2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS) | 979-8-3315-1674-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/FOCS61266.2024.00071

Quantum computational advantage with constant-temperature Gibbs sampling

Thiago Bergamaschi
Department of EECS
University of California, Berkeley
Berkeley, CA, USA
thiagob@berkeley.edu

Chi-Fang Chen
Department of Physics
California Institute of Technology
Pasadena, CA, USA
chifang@caltech.edu

Yunchao Liu
Department of EECS
University of California, Berkeley
Berkeley, CA, USA
yunchaoliu@berkeley.edu

***Abstract*—A quantum system coupled to a bath at some fixed, finite temperature converges to its Gibbs state. This *thermalization* process defines a natural, physically-motivated model of quantum computation. However, whether quantum computational advantage can be achieved within this realistic physical setup has remained open, due to the challenge of finding systems that thermalize quickly, but are classically intractable. Here we consider sampling from the measurement outcome distribution of quantum Gibbs states at constant temperatures, and prove that this task demonstrates quantum computational advantage. We design a family of commuting local Hamiltonians (parent Hamiltonians of shallow quantum circuits) and prove that they rapidly converge to their Gibbs states under the standard physical model of thermalization (as a continuous-time quantum Markov chain). On the other hand, we show that no polynomial time classical algorithm can sample from the measurement outcome distribution by reducing to the classical hardness of sampling from noiseless shallow quantum circuits. The key step in the reduction is constructing a fault-tolerance scheme for shallow IQP circuits against input noise.**

***Index Terms*—thermalization, quantum computational advantage, Gibbs sampling, commuting local Hamiltonians**

I. Introduction

A major goal of today's quantum computing efforts is to realize quantum computational advantage in realistic physical setups. One such setup is open system thermalization, where a quantum many-body system is specified by a Hamiltonian H and then coupled to a bath at finite (constant) temperature β, and the system converges to the Gibbs state $\rho_\beta \propto e^{-\beta H}$. Under physical assumptions,[1] this thermalization process can be described by a *thermal Lindbladian* (a continuous-time quantum Markov chain), most notably the Davies generator [1] and its variants (e.g. [2]). This setup is especially relevant for physical platforms in which implementing digital quantum circuits is difficult. However, there has been no complexity-theoretic evidence showing that quantum computational advantage can be achieved in this model (see Section I-A for a discussion).

In this paper, we provide such evidence by showing that quantum computational advantage can be achieved for the task of sampling from the measurement outcome distribution of Gibbs states at constant temperatures. In particular, we construct a family of commuting local Hamiltonians and show that its thermalization process (described by the Davies generator) is rapidly mixing. Meanwhile, its Gibbs state is classically intractable to sample from.

Theorem I.1 (Main result)**.** *For any constant inverse-temperature $\beta = \Theta(1)$, there exists a family of n-qubit commuting $O(1)$-local Hamiltonians, such that the n-qubit Gibbs state ρ_β is both*

1) Rapidly Thermalizing. *It can be prepared within small trace distance by the Davies generator (a quantum Markov chain describing thermalization), in time $n^{o(1)}$. In addition, this process can be simulated on a quantum computer in time $n^{1+o(1)}$. And yet,*
2) Classically Intractable. *Under certain complexity-theoretic assumptions, there is no polynomial time classical algorithm to sample from the measurement outcome distribution $p(x) = \langle x| \rho_\beta |x\rangle$ within small total variation distance.*

The classical hardness is based on the hardness of approximate sampling from the output distribution of ideal shallow quantum circuits. The main result, therefore, places the hardness of rapidly mixing thermalization to the same level as ideal sampling-based quantum supremacy experiments (see Section VII for more details).

A more general version of the result (Theorem IX.1) is given in Section IX, where we generalize the above and show how to trade-off locality for mixing time, including a family of $O(\log\log n)$-local Hamiltonians which thermalizes in $\mathrm{polylog}(n)$ time.[2]

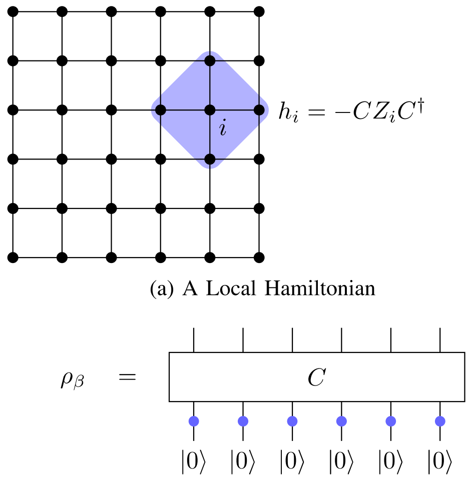
a) Our approach: The family of Hamiltonians we consider is the class of "parent" Hamiltonians of shallow quantum circuits (Fig. 1a). Starting from a trivial, non-interacting Hamiltonian $H_{\mathsf{NI}} = -\sum_i Z_i$ consisting of single-qubit Pauli-Z terms, we consider the family of Hamiltonians that are related to H_{NI} by a low depth circuit,

$$\mathscr{H} = \left\{ H : \exists \text{ low-depth circuit } C,\ H = C H_{\mathsf{NI}} C^\dagger \right\}. \quad (1)$$

[1] The bath is Markovian and the coupling is weak.

[2] In the initial posting of this work, we stated this latter construction as our main result. We thank James Watson and Joel Rajakumar for the observation that under appropriate parameter choices, our construction in fact has constant locality (see Section II-B2 and [3]).

2575-8454/24/$31.00 ©2024 IEEE
DOI 10.1109/FOCS61266.2024.00071


$h_i = -CZ_iC^\dagger$
i
$\rho_\beta \;=$
C
$|0\rangle$ $|0\rangle$ $|0\rangle$ $|0\rangle$ $|0\rangle$ $|0\rangle$

(a) A Local Hamiltonian

(b) The Noise Model

Fig. 1: (a) We consider local Hamiltonians $H = \sum_i h_i$ which are parent Hamiltonians of shallow quantum circuits. (b) The Gibbs states of these Hamiltonians $\rho_\beta \propto e^{-\beta H}$ are equivalent to the output state of C, where the input qubits are subject to bit-flip errors (blue dots) of rate $(1 + e^{2\beta})^{-1}$.

Each $H \in \mathscr{H}$ is local, commuting, and it encodes the computation C in the sense that its ground state is the output of the circuit $C\,|0^n\rangle$. The reason that these Hamiltonians are good candidates for quantum advantage at constant temperatures lies in the following key observation:

> *The Gibbs state of each $H \in \mathscr{H}$ is a noisy version of the underlying computation, where random bit-flip errors are applied to the input qubits (Fig. 1b).*

This is a clean example of the general intuition that constant-temperature Gibbs states are very noisy and far from ground states. To encode computational hardness into the Gibbs states of $H \in \mathscr{H}$, it then suffices to design a shallow quantum circuit which is classically intractable to simulate even under input noise. Our main result then follows from two key technical contributions:

1) **A construction of classically-hard shallow quantum circuits that are fault-tolerant against input noise.** Standard techniques in quantum fault-tolerance blow up the circuit depth, and in turn, the locality of the parent Hamiltonian[3]. We start from a specific family of classically-hard shallow circuits (namely, IQP circuits [4], [5]), and then design a low-overhead fault-tolerance scheme tailored to IQP circuits and the input noise model.
2) **A proof that these Hamiltonians thermalize rapidly, via a modified log-Sobolev inequality.** We prove a rapid mixing bound for Hamiltonians in $\mathscr{H}$ which leverages the structure of the thermal Lindbladian (the quantum Markov chain describing thermalization), in combination with a carefully constructed lightcone argument for shallow quantum circuits.

A. *Related work*

a) Complexity of Gibbs states: Establishing quantum computational advantage with constant-temperature Gibbs sampling faces inherent difficulties. After all, at high enough temperatures, Gibbs states are expected to be essentially classical objects; in particular, sampling from these Gibbs states is efficient to simulate on a classical computer[4] [6], [7]. On the other hand, in the low temperature regime, preparing Gibbs states is expected to be hard in general even for a quantum computer;[5] in particular, the thermalization process may take exponential time.

Nevertheless, a path exists to circumvent these issues, by embedding some classically hard quantum computation into a local Hamiltonian. It is reasonable to hope that the nature of this embedding ensures that producing the Gibbs state is still tractable for quantum computers[6] (e.g. [9], [10]), and one can further hope that the Gibbs state is classically hard. But there is yet another issue: standard means to embed quantum circuits into Hamiltonians [11] typically encode the quantum computation into its ground state. However, Gibbs states at constant temperatures are understood to be very noisy, and far from the ground state. In this manner, to argue that this noisy version of the ground state remains classically hard, there must be an inherent *fault-tolerance* to the circuit-to-Hamiltonian mapping. Our approach can be viewed as a clean example that satisfies all of the above criteria.

b) Gibbs samplers and rapid mixing: Preparing Gibbs states (or Gibbs sampling) is a candidate application of quantum computers as well as an important quantum algorithmic primitive. While there are many proposed quantum Gibbs samplers, recent developments have focused on an approach of simulating open system (Lindbladian) dynamics, in particular the Davies generator and its variants which mimic thermalization in nature [12], [13], [14].

The key missing ingredient to the efficiency of these quantum simulation algorithms is a bound on the mixing time of the underlying quantum Markov chain. The standard approach, via a bound on the spectral gap, gives a mixing time that has intrinsic polynomial dependence in n [15]. A much stronger approach known as (quantum) *log-Sobolev inequalities* consists of a decay of the relative entropy, and results in only $\mathrm{polylog}(n)$ mixing time, a phenomenon known as *rapid mixing*. These stronger inequalities are notoriously

[3] Our interest in decreasing the locality stems both from the practical challenges behind engineering systems with many-body interactions, and a complexity-theoretic understanding of the role of locality in the hardness of Gibbs sampling.

[4] This does not contradict our result which holds for arbitrary constant temperature, due to the order of quantifiers; see Remark IX.1.

[5] Indeed, NP-hard due to the classical PCP theorem [8].

[6] In fact, we desire something even stronger: that the Hamiltonian is rapidly thermalizing.

hard to prove: examples have only been shown for certain commuting systems, in 1D [16], [17] or on lattices above a threshold temperature [18]. Our rapid mixing bound uses the lightcone structure of shallow quantum circuits, and does not require geometric locality or a temperature threshold.

c) Shallow quantum circuits and fault-tolerance: Shallow quantum circuits are widely used in quantum algorithms for near-term devices and quantum supremacy experiments. The hardness of sampling from the output distribution of shallow quantum circuits provides the complexity foundation for these experiments (see [19] for a review). We focus on constant-depth *instantaneous quantum polynomial time* (IQP) circuits $C = H^{\otimes n} D H^{\otimes n}$ where D is a constant-depth diagonal unitary, which provides hardness due to the universality of measurement-based quantum computation [4], [5]. However, these circuits are not noise-robust and become classically simulable under noise [20], [21]; fault-tolerance techniques are therefore necessary for classical hardness in our context.

There is a tension between shallow quantum circuits and the overhead of quantum fault-tolerance.[7] Standard techniques encode a constant depth quantum circuit into a fault-tolerant circuit of $\mathrm{polylog}(n)$ depth [24], and fault-tolerance with constant circuit depth overhead is only known for shallow Clifford circuits [25]. Ref. [20] devised a fault-tolerance scheme specialized to IQP circuits and the input noise model, and we design a new scheme in this setting which achieves a significantly smaller overhead.

B. Our Contributions

1) Efficient quantum Gibbs sampling via rapid mixing: Our first result is a quantum algorithm for preparing the Gibbs states of $H \in \mathscr{H}$, given only a description of its local terms $H = \sum_i h_i$ (as $2^\ell \times 2^\ell$ matrices).[8]

Lemma I.2 (Gibbs State Preparation). *Fix $\beta > 0$, and let $H \in \mathscr{H}$ be the parent Hamiltonian of a quantum circuit on n qubits, of depth d and lightcone size ℓ. Then, there exists a quantum algorithm which can prepare the Gibbs state of H at inverse-temperature β up to an error ε in trace distance in time $O(2^{4\ell} \cdot 2^d \cdot e^{2\beta} \cdot n \cdot \mathrm{poly}(\log \frac{n}{\varepsilon}, \ell, \beta))$.*

In general, the lightcone size ℓ is upper bounded by $\ell \leq 2^d$. We emphasize we do not make any assumptions on the temperature or geometric locality. This is important as our fault-tolerant circuits (Lemma I.3) are not naturally defined on a lattice.

The algorithm in Lemma I.2 follows from a two-step argument. The first step is the design and analysis of a particular family of Davies generators [1], a family of dissipative Lindbladians whose local jumps (or transitions) are engineered to resemble the connectivity of the Hamiltonian. In Lemma III.4, we prove that the mixing time of our Lindbladians is $t_{mix} = O(4^\ell \log n)$ via a modified log-Sobolev inequality. In principle, this step is already a *thermal algorithm*, in the sense that "placing the system in a fridge" would drive it to the Gibbs state in time $t_{mix} \cdot \log(1/\varepsilon)$.

The second step is the simulation of the dissipative (non-unitary) dynamics on a quantum computer. We employ the block-encoding framework of [13] which we significantly simplify as our family of Hamiltonians is commuting and has integer spectra. The quantum simulation adds a factor of n to the running time, which may be hard to improve due to the absence of geometric locality. In Section V, we discuss an alternative method for Gibbs state preparation assuming finite dimensional lattice geometry using the framework of [26].

2) Fault-tolerance of shallow IQP circuits: The key ingredient for the classical hardness of sampling from quantum Gibbs states is to produce a shallow quantum circuit which is hard to sample from even under input noise. For this purpose, we design a fault-tolerance scheme for shallow IQP circuits [4], [5] since their gate set works nicely with fault tolerance techniques. Our result ensures that any IQP circuit can be made robust to input noise with only a small additive blow-up to the circuit depth (see Lemma VIII.1 for a more general statement).

Lemma I.3. *Let $p < \frac{1}{2}$ be a constant bit-flip error rate, and let C be an n qubit IQP circuit of depth d. Then, there exists an $O(n \log \frac{n}{\varepsilon})$ qubit circuit $\tilde{C}$ of depth $d + o(\log \frac{n}{\varepsilon})$, such that a sample from $\tilde{C}$ under input noise (Fig. 1b) can be efficiently post-processed into a sample within ε total variation distance to the output distribution of C.*

This result significantly reduces the blow-up in circuit depth compared to a prior fault-tolerance scheme of [20]. Moreover, the locality of the resulting parent Hamiltonian $H = -\sum_i \tilde{C} Z_i \tilde{C}^\dagger$ is only a constant. Our key idea is a non-adaptive state distillation scheme, drawing inspiration from magic state distillation [27]: distilling a near-perfect initial state from noisy initial states, up to a known but uncorrected Pauli error. The error is propagated through the circuit and corrected in post-processing, similar to [25]. Propagating Pauli errors through non-Clifford circuits is hard in general, but here it works thanks to the structure of IQP circuits.

3) Applications:

a) BQP Completeness under adaptive single-qubit measurements: In addition to quantum advantage, using our techniques we can show that constant-temperature Gibbs states do have some inherent form of universality for quantum computation. In Section X we prove that there exist local Hamiltonians whose Gibbs states are universal resource states for quantum computation, in the sense that they can be used for universal measurement-based quantum computation.

Theorem I.4. *Fix an inverse-temperature $\beta = \Theta(1)$. Then, there exists an n-qubit, $O(1)$-local commuting Hamiltonian, whose Gibbs state at inverse-temperature β is a universal*

[7] Note that some models of fault-tolerance assume instant classical computation and feedforward within a quantum circuit [22], [23]. This is not allowed in our setting: all operations must be realized by quantum gates.

[8] Although H has a simple structure by definition, the underlying global structure (the low-depth circuit C) is hidden among the local terms, and is not directly accessible. See Remark IX.2 for a discussion.

resource state for quantum computation and is efficiently preparable on a quantum computer.

Theorem I.4 is based on the universality of cluster-states for measurement-based quantum computation. That is to say, any quantum computation of bounded size can be implemented using adaptive single-qubit measurements on top of a fixed 2D cluster-state (e.g. [28]). We design a Hamiltonian whose Gibbs state resembles a noisy version of a cluster-state, such that under adaptive single-qubit operations, one can nevertheless correct and distill out computation.

b) Gibbs sampling under measurement errors: An interesting question is whether the thermal quantum advantage demonstrated in this paper, is itself robust to noise. That is, in realistic physical platforms, we expect imperfect state preparation, noisy system-bath couplings, and erroneous measurements. As a starting point to this problem, we consider a model where the Gibbs state preparation is ideal, but there are random bit-flip errors in the measurement outcome.

We show that the quantum advantage survives in this model, albeit at a higher locality.

Theorem I.5. *Fix an inverse temperature $\beta = \Theta(1)$, and a measurement error rate $p < \frac{1}{2}$. There exists a family of n-qubit, $O(\log n)$-local Hamiltonians, such that sampling from their Gibbs state at inverse-temperature β, under measurement errors of rate p, is classically intractable under certain complexity-theoretic assumptions. Moreover, there exists a $\mathrm{poly}(n)$ time quantum algorithm to produce said Gibbs state.*

The Hamiltonians of Theorem I.5 are similar to that of Theorem I.1, in the sense that they are parent Hamiltonians of fault-tolerant IQP circuits. However, to ensure classical hardness under measurement errors, our quantum circuits now need to be fault-tolerant against both input and output errors. (Recall that the "input errors" come from temperature, while "output errors" come from actual physical noise in measurements.) To do so, in Section XI we appeal to an optimized construction of a prior fault-tolerance scheme by [20], at the cost of an increase to the locality of the Hamiltonians, which also changes the mixing time from $n^{o(1)}$ to $\mathrm{poly}(n)$.

C. Discussion

We conclude by discussing two future directions, broadly related to the complexity of Gibbs sampling. The first of which concerns the BQP Completeness of Gibbs sampling (without adaptivity).

