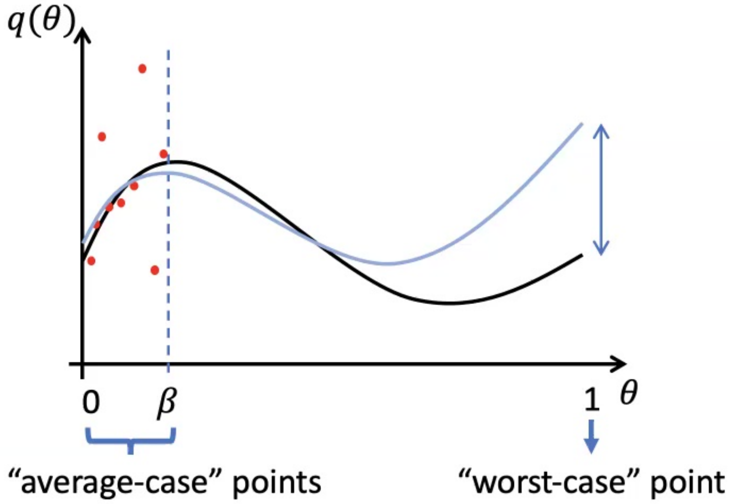


On robustness to *imprecision* [BFLL'21]

- **Recall:** I claimed it's **#P**-hard to:
  - Output an estimate $|y - p_0(C)| \leq 2^{-O(m \log m)}$ w.p 2/3 over C
- [*Main technical lemma*] Given:
  - $\mathrm{O}(m^2)$ noisy evaluation points $\{(\theta_i, y_i)\}$ to a polynomial $\mathrm{q}(\theta)$ of degree m where:
    1. θ_i are equally spaced in the interval $[0, \beta = 1/\mathrm{m}]$
    2. And we know ***at least 2/3*** of y_i are δ***-close*** to $\mathrm{q}(\theta_i)$

On robustness to *imprecision* [BFLL’21]

- **Recall:** I claimed it’s **#P**-hard to:
  - Output an estimate $|y - p_0(C)| \leq 2^{-O(m \log m)}$ w.p 2/3 over C
- [*Main technical lemma*] Given:
  - $O(m^2)$ noisy evaluation points $\{(\theta_i, y_i)\}$ to a polynomial q(θ) of degree m where:
    1. θ_i are equally spaced in the interval $[0, \beta = 1/\mathrm{m}]$
    2. And we know ***at least** 2/3* of y_i are δ***-close*** to $\mathrm{q}(\theta_i)$
- Then there’s an algorithm that outputs z:
  - $|z - q(1)| \leq \delta 2^{O(m \log \beta^{-1})} = \delta 2^{O(m \log m)}$ whp
- **Upshot:** if δ is small enough then we can use these noisy evaluation points (corresponding to random circuits) to estimate q(1), which is hard!

