Gibbs Sampling gives Quantum Advantage at Constant Temperatures with $O(1)$-Local Hamiltonians

Joel Rajakumar[1,2] and James D. Watson[1,2]

[1] Joint Center for Quantum Information & Computer Science, National Institute of Standards & Technology and University of Maryland, College Park

[2]Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park

Sampling from Gibbs states — states corresponding to system in thermal equilibrium — has recently been shown to be a task for which quantum computers are expected to achieve super-polynomial speed-up compared to classical computers, provided the locality of the Hamiltonian increases with the system size [1]. We extend these results to show that this quantum advantage still occurs for Gibbs states of Hamiltonians with $O(1)$-local interactions at constant temperature by showing classical hardness-of-sampling and demonstrating such Gibbs states can be prepared efficiently using a quantum computer. In particular, we show hardness-of-sampling is maintained even for 5-local Hamiltonians on a 3D lattice. We additionally show that the hardness-of-sampling is robust when we are only able to make imperfect measurements.

1 Introduction

Gibbs states are fundamental objects of interest in many-body physics and chemistry, where they correspond to the state a system equilibrates to at a fixed temperature, and play an important role in semidefinite program solving and machine learning. One of the key uses cases for quantum computers has been to simulate many-body quantum systems, and with this in mind, a variety of quantum algorithms for Gibbs state preparation and sampling have been proposed[1], including a recent promising quantum generalisation of Metropolis-Hastings algorithm [17, 18].

At a given inverse temperature β and Hamiltonian H, we define the Gibbs state as:

$$\rho(H,\beta) = \frac{e^{-\beta H}}{\mathrm{tr}\left[e^{-\beta H}\right]}.$$

Despite their importance, the computational complexity of calculating quantum Gibbs state properties, and whether there is a quantum advantage, has remained poorly understood — particularly at temperature independent of system size $\beta = \Theta(1)$. For Gibbs states of both classical and quantum Hamiltonians, efficient algorithms are known to exist above certain critical temperatures [19, 20, 21, 22, 23, 24], and at sufficiently low $O(1)$ temperatures Gibbs state properties of classical Hamiltonians are known to be NP-hard and even MA-complete to compute [25, 26, 27]. At even lower temperatures, $\beta = \Omega(\mathrm{poly}(n))$, the Gibbs state has high overlap with the system's ground state, and so for quantum Hamiltonians we can argue that cooling to these temperatures must be at least QMA-hard. Indeed, for quantum Hamiltonians

[1][2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. See table 1 of [11] for a discussion of Gibbs state algorithms and their limitations. Beyond this, many algorithms for computing thermal expectation values by relation to the microcanonical ensemble [14, 15, 16]

at $\beta = \Omega(\text{poly}(n))$, computing expectation values of local observables measured on Gibbs states has been shown to be hard for a class QXC, "Quantum Approximate Counting" [28, 29][2]. But exactly how this class relates to other classical and quantum complexity classes, and whether we expect similar hardness results for $\beta = \Theta(1)$, remains poorly understood. The question of the complexity of Gibbs state properties is intimately related to the question of mixing times for preparing Gibbs states — a rapid mixing time for an algorithm preparing a particular Gibbs state implies that preparing the Gibbs state in question is computationally tractable — and many results have been proven for quantum a variety of Hamiltonians and algorithms [31, 32, 33, 34, 35]. Gibbs sampling is also related to other important problems which have shown quantum speed-up including sampling log-concave distributions and volume estimation [36, 37].

Recent work by Bergamaschi et al. consider the task of sampling bitstrings from quantum Gibbs states [1]. That is, sampling from a distribution close in total variation distance to

$$P(x) = \frac{\langle x| e^{-\beta H} |x\rangle}{\text{tr}\left[e^{-\beta H}\right]}$$

for a bitstring $x \in \{0,1\}^n$. Bergamaschi et al. demonstrate that sampling bitstrings from families of Gibbs states of quantum Hamiltonians with $O(\log\log(n))$-local terms at temperature $\beta = \Theta(1)$ is classically intractable under complexity-theoretic conjectures (e.g. unless the polynomial hierarchy collapses to the third level) [1]. They further demonstrate that such Gibbs states can be prepared in polynomial time using a quantum computer, and hence the task of Gibbs sampling can be done efficiently with the aid of a quantum computer. Not only does this demonstrate a quantum advantage in Gibbs state preparation over classical computers — and thus may be a good test of so-called quantum supremacy — but it is arguably the first convincing demonstration that Gibbs states retain non-trivial quantum computational properties at $\beta = \Theta(1)$ temperatures.

In this work we build upon the work in Bergamaschi et al. and demonstrate there exist families of 5-local and 6-local Hamiltonians such that sampling from Gibbs states at $\beta = \Theta(1)$ temperatures remains classically intractable under complexity-theoretic conjectures.

Theorem 1 ((Informal) Classically Intractable Gibbs Sampling)**.** *There exist two families $\mathcal{F}_1, \mathcal{F}_2$ of efficiently constructable Hamiltonians such that sampling from a probability distribution $Q(x)$ satisfying:*

$$\left\| Q(x) - \frac{\langle x| e^{-\beta H} |x\rangle}{\text{tr}\left[e^{-\beta H}\right]} \right\|_1 \leq \epsilon,$$

is not possible for randomised classical algorithms under complexity-theoretic conjectures, for the following parameter regimes:

- *$\mathcal{F}_1$: the Hamiltonians are 5-local, nearest-neighbour Hamiltonians on a 3D cubic lattice, where each qubit is involved in at most 4 Hamiltonian terms, and $\epsilon = O(\exp(-n))$. Furthermore, we allow each single-qubit measurement outcome to be incorrect with an $O(1)$ probability.*
- *$\mathcal{F}_2$: the Hamiltonians are 6-local, and $\epsilon = O(1)$.*

Sampling from a $Q(x)$ close to either of these distributions can be done efficiently using a quantum computer.

We prove our results by combining results from the hardness of sampling IQP circuits with results from measurement-based quantum computation and error-detection protocols. In particular, we utilise work

[2]For $\beta = \Omega(\text{poly}(n))$ it is also true that estimating the *normalised* partition function is DQC1-hard [30]

on the hardness of sampling from IQP circuits developed by Fujii and Tamate [38] and Bremner et al. [39] to construct families of Hamiltonians whose corresponding Gibbs states inherent this hardness of sampling. We anticipate that the families of Hamiltonians in $\mathcal{F}_1, \mathcal{F}_2$ are easier to realise experimentally than those presented in previous works.

Paper Outline In Section 2 we give the preliminaries and notation. Section 3 contains an overview of the proof techniques, as well as our main results: Section 3.2 proves hardness of sampling Gibbs states of 5-local Hamiltonians on a 3D lattice and Section 3.3 proves hardness for 6-local Hamiltonians. In Section 4, we show how these hardness results can be made robust to errors in measurement. In Section 5 we suggest a heuristic method of demonstrating that we have prepared the desired Gibbs state. Finally, we discuss this work and open questions in Section 6.