Question 1 (BQP Completeness of Gibbs Sampling)**.** *For every n qubit, $\mathrm{poly}(n)$ depth quantum circuit C, does there exist a Hamiltonian H and a constant inverse-temperature $\beta > 0$ such that by sampling from its Gibbs state one can recover the output of the quantum computation C?*

Partial progress on this question has recently been made by [10], albeit, only at very low temperatures where the Gibbs state approximates the ground state. In particular, they showed how to embed an arbitrary quantum computation into a (modified) Feynman-Kitaev circuit-to-Hamiltonian mapping, which could be efficiently prepared by a Lindbladian evolution. Whether similar ideas could work at constant temperatures remains an open problem.

Another interesting direction lies in the time overhead for fault-tolerance, and for quantum advantage using shallow circuits which are robust to noise.

Question 2 (Quantum Advantage in Noisy Shallow Circuits)**.** *Does there exist a family of constant depth quantum circuits (using only quantum gates) which is classically hard to sample from in the presence of depolarizing noise on each gate?*

Its main motivation lies in the design of quantum advantage experiments, which can be implemented on near-term devices. Depolarizing noise on each gate of the circuit, however, is naturally a significantly more general noise model than input noise. Nevertheless, the same question with input noise remains open as well.

II. Technical Overview

In this section we give a sketch of our two main technical contributions: (1) A proof of a modified log-Sobolev inequality for a family of Davies generators, via a lightcone argument (Section II-A); and (2) A fault-tolerance scheme for shallow IQP circuits against input noise, via non-adaptive state distillation (Section II-B). We begin by presenting some basic notation and background on thermal Lindbladians.

A. Gibbs state preparation via rapid mixing

Fix a Hamiltonian $H \in \mathscr{H}$. By definition, there exists a shallow circuit C such that

$$H = \sum_{i\in[n]} h_i, \quad \text{where} \quad h_i = C\left(|1\rangle\langle 1|_i \otimes \mathbb{I}_{[n]\setminus i}\right) C^\dagger, \tag{2}$$

and each $|1\rangle\langle 1|_i$ is a single-qubit projection. Note that Eq. (2) is equivalent to Eq. (1) up to a shift. The eigenstates of H of energy $k \in [n]$ are all the states $C|x\rangle$, where $x \in \{0,1\}^n$ has Hamming weight $|x| = k$. We denote the projection Π_k onto the eigenspace of H of energy k as

$$\Pi_k = C\bigg(\sum_{|x|=k} |x\rangle\langle x|\bigg)C^\dagger. \tag{3}$$

We consider two notions of locality for C and H respectively:

- The circuit lightcone. The *lightcone* L_i of qubit i is the set of qubits that can be reached by i via gates in C, and we define the *lightcone size* as $\ell = \max_i |\mathsf{L}_i|$.
- The Hamiltonian locality. Let $\mathsf{S}_i = \mathrm{supp}(h_i) = \mathrm{supp}(CZ_iC^\dagger)$ be the set of qubits that h_i acts nontrivially on. The *locality* of the Hamiltonian H is defined as $r = \max_i |\mathsf{S}_i|$.

Note that S_i is related to the propagation of Z_i under C, and thus by definition we have $\mathsf{S}_i \subseteq \mathsf{L}_i$. In fact, $r \ll \ell$ for the family of circuits we consider.

1) Our Davies generators: We determine our family of thermal Lindbladians, or Davies generators, by specifying two ingredients: a set of jump operators, and transition weights. Technically, general thermal Lindbladians for noncommuting Hamiltonians need not take the Davies' form (see, e.g, [13], [14]), but for commuting Hamiltonians, the Davies' generator is nonetheless sufficient for all our discussions.

- **Jump Operators.** To generate the transitions, we consider the set of jump operators which are local, ℓ-qubit Pauli operators on the support of each lightcone L_i,[9]
$$\{A^a\}_{a\in\mathcal{A}} = 2^{-\ell}\cdot\left\{P_{\mathsf{L}_i}\otimes\mathbb{I}_{[n]\setminus\mathsf{L}_i} : i\in[n], P\in\mathcal{P}_\ell\right\}, \quad (4)$$
where $\mathcal{P}_\ell = \{I,X,Y,Z\}^{\otimes\ell}$.[10] In contrast to classical Markov Chain transitions, these quantum jumps will change the energy of the system in superposition. Thereby, it will be convenient to decompose the jump operators into the energy basis:
$$A^a_\nu := \sum_{k\in[n]}\Pi_{k+\nu}A^a\Pi_k \quad \text{such that} \quad \sum_{\nu\in[-n,n]} A^a_\nu = A^a. \tag{5}$$
- **Transition Weights.** The transition weight is selected to be the Glauber dynamics weight, $\gamma(\nu) = 1/(1+e^{-\beta\nu})$ for all $\nu\in[-n,n]$.

Put together, the associated family of Davies generators $\mathcal{L}$ can be written down as[11]
$$\mathcal{L}[\rho] = \sum_{a\in\mathcal{A}}\sum_{\nu}\gamma(\nu)\left(A^a_\nu\rho(A^a_\nu)^\dagger - \frac{1}{2}\left\{(A^a_\nu)^\dagger A^a_\nu,\rho\right\}\right). \tag{6}$$

This construction satisfies the *quantum detailed balance* condition, which implies that the desired Gibbs state is a fixed point $\mathcal{L}[\rho_\beta] = 0$ of the evolutions (see e.g. [29], or Fact III.1). It remains to show that the Lindblad dynamics, governed by the exponential map
$$\frac{d}{dt}\rho = \mathcal{L}[\rho] \Rightarrow \rho(t) = e^{\mathcal{L}t}[\rho_0], \tag{7}$$
converges quickly to ρ_β. This is achieved by presenting a bound on the mixing time of $\mathcal{L}$, which is the shortest time t_{mix} such that
$$\left\|e^{\mathcal{L}t_{mix}}[\rho-\sigma]\right\|_1 \le \frac{1}{2}\left\|\rho-\sigma\right\|_1, \quad \text{for all density matrices} \quad \rho,\sigma. \tag{8}$$

2) A lightcone argument for the modified log-Sobolev inequality: To study the mixing time of our algorithm, our starting point is first to study the trivial non-interacting Hamiltonian $H_{\mathsf{NI}} = \sum_{i\in[n]}|1\rangle\langle 1|_i\otimes\mathbb{I}_{[n]\setminus\{i\}}$, and prove a rapid mixing bound for the associated Davies generator $\mathcal{L}_{\mathsf{NI}}$. Subsequently, we argue that the mixing time of $\mathcal{L}$ can be *compared* with that of $\mathcal{L}_{\mathsf{NI}}$. This is achieved by leveraging the lightcone structure of shallow quantum circuits. We begin by presenting basic definitions of Log-Sobolev bounds.

a) Mixing time bounds via log-Sobolev inequalities: There are two general purpose methods to bound the mixing time of Lindbladian evolution. The first of which consists of a bound on the spectral gap of $\mathcal{L}$. Unfortunately, a spectral gap bound comes at an inherent polynomial overhead to the mixing time, see Section III. Instead, we make use of a much sharper notion of convergence known as a modified log-Sobolev inequality (MLSI) [15]. Informally, a MLSI quantifies the rate of decay of the relative entropy,[12] by relating it to the relative entropy itself:
$$\frac{d}{dt}\bigg|_{t=0} D\big(e^{t\mathcal{L}}[\rho]||\rho_\beta\big) \le -\alpha\cdot D\big(\rho||\rho_\beta\big) \qquad \text{(MLSI)}$$
for every density matrix ρ, where α is known as the MLSI constant. This clearly implies an exponential decay of $D(e^{\mathcal{L}t}[\rho]||\rho_\beta) \le e^{-\alpha t}\cdot D(\rho||\rho_\beta)$. Which, in turn, tells us the mixing time is bounded by $t_{mix}\le\alpha^{-1}\cdot O(\log n)$ (Pinsker's inequality). This logarithmic mixing time bound is known as *rapid mixing*, and proving good lower bounds on the constant α has proven to be quite challenging in the quantum setting.

b) The non-interacting Lindbladian: The simplest Hamiltonian in the family $\mathscr{H}$ is the non-interacting system H_{NI}. Its Gibbs state is the tensor product state $\sigma_\beta \propto \left(e^{-\beta|1\rangle\langle 1|}\right)^{\otimes n}$. Under our framework (described in Eq. (6)), its associated Lindbladian $\mathcal{L}_{\mathsf{NI}}$ has the same form as $\mathcal{L}$, except that the circuit C has been replaced by the identity. In this manner, $\mathcal{L}_{\mathsf{NI}}$ itself can also be written as a sum of non-interacting, single-qubit components:
$$\mathcal{L}_{\mathsf{NI}} = \sum_{i\in[n]}\mathcal{L}^i_{single}\otimes\mathbb{I}_{[n]\setminus\{i\}} \tag{9}$$
Since each single qubit Lindbladian $\mathcal{L}_{single}$ is highly explicit (it acts on 2×2 matrices), in Section III, following now standard techniques, we are able to prove simple bounds on its MLSI constant.

Claim II.1. *$\mathcal{L}_{\mathsf{NI}}$ satisfies a MLSI with constant $\Omega(e^{-\beta})$.*

c) The convex combination argument: The main technical challenge in our analysis lies in relating $\mathcal{L}_{\mathsf{NI}}$ with our family of Davies generators $\mathcal{L}$ from Eq. (6), in order to inherit the rapid mixing properties from the former. The crux of our proof lies in analyzing $\mathcal{L}$ in a basis rotated by C, to show that the rotated Davies generator is a *convex combination* of $\mathcal{L}_{\mathsf{NI}}$ and some other Davies generator. This involves a delicate lightcone argument shown in Fig. 2, and discussed shortly.

Claim II.2. *In a basis rotated by C, the Lindbladian $\mathcal{L}$ from Eq. (6) can be written as a convex combination*
$$\tilde{\mathcal{L}} \equiv C^\dagger\mathcal{L}[C\cdot C^\dagger]C = q\cdot\mathcal{L}_{\mathsf{NI}}[\cdot] + (1-q)\cdot\mathcal{L}_{rest}[\cdot], \tag{10}$$
where both $\mathcal{L}_{\mathsf{NI}}, \mathcal{L}_{rest}$ share the fixed point σ_β, and $q = 4^{1-\ell}$.

[9] This set of jump operators which "drive" the transition can be essentially arbitrary, however, this choice resembling the connectivity of the underlying Hamiltonian will play an important role in our analysis.

[10] Note that there are $|\mathcal{A}| = n\cdot 4^\ell$ jump operators.

[11] Where $\{A,B\} := AB + BA$ is the anti-commutator.

[12] The quantum relative entropy between two density matrices ρ,σ is given by $D(\rho||\sigma) = \mathrm{Tr}\big[\rho\cdot(\log\rho - \log\sigma)\big]$.

We emphasize that the parameter $q \in [0,1]$ only depends on the lightcone size of C. Moreover, while $\mathcal{L}_{\mathsf{NI}}$ is the well-understood non-interacting system discussed previously, $\mathcal{L}_{rest}$ may apriori be arbitrary. However, at the very least we know it shares a fixed point with $\mathcal{L}_{\mathsf{NI}}$.

Neverthless, in Section II-A we show that convexity is precisely enough[13] to exhibit an MLSI for $\mathcal{L}$, with a constant which is only a multiplicative factor of q off of that of $\mathcal{L}_{\mathsf{NI}}$.

Lemma II.3. *$\mathcal{L}$ satisfies a MLSI with constant $\Omega(4^{-\ell}e^{-\beta})$.*

d) The lightcone argument: We conclude by discussing the key technical step: a proof of Claim II.2 via a lightcone argument. The starting point is to examine the Davies generator $\mathcal{L}$ (Eq. (6)) and its rotated version $\tilde{\mathcal{L}} = C^\dagger\mathcal{L}[C\cdot C^\dagger]C$. The goal is to show that "a fraction of" $\tilde{\mathcal{L}}$ equals the Davies generator of the non-interacting Hamiltonian H_{NI}, that is, within that fraction, the effect of C is erased.

To begin, note that the Davies generator $\mathcal{L}$ (Eq. (6)) only depends on the circuit C through the jump operators decomposed into the frequency basis, A^a_ν, which we recollect can be written as

$$\begin{aligned} A^a_\nu &= \sum_k \Pi_{k+\nu} A^a \Pi_k \\ &= \sum_k C\Big(\sum_{|y|=k+\nu} |y\rangle\langle y|\Big)C^\dagger A^a C\Big(\sum_{|x|=k}|x\rangle\langle x|\Big)C^\dagger. \end{aligned} \tag{11}$$

Crucially, due to the rotation of $\mathcal{L}$ to $\tilde{\mathcal{L}}$,[14] we observe that the dependence of C within $\tilde{\mathcal{L}}$ is only through second-moment operators of the form

$$\mathop{\mathbb{E}}_{P\sim\mathcal{P}_\ell}[C^\dagger PC \otimes C^\dagger PC], \tag{12}$$

where we consider the sum of all jump operators that act on a specific lightcone L_i of size ℓ, and recall that the jump operators are ℓ-qubit Pauli operators. It remains to express this operator as a convex combination, as shown in Fig. 2. A sketch of the argument follows:

- Step (i): Uses the identity $\mathbb{E}_{P\sim\mathcal{P}_\ell}[P\otimes P] = \frac{1}{2^\ell}\cdot\mathrm{SWAP}$, and linearity of expectation.
- Step (ii): Since C is a low-depth circuit, one can cancel the quantum gates within the lightcone of qubit i with their inverse.
- Step (iii): Uses the identity $\mathbb{E}_{P\sim\mathcal{P}_\ell}[P\otimes P] = \frac{1}{2^\ell}\cdot\mathrm{SWAP}$ again, but in the other direction.
- Step (iv): Re-writes the expectation into two parts: the first part is over the 4 single-qubit Paulis that act only on qubit i, and the second part, is all the remaining ℓ-qubit Paulis.

The crux of the argument lies in noting that in the first part of Step (iv), the ith qubit has been completely disentangled from the remaining circuit. Thereby, the single-qubit Pauli acts on a disentangled wire, and all remaining gates cancel with each other.

This gives the desired convex combination, where the first term corresponds to the non-interacting Hamiltonian with single-qubit jump operators. Finally, note that our choice of the jump operators (as ℓ-qubit Paulis acting on each lightcone) is crucial for this argument, and it is unclear if an arbitrary choice of jump operators would suffice.

3) Efficient implementation on a quantum computer: Rapid mixing of the Davies generator implies that the Gibbs state can be efficiently prepared in the thermal model of computation, described by coupling the quantum system to a thermal bath [31]. Next, we briefly discuss how to simulate the dissipative Lindbladian evolution $e^{\mathcal{L}t}$ on a quantum computer.

We leverage the "continuous-time quantum Gibbs sampler" framework of [13]. They show that to implement the map $e^{\mathcal{L}t}$ one requires $\tilde{O}(t)$ black-box invocations to a *unitary block-encoding* of the Lindblad operators ([13], Theorem I.1). In turn, to implement such a block-encoding for Hamiltonians H of integer spectra, it suffices to design quantum circuits which implement the Hamiltonian simulation of H, a block-encoding for the jump operators A^a, as well as a certain"frequency filter" which implements the Glauber dynamics weight. In Section IV, we discuss circuit implementations of all these ingredients, summarized in the following Lemma.

Lemma II.4 (Dissipative Lindbladian Implementation)**.** *Fix parameters $t \geq 1$ and $\varepsilon \leq \frac{1}{2}$. Let $\mathcal{L}$ denote the Lindbladian of Eq. (6), defined by a quantum circuit C on n qubits of depth d and lightcone size ℓ. Then, we can simulate the map $e^{t\mathcal{L}}$ to error ε in diamond norm using a quantum circuit of depth $O(t\cdot n\cdot 4^\ell\cdot 2^d\cdot \mathrm{poly}(\ell, \log n, \log\frac{1}{\varepsilon}, \log t))$.*

Put together with our bound on the mixing time, we arrive at our main statement on Gibbs state preparation in Lemma I.2.

B. Classical hardness of Gibbs sampling

As discussed in Section I, to obtain the classical intractability of quantum Gibbs sampling it suffices to construct a family of low depth quantum circuits which are hard to sample from even in the presence of input errors (Fig. 1b). The reason this imposes a challenge is two-fold. First, it is known that many classically hard shallow quantum circuits actually become classically simulable under input noise [20], thereby suggesting a need for fault-tolerance techniques. However, standard fault-tolerance techniques [24] often come with a prohibitive circuit depth overhead, which blows up the locality of the parent Hamiltonian. We address these challenges by designing a fault-tolerance scheme tailored to the input noise model with small overhead.

Our plan is to focus on IQP circuits, which are known to be already classically hard at constant depth. We show that their commuting structure plays an important role in our fault-tolerance techniques at low overhead.

1) Quantum computational advantage with shallow IQP circuits: Recall that IQP circuits can be written as $C =$

[13] This is inspired by [30], who leveraged the concavity of the spectral gap to prove mixing properties of stochastic Hamiltonians.

[14] And the fact that our jump operators are Pauli operators.

