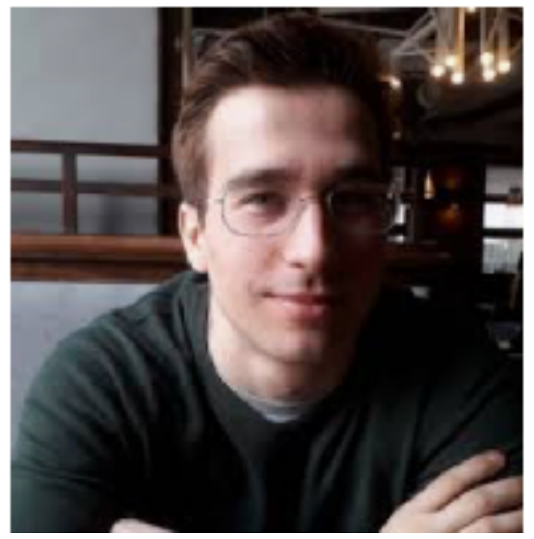
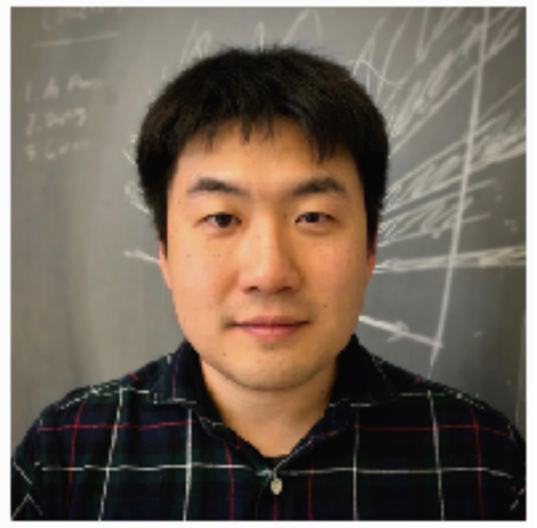
Limitations of Linear Cross-Entropy as a Measure for Quantum Advantage





Xun Gao, Marcin Kalinowski, Chi-Ning Chou, Misha Lukin, Boaz Barak — Harvard University

Soonwon Choi — MIT

arXiv:2112.01657

A Huge Gap Between Theories and Practices

Quantum Computational Advantage

A Computation Problem

Physical Implementations

- Google.
- USTC.
- More to come…

Classical Hardness

- Complexity-theoretic foundations.
- Heuristic arguments.

Practical motivation: constituting a milestone for quantum technology.

Theoretical motivation: challenging the extended Church-Turing thesis.

Example: Google's Quantum Supremacy Claim

nature

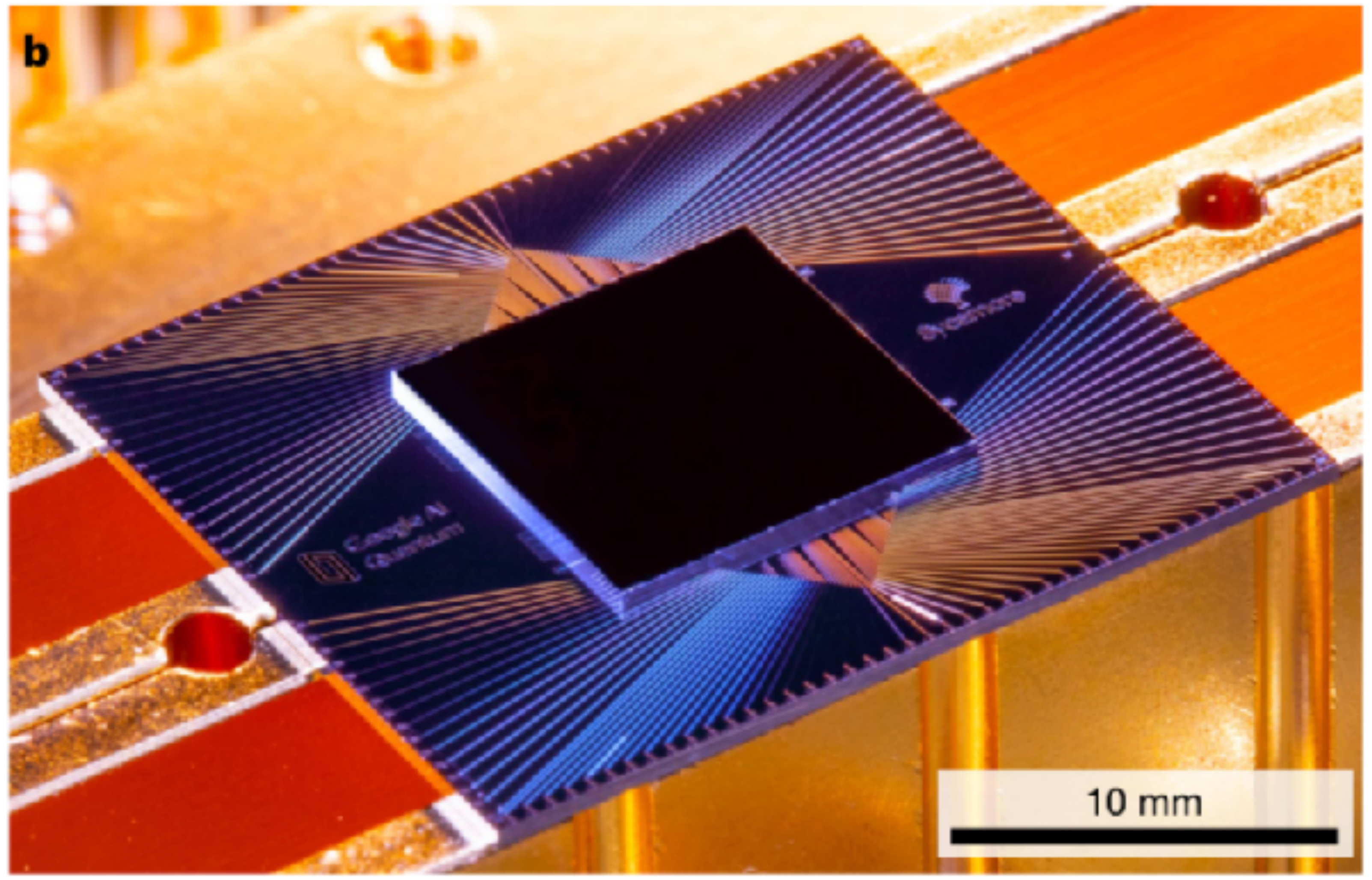
Explore content · About the journal · Publish with us · Subscribe

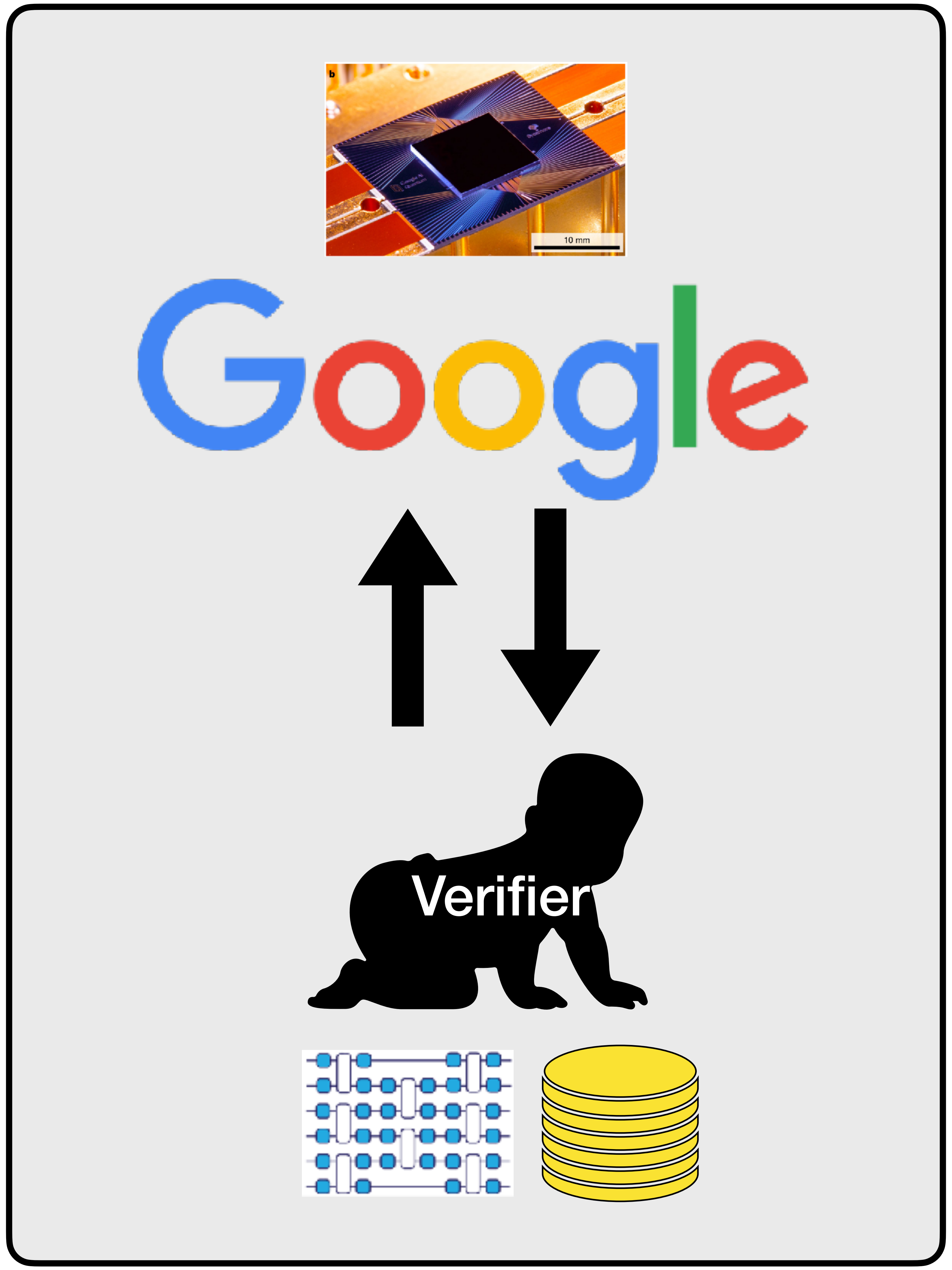
nature > news > article

NEWS | 23 October 2019

Hello quantum world! Google publishes landmark quantum supremacy claim

The company says that its quantum computer is the first to perform a calculation that would be practically impossible for a classical machine.





Example: Google's Quantum Supremacy Claim


nature

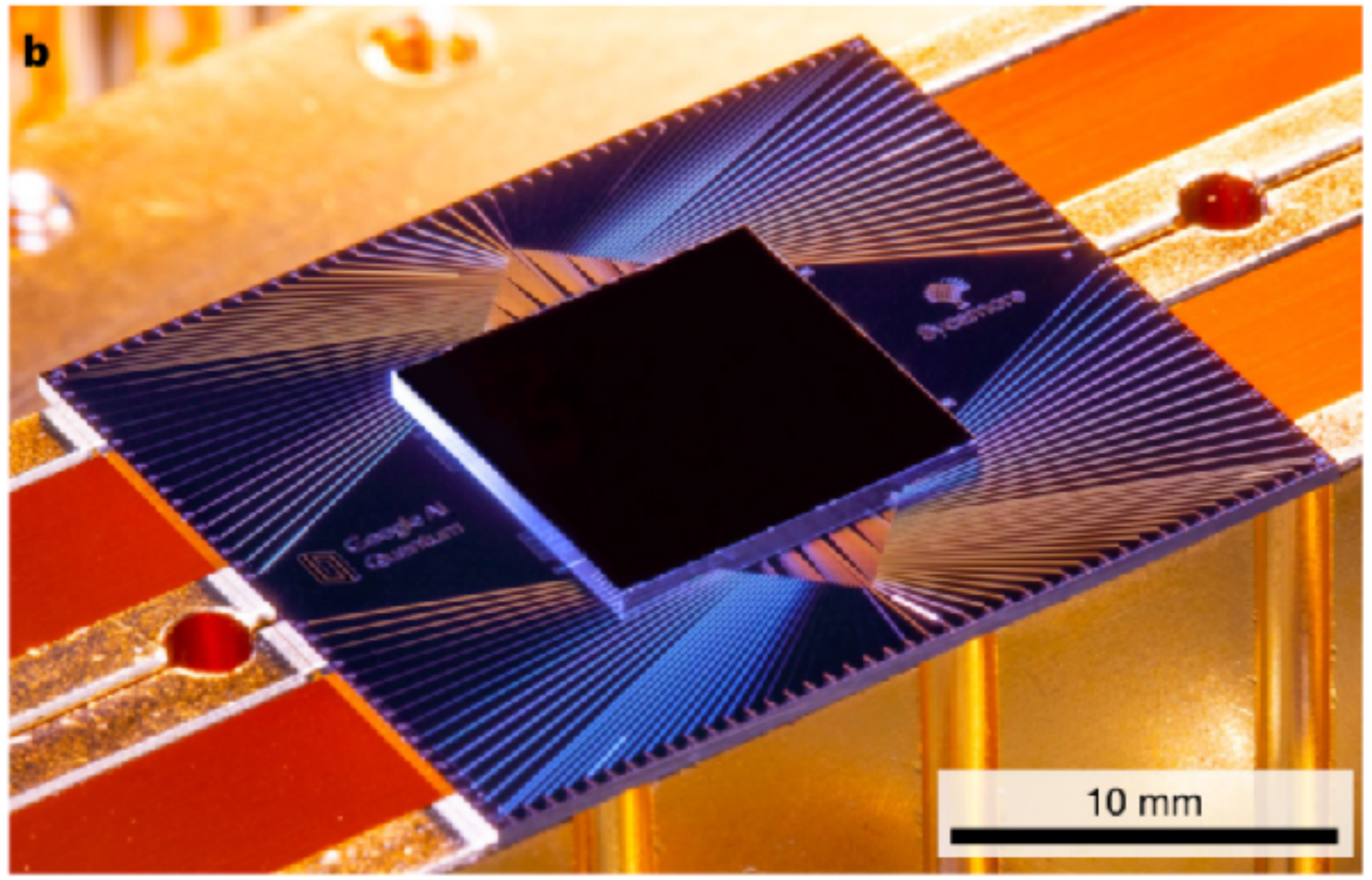
Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

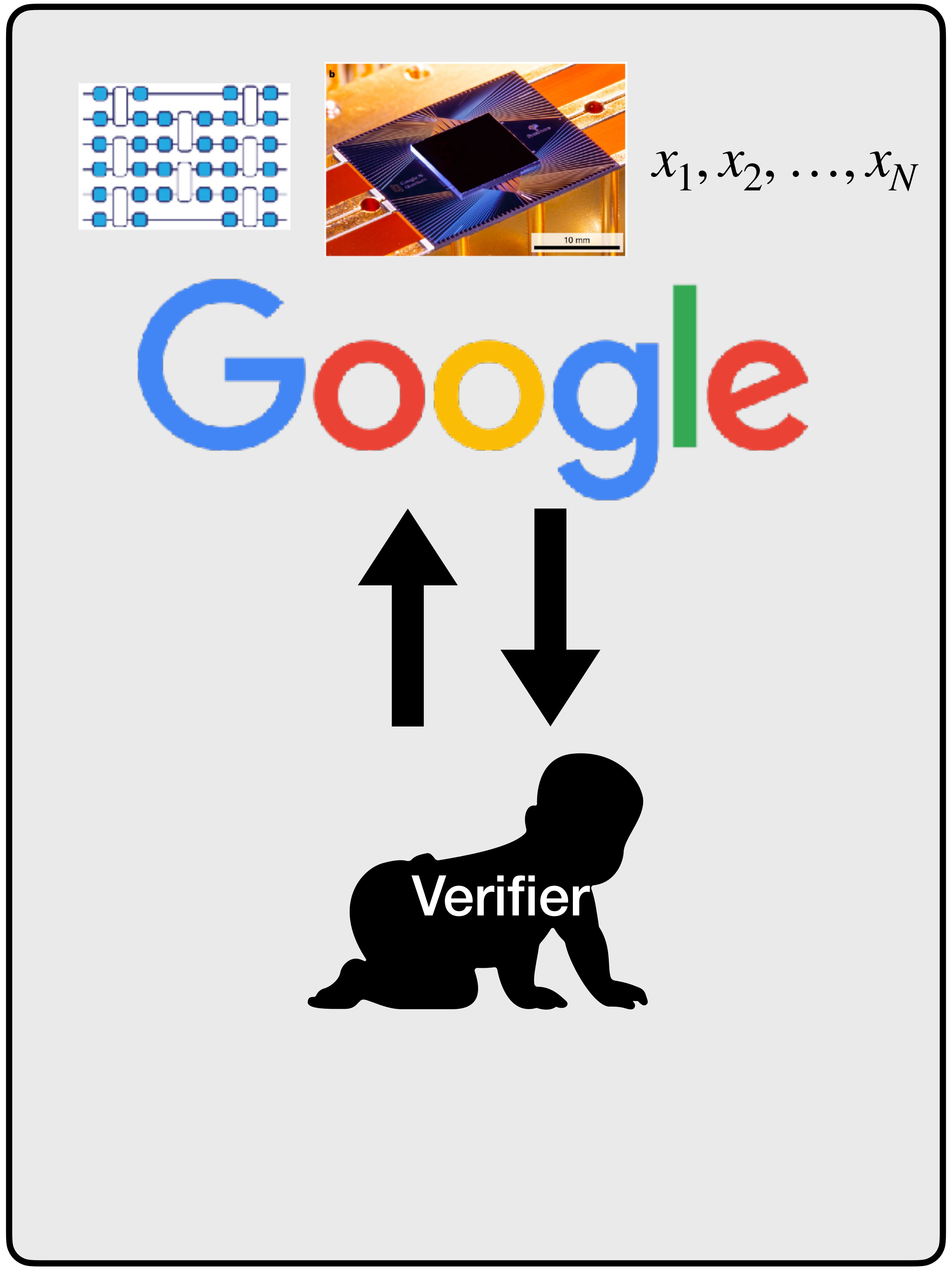
nature > news > article

NEWS | 23 October 2019

Hello quantum world! Google publishes landmark quantum supremacy claim

The company says that its quantum computer is the first to perform a calculation that would be practically impossible for a classical machine.





Example: Google's Quantum Supremacy Claim

nature

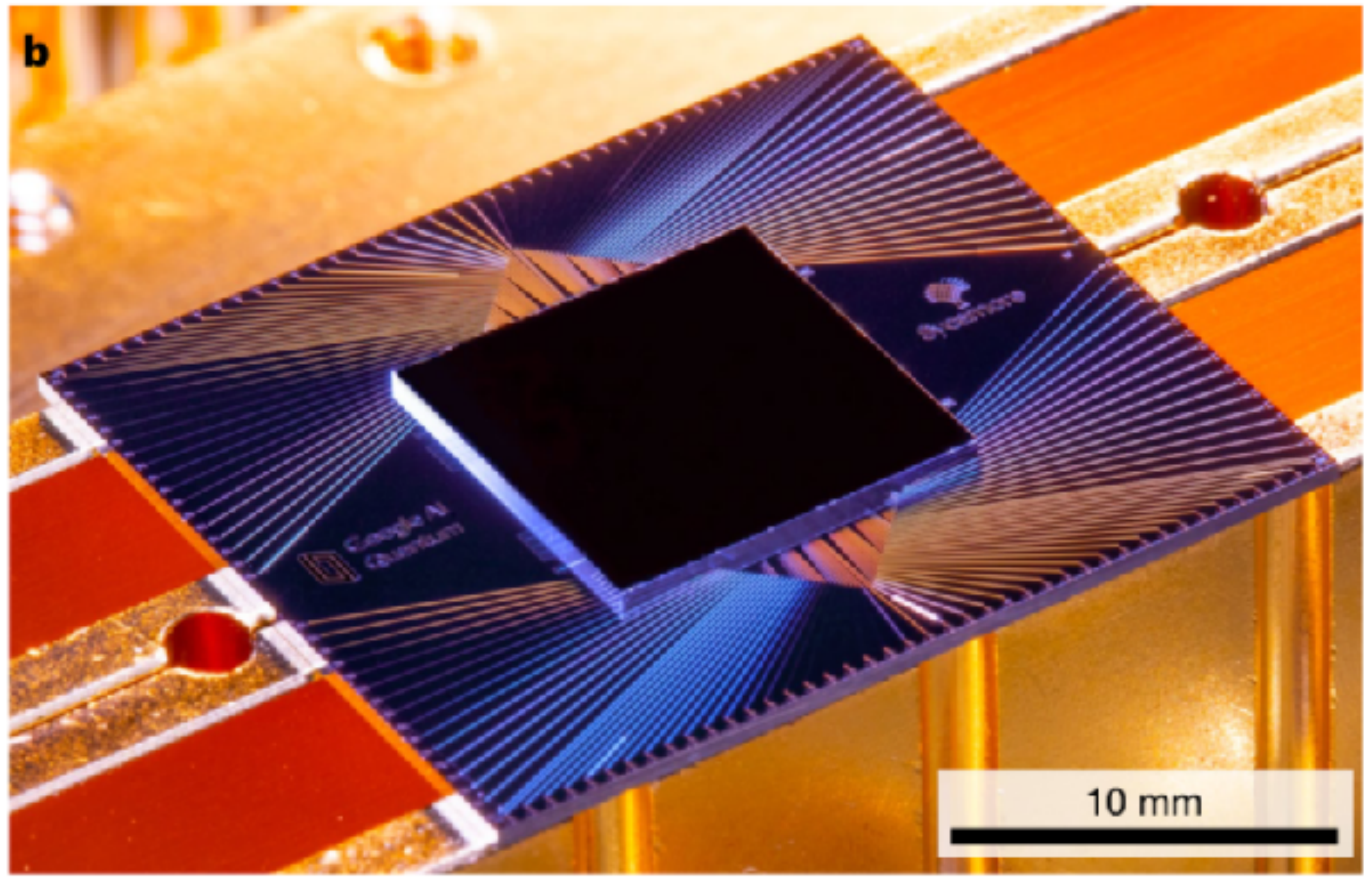
Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature > news > article

NEWS | 23 October 2019

Hello quantum world! Google publishes landmark quantum supremacy claim

The company says that its quantum computer is the first to perform a calculation that would be practically impossible for a classical machine.



