Quantum computational advantage with constant-temperature Gibbs sampling

Yunchao Liu (UC Berkeley → Harvard)

with Thiago Bergamaschi and Chi-Fang Chen, 2404.14639v2, FOCS 2024

Quantum computational advantage: what's next?

- Demonstrate computational advantage on new physical platforms
  - E.g. analog devices

- Develop quantum algorithms toward useful quantum advantage

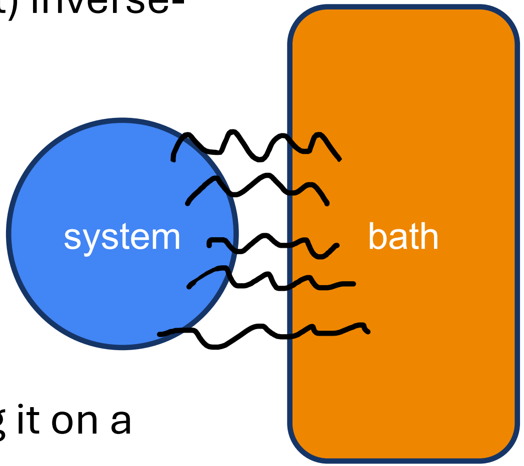
Here's a quantum algorithm

A quantum system is coupled to a thermal bath at finite (constant) inverse-temperature β

1. Engineer the system in a desired Hamiltonian H
2. Wait for the system to converge to its Gibbs state

$$\rho_\beta \propto e^{-\beta H}$$

3. Measure in the standard basis, obtain sample



Think of actually implementing this process, as well as simulating it on a quantum computer

This talk: complexity theoretic evidence of quantum computational advantage in this model

What makes this challenging?

- At high enough temperatures, sampling from Gibbs states is classically simulable
    - “High-Temperature Gibbs States are Unentangled”
- At low enough temperatures, this task is hard in general even for quantum computers
    - At least NP-hard due to classical PCP theorem

Hamiltonians which are “classically hard, but quantumly easy” are a sweet spot: How to make it classically hard, but not too hard?

Construction: the second simplest example you can think of

$$H = -C\left(\sum_{i=1}^{n} Z_i\right)C^{\dagger}$$

Shallow quantum circuit

- Classically, hard to sample from Gibbs state: uses hardness of shallow quantum circuits + fault tolerance
- Quantumly, thermalization process is rapidly mixing: uses lightcone structure of shallow quantum circuits
- This is an example of a “sweet spot”

Efficiently samplable, but classically intractable Gibbs states

Task: Given a local Hamiltonian H & an inverse-temperature β, approximately sample from

$$p(x) = \langle x| \rho_\beta |x\rangle \,, \text{ where } \rho_\beta = \frac{e^{-\beta H}}{\operatorname{Tr} e^{-\beta H}}$$

Theorem. There exists a family of n qubit, $O(1)$ local Hamiltonians at any finite temperature β, which is

- Rapidly thermalizing (and thus efficiently samplable), in time $n^{o(1)}$
  Can be simulated on a quantum computer in time $n^{1+o(1)}$
- Classically intractable under standard complexity-theoretic assumptions

Our approach

Goal: construct a family of local Hamiltonians, which is both

Classically Intractable by embedding computation into its Gibbs state

- Gibbs states are typically “noisy” versions of the ground state.
- Use fault tolerance to correct the “noise”

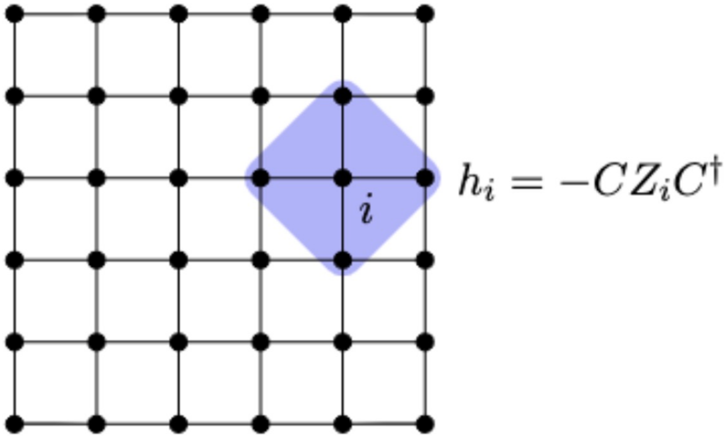
Rapidly Thermalizing i.e. converging to the Gibbs state in less than polynomial time

- Can be quite challenging, even in commuting systems

Parent Hamiltonians of shallow quantum circuits

Starting with a non-interacting system,

$$H_{\mathsf{NI}} = -\sum_i Z_i$$



Consider the class of "parent" Hamiltonians

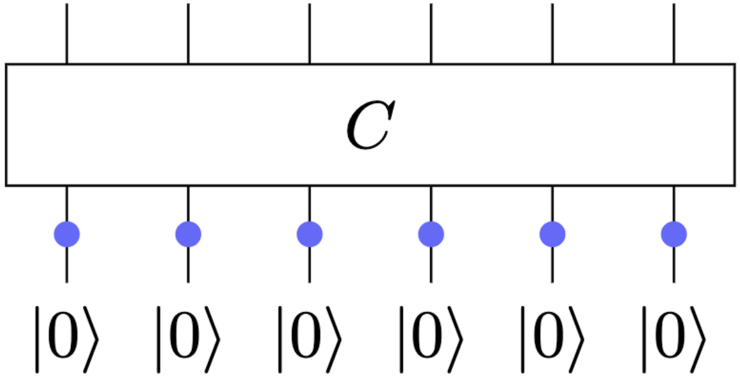
$$\mathscr{H} = \left\{ H : \exists \text{ low-depth circuit } C,\ H = CH_{\mathsf{NI}}C^\dagger \right\}$$

Nice Properties: Local, commuting, integer spectra, and its ground-state is $C\,|0\rangle^{\otimes n}$

The input noise model

Their Gibbs state resembles a noisy version of the circuit

$$\rho_\beta = \frac{e^{-\beta H}}{\operatorname{Tr} e^{-\beta H}} =$$

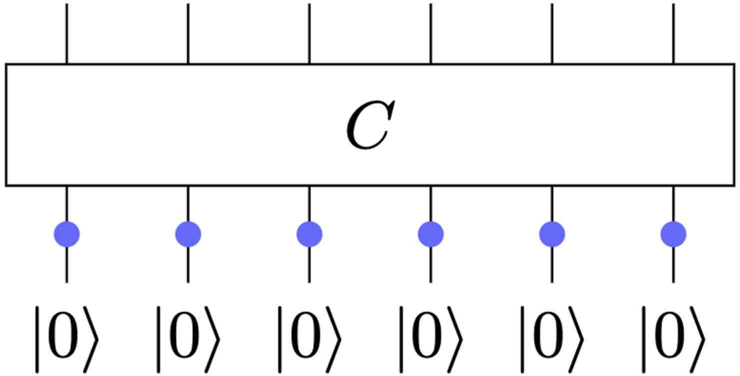


The input noise model

$$\boxed{\mathcal{D}_p(\sigma) = (1-p)\cdot\sigma + p\cdot\mathrm{Tr}\,\sigma\cdot\frac{\mathbb{I}}{2}.}$$

Their Gibbs state resembles a noisy version of the circuit

$$\rho_\beta = \frac{e^{-\beta H}}{\mathrm{Tr}\,e^{-\beta H}} =$$



Note $$e^{-\beta H_{\mathrm{NI}}} = \otimes_i e^{\beta Z_i} \propto \Big(D_p(|0\rangle\langle 0|)\Big)^{\otimes n}$$