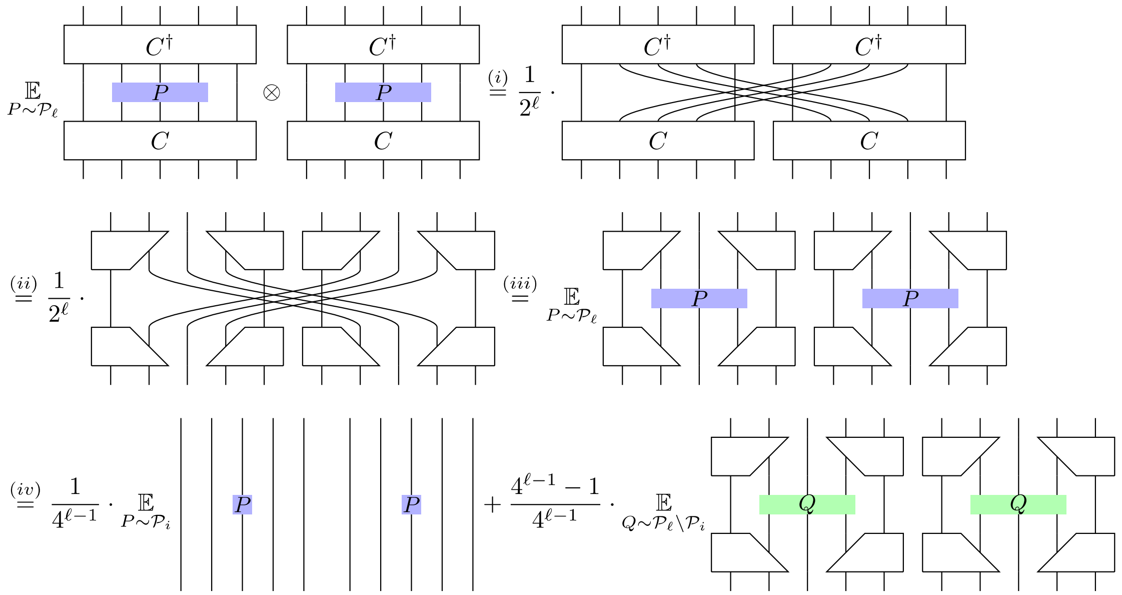

$\mathbb{E}_{P\sim\mathcal{P}_\ell}$
$C^\dagger$
P
C
$\otimes$
$\overset{(i)}{=} \frac{1}{2^\ell} \cdot$
$\overset{(ii)}{=} \frac{1}{2^\ell} \cdot$
$\overset{(iii)}{=} \mathbb{E}_{P\sim\mathcal{P}_\ell}$
$\overset{(iv)}{=} \frac{1}{4^{\ell-1}} \cdot \mathbb{E}_{P\sim\mathcal{P}_i}$
$+ \frac{4^{\ell-1}-1}{4^{\ell-1}} \cdot \mathbb{E}_{Q\sim\mathcal{P}_\ell\setminus\mathcal{P}_i}$
Q

Fig. 2: A lightcone argument for proving the modified log-Sobolev inequality.

$H^{\otimes n} D H^{\otimes n}$, where D is a diagonal unitary. The induced probability distribution $p(x) = |\langle x|C|0^n\rangle|^2$ is hard to sample from classically in general [32]. While any family of constant-depth and classically-hard IQP circuits suffices for our purpose, in this paper we use the concrete example of cluster states on regular lattices composed with random Z-rotations[15], which have become the basis for various proposals of sampling-based quantum supremacy using low-depth circuits [32], [20], [4], [5], [33], [34].

We present the structure of these circuits in more detail in Section VII, where we additionally present a comprehensive discussion on the foundations of their hardness. As a brief overview, note that 2D cluster states with single-qubit Z rotations is a universal resource state for measurement-based quantum computation (MBQC) [28]. This implies that *exactly* sampling from their output distribution is hard in the worst-case [35]. The hardness of approximate sampling from these architectures are based on further assumptions [36], [4], [33], which we rigorously define in Section VII. The following theorem thus provides the complexity-theoretic basis of our hardness arguments.

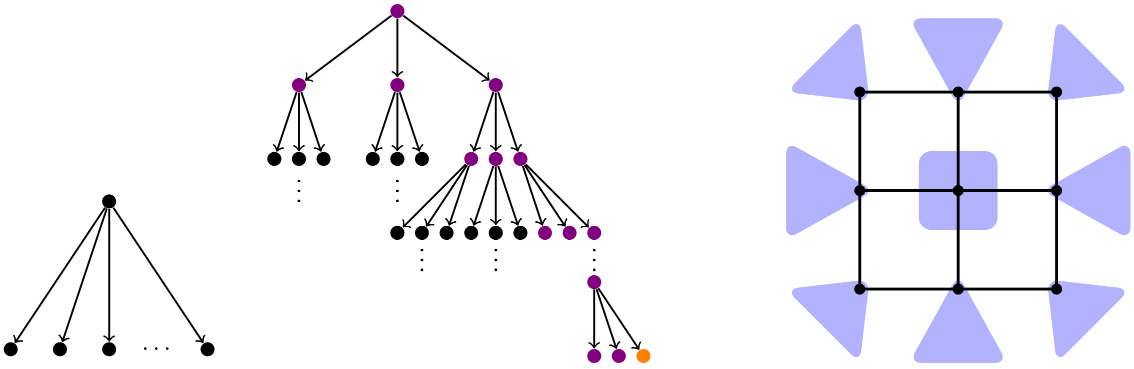
Theorem II.5 (Complexity of constant-depth IQP sampling [4], [5])**.** *There exists a constant $\delta > 0$, and a family of constant depth IQP circuits $\{C_n\}_{n\geq 1}$ on n qubits, such that no randomized classical polynomial-time algorithm can sample from the output distribution of C_n up to additive error δ in total variation distance, assuming the average-case hardness of computing a fixed family of partition functions (Conjecture 1), and the non-collapse of the Polynomial Hierarchy.*

To establish classical hardness of the Gibbs sampling task, it suffices to map the above circuit C to a fault-tolerant circuit $\tilde{C}$, such that a sample from the output distribution of $\tilde{C}$ under input noise can be efficiently post-processed into an ideal sample from C. The key challenge is to reduce the fault-tolerance overhead in $\tilde{C}$, so that the corresponding parent Hamiltonian has small locality.

2) Fault-tolerance of IQP circuits against input noise: The starting point in our approach is the observation that it suffices to *error-detect* the random inputs bits, instead of correcting them, to preserve the hardness-of-sampling of C. Indeed, bit-flip errors (which are Pauli-X errors) on the input of IQP circuits, become phase-flip errors after the first layer of Hadamard gates, and thus commute with the entire IQP circuit. In this manner, they are equivalent to bit-flip errors on the measured output string. Therefore, if we could identify the computational basis state $|r\rangle = \otimes_i |r_i\rangle$ fed into the IQP circuit, we would be able to correct the measured output sample by simply subtracting $r \in \{0,1\}^n$. Indeed, we emphasize we don't intend to correct the input error within the quantum circuit at all, as this would require decoding and feedforward, and potentially a much deeper circuit. Instead, we correct the error only during classical post-processing (that is, after all qubits are measured).

The crux of our approach is the design of a "distillation" gadget, which independently pre-processes each input bit r_i into k others in such a manner which enables us to reconstruct r_i (with high probability) given only the other $k-1$ noisy bits. We illustrate this task with a simple example, based on the repetition code.

[15] In the literature, these circuits are also known as the "evolution (quench) of an nearest-neighbor, translationally invariant (NNTI) Hamiltonian".


(a) Repetition gadget

(b) Recursive concatenation

(c) Fault-tolerant circuit

Fig. 3: Fault-tolerance via state distillation gadgets. (a) The repetition code gadget. (b) A B-Tree and the recursive concatenation scheme. Arrows denote the direction of CNOT gates. (c) Pre-processing the circuit using distillation gadgets.

a) A distillation gadget based on the repetition code: Recall that all input bits are noisy: each of them is flipped from 0 to 1 with probability q. Given k bits drawn from $s \leftarrow \mathsf{Bern}^k(q)$, suppose we designate the k-th bit as the "root" and apply a CNOT gate from it to the other $k-1$ bits (Fig. 3a). During the decoding stage, we would like to reconstruct the root bit given the other $k-1$ bits. To do this we simply compute the majority of the "leaves":

$$\hat{s}_k = \mathsf{Maj}(s_1 \oplus s_k, s_2 \oplus s_k, \cdots, s_{k-1} \oplus s_k). \tag{13}$$

We show that the probability of failure (when $\hat{s}_k \neq s_k$) equals $\delta = q^{\Omega(k)}$.

To highlight how these gadgets can be used for fault-tolerance, given an n-qubit IQP circuit C, we begin by pre-processing each of n input bits independently into a distillation gadget of size k, resulting in a circuit on $n \cdot k$ bits. Each of the n "root" bits are then fed into C (Fig. 3c). Note that the $n \cdot (k-1)$ remaining bits are untouched by C. In the end, after all qubits are measured, we can use the $n \cdot (k-1)$ ancilla bits to infer if an error had happened on each of the "root" bits fed into the circuit. As argued earlier, if an error did happen, it can be corrected by simply flipping the measurement outcome since the error commutes with the circuit. If we choose $k = \Theta(\log n)$, then the entire error correction process succeeds with high probability.

b) Recursive concatenation and B-Trees: In effect, the scheme above distills the "root" bit s_k with an effective bit-flip error rate $q^{\Omega(k)}$, using $k-1$ redundant "syndrome" bits of error rate q. Note that it used no information about the distribution of s_k, only that of the "leaves" $s_1, \cdots, s_{k-1}$.

To improve on this example, we bootstrap the above technique by recursively preparing "syndrome" bits of better and better fidelity.[16] Suppose we organize k bits into a tree of arity B and depth 2, such $k = 1 + B + B^2$. Moreover, apply the repetition code gadget on each layer, from leaves to root of the tree, by applying a CNOT gate from each parent bit to their respective children bits in the tree. In doing so, by the previous analysis we can identify each bit at the middle layer, just using the bits at the leaves, with error probability $q^{\Omega(B)}$. By performing majority again at the middle layer, we are now able to identify the bit at the root of this two-layer tree with error rate $(q^{\Omega(B)})^{\Omega(B)} = q^{\Omega(B^2)}$. By recursively applying this approach on a B-tree of depth d, the error probability at the root of the tree scales doubly-exponentially with the depth d, $q^{B^{\Omega(d)}}$.

At face value, it may seem that we haven't gained anything over the repetition code, as the error probability still only decays exponentially with the size of the gadget. The advantage lies instead in the *locality* of the gadget. Indeed, consider the lightcone of the orange qubit u at the leaf of the tree in Fig. 3b. By examining the causal influence of this qubit, we conclude that only the qubits in the neighborhood of its path to the root (the purple nodes in Fig. 3b) can lie in its lightcone. That is, if

$$u = u_0 \rightarrow u_1 \rightarrow \cdots \rightarrow u_d \equiv \text{ root} \tag{14}$$

denotes the path from leaf to root, then the *lightcone* of u is contained the union of the neighborhoods $\mathsf{L}_u \equiv \cup_i^d N(u_i)$. Therefore, $|\mathsf{L}_u| \leq O(B \cdot d)$, which is a linear function of the depth of the tree. By further studying the propagation of Z Pauli's through the gadget, we analogously show that that the locality of the parent Hamiltonian of the distillation circuit is $|\mathsf{S}_u| \leq d$; precisely the nodes on the path from leaf to root. Lemma I.3 then follows from a careful choice of B and d.

C. Organization

We organize the rest of this work as follows.

Gibbs State Preparation. In Section III, we prove our rapid mixing bounds for Davies Generators, and in Section IV discuss their simulation on a quantum computer. In Section V, we discuss alternative state preparation algorithms for parent

[16] This construction is largely inspired by recursive magic state distillation schemes.

Hamiltonians of circuits which lie on lattices.

Classical Intractability of Gibbs Sampling. In Section VI, we prove that the constant temperature Gibbs states of the Hamiltonians in $\mathscr{H}$, can be interpreted as the output of noisy circuits. In Section VII, we present an overview of the computational complexity of shallow IQP sampling. In Section VIII, we present our fault-tolerance scheme based on state distillation.

Finally, in Section IX, we put everything together and prove our main result (Theorem I.1).

Applications. In Section XI, we present our results on Gibbs sampling with measurement errors, and in Section X, we discuss the BQP completeness of Gibbs sampling with adaptive single-qubit measurements.

III. Rapid Mixing and Efficient Gibbs State Preparation

We dedicate this section to a proof of the rapid convergence of our dissipative Lindbladians. We defer a discussion on its implementation using quantum circuits to Section IV. For simplicity, henceforth we re-scale the class of parent Hamiltonians,[17]

$$H = \sum_{i\in[n]} h_i = \sum_i C\big(|1\rangle\langle 1| \otimes \mathbb{I}_{[n]\setminus i}\big)C^\dagger \qquad (15)$$

to ensure frustration-freeness and positive integer spectra $[n] = \{0, \cdots, n\}$. Recall this Hamiltonian is commuting, and its eigenstates are given by $\{C\,|x\rangle : x \in \{0,1\}^n\}$. Let ℓ be the lightcone size of C. The jump operators of our Davies generator are ℓ-qubit Pauli operators on the support of the circuit lightcone. We refer the reader to Section II-A, Eq. (6) for a description of our Lindbladian.

Remark III.1. *The support S_i of each Hamiltonian term h_i is contained within the lightcone L_i, see below Eq. (3). In general, h_i acts nontrivially on the entire lightcone (for example, when C uses Haar random gates), but for our hard instances $|\mathsf{S}_i| \ll |\mathsf{L}_i|$. To ensure the quantum algorithm always "knows" $\{\mathsf{L}_i\}$ when it only sees $\{h_i\}$, we assume this information is encoded in the definition of the family of Hamiltonians, and is given to the algorithm.*

A. Preliminaries on thermal Lindbladians and their convergence

We dedicate this subsection to background on the evolution and convergence of open quantum systems described by a Lindbladian. Recall, a general Lindbladian is a continuous-time Markov chain acting on density operators:

$$\mathcal{L}[\rho] = \sum_j J_j\rho J_j^\dagger - \frac{1}{2}\{J_jJ_j^\dagger, \rho\} \qquad (16)$$

for some set of Lindblad operators $\{J_j\}_j$, which generates a family of completely positive and trace-preserving map

$$e^{\mathcal{L}t}[\rho] \quad \text{for each} \quad t \geq 0. \qquad (17)$$

Our Lindbladians of interest satisfy a particular property known as detailed balance.

Definition III.1 (s-Inner Product)**.** *Fix a full rank density matrix σ and $s \in [0,1]$. We define the weighted Hilbert-Schmidt inner product: for any A, B,*

$$\langle A, B\rangle_s = \langle A, \sigma^{1-s}B\sigma^s\rangle = \mathrm{Tr}\big[A^\dagger\sigma^{1-s}B\sigma^s\big]. \qquad (18)$$

Definition III.2 (s-Detailed Balance)**.** *A Lindbladian $\mathcal{L}$ is s-*Detailed Balance *with respect to σ if $\mathcal{L}^\dagger$ is self-adjoint with respect to $\langle\cdot,\cdot\rangle_s$:*

$$\forall A, B : \langle A, \mathcal{L}^\dagger[B]\rangle_s = \langle\mathcal{L}^\dagger[A], B\rangle_s \qquad (19)$$

There are two important structural consequences of this detailed balance condition. The first is that the density operator σ is a fixed point of Lindbladian evolution:

$$\mathcal{L}[\sigma] = 0 \qquad (20)$$

The second, as discussed shortly, is that it implies a powerful means to understand the convergence of the mixing process. For the reader most familiar with classical Markov chains, the detailed balance condition above is an analog to its classical counterpart, however, with an additional degree of freedom $0 \leq s \leq 1$ which arises due to non-commutativity.

Two special cases of the above are the GNS (where $s = 1$) and KMS ($s = 1/2$) detailed balance conditions. Fortunately, under minor constraints on the family of Lindbladians (which our Lindbladian satisfies), all these definitions collapse. We refer the reader back to Eq. (6) for the definition of the family of Lindbladians we consider, Davies Generators.

Fact III.1 (Davies' generators are detailed balanced)**.** *Consider the Davies generator $\mathcal{L}$ described in Eq. (6), subject to the constraint that the transition weights satisfy $\forall\nu$: $\gamma(\nu)/\gamma(-\nu) = e^{-\beta\nu}$, and the jump operators contain their adjoints $\{A_a\} = \{A_a^\dagger\}$. Then, $\mathcal{L}$ satisfies s-DB $\forall s \in [0,1]$ w.r.t. the Gibbs state $\rho_\beta \propto e^{-\beta H}$.*

In this manner, the Gibbs state $\rho_\beta \propto e^{-\beta H}$ is a fixed point of the Davies generator we designed in Section II-A. However, it may not be the unique stationary state, nor may its evolution converge rapidly. To understand the rate of convergence of this process, we need a bound on its mixing time $t_{mix}(\mathcal{L})$. Physically, the mixing time provides an estimate for the thermalization time of the system.

Definition III.3 (Mixing time)**.** *The mixing time $t_{mix}(\mathcal{L})$ of a Lindbladian $\mathcal{L}$ is the smallest time $t \geq 0$ for which*

$$\|e^{t\mathcal{L}}(\rho_1 - \rho_2)\|_1 \leq \frac{1}{2}\|\rho_1 - \rho_2\|_1 \quad \textit{for any two states } \rho_1, \rho_2. \qquad (21)$$

In what remains of this subsection, we describe two means

[17] Note that we are simply applying an affine transformation, $H = \frac{1}{2}\big(n \cdot \mathbb{I} + C\sum Z_iC^\dagger\big)$, such that the Gibbs state of H at temperature β is the same as that of $C\sum Z_iC^\dagger$ at temperature $\beta/2$.

to analyze t_{mix}. The first of which consists of a bound on the spectral gap of $\mathcal{L}$. Apriori, however, the super-operator $\mathcal{L}$ is not even Hermitian, and its spectral gap may not even be well-defined. Fortunately, under an appropriate similarity transformation, we can appeal to a related Hermitian quantity known as *the discriminant*:

Definition III.4 (Quantum discriminant)**.** *Fix $s \in [0, 1]$ and a full-rank density matrix σ. The discriminant $\mathcal{K}_s$ of $\mathcal{L}$ consists of the super-operator*

$$\mathcal{K}_s(\cdot) = \sigma^{-\frac{1-s}{2}} \mathcal{L}\left(\sigma^{\frac{1-s}{2}} \cdot \sigma^{\frac{s}{2}}\right) \sigma^{-\frac{s}{2}}. \tag{22}$$

Lemma III.1 ([29], Lemma 5 and 7)**.** *The discriminant $\mathcal{K}_s$ of $\mathcal{L}$ satisfies the following properties*

1) *$\mathcal{L}$ satisfies s-DB if and only if $\mathcal{K}_s$ is Hermitian.*
2) *If $\mathcal{L}$ satisfies s-DB, then the eigenvalues of $\mathcal{L}$ are the same as that of $\mathcal{K}_s$, which are real.*
3) *If $\mathcal{L}$ is a Davies generator satisfying the constraints of Fact III.1, then $\mathcal{K} \equiv \mathcal{K}_s$ is independent of $s \in [0, 1]$.*

The spectral gap $\Delta(\mathcal{L}) = \Delta(\mathcal{K}_s)$ of a given Lindbladian is defined to be that of the associated discriminant. Analyzing this gap can be a challenging task, and concrete bounds are often case-dependent. Nevertheless, it provides a powerful means to control the convergence of the time-evolution.