Example: Google’s Quantum Supremacy Claim

nature

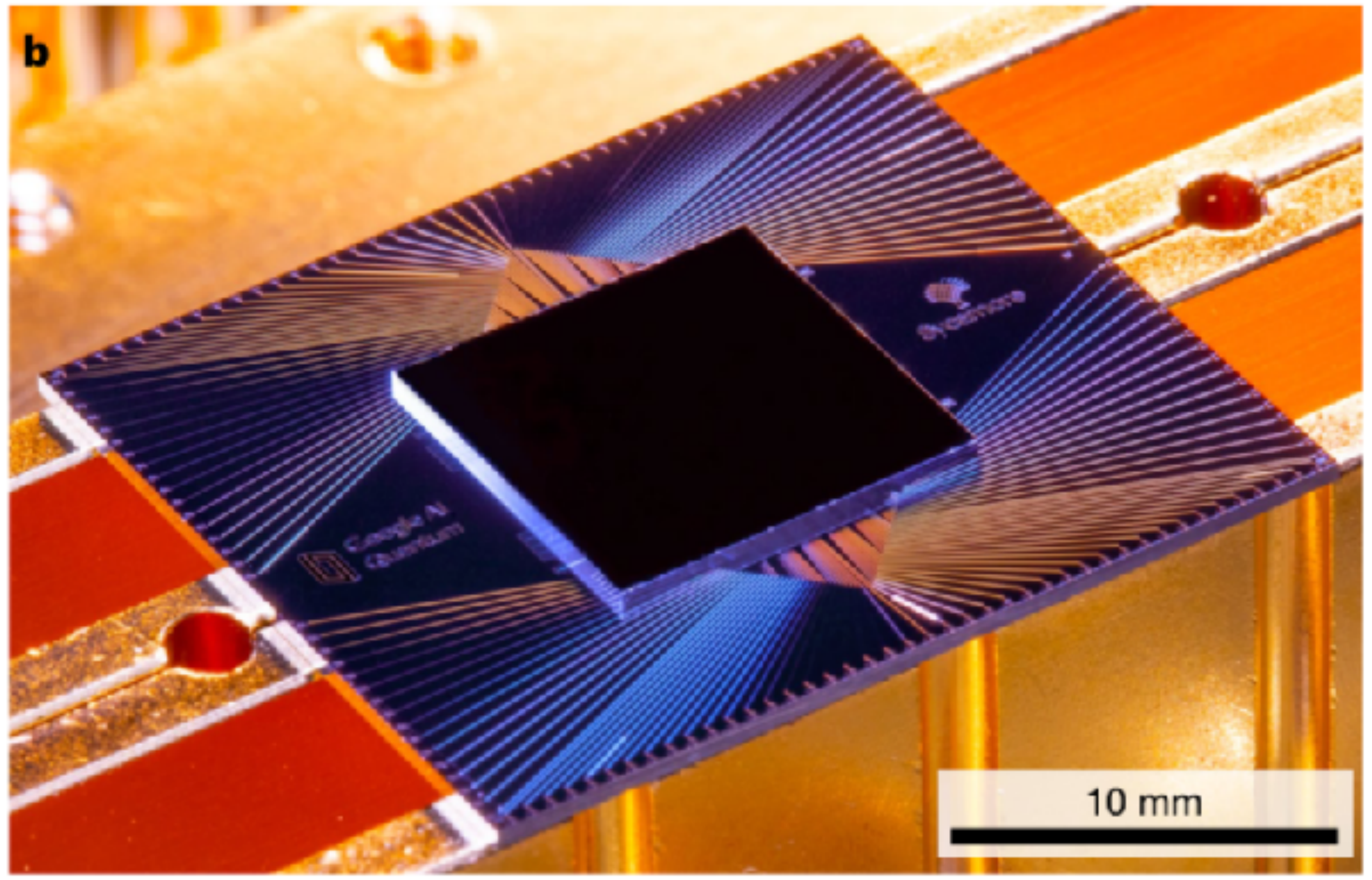
Explore content ⌄ About the journal ⌄ Publish with us ⌄ | Subscribe

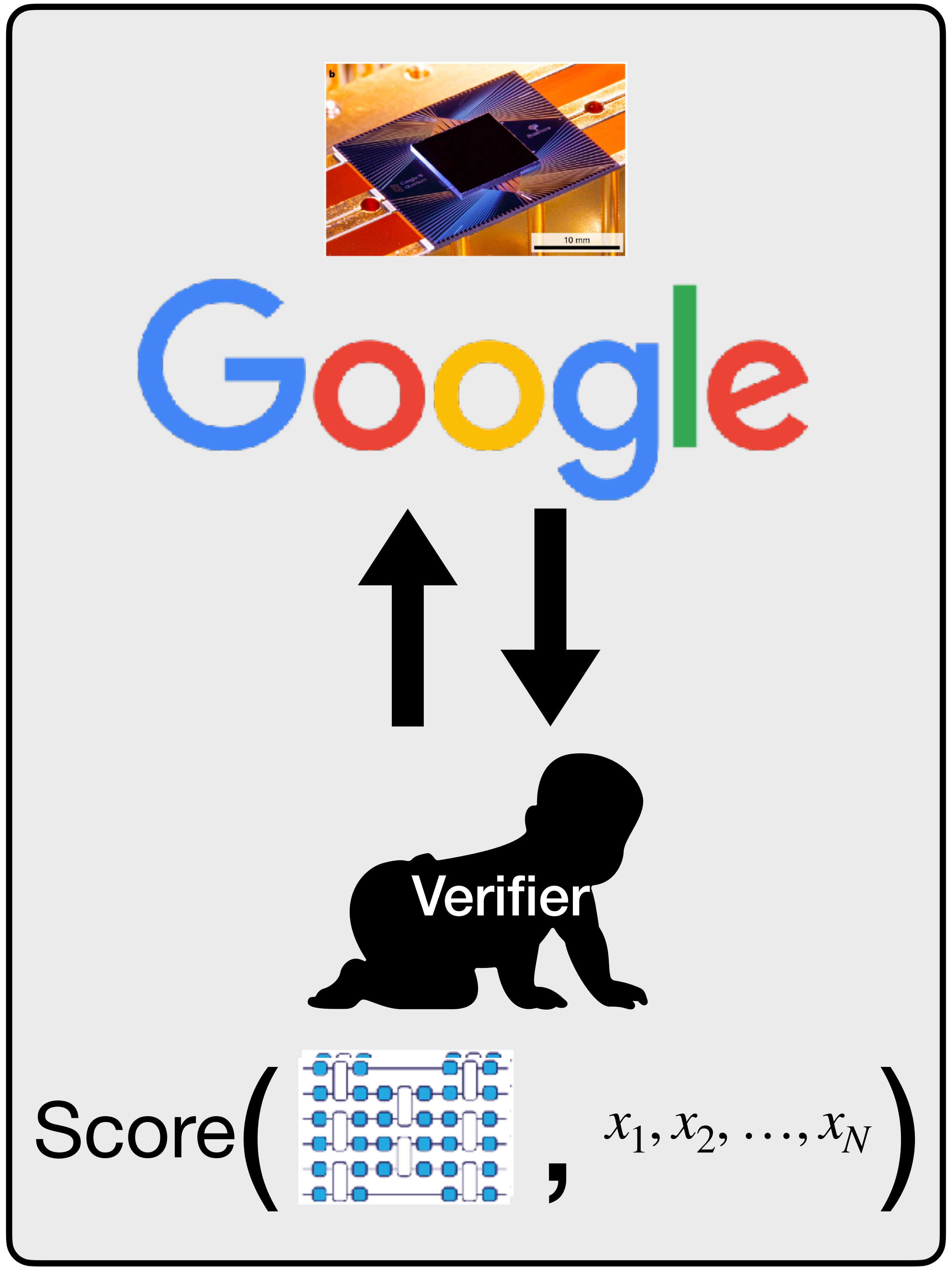
nature > news > article

NEWS | 23 October 2019

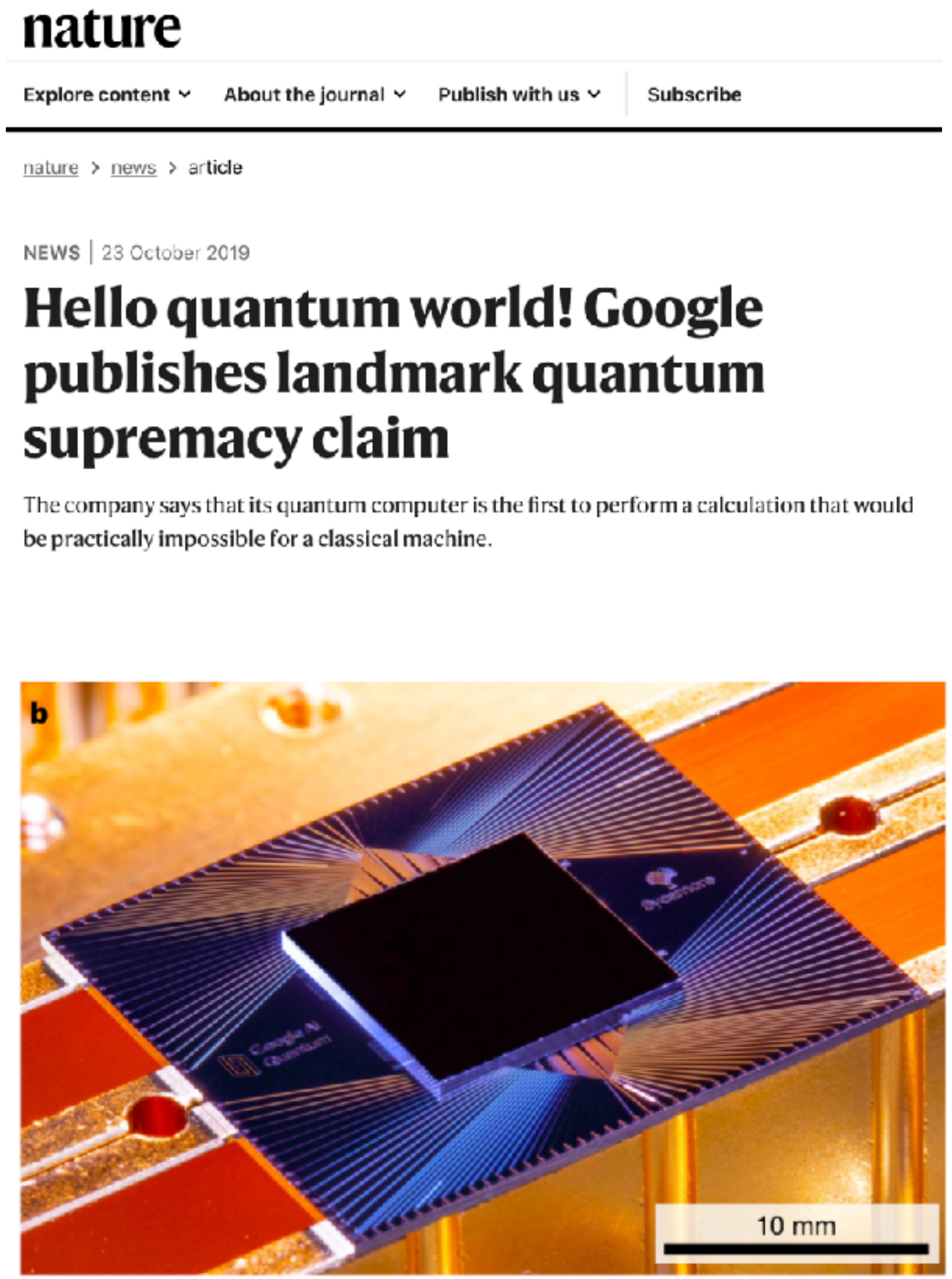
Hello quantum world! Google publishes landmark quantum supremacy claim

The company says that its quantum computer is the first to perform a calculation that would be practically impossible for a classical machine.





Example: Google's Quantum Supremacy Claim


nature

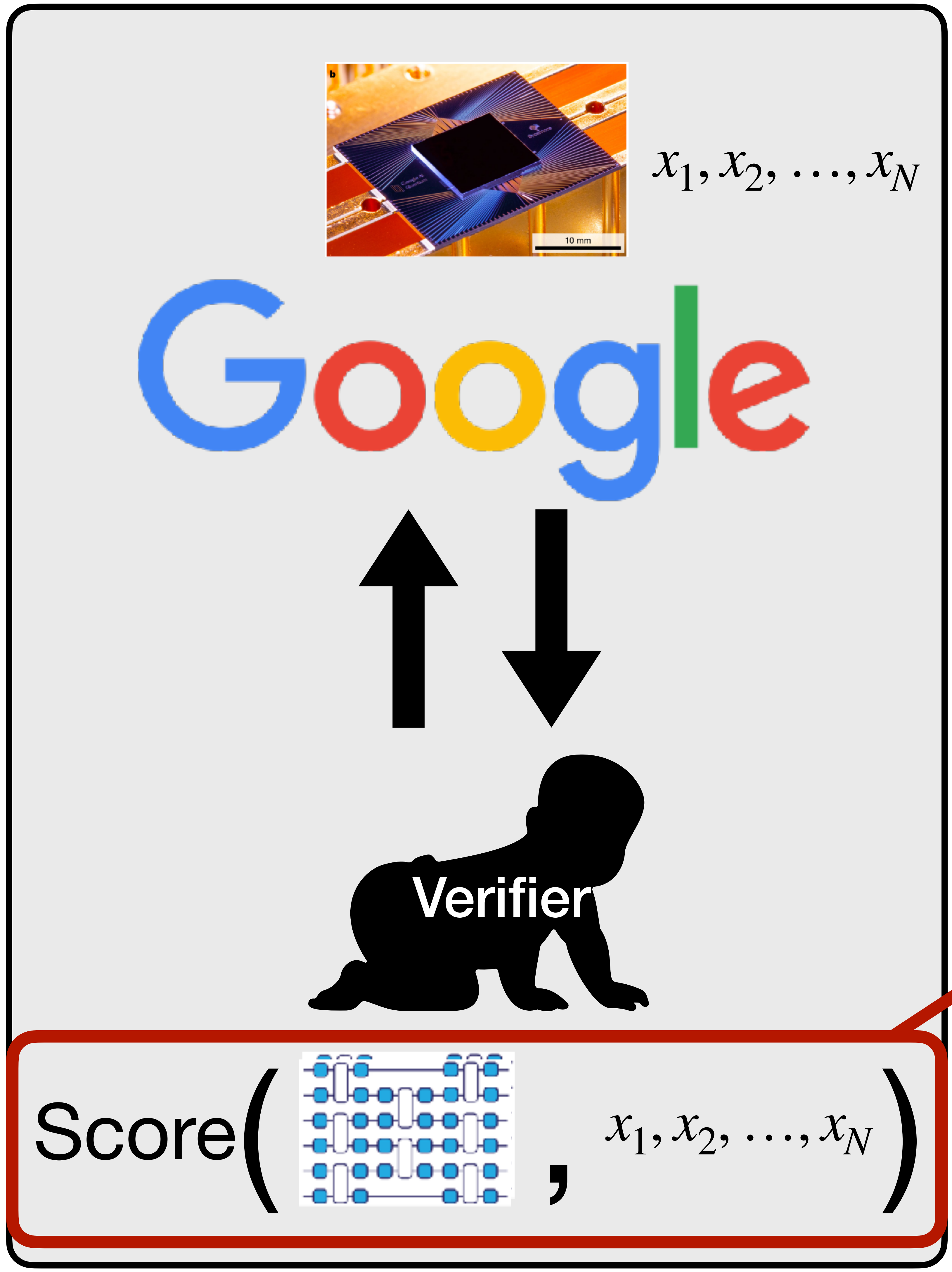
Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

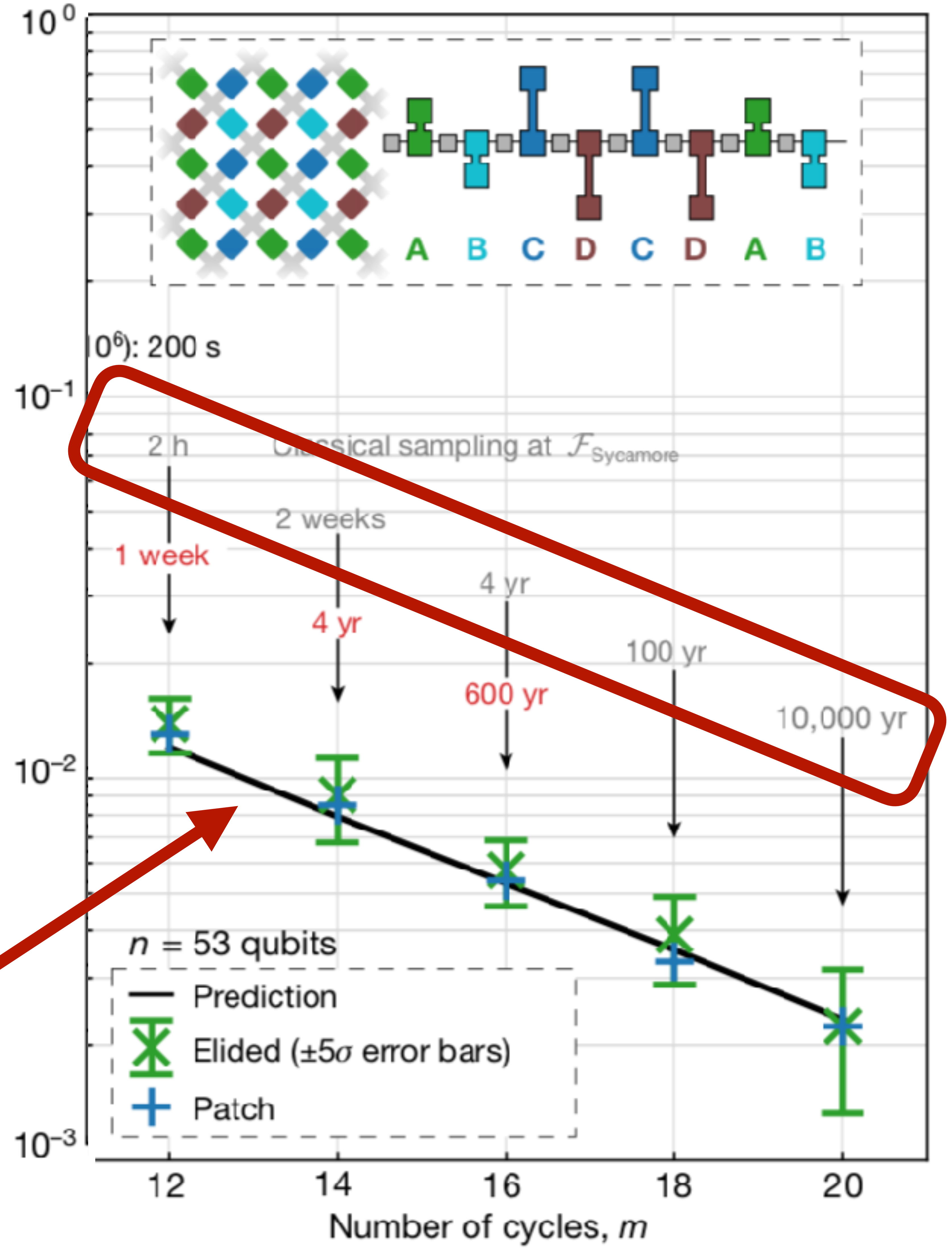
nature > news > article

NEWS | 23 October 2019

Hello quantum world! Google publishes landmark quantum supremacy claim

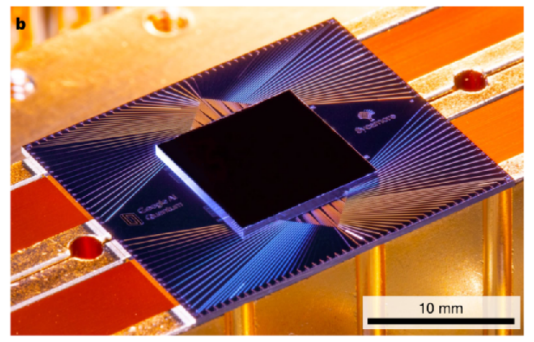
The company says that its quantum computer is the first to perform a calculation that would be practically impossible for a classical machine.





Have We Demonstrated Quantum Computational Advantage!?

Almost There! But Maybe We Have to be More Careful?

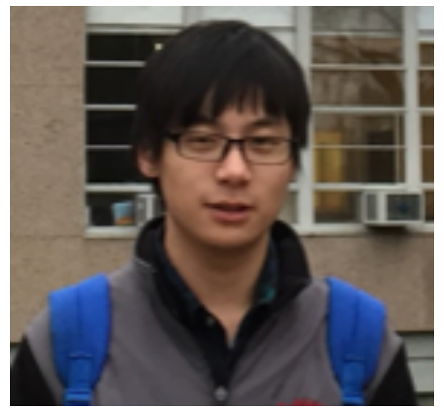


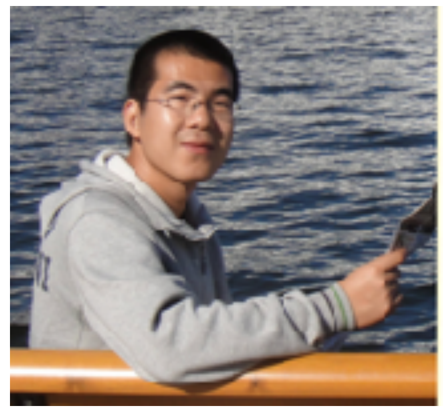
It requires 10,000 years for the best supercomputer!

20 days on the supercomputer Summit is enough!

[Huang et al., arXiv 2005.06787]

[Pan-Chen-Zhang, arXiv 2111.03011]





15 hours using 512 GPUs is enough!

We reexamine the foundation of the “score” in Google’s experiment!