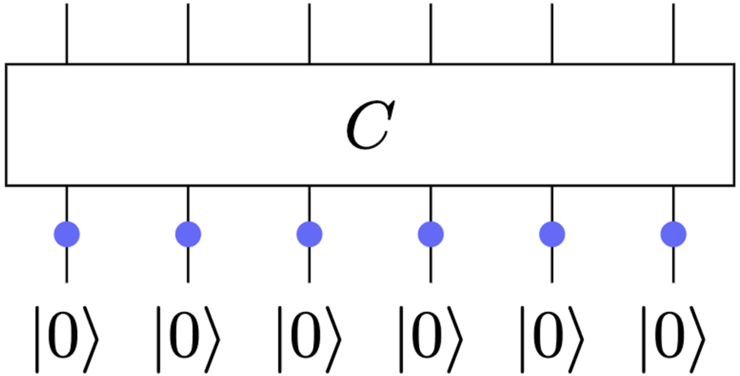
$$p \approx e^{-2\beta}$$

The input noise model

$$Ve^{A}V^{\dagger} = e^{VAV^{\dagger}}$$

Their Gibbs state resembles a noisy version of the circuit

$$\rho_\beta = \frac{e^{-\beta H}}{\operatorname{Tr} e^{-\beta H}} =$$



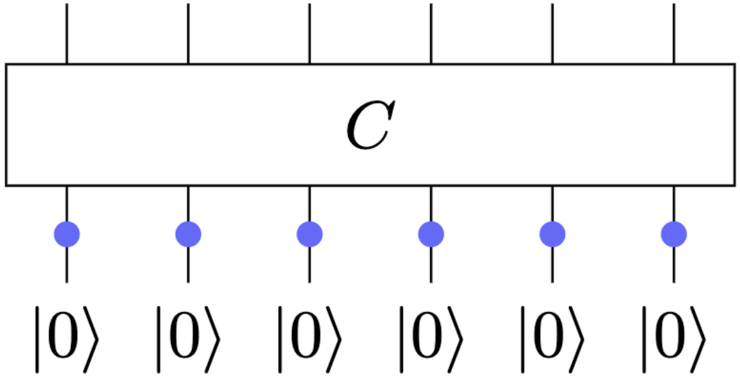
Note $$e^{-\beta H_{\mathsf{NI}}} = \otimes_i e^{\beta Z_i} \propto \Big(D_p(|0\rangle\langle 0|)\Big)^{\otimes n} \qquad e^{-\beta H} = C e^{-\beta H_{\mathsf{NI}}} C^{\dagger} \propto C\Big(\mathcal{D}_p(|0\rangle\langle 0|)\Big)^{\otimes n} C^{\dagger}$$

$$p \approx e^{-2\beta}$$

The input noise model

Their Gibbs state resembles a noisy version of the circuit

$$\rho_\beta = \frac{e^{-\beta H}}{\mathrm{Tr}\, e^{-\beta H}} =$$



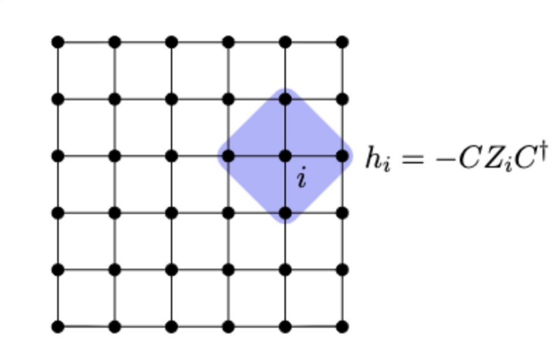
[BMS16] Many classically-hard shallow circuits become simulable under input noise

Need to embed some form of fault-tolerance into the circuit

Outline

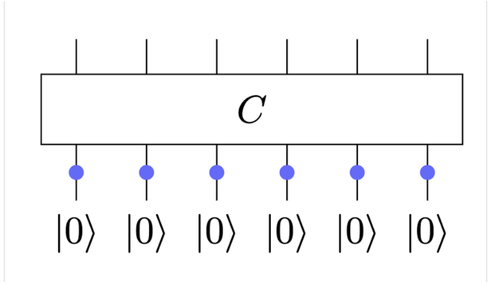
I. Efficient Gibbs sampling

Rapid mixing bounds for Lindbladians,
via lightcone arguments



II. Fault tolerance of IQP circuits

Designing fault-tolerant circuits which are
hard-to-sample from under input noise



Thermalization

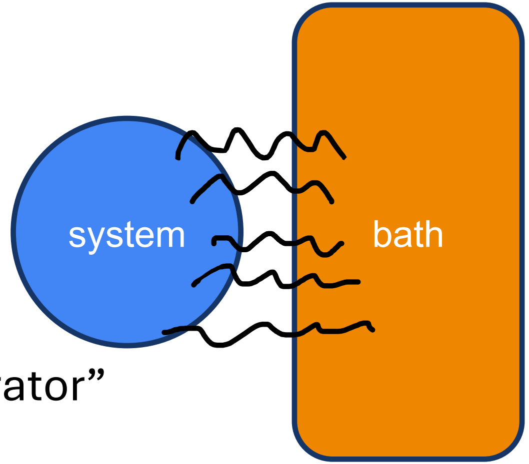
1. What do we mean by "a system is coupled to a bath", or "put a quantum system in a fridge"

2. What is needed to prove rapid mixing for thermalization

3. How to prove it for our Hamiltonians

4. (skipped) How to simulate this process on a digital quantum computer

“A system is coupled to a bath”



- System and bath in a joint unitary evolution
- Trace out the bath, focus on the system dynamics
- Described by a specific Lindbladian called “Davies generator”

- Intuition: a continuous-time quantum Markov chain, jumping around in the system Hamiltonian eigenbasis

- No matter the initial state, the system always converges to the Gibbs state $\rho_\beta \propto e^{-\beta H}$
    - Need to bound mixing time: how fast does it converge

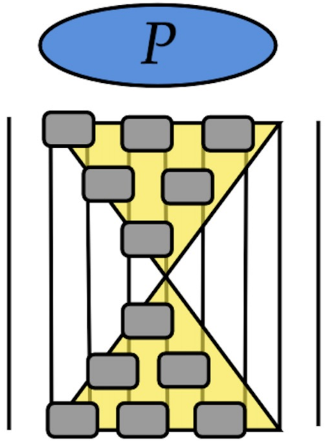
Thermal Lindbladians and Davies Generators

A set of jump operators

$$\{A^a\}_{a\in\mathcal{A}} \propto \left\{\ell\text{-local Paulis } P \in \mathcal{P}_\ell \text{ on each lightcone } i \in [n]\right\}$$

And transition weights

$$\gamma_\beta(\omega) \equiv \gamma(\omega) = 1/(1+e^{-\beta\omega})$$



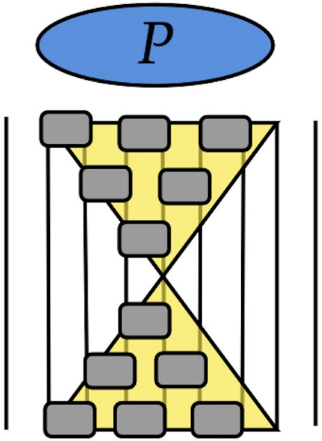
Thermal Lindbladians and Davies Generators

A set of jump operators

$$\{A^a\}_{a\in\mathcal{A}} \propto \left\{\ell\text{-local Paulis } P \in \mathcal{P}_\ell \text{ on each lightcone } i \in [n]\right\}$$

And transition weights

$$\gamma_\beta(\omega) \equiv \gamma(\omega) = 1/(1+e^{-\beta\omega})$$



Define a Davies Generator

$$\mathcal{L}[\rho] = \sum_{a\in\mathcal{A}}\sum_{\nu}\gamma(\nu)\left(A^a_\nu\rho(A^a_\nu)^\dagger - \frac{1}{2}\left\{(A^a_\nu)^\dagger A^a_\nu, \rho\right\}\right)$$