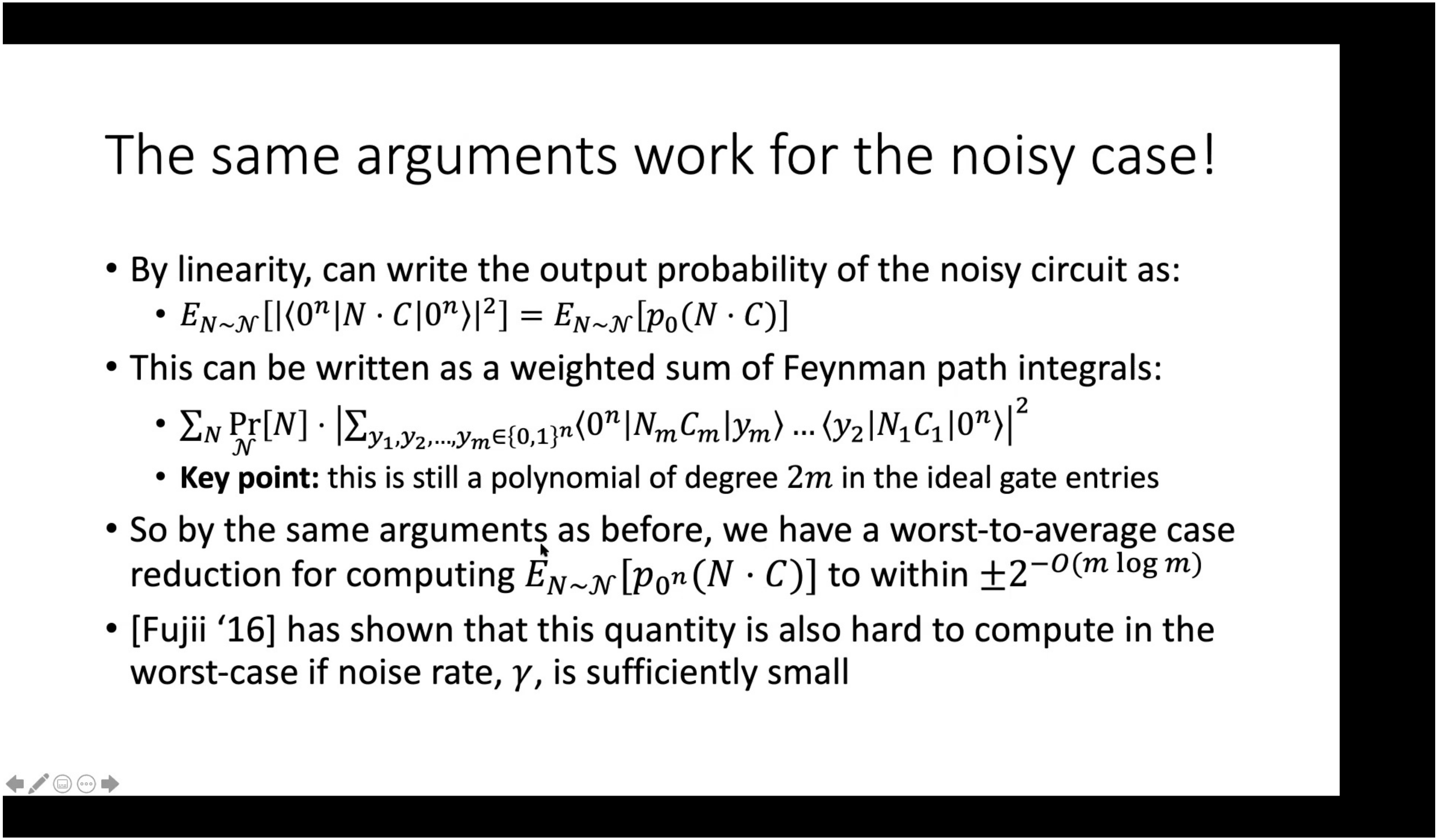


“average-case” points

“worst-case” point

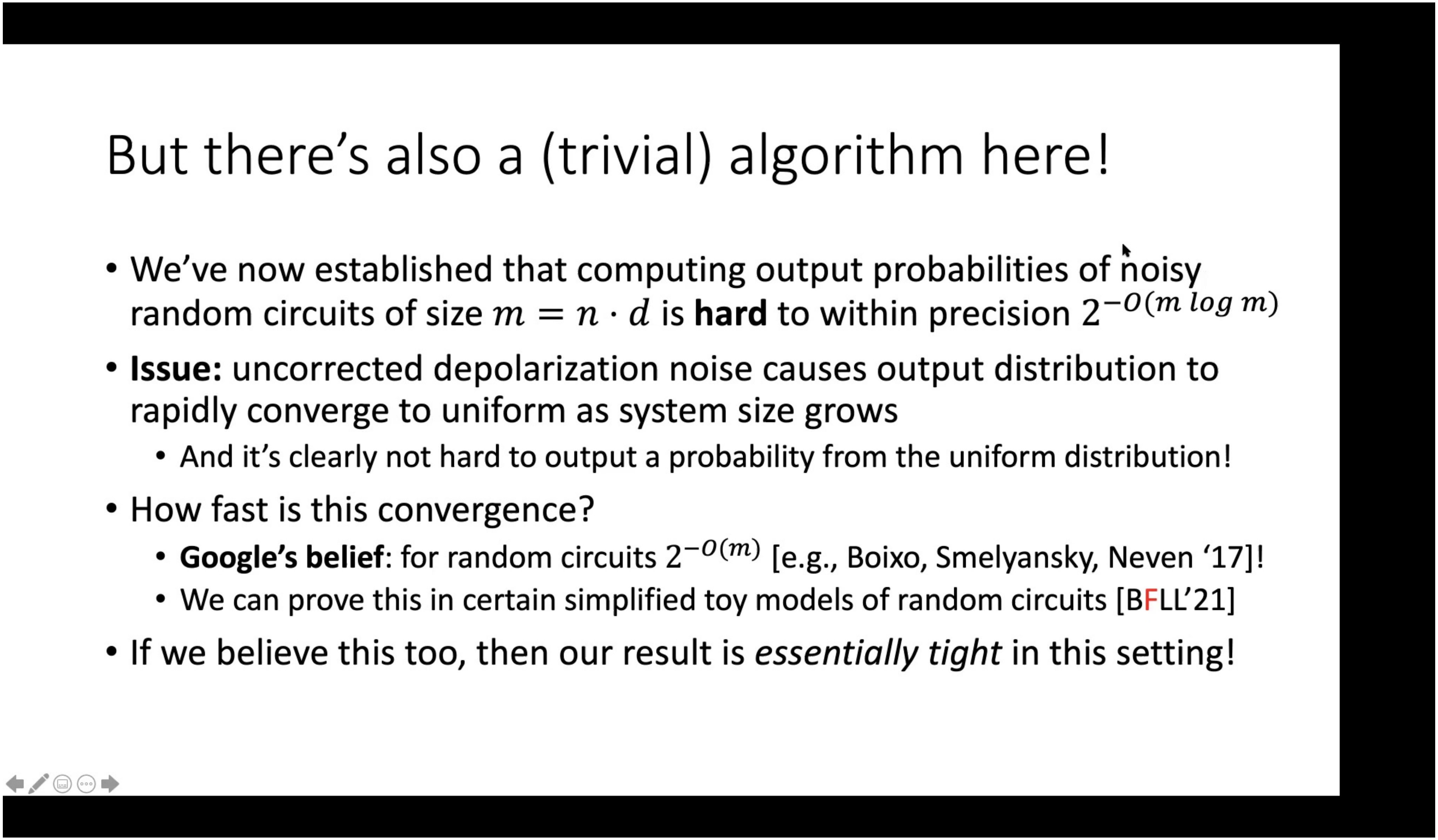
Hardness of noisy random circuits [BFLL’21]

- Without error-correction noise eventually overwhelms
    - e.g., Google’s RCS experiment ~0.2% fidelity and 99.8% noise
- How can we model this theoretically for RCS?
- Each random gate C_i is followed by two qubit depolarizing noise channel:
    - $\mathcal{E}_i = (1-\gamma)\rho + \frac{\gamma}{15}\sum_{\alpha,\beta\in\mathcal{P}\times\mathcal{P}-(I,I)}(\sigma_\alpha \otimes \sigma_\beta)\rho(\sigma_\alpha \otimes \sigma_\beta)$
- That is, we can think about choosing a noisy random circuit by:
    - First pick ideal circuit $C = C_m C_{m-1} \dots C_1$ from the random circuit distribution
    - Then environment chooses operators N, from a distribution $\mathcal{N}$ (i.e., via $\mathcal{E}_i$)
    - We get a sample from output distribution of $N \cdot C$ without learning the noise operators