Roadmap

Summary & Future Directions

The Complexity-Theoretic Aspect

Our Analytical Model Linear XEB and Fide

-Eye View on r Results

Basic Setup:

- Random circuits Sampling (R
- Linear Cross-Entropy (XEB)

Basic Setup

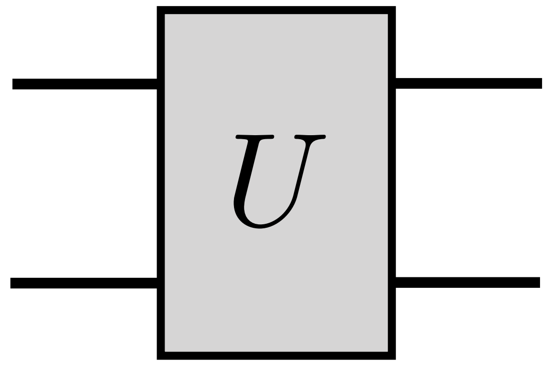
Background: Quantum Circuits

- Quantum states $|\psi\rangle$.

$$|\psi\rangle = \begin{bmatrix} \alpha_{0\dots00} \\ \alpha_{0\dots01} \\ \vdots \\ \alpha_{1\dots11} \end{bmatrix}$$

n-qubit state — length 2^n unit complex vector

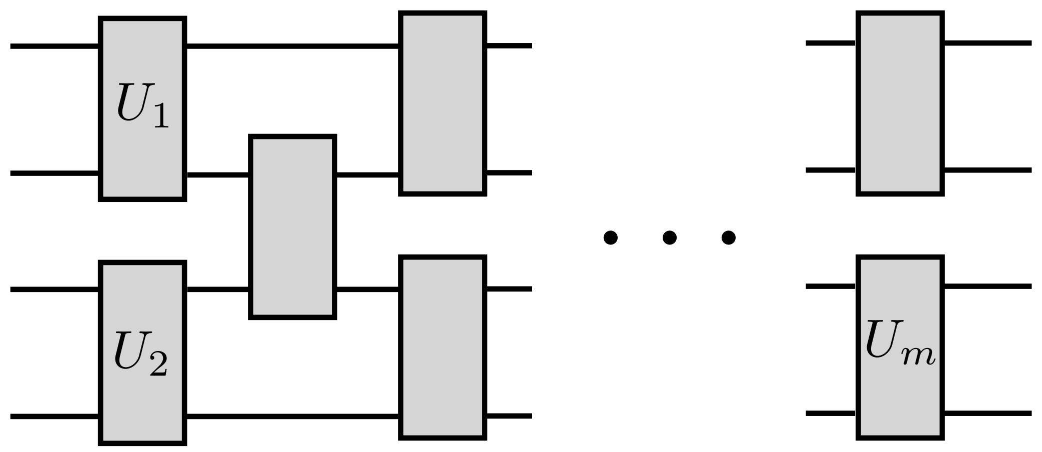
- Quantum gates U.



$$= U \otimes I_{2^{n-2}}$$

2-qubit gate — 2^n by 2^n unitary matrix

- Quantum circuits C.



$$= C$$

circuit with 1- and 2-qubit gates — 2^n by 2^n unitary matrix

- Output distribution q_C.

$$C|\psi\rangle = \begin{bmatrix} \alpha_{0\dots00} \\ \alpha_{0\dots01} \\ \vdots \\ \alpha_{1\dots11} \end{bmatrix}$$

length 2^n unit complex vector

$$q_C(x) = |\alpha_x|^2$$

distribution over n-bit strings

Background: Quantum Circuits

Linear algebra over complex numbers

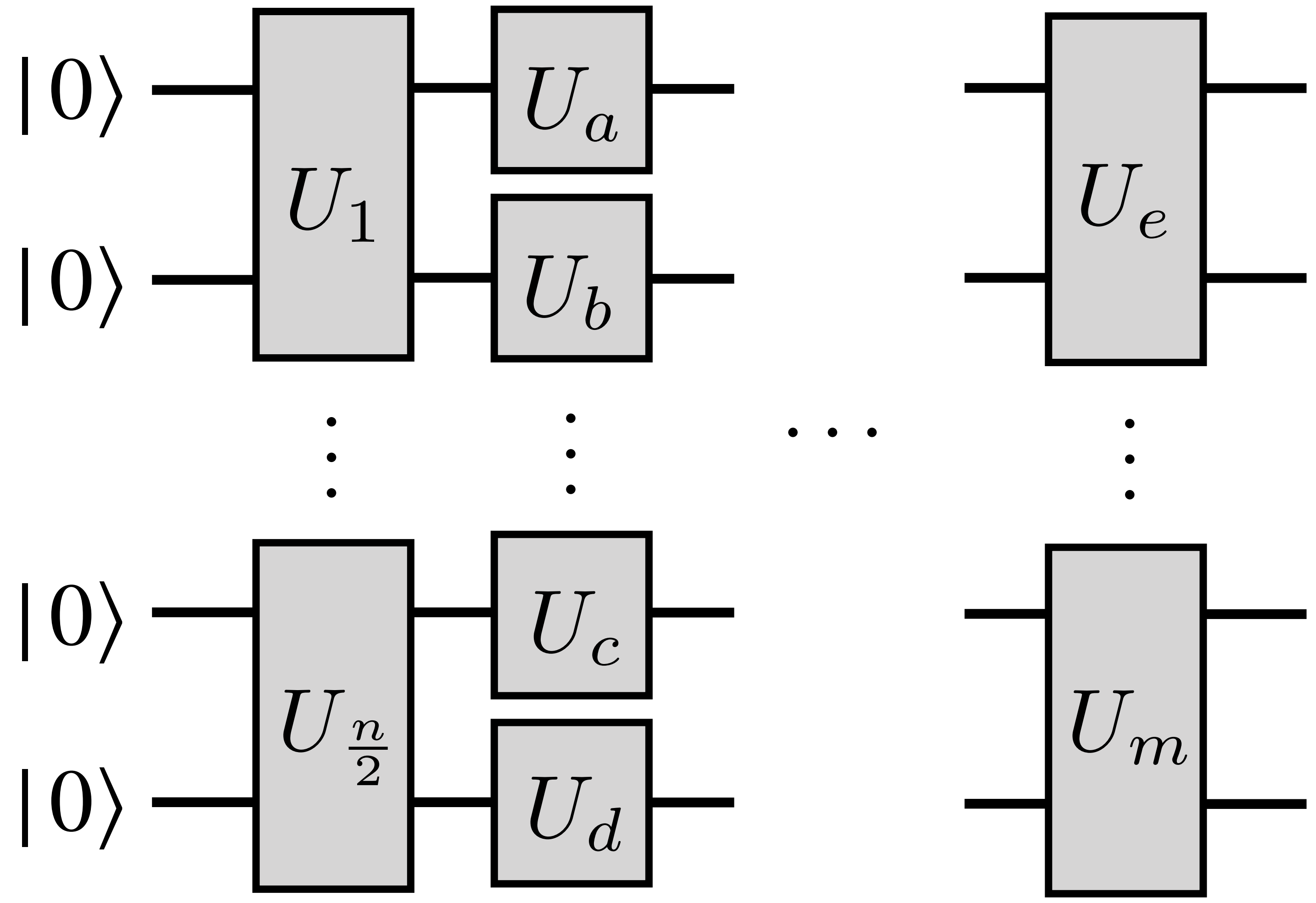
&

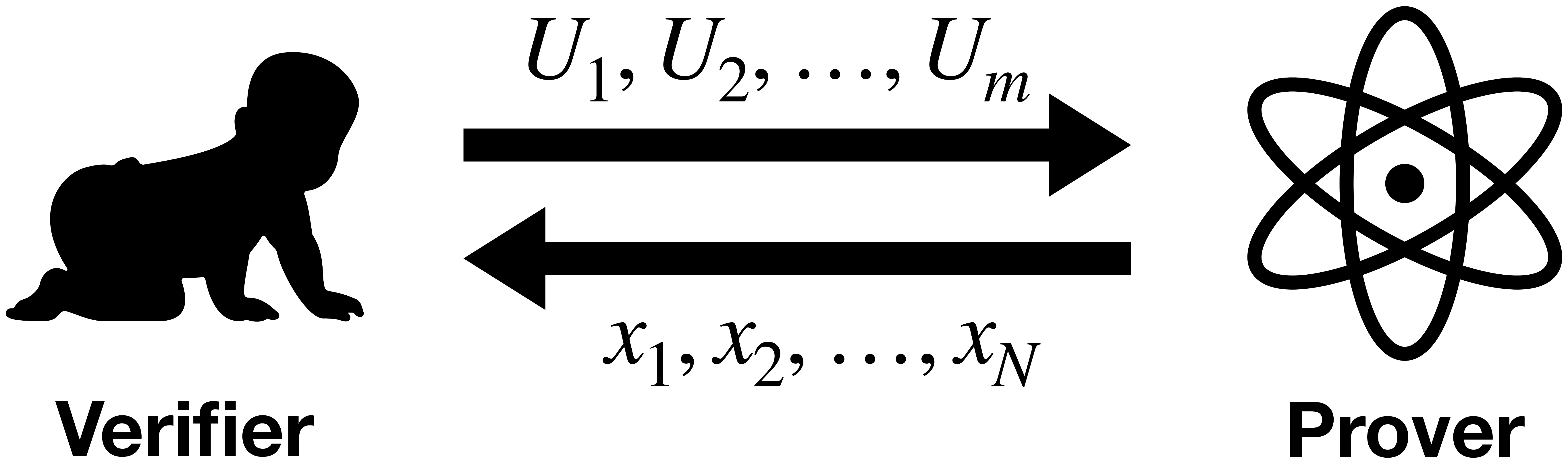
A quantum circuit computes a distribution over n-bit strings

A direct classical simulation of a quantum circuit takes exponential time*!

* I'll elaborate on this soon!

Random Circuits Sampling (RCS) Based Quantum Advantage





Step 0: Both parties agree on a circuit architecture.

Step 1: Verifier random samples gates $U_1, U_2, \ldots, U_m$ and sends to Prover.

Step 2: Prover simulates the quantum circuits, produces strings $x_1, x_2, \ldots, x_m$, and send back.

Step 3: Verifier tries to figure out if the strings were sampled correctly.

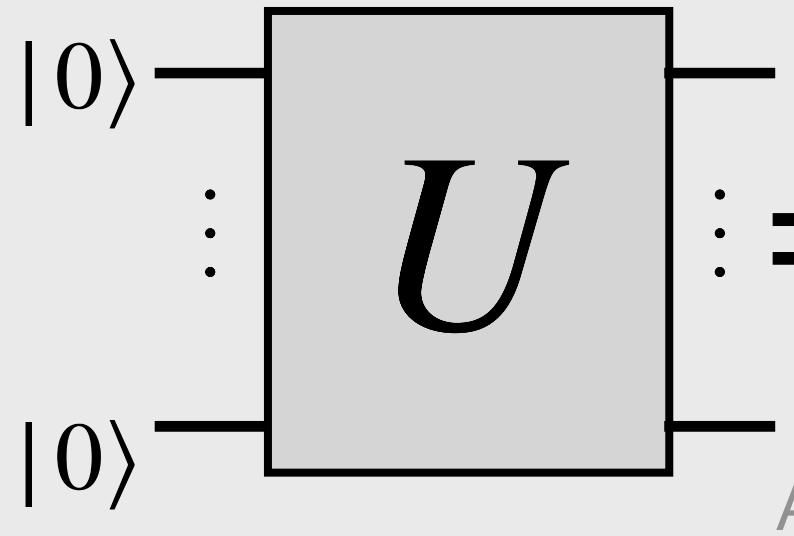
Intuition: A classical prover cannot extract useful information in poly-time!

* Here I describe the “adversarial setting” while some physicists focus on the “benign setting”.

Fidelity & Linear Cross-Entropy (XEB) Benchmark

Q: How to classically verify if the sampled distribution is close to the right one?

Ideal circuit



$= |\psi_U\rangle \xrightarrow{\text{Corresponds to}} p_U$

A (pure) quantum state — A distribution over n-bit strings

A sampled distribution

$\rho \xrightarrow{\text{Corresponds to}} q$

A (mixed) quantum state — A distribution over n-bit strings

Fidelity

$$\langle \psi_U | \rho | \psi_U \rangle$$

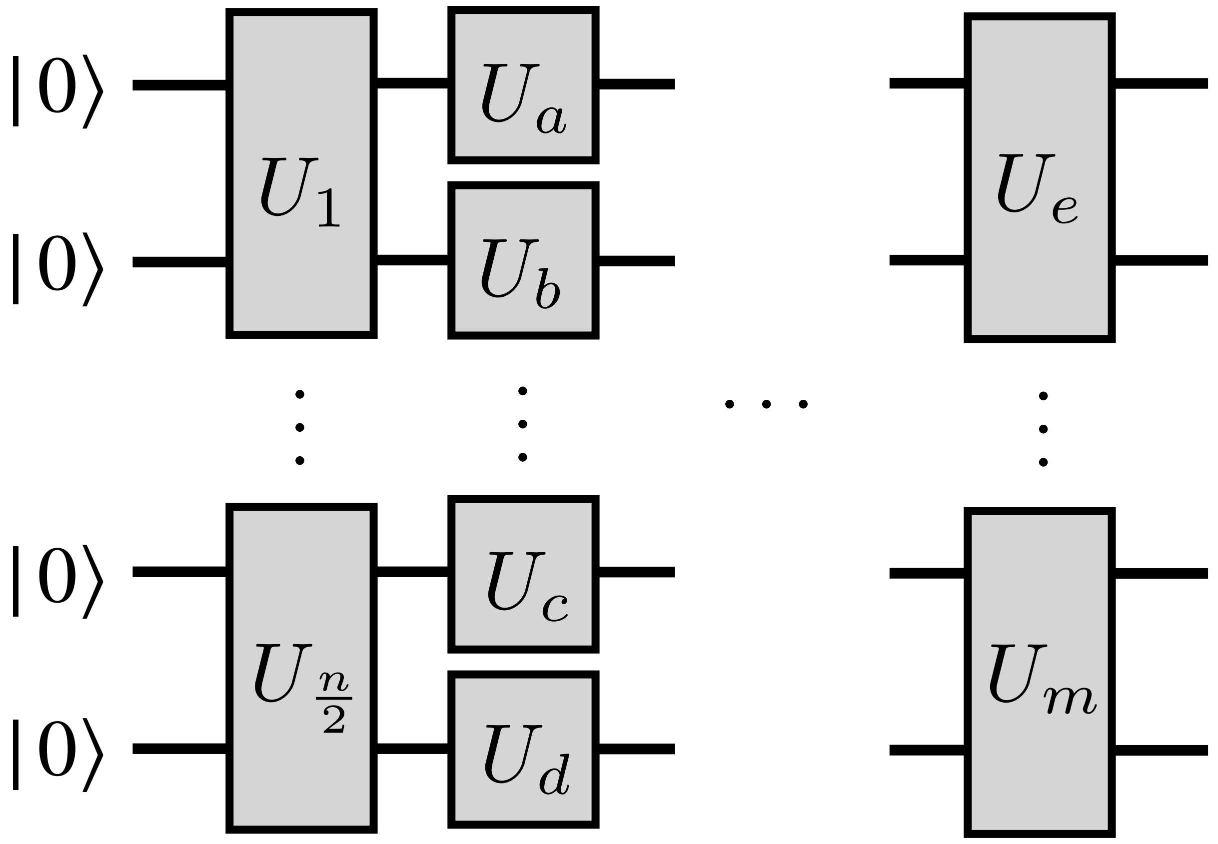
A common measure for the “closeness” between quantum states. However, **hard to estimate** in general.

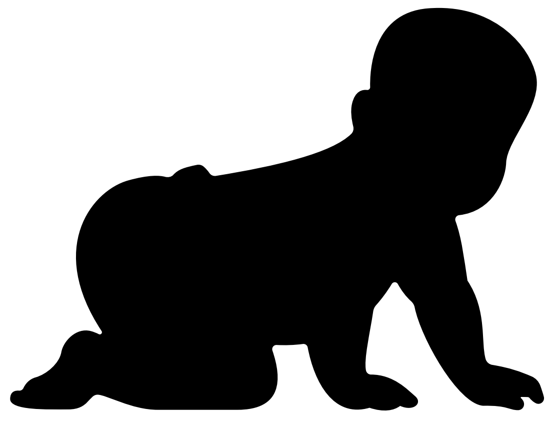
Linear Cross-Entropy (XEB)

$$\mathbb{E}_{x \sim q} \left[2^n p_U(x) - 1 \right]$$

An **empirical “proxy” for fidelity** and is used in RCS-based quantum advantage.

Ideal Circuits, Noisy Circuits, and Classical Simulations

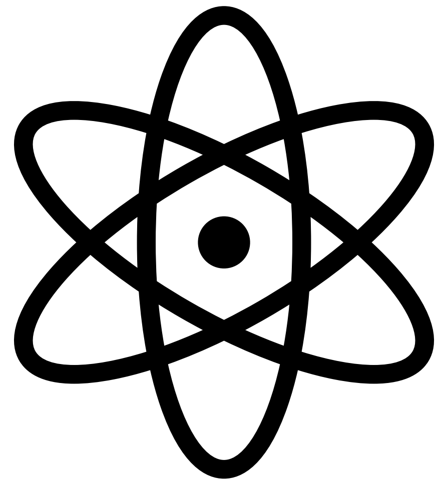




Verifier

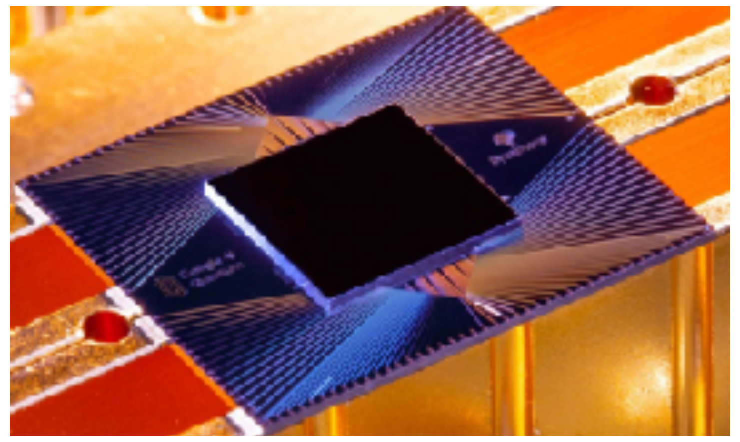
$U_1, U_2, \ldots, U_m$

$x_1, x_2, \ldots, x_N$



Ideal Prover

$\mathbb{E}_U[\chi_U(p_U)] = 1$



Noisy Prover

$\mathbb{E}_U[\chi_U(\mathcal{N}_\epsilon)] > 0$

ϵ is the "gate fidelity" which describes how much noise per gate.

Classical Prover

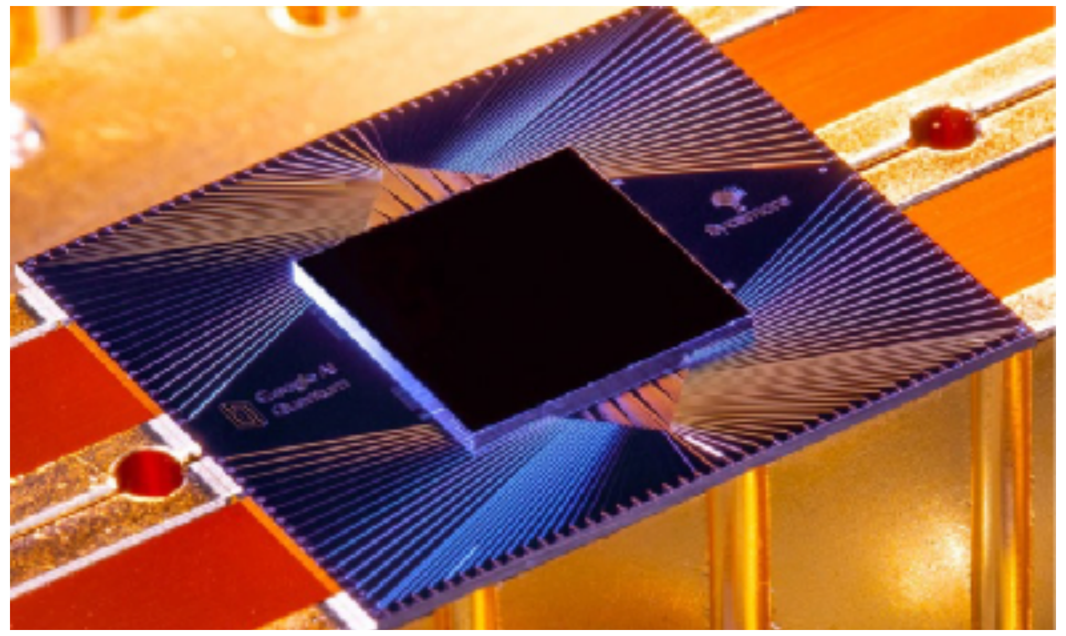
$\mathbb{E}_U[\chi_U(C)] \approx 0$

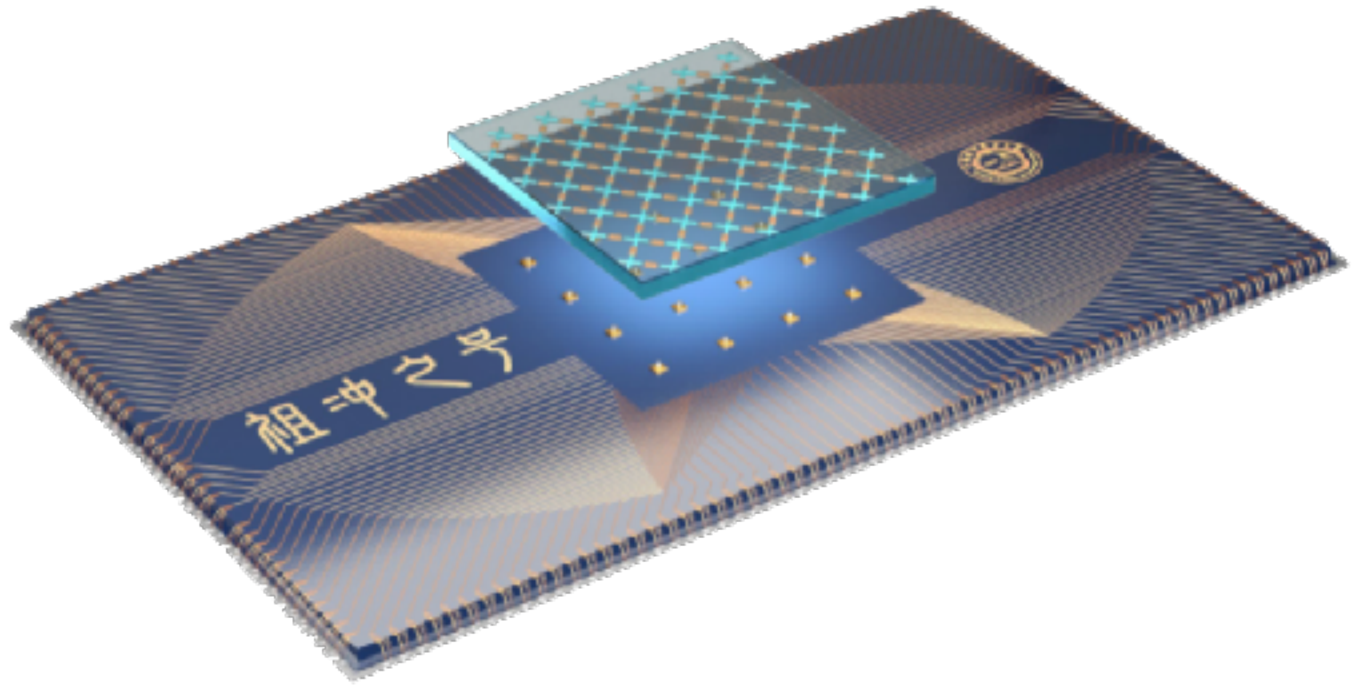
Conjectured

XEB

$$\chi_U(q) = \mathop{\mathbb{E}}_{x \sim q}[2^n p_U(x) - 1]$$

RCS-Based Quantum Advantage Using the XEB Benchmark





Sycamore (**53 qubits, 20 depth**)
by Google, Oct. 2019

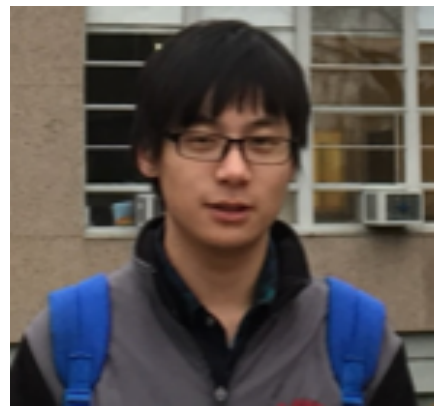
$\chi_U(\mathcal{N}_\epsilon) \approx 2.24 \times 10^{-3}$

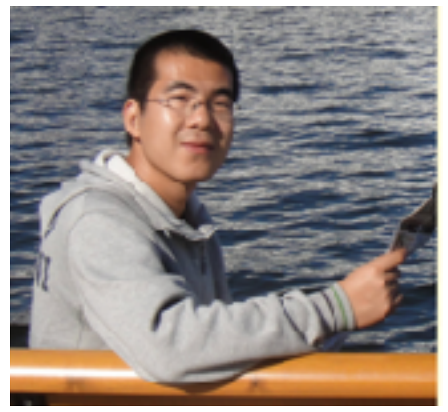
Zuchongzhi (**56 qubits, 20 depth**)
by USTC, Jun. 2021

$\chi_U(\mathcal{N}_\epsilon) \approx 6.62 \times 10^{-4}$

Zuchongzhi-2 (**60 qubits, 24 depth**)
by USTC, Sep. 2021

$\chi_U(\mathcal{N}_\epsilon) \approx 3.66 \times 10^{-4}$





512 GPUs & 15 hours for Sycamore circuits
(**53 qubits, 20 depth**)
Nov. 2021

$\chi_U(C) \approx 3.7 \times 10^{-3}$

The current finite size regime has been challenged!