2 Preliminaries

2.1 Notation

Consider a set of n particles, each with local Hilbert space $\mathbb{C}^d$, then the full Hilbert space is $(\mathbb{C}^d)^{\otimes n}$. For a Hilbert space $\mathcal{H}$ we denote the set of bounded linear operators as $\mathcal{B}(\mathcal{H})$. A Hamiltonian acting on n qudits is a Hermitian operator $H \in \mathcal{B}((\mathbb{C}^d)^{\otimes n})$. For the rest of this work will we consider the case $d = 2$, i.e. the Hamiltonian acts on qubits. $\|\cdot\|$ will denote the operator norm, $\|\cdot\|_1$ will denote the Schatten 1-norm when acting on operators, and if acting on probability distributions is the standard $L1$ distance between probability distributions.

Hamiltonian Parameters. The Hamiltonian is called a k-local Hamiltonian if we can write it as:

$$H = \sum_i h_i$$

if each h_i acts on at most k many qubits. Given a local Hamiltonian, we can define a interaction (hyper)graph, where the vertices are given by the qubits and the (hyper)edges are given by the qubits h_i acts non-trivially on. We denote the maximum degree of the interaction graph as Δ. We assume that $\|h_i\| = O(1)$ for all local terms in the Hamiltonian.

Thermal Physics. Given a temperature $T \in \mathbb{R}^+$, we will work with the inverse temperature $\beta = 1/T$. For a given Hamiltonian H with eigenvalues $\{\lambda_i\}_i$, and a given inverse temperature β, the partition function is define as:

$$Z = \text{tr}\left[e^{-\beta H}\right] = \sum_i e^{-\beta \lambda_i},$$

and the associated Gibbs state is:

$$\rho(H, \beta) = \frac{1}{Z} e^{-\beta H}.$$

Noisy Circuits. We will often discuss circuits by directly referring to their unitary matrix C, with the convention that $C|0\rangle^{\otimes n}$ is the output state after applying C on $|0\rangle^{\otimes n}$. We will be interested in bit-flip noise in our circuit, and define the single-qubit channel:

$$\mathcal{D}_p(\rho) = (1-p)\rho + pX\rho X$$

We use the following notation to refer to the output distribution of a circuit with noise on the input state:

Definition 2 (Circuit Sampling Distributions)**.** *For any circuit C we will use P_C to denote the output distribution of sampled bitstrings from C,*

$$P_C(x) = |\langle x| C |0^n\rangle|^2,$$

and $P_{C,q}$ to denote the output distribution of sampled bitstrings from C with independent single-qubit bit-flips applied on every input qubit with probability q:

$$P_{C,q}(x) = \mathrm{tr}\left[|x\rangle\langle x| \cdot C(\mathcal{D}_q^{\otimes n}(|0\rangle\langle 0|^{\otimes n}))C^\dagger\right].$$

2.2 Parent Hamiltonian Construction for Quantum Circuits

Initially consider a non-interacting Hamiltonian on n qubits:

$$H_{\mathrm{NI}} = \sum_{i=1}^{n} \frac{1}{2}(\mathbb{1} - Z_i).$$

It can be seen that the eigenstates of this Hamiltonian correspond to bitstrings $|x\rangle$ for $x \in \{0,1\}^n$, where the energy of eigenstate $|x\rangle$ is $\lambda_x = \mathrm{HW}(x)$, where HW denotes the Hamming weight of a string x. The zero-energy ground state is $|0\rangle^{\otimes n}$. The corresponding Gibbs state of the Hamiltonian is:

$$\frac{1}{Z}e^{-\beta H} = \frac{1}{Z}e^{-\beta\sum_{x\in\{0,1\}^n}\lambda_x|x\rangle\langle x|}$$

and $Z = \sum_{x\in\{0,1\}^n} e^{-\beta\lambda_x}$. We see that Z is the partition function for n non-interacting spins, and hence $Z = (1+e^{-\beta})^n$. Now, given a circuit C on n qubits, we can define a *parent Hamiltonian:*

$$H_C = CH_{\mathrm{NI}}C^\dagger.$$

2.3 Gibbs States of Parent Hamiltonians

We use the following two lemmas in our work. The first (implicitly used in [1]) shows that for the set of parent Hamiltonians formed using quantum circuits, their Gibbs states are equivalent to the circuit with bit-flip noise acting on the input. Here the input noise strength is related to the temperature of the Gibbs state. We provide an explicit proof in Appendix B for completeness. Here the temperature of the Gibbs state corresponds to the strength of the bit-flip noise.

Lemma 3. *For any circuit C constructed from k-local gates of depth d, there exists a $O(kd)$-local parent Hamiltonian H_C, such that:*

$$\frac{1}{Z}e^{-\beta H_C} = C(\mathcal{D}_q^{\otimes n}(|0\rangle\langle 0|))C^\dagger$$

for $q = \frac{e^{-\beta}}{1+e^{-\beta}}$.

The second lemma, which is a large portion of the technical work in [1], shows that a quantum computer can efficiently prepare the Gibbs States of these parent Hamiltonians.

Lemma 4 (Efficient Gibbs State Preparation for Parent Hamiltonians, Lemma 1.2 of [1])**.** *Fix $\beta > 0$, and let H_C be the parent Hamiltonian of a quantum circuit C on n qubits, of depth d and locality 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 $O(2^{4\ell}2^d e^\beta n\,\mathrm{poly}(\log\frac{n}{\epsilon}, \ell, \beta))$.*

At a high level, this result proves a rapid mixing property of a particular Lindbladian evolution that converges to the Gibbs state. The analysis takes advantage of the fact that the Hamiltonian is the parent Hamiltonian of a $O(\log(n))$-depth circuit, which means that the 'lightcones' of any qubit can be bounded.

3 Hardness of Sampling Gibbs States with *O(1)*-Local Hamiltonians

3.1 Overview of Techniques

We wish to prove hardness of classically sampling the from Gibbs states of a Hamiltonian. To do so, we turn to Lemma 3: our goal is to find a family of circuits C with corresponding parent Hamiltonians whose Gibbs states are hard to sample from with a classical computer but efficiently sampleable with a quantum computer. C must satisfy the following requirements,

1. For some noise strength $q = O(1)$, $P_{C,q}$ must be hard to sample from with either inverse exponential or additive error.
2. H_C must be $O(1)$-local.
3. C must have $O(\log n)$ depth.

In this work, we will choose C to be from the family of IQP circuits with the aim of maintaining classical hardness-of-sampling while ensuring the associated parent Hamiltonian remains $O(1)$-local. The main reason for choosing these circuits is that there are known constant depth IQP circuits which are hard to sample from Refs. [38, 40], and the shallowness of the circuit is useful to obtain $O(1)$-locality of the parent Hamiltonian. In particular, we need to choose a family of IQP circuits whose hardness of sampling is robust to bit-flip noise on the input. This is non-trivial since in many cases the presence of noise may make the IQP circuit classically easy to sample [41, 42].

To prove our results, we use two separate error mitigation techniques. First, we use a result from Fujii and Tamate [38]. Fujii and Tamate make use of topologically protected circuits to ensure that, provided the circuit noise is below a certain threshold, exact sampling from the circuit remains difficult. Modifying this proof allows us to obtain hardness of sampling with inverse exponential error, and this allows us to arrive at $\mathcal{F}_1$ of Theorem 1, a family of 5-local Hamiltonians with degree 5, which are hard to sample from with inverse exponential error.