Lemma III.2 (Mixing time from the Spectral Gap, [15])**.** *If a Lindbladian $\mathcal{L}$ satisfies KMS reversibility with fixed point σ, then*

$$t_{mix}(\mathcal{L}) \leq \frac{\log\left(2\|\sigma^{-1/2}\|\right)}{\Delta(\mathcal{L})}, \tag{23}$$

We remark that the dependence on $\log \|\sigma^{-1/2}\| \approx O(\beta n)$ often-times incurs a polynomial overhead to the mixing time. The notion of a (modified) Log Sobolev inequality provides a significantly stronger means of analyzing the mixing time. To formalize this method, we first require the definition of the conditional expectation of an operator X, $\mathcal{E}[X] = \lim_{t\to\infty} e^{t\mathcal{L}}[X]$.

Definition III.5 (Modified Logarithmic Sobolev inequality)**.** *The Markov semigroup $(e^{t\mathcal{L}})_{t\geq 0}$ satisfies a* Modified Logarithmic Sobolev inequality *(MSLI) with constant α if for any ρ,*

$$\begin{aligned} \frac{d}{dt} D\left(e^{t\mathcal{L}}[\rho] || \mathcal{E}[\rho]\right)\Big|_{t=0} &= \operatorname{Tr} \mathcal{L}[\rho](\log\rho - \log\mathcal{E}[\rho]) \\ &\leq -\alpha \cdot D(\rho||\mathcal{E}[\rho]), \end{aligned} \tag{24}$$

where $D(\rho||\sigma) = \operatorname{Tr} \rho(\log\rho - \log\sigma)$ is the quantum relative entropy.

In other words, a MLSI quantifies the decay of the relative entropy, which converts to a bound on the mixing time through Pinsker's inequality.

Lemma III.3 (Mixing time from MLSI, [15])**.** *If a Lindbladian $\mathcal{L}$ satisfies KMS-detailed balance with fixed point σ and a MLSI with constant α, then*

$$t_{mix}(\mathcal{L}) \leq \frac{2 \cdot \log\left(4 \cdot \log\|\sigma^{-1}\|\right)}{\alpha} \tag{25}$$

This *polylogarithmic* overhead in system size is known as *rapid mixing*. Moreover, if given an additional entangled reference system R the semigroup $(e^{t\mathcal{L}} \otimes \mathbb{I}_R)_{t\geq 0}$ satisfies an MSLI, then $\mathcal{L}$ is said to satisfy a *complete* modified logarithmic Sobolev inequality (CMLSI).

B. Analysis

The main result of this subsection is a bound on the mixing time of our family of Lindbladians,

Lemma III.4. *The mixing time of our family of Lindbladians $\mathcal{L}$ defined in Eq. (6) is bounded by*

$$t_{mix}(\mathcal{L}) = O(4^\ell \cdot e^\beta \cdot \log n). \tag{26}$$

The starting point of our analysis is based on that of a much simpler Lindbladian, namely, that corresponding to the trivial circuit $C = \mathbb{I}$. In this setting, both the associated parent Hamiltonian, and the associated Lindbladian, are a sum over non-interacting, single-qubit terms:

$$H_{\mathsf{NI}} = \sum_i |1\rangle\langle 1|_i \text{ and } \mathcal{L}_{\mathsf{NI}} = \sum_{i\in[n]} \mathcal{L}^i_{single}, \tag{27}$$

where $\mathcal{L}^i_{single}[\sigma^i_\beta] = 0$ and $\sigma^i_\beta \propto e^{-\beta|1\rangle\langle 1|_i}$. The jump operators of $\mathcal{L}^i_{single}$ are simply single-qubit Pauli operators, and the single-qubit Gibbs state σ^i_β is its fixed point. Using now standard techniques, one can prove that this non-interacting Lindbladian is both gapped and mixes rapidly:

Claim III.5 (The Non-Interacting Lindbladian is rapidly mixing)**.** *The non-interacting Lindbladian $\mathcal{L}_{\mathsf{NI}}$ has a constant spectral gap $\Delta(\mathcal{L}_{\mathsf{NI}}) \geq 4^{-1}$ and satisfies a MSLI with constant $\alpha_{\mathsf{NI}} = \Omega(e^{-\beta})$.*

The *unique* fixed point of $\mathcal{L}_{NI}$ is thus the tensor product state $\sigma_\beta = \otimes_i \sigma^i_\beta \propto e^{-\beta H_{\mathsf{NI}}}$. We defer a proof of Claim III.5 to the next subsection. In the rest of this subsection, we show how to relate our Lindbladian $\mathcal{L}$ of Eq. (6) (implicitly defined by the quantum circuit C), to $\mathcal{L}_{\mathsf{NI}}$, and moreover how to inherit its rapid mixing properties.

Claim III.6 (A Convex Combination of Lindbladians)**.** *In a basis rotated by C, the Lindbladian $\mathcal{L}$ can be written as the convex combination*

$$C^\dagger \mathcal{L}[C \cdot C^\dagger] C = q \cdot \mathcal{L}_{\mathsf{NI}}[\cdot] + (1-q) \cdot \mathcal{L}_{rest}[\cdot], \tag{28}$$

of two Lindbladians $\mathcal{L}_{\mathsf{NI}}, \mathcal{L}_{rest}$ which share the fixed point $\sigma_\beta = \otimes_i \sigma^i_\beta$. Moreover, the parameter $q = 4^{1-\ell}$ depends only on the lightcone size of C.

A proof of which we also defer to a future subsection. The convex combination claim above is the heart of our analysis, as it enables us to inherit the gap and mixing properties of $\mathcal{L}_{NI}$, without knowing properties of $\mathcal{L}_{rest}$ except for its (common) fixed point. To conclude this subsection, we present a proof of the MLSI of $\mathcal{L}$:

Claim III.7 (The Modified Log-Sobolev Inequality)**.** *The Lindbladian $\mathcal{L}$ satisfies a MSLI with constant $\alpha \geq q \cdot \alpha_{NI} = \Omega(4^{-\ell} \cdot e^{-\beta})$.*

Proof. [of Claim III.7] From Claim III.6, we can write our Lindbladian $\mathcal{L}$ in a basis rotated by the circuit C as a convex combination

$$\tilde{\mathcal{L}} = C^\dagger \mathcal{L}[C \cdot C^\dagger] C = q \cdot \mathcal{L}_{\mathsf{NI}}[\cdot] + (1-q) \cdot \mathcal{L}_{rest}[\cdot] \tag{29}$$

Since relative entropy is basis independent, proving a MLSI for $\tilde{\mathcal{L}}$ similarly implies one for $\mathcal{L}$ with the same constant. To do so, we begin by expressing the "entropy production rate" as a convex combination.

$$\begin{aligned} \mathsf{EP}_{\tilde{\mathcal{L}}}(\rho) &= \mathrm{Tr}\Big[\tilde{\mathcal{L}}[\rho](\log\rho - \log\sigma_\beta)\Big] \\ &= q \,\mathrm{Tr}[\mathcal{L}_{\mathsf{NI}}[\rho](\log\rho - \log\sigma_\beta)] \\ &\quad + (1-q)\,\mathrm{Tr}[\mathcal{L}_{rest}[\rho](\log\rho - \log\sigma_\beta)] \end{aligned} \tag{30}$$

To the first term on the RHS above, we can simply apply the MLSI for the non-interacting Lindbladian Claim III.5:

$$\mathrm{Tr}[\mathcal{L}_{\mathsf{NI}}[\rho](\log\rho - \log\sigma_\beta)] \leq -\alpha_{\mathsf{NI}} \cdot D(\rho||\sigma_\beta). \tag{31}$$

In turn, we claim that the second term on the RHS above is non-positive. Indeed, note that by Claim III.6, σ_β is a fixed point of $\mathcal{L}_{rest}$. The data-processing inequality for the relative entropy then tells us that

$$\begin{aligned} &\mathrm{Tr}[\mathcal{L}_{rest}[\rho](\log\rho - \log\sigma_\beta)] \\ &= \frac{d}{dt} D(e^{t\mathcal{L}_{rest}}[\rho]||\sigma_\beta)\Big|_{t=0} \\ &= \frac{d}{dt} D(e^{t\mathcal{L}_{rest}}[\rho]||e^{t\mathcal{L}_{rest}}[\sigma_\beta])\Big|_{t=0} \leq 0. \end{aligned} \tag{32}$$

Put together, we conclude $\mathsf{EP}_{\tilde{\mathcal{L}}}(\rho) \leq -q \cdot \alpha_{\mathsf{NI}} \cdot D(\rho||\sigma_\beta)$. □

C. The non-interacting Lindbladian is gapped (Claim III.5)

We dedicate this subsection to an analysis of the non-interacting Lindbladian $\mathcal{L}_{NI}$ (Claim III.5).

Lemma III.8. *The spectral gap of the single-qubit Lindbladian $\mathcal{L}_{single}$ is $\Delta(\mathcal{L}_{single}) \geq 4^{-1}$.*

To understand this spectral gap, we revisit the (Hermitian) Discriminant super-operator $\mathcal{K}$ defined in Definition III.4. Recall, from Lemma III.1, that (under detailed balance) this super-operator has the same eigenvalues of $\mathcal{L}$. In turn, to understand the spectral gap of $\mathcal{K}$, we vectorize this super-operator (on 2×2 matrices) into an operator (a 4×4 matrix).

$$\mathcal{K}[\cdot] = \sum_j A_j[\cdot]B_j \rightarrow \mathbf{K} = \sum_j A_j \otimes B_j^T. \tag{33}$$

Proof. To analyze the gap, we consider the discriminant $\mathcal{K}_{single}$ of the Lindbladian $\mathcal{L}_{single}$, and in particular its vectorization:

$$\begin{aligned} \mathbf{K}_{single} = \sum_{a\in\mathcal{A}} \sum_{\nu\in[-n,n]} & -\sqrt{\gamma(\nu)\gamma(-\nu)} \cdot A^a_\nu \otimes (A^a_\nu)^* \\ & + \frac{\gamma(\nu)}{2}\Big((A^a_\nu)^\dagger A^a_\nu \otimes \mathbb{I} + \mathbb{I} \otimes (A^a_\nu)^T (A^a_\nu)^*\Big) \end{aligned} \tag{34}$$

which is PSD, frustration free, and preserves the eigenvalues of $\mathcal{L}_{single}$ (up to a factor of -1) Lemma III.1. Moreover, via detailed balance, the purified Gibbs state $\left|\sqrt{\sigma_\beta}\right\rangle \propto |00\rangle + e^{-\beta/2}|11\rangle$ is a ground state of $\mathbf{K}_{single}$. Since the jump operators are single qubit Pauli operators $\{\mathbb{I}, X, Y, Z\}$, they can be written in the energy basis as

$$A_0^{\mathbb{I}} \propto \mathbb{I} \quad \text{and} \quad A_0^Z \propto Z_i \quad \text{and} \quad A_1^X = (-i)A_1^Y \propto \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{35}$$

and such that the conjugates can be inferred from the identity $A^a_\nu = A^{a\dagger}_{-\nu}$. The 4×4 vectorized discriminant can therefore be written as

$$\mathbf{K}_{single} = \tfrac{1}{2}\begin{bmatrix} \gamma(1) & 0 & 0 & -\sqrt{\gamma(1)\gamma(-1)} \\ 0 & \frac{\gamma(1)+\gamma(-1)}{2} + \gamma(0) & 0 & 0 \\ 0 & 0 & \frac{\gamma(1)+\gamma(-1)}{2} + \gamma(0) & 0 \\ -\sqrt{\gamma(1)\gamma(-1)} & 0 & 0 & \gamma(-1) \end{bmatrix} \tag{36}$$

which we identify to be frustration free and have spectral gap $\frac{\gamma(1)}{2} \cdot \min(1 + e^\beta, \frac{1+e^\beta}{2} + \frac{\gamma(0)}{\gamma(1)}) = \frac{\gamma(1)}{2} \geq \frac{1}{4}$ under Glauber Dynamics, where $\gamma(\nu) = (1 + e^{-\beta\nu})^{-1}$. □

The positivity of the spectral gap can be used to show a complete MLSI, as shown by [37]. This conversion comes at the cost of factors of the local dimension of the Lindbladian - which in the case of $\mathcal{L}_{single}$, is just 2.

Theorem III.9 (CMLSI from the Spectral Gap, [37] Theorem 4.3)**.** *Suppose a Lindbladian $\mathcal{G}$, acting on a D-dimensional Hilbert space, is GNS-symmetric w.r.t a fixed state $\sigma > 0$. Then, it satisfies a CMLSI with constant*

$$\alpha_c \geq \Delta(\mathcal{G}) \cdot \frac{\|\sigma^{-1}\|^{-1}}{D^2}. \tag{37}$$

In this manner, $\mathcal{L}_{single}$ satisfies a CMLSI with constant $\alpha_{single} = \frac{1}{16\cdot(1+e^\beta)}$.[18] We are now in a position to prove the MLSI for $\mathcal{L}_{NI}$.

Proof. [of Claim III.5] We begin by leveraging the Complete MLSI, on the local Lindbladians (Theorem III.9 and Lemma III.8):

$$\begin{aligned} &\mathrm{Tr}[\mathcal{L}_{NI}[\rho](\log(\rho) - \log(\sigma_\beta))] \\ &= \sum_{i\in[n]} \mathrm{Tr}\big[\mathcal{L}^i_{single}[\rho](\log(\rho) - \log(\sigma_\beta))\big] \\ &= \sum_{i\in[n]} \mathrm{Tr}\big[\mathcal{L}^i_{single}[\rho](\log(\rho) - \log(\mathcal{E}_i[\rho]))\big] \\ &\leq -\alpha_{single} \sum_{i\in[n]} D(\rho||\mathcal{E}_i[\rho]) \end{aligned} \tag{38}$$

[18] As remarked by an anonymous reviewer, it may be possible to remove the β dependence, since the Lindbladian is non-interacting, following the results of [38], [39], [40] or [41].

where $\mathcal{E}_i$ is the conditional expectation of the ith semigroup $e^{t\mathcal{L}^i_{single}}$. Next, we leverage the strong subadditivity of non-interacting conditional expectations [37] (eq. 5), see also [42]:

$$\sum_{i\in[n]} D(\rho||\mathcal{E}_i[\rho]) \geq D(\rho||\prod_i \mathcal{E}_i[\rho]). \tag{39}$$

To conclude, observe that for any ρ, the collection of conditional expectation $\prod_{i\in[n]}\mathcal{E}_i[\rho] = \sigma_\beta$ maps to the unique stationary state. □

D. The convex combination claim (Claim III.6)

Recall that the parent Hamiltonian H has solvable eigenstates given by $\{C\,|x\rangle : x \in \{0,1\}^n\}$, with energies given by the Hamming weight $|x| \in [n]$.

Proof. We begin by explicitly writing down each jump operator in the frequency basis:

$$\begin{aligned} A^a_\nu &= \sum_k \Pi_{k+\mu} A^a \Pi_k \\ &= C\Bigg(\sum_{k\in[n]} \sum_{\substack{|x|=k\\|y|=k+\nu}} |y\rangle\langle x| \cdot \langle y|\, C^\dagger A^a C\,|x\rangle\Bigg)C^\dagger \\ &=: C\Bigg(\sum_{k\in[n]} \sum_{\substack{|x|=k\\|y|=k+\nu}} |y\rangle\langle x| \cdot N^a_{x,y}\Bigg)C^\dagger. \end{aligned} \tag{40}$$

For conciseness, we have denoted the coefficient by $N^a_{x,y} = \langle y|\, C^\dagger A^a C\,|x\rangle$. Our Lindbladian $\mathcal{L}$ of Eq. (6) can thus be written in a basis rotated by the circuit C, in terms of the second moment of these coefficients:

$$\begin{aligned} C^\dagger\mathcal{L}[C\rho C^\dagger]C = \sum_\nu \gamma_\nu \cdot \sum_a \Bigg(&\sum_{k,k'\in[n]} \sum_{\substack{|x|=k\\|y|=k+\nu}} \sum_{\substack{|x'|=k'\\|y'|=k'+\nu}} N^a_{x,y}\cdot(N^a_{x',y'})^* \cdot |y\rangle\langle x|\,\rho\,|x'\rangle\langle y'| \\ &- \frac{1}{2}\sum_{k\in[n]} \sum_{\substack{|x|,|x'|=k\\|y|=k+\nu}} N^a_{x,y}\cdot(N^a_{x',y})^* \cdot \Big\{ |x'\rangle\langle x|, \rho\Big\}\Bigg). \end{aligned} \tag{41}$$

For $i \in [n]$, consider the subset of jump operators $\mathcal{A}_i$, centered around the i-th lightcone L_i:

$$\mathcal{A}_i = 2^{-\ell}\cdot\left\{P_{\mathsf{L}_i}\otimes \mathbb{I}_{[n]\setminus\mathsf{L}_i} : P \in \mathcal{P}_\ell\right\} \tag{42}$$

By definition, these subsets are disjoint, and form a partition $\cup_i\mathcal{A}_i = \mathcal{A}$. We claim that we can rotate the jump operators in each subset, by substituting

$$\mathcal{A}_i \to \mathcal{A}'_i = U_i\mathcal{A}_iU_i^\dagger \tag{43}$$

for an *arbitrary* choice of unitary U_i of support contained in L_i, while keeping the Lindbladian $\mathcal{L}$ invariant. Essentially, this is because the Lindbladian is only defined by the second moments of the jump operators, and that the second moment of random Pauli operators is Haar random (via the 1-design property) and thus invariant under unitary conjugation:

$$\sum_{a\in\mathcal{A}_i} A^a[\cdot]A^a = \frac{1}{2^\ell}\operatorname{tr}_{\mathsf{L}_i}[\cdot] = \sum_{a\in\mathcal{A}_i} U_i^\dagger A^a U_i[\cdot]U_i^\dagger A^a U_i \tag{44}$$

for any U_i supported on L_i. Indeed, for every choice of basis elements $x, x', y, y' \in \{0,1\}^n$, the pre-factor

$$\begin{aligned} &\sum_{a\in\mathcal{A}_i} N^a_{x,y}\cdot(N^a_{x',y'})^* \\ &= \sum_{a\in\mathcal{A}_i} \langle y|\,C^\dagger A_a C\,|x\rangle\langle x'|\,C^\dagger A_a C\,|y'\rangle \\ &= \sum_{a\in\mathcal{A}_i} \langle y|\,C^\dagger U_i A_a U_i^\dagger C\,|x\rangle\langle x'|\,C^\dagger U_i A_a U_i^\dagger C\,|y'\rangle \\ &= \sum_{a'\in\mathcal{A}'_i} N^{a'}_{x,y}\cdot(N^{a'}_{x',y'})^* \end{aligned} \tag{45}$$

is preserved, whether in $\mathcal{A}_i$ or $\mathcal{A}'_i$.