However, these classical algorithms do not scale up.

Q: How does the XEB of a noisy simulation scale with #qubits, the noise per gate ϵ, and the depth d?

Q: What's the XEB a scalable classical algorithm can achieve?

Interlude

Why People Think It's Hard to Spoof XEB?

Two Direct Classical Simulations for a Quantum Circuit

Schrödinger's Algorithm

Store the whole quantum state and update it gate by gate.

- Time: $O(m2^n)$
- Space: $O(2^n)$.

$$|\Psi\rangle = U_m \cdots U_2 U_1 |0^n\rangle$$

Define $|\Psi_i\rangle = U_i \cdots U_2 U_1 |0^n\rangle$ and compute them one by one.

Feynman's Algorithm

Express the final quantum state as a sum of all possible path from the input layer.

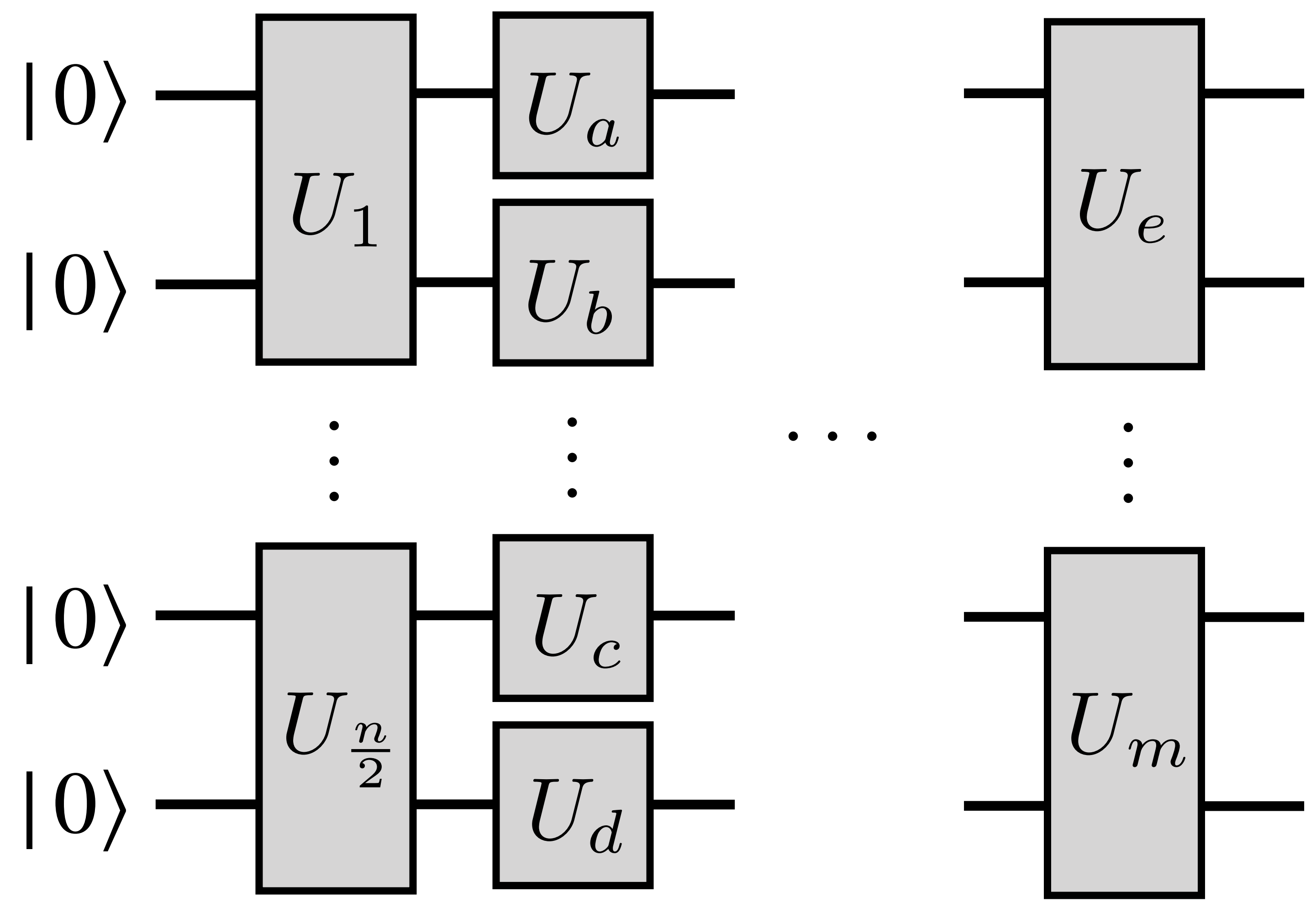
- Time: $O(4^m)$
- Space: $O(m+n)$.

$$|\Psi\rangle = \sum_{x_1,\ldots,x_{m-1} \in \{0,1\}^4} U_m |x_{m-1}\rangle \cdots |x_2\rangle \underbrace{\langle x_2 | U_2 | x_1\rangle}_{\text{A scalar}} \underbrace{\langle x_1 | U_1 | 0^n\rangle}_{\text{A scalar}}$$

Enumerate all $x_1, \ldots, x_{m-1} \in \{0,1\}^4$ and compute the summation.

People believe this is the best one can do for reasonably complicated circuits!

Why People Think Spoofing XEB Could be Classically Hard?



XEB

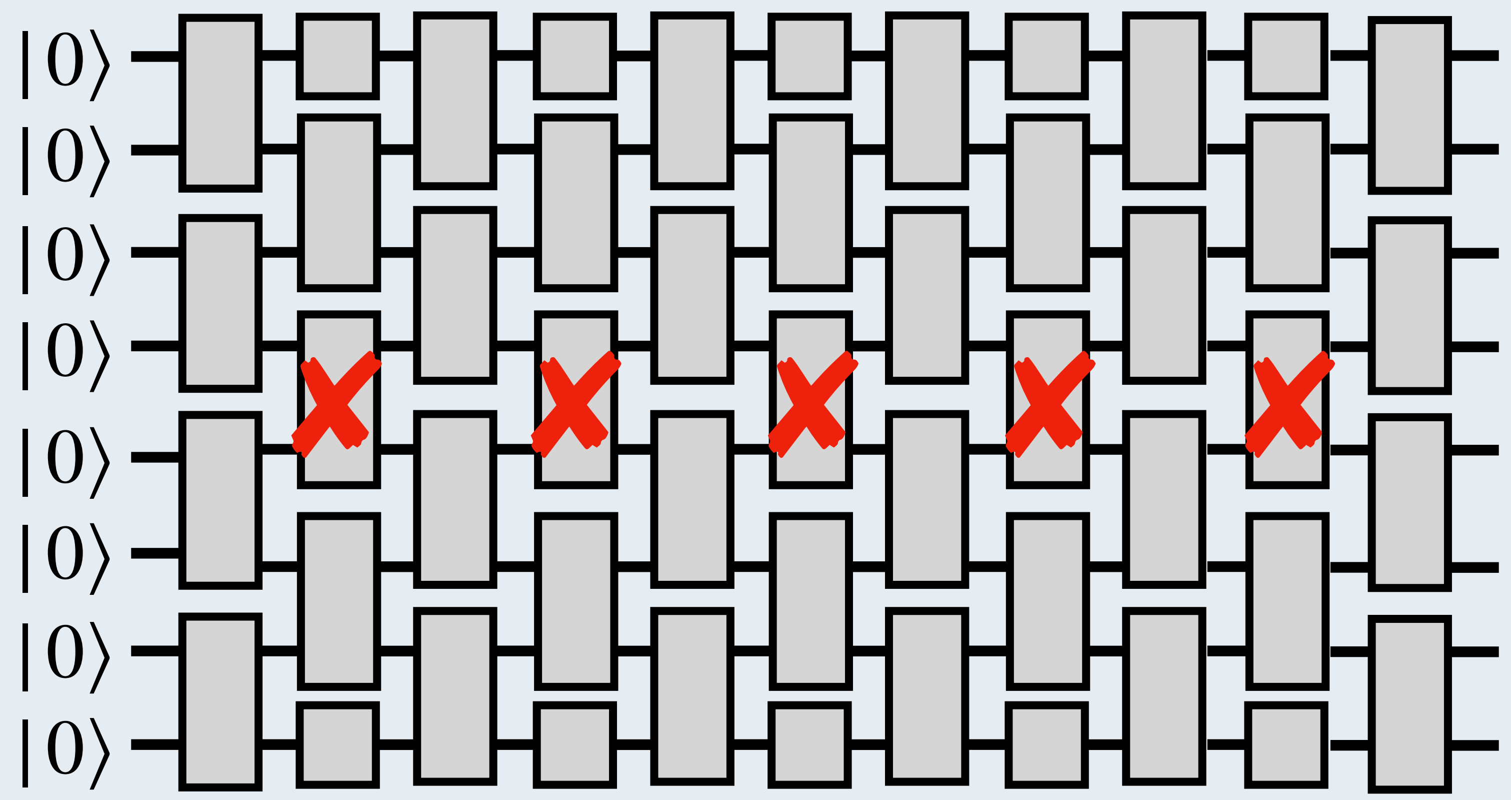
$$\chi_U(q) = \mathop{\mathbb{E}}_{x\sim q}[2^n p_U(x) - 1]$$

- Essentially one needs to be able to sample from a distribution q that is close to the ideal distribution p_U.
- Such q can be used to "estimate" $p_U(0^n)$ [Aaronson-Chen 2017] [Aaronson-Gunn 2020].
- **Conjecture**: the best one can do classically is to "estimate" $p_U(0^n)$ by running either Schrödinger's algorithm or Feynman's algorithm [Aaronson-Chen 2017] [Aaronson-Gunn 2020].
- **Spoiler**: For realistic circuit architectures, we show that the conjecture is false.

A Bird-Eye View on Our Results

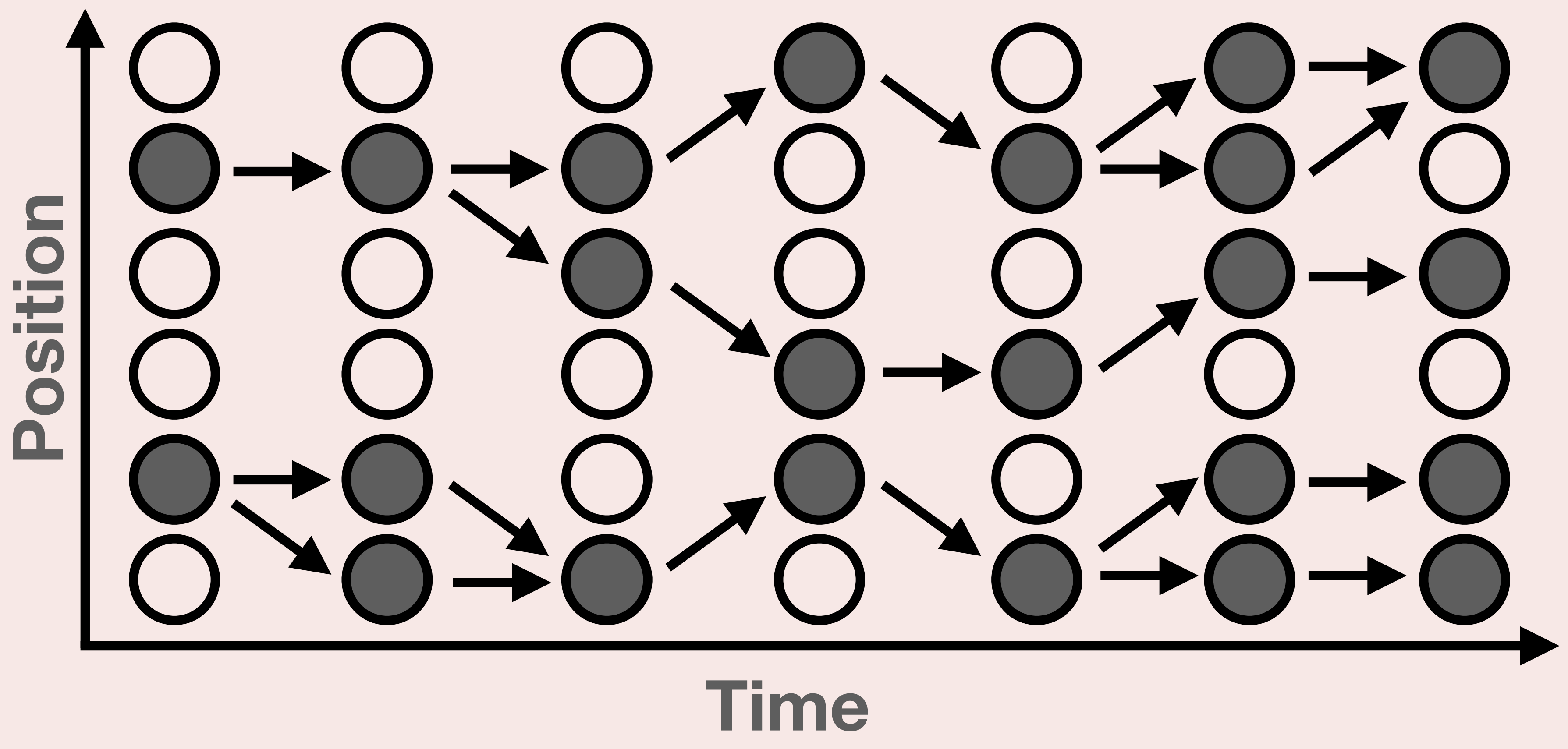
Limitations of Linear XEB as a Measure for Quantum Advantage

Classical Algorithms Spoofing XEB



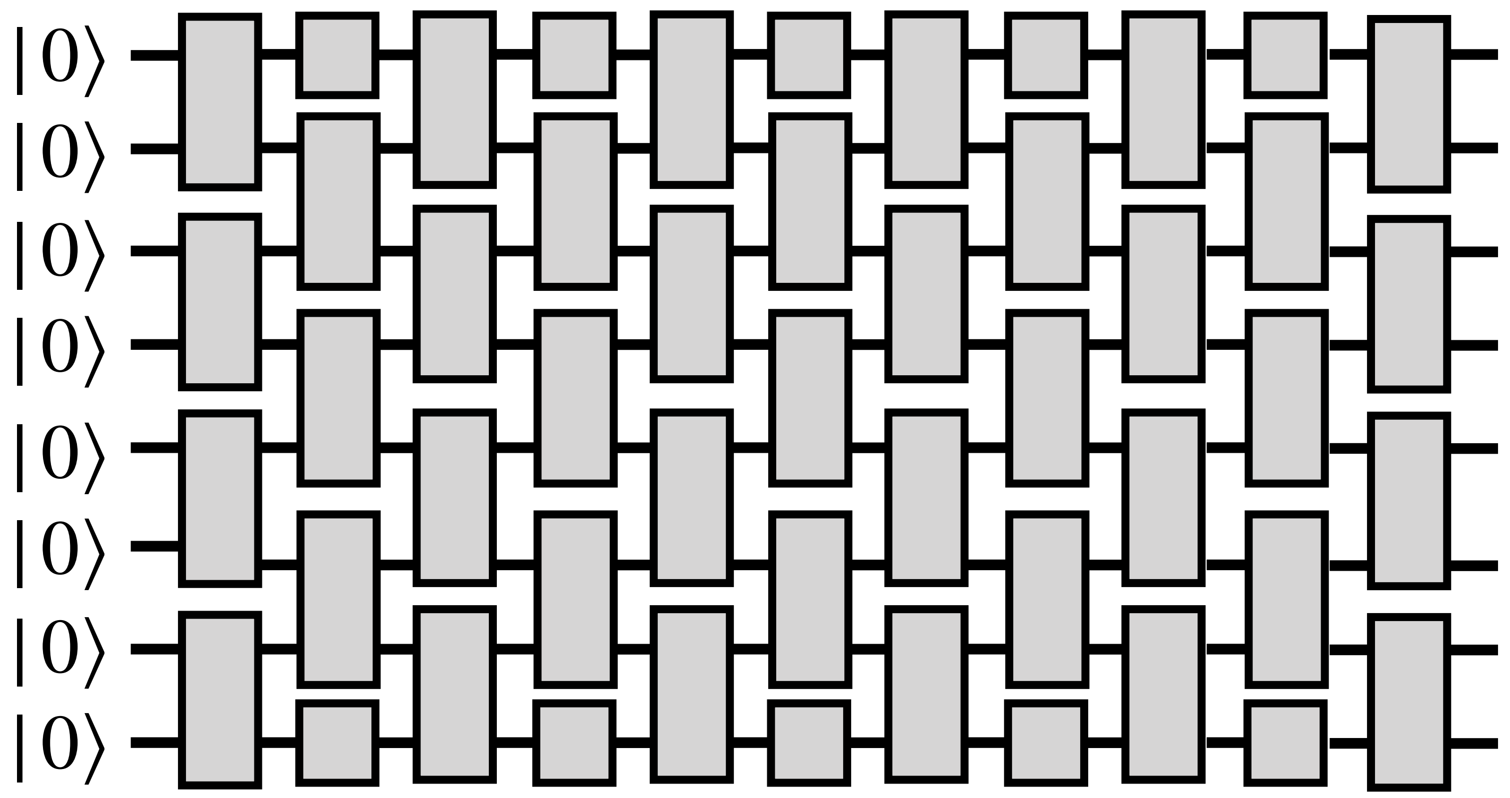
1. **Complexity-theoretically**, refute the XQUATH in realistic circuit architectures.
2. **Experimentally**, achieve 2%~12% of Google's and USTC's XEB in ~1s on 1 GPU.
3. **Asymptotically**, our algorithms are scalable.