Second, we implement an error-detection strategy inspired by the techniques of Refs. [39, 1], where parts of the circuit are repeated and we simultaneously introduce flag qubits which flip to show where an error may have occurred. This allows us to arrive at hardness of sampling from with additive error for $\mathcal{F}_2$, a family of Hamiltonians with $O(1)$ locality. Our hardness-of-sampling proof follows the similar outline as Ref. [1], but here we prove hardness for different sets of circuits which allows us to obtain hardness for families of Hamiltonians with $O(1)$-locality rather than $O(\log\log(n))$-locality. We also note similar work on thermal states of measurement-based thermal states in [43].

Finally, by combining these two techniques, we obtain a family of 6-local Hamiltonians with max degree Δ, such that their Gibbs states are hard to sample from with inverse exponential error when $\beta = O(1/\sqrt{\Delta})$, which complements easiness results from [24].

3.2 Hardness of Sampling Gibbs States on 3D Lattices

Fujii and Tamate demonstrate a family of constant-depth, geometrically local IQP circuits which are hard to *exactly* sample from in the presence of noise of constant strength [38]. At a high level, Fujii and Tamate give a protocol in which a cluster state is constructed using a depth 4 IQP circuit. If the noise level is sufficiently smaller than the threshold value for topologically protected MBQC, then the output distribution of the circuit cannot be exactly sampled from. Provided the noise is restricted to bit-flip noise and kept below a threshold strength of ~ 0.134, then T-gate synthesis is still possible in the circuit, and postselecting on error-free outcomes allows one to decide PostBQP-complete problems. Since PostBQP $=$ PP and it is believed that PostBQP $\not\subseteq$ PostBPP unless the polynomial hierarchy collapses to the third level, then sampling from the noisy circuit is still hard [44].

By taking this set of topologically protect circuits from Fujii and Tamate we prove the following lemma in Appendix A.

Lemma 5. *There exists a family of IQP circuits $\mathcal{C}$, constructed on a 3D cubic lattice, consisting of a single layer of $e^{iZ\pi/8}$ gates and 4 layers of nearest-neighbour $e^{iZZ\pi/4}$ gates, such that sampling from any probability distribution Q over bitstrings which satisfies $\|P_{C,q} - Q\|_1 < (1-q)^n 2^{-6n-4}/5$ is not possible with a classical polynomial algorithm, unless the polynomial hierarchy collapses to the third level, when $q < 0.134$.*

We now want to use this family of classically hard-to-sample circuits to generate hard-to-sample Gibbs states. In particular, we combine this hard-to-sample from family of IQP circuits with the fact that the Gibbs states of IQP parent Hamiltonians look like the output of noisy IQP circuits, as per Lemma 3, and we get the following.

Theorem 6 (Classical Hardness of Gibbs Sampling)**.** *One can efficiently construct a family of 5-local Hamiltonians on a 3D cubic lattice of qubits, $\{H_C\}_C$, whose Gibbs states at $\beta = O(1)$ are hard to classically sample from with 1-error $(1-q)^n 2^{-6n-4}/5$, unless the polynomial hierarchy collapses to the 3rd level.*

Proof. We choose a hard-to-sample IQP circuit C from Lemma 5 and specify the new Hamiltonian as $H_C = C H_{\mathrm{NI}} C^\dagger$. By Lemma 3, the output distribution of the Gibbs states of H_C is exactly $P_{C,q}$ Choosing:

$$q = \frac{e^{-\beta}}{1+e^{-\beta}} \leq 0.134 \tag{1}$$

and we see that the output distribution must be hard to sample from, as per Lemma 5. Thus we see that for $\beta \gtrsim 1.87 = O(1)$, it is hard to sample this distribution.

To see that the locality of H_C is 5, we note that for each qubit i, there are at most 4 gates acting on qubit i in C, each of which introduces interaction with at most 4 other qubits. Thus, when conjugating the Z_i term of H_{NI} with the gates of C, it becomes an at most 5-local term. □

Notably, in order to prove hardness of sampling with inverse exponential error from a noisy IQP circuit, a postselection argument suffices, rather than a full fault-tolerant encoding. Thus, there is no added overhead needed to give hardness of sampling, which ensures that the locality of the parent Hamiltonian is purely determined by the logical IQP circuit structure, which has low depth and locality. This is not true for the circuits and parent Hamiltonians in the next section, or in Ref. [1] where additional elements need to be added to the circuit to allow for error detection.

3.3 Hardness of Gibbs Sampling to $O(1/\mathrm{poly}(N))$ Error

While the hardness of sampling from noiseless IQP circuits with inverse exponential error can be obtained through a postselection argument, proving hardness of sampling with additive error typically requires a few more ingredients (e.g. anticoncentration, average-to-worst case reduction). Here, we use existing results from Hangleiter et al. [40] which construct a family of constant-depth IQP circuits that exhibit such hardness (up to some complexity-theoretic conjectures).

Lemma 7. *(Corollary 12 of [40]) There exists a family of constant depth IQP circuits $\{C_n\}_n \geq 1$ on a 2D square lattice of n qubits, such that no randomised classical polynomial-time algorithm can sample from the output distribution of C_n up to additive 1 error of $\delta = 1/192$, assuming the average-case hardness of computing a fixed family of partition functions, and the non-collapse of the polynomial hierarchy.*

Now we would like to show that the hardness of sampling from this circuit is robust to input noise, and without postselection (as this would make it more difficult to show hardness of additive error sampling). This task has been explored in Refs. [39, 1] for IQP circuits, and they each provide a method of encoding an arbitrary IQP circuit into a larger, noise-robust circuit. Here, we describe an encoding method that is heavily inspired by these techniques.

Our construction is as follows. Suppose the initial circuit acts on n qubits. Fix some positive integer r. For each 'logical' qubit i in the circuit of Lemma 7 (initialised in the $|0\rangle$ state), initialise a block of r 'physical' qubits in the $|0\rangle$ state, and label them $i_1, \ldots, i_r$. For each block i, apply $r-1$ CNOT gates which are all controlled on i_1, and targeting a qubit in $i_2, \ldots, i_r$. Let the unitary corresponding to this CNOT network be B. Then, apply the logical circuit (from Lemma 7) amongst the first qubits of each block (qubits labelled $1_1, 2_1, \ldots, n_1$). Finally, measure all qubits in the computational basis. See Fig. 1 for an example of such a circuit with $r = 2$.

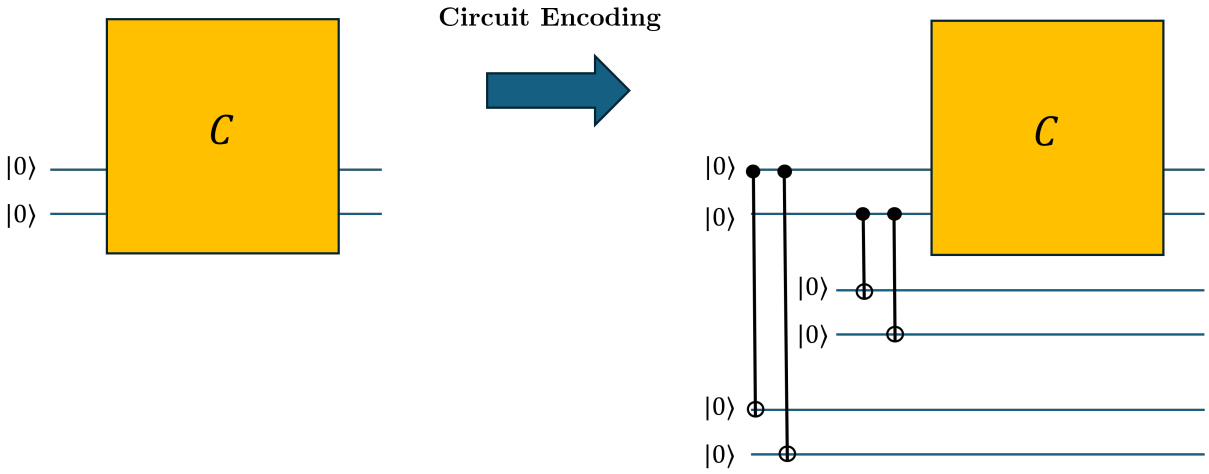

Figure 1: The encoding of the IQP circuit C with $r = 3$. For clarity, only two input qubits are shown explicitly, but the procedure applies to all qubits present.