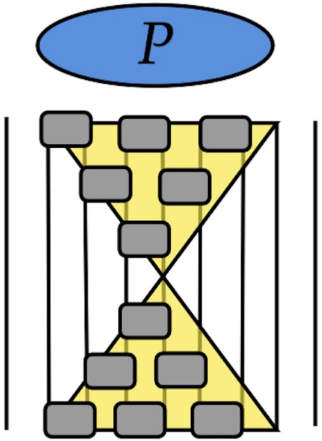
Thermal Lindbladians and Davies Generators

A set of jump operators

$$\{A^a\}_{a\in\mathcal{A}} \propto \left\{ \ell\text{-local Paulis } P \in \mathcal{P}_\ell \text{ on each lightcone } i \in [n] \right\}$$

And transition weights

$$\gamma_\beta(\omega) \equiv \gamma(\omega) = 1/(1+e^{-\beta\omega})$$



Define a Davies Generator

$$\mathcal{L}[\rho] = \sum_{a\in\mathcal{A}} \sum_{\nu} \gamma(\nu) \left(A^a_\nu \rho (A^a_\nu)^\dagger - \frac{1}{2}\left\{ (A^a_\nu)^\dagger A^a_\nu, \rho \right\} \right)$$

Where $A^a_\nu = \sum_k \Pi_{k+\nu} A^a \Pi_k$

jumping between the Hamiltonian eigenbasis

Detailed balance of Davies generators

The Davies Generator defines a continuous-time dynamics

$$\frac{d}{dt}\rho = \mathcal{L}[\rho] \Rightarrow \rho(t) = e^{\mathcal{L}t}[\rho_0]$$

Under modest constraints, the DG satisfies detailed balance

$$\Rightarrow \mathcal{L}[\rho_\beta] = 0$$

That is, it converges to the Gibbs state, but it may not converge quickly.

Convergence time of Lindbladian evolution

The mixing time is the smallest time t for which

$$\|e^{t\mathcal{L}}(\rho_1 - \rho_2)\|_1 \leq \frac{1}{2}\|\rho_1 - \rho_2\|_1, \text{ for any two states } \rho_1, \rho_2$$

Standard approach is a bound on the spectral gap,

$$t_{mix} \leq \frac{n}{\lambda(\mathcal{L})} \cdot (1+\beta).$$

However, inherently comes at a polynomial overhead in system size

(Modified) Log Sobolev Inequalities (MLSI)

$$D(\rho||\sigma) = \mathrm{Tr}[\rho(\log\rho - \log\sigma)]$$

A MLSI quantifies the rate of decay of the relative entropy

$$\mathsf{EP}_{\mathcal{L}}(\rho) \equiv \left.\frac{d}{dt}\right|_{t=0} D(e^{t\mathcal{L}}[\rho]||\rho_\beta)$$

(Modified) Log Sobolev Inequalities (MLSI)

$$\boxed{D(\rho||\sigma) = \mathrm{Tr}[\rho(\log\rho - \log\sigma)]}$$

A MLSI quantifies the rate of decay of the relative entropy

$$\mathsf{EP}_{\mathcal{L}}(\rho) \equiv \left.\frac{d}{dt}\right|_{t=0} D(e^{t\mathcal{L}}[\rho]||\rho_\beta)$$

[KT13] If there exists a constant α s.t.,

$$\boxed{\forall\rho : \mathsf{EP}_{\mathcal{L}}(\rho) \leq -\alpha \cdot D(\rho||\rho_\beta)}$$

(Modified) Log Sobolev Inequalities (MLSI)

$$\boxed{D(\rho||\sigma) = \mathrm{Tr}[\rho(\log\rho - \log\sigma)]}$$

A MLSI quantifies the rate of decay of the relative entropy

$$\mathsf{EP}_{\mathcal{L}}(\rho) \equiv \frac{d}{dt}\bigg|_{t=0} D(e^{t\mathcal{L}}[\rho]||\rho_\beta)$$

[KT13] If there exists a constant α s.t.,

$$\boxed{\forall\rho : \mathsf{EP}_{\mathcal{L}}(\rho) \leq -\alpha \cdot D(\rho||\rho_\beta)} \Rightarrow D(e^{t\mathcal{L}}[\rho]||\rho_\beta) \leq O(n) \cdot e^{-\alpha \cdot t}$$

(Modified) Log Sobolev Inequalities (MLSI)

A MLSI quantifies the rate of decay of the relative entropy

$$\mathsf{EP}_{\mathcal{L}}(\rho) \equiv \frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}}[\rho]||\rho_\beta)$$

[KT13] If there exists a constant α s.t.,

$$\boxed{\forall\rho : \mathsf{EP}_{\mathcal{L}}(\rho) \leq -\alpha \cdot D(\rho||\rho_\beta)} \quad \Rightarrow D(e^{t\mathcal{L}}[\rho]||\rho_\beta) \leq O(n) \cdot e^{-\alpha\cdot t}$$

Then, by Pinsker's inequality, $t_{mix}(\mathcal{L}) = O(\alpha^{-1}\log n)$ Rapid Mixing

Next: Modified Log Sobolev Inequality for our Hamiltonians

idea: prove this for the trivial Hamiltonian, then "inherit" this to our Hamiltonians using lightcone arguments

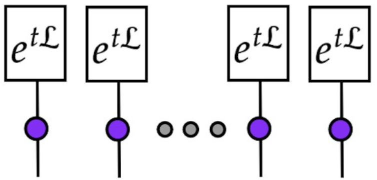
Step 1: The non-interacting system

Let's first consider the trivial Hamiltonian

$$H_{\mathsf{NI}} = -\sum_i Z_i \text{ and } \sigma_\beta \propto e^{\beta Z_1} \otimes e^{\beta Z_2} \otimes \cdots e^{\beta Z_n}$$

Jump operators are just single qubit Paulis, and the Lindbladian is non-interacting

$$\mathcal{L}_{\mathsf{NI}} = \sum_{i \in [n]} \mathcal{L}^i_{single} \otimes \mathbb{I}_{[n]\setminus i}$$



Claim $\mathcal{L}_{\mathsf{NI}}$ satisfies a MLSI with constant $\alpha_{\mathsf{NI}} = \Omega(e^{-2\beta})$

Step 2: “Inherit” the mixing time

Goal: Can we inherit the fast mixing of the non-interacting case?

Idea: Our Hamiltonian is just a rotation of the trivial Hamiltonian;
The Lindbladian is quite complicated, but we can look at it in a rotated basis

Step 2: "Inherit" the mixing time

Goal: Can we inherit the fast mixing of the non-interacting case?

Idea: In a rotated basis, our Lindbladian is a convex combination of D.G.s:

$$C^\dagger \mathcal{L}[C \cdot C^\dagger]C = q \cdot \mathcal{L}_{\mathsf{NI}}[\cdot] + (1-q) \cdot \mathcal{L}_{rest}[\cdot],$$

which satisfy detailed balance, and $q = 4^{1-\ell}$.

$\mathcal{L}_{rest}$ is quite complicated, but at least it fixes the Gibbs state

Step 2: "Inherit" the mixing time

Goal: Can we inherit the fast mixing of the non-interacting case?

Idea: In a rotated basis, our Lindbladian is a convex combination of D.G.s:

$$C^\dagger \mathcal{L}[C \cdot C^\dagger]C = q \cdot \mathcal{L}_{\mathsf{NI}}[\cdot] + (1-q) \cdot \mathcal{L}_{rest}[\cdot],$$

which satisfy detailed balance, and $q = 4^{1-\ell}$.

Claim: This gives us a MLSI for $\mathcal{L}$, with constant $\Omega(4^{-\ell} e^{-2\beta})$

Finally: How do we prove the convex combination?