A Better Understanding on XEB and Fidelity

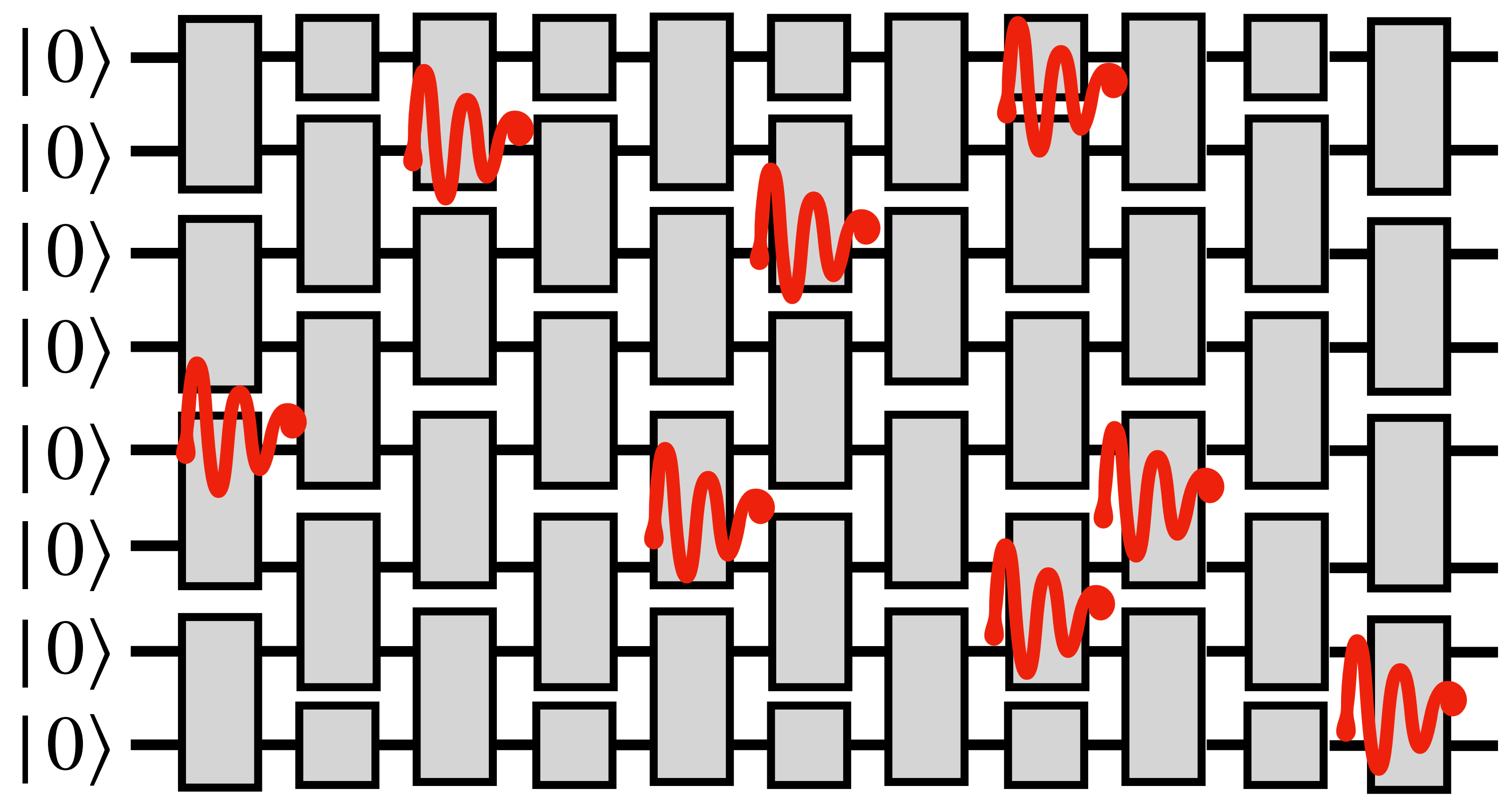


1. **XEB can overestimate fidelity** in both our algorithms and noisy simulation.
2. **XEB has a limited utility as a benchmark** for quantum advantage.

The Template for Our Spoofing Algorithms



Idea circuit

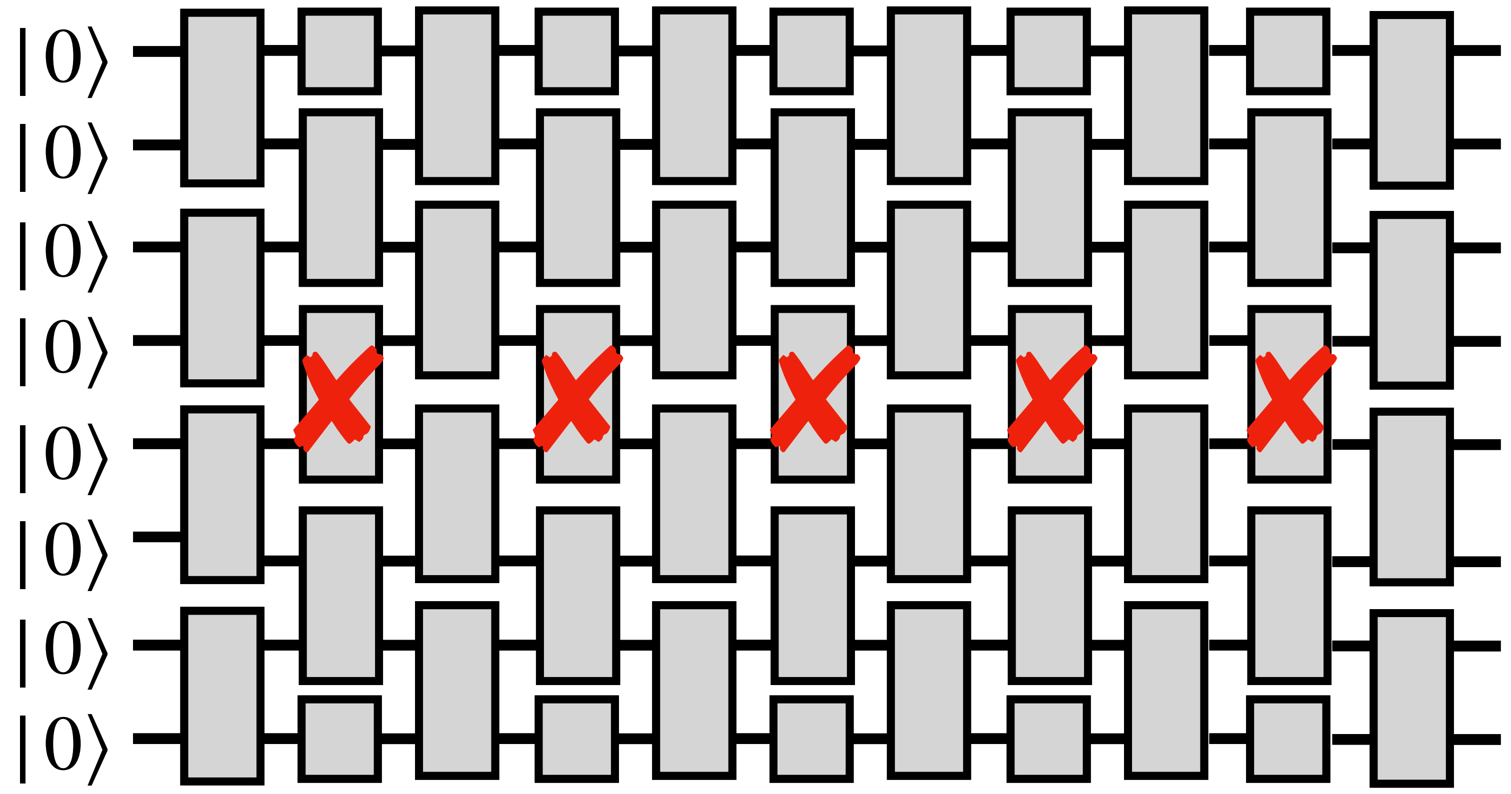


A noisy simulation

Inefficient for a classical algorithm to fully simulate the ideal circuit!

A quantum simulation is necessarily noisy and hence won’t achieve high XEB!

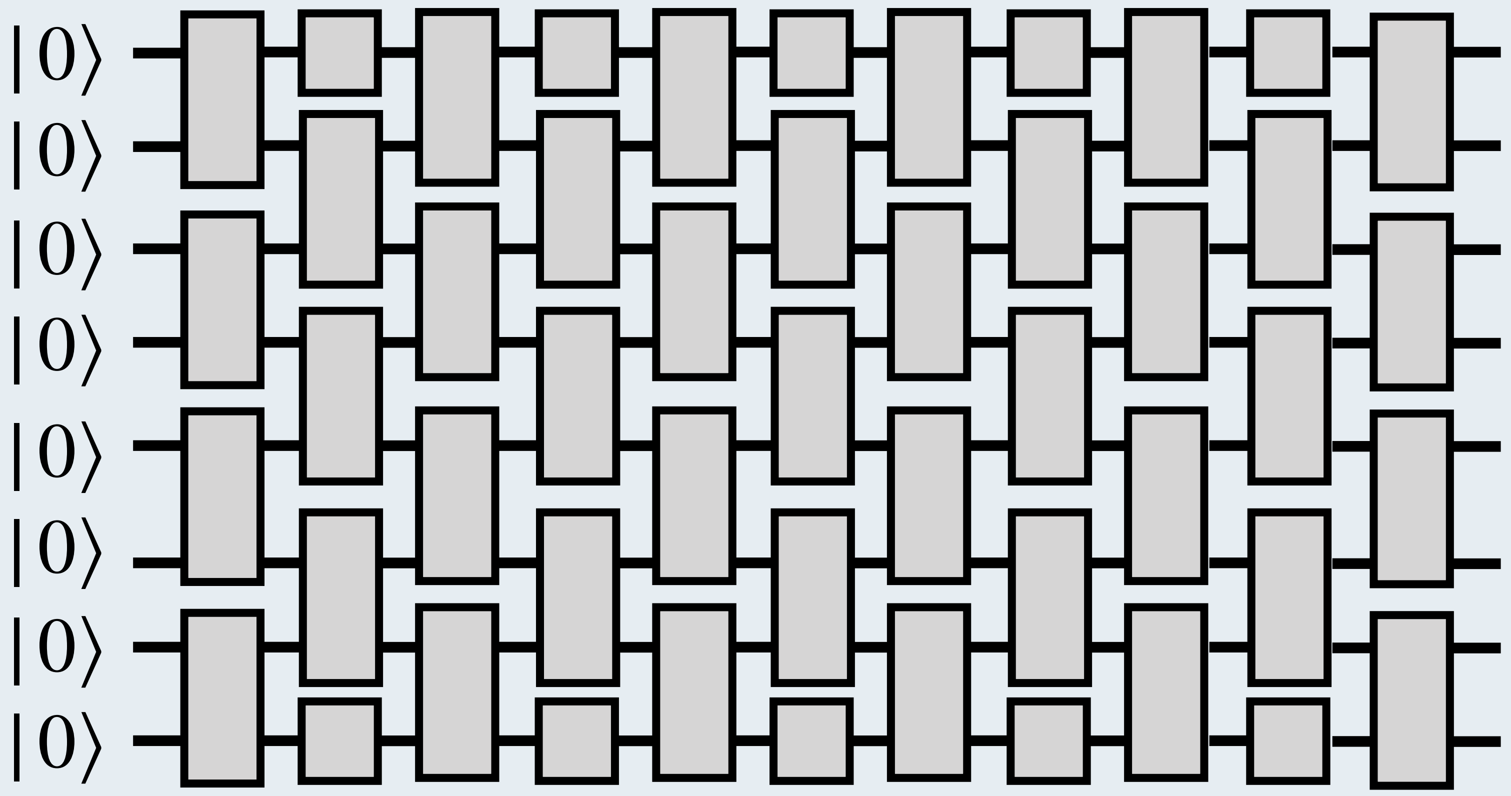
Easy to fully simulate each sub-circuit



Hope: the #removed gate ≈ amount of noise

Theoretical Results

1 Dimensional Circuits

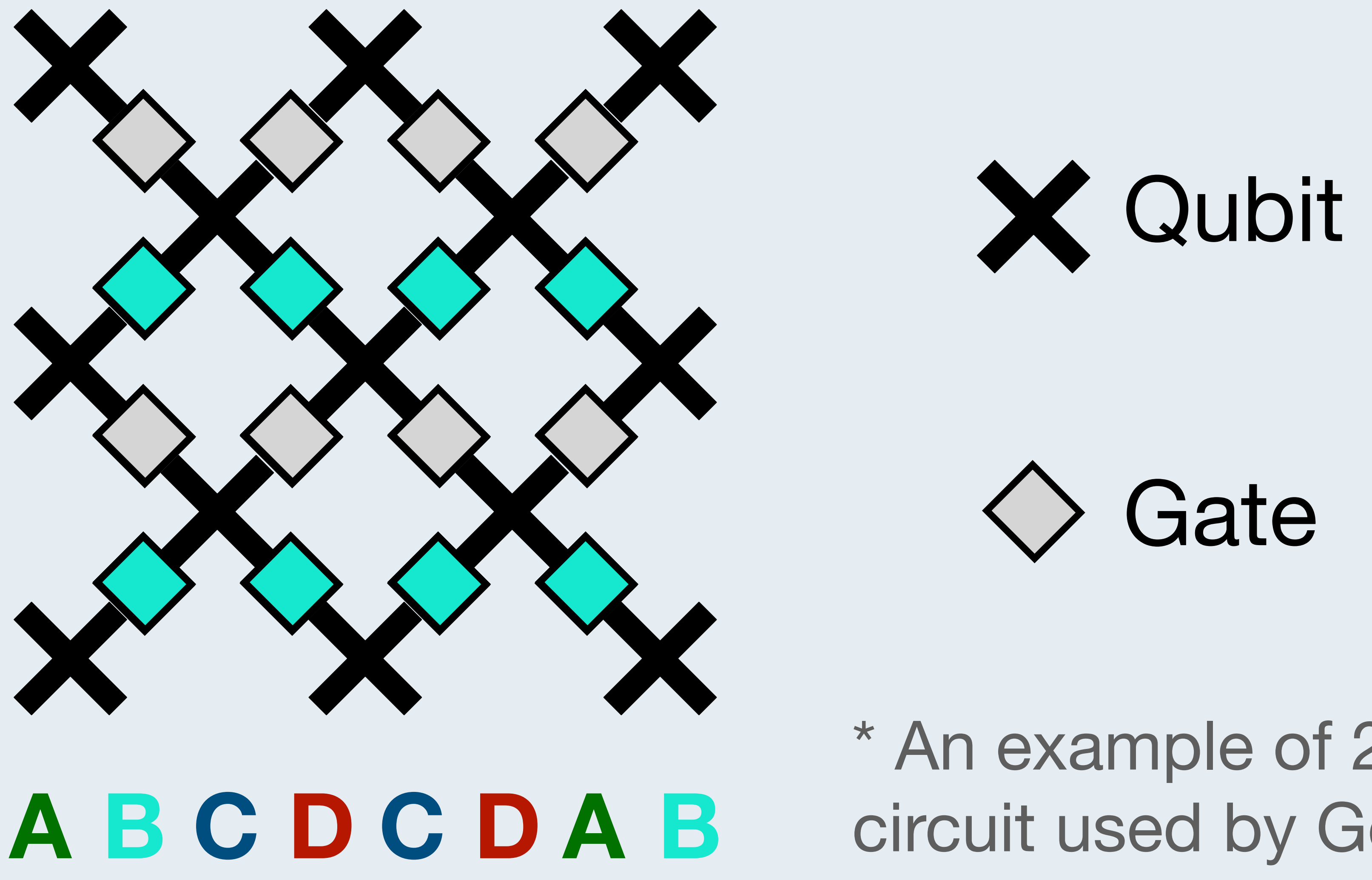


For every constant $\epsilon > 0$ and $N = \Omega(1/\epsilon)$, our algorithm C runs in linear time and

$$\mathbb{E}_U[\chi_U(C)] \geq \mathbb{E}_U[\chi_U(\mathcal{N}_\epsilon)]\,.$$

Q: $\sqrt{\mathrm{Var}_U(\chi_U(C))} \approx \mathbb{E}_U[\chi_U(\mathcal{N}_\epsilon)]$?

Constant Dimensional Circuits



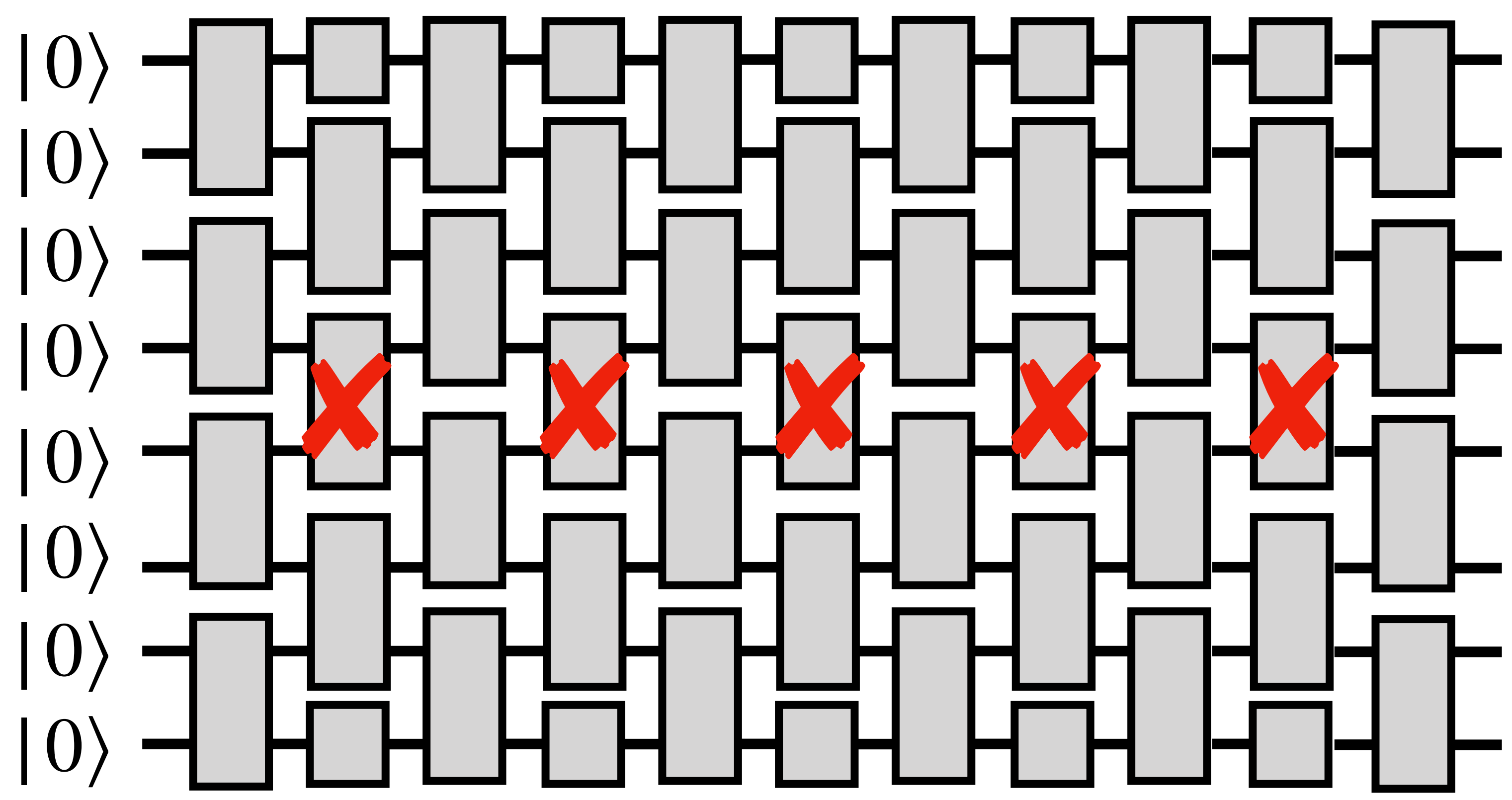
* An example of 2-dim circuit used by Google.

We refute XQUATH [Aaronson-Gunn 20], which is the complexity-theoretic foundation for the classical hardness of XEB-based quantum advantage.

Numerical Results

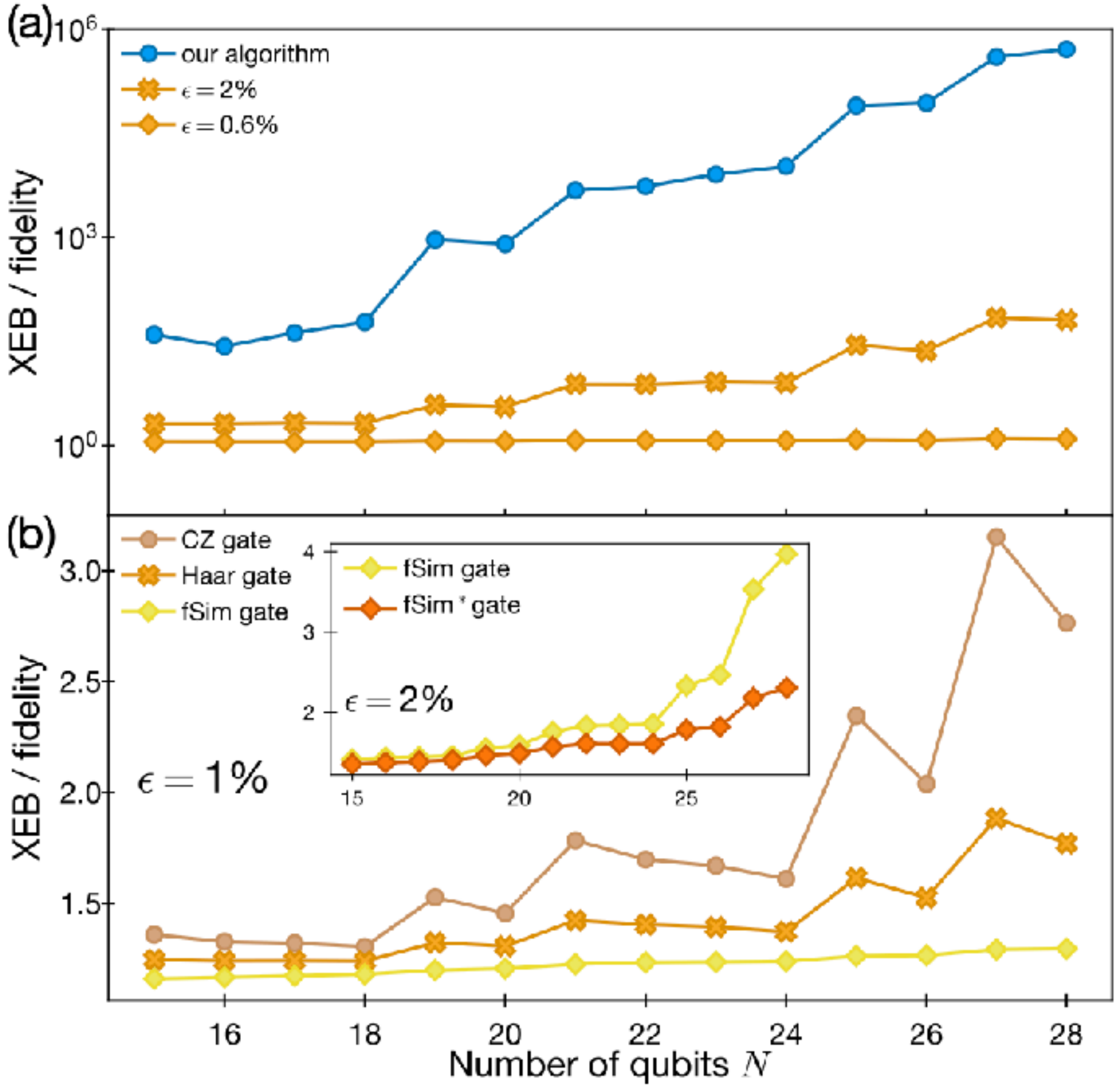
	Google [5]	USTC-1 [6]	USTC-2 [7]
system size	53 qubits, 20 depth	56 qubits, 20 depth	60 qubits, 24 depth
claimed running time on supercomputer [7]	15.9d	8.2yr	4.8×10^4yr
running time on quantum processor	600s	1.2h	4.2h
experimental XEB	2.24×10^{-3}	6.62×10^{-4}	3.66×10^{-4}
running time of our algorithm (1 GPU$^{(a,b)}$)	0.6s	0.6s	1.5s
XEB of our algorithm$^{(b)}$	1.85×10^{-4}	8.18×10^{-5}	7.75×10^{-6}
ratio of ours to experimental XEB	8.26%	12.4%	2.12%