We see that if the initial logical qubit is flipped by noise at the start of the circuit, the CNOT network associated with that qubit should flip all the associated ancillary qubits to the $|1\rangle$ state, thus flagging the error which can be corrected in postprocessing. We require r auxiliary qubits to flag the error as the auxiliary qubits themselves may have noise incident on them, and so the multiple copies acts a repetition code to suppress this case.

We show in Appendix C that with classical post-processing, one can recover the logical output distribution with an error rate that is exponentially suppressed in r. Crucially, we only require $r = O(1)$ which means the parent Hamiltonian remains $O(1)$-local. Our proof works by examining the propagation of Pauli's through the CNOT network.

Lemma 8. *Let C be an arbitrary IQP circuit constructed with 2-qubit gates of depth d on n qubits. Then, for any integer parameter $r \geq 1$, there is an encoded IQP circuit C^* constructed with 2-qubit gates of depth $d + r$ on nr qubits, and a decoding algorithm A^* such that,*

$$A^*(P_{C^*,q}) = P_{C,p_{fail}} \tag{2}$$

where $p_{fail} \leq (4q(1-q))^{r/2}$. Furthermore, the parent Hamiltonian of C^* has locality $k \leq d+2$ and degree $\Delta \leq r(d+1)$

Thus, given an initial circuit C, obtain a new circuit composed of an initial set of CNOT layers, followed by an IQP circuit. The output of the error-free circuit can be extracted from classical post-processing. We can put this all together to obtain the following theorem.

Theorem 9. *One can efficiently construct a family of* 6*-local Hamiltonians,* $\{H_C\}_C$*, whose Gibbs states at* $\beta = O(1)$ *are hard to classically sample from with constant* 1*-error up to* $1/192$*, under the assumptions of Lemma 7.*

Proof. We choose a hard-to-sample IQP circuit C from Lemma 7. Then, we encode C into C^* as in Lemma 8 (with parameter r to be set later), and specify the new Hamiltonian as $H_{C^*} = C^* H_{\mathrm{NI}} C^{*\dagger}$. By Lemma 3, the output distribution of the Gibbs states of H_{C^*} at constant temperature is exactly $P_{C^*,q}$. Using A^*, we can postprocess this to obtain $P_{C,p_{\mathrm{fail}}}$ where $p_{\mathrm{fail}} \leq (4q(1-q))^{r/2}$. Therefore, it suffices to choose $r = O(\log n)$ in order to set the probability of an error happening on any qubit to be $O(1/n)$. By increasing r, we can set the approximation error arbitrarily lower than $1/192$. Thus, it is hard to sample from the Gibbs states of H_{C^*} with constant 1 error up to $\delta = 1/192$. □

Additionally, in Appendix C we show that the output distributions of this work and previous noise-robust IQP encodings (specifically those presented in here and in Refs. [39, 1]) are equivalent, up to reversible classical postprocessing. This highlights the fact that improvements in Hamiltonian locality have not been obtained by considering very different output distributions, but rather, different Hamiltonian constructions.

3.4 Quantum Advantage from Gibbs Sampling of $O(1)$-Local Hamiltonians

So far we have proven hardness of sampling from Gibbs states for the families of parent Hamiltonians in Section 3.2 and Section 3.3. Here we employ the Gibbs state preparation algorithm of Ref. [11] in conjunction with Lemma 4.

Theorem 10 (Quantum Advantage with $O(1)$-Local Hamiltonians)**.** *The Gibbs states associated with the families of parent Hamiltonians in Theorem 6 and Theorem 9 can be sampled from efficiently using a quantum computer in the same parameter regimes for which they are classically hard.*

Proof. For both of the families of Hamiltonians in Theorem 6 and Theorem 9 the Hamiltonians are $O(1)$-local, hence the runtime of the Gibbs state preparation algorithm in Lemma 4 with $\beta = \Theta(1)$, $\ell = O(1)$ and $\epsilon = \frac{1}{9}(1+e^{-\beta})^{-n}2^{-6n-4}$ gives a total runtime of $O(2^{4\ell}2^d e^{\beta} n\,\mathrm{poly}(\log \frac{n}{\epsilon}, \ell, \beta)) = O(\mathrm{poly}(n))$. Thus sampling from these distributions can be done in polynomial time for a quantum computer.

On the other hand, as shown in Theorem 6 and Theorem 9, sampling from these Gibbs states in this parameter regime is classically intractable under complexity-theoretic conjectures. □

4 Measurement Errors in Gibbs State Sampling

When trying to show a quantum advantage in Gibbs state sampling, there are two sources of error: *(a)* we can only approximately prepare the true Gibbs state, and *(b)* when sampling, we expect our measurements to have an error associated with them. Here we show that, even with these two sources of error, we are sampling from a probability distribution that we do not expect to be able to sample from classically.

Theorem 11. *Consider a family of IQP circuits* $\mathcal{C}$*, the associated parent Hamiltonians* $\{H_C\}_{C\in\mathcal{C}}$ *and an inverse temperature* $\beta = \Theta(1)$*. There are a family of efficiently preparable states* $\{\tilde{\rho}_C\}_{C\in\mathcal{C}}$ *in time* $\mathrm{poly}(n)$ *on a quantum computer, such that for* $\epsilon = \frac{1}{9}(1+e^{-\beta})^{-n}2^{-6n-4}$*,*

$$\|\tilde{\rho}_C - \rho(H_C, \beta)\|_1 \leq \epsilon$$

and sampling from $\mathcal{D}_q^{\otimes n}(\tilde{\rho}_C)$ (i.e. with imperfect measurements with single qubit measurement error $q = \Theta(1)$) is not possible using a randomised classical algorithm unless the polynomial hierarchy collapses to the 3rd level.