Goal $$C^\dagger \mathcal{L}[C \cdot C^\dagger]C = q \cdot \mathcal{L}_{\mathsf{NI}}[\cdot] + (1-q) \cdot \mathcal{L}_{rest}[\cdot],$$

Proof by picture

$$\mathcal{L}[\rho] = \sum_{a \in \mathcal{A}} \sum_{\nu} \gamma(\nu) \left(A_\nu^a \rho (A_\nu^a)^\dagger - \frac{1}{2} \left\{ (A_\nu^a)^\dagger A_\nu^a, \rho \right\} \right)$$

Goal $$C^\dagger \mathcal{L}[C \cdot C^\dagger] C = q \cdot \mathcal{L}_{\mathsf{NI}}[\cdot] + (1-q) \cdot \mathcal{L}_{rest}[\cdot],$$

Proof by picture

$$\mathcal{L}[\rho] = \sum_{a\in\mathcal{A}}\sum_{\nu}\gamma(\nu)\left(A^a_\nu\rho(A^a_\nu)^\dagger - \frac{1}{2}\left\{(A^a_\nu)^\dagger A^a_\nu, \rho\right\}\right)$$

Goal $$C^\dagger\mathcal{L}[C\cdot C^\dagger]C = q\cdot\mathcal{L}_{\mathsf{NI}}[\cdot] + (1-q)\cdot\mathcal{L}_{rest}[\cdot],$$

Suffices to look at the jump operators in the Hamiltonian eigenbasis

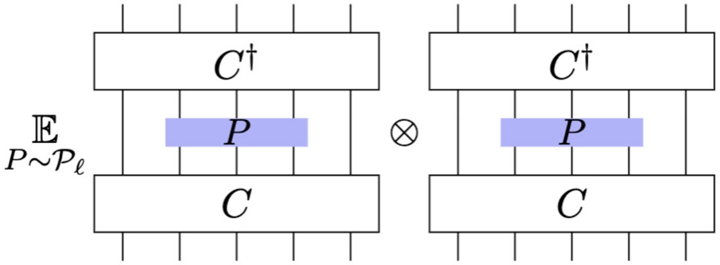
$$\mathbb{E}_a A^a_\nu \otimes (A^a_\nu)^\dagger,\ \text{for different } \nu\in[-n,n] \qquad A^a_\nu = \sum_k \Pi_{k+\nu}A^a\Pi_k$$

Proof by picture

Goal $$C^\dagger \mathcal{L}[C \cdot C^\dagger] C = q \cdot \mathcal{L}_{\mathsf{NI}}[\cdot] + (1-q) \cdot \mathcal{L}_{rest}[\cdot],$$

Suffices to look at the jump operators in the Hamiltonian eigenbasis

$$\mathbb{E}_a A^a_\nu \otimes (A^a_\nu)^\dagger, \text{ for different } \nu \in [-n, n]$$



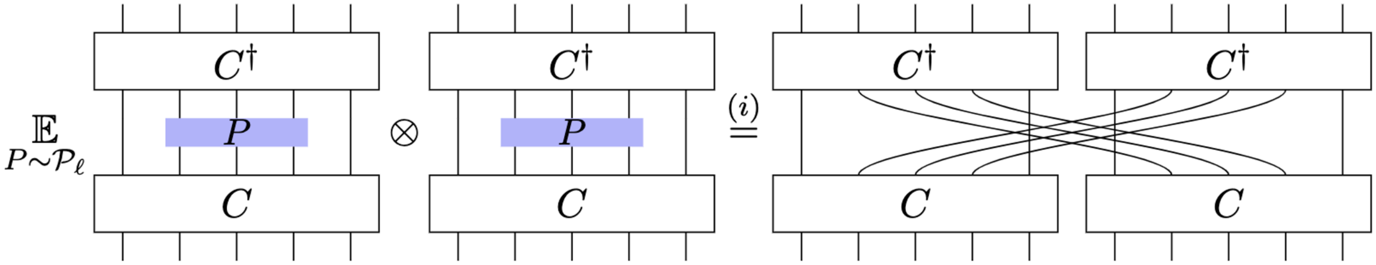
Proof by picture

$\mathbb{E}_P P \otimes P \propto \text{SWAP}$

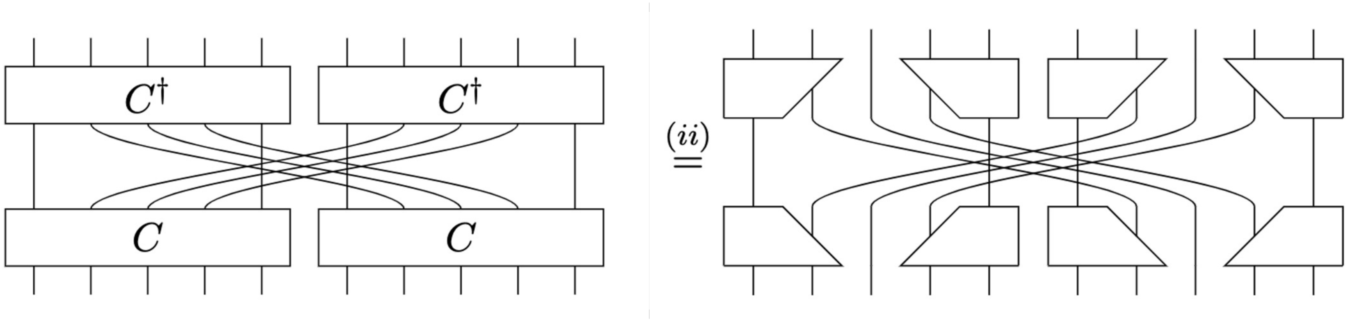
Goal $$C^\dagger \mathcal{L}[C \cdot C^\dagger] C = q \cdot \mathcal{L}_{\mathsf{NI}}[\cdot] + (1-q) \cdot \mathcal{L}_{rest}[\cdot],$$

Suffices to look at the jump operators in the Hamiltonian eigenbasis

$$\mathbb{E}_a A_\nu^a \otimes (A_\nu^a)^\dagger, \text{ for different } \nu \in [-n, n]$$