Finally, let U_i be the gates in C contained in the lightcone of the ith qubit. The sum over jump operators $a \in \mathcal{A}$ can be written as an expectation over random ℓ-qubit Paulis, $P \in \mathcal{P}_\ell$, and then an expectation over the center $i \in [n]$ in which to place P. With probability $4/4^\ell$, P is a single qubit Pauli P_i centered at i. Moreover, for a single-qubit Pauli P_i centered at i, the choice of U_i *exactly cancels with the circuit* $C^\dagger$:

$$\langle y|\,C^\dagger U_i\Big(P_i\otimes\mathbb{I}_{[n]\setminus\{i\}}\Big)U_i^\dagger C\,|x\rangle = \langle y|\Big(P_i\otimes\mathbb{I}_{n\setminus\{i\}}\Big)|x\rangle, \tag{46}$$

$$\text{for each} \quad x, y \in \{0,1\}^n \quad \text{and} \quad P_i \in \{\mathbb{I}, X, Y, Z\}_i. \tag{47}$$

We note that these are precisely the jump operators we expect in the non-interacting case $\mathcal{L}_{\mathsf{NI}}$, where the circuit is replaced by the trivial circuit $C = \mathbb{I}$. In this manner, we conclude that the rotated Lindbladian can be written a convex combination:

$$C^\dagger\mathcal{L}[C\cdot C^\dagger]C = 2^{2(1-\ell)}\cdot\mathcal{L}_{\mathsf{NI}}[\cdot] + (1-2^{2(1-\ell)})\cdot\mathcal{L}_{rest}[\cdot], \tag{48}$$

where both $\mathcal{L}_{\mathsf{NI}}[\cdot]$, $\mathcal{L}_{rest}[\cdot]$ are Davies' generators defined on disjoint sets of jump operators. Both of them satisfy detailed balance and share the Gibbs state as the stationary state. □

IV. Circuit Implementation of the Dissipative Lindbladian

The main claim of this section is an efficient implementation of the Lindbladian time-evolution using a quantum circuit. Put together with our bound on the mixing time of our Lindbladians, this all but concludes the proof of the preparation of Gibbs states of the parent Hamiltonians of quantum circuits.

Lemma IV.1 (Dissipative Lindbladian Implementation)**.** *Fix parameters $t \geq 1$ and $\varepsilon \leq \frac{1}{2}$. Let $\mathcal{L}$ denote the Lindbladian of Eq. (6), defined by a quantum circuit on n qubits, of lightcone size ℓ and depth d. Then, we can simulate the map $e^{t\mathcal{L}}$ to error ε in diamond norm using a quantum algorithm of depth $O(t\cdot n\cdot 2^{2\ell}\cdot 2^d\cdot \mathrm{poly}(\ell, \log n, \log\frac{1}{\varepsilon}, \log t))$.*

We dedicate Section IV-A to presenting the required background results on implementing Lindbladian evolution using quantum circuits. In the ensuing section Section IV-B, we discuss optimizations both particular to our systems, and generic, to the runtime of our algorithms.

A. Preliminaries on simulating Lindbladian evolution

Our implementation of the map $e^{t\mathcal{L}}$ follows the framework of [13], in reducing the task to constructing a block-encoding of the Lindblad operators. To implement their scheme it is suitable to renormalize the time-scale and Lindblad operators $\{L_j\}_{j\in\mathcal{A}}$ such that the resulting Lindbladian has norm 1:

$$t \to t\cdot \|\sum_j L_j^\dagger L_j\| \text{ and } L_j \to L_j \cdot \|\sum_j L_j^\dagger L_j\|^{-1/2} \quad (49)$$

Given the choice of jump operators from Eq. (6), under this normalization we have $t \to n\cdot t$.

Definition IV.1 (Unitary block encoding for Lindblad Operators [13, Definition I.2])**.** *Given a purely irreversible Lindbladian determined by the Lindblad operators $\{L_j\}_{j\in\mathcal{A}}$, a unitary U is said to be block-encoding of the Lindblad operators if*

$$\big(\langle 0|^b \otimes \mathbb{I}\big) U \big(|0\rangle^c \otimes \mathbb{I}\big) = \sum_{a\in\mathcal{A}} |a\rangle \otimes L_j \textit{ for } b, c\in\mathbb{N} \quad (50)$$

Given a black-box circuit corresponding to a block-encoding of $\mathcal{L}$, the following theorem stipulates that one can simulate the corresponding Lindbladian evolution for time t using just $\tilde{O}(t)$ invocations of the black-box:

Theorem IV.2 (Theorem I.2, [13])**.** *Suppose U is a unitary block-encoding of the Lindbladian $\mathcal{L}$ as in Definition IV.1. Let time $t\geq 1$ and error $\varepsilon \leq \frac{1}{2}$, then we can simulate the map $e^{t\mathcal{L}}$ to error ε in diamond norm using*

1) *$O((c+\log\frac{t}{\varepsilon})\log\frac{t}{\varepsilon})$ resettable ancilla qubits,*
2) *$\tilde{O}(t)$ controlled uses of U and $U^\dagger$, and*
3) *$\tilde{O}(t+c)$ other 2-qubit gates.*

We remark that since our Hamiltonian has a integer spectra $[n]$, one can exactly implement the projection of the jump operators $\{A^a\}$ onto the energy eigenbasis by performing an operator Fourier transform with uniform weights:

$$A^a_\nu \propto \sum_{\bar{t}\in S_{\pi/n}} e^{i\nu\bar{t}} e^{iH\bar{t}} A^a e^{-iH\bar{t}} \quad (51)$$

where $S_{\pi/n} = \frac{\pi}{n}\cdot\{-n, -(n-1), \cdots, -1, 0, 1, \cdots, n\}$. In this setting, we can now apply a lemma on the efficient implementation of block-encodings from [13], simplified to the context of integer spectra Hamiltonians.

Lemma IV.3 (Lemma I.1, [13])**.** *In the setting of Theorem IV.2, a unitary block encoding for the Lindblad operators corresponding to a Hamiltonian H of integer spectra $[n]$ can be created using $O(n+\log|A|)$ ancilla qubits, as well as one query to*

1) *The controlled Hamiltonian simulation: $\sum_{\bar{t}\in S_{\pi/n}} |\bar{t}\rangle\langle\bar{t}| \otimes e^{\pm iH\bar{t}}$,*
2) *A block-encoding of the jump operators: $\sum_{a\in\mathcal{A}} |a\rangle \otimes A^a$,*
3) *$O(\log n)$ qubit Quantum Fourier transform: $|\bar{t}\rangle \to (2n)^{-1/2}\sum_{\omega\in[-n,\cdots n]} e^{-i\omega\bar{t}}|\omega\rangle$*
4) *And a controlled filter for the Boltzmann factors:*

$$W = \sum_{\omega\in[-n,\cdots,n]} \begin{bmatrix} \sqrt{\gamma(\omega)} & -\sqrt{1-\gamma(\omega)} \\ \sqrt{1-\gamma(\omega)} & \sqrt{\gamma(\omega)} \end{bmatrix} \otimes |\omega\rangle\langle\omega| \quad (52)$$

B. Optimizing the circuit implementation

In light of Theorem IV.2 and Lemma IV.3, in what remains of this section, we describe how to implement the controlled Hamiltonian simulation (Claim IV.5), the block-encoding of the jump operators (Claim IV.6), and the controlled Boltzmann filter (Claim IV.7), in circuit depth $O(4^\ell\cdot 2^d\cdot\text{poly}(\log n, \log\frac{1}{\varepsilon}, \ell))$. While the first two optimizations are particular to our family of Hamiltonians, the latter may find independent application to the framework of [13].

We begin with a simple lemma which "colors" the interaction graph of the Hamiltonian, partitioning the interactions into disjoint subsets $S_1, S_2\cdots S_\Delta \subset [n]$ such that no two terms h_i, h_j of the same subset have overlapping support.

Lemma IV.4. *Any parent Hamiltonian $H\in\mathscr{H}$ defined by a quantum circuit of depth d and lightcone size ℓ can be Δ-colored with $\Delta\leq\ell\cdot 2^d+1$ colors.*

Proof. Two interactions h_i, h_j overlap at a qubit only if their lightcones intersect in the underlying circuit C, which determines H. Let ℓ_r denote the maximum "reverse lightcone" size of the circuit, that is, the maximum number of qubits which have a given qubit in their lightcone. Since the Hamiltonian is at most ℓ-local, any interaction h_i overlaps with at most $\ell\cdot\ell_r$ other terms, which in turn tells us the interactions can then be partitioned using $\Delta\leq\ell\cdot\ell_r+1$ different colors. If the depth of the circuit C as measured by layers of 2-qubit gates is d, then $\ell_r\leq\min(n, 2^d)$. □

Claim IV.5. *The controlled time-evolution of an n qubit parent Hamiltonian of a quantum circuit with lightcone size ℓ, can be implemented using a quantum circuit of depth $O(4^\ell\cdot\Delta\cdot\log n)$ and size $O(4^\ell\cdot\Delta\cdot n\cdot\log n)$.*

At a high level, the circuit of Claim IV.5 partitions the terms of the (commuting) Hamiltonian into disjoint subsets of non-overlapping terms, which can be implemented in parallel. However, since we need to implement the *controlled* Hamiltonian simulation, all of these Hamiltonian terms need to act conditioned on the time-register, which is a sequential bottleneck to the circuit depth. In order to further compress the depth, we parallelized the access to the time-register by encoding it into a GHZ state.

Proof. Since the Hamiltonian is commuting, let us restrict our attention to a fixed subset S in the partition guaranteed by Lemma IV.4. It suffices to prove how to implement the controlled time evolution of each subset of non-overlapping terms $H_S = \sum_{i\in S} h_i$. For this purpose, we begin by parallelizing the access to the clock register: $|\bar{t}\rangle \to |\bar{t}\rangle^{\otimes n}$, using a $O(\log n)$ depth circuit of CNOT gates, of size $O(n\log n)$.

Next, controlled on the jth clock register, we apply the time evolution of the jth interaction. Although this gate acts on

$O(\log n) + \ell$ qubits, it can be implemented via a sequence of $O(\log n)$ gates acting only on $\ell + 1$ qubits, by applying a binary expansion of the time register $\bar{t} = \frac{\pi}{n} \cdot \sum_k 2^k \bar{t}_k$:

$$\textstyle\sum_{\bar{t}} |\bar{t}\rangle\langle\bar{t}| \otimes e^{i\bar{t}h_j} = \prod_k \mathbb{I}_{\setminus k} \otimes \sum_{\bar{t}_k \in \{0,1\}} |\bar{t}_k\rangle\langle\bar{t}_k| \otimes \exp\big[i\frac{\pi}{n} \cdot 2^k \bar{t}_k \cdot h_j\big]. \tag{53}$$

In turn, each unitary on $\ell + 1$ qubits can generically be implemented in $O(4^\ell)$ size and depth. After all the colors have concluded, we revert the copies of the clock register. □

Claim IV.6. *A block encoding of the jump operators can be implemented using a quantum circuit of depth $O(\Delta \cdot (\ell + \log n))$ and size $O(n \cdot \Delta \cdot (\ell + \log n))$.*

Proof. There are $|A| = n \cdot 4^\ell$ jump operators, which we represent by indexing them using a pair $a = (i, P)$ in terms of the center of its support $i \in [n]$, using $\log n$ qubits, as well as the ℓ-local Pauli P, using $2 \cdot \ell$ qubits. Next we proceed using similar techniques to Claim IV.5. We begin by partitioning the jump operators into Δ disjoint subsets using Lemma IV.4, where any two jump operators in the same subset either have the same center, or do not intersect. Our implementation proceeds by addressing each color $c \in [\Delta]$ independently.

First, we create copies of the control register $|a\rangle \to |a\rangle^{\otimes n}$, to parallelize access to it. Suppose all the jump operators centered at j have been colored c. Our goal is to coherently apply all the controlled jump operators of the form (j, P) by acting only on the support (centered at j) and the jth control register $|a = (i, Q)\rangle_j$. For this purpose, we first check whether $i = j$, and controlled on the one-qubit outcome, we apply the Pauli Q. The check can be implemented using $O(\log n)$ size and depth, and the controlled Pauli in $O(\ell)$ size and depth. We conclude by inverting the checking and copying steps. □

Claim IV.7. *The controlled filter W can be implemented up to error ε in spectral norm using a circuit of size $O(\mathrm{polylog}(\frac{n}{\varepsilon}))$ 2-qubit gates.*

Proof. Let us denote $n_\delta = \beta^{-1} \log \frac{1}{\delta}$. Then, the Glauber dynamics weight $\gamma(\nu) = (1 + e^{-\beta\nu})^{-1}$ satisfies

$$\gamma(\nu) \leq \delta \qquad \text{if} \quad \nu \leq -n_\delta, \tag{54}$$

$$\text{and } \gamma(\nu) \geq 1 - \delta \quad \text{if} \quad \nu \geq n_\delta. \tag{55}$$

We claim that the W gate can be replaced by a truncation W_δ,

$$W_\delta = \sum_{\omega \in [-n, \cdots, n]} \begin{bmatrix} \sqrt{\tilde{\gamma}(\omega)} & -\sqrt{1 - \tilde{\gamma}(\omega)} \\ \sqrt{1 - \tilde{\gamma}(\omega)} & \sqrt{\tilde{\gamma}(\omega)} \end{bmatrix} \otimes |\omega\rangle\langle\omega| ,$$
$$\tilde{\gamma}(\omega) := \begin{cases} \gamma(\omega) & \text{if } \omega \in [-n_\delta, n_\delta] \\ 1 & \text{if } \omega > n_\delta \\ 0 & \text{if } \omega < -n_\delta \end{cases} . \tag{56}$$

Indeed, the truncation error is controlled by

$$\begin{aligned} &\|W - W_\delta\| \\ &\leq \sum_{j \in [-n, -n_\delta]} \| \begin{bmatrix} \sqrt{\gamma(\omega)} & 1 - \sqrt{1 - \gamma(\omega)} \\ \sqrt{1 - \gamma(\omega)} - 1 & \sqrt{\gamma(\omega)} \end{bmatrix} \| \\ &+ \sum_{j \in [n_\delta, n]} \| \begin{bmatrix} \sqrt{\gamma(\omega)} - 1 & -\sqrt{1 - \gamma(\omega)} \\ \sqrt{1 - \gamma(\omega)} & \sqrt{\gamma(\omega)} - 1 \end{bmatrix} \| \\ &\leq 2n \cdot (2\sqrt{\delta} + 2\delta) \leq 8n \cdot \sqrt{\delta}, \end{aligned} \tag{57}$$

where the last line uses that $1 - \sqrt{1 - x} \leq x$ when $x \in [0, 1/2]$.

It only remains now analyze the gate complexity of implementing W_δ. Following [13] (pg. 25, footnote 33), the W_δ filter for the Glauber weight between $[-n_\delta, n_\delta]$ can be implemented using the QSVT up to error ε using $\tilde{O}((1 + \beta n_\delta)\mathrm{polylog}\frac{1}{\varepsilon})$ 2-qubit gates. With the choice $\delta = O(\frac{\varepsilon}{n})$, we arrive at the advertised bounds by combining with the trivial cases $\omega \notin [-n_\delta, n_\delta]$. □

We remark that this error in spectral norm between unitaries is equivalent to the channel diamond norm distance, up to a constant: $\|U - V\|_\diamond \leq 2 \cdot \|U - V\|$.

Put together, Claim IV.5, Claim IV.6 and Claim IV.7 imply Lemma IV.1.

V. Low-Depth State Preparation on Lattices

We dedicate this section to a proof of Theorem V.1, on the preparation of Gibbs states of parent Hamiltonians of quantum circuits defined on lattices.

Theorem V.1. *Fix an inverse-temperature $\beta > 0$, and let H be the parent Hamiltonian of an n qubit, depth d, quantum circuit comprised of 2-qubit nearest neighbor gates in D dimensions. Then, there exists a quantum algorithm which prepares the Gibbs state of H up to error ε in depth $2^{O(d^D)} \cdot \mathrm{polylog}\frac{n}{\varepsilon}$.*

The algorithm of Theorem V.1 based on that of [26], who showed that if the Gibbs state satisfies two structural decay-of-correlations properties, then it can be efficiently prepared by a quantum computer. In particular, a decay of the conditional mutual information (CMI) and a certain clustering of correlations. In this section we show that parent Hamiltonians of low depth circuits satisfy strengthened versions of both of these properties, giving rise to a state preparation circuit of nearly constant depth.