Proof. From Lemma 4, we see that preparing $\tilde{\rho}_C$ to this precision is possible in $O(\text{poly}(n))$ time using a quantum computer. The distribution we are sampling corresponding to faulty measurements with probability q on the prepared state $\tilde{\rho}$ approximating $\rho(H_C, \beta)$ is:

$$\tilde{P}(x) = \text{tr}\left[|x\rangle\langle x| \cdot \mathcal{D}_q^{\otimes n}(\tilde{\rho}_C)\right].$$

We using Hölder's inequality:

$$\left\|\tilde{P}(x) - \text{tr}\left[|x\rangle\langle x| \mathcal{D}_q^{\otimes n}(\rho(H_C, \beta))\right]\right\|_1 \leq \left\|\mathcal{D}_q^{\otimes n}(\tilde{\rho}_C) - \mathcal{D}_q^{\otimes n}(\rho(H_C, \beta))\right\|_1.$$

Since $\|\tilde{\rho}_C - \rho(H_C, \beta)\|_1 \leq \epsilon$, then we see:

$$\begin{aligned}\left\|\mathcal{D}_q^{\otimes n}(\tilde{\rho}_C - \rho(H_C, \beta))\right\|_1 &\leq \|\tilde{\rho}_C - \rho(H_C, \beta)\|_1 \\ &\leq \epsilon,\end{aligned}$$

where we have used that $\mathcal{D}^{\otimes n}(\cdot)$ is a CPTP map. Thus we see that sampling from $\tilde{P}(x)$ approximates sampling from the Gibbs state $\mathcal{D}^{\otimes n}(\rho(H_C, \beta))$ with imperfect measurements.

Now, we show that $\mathcal{D}^{\otimes n}(\rho(H_C, \beta))$ is also hard to classically sample from. We know from Lemma 3 that we can write Gibbs states of parent Hamiltonians as equivalent to the output of a noisy IQP circuit:

$$\begin{aligned}\mathcal{D}_q^{\otimes n}\left(\frac{e^{-\beta H_C}}{Z}\right) &= \mathcal{D}_q^{\otimes n}\mathcal{D}^{\otimes n}_{\frac{e^{-\beta}}{1+e^{-\beta}}}(C|0\rangle\langle 0|^{\otimes n} C^\dagger) \\ &= \mathcal{D}_{q'}^{\otimes n}(C|0\rangle\langle 0|^{\otimes n} C^\dagger) \\ &= \frac{e^{-\beta' H_C}}{Z}\end{aligned}$$

where we have defined $q' = (\frac{e^{-\beta}}{1+e^{-\beta}})(1-q) + q(1 - \frac{e^{-\beta}}{1+e^{-\beta}})$ and $q' = \frac{e^{-\beta'}}{1+e^{-\beta'}}$.

Provided we keep β' sufficiently small (but still $\Theta(1)$), we see that sampling from $\rho(H_C, \beta')$ is classically intractable from Theorem 6, hence $\mathcal{D}_q^{\otimes n}(\rho(H_C, \beta))$ is classically hard to sample too. In particular, we can choose values of $\beta, q = \Theta(1)$ simultaneously. □

5 Heuristic Verification Procedure

One the of the major limitations with many quantum supremacy experiments is that verifying that the desired experiment has actually been implemented — and not ruined by noise or an adversary — is not trivial [45]. For example, in the case of Random Circuit Sampling and IQP sampling, cross-entropy benchmarking is often used as a proxy for measuring fidelity between the actual state and the ideal state [46, 47]. However, although cross-entropy benchmarking only requires a polynomial number of samples, it requires computing samples of the ideal distribution which is exponentially expensive, and it appears to be spoofable for many classes of circuits. Thus verifying that one is sampling from a distribution close one taken from random circuits is highly non-trivial.

A Heuristic Verification Protocol Suppose we wish to implement the Gibbs sampling procedure using the Hamiltonians in this work as a quantum supremacy test, then we have a similar problem — how do we verify that we are sampling from the correct distribution? Here we suggest a heuristic test for correctness — we ask to verify that the Gibbs state we are sampling from has the correct Hamiltonian. That is, suppose for part of our sampling routine, we wish to sample from the state $\rho(H_P, \beta)$, where H_P is a Hamiltonian formed from a sum of Pauli operators and can be written as:

$$H_P = \sum_{i=1}^{m} \mu_i P_i$$

where each P_i is a k-local Pauli string. To verify that we are correctly preparing Gibbs states, we take N copies of the Gibbs state $\rho(H_P, \beta)$. We then use the Hamiltonian learning algorithm proposed in [48, Theorem 6.1] to learn the coefficients of the Hamiltonian. In particular, if we want to learn the parameters to precision $(\tilde{\mu}_i - \mu_i)^2 \leq \epsilon$ with probability $> 1 - \delta$, we need a number of sample N:

$$N = O\left(\frac{m^6}{\epsilon^c} + \frac{c}{\beta^2\epsilon^2}\log\left(\frac{m}{\delta}\right)\right),$$

for a constant c which depends on the locality and maximum degree of the interaction graph. Then we check closeness for all parameters μ_i which appear in our Hamiltonians. We can then use the following lemma to argue that if the measured Hamiltonians are close, then the sampled states are close:

Lemma 12 (Lemma 16 of [49]). *Let H, H' be Hermitian matrices. Then:*

$$\left\| \frac{e^{-H_1}}{\operatorname{tr}\left[e^{-H_1}\right]} - \frac{e^{-H_2}}{\operatorname{tr}\left[e^{-H_2}\right]} \right\|_1 \leq 2(e^{\|H_1 - H_2\|_\infty} - 1). \tag{3}$$

In particular, if $\beta\|H_1 - H_2\| \leq \epsilon \ll 1$, then $\|\rho(H_1, \beta) - \rho(H_2, \beta)\|_1 \leq O(\epsilon)$.

In particular, if we wish to verify that the actual observed Hamiltonian is close to the ideal Hamiltonian on all k-local terms, then this reduces to learning a Hamiltonian with $4^k\binom{n}{k} = O(\text{poly}(n))$-many terms. This can be done efficiently using the algorithms of [48].

Limitations of the Protocol However, our measurement approach is limited in the sense that the state prepared may not have been the Gibbs state of the desired form, but some other state ρ. We can always write $\rho = \frac{1}{Z}\sum_i p_i |\psi_i\rangle\langle\psi_i| = \frac{1}{Z}\sum_i e^{-\beta\lambda_i}|\psi_i\rangle\langle\psi_i|$ where we have written $p_i = \frac{e^{-\beta\lambda_i}}{Z}$ for some "false" energies λ_i. This gives and effective Hamiltonian $H_{\text{eff}} = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i| = \sum_j \mu'_j P_j$. Thus if the state is incorrect, and not close to the Gibbs state, then when we apply our learning algorithm we will learn the parameters of H_{eff}. Although the μ'_i associated with H_{eff} may be close to H_P on all the parameters μ_i, there may be some μ'_i which are non-zero (and large) when μ_i is zero. Thus H_{eff} will look close to H_P for the Pauli coefficients that we have measured. In general, to distinguish H_{eff} and H_P, we may have to make 4^n many measurements.

Despite this, it remains an open question whether Hamiltonian structure learning can be done efficiently for quantum Gibbs states. That is, can we not only estimate the values of the parameters μ_i, but determine the entire Hamiltonian structure? Classically, this can be done efficiently [50, 51].

Finally, we note that although the above Hamiltonian learning procedure may be spoofed if the measurements are performed by an untrustworthy source, they may be useful for verification in a laboratory setting where measurements are trusted.