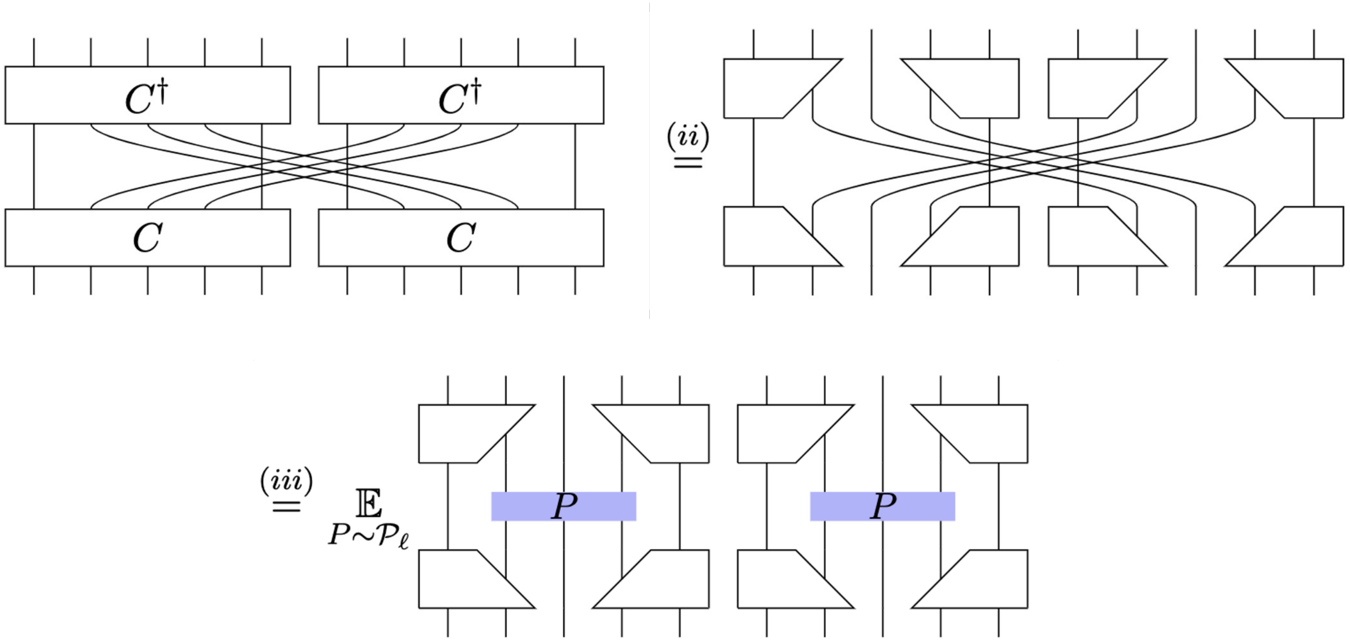


Proof by picture



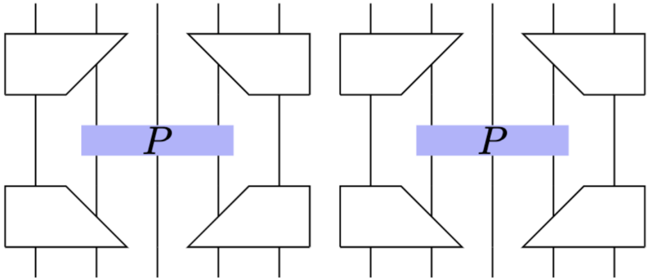
Proof by picture

$\mathbb{E}_P P \otimes P \propto \text{SWAP}$



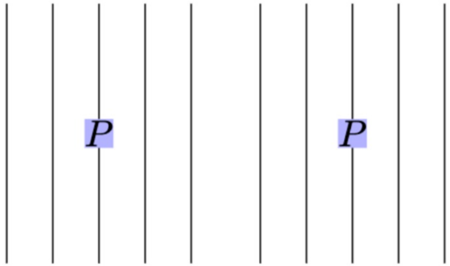
Proof by picture

$$\mathop{\mathbb{E}}_{P \sim \mathcal{P}_\ell}$$

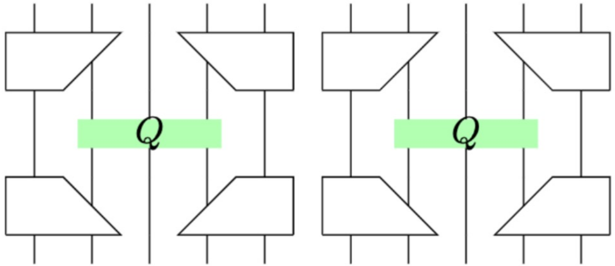


$$\overset{(iv)}{=}$$

$$\frac{1}{4^{\ell-1}} \cdot \mathop{\mathbb{E}}_{P \sim \mathcal{P}_i}$$



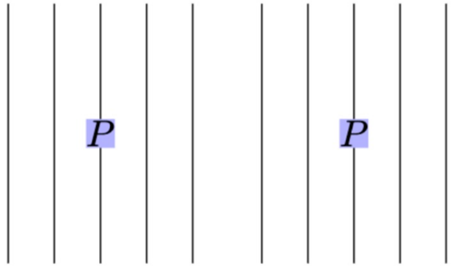
$$+ \frac{4^{\ell-1}-1}{4^{\ell-1}} \cdot \mathop{\mathbb{E}}_{Q \sim \mathcal{P}_\ell \setminus \mathcal{P}_i}$$



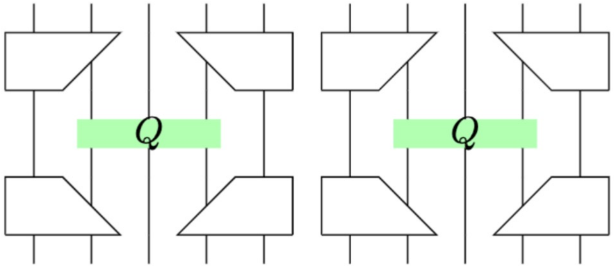
Proof by picture

In plain English, the second moment of the jump operators is a convex combination of two sets of jump operators

$\frac{1}{4^{\ell-1}} \cdot \mathbb{E}_{P\sim\mathcal{P}_i}$



$+ \frac{4^{\ell-1}-1}{4^{\ell-1}} \cdot \mathbb{E}_{Q\sim\mathcal{P}_\ell\setminus\mathcal{P}_i}$



non-interacting system

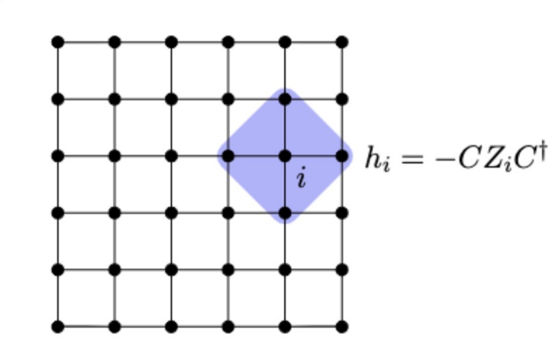
rest

Outline

I. Efficient Gibbs sampling

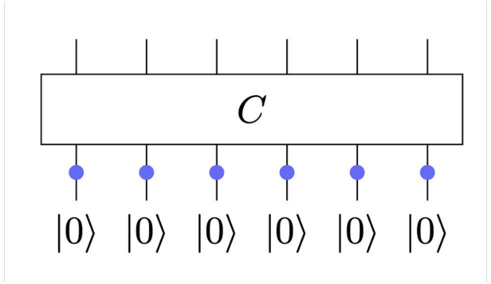
Rapid mixing bounds for Lindbladians,
via light-cone arguments

Key idea: “Inherit” the mixing time from trivial system



II. Fault tolerance of IQP circuits

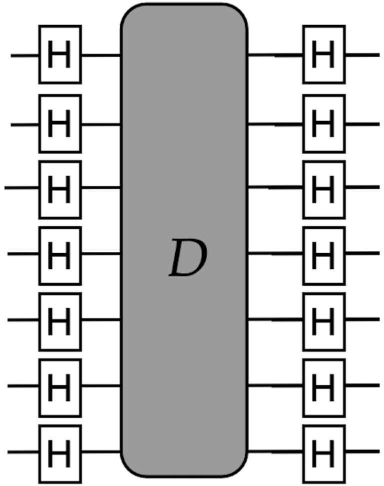
Designing fault-tolerant circuits which are
hard-to-sample from under input noise



Our approach

Goal: Construct a low-depth quantum circuit, hard to sample from under input noise

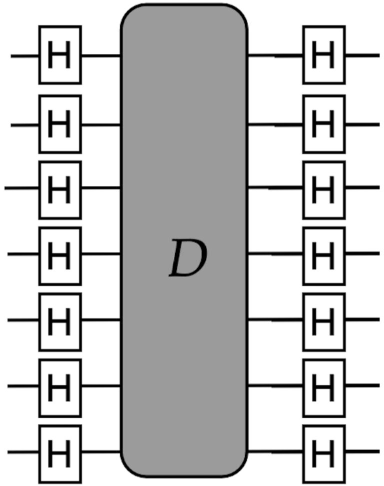
Idea: Start from shallow IQP circuits, and then make them fault tolerant



Hardness of shallow IQP circuits

Goal: Construct a low-depth quantum circuit, hard to sample from under input noise

Idea: Start from shallow IQP circuits, and then make them fault tolerant



D: diagonal gates (Z, CZ and T gates), all commuting

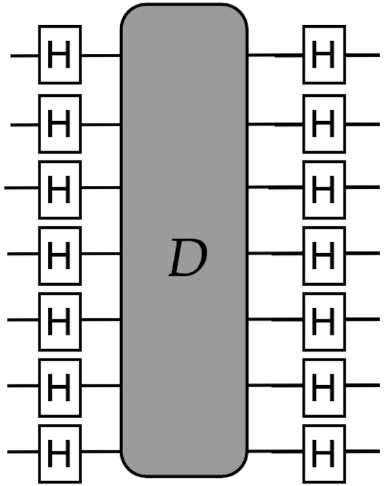
Hardness of ideal IQP circuits: [GWD16, BHS+16] There is a family of shallow IQP circuits which is hard to sample within constant TVD, assuming the average-case hardness of computing certain partition functions