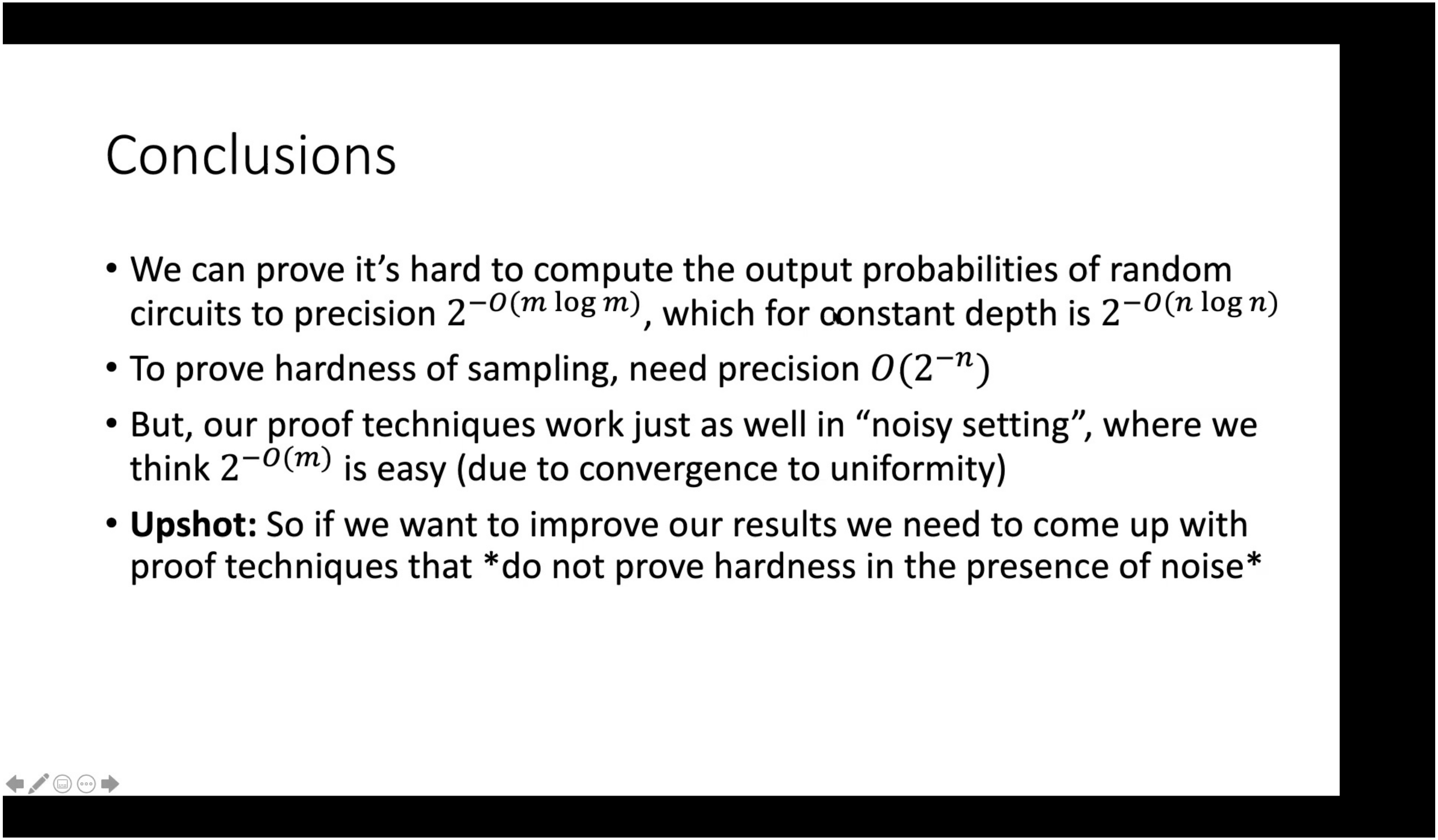
The same arguments work for the noisy case!

- By linearity, can write the output probability of the noisy circuit as:
  - $E_{N\sim\mathcal{N}}[|\langle 0^n|N \cdot C|0^n\rangle|^2] = E_{N\sim\mathcal{N}}[p_0(N \cdot C)]$
- This can be written as a weighted sum of Feynman path integrals:
  - $\sum_N \Pr_{\mathcal{N}}[N] \cdot \left|\sum_{y_1,y_2,\ldots,y_m\in\{0,1\}^n}\langle 0^n|N_m C_m|y_m\rangle \ldots \langle y_2|N_1 C_1|0^n\rangle\right|^2$
  - **Key point:** this is still a polynomial of degree $2m$ in the ideal gate entries
- So by the same arguments as before, we have a worst-to-average case reduction for computing $E_{N\sim\mathcal{N}}[p_{0^n}(N \cdot C)]$ to within $\pm 2^{-O(m\log m)}$
- [Fujii '16] has shown that this quantity is also hard to compute in the worst-case if noise rate, γ, is sufficiently small