Gibbs Sampling as a Test for Quantum Advantage We briefly note here that the algorithms considered in this work for Gibbs state preparation are well beyond the power of NISQ devices, and likely require a full fault tolerant quantum computer to implement. As such, we believe that Gibbs sampling as a test of quantum advantage may be more appropriate for analogue or dissipative models of computation which naturally approach a Gibbs state in certain conditions. Alternatively, we suggest that, for the specific case of IQP parent Hamiltonians, there may be much simpler algorithms which guarantee convergence to the Gibbs state. Furthermore, because Gibbs states are fixed points of Lindbladians, we might expect a certain level of robustness to external noise. For example, one might hope that Trotterizing a parent Hamiltonian and running its time-evolution on a weakly-noisy device may result in the state converging to the Gibbs state.

6 Discussion and Future Work

In this paper we have constructed two families of $O(1)$-local Hamiltonians for which their corresponding Gibbs states are classically hard to sample at $O(1)$ temperatures. Furthermore, these Gibbs states can be efficiently prepared and sampled from using a quantum computer. By showing that our hardness-of-sampling results hold for $\beta = \Omega(1/\sqrt{\Delta})$, we have placed bounds on where classical sampling algorithms can be efficient. Finally, we have suggested some heuristic routes for verifying quantum advantage via Gibbs sampling.

Beyond the work studied here, there exist many potential future routes for improvement.

1. Although the Hamiltonians here have constant locality, they do not necessarily correspond to Hamiltonians seen in nature, or to distributions we are interested in sampling from in computer science. A natural question to ask it whether sampling from Gibbs states of geometrically local Hamiltonians is still hard at constant temperature, even considering hardness with additional constraints such as translationally invariant terms or certain symmetries or restricted families of interaction terms. Remarkably, in the case of ground states, predicting local observables is known to be hard on 2D lattice and 1D translationally invariant chains [52, 53, 54]. While the properties of Gibbs states in 1D are well understood to be easy-to-sample, for 2D and beyond is not.

2. We also note that the parent Hamiltonian here has same partition function as a non-interacting Hamiltonian. Since non-interacting Hamiltonians do not undergo thermally-driven phase transitions, the parent Hamiltonian here does not either. Yet there is an apparent transition in temperature where the system moves from easy-to-sample to hard-to-sample. It would be interesting to understand if this transition coincides with other physical changes in the system, or if there is some way of characterising the change in sampleability in terms of other physics properties such as entanglement, entropies, etc.

3. The bounds on efficient sampling of Gibbs states by a classical computer leaves a gap open relative to the boundary showing classical easiness of sampling in Bakshi et al. [24], i.e. $O(\sqrt{\Delta})$ vs $\Omega(\Delta)$. There is future work to be done to understand where exactly this boundary lies.

4. Recently there has been much discussion of learning Hamiltonians with sample access to Gibbs states [55, 56, 57, 58, 59, 60, 61, 24]. Universally, the sample complexity of these algorithms increases as the temperature drops. However, it is notable that some of these algorithms are only efficient in the high temperature regime, and their costs blows up exponentially past some critical threshold. It would be desirable to understand how these thresholds relate the onset of sampling complexity (if there is any relation at all).

5. As far as we are aware, classical hardness-of-sampling results for noisy IQP circuits from Ref. [38] are only known for sampling to exponentially high precision in total variation distance (i.e.

multiplicative error). It remains an open question whether the output distributions of these noisy circuits can be shown to anti-concentrate. If they do anti-concentrate, then this would imply hardness of classical sampling to $1/\operatorname{poly}(n)$ error in total variation distance, which then implies hardness of sampling Gibbs states with $O(1/\operatorname{poly}(n))$ error for the circuits in Section 3.2.

Acknowledgements

The authors thank Atul Mantri and Dominik Hangleiter for useful discussions about IQP circuits. The authors also thank Yi-Kai Liu and Anirban Chowdhury for useful feedback and discussions, as well as Ewin Tang for correcting some of our errors in the initial manuscript.

JR acknowledges support from the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1840340. JDW acknowledges support from the United States Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Accelerated Research in Quantum Computing program, and also NSF QLCI grant OMA-2120757.

Author Contribution Statement

Both authors contributed equally and are listed alphabetically.

References

[1] Thiago Bergamaschi, Chi-Fang Chen, and Yunchao Liu. "Quantum computational advantage with constant-temperature Gibbs sampling". *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*. 2024. Pp. 1063–1085.

[2] Kristan Temme, Tobias J Osborne, Karl G Vollbrecht, David Poulin, and Frank Verstraete. "Quantum metropolis sampling". In: *Nature* 471.7336 (2011), pp. 87–90.

[3] Fernando GSL Brandao and Michael J Kastoryano. "Finite correlation length implies efficient preparation of quantum thermal states". In: *Communications in Mathematical Physics* 365 (2019), pp. 1–16.

[4] Jonathan E Moussa. "Low-depth quantum metropolis algorithm". In: *arXiv preprint arXiv:1903.01451* (2019).

[5] Mario Motta et al. "Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution". In: *Nature Physics* 16.2 (2020), pp. 205–210.

[6] Evgeny Mozgunov and Daniel Lidar. "Completely positive master equation for arbitrary driving and small level spacing". In: *Quantum* 4 (2020), p. 227.

[7] Chi-Fang Chen and Fernando GSL Brandão. "Fast thermalization from the eigenstate thermalization hypothesis". In: *arXiv preprint arXiv:2112.07646* (2021).

[8] Pawel Wocjan and Kristan Temme. "Szegedy walk unitaries for quantum maps". In: *Communications in Mathematical Physics* 402.3 (2023), pp. 3201–3231.

[9] Oles Shtanko and Ramis Movassagh. "Preparing thermal states on noiseless and noisy programmable quantum processors". In: *arXiv e-prints* (2021), arXiv–2112.

[10] Daniel Zhang, Jan Lukas Bosse, and Toby Cubitt. "Dissipative quantum Gibbs sampling". In: *arXiv preprint arXiv:2304.04526* (2023).

[11] Chi-Fang Chen, Michael J Kastoryano, Fernando GSL Brandão, and András Gilyén. "Quantum thermal state preparation". In: *arXiv preprint arXiv:2303.18224* (2023).

[12] Jiaqing Jiang and Sandy Irani. "Quantum Metropolis Sampling via Weak Measurement". In: *arXiv preprint arXiv:2406.16023* (2024).

[13] Hongrui Chen, Bowen Li, Jianfeng Lu, and Lexing Ying. "A randomized method for simulating Lindblad equations and thermal state preparation". In: *Quantum* 9 (2025), p. 1917.

[14] Sirui Lu, Mari Carmen Bañuls, and J Ignacio Cirac. "Algorithms for quantum simulation at finite energies". In: *PRX Quantum* 2.2 (2021), p. 020321.

[15] Alexander Schuckert, Annabelle Bohrdt, Eleanor Crane, and Michael Knap. "Probing finite-temperature observables in quantum simulators of spin systems with short-time dynamics". In: *Physical Review B* 107.14 (2023), p. L140410.

[16] Khaldoon Ghanem, Alexander Schuckert, and Henrik Dreyer. "Robust extraction of thermal observables from state sampling and real-time dynamics on quantum computers". In: *Quantum* 7 (2023), p. 1163.

[17] Chi-Fang Chen, Michael J Kastoryano, and András Gilyén. "An efficient and exact noncommutative quantum gibbs sampler". In: *arXiv preprint arXiv:2311.09207* (2023).

[18] András Gilyén, Chi-Fang Chen, Joao F Doriguello, and Michael J Kastoryano. "Quantum generalizations of Glauber and Metropolis dynamics". In: *arXiv preprint arXiv:2405.20322* (2024).