Becomes classically simulable under noise [BMS16]

Hardness of shallow IQP circuits

Goal: Construct a low-depth quantum circuit, hard to sample from under input noise

Idea: Start from shallow IQP circuits, and then make them fault tolerant

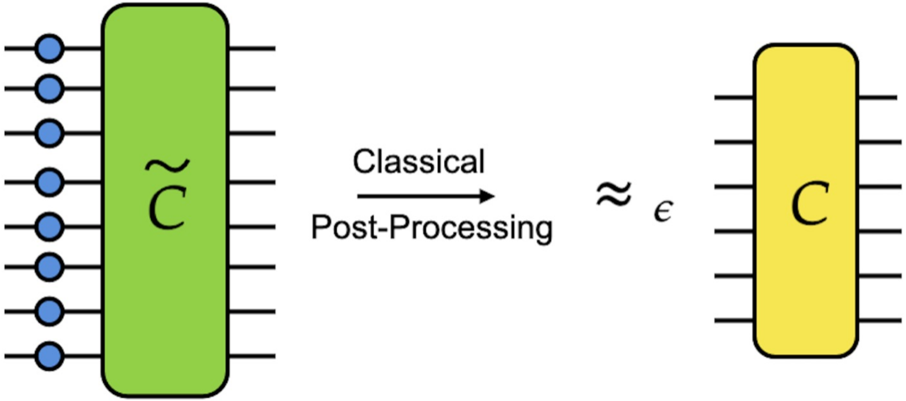


D: diagonal gates (Z, CZ and T gates), all commuting

Want hardness? Need fault tolerance!

Fault tolerance for shallow IQP circuits

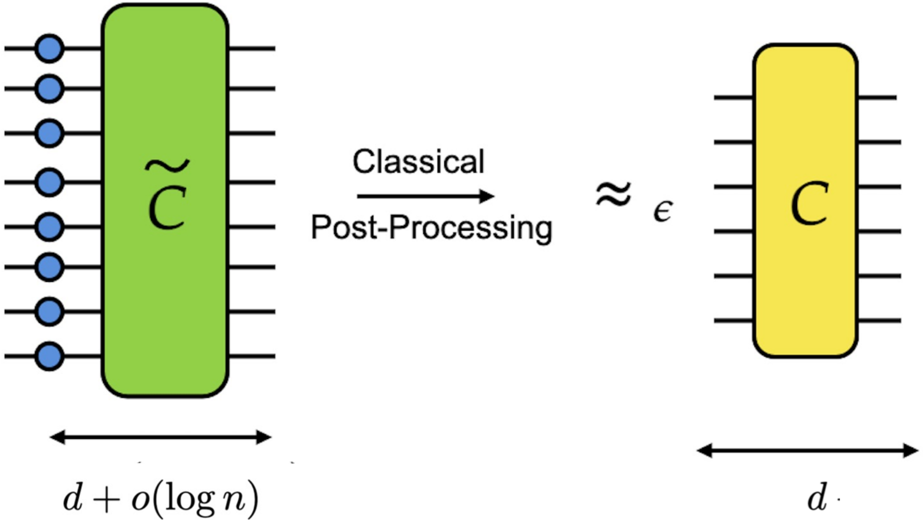
Lemma Fix a noise rate $p < 1$. Every IQP circuit can be encoded into a *slightly* bigger circuit, s.t. the new circuit is robust to input noise



Fault tolerance for shallow IQP circuits

Lemma Fix a noise rate p < 1. Every IQP circuit can be encoded into a *slightly* bigger circuit, s.t. the new circuit is robust to input noise

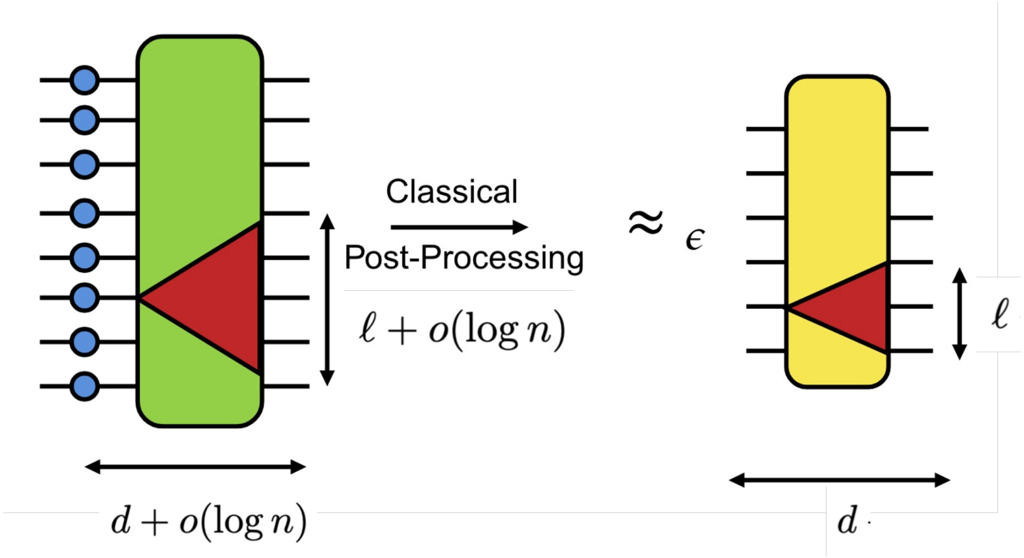
The depth,



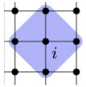
Fault tolerance for shallow IQP circuits

Lemma Fix a noise rate p < 1. Every IQP circuit can be encoded into a *slightly* bigger circuit, s.t. the new circuit is robust to input noise

The depth, the lightcone size,



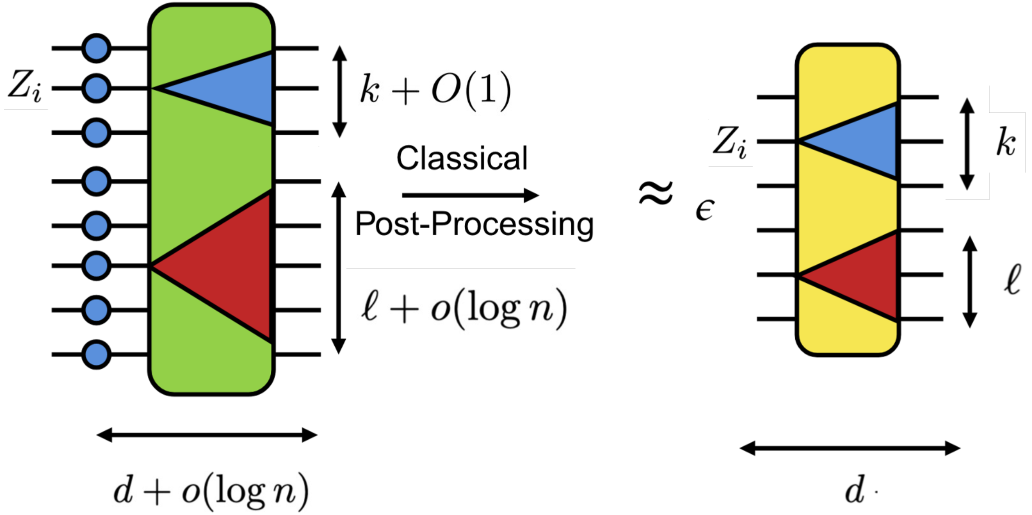
$h_i = -CZ_iC^\dagger$



Fault tolerance for shallow IQP circuits

Lemma Fix a noise rate p < 1. Every IQP circuit can be encoded into a *slightly* bigger circuit, s.t. the new circuit is robust to input noise

The depth, the lightcone size, and the Hamiltonian locality are slightly increased



We have a local Hamiltonian, because the Z-propagation in the encoded circuit is local (Joel Rajakumar & James Watson)

How to deal with input noise?