- About 1 second on 1 GPU.
- Achieve 2~12% XEB of Google and USTC.
- Our algorithms haven't been fully optimized!
- The choice of "gate set" matters…

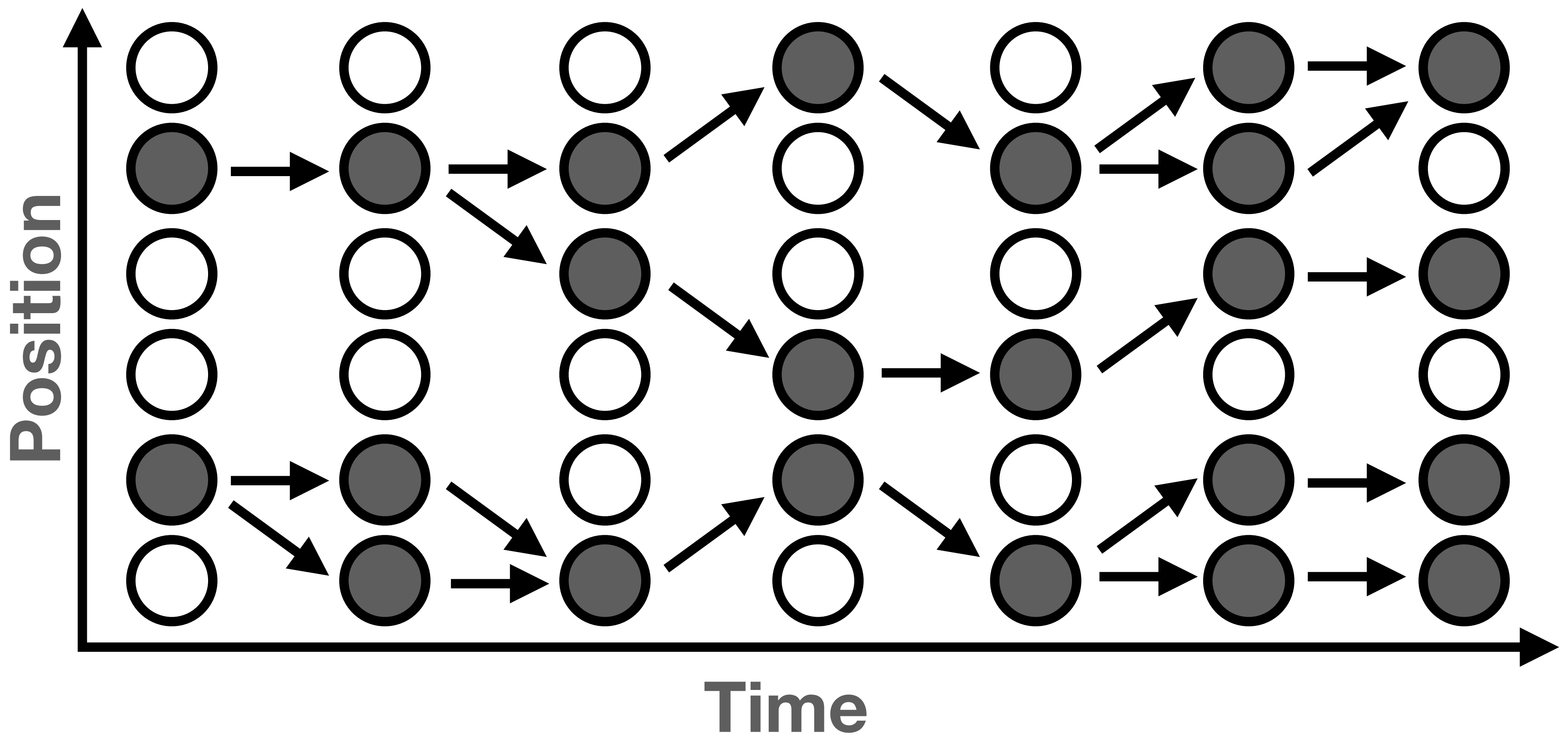


A Better Understanding on XEB and Fidelity

Recall: Fidelity captures how well a simulation is but it is hard to estimate. In practice, XEB serves as a proxy for fidelity.



- XEB can deviate from fidelity! Both for our classical algorithms and noisy simulations.
- We also develop an analytical model to understand when and how do such deviations could happen.



A Glimpse into Our Analytical Models for XEB & Fidelity

Overview of Our Analytical Models for XEB & Fidelity

$$\mathbb{E}_U[\langle \psi_U | \rho_C | \psi_U \rangle]$$

Fidelity of our algorithm

$$1 + \mathbb{E}_U[\chi_U(C)]$$

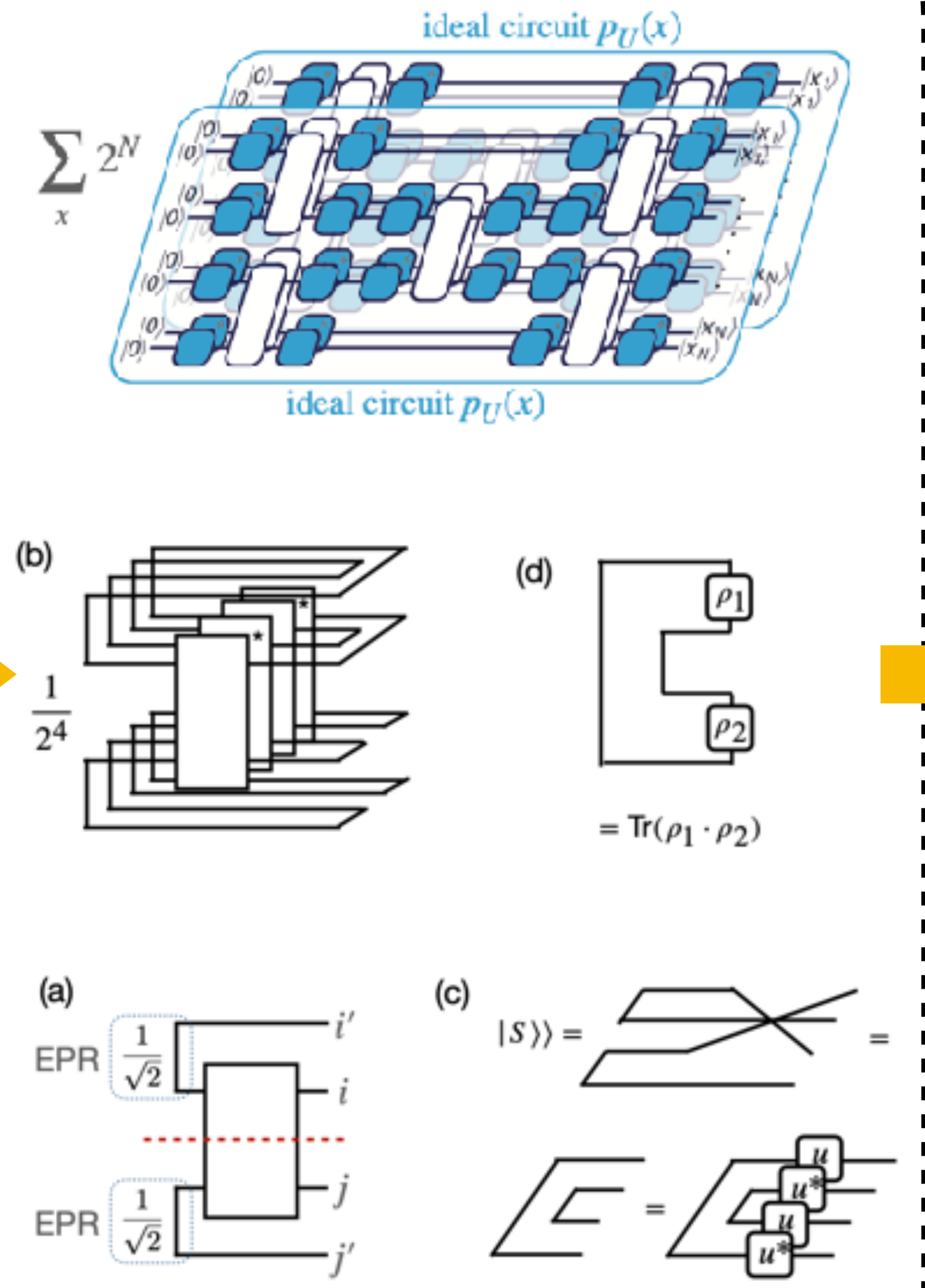
XEB of our algorithm

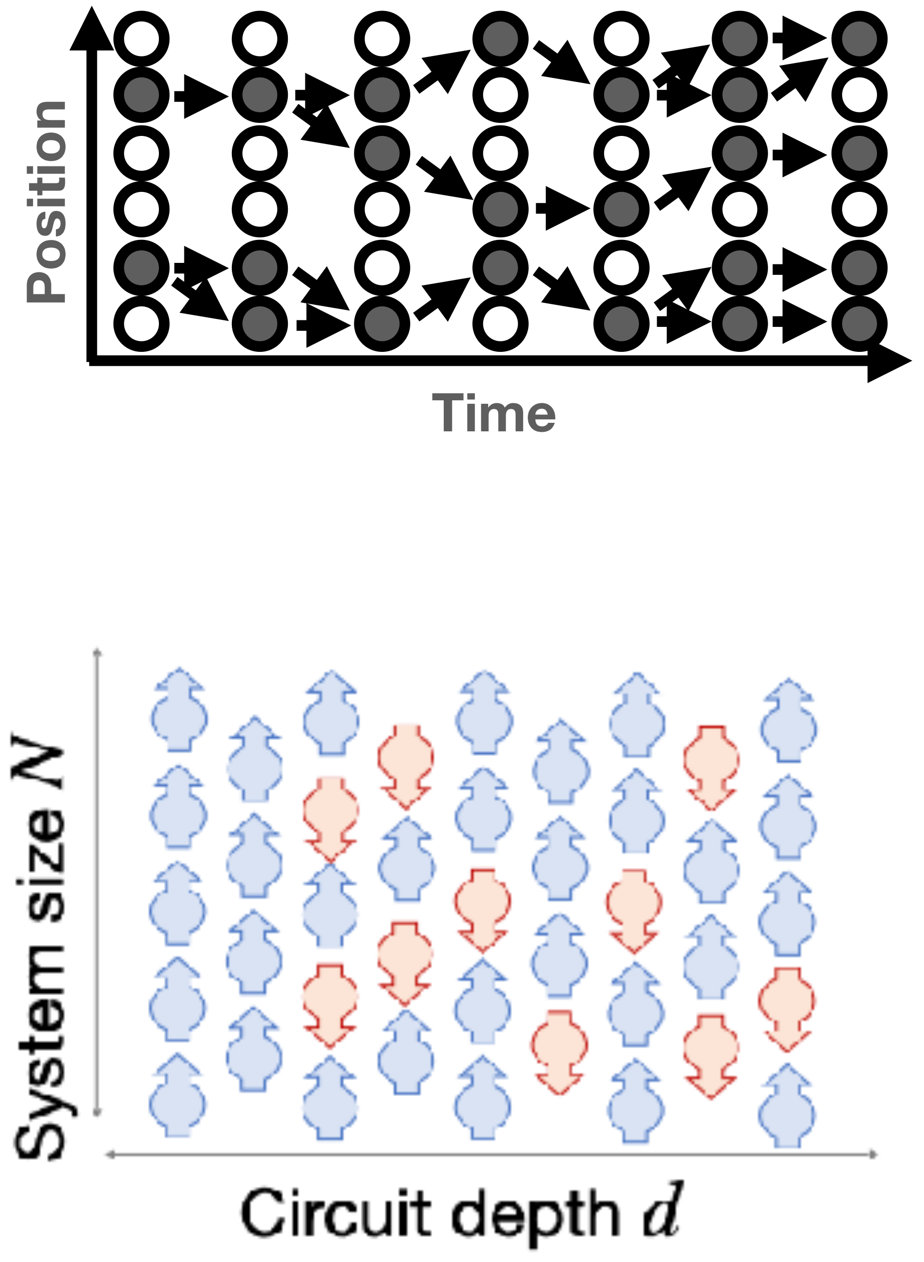
$$\mathbb{E}_U[\langle \psi_U | \rho_\epsilon | \psi_U \rangle]$$

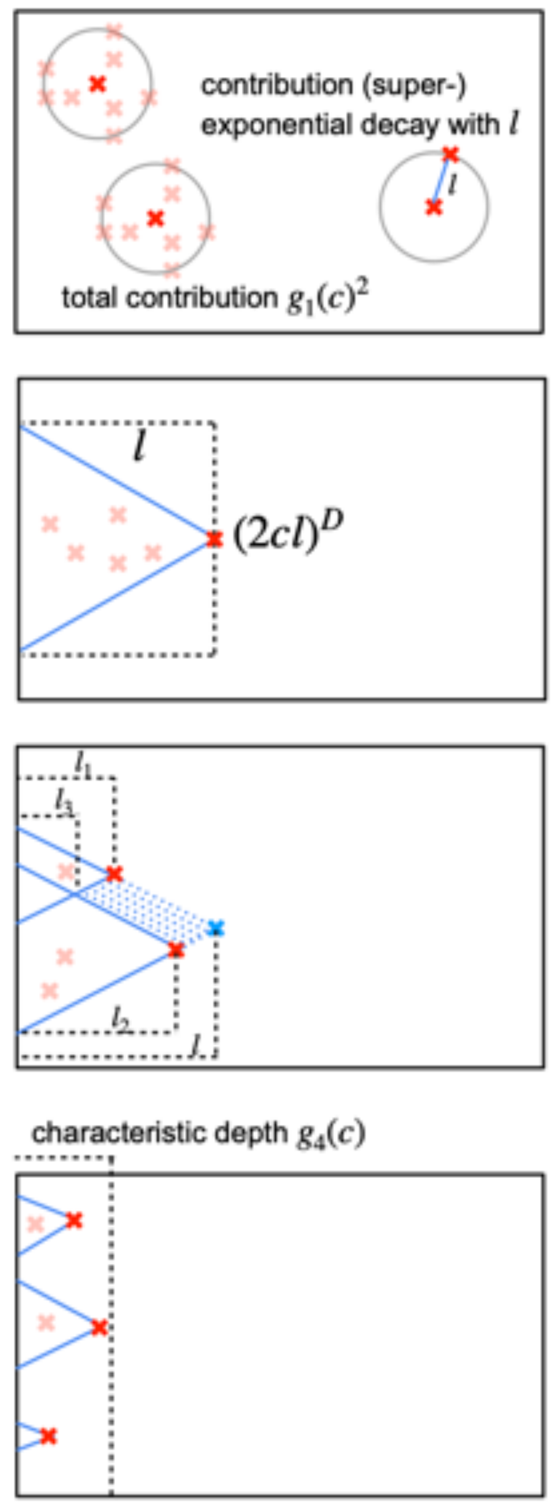
Fidelity of a noisy simulation

$$1 + \mathbb{E}_U[\chi_U(\mathcal{N}_\epsilon)]$$

XEB of a noisy simulation







Step 0:
Expected value

Step 1:
Tensor networks

Step 2:
Statistical mechanics models

Step 3:
Quantitative analysis

Overview of Our Analytical Models for XEB & Fidelity

$$\mathbb{E}_U[\langle \psi_U | \rho_C | \psi_U \rangle]$$

Fidelity of our algorithm

$$1 + \mathbb{E}_U[\chi_U(C)]$$

XEB of our algorithm

$$\mathbb{E}_U[\langle \psi_U | \rho_\epsilon | \psi_U \rangle]$$

Fidelity of a noisy simulation

$$1 + \mathbb{E}_U[\chi_U(\mathcal{N}_\epsilon)]$$

XEB of a noisy simulation

Step 0:
Expected value

→

Graphical representations for complicated (multi-)linear algebra and beyond

Step 1:
Tensor networks

→

Statistical physics models the **exactly** captures the expected value of XEB and fidelity.

Step 2:
Statistical mechanics models

→

Intuitive models that captures the **quantitative behavior** of XEB and fidelity under noises and gate removal.

Step 3:
Quantitative analysis

Overview of Our Analytical Models for XEB & Fidelity

$\mathbb{E}_U[\langle\psi_U|\rho_C|\psi_U\rangle]$

Fidelity of our algorithm

$1+\mathbb{E}_U[\chi_U(C)]$

XEB of our algorithm

$\mathbb{E}_U[\langle\psi_U|\rho_\epsilon|\psi_U\rangle]$

Fidelity of a noisy simulation

$1+\mathbb{E}_U[\chi_U(\mathcal{N}_\epsilon)]$

XEB of a noisy simulation

Step 0:
Expected value

Graphical representations for complicated (multi-)linear algebra and beyond

Step 1:
Tensor networks

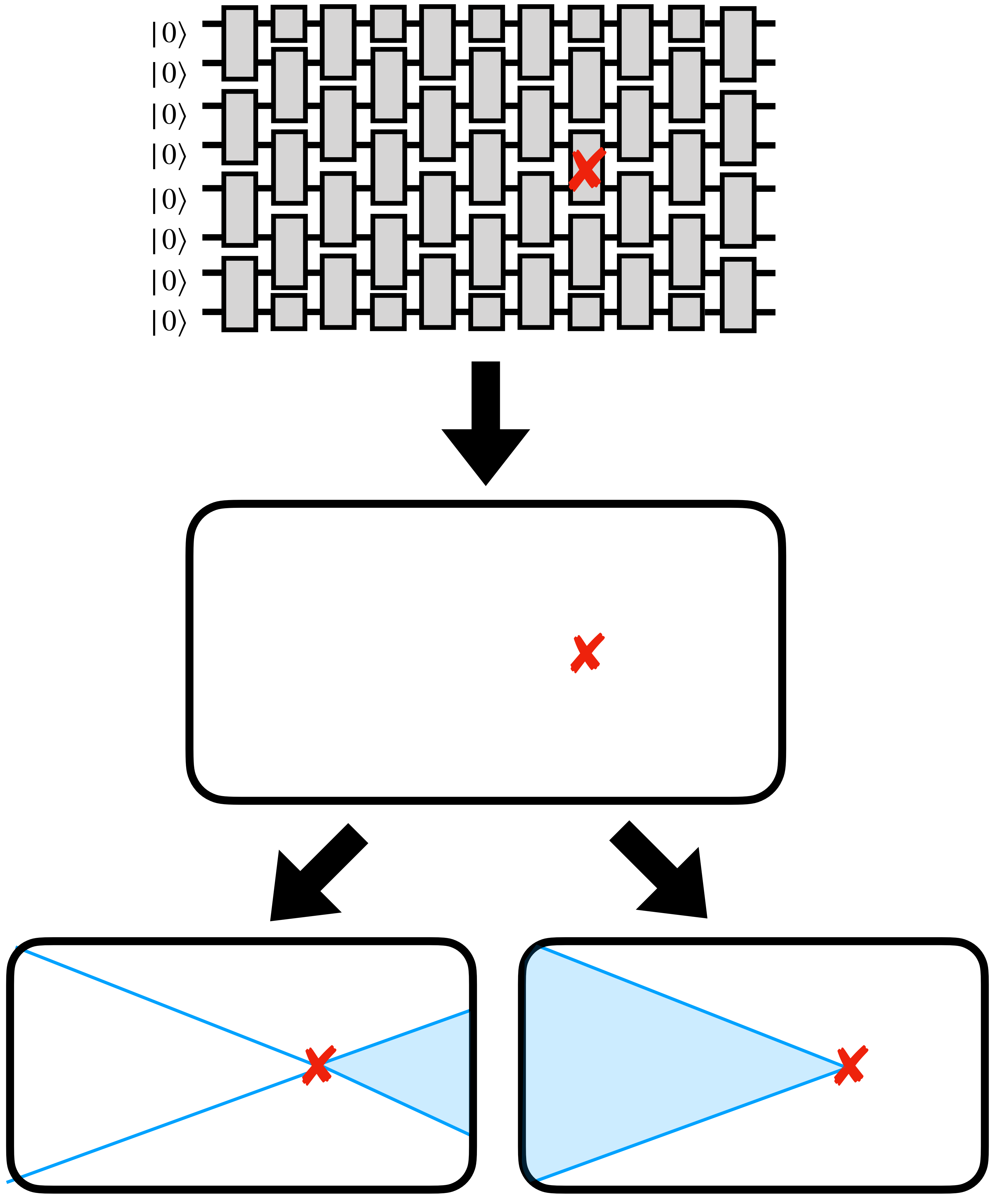
Statistical physics models the **exactly** captures the expected value of XEB and fidelity.