[19] Daniel Stilck França. "Perfect sampling for quantum gibbs states". In: *Quantum Info. Comput.* 18.5–6 (May 2018), 361–388.

[20] Aram W Harrow, Saeed Mehraban, and Mehdi Soleimanifar. "Classical algorithms, correlation decay, and complex zeros of partition functions of quantum many-body systems". *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*. 2020. Pp. 378–386.

[21] Ryan L Mann and Tyler Helmuth. "Efficient algorithms for approximating quantum partition functions". In: *Journal of Mathematical Physics* 62.2 (2021).

[22] Chao Yin and Andrew Lucas. "Polynomial-time classical sampling of high-temperature quantum Gibbs states". In: *arXiv preprint arXiv:2305.18514* (2023).

[23] Cambyse Rouzé, Daniel Stilck França, and Álvaro M Alhambra. "Efficient thermalization and universal quantum computing with quantum Gibbs samplers". *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*. 2025. Pp. 1488–1495.

[24] Ainesh Bakshi, Allen Liu, Ankur Moitra, and Ewin Tang. "High-temperature Gibbs states are unentangled and efficiently preparable". *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*. 2024. Pp. 1027–1036.

[25] Allan Sly. "Computational transition at the uniqueness threshold". *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*. 2010. Pp. 287–296.

[26] Isaac J Crosson, Dave Bacon, and Kenneth R Brown. "Making classical ground-state spin computing fault-tolerant". In: *Physical Review E* 82.3 (2010), p. 031106.

[27] Allan Sly and Nike Sun. "The computational hardness of counting in two-spin models on d-regular graphs". *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*. 2012. Pp. 361–369.

[28] Sergey Bravyi, Anirban Chowdhury, David Gosset, and Pawel Wocjan. "Quantum Hamiltonian complexity in thermal equilibrium". In: *Nature Physics* 18.11 (2022), pp. 1367–1370.

[29] Sergey Bravyi, Anirban Chowdhury, David Gosset, Vojtěch Havlíček, and Guanyu Zhu. "Quantum complexity of the Kronecker coefficients". In: *PRX Quantum* 5.1 (2024), p. 010329.

[30] Fernando GSL Brandão. "Entanglement theory and the quantum simulation of many-body physics". In: *arXiv preprint arXiv:0810.0026* (2008).

[31] David Gamarnik, Bobak T. Kiani, and Alexander Zlokapa. "Slow Mixing of Quantum Gibbs Samplers". In: *arXiv e-prints*, arXiv:2411.04300 (Nov. 2024), arXiv:2411.04300. arXiv: 2411.04300 [quant-ph].

[32] Joao Basso, Chi-Fang Chen, and Alexander M. Dalzell. "Optimizing random local Hamiltonians by dissipation". In: *arXiv e-prints*, arXiv:2411.02578 (Nov. 2024), arXiv:2411.02578. arXiv: 2411.02578 [quant-ph].

[33] Chi-Fang Chen, Hsin-Yuan Huang, John Preskill, and Leo Zhou. "Local minima in quantum systems". *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*. 2024. Pp. 1323–1330.

[34] Štěpán Šmíd, Richard Meister, Mario Berta, and Roberto Bondesan. "Rapid Mixing of Quantum Gibbs Samplers for Weakly-Interacting Quantum Systems". In: *arXiv e-prints*, arXiv:2510.04954 (Oct. 2025), arXiv:2510.04954. arXiv: 2510.04954 [quant-ph].

[35] Thiago Bergamaschi and Chi-Fang Chen. "Quantum Spin Chains Thermalize at All Temperatures". In: *arXiv e-prints*, arXiv:2510.08533 (Oct. 2025), arXiv:2510.08533. arXiv: 2510.08533 [quant-ph].

[36] Andrew M Childs, Tongyang Li, Jin-Peng Liu, Chunhao Wang, and Ruizhe Zhang. "Quantum algorithms for sampling log-concave distributions and estimating normalizing constants". In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 23205–23217.

[37] Shouvanik Chakrabarti et al. "Quantum algorithm for estimating volumes of convex bodies". In: *ACM Transactions on Quantum Computing* 4.3 (2023), pp. 1–60.

[38] Keisuke Fujii and Shuhei Tamate. "Computational quantum-classical boundary of noisy commuting quantum circuits". In: *Scientific Reports* 6.1 (May 2016), p. 25598.

[39] Michael J. Bremner, Ashley Montanaro, and Dan J. Shepherd. "Achieving quantum supremacy with sparse and noisy commuting quantum computations". In: *Quantum* 1 (Apr. 2017), p. 8.

[40] Dominik Hangleiter, Juan Bermejo-Vega, Martin Schwarz, and Jens Eisert. "Anticoncentration theorems for schemes showing a quantum speedup". In: *Quantum* 2 (May 2018), p. 65.

[41] Michael J Bremner, Ashley Montanaro, and Dan J Shepherd. "Average-case complexity versus approximate simulation of commuting quantum computations". In: *Physical review letters* 117.8 (2016), p. 080501.

[42] Joel Rajakumar, James D Watson, and Yi-Kai Liu. "Polynomial-time classical simulation of noisy iqp circuits with constant depth". *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2025. Pp. 1037–1056.

[43] Keisuke Fujii. *Quantum Computation with Topological Codes: from qubit to topological fault-tolerance*. Vol. 8. Springer, 2015.

[44] Scott Aaronson. "Quantum computing, postselection, and probabilistic polynomial-time". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 461.2063 (2005), pp. 3473–3482.

[45] Dominik Hangleiter, Martin Kliesch, Jens Eisert, and Christian Gogolin. "Sample complexity of device-independently certified "quantum supremacy"". In: *Physical review letters* 122.21 (2019), p. 210502.

[46] Sergio Boixo et al. "Characterizing quantum supremacy in near-term devices". In: *Nature Physics* 14.6 (2018), pp. 595–600.

[47] Dominik Hangleiter et al. "Fault-Tolerant Compiling of Classically Hard Instantaneous Quantum Polynomial Circuits on Hypercubes". In: *PRX Quantum* 6.2 (2025), p. 020338.

[48] Ainesh Bakshi, Allen Liu, Ankur Moitra, and Ewin Tang. "Learning quantum Hamiltonians at any temperature in polynomial time". *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*. 2024. Pp. 1470–1477.

[49] Fernando GSL Brandao and Krysta M Svore. "Quantum speed-ups for solving semidefinite programs". *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*. 2017. Pp. 415–426.

[50] Marc Vuffray, Sidhant Misra, Andrey Lokhov, and Michael Chertkov. "Interaction screening: Efficient and sample-optimal learning of Ising models". In: *Advances in neural information processing systems* 29 (2016).

[51] Adam Klivans and Raghu Meka. "Learning graphical models using multiplicative weights". *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*. 2017. Pp. 343–354.

[52] Daniel Gottesman and Sandy Irani. "The quantum and classical complexity of translationally invariant tiling and Hamiltonian problems". *2009 50th Annual IEEE Symposium on Foundations of Computer Science*. 2009. Pp. 95–104.

[53] Sevag Gharibian, Stephen Piddock, and Justin Yirka. "Oracle Complexity Classes and Local Measurements on Physical Hamiltonians". *37th International Symposium on Theoretical Aspects of Computer Science (STACS 2020)*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik2020. 20:1–20:37.