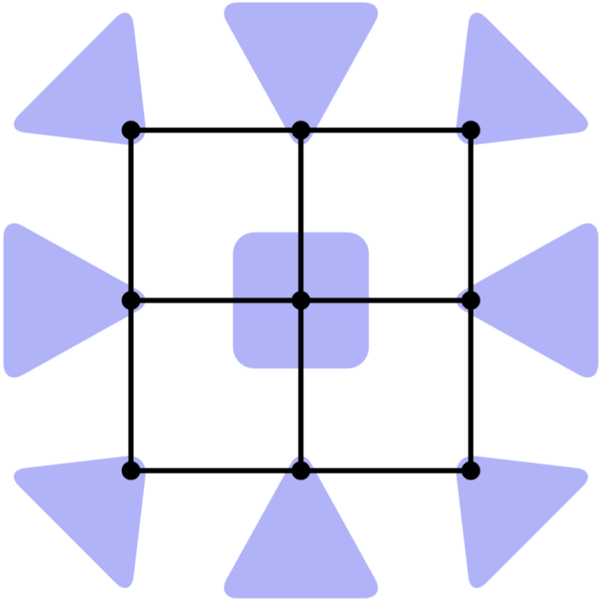
Basic idea: we take a bunch of noisy 0's, compute the majority, get a less noisy 0.

Fault tolerance construction

All we need to do is distill clean input qubits.

1. Place a gadget on each of n input qubits.

2. Input one "root" qubit per gadget into C.



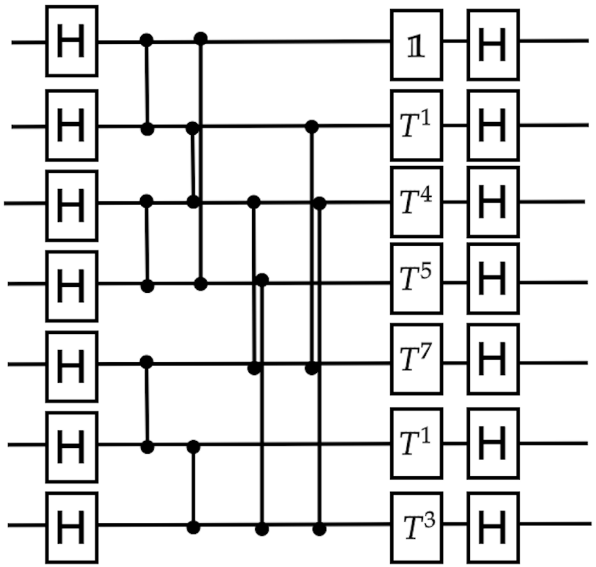
To achieve this at low overhead...

Idea 1 Suffices to detect the input error, instead of correcting it

Idea 2 Recursive state distillation

Fault tolerance of IQP circuits (against input noise)

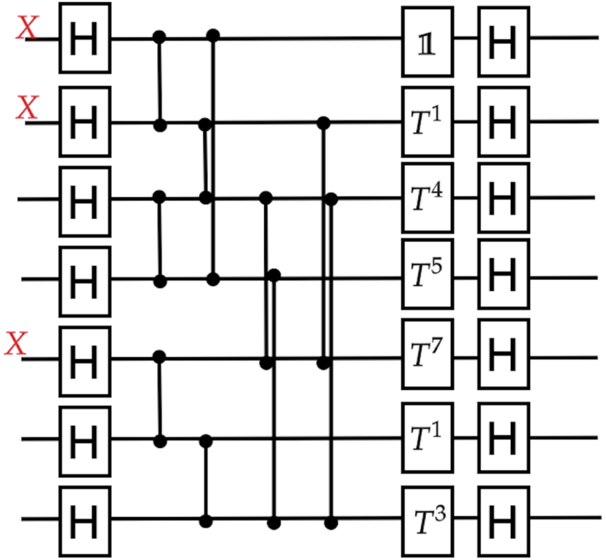
Idea 1. It suffices to detect the input error, instead of correcting it.



Fault tolerance of IQP circuits (against input noise)

Idea 1. It suffices to detect the input error, instead of correcting it.

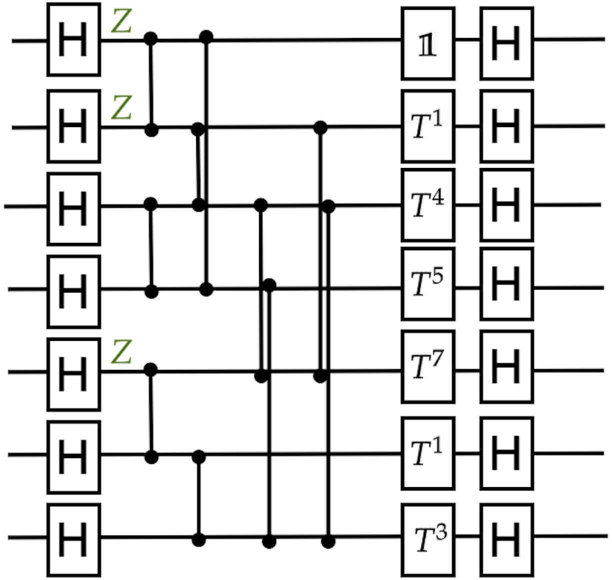
Suppose at the end of the computation, we knew that initially there were X errors on qubits 1, 2 and 5 ...



Fault tolerance of IQP circuits (against input noise)

Idea 1. It suffices to detect the input error, instead of correcting it.

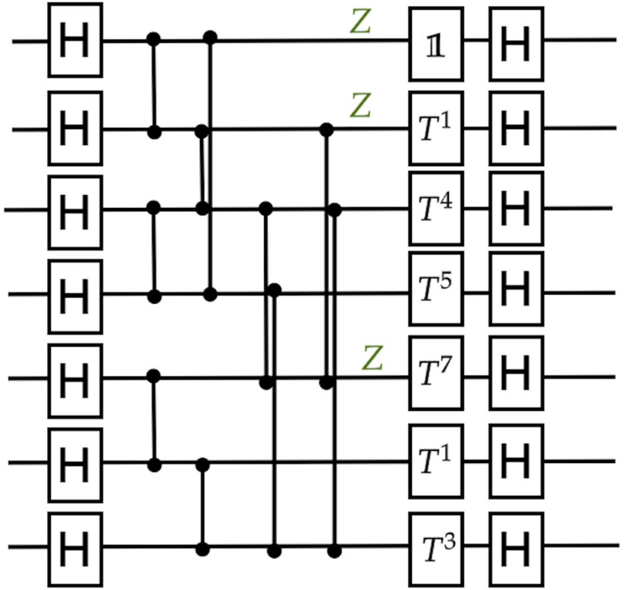
Suppose at the end of the computation, we knew that initially there were X errors on qubits 1, 2 and 5 ...



Fault tolerance of IQP circuits (against input noise)

Idea 1. It suffices to detect the input error, instead of correcting it.