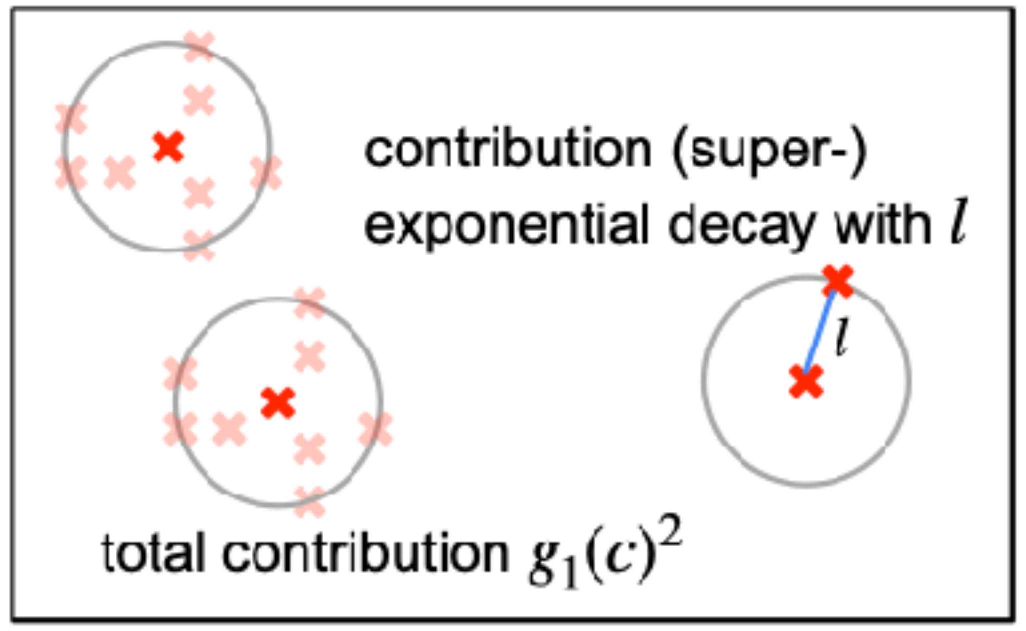
Step 2:
Statistical mechanics models

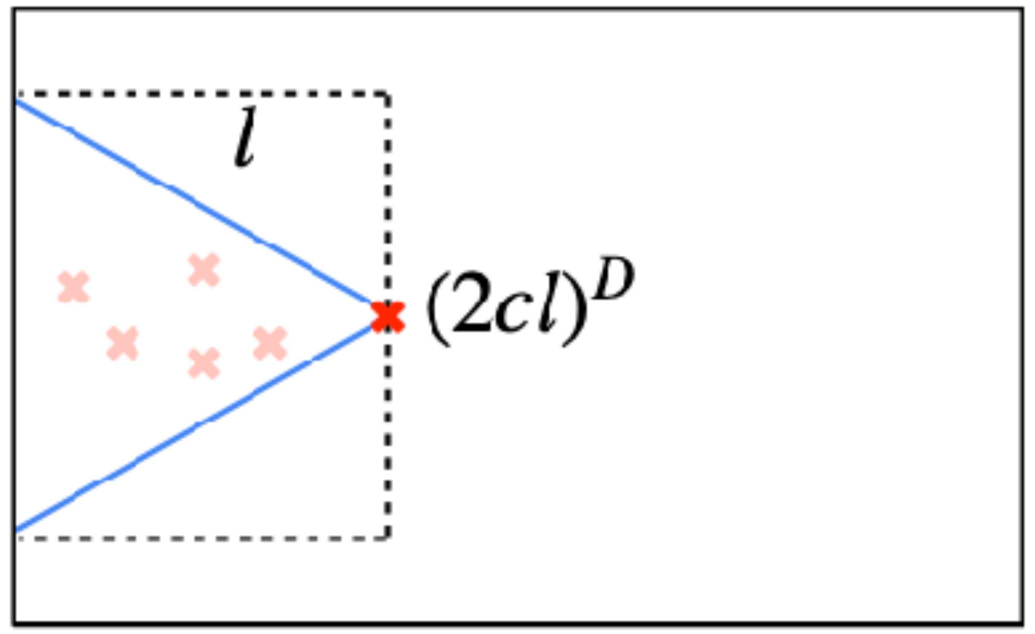
Intuitive models that captures the **quantitative behavior** of XEB and fidelity under noises and gate removal.

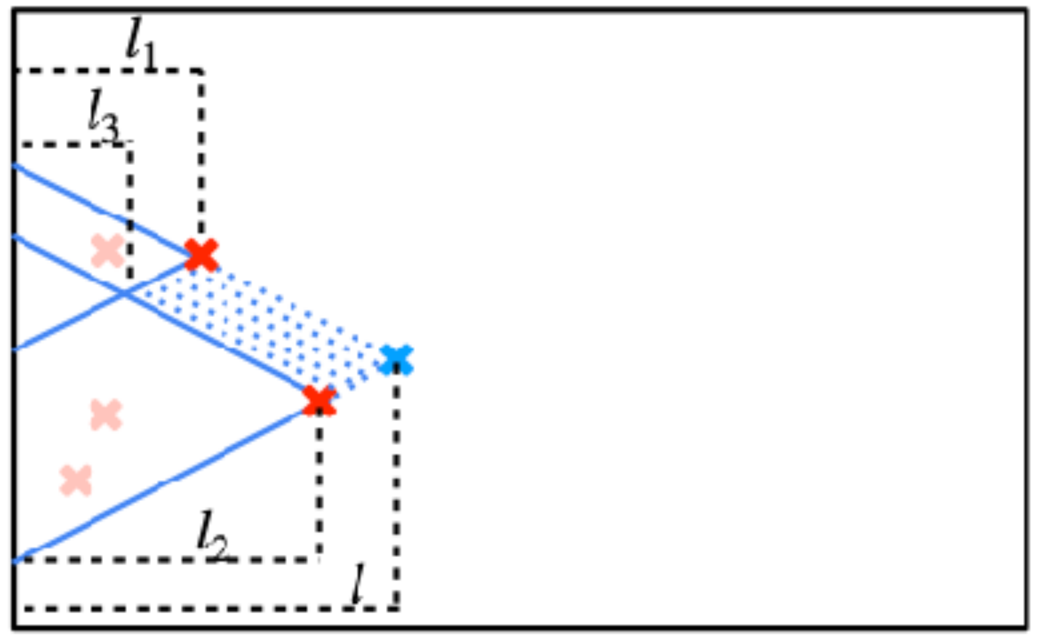
Step 3:
Quantitative analysis

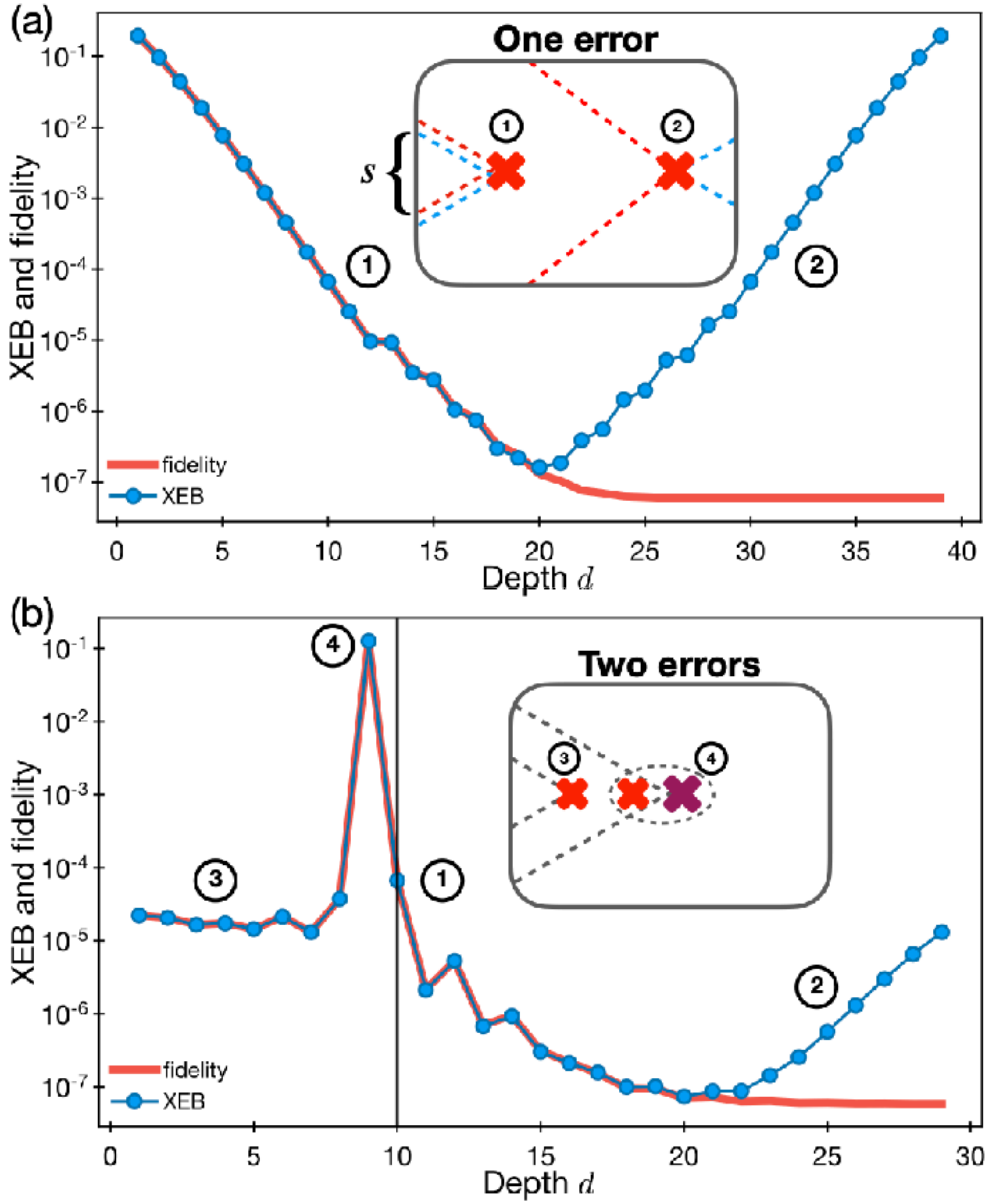
Our Quantitative Analysis for XEB & Fidelity (Step 3)



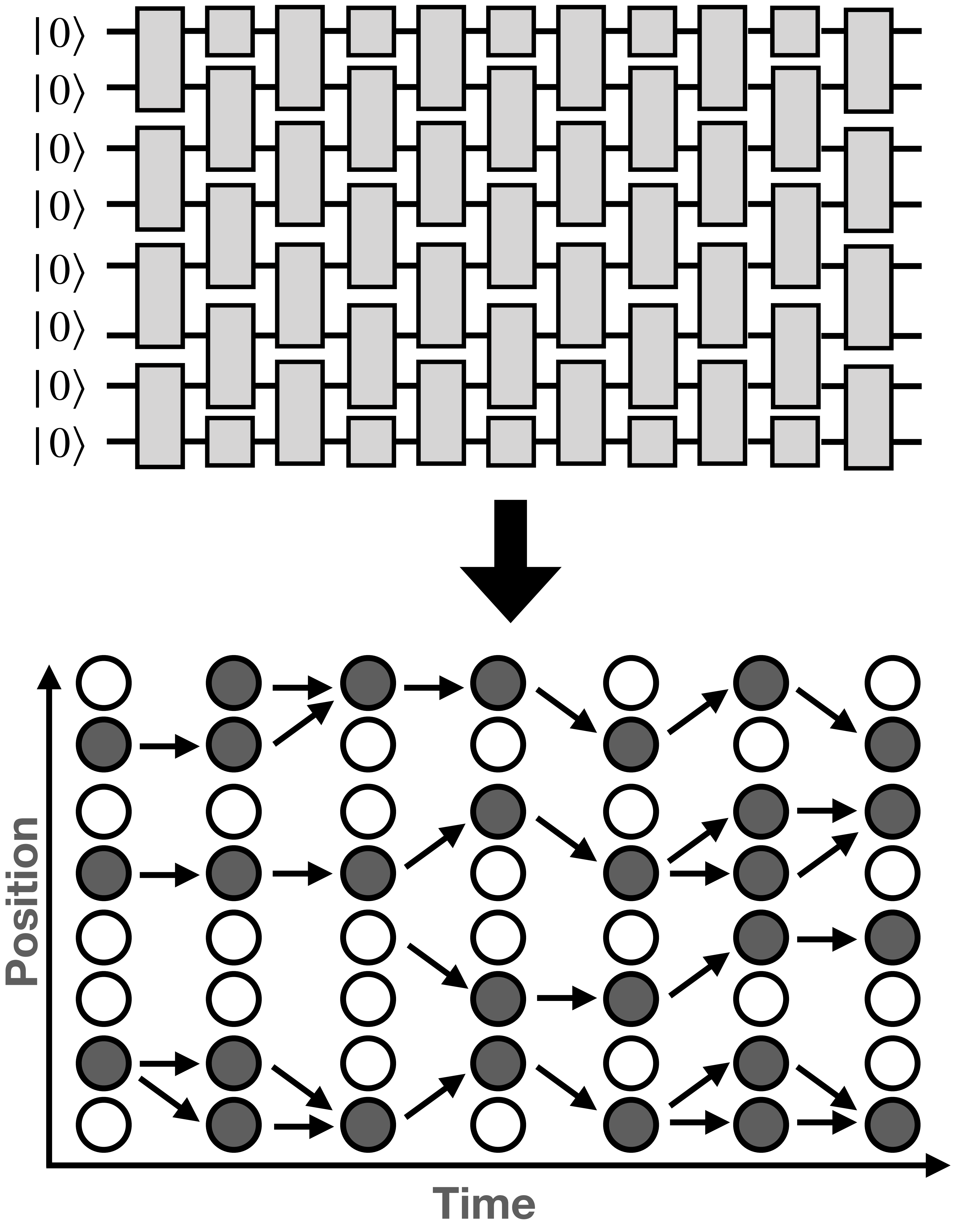




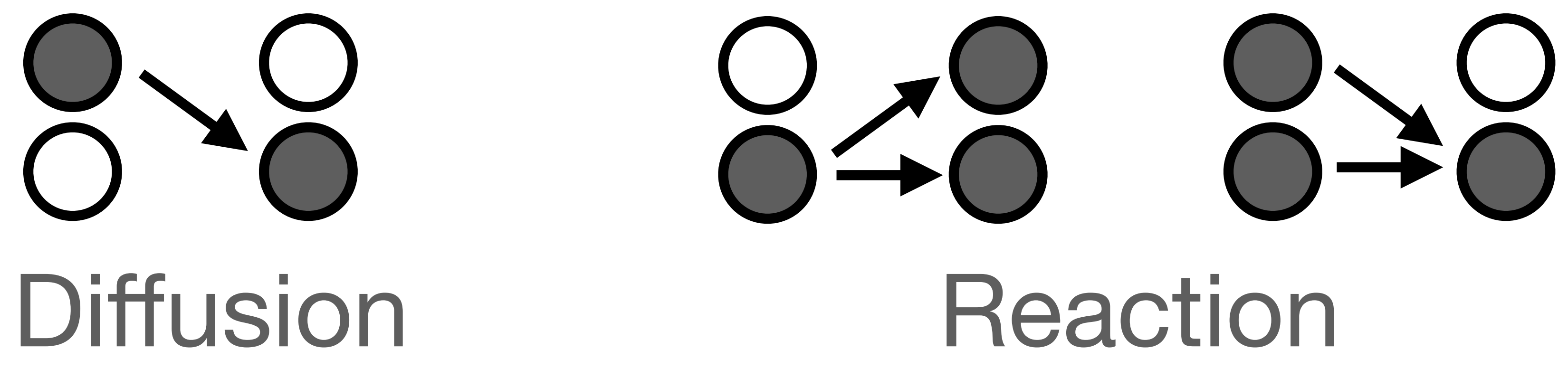




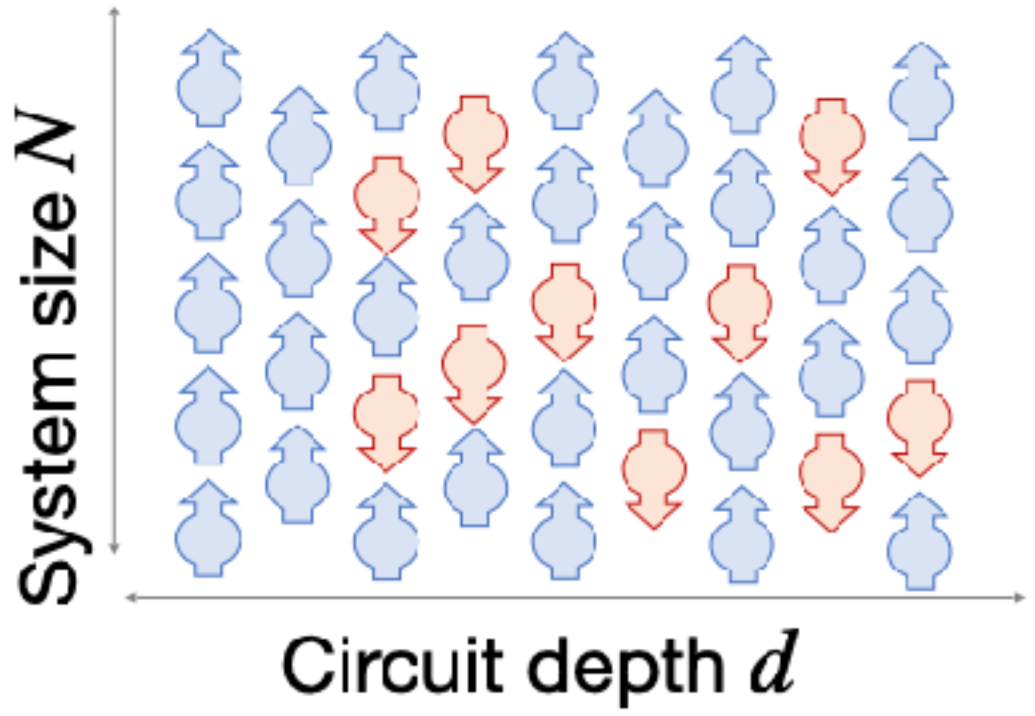
Our Statistical Physics Models for XEB & Fidelity (Step 2)



- Both XEB & fidelity are (exactly) mapped to statistics in a diffusion-reaction model!

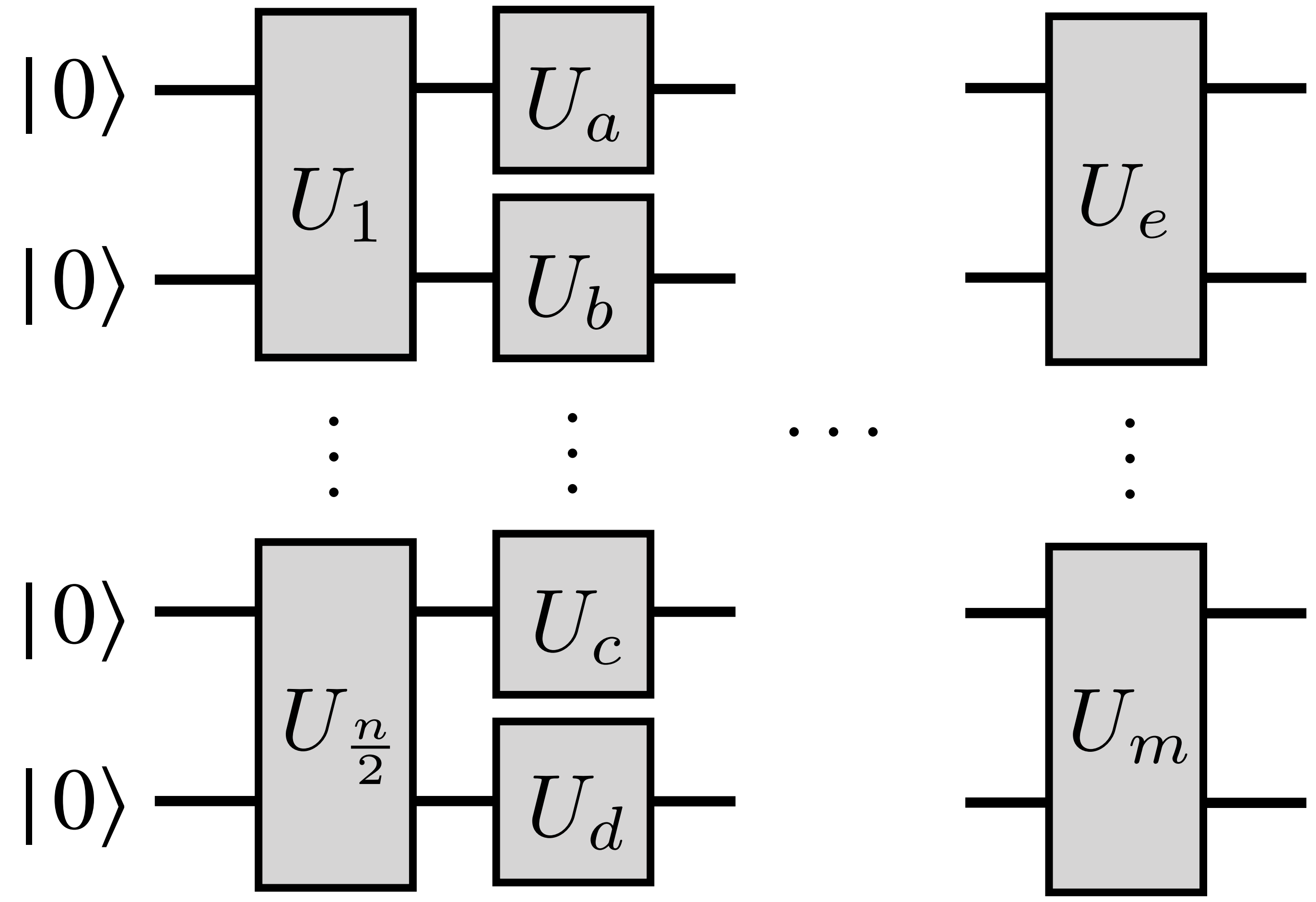


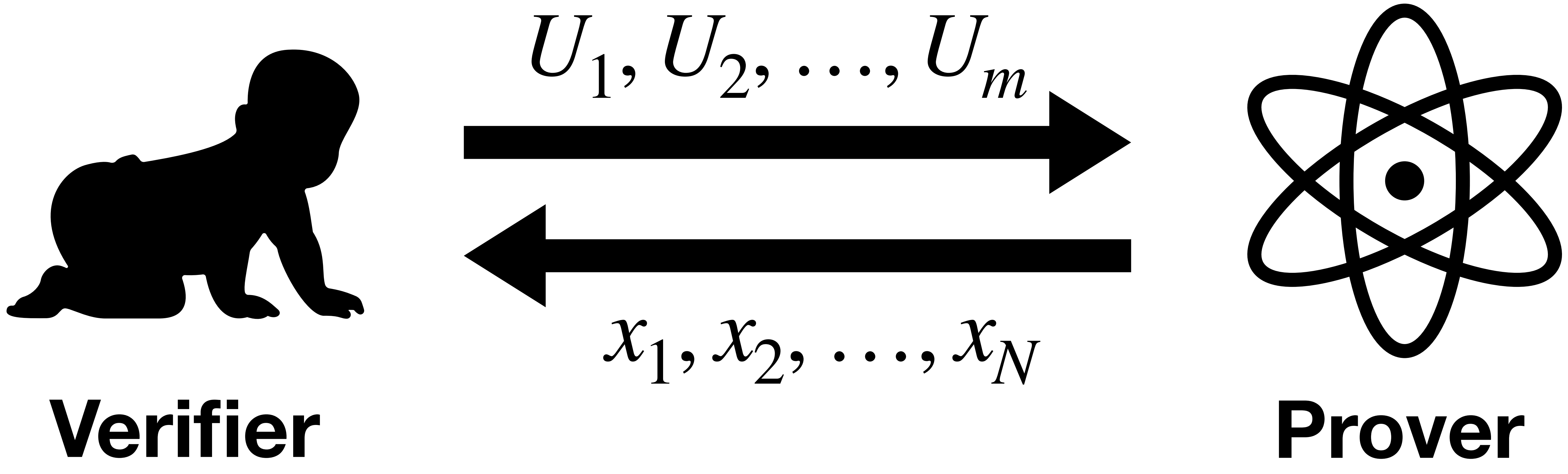
- Noise (from a noisy simulation) and gate removal (from our algorithms) change the transition probability differently.
- For 1D circuits, we further have a 2D Ising model that gives us more quantitative results!