A. The Markov Property

The condition mutual information is an information-theoretic measure of the correlations in a tripartite state ρ:

$$I(A : C|B)_\rho = S(AB)_\rho + S(BC)_\rho - S(ABC)_\rho - S(B)_\rho, \tag{58}$$

where $S(A)_\rho = -\operatorname{Tr} \rho_A \log \rho_A$ is the von Neumann entropy. Roughly, this quantity captures the mutual information between systems A and C, conditioned on system B.

The quantum states with vanishing conditional mutual information are known as quantum Markov chains, which is equivalent to the existence of a local recovery map that reconstructs ρ from just its subsystems. The following results also handle the cases with errors.

Theorem V.2 ([43])**.** *For any state ρ on ABC, there exists a quantum channel $\mathcal{R}_{B\to BC}$ such that*

$$I(A:C|B)_\rho \geq \frac{1}{4\ln 2}\|\rho - [\mathbb{I}_A \otimes \mathcal{R}_{B\to BC}](\rho_{AB})\|_1^2. \quad (59)$$

In the special case of $I(A:C|B)_\rho = 0$, the recovery map is explicit and can be chosen to be the Petz (or transpose) map [42], [44]:

$$\mathcal{R}_{B\to BC}[X] = \rho_{BC}^{1/2}\big(\rho_B^{-1/2} X \rho_B^{-1/2} \otimes \mathbb{I}_C\big)\rho_{BC}^{1/2} \quad (60)$$

In the context of this work, we require characterization of the CMI of the Gibbs states of commuting Hamiltonians. Fix a local commuting Hamiltonian H defined on a set of vertices (the qubits) and hyper-edges (the interactions). Given three disjoint subsets A, B, C of the qubits of H, we say that B *shields* A from C if all paths on the hyper-edges from A to C must pass through B. For instance, the boundary of a region C, comprised of the qubits which share an interaction with C, shield C from the rest of the lattice/qubits.

The Quantum Hammersley-Clifford Theorem states that the Gibbs states of commuting local Hamiltonians form quantum Markov Chains, for *shielding* tripartitions.

Fact V.1 (The Quantum Hammersley-Clifford Theorem, [45])**.** *Let H denote a commuting local Hamiltonian and A, B, C three disjoint subsets of the qubits of H such that B shields A from C. Then,*

$$I(A:C|B)_\rho = 0, \text{ where } \rho = e^{-H}/\operatorname{Tr} e^{-H}. \quad (61)$$

We remark that parent Hamiltonians of depth d quantum circuits have local interactions of range $2\cdot d$. In this manner, the boundary of any region C on the lattice is comprised of all the qubits at distance $\leq 2\cdot d$ from C, which aren't already in C.

B. Local Indistinguishability

Given a local Hamiltonian H defined on a lattice Λ, and a subset of said lattice $X \subset \Lambda$, we refer to the subsystem Hamiltonian H_X as all the interactions contained *entirely* in X[19]:

$$H_X = \sum_{e\subset X} h_e. \quad (62)$$

Henceforth let us denote as $\rho \propto e^{-\beta H}$ the Gibbs state of H, and $\rho^X \propto e^{-\beta H_X}$ the Gibbs state of the subsystem Hamiltonian.

Definition V.1 (Local indistinguishability)**.** *Let H be a local Hamiltonian defined on a lattice Λ, β an inverse temperature, and $ABC = X \subset \Lambda$ be a partition of a subset of the lattice. Then the Gibbs state of H is said to satisfy* local indistinguishability *on ABC if*

$$\operatorname{Tr}_{BC}\rho^X = \operatorname{Tr}_B \rho^{AB}. \quad (63)$$

That is, the thermal expectation of local observables for a global Hamiltonian can be evaluated using only knowledge of a local Hamiltonian patch. This notion is closely related to the stability of gapped phases [46]. [26] show that an approximate version of the above follows if one assumes a clustering property for the correlations in the Gibbs state of H.

In the context of this work, we claim that the structure of parent Hamiltonians of low depth circuits implies an exact local-indistinguishability property for their associated Gibbs states.

Claim V.3 (Exact local indistinguishability for our parent Hamiltonians)**.** *Let C be a quantum circuit comprised of d layers of nearest neighbor gates on a D dimensional lattice Λ, and H its associated parent Hamiltonian. Then, the Gibbs state of H satisfies local indistinguishability for arbitrary $ABC = X \subset \Lambda$ such that the distance $d(A, C) \geq 4\cdot d + 1$.*

Here we define the distance between subsets of the lattice $d(A, C) = \min_{i\in A, j\in C} d(i, j)$, defined by the path-length over edges of the lattice.

Proof. We claim that the unitary U comprised of all the gates in the reverse-lightcone of B i.e, the union of all the lightcones L_i which are entirely contained in B, disentangles A from C in the Gibbs state:

$$(U^\dagger \otimes \mathbb{I})\rho^X(U \otimes \mathbb{I}) = \sigma_{AB_1} \otimes \gamma_{B_2C}. \quad (64)$$

This decoupling implies the desired property. To begin, note that the depth d of the circuit C implies a bound of $\leq 2d$ on the range of interaction of the terms in the Hamiltonian H. In this manner, if $d(A, C) \geq 2\cdot d + 1$ then B screens A from C in the subsystem Hamiltonian H_X.

Suppose we partition the qubits of B into B_A, all qubits of B at distance $\leq d$ from A, B_C, those at distance $\leq d$ from C, and $B_- = B \setminus (B_A \cup B_C)$. Since $d(A, C) \geq 4d + 1$, no two terms h_i, h_j, $i \in A \cup B_A$ and $j \in C \cup B_C$ have intersecting support. Moreover, for $k \in B_-$ such that the lightcone $L_k \subset B$,

$$U^\dagger h_k U = U^\dagger C(|1\rangle\langle 1|_k \otimes \mathbb{I}_{[n]\setminus\{k\}})C^\dagger U = |1\rangle\langle 1|_k, \quad (65)$$

and therefore other $U^\dagger h_i U$ has support on any element of B_-. Thus, the unitary U partitions the Hamiltonian into non-interacting components:

$$U^\dagger H_X U = H_{AB_A} + H_{B_-B_CC}. \quad (66)$$

To conclude, we observe that $U^\dagger \rho^X U$ is the Gibbs state of $U^\dagger H_X U$. □

We remark that the claim above can alternatively be seen as a consequence of both clustering of correlations and the quantum belief propagation equations with 0 error [26], [47].

C. The Algorithm

We can now invoke a much simplified version of the algorithm of [26], on the preparation of Gibbs states of Hamiltonians of D-dimensional lattices. Their algorithm is based on

[19] i.e., the interactions act trivially outside subsystem X

the operation of recovery maps on finite-sized regions, namely up to a certain correlation length-scale ℓ_0. To instantiate their algorithm, it remains to show how to perform these recovery maps.

Fortunately, we can appeal to the explicit structure of the Petz map of Eq. (60). Via a Stinespring dilation[20], each such recovery channel can be purified into a quantum circuit on a doubled Hilbert space. In turn, said circuit can be implemented in exponential time in the volume (the number of qubits) of the region, up to factors of $\mathrm{polylog}\frac{n}{\varepsilon}$. See also [49] for a more general discussion on implementing Petz recovery maps.

Theorem V.4 ([26], Theorem 6, Simplified)**.** *Fix $\beta > 0$, let H be a bounded local Hamiltonian defined on D dimensional lattice Λ, and let ρ be its Gibbs state at inverse-temperature β. If ρ is both locally indistinguishable and exactly Markov for all tripartitions $ABC = X \subset \Lambda$ such that $d(A, C) \geq \ell_0$, then a purification of ρ can be prepared up to error ε using a quantum circuit of depth $2^{O(\ell_0^D)} \cdot \mathrm{polylog}\frac{n}{\varepsilon}$.*

In full generality, [26]'s result applies to Gibbs states which are approximately Markov and approximately locally indistinguishable. However, the typically exponentially-decaying tail in the approximation error implies quasi-polynomial-depth state preparation algorithms. By appealing to their exact counterparts, we arrive at a polylog depth state-preparation algorithm. To conclude Theorem V.1, we note that $\ell_0 = O(d)$ from Fact V.1 and Claim V.3.

VI. The Input Noise Model and Gibbs States of Quantum Circuits

In this section, we show that the Gibbs states of parent Hamiltonians of quantum circuits correspond to noisy versions of the output of the quantum circuit, under a certain input noise model. To begin, let us recollect the noise model. Fix a noise rate $p \in (0, 1)$. The single-qubit bit-flip error channel consists of the superoperator

$$\mathcal{D}_p(\sigma) = (1-p)\cdot\sigma + p\cdot X\sigma X. \tag{67}$$

Given a quantum circuit C on n qubits, the input noise model consists of independent applications of the bit-flip error channel on the input wires of C. In particular, the mixed state given by the output of the noisy circuit is:

$$\rho = C\Big(\mathcal{D}_p(|0\rangle\langle 0|)\Big)^{\otimes n} C^\dagger \tag{68}$$

For a fixed n qubit quantum circuit C, recall that we refer to the parent Hamiltonian of C as

$$H_C = C\Big(\sum_{i\in[n]} |1\rangle\langle 1|_i \otimes \mathbb{I}_{[n]\setminus i}\Big)C^\dagger \tag{69}$$

Lemma VI.1. *Fix $\beta > 0$, and let H_C be the parent Hamiltonian of a quantum circuit C. The Gibbs state of H_C at inverse-temperature β is given by the output of the circuit C under input level noise with probability $p = (1+e^\beta)^{-1}$:*

$$\rho_\beta = \frac{e^{-\beta H_C}}{\mathrm{Tr}\, e^{-\beta H_C}} = C\Big(\mathcal{D}_p(|0\rangle\langle 0|)\Big)^{\otimes n} C^\dagger \tag{70}$$

Proof. It suffices to consider the Gibbs state σ_β of the Hamiltonian $H = \sum_{i\in[n]} |1\rangle\langle 1|_i$, as $\rho_\beta = C\sigma_\beta C^\dagger$. Since H is commuting, the partition function can be written as:

$$\mathrm{Tr}\, e^{-\beta H_C} = \mathrm{Tr}\, e^{-\beta H} = \sum_{x\in\{0,1\}^n} \prod_i^n \langle x_i| e^{-\beta |1\rangle\langle 1|_i} |x_i\rangle \tag{71}$$

$$= \prod_i^n \sum_{x_i\in\{0,1\}} \langle x_i| e^{-\beta |1\rangle\langle 1|_i} |x_i\rangle = (1+e^{-\beta})^n. \tag{72}$$

Therefore, the Gibbs state of H can be expressed as the outcome of the depolarizing channel:

$$\begin{aligned}\sigma_\beta &= (1+e^{-\beta})^{-n}\cdot e^{-\beta H} \\ &= \bigotimes_i^n \Big(\frac{|0\rangle\langle 0|}{1+e^{-\beta}} + \frac{|1\rangle\langle 1|}{1+e^{\beta}}\Big) \\ &= \Big(\mathcal{D}_p(|0\rangle\langle 0|)\Big)^{\otimes n},\end{aligned} \tag{73}$$

with $p = (1+e^\beta)^{-1}$.

□

VII. Computational Complexity of Shallow IQP Sampling

In recent years several architectures have been proposed for achieving a quantum speedup, based on quantum processes which resemble or are equivalent to the IQP Circuit Sampling task discussed in Section II-B. The basis for these speedups is on standard complexity-theoretic conjectures, including the non-collapse of the Polynomial Hierarchy, often in addition to strong assumptions on the hardness of computing permanents or partition functions. We dedicate this section to a discussion on the background behind Theorem II.5, as well as a comparison to related statements in the literature.

To begin, let us recollect the circuit described in Section II-B, comprised of a 2D cluster state and random phase gates [5], [4] (see Fig. 4).

If instead of *random* powers of single-qubit T gates, the powers were chosen adaptively given partial measurements of the circuit, this scheme would implement measurement-based quantum computation [28]. The universality of MBQC under adaptivity (or post-selection) implies the hardness of *exactly* sampling from the output distribution, unless the polynomial hierarchy collapses to the third level [32], [4]. To reproduce their argument, universality implies

$$\mathsf{PostIQP} \underbrace{=}_{[32]} \mathsf{PostBQP} \underbrace{=}_{[50]} \mathsf{PP}. \tag{75}$$

If we now assume there existed a classical algorithm to exactly sample from arbitrary IQP circuits, that would imply $\mathsf{PP} =$

[20] See e.g. Chapter 2 of [48]

1) Prepare an n qubit cluster state on a 2D rectangular lattice using a layer of Hadamard gates and 4 layers of CZ gates.
2) Sample a random string $b \in [7]^n$, and apply powers of T gates to each qubit:

$$\bigotimes_{i\in[n]} T^{b_i} \left(\prod_{<i,j>} CZ_{i,j} \left|+\right\rangle^{\otimes n} \right) = C_b \left|0\right\rangle^{\otimes n}, \tag{74}$$

where $T = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$.
3) Finally, measure the output in the X basis.

Fig. 4: A family of random IQP circuits, $\{C_b\}$.

$\mathsf{PostIQP} \subseteq \mathsf{PostBPP}$, which in turn gives us a collapse of the Polynomial Heirarchy (henceforth, PH):

$$\mathsf{PH} \underbrace{=}_{\text{Toda's Theorem}} P^{\mathsf{PP}} \underbrace{=}_{\text{By assumption}} P^{\mathsf{PostBPP}} = \Sigma_3. \tag{76}$$

In fact, by similar reasoning [32] (Theorem 2) showed that no classical algorithm can even *weakly* approximately sample from IQP circuits - i.e. up to some fixed multiplicative error. To extend these hardness results to approximate sampling (up to some additive error) in total variation distance, we require stronger assumptions.

[36] were the first to show that, assuming an additional complexity-theoretic conjecture on the average-case hardness of computing partition functions, approximately sampling from the output of IQP circuits remains classically intractable even up to small total variation distance. They noted that the output distribution of IQP circuits,

$$p_x = |\left\langle x\right| C \left|0\right\rangle^{\otimes n}|^2 = 2^{-n} \cdot \left|\mathcal{Z}_x\right|^2, \tag{77}$$

precisely resembles a complex-valued partition function, defined by $w_{u,v}, w_u$ real-valued edge and vertex weights on some underlying architecture graph G:

$$\mathcal{Z}_x = \sum_{z\in\{\pm1\}^n} \exp\left[i\left(\sum_{<u,v>} w_{uv} z_u z_v + \sum_u (\pi \cdot x_u + w_u) z_u \right]. \tag{78}$$

They prove that approximating $\left|\mathcal{Z}_x\right|^2$, and therefore p_x, up to multiplicative error is $\#P$ hard in the worst-case, and pose as a conjecture its hardness in the average case over x. Under this conjecture, [36] show that the existence of an efficient classical algorithm to approximately sample from $\{p_x\}$, even up to constant TVD, would imply a collapse of the polynomial hierarchy.

However, the original results of [36] referred to a complete graph G, which, roughly speaking, correspond to IQP circuits of some polynomial depth. In follow up work by the same authors [20], they reduced the circuit depth to logarithmic under a sparsified version of the graph G. It was only in [4] and [5] that the $\#P$ hardness of approximately computing p_x on 2D circuit architectures was established (in the worst-case), corresponding to constant depth IQP circuits in 2D. Their analogous average-case conjecture for approximately computing p_x on 2D circuits, is reproduced below:

Conjecture 1 ([4]). *There exists a choice of vertex and edge weights $\{w_{uv}, w_u\}_{u,v\in[n]}$ on a 2D lattice G, and constants ε, δ, such that approximating the measurement distribution $\{p_x\}$ to the following mixture of multiplicative and additive errors*

$$|\tilde{p}_x - p_x| \leq \frac{1}{\text{poly}(n)} \cdot p_x + \frac{\varepsilon}{\delta \cdot 2^n} \tag{79}$$

is $\#P$ hard for any $1-\delta$ fraction of instances x.

[4] show that Conjecture 1 implies Theorem II.5:

Theorem VII.1 ([4], restatement of Theorem II.5)**.** *Assuming Conjecture 1, simulating the distribution $\{p_x\}$ up to ε total variation distance is classically intractable, assuming* PH *doesn't collapse.*

A related result was shown by [5]. They start from the (weaker) conjecture that computing $p(x)$ up to a multiplicative factor is hard-on-average, and combine it with a further conjecture on the anti-concentration of the output distribution of random linear-depth IQP circuits. Put together, they also arrive at Theorem II.5.

VIII. Fault Tolerance of IQP Circuits under Input Noise

We dedicate this section to a proof of Lemma I.3, on the fault tolerance of IQP circuits under input noise.

Lemma VIII.1. *Fix an input noise rate $p < \frac{1}{2}$ and a positive integer D. Let C be an n qubit IQP circuit with depth d and lightcone size ℓ. Then, there exists another quantum circuit $\tilde{C}$, such that a sample from the output of $\tilde{C}$ under input bit-flip errors can be post-processed using an efficient classical algorithm into a sample ε-close to the output distribution of C. The circuit $\tilde{C}$*

1) *acts on $O(n \log \frac{n}{\varepsilon})$ qubits,*
2) *has lightcone size $\ell + O\left(D \log^{1/D}\left(\frac{n}{\varepsilon}\right)\right)$,*
3) *depth $d + O\left(D \log^{1/D}\left(\frac{n}{\varepsilon}\right)\right)$, and*
4) *the locality of its parent Hamiltonian is $\ell + O(D)$.*

For any sufficiently large constant D, we recover the claimed fault-tolerance result of Lemma I.3. If $D = O(\log\log \frac{n}{\varepsilon})$, then the circuit depth, lightcone size, and locality are all increased by an additive $O(\log\log \frac{n}{\varepsilon})$ factor.