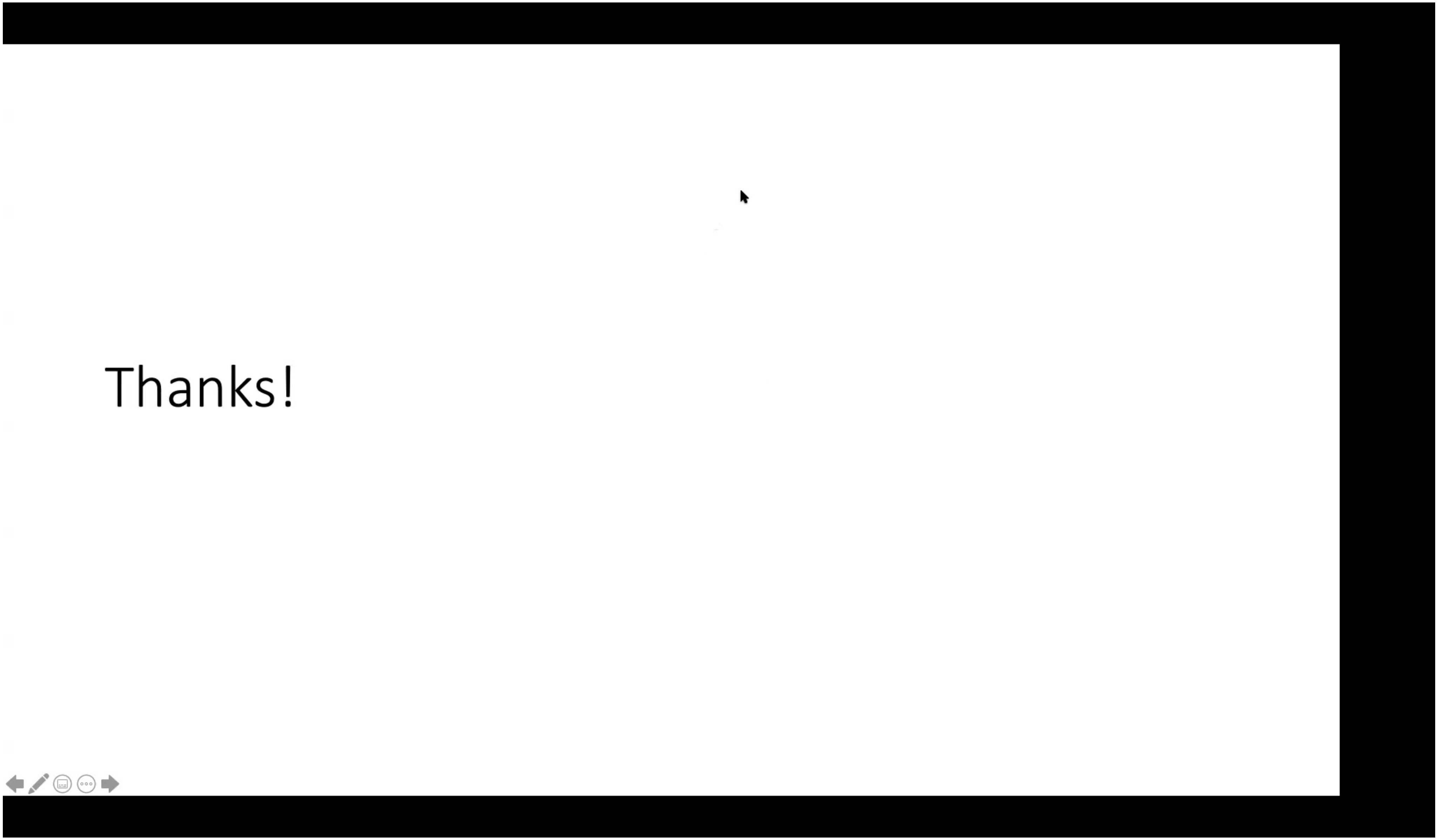
But there’s also a (trivial) algorithm here!

- We’ve now established that computing output probabilities of noisy random circuits of size $m = n \cdot d$ is **hard** to within precision $2^{-O(m \log m)}$
- **Issue:** uncorrected depolarization noise causes output distribution to rapidly converge to uniform as system size grows
  - And it’s clearly not hard to output a probability from the uniform distribution!
- How fast is this convergence?
  - **Google’s belief**: for random circuits $2^{-O(m)}$ [e.g., Boixo, Smelyansky, Neven ‘17]!
  - We can prove this in certain simplified toy models of random circuits [BFLL’21]
- If we believe this too, then our result is *essentially tight* in this setting!



Conclusions

- We can prove it’s hard to compute the output probabilities of random circuits to precision $2^{-O(m\log m)}$, which for constant depth is $2^{-O(n\log n)}$
- To prove hardness of sampling, need precision $O(2^{-n})$
- But, our proof techniques work just as well in “noisy setting”, where we think $2^{-O(m)}$ is easy (due to convergence to uniformity)
- **Upshot:** So if we want to improve our results we need to come up with proof techniques that *do not prove hardness in the presence of noise*



Thanks!