The Complexity-Theoretic Aspect

Linear XEB as a Computational Problem





Linear Cross-Entropy Quantum Threshold Assumption [AG20]

There's a constant $c > 0$ such that there's no polynomial time classical algorithm to produce an estimation p for $p_U(0^n)$ with

$$\mathop{\mathbb{E}}_{U}[(p_U(0^n) - p)^2] = \mathop{\mathbb{E}}_{U}[(p_U(0^n) - 2^{-n})^2] - \Omega(2^{-cn}) .$$

Refuting XQUATH for Constant Dimensional Circuits

Linear Cross-Entropy Quantum Threshold Assumption [AG20]

There's a constant $c > 0$ such that there's no polynomial time classical algorithm to produce an estimation p for $p_U(0^n)$ with

$$\mathop{\mathbb{E}}_{U}[(p_U(0^n) - p)^2] = \mathop{\mathbb{E}}_{U}[(p_U(0^n) - 2^{-n})^2] - \Omega(2^{-cn})\,.$$

Theorem

There's a polynomial-time classical algorithm that produces an estimation p for $p_U(0^n)$ in D-dimensional random circuits with

$$\mathop{\mathbb{E}}_{U}[(p_U(0^n) - p)^2] = \mathop{\mathbb{E}}_{U}[(p_U(0^n) - 2^{-n})^2] - \Omega(2^{-cd})$$

where d is the depth of the circuit and $c > 0$ is a constant.

Intuitions

Recall: People believe that Schrödinger's algorithm & Feynman's algorithm are the best.

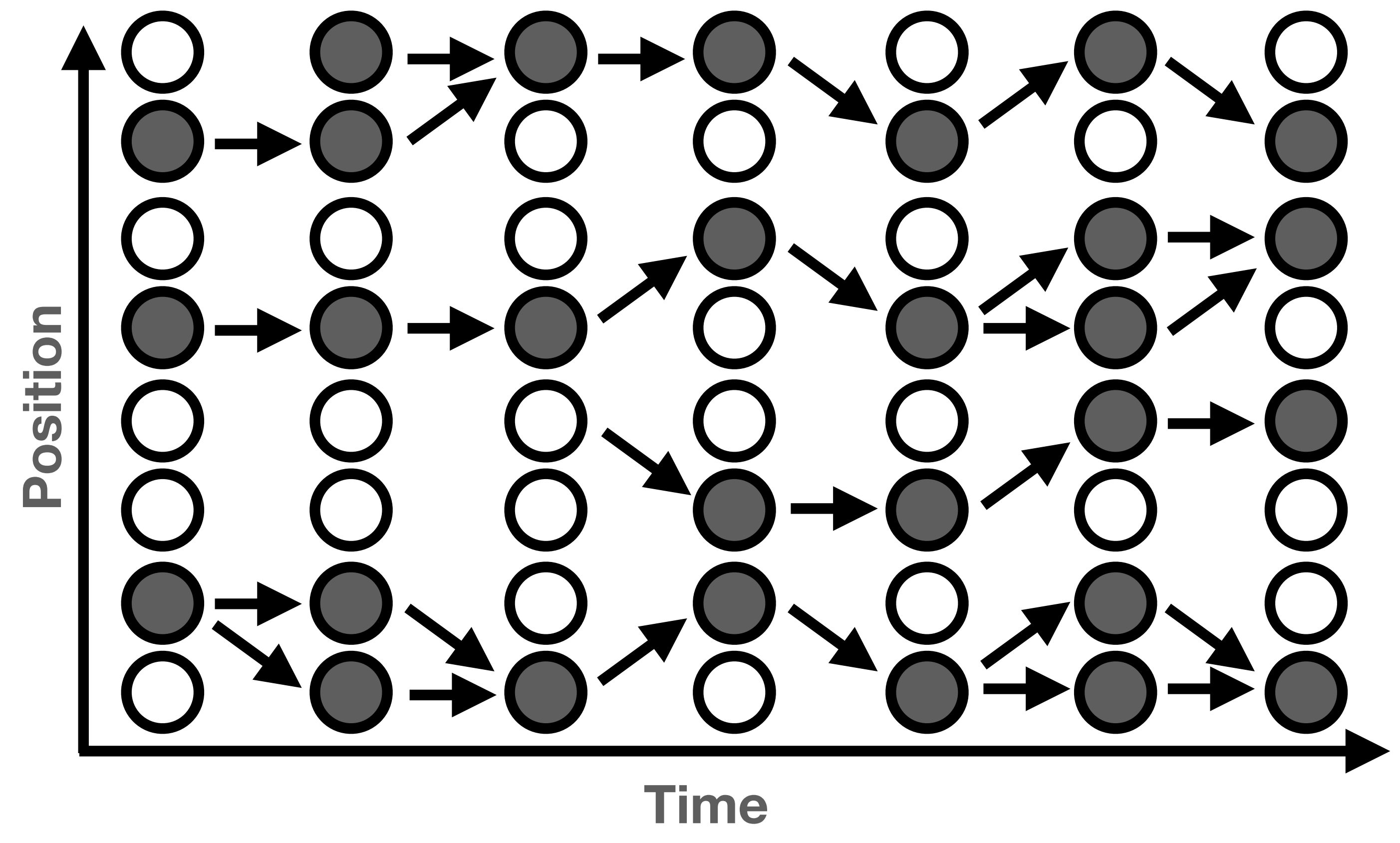
Feynman's Algorithm

Express the final quantum state as a sum of all possible path from the input layer.

- Time: $O(4^m)$.
- Space: $O(m+n)$.

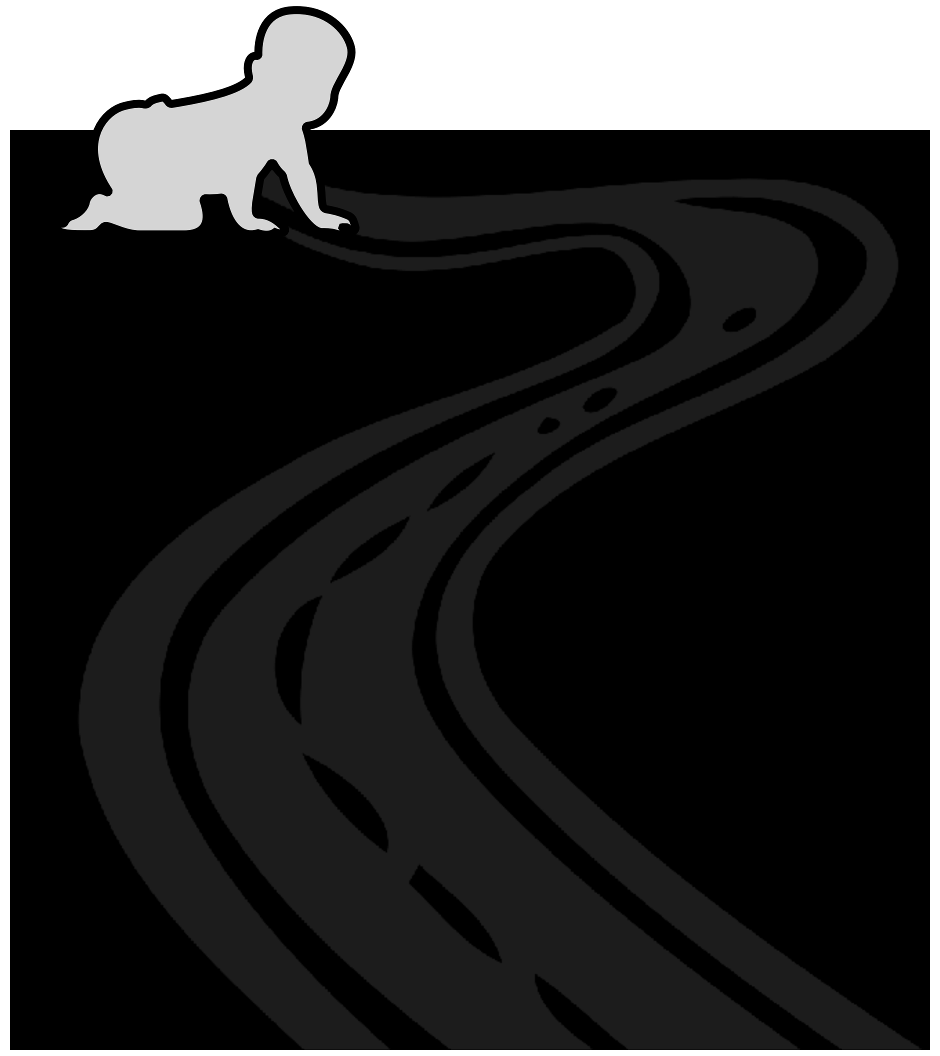
$$|\Psi\rangle = \sum_{x_1,\ldots,x_{m-1}\in\{0,1\}^4} U_m|x_{m-1}\rangle\cdots|x_2\rangle\underbrace{\langle x_2|U_2|x_1\rangle}_{\text{A scalar}}\underbrace{\langle x_1|U_1|0^n\rangle}_{\text{A scalar}}$$

Enumerate all $x_1,\ldots,x_{m-1}\in\{0,1\}^4$ and compute the summation.



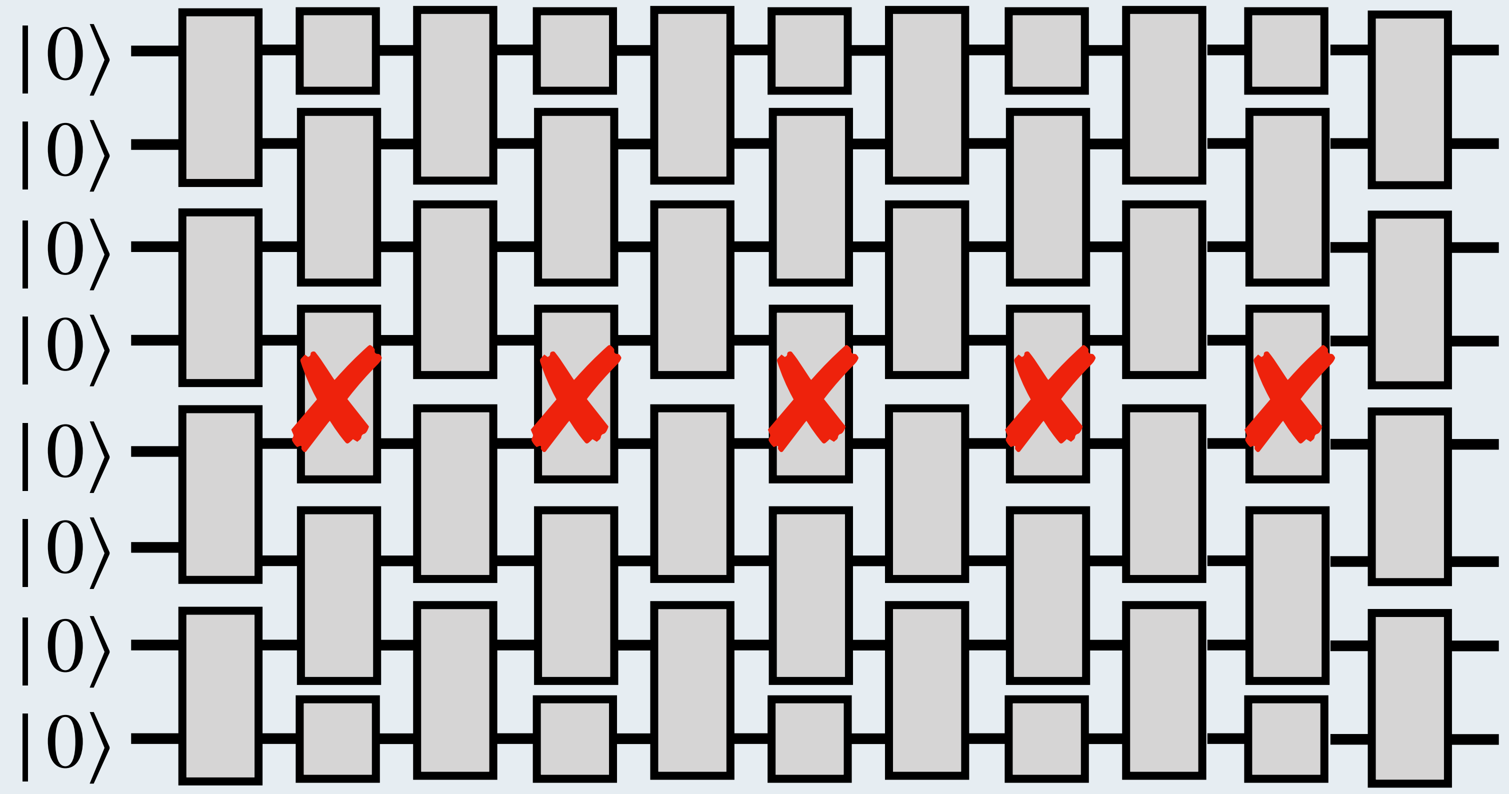
Our diffusion-reaction model shows that path integration in a different basis can leads to efficient classical algorithms for constant dimensional random circuits!

Summary & Future Directions



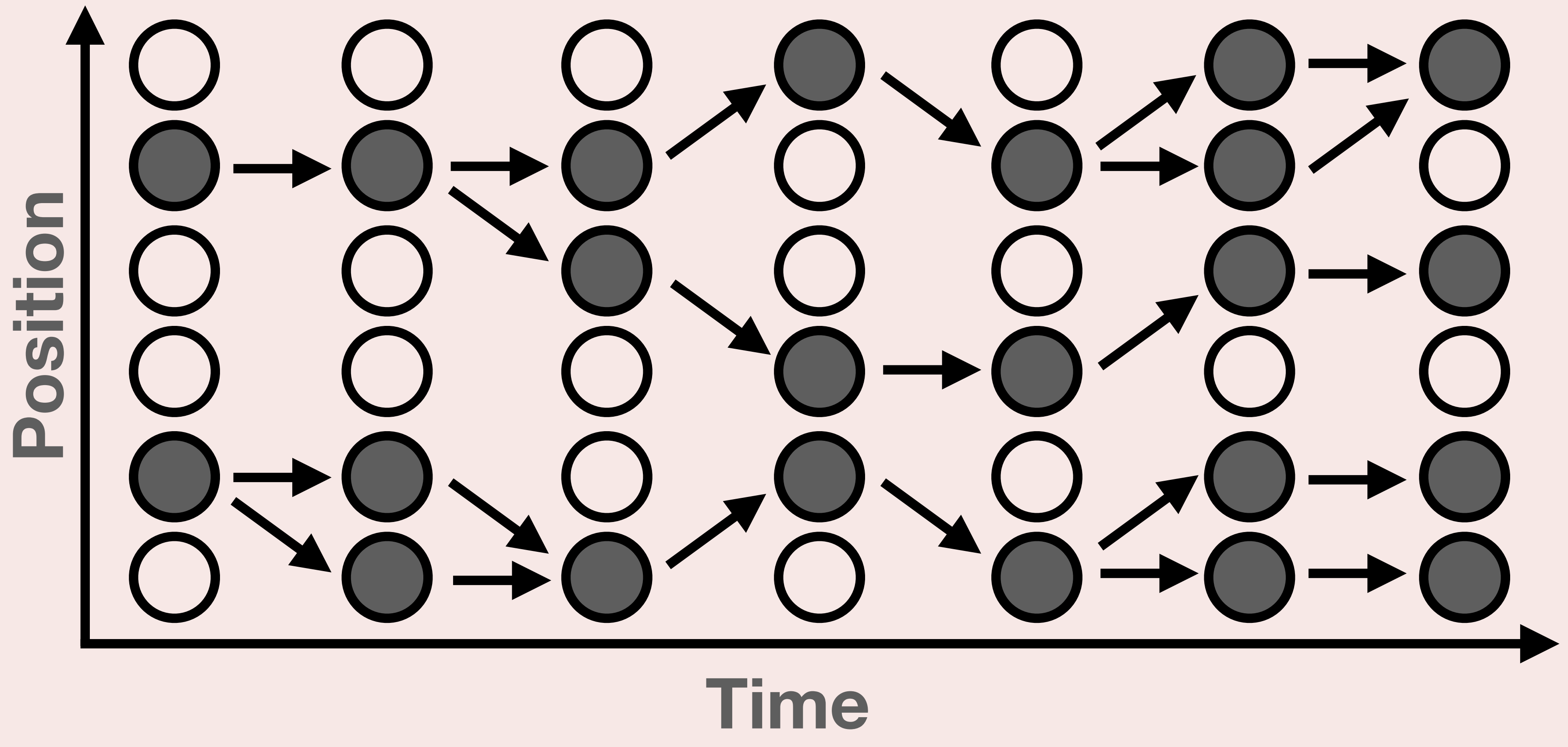
Summary

Classical Algorithms Spoofing XEB



1. **Complexity-theoretically**, refute the XQUATH in realistic circuit architectures.
2. **Experimentally**, achieve 2%~12% of Google's and USTC's XEB in ~1s on 1 GPU.
3. **Asymptotically**, our algorithms are scalable.

A Better Understanding on XEB and Fidelity

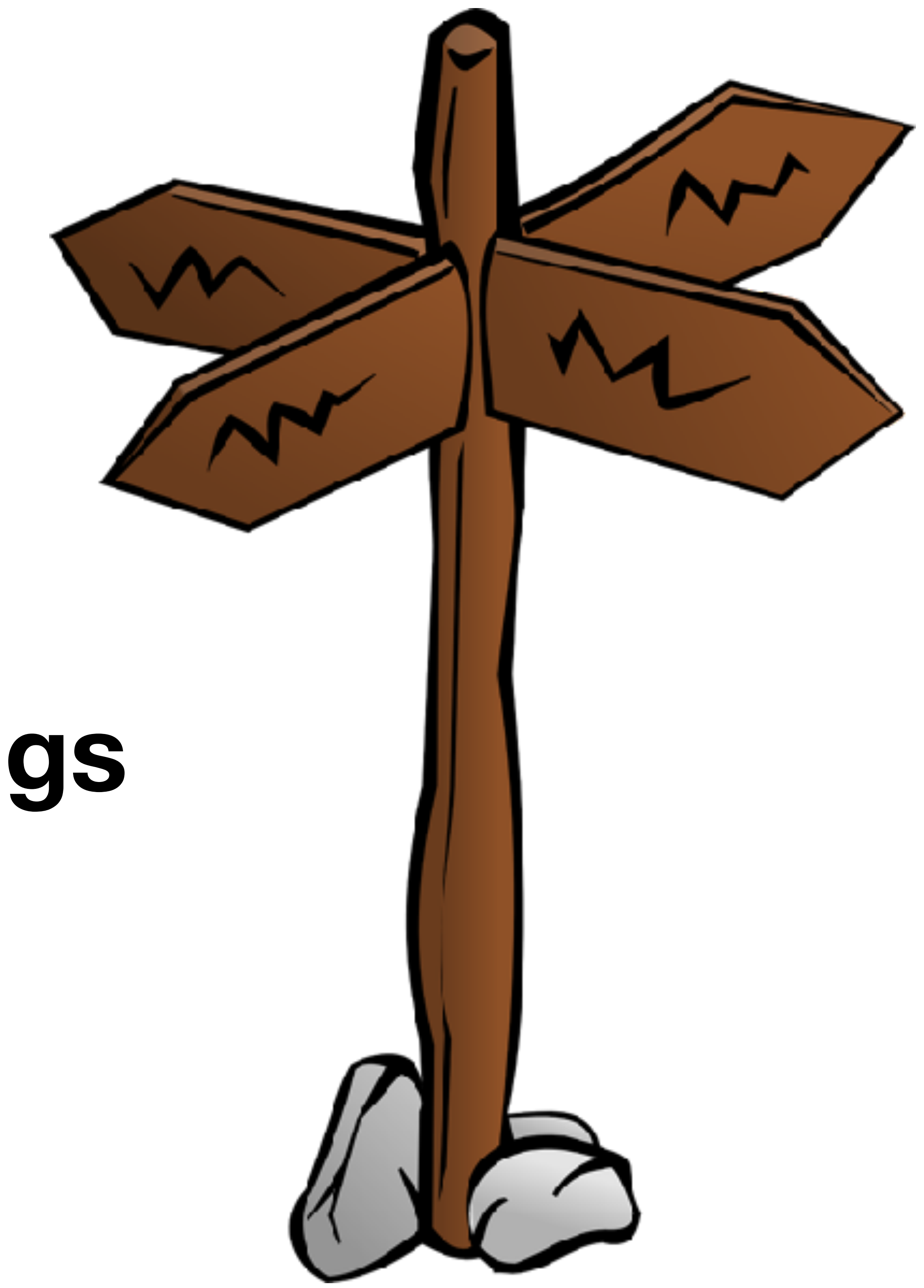


1. **XEB can overestimate fidelity** in both our algorithms and noisy simulation.
2. **XEB has a limited utility as a benchmark** for quantum advantage.

Future Directions

Ask me offline!

Improving our spoofing algorithms?



New quantum advantage proposal?

Fine-grained understandings in what's the applicable regime for XEB?

New complexity-theoretic foundation for RCS-based quantum advantage?

Thanks for your attention 🙂

arXiv:2112.01657