At a high level, our approach is based on pre-processing each of the n input bits into "code-blocks" or gadgets of size $k = O(\log \frac{n}{\varepsilon})$ bits, where each gadget has a designated "root" bit. The n root bits are then input into the IQP circuit C. Since bit-flip errors commute with the IQP circuit, to be able to sample from the original output distribution of C, it suffices to *identify* these root bits. Indeed, we emphasize that we do not use the encoding to *correct* the errors within the circuit, as this would require adaptivity and an increase in circuit depth, and instead perform the correction only in post-processing.

A. The Distillation Gadget

We place the noisy bits into a tree of arity B (a "B-tree") of depth D. For notational convenience, let us partition the nodes in the tree into disjoint subsets, $L_1 \cup L_2 \cdots \cup L_D = [k]$, the "layers" of the tree. Moreover, for each node u in the tree, let the subset N_u denote its children or (downwards) neighbors in the tree. The encoding circuit proceeds over the layers from the leaves to the root, where at the ith layer L_i of the tree, a CNOT gate is applied from each parent bit to each of its children.

Note that the size of the tree k is implicitly defined by B and d: $k = \sum_{j=0}^{D-1} B^j = \Theta(B^D)$.

Algorithm 1: The Distillation Gadget U

```
Input: $k$ qubits in the computational basis $|s\rangle$, where
        $s \leftarrow \mathrm{Ber}^k(p/2)$.
For each layer $i \in [2, \cdots, D]$ from leaves to root,
For each child $c \in N_p$ of a parent node $p$, apply a
    $\mathsf{CNOT}$ gate from $p$ to $c$.
```

$$\prod_{i\in[D]} \bigotimes_{p\in L_i} \left(\prod_{c\in N_p} \mathsf{CNOT}_{p,c} \right) |s\rangle \equiv U\,|s\rangle\,.$$

We emphasize that the ordering of operations, from leaves to root, matters crucially. In this manner, the ith layer acts as a "parity check syndrome" for the $(i+1)$st. When implemented using 2-qubit gates, the depth of the distillation circuit is $B \cdot D$, as the CNOT gates at the same layer but operating on different subtrees can be performed in parallel, but the B CNOT gates which act on the same parent must be performed sequentially.

B. The Decoding Algorithm

Next, suppose that all the qubits of $U\,|s\rangle$ except for that at the root of the tree have been measured, resulting in bits $b_2, \cdots b_k$. Can we reconstruct s_1, the bit at the root? The decoding algorithm below traverses the tree layer by layer, from leaves to root, attempting to reconstruct the bit s_p of the next layer.

Algorithm 2: The Decoding Algorithm

```
Input: $(k-1)$ bits $b_2, \cdots, b_k$, organized into a
        $B$-tree, where the root bit has been removed.
Output: A single bit $\tilde{s}_1$, a guess for the bit at the root.
At the leaves $L_1 \subset [k]$, let us denote $\tilde{b}_u = b_u$ for $u \in L_1$.
For each layer $i \in [2, \cdots, D]$, from leaves to root,
For each parent node $p \in L_i$, let $\tilde{s}_p = \mathrm{Maj}(\tilde{b}_c : c \in N_p)$
    be the majority of its children bits.
If the root hasn't been reached, update $\tilde{b}_p \leftarrow \tilde{s}_p \oplus b_p$.
    Otherwise, output $\tilde{s}_1$.
```

The decoding algorithm above maintains the invariant that $\tilde{s}_u$ is a "guess" for the original noisy bit s_u input into the distillation gadget. Since U acts from leaves to root, the children in each layer contain (with high probability) the necessary information to reconstruct the parents' bit s_p. Together with the measurement outcome b_p - which reveals information about the layer above - we can continue the reconstruction up the tree.

C. Analysis

We divide the analysis into three claims, which consider the correctness, the lightcone size of the circuit, and the "Z-locality" of the distillation gadget which determines the locality of the parent Hamiltonian.

Claim VIII.2 (Correctness)**.** *Fix any noise rate $p \leq \frac{1-\delta}{2}$ and let $B = \Omega(\delta^{-2})$. Then, the effective bit-flip error rate at the root of the depth d B-tree is $\leq 2^{-B^{\Omega(d)}}$.*

Proof. We prove inductively that the *effective* bit-flip error rate p_i at the ith layer, i.e.,

$$p_i \equiv \mathbb{P}_{s \leftarrow \mathsf{Bern}^k(p)}[s_u \neq \tilde{s}_u] \text{ for each node } u \in L_i, \tag{80}$$

decays doubly-exponentially with the layer index $i > 2$. As the base case, $p_1 = p$ is the probability of a bit-flip error on the leaves. Suppose $p = \frac{1-\delta}{2}$. Then, after the first layer, the probability the majority vote of the children bits is incorrect is

$$\begin{aligned} p_2 &\leq \sum_{j=B/2}^{B} \binom{B}{j} \big(p_1\big)^j \big(1-p_1\big)^{B-j} \\ &= 2^{-B} \sum_{j=B/2}^{B} \binom{B}{j} \big(1-\delta\big)^j \big(1+\delta\big)^{B-j} \\ &\leq \big(1-\delta\big)^{B/2} \cdot \big(1+\delta\big)^{B/2} \leq \big(1-\delta^2\big)^{B/2} \leq \frac{1}{16}, \end{aligned} \tag{81}$$

so long as B is chosen to be $\Omega(\delta^{-2})$. For each layer $i \geq 2$, the effective bit-flip error rate on the $(i+1)$st layer is

$$p_{i+1} \leq \textstyle\sum_{j=B/2}^{B} \binom{B}{j} \big(p_i\big)^j \cdot \big(1-p_i\big)^{B-j} \leq 2^B (p_i)^{B/2} \leq p_i^{B/4}. \tag{82}$$

In this manner, $p_{i+1} \leq 2^{-(B/4)^i}$ for $i \geq 1$.

□

Claim VIII.3 (Circuit lightcone size)**.** *The circuit lightcone size of the distillation scheme is $\leq B \cdot D$.*

Proof. The lightcone size of the quantum circuit U is upper bounded by the size of the lightcone of the qubits at the leaves of the tree. Crucially, we claim that if

$$u = u_1 \rightarrow u_2 \rightarrow u_3 \cdots \rightarrow u_D = \text{root} \tag{83}$$

denotes the path from a leaf $u \in L_1$ to the root, then only the children of these nodes can be in the lightcone of u. Indeed, this is since the CNOT gates in Line 1 are applied layer by layer in increasing order, so the only nodes which are causally connected to u in the circuit are its immediate ascendants or their neighbors. In turn, the size of this set is bounded by $B \cdot D$.

□

The last key claim makes reference to the locality of the parent Hamiltonian of the distillation circuit, that is, the size of the support of the operator $U(Z_i \otimes \mathbb{I})U^\dagger$, maximized over bits i in the gadget. We thank Joel Rajakumar and James Watson for the observation that the locality of the Hamiltonian is only related to the propagation of Pauli-Z instead of the full circuit lightcone (see also [3]).

Claim VIII.4 (Parent Hamiltonian Locality)**.** *The locality of parent Hamiltonian of the distillation circuit is $\leq D$.*

Proof. The following two circuit identities describe how Pauli Z operators propagate through CNOT gates.

$$\mathsf{CNOT}_{i,j}(Z_i \otimes \mathbb{I})\mathsf{CNOT}_{i,j} = Z_i \otimes \mathbb{I} \tag{84}$$

$$\mathsf{CNOT}_{i,j}(\mathbb{I} \otimes Z_j)\mathsf{CNOT}_{i,j} = Z_i \otimes Z_j \tag{85}$$

Crucially, the locality only increases (or propagates) from the target qubit to the control qubit. Applied to our gadget in Line 1, we conclude that the qubits in the Z-lightcone of any qubit i in the tree, are precisely the ancestors of i. Thus, $|\text{supp}(U(Z_i \otimes \mathbb{I})U)| \leq D$, the depth of the tree. □

We are now in a position to conclude the proof of Lemma VIII.1.

Proof. [of Lemma VIII.1] By Claim VIII.2, if $p \leq \frac{1}{2}(1-\delta)$, then, so long as

$$B = \max\left(\Theta(\delta^{-2}), \log^{1/D}\left(\frac{n}{\varepsilon}\right)\right) \tag{86}$$

the probability the decoding algorithm incorrectly outputs the bit at the root of the tree is $\leq \varepsilon n^{-1}$. By a union bound, all the gadgets succeed with probability $\geq 1-\varepsilon$. Conditioned on this event, the output distribution of $\tilde{C}$ corrected by the output of the n decoding algorithms is exactly that of C, which implies the bound on the TV distance. To conclude, the locality parameters are then implied by Claim VIII.3 and Claim VIII.4

□

IX. Quantum Advantage in Gibbs Sampling

We dedicate this section to combining all the aforementioned ingredients and concluding the proof of our main result in Theorem I.1.

Theorem IX.1 (General version of Theorem I.1)**.** *For any constant inverse-temperature $\beta = \Theta(1)$ and integer L, there exists a family of n-qubit commuting $O(L)$-local Hamiltonians, such that the n-qubit Gibbs state ρ_β is both*

1) Rapidly Thermalizing. *It can be prepared within small trace distance by the Davies generator (Eq. (6)) which has mixing time $e^{O(L\cdot\log^{1/L}(n))}$. In addition, this process can be simulated on a quantum computer in time $n \cdot e^{O(L\cdot\log^{1/L}(n))}$. And yet,*
2) Classically Intractable. *Under Conjecture 1, there is no polynomial time classical algorithm to sample from the measurement outcome distribution $p(x) = \langle x|\,\rho_\beta\,|x\rangle$ within small constant total variation distance.*

In particular, the choice of a sufficiently large constant L recovers our main result of Theorem I.1. When $L = \log\log n$, we obtain a mixing time of $\text{polylog}(n)$.

Proof of Theorem IX.1. To begin our proof, let us fix an inverse-temperature $\beta = \Theta(1)$, and consider the equivalent bit-flip error rate

$$p = (1+e^\beta)^{-1} < \frac{1}{2}, \tag{87}$$

as guaranteed by Lemma VI.1.

a) Classical Intractability: Consider the family of constant-depth, classically intractable, n-qubit IQP circuits C guaranteed by Theorem II.5 (Conjecture 1). Using Lemma I.3, let us fix a depth parameter L, and embed each circuit in said family into a new circuit $\tilde{C}$, which is fault tolerant to input noise of rate $p = \frac{1}{2}(1-\Theta(1))$. $\tilde{C}$ now has Z-locality $O(L)$, circuit depth and lightcone size $O(L\log^{1/L}(\frac{n}{\varepsilon}))$; and a noisy sample from $\tilde{C}$ can be efficiently classically post-processed into a sample ε-close in trace distance to an ideal sample from C.

Now, consider the family of parent Hamiltonians defined by the family of Fault-Tolerant circuits $\tilde{C}$,

$$H = \sum_i \tilde{C}\Big(Z_i \otimes \mathbb{I}_{[n]\setminus i}\Big)\tilde{C}^\dagger. \tag{88}$$

The support size of each term is given by the Z-locality of the fault-tolerant circuit $\tilde{C}$, which is $O(L)$.

If, by assumption, there was a polynomial time classical algorithm $\mathcal{A}$ to sample from the Gibbs state of H at inverse-temperature β, then we could construct a polynomial time classical algorithm to sample from a distribution ε-close to the ideal distribution of C, as follows: First, construct $\tilde{C}$ and thus the local terms of H from C. Then, leverage $\mathcal{A}$ to sample from $\propto e^{-\beta H}$. Finally, process the output sample using the post-processing algorithm from the fault-tolerance statement of Lemma I.3.

b) Rapid Thermalization: To conclude, via Lemma I.2, the Gibbs state of H can be prepared using the Davies generator of Eq. (6) of mixing time exponential in the circuit lightcone size, $\log n \cdot \exp\Big(O(L\cdot\log^{1/L}(n))\Big) = \exp\Big(O(L\cdot\log^{1/L}(n))\Big)$. To simulate this process on a quantum computer, the overall runtime $n\cdot\exp\Big(O(L\cdot\log^{1/L}(n))\Big)$ has an additional quasi-linear overhead. □

Remark IX.1. *Theorem I.1 asserts that for every constant temperature, there exists a Hamiltonian H which is classically hard-to-sample from. Conversely, results by [6] and [7] show that every local Hamiltonian (of fixed degree) has a critical* *temperature, such that above said threshold one can efficiently classically sample from their Gibbs state. The resolution to this apparent contradiction lies in the order of quantifiers. The degree/locality of our Hamiltonians increases with the temperature, see Section VIII for their dependence on the noise rate.*

Remark IX.2. *Since the Gibbs state is determined by a low depth quantum circuit C, with access to a description of C, one could trivially produce it on a quantum computer. However, if given access only to the local Hamiltonian terms $\{h_i\}_i = \{-CZ_iC^\dagger\}_i$, we do not believe it to be computationally efficient to recover the global structure of C, in general. While this is not a rigorous statement, we only know how to do so for 1D circuits, via dynamic programming. It is worthwhile to contrast this to the Feynman-Kitaev circuit-to-Hamiltonian mapping [11], wherein the gates of the circuit can be exactly read-off from the local Hamiltonian interactions.*

X. BQP Completeness with Adaptive Single-Qubit Measurements

We dedicate this section to a proof of Theorem I.4, on the BQP completeness of Gibbs Sampling with adaptive measurements.

Theorem X.1. *Fix an inverse-temperature $\beta = \Theta(1)$. Then, there exists an n-qubit $O(1)$-local Hamiltonian, whose Gibbs state at inverse-temperature β is a universal resource state for quantum computation and is efficiently preparable on a quantum computer.*

This result is all but a corollary of our fault tolerance techniques for IQP circuits, applied to measurement-based quantum computation. Indeed, it is well known that 2D cluster states, in addition to single-qubit measurements in adaptively chosen basis on the $X - Y$ plane, is universal for quantum computation. The following lemma shows that one can produce said cluster state out of the Gibbs state of a local Hamiltonian, so long as we are allowed to measure a subset of the qubits, and subsequently apply a Pauli correction to "distill" out the cluster state.

Lemma X.2. *There exists a n-qubit, $O(1)$-local commuting Hamiltonian, whose Gibbs state at inverse-temperature β can be used to prepare a cluster state. That is, by measuring a subset of the qubits of the Gibbs state, and then with 1 round of adaptive Pauli correction, one can produce a 2D cluster state on $O(n/\log\frac{n}{\varepsilon})$ qubits with probability $1-\varepsilon$.*

Proof. Let C be the circuit which prepares a 2D cluster state on m qubits, comprised of Hadamard gates and CZ gates. Let $\tilde{C}$ be the $n = \Theta(m\log\frac{m}{\varepsilon})$ qubit circuit defined by the fault tolerance scheme of Lemma VIII.1, which is robust to input errors of finite probability $< \frac{1}{2}$. Then, consider the parent Hamiltonian H of $\tilde{C}$, on n qubits and with locality $O(1)$.

By construction, its Gibbs state is a quantum-classical state, of classical bits lying in the fault-tolerance gadget of Lemma VIII.1, and qubits comprising a cluster-state under input noise. Again, recall that input bit-flip errors are equivalent to output Z errors, due to the gate structure of C. From Lemma VIII.1, by measuring the classical bits of the fault-tolerance gadget, one can recover the output Z error with probability $1-\varepsilon$. □

We remark that the adaptively chosen $X-Y$ measurements can be performed simultaneously with the Pauli corrections. In this manner, after producing the desired resource Gibbs state, it suffices to perform adaptively chosen single-qubit measurements to achieve universal measurement based quantum computation.

XI. Addressing Output Measurement Errors

In this section, we prove Theorem I.5 on sampling from finite-temperature Gibbs states subject to measurement errors.

Lemma XI.1. *Fix an inverse temperature $\beta = \Theta(1)$, and a measurement error rate $p < \frac{1}{2}$. There exists a family of n-qubit, $O(\log n)$-local Hamiltonians, such that sampling from their Gibbs state at inverse-temperature β, under measurement errors of rate p, is classically intractable under Theorem II.5. Moreover, there exists a $\mathrm{poly}(n)$ time quantum algorithm to produce said Gibbs state.*

Our construction of Lemma XI.1 is similarly based on the parent Hamiltonians of fault-tolerant IQP circuits, which are hard-to-sample from in the ideal case. We note that the distribution defined by sampling from the Gibbs state of the parent Hamiltonian of a quantum circuit C, given measurement errors, corresponds exactly to sampling from C under both input and output noise, albeit with different noise rates. Unfortunately, to address this mixed noise model, we do need to appropriately modify our fault-tolerance scheme. For this purpose, we appeal to prior work by [20], at the cost of a higher locality.

A. Overview

To model the noise in this section, recall the definition of the bit-flip error channel $\mathcal{D}_p$ in Eq. (67). Given a quantum circuit C on n qubits, and fixed noise rates $p_{in}, p_{out} \in [0, \frac{1}{2})$, the noisy output distribution of C given input and output noise is given by

$$p_{C,p_{in},p_{out}}(x) = \mathrm{Tr}\left[|x\rangle\langle x| \cdot \mathcal{D}_{p_{out}}^{\otimes n} \circ C\Big(\mathcal{D}_{p_{in}} \circ (|0\rangle\langle 0|)\Big)^{\otimes n} C^\dagger\right]. \quad (89)$$

If $\mathcal{A} : \{0,1\}^n \to \{0,1\}^{n'}$ is a deterministic classical post-processing algorithm, we denote as $\mathcal{A} \circ p$ the distribution given by sampling $x \leftarrow p$ and outputting $\mathcal{A}(x)$. The following lemma is a fault-tolerance statement for IQP circuits against this input/output noise model.