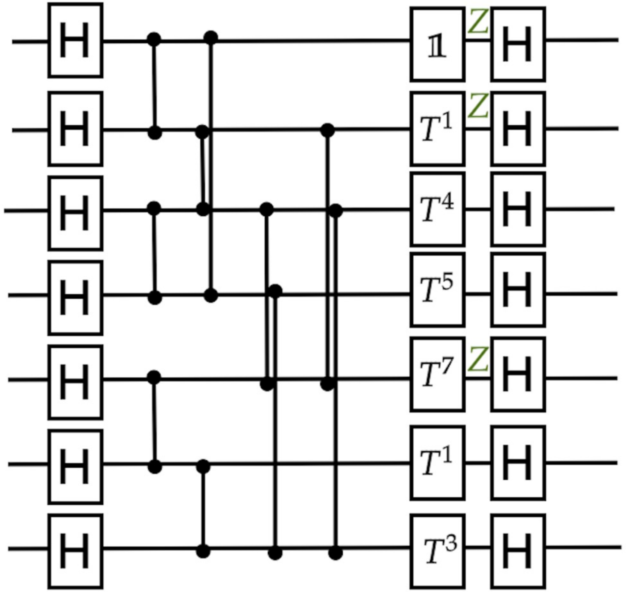
Suppose at the end of the computation, we knew that initially there were X errors on qubits 1, 2 and 5 ...



Fault tolerance of IQP circuits (against input noise)

Idea 1. It suffices to detect the input error, instead of correcting it.

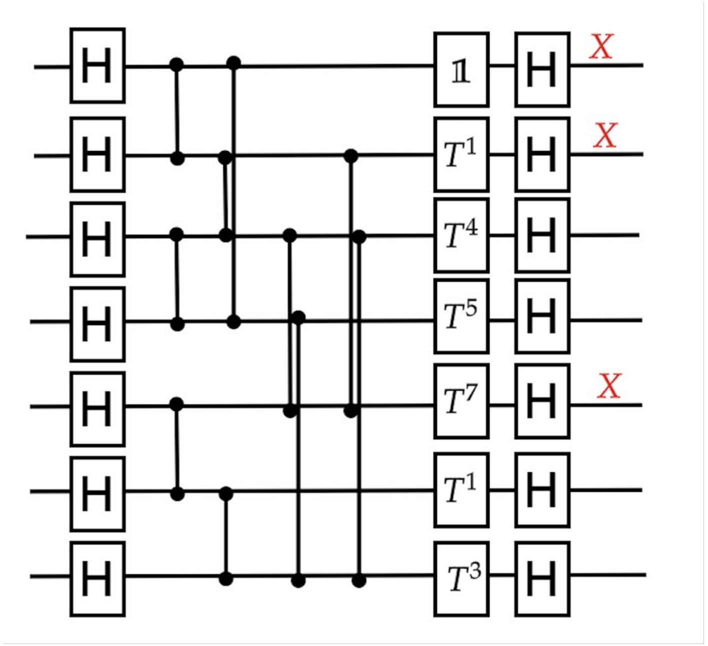
Suppose at the end of the computation, we knew that initially there were X errors on qubits 1, 2 and 5 ...



Fault tolerance of IQP circuits (against input noise)

Idea 1. It suffices to detect the input error, instead of correcting it.

Equivalent to bit-flip errors on the output string, which we can correct classically



Fault tolerance construction

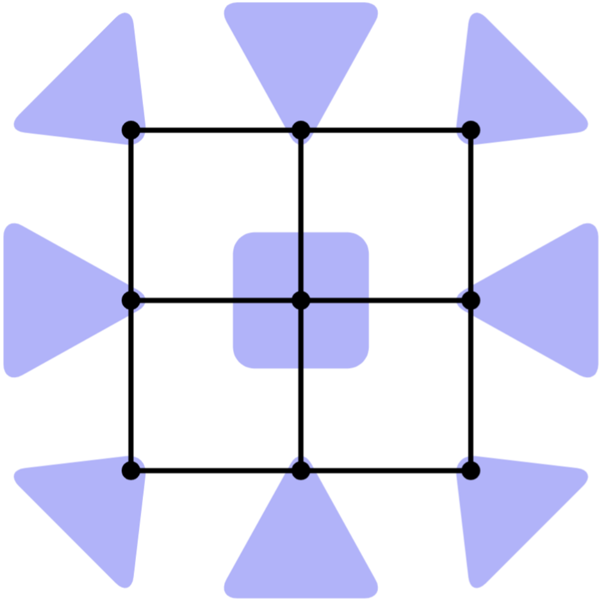
All we need to do is distill clean input qubits.

Deferring the decoding overhead into classical postprocessing

To achieve this at low overhead...

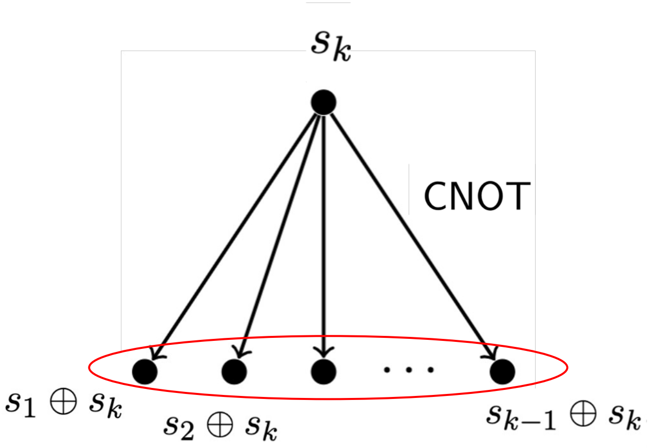
Idea 1 Suffices to detect the input error, instead of correcting it

Idea 2 Recursive state distillation



Next: how to detect, and how to further reduce overhead using recursion

Error detection

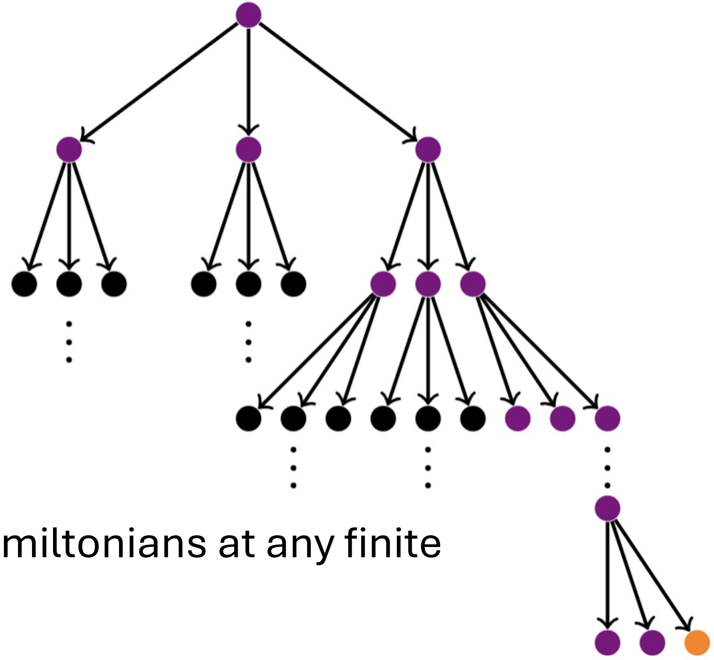


- Every black dot is a noisy bit (ideally 0, could be flipped to 1)
- Apply CNOT from root to every leaf
- Majority of leaves equals root whp

- Compute Majority at the end of the computation
- If there was an error on the root, it is propagated to the end and corrected classically

Recursion

- Use Majority of Majority of Majority...
- Causal influence only travel upwards
- Reduces the lightcone blowup of the construction



Theorem. There exists a family of n qubit, $O(1)$ local Hamiltonians at any finite temperature β, which is

- Rapidly thermalizing in time $n^{o(1)}$
- Classically intractable under standard complexity-theoretic assumptions

Future directions

- Noise robustness?
    - We can handle measurement noise with a larger blowup
    - The hope is that the Gibbs state is already a natural “noise model”

- Complexity of temperature
- Can we embed more general quantum computation into a constant temperature Gibbs state?
    - Resource state for universal MBQC
    - Universal quantum computation by directly sampling from Gibbs states?

Most important open question in NISQ/Early-FT

- Quantum computational advantage with noisy shallow circuits:

- Is there a family of constant (or $O(\log n)$) depth circuits which is classically hard-to-sample from (within 1/poly TVD) under depolarizing noise?

Thanks!