Lemma XI.2. *Let C be an n qubit IQP circuit of depth d and lightcone size ℓ, and fix input and output bit-flip error rates $p_{in}, p_{out} \in [0, \frac{1}{2})$. Then, for every $r \in \mathbb{N}$ there exists a quantum circuit C_r and a deterministic, $O(n_r)$-time decoding* *algorithm $\mathcal{A}_r : \{0,1\}^{n_r} \to \{0,1\}^n$, such that in the presence of input and output noise, the statistical distance*

$$\|\mathcal{A}_r \circ p_{C_r,p_{in},p_{out}} - p_{C,0,0}\|_1 \leq n \cdot (4q(1-q))^{r/2}, \tag{90}$$

where $q = p_{in}(1-p_{out}) + p_{out}(1-p_{in}) < \frac{1}{2}$. Moreover, C_r is defined on $n_r = n \cdot r$ qubits, has depth $d_r = d \cdot \log r$ and lightcone size $\leq \ell \cdot r$.

In other words, noisy samples from C_r can be post-processed into nearly-ideal samples from C. Note that $q < \frac{1}{2}$ implies the total variation distance above decays exponentially with r.

Corollary XI.3. *Fix input and output bit-flip error rates $< \frac{1}{2}$. Then, any IQP circuit on n qubits and constant depth can be efficiently transformed into a quantum circuit of $O(\log\log n)$ depth and $O(\log n)$ lightcone size, robust to input and output noise with error $n^{-\Omega(1)}$.*

Starting from the hard-to-sample IQP circuits ensured by Theorem II.5, we can construct circuits fault-tolerant to input and output noise via the Corollary above. In turn, these fault-tolerant circuits define a parent Hamiltonian, which is rapidly thermalizing (via Lemma I.2), and yet, classically hard to sample from. Put together, we prove Lemma XI.1.

B. Analysis

We remark that if the circuit C itself is an IQP circuit, then the bit-flip noise model $\mathcal{B}_p$ commutes with the circuit, and thus the input/output noise is equivalent to input noise at a higher rate: $p_{C,p_{in},p_{out}}(x) = p_{C,q,0}(x)$, with

$$q = p_{in}(1-p_{out}) + p_{out}(1-p_{in}) < \frac{1}{2} \tag{91}$$

To leverage this equivalence, however, we need to design a fault-tolerant circuit which itself is an IQP circuit. Fortunately, here we can appeal to [20], who achieved precisely that. To summarize their construction, their circuit embedding leverages the following property of IQP circuits. The diagonal part D of any IQP circuit can be expressed as a matrix-exponential of a polynomial of Z Pauli matrices:

$$D = \exp\left[i \sum_{j\in[m]} \theta_j \bigotimes_{i\in[n]} Z_i^{M_{ji}}\right], \tag{92}$$

for real coefficients $\{\theta_j\}$, and a boolean matrix $M \in \mathbb{F}_2^{m\times n}$. If D is comprised of 2-qubit gates, then the weight of any row of M is ≤ 2. Now, suppose $G \in \mathbb{F}_2^{(n\cdot r)\times n}$ is the generator matrix of a repetition code, on $n' = n \cdot r$ bits and rate $n/n' = \frac{1}{r}$. [20] observe that the new IQP circuit defined by mapping $M \to \tilde{M} = M \cdot G^T$ is robust to input noise, up to (roughly) the random-error-correction capacity of G. Indeed, this follows from the fact that

$$\left\langle G^T x\right| D \left|G^T x\right\rangle = \langle x| \tilde{D} |x\rangle, \forall x \in \{0,1\}^{n'}. \tag{93}$$

Therefore, the output distribution of the new circuit $\tilde{C}$ under input (or output) noise is the same as sampling $y \in \{0,1\}^n$ from C, encoding y into the code $\tilde{y} = Gy \in \{0,1\}^{n'}$, and finally flipping each entry of $\tilde{y}$ independently with probability q. If the repetition code can tolerate random bit-flip errors with rate q, then one can approximately sample from C using noisy samples from $\tilde{C}$.

The caveat in their approach is that the resulting IQP circuits maybe polynomially larger. Indeed, each two qubit gate in the original circuit C, is mapped to a $2 \cdot r$ multi-qubit gate in $\tilde{C}$:

$$e^{i\theta Z_a \otimes Z_b} \to e^{i\theta Z_a^1 \otimes Z_a^2 \cdots Z_a^r \otimes Z_b^1 \cdots Z_b^r} \tag{94}$$

which is complex to implement using only diagonal operations. Instead, we dispense with the requirement that the intermediate gates in the circuit be diagonal (and thus the circuit is not an IQP circuit), however, globally it is equivalent to the same (IQP) unitary operation.

Definition XI.1. *A* k-local Pauli rotation gate *is the k qubit unitary U defined by an angle $\theta \in [0, 2\pi]$ and a k-qubit Pauli P where $U = e^{i\theta P}$.*

Of particular note to us are multi-controlled Z rotations, where $P = Z_1 \otimes Z_2 \cdots Z_k$.

Claim XI.4. *Any k-local Pauli rotation gate can be implemented using an $\leq \log k$ depth circuit on a fully connected architecture of 2-qubit gates.*

For simplicity, we prove the above for multi-qubit Z Paulis, as the general case is analogous.

Proof. Let U be a k-local Z rotation gate, and V be any unitary. Then, the identity $Ve^{i\theta P}V^\dagger = e^{i\theta VPV^\dagger}$ tells us that it suffices to find a depth $d \leq \log k$ Clifford circuit V such that $V(\otimes_i^k Z_i)V^\dagger = Z_1 \otimes \mathbb{I}_{[k]\setminus 1}$. We claim that this can be done recursively, where each layer of V halves the weight of the remaining Z's. Indeed, since $(\mathbb{I} \otimes Z) = \mathsf{CNOT}(Z \otimes Z)\mathsf{CNOT}^\dagger$, layers of CNOT gates on a matching of the remaining Z's will suffice. □

To prove our statement, we instantiate the lemma below with our implementation of multi-controlled Z gates.

Lemma XI.5 ([20]). *Let C be an n qubit IQP circuit of depth d. Then, for every $r \in \mathbb{N}$, there exists a deterministic, $O(n\cdot r)$-time decoding algorithm $\mathcal{A}_r : \{0,1\}^{n\cdot r} \to \{0,1\}^n$, and a quantum circuit C_r on $n_r = n \cdot r$ qubits, comprised only of Hadamard gates and $O(d)$ layers of $\leq 2r$-local Z rotation gates, satisfying*

1) *In the absence of noise, the distribution $\mathcal{A}_r \circ p_{C_r,0,0}$ given by sampling $y \leftarrow p_{\tilde{C}_r,0,0}$ from the output of C_r, and outputting $\mathcal{A}_r(y)$, is the same as sampling from C.*
2) *In the presence of input-level noise with probability q, the statistical distance*

$$\|\mathcal{A}_r \circ p_{\tilde{C}_r,q,0} - p_{C,0,0}\|_1 \leq n \cdot (4 \cdot q \cdot (1-q))^{r/2}. \tag{95}$$

Acknowledgements

We thank Anurag Anshu, Ángela Capel, Lijie Chen, Daniel Stilck França, Sevag Gharibian and Umesh Vazirani for helpful discussions. This work was done in part while the authors were visiting the Simons Institute for the Theory of Computing. T.B. acknowledges support by the National Science Foundation under Grant No. DGE 2146752. Y.L. is supported by DOE Grant No. DE-SC0024124, NSF Grant No. 2311733, and DOE Quantum Systems Accelerator.

References

[1] E. B. Davies, "Markovian master equations," *Communications in Mathematical Physics*, vol. 39, no. 2, pp. 91–110, 6 1974. [Online]. Available: https://doi.org/10.1007/BF01608389

[2] E. Mozgunov and D. Lidar, "Completely positive master equation for arbitrary driving and small level spacing," *Quantum*, vol. 4, p. 227, 2020.

[3] J. Rajakumar and J. D. Watson, "Gibbs sampling gives quantum advantage at constant temperatures with $o(1)$-local hamiltonians," 2024. [Online]. Available: https://arxiv.org/abs/2408.01516

[4] X. Gao, S.-T. Wang, and L. Duan, "Quantum supremacy for simulating a translation-invariant ising spin model." *Physical review letters*, vol. 118 4, p. 040502, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:19828403

[5] J. Bermejo-Vega, D. Hangleiter, M. Schwarz, R. Raussendorf, and J. Eisert, "Architectures for quantum simulation showing a quantum speedup," *arXiv: Quantum Physics*, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:119207848

[6] C. Yin and A. Lucas, "Polynomial-time classical sampling of high-temperature quantum gibbs states," 2023.

[7] A. Bakshi, A. Liu, A. Moitra, and E. Tang, "High-temperature gibbs states are unentangled and efficiently preparable," 2024.

[8] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy, "Proof verification and the hardness of approximation problems," *J. ACM*, vol. 45, no. 3, pp. 501–555, 1998. [Online]. Available: https://doi.org/10.1145/278298.278306

[9] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, "Adiabatic quantum computation is equivalent to standard quantum computation," *SIAM Journal on Computing*, vol. 37, no. 1, pp. 166–194, 2007. [Online]. Available: https://doi.org/10.1137/S0097539705447323

[10] C.-F. Chen, H.-Y. Huang, J. Preskill, and L. Zhou, "Local minima in quantum systems," 2023.

[11] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation*, ser. Graduate studies in mathematics. American Mathematical Society, 2002, vol. 47. [Online]. Available: https://bookstore.ams.org/gsm-47/

[12] P. Rall, C. Wang, and P. Wocjan, "Thermal State Preparation via Rounding Promises," *Quantum*, vol. 7, p. 1132, Oct. 2023. [Online]. Available: https://doi.org/10.22331/q-2023-10-10-1132

[13] C.-F. Chen, M. J. Kastoryano, F. G. S. L. Brandão, and A. Gilyén, "Quantum thermal state preparation," 2023.

[14] C.-F. Chen, M. J. Kastoryano, and A. Gilyén, "An efficient and exact noncommutative quantum gibbs sampler," 2023. [Online]. Available: https://arxiv.org/abs/2311.09207

[15] M. J. Kastoryano and K. Temme, "Quantum logarithmic Sobolev inequalities and rapid mixing," *Journal of Mathematical Physics*, vol. 54, no. 5, p. 052202, 05 2013. [Online]. Available: https://doi.org/10.1063/1.4804995

[16] I. Bardet, A. Capel, L. Gao, A. Lucia, D. Pérez-García, and C. Rouzé, "Rapid thermalization of spin chain commuting hamiltonians," *Phys. Rev. Lett.*, vol. 130, p. 060401, 2 2023. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.130.060401

[17] I. Bardet, Á. Capel, L. Gao, A. Lucia, D. Pérez-García, and C. Rouzé, "Entropy decay for davies semigroups of a one dimensional quantum lattice," *Communications in Mathematical Physics*, vol. 405, no. 2, p. 42, 2 2024. [Online]. Available: https://doi.org/10.1007/s00220-023-04869-5

[18] Ángela Capel, C. Rouzé, and D. S. França, "The modified logarithmic sobolev inequality for quantum spin systems: classical and commuting nearest neighbour interactions," 2021.

[19] D. Hangleiter and J. Eisert, "Computational advantage of quantum random sampling," *Rev. Mod. Phys.*, vol. 95, p. 035001, 7 2023. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.95.035001

[20] M. J. Bremner, A. Montanaro, and D. J. Shepherd, "Achieving quantum supremacy with sparse and noisy commuting quantum computations," *Quantum*, vol. 1, p. 8, 4 2017. [Online]. Available: https://doi.org/10.22331/q-2017-04-25-8

[21] J. Rajakumar, J. D. Watson, and Y.-K. Liu, "Polynomial-time classical simulation of noisy iqp circuits with constant depth," 2024. [Online]. Available: https://arxiv.org/abs/2403.14607

[22] R. Mezher, J. Ghalbouni, J. Dgheim, and D. Markham, "Fault-tolerant quantum speedup from constant depth quantum circuits," *Phys. Rev. Res.*, vol. 2, p. 033444, 9 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevResearch.2.033444

[23] L. Paletta, A. Leverrier, A. Sarlette, M. Mirrahimi, and C. Vuillot, "Robust sparse IQP sampling in constant depth," *Quantum*, vol. 8, p. 1337, May 2024. [Online]. Available: https://doi.org/10.22331/q-2024-05-06-1337

[24] D. Aharonov and M. Ben-Or, "Fault-tolerant quantum computation with constant error rate," *SIAM Journal on Computing*, vol. 38, no. 4, pp. 1207–1282, 2008. [Online]. Available: https://doi.org/10.1137/S0097539799359385

[25] S. Bravyi, D. Gosset, R. König, and M. Tomamichel, "Quantum advantage with noisy shallow circuits," *Nature Physics*, vol. 16, no. 10, pp. 1040–1045, 10 2020. [Online]. Available: https://doi.org/10.1038/s41567-020-0948-z

[26] F. G. S. L. Brandão and M. J. Kastoryano, "Finite correlation length implies efficient preparation of quantum thermal states," *Communications in Mathematical Physics*, vol. 365, no. 1, pp. 1–16, 1 2019. [Online]. Available: https://doi.org/10.1007/s00220-018-3150-8

[27] S. Bravyi and A. Kitaev, "Universal quantum computation with ideal clifford gates and noisy ancillas," *Phys. Rev. A*, vol. 71, p. 022316, 2 2005. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.71.022316

[28] A. Broadbent, J. Fitzsimons, and E. Kashefi, "Universal blind quantum computation," *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pp. 517–526, 2008. [Online]. Available: https://api.semanticscholar.org/CorpusID:650251

[29] P. Wocjan and K. Temme, "Szegedy walk unitaries for quantum maps," *Communications in Mathematical Physics*, vol. 402, pp. 3201–3231, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:235899246

[30] E. Onorati, O. Buerschaper, M. Kliesch, M. Kliesch, W. D. Brown, A. H. Werner, A. H. Werner, and J. Eisert, "Mixing properties of stochastic quantum hamiltonians," *Communications in Mathematical Physics*, vol. 355, pp. 905 – 947, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:30986363

[31] M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano, and J. Eisert, "Dissipative quantum church-turing theorem," *Physical Review Letters*, vol. 107, no. 12, Sep. 2011. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.107.120501

[32] M. J. Bremner, R. Jozsa, and D. J. Shepherd, "Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 467, pp. 459 – 472, 2010. [Online]. Available: https://api.semanticscholar.org/CorpusID:12301677

[33] D. Hangleiter, J. Bermejo-Vega, M. Schwarz, and J. Eisert, "Anticoncentration theorems for schemes showing a quantum speedup," *arXiv: Quantum Physics*, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:51689113

[34] L. Novo, J. Bermejo-Vega, and R. Garc'ia-Patr'on, "Quantum advantage from energy measurements of many-body quantum systems," *Quantum*, vol. 5, p. 465, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:209370723

[35] S. Aaronson and A. Y. Arkhipov, "The computational complexity of linear optics," *ArXiv*, vol. abs/1011.3245, 2010. [Online]. Available: https://api.semanticscholar.org/CorpusID:681637

[36] M. J. Bremner, A. Montanaro, and D. J. Shepherd, "Average-case complexity versus approximate simulation of commuting quantum computations," *Physical review letters*, vol. 117 8, p. 080501, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:8590553

[37] L. Gao and C. Rouzé, "Complete entropic inequalities for quantum markov chains," *Archive for Rational Mechanics and Analysis*, vol. 245, no. 1, pp. 183–238, 2022.

[38] A. Muller-Hermes, D. S. França, and M. M. Wolf, "Entropy production of doubly stochastic quantum channels," *arXiv: Quantum Physics*,

2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:119321308

[39] ——, "Relative entropy convergence for depolarizing channels," *arXiv: Quantum Physics*, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:119314729

[40] Á. Capel, A. Lucia, and D. Pérez-García, "Quantum conditional relative entropy and quasi-factorization of the relative entropy," *Journal of Physics A: Mathematical and Theoretical*, vol. 51, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:13749880

[41] S. Beigi, N. Datta, and C. Rouzé, "Quantum reverse hypercontractivity: Its tensorization and application to strong converses," *Communications in Mathematical Physics*, vol. 376, pp. 753–794, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:13741288

[42] D. Petz, "On certain properties of the relative entropy of states of operator algebras," *Mathematische Zeitschrift*, vol. 206, pp. 351–361, 1991. [Online]. Available: https://api.semanticscholar.org/CorpusID:121075061

[43] O. Fawzi and R. Renner, "Quantum conditional mutual information and approximate markov chains," *Communications in Mathematical Physics*, vol. 340, pp. 575 – 611, 2014. [Online]. Available: https://api.semanticscholar.org/CorpusID:118848654

[44] M. Junge, R. Renner, D. Sutter, M. M. Wilde, and A. J. Winter, "Universal recovery from a decrease of quantum relative entropy," *ArXiv*, vol. abs/1509.07127, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:1955738

[45] W. Brown and D. Poulin, "Quantum markov networks and commuting hamiltonians," *arXiv: Quantum Physics*, 2012. [Online]. Available: https://api.semanticscholar.org/CorpusID:55999418

[46] S. Michalakis and J. P. Zwolak, "Stability of frustration-free hamiltonians," *Communications in Mathematical Physics*, vol. 322, pp. 277–302, 2011. [Online]. Available: https://api.semanticscholar.org/CorpusID:39371867

[47] M. B. Hastings, "Quantum belief propagation: An algorithm for thermal quantum systems," *Physical Review B*, vol. 76, p. 201102, 2007. [Online]. Available: https://api.semanticscholar.org/CorpusID:117701192

[48] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 2018.

[49] A. Gilyén, S. Lloyd, I. Marvian, Y. Quek, and M. M. Wilde, "Quantum algorithm for petz recovery channels and pretty good measurements," *Physical Review Letters*, vol. 128, no. 22, p. 220502, 2022.

[50] S. Aaronson, "Quantum computing, postselection, and probabilistic polynomial-time," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 461, pp. 3473 – 3482, 2004. [Online]. Available: https://api.semanticscholar.org/CorpusID:1770389