[54] James D. Watson, Johannes Bausch, and Sevag Gharibian. "The Complexity of Translationally Invariant Problems Beyond Ground State Energies". *40th International Symposium on Theoretical Aspects of Computer Science (STACS 2023)*. 2023. 54:1–54:21.

[55] Anurag Anshu, Srinivasan Arunachalam, Tomotaka Kuwahara, and Mehdi Soleimanifar. "Sample-efficient learning of interacting quantum systems". In: *Nature Physics* 17.8 (2021), pp. 931–935.

[56] Anurag Anshu, Srinivasan Arunachalam, Tomotaka Kuwahara, and Mehdi Soleimanifar. "Efficient learning of commuting Hamiltonians on lattices". In: *Electronic notes* (2021).

[57] Jeongwan Haah, Robin Kothari, and Ewin Tang. "Optimal learning of quantum Hamiltonians from high-temperature Gibbs states". *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*. 2022. Pp. 135–146.

[58] Álvaro M Alhambra. "Quantum many-body systems in thermal equilibrium". In: *PRX Quantum* 4.4 (2023), p. 040201.

[59] Cambyse Rouzé, Daniel Stilck França, Emilio Onorati, and James D Watson. "Efficient learning of ground and thermal states within phases of matter". In: *Nature Communications* 15.1 (2024), p. 7755.

[60] Cambyse Rouzé and Daniel Stilck França. "Learning quantum many-body systems from a few copies". In: *Quantum* 8 (2024), p. 1319.

[61] Luis Pedro García-Pintos et al. "Estimation of Hamiltonian parameters from thermal states". In: *Physical Review Letters* 133.4 (2024), p. 040802.

[62] Dominik Hangleiter and Jens Eisert. "Computational advantage of quantum random sampling". In: *Rev. Mod. Phys.* 95 (3 July 2023), p. 035001.

[63] Dan Shepherd and Michael J. Bremner. "Temporally unstructured quantum computation". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 465.2105 (2009), pp. 1413–1439. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.2008.0443.

[64] Dan Shepherd. *Binary Matroids and Quantum Probability Distributions*. 2010. arXiv: 1005.1744 [cs.CC].

[65] Daniel James Shepherd. "Quantum Complexity: restrictions on algorithms and architectures". In: *arXiv preprint arXiv:1005.1425* (2010).

[66] David Gross and Dominik Hangleiter. *Secret extraction attacks against obfuscated IQP circuits*. 2023. arXiv: 2312.10156 [quant-ph].

Appendix

A Hardness of Noisy IQP Sampling to Inverse Exponential Approximation

We make modifications to the proof from Ref. [38] to show that noisy IQP circuits are hard to sample with, even with inverse exponential error. First, we introduce the following lemma (derived in a different form in Ref. [38]) which allows us to bound the probability that postselected probabilities are close.

Lemma 13. *Let P and P' be any two probability distributions over bitstrings of length n. Suppose further that the bitstrings start with a decision bit x, and are followed by a 'postselection' register y of $n-1$ bits. For any $0 < \delta < 1/2$, if $\|P' - P\|_1 < \frac{\delta}{2+\delta} P(y=0)$, then*

$$|P'(x|y=0) - P(x|y=0)| < \delta \tag{4}$$

Proof. We can lower bound $P'(y=0)$ as

$$P'(y=0) > P(y=0) - \|P' - P\|_1 > P(y=0)\frac{2}{2+\delta} \tag{5}$$

Then, we have

$$|P'(x|y=0) - P(x|y=0)| = \left|\frac{P'(x,y=0)}{P'(y=0)} - \frac{P(x,y=0)}{P(y=0)}\right| \tag{6}$$

$$\leq \left|\frac{P'(x,y=0)}{P'(y=0)} - \frac{P(x,y=0)}{P'(y=0)}\right| + \left|\frac{P(x,y=0)}{P'(y=0)} - \frac{P(x,y=0)}{P(y=0)}\right| \tag{7}$$

$$\leq \frac{1}{P'(y=0)}\left|P'(x,y=0) - P(x,y=0)\right| + P(x,y=0)\left|\frac{1}{P'(y=0)} - \frac{1}{P(y=0)}\right| \tag{8}$$

$$\leq \frac{\|P'-P\|_1}{P'(y=0)} + P(x,y=0)\left|\frac{P(y=0)-P'(y=0)}{P'(y=0)P(y=0)}\right| \tag{9}$$

$$\leq \frac{\|P'-P\|_1}{P'(y=0)}\left(1 + \frac{P(x,y=0)}{P(y=0)}\right) \tag{10}$$

$$\leq \frac{2\|P'-P\|_1}{P'(y=0)} \tag{11}$$

$$\leq \frac{2\delta}{(2+\delta)}\frac{P(y=0)}{P'(y=0)} \tag{12}$$

$$< \delta \tag{13}$$

Where we have used the fact that $P(x, y = 0) \leq P(y = 0)$, and substituted expressions for $\|P' - P\|_1$ and $P'(y = 0)$ in terms of $P(y = 0)$ in the final steps. □

Definition 14. *(PostBQP) A language L is in the class PostBQP iff there exists a uniform family of postselected circuits $\{C_w\}$ with a decision port x and a postselection port y, and*

$$\text{if } w \in L, P_{C_w}(x=1|y=0) \geq 1/2 + \delta \tag{14}$$

$$\text{if } w \notin L, P_{C_w}(x=1|y=0) \leq 1/2 - \delta \tag{15}$$

where δ can be chosen arbitrary such that $0 < \delta < 1/2$. Further, the class PostBQP* is defined with the added restriction that $P_{C_w}(y=0) > 2^{-6m-4}$, where m is the number of qubits, and [38] show that PostBQP* = PostBQP = PP.

Lemma 15. *(Restatement of Lemma 5) There exists a family of IQP circuits $\mathcal{C}$, constructed on a 3D cubic lattice, consisting of a single layer of $e^{iZ\pi/8}$ gates and 4 layers of nearest-neighbour $e^{iZZ\pi/4}$ gates, such that sampling from any probability distribution Q over bitstrings which satisfies $\|P_{C,q} - Q\|_1 < (1-q)^n 2^{-6n-4}/5$ is not possible with a classical polynomial algorithm, unless the polynomial hierarchy collapses to the third level, when $q < 0.134$.*

Proof. First, we will outline the proof in [38] for review, then we will make our modifications. As in [38], by topologically protected MBQC, any quantum circuit can be encoded into a fault-tolerant IQP circuit of the form mentioned in Lemma 15, such that postselection recovers the original output distribution. Explicitly, for any quantum circuit C_w on m qubits and error parameter ϵ, one can construct C on $n = \text{poly}(m, \log(1/\epsilon))$ qubits, such that

$$\|P_{C,q|y'=0} - P_{C_w}\|_1 \leq \epsilon \tag{16}$$

where y' is an error-detection register. This is essentially a variant of the threshold theorem, for the case of postselected error detection, rather than error correction (instead of doing fault-tolerant MBQC which requires adaptive measurement, we post-select on outcomes corresponding to 'no error-detected').

Suppose we choose some uniform family of quantum circuits C_w which decide some PP-complete language L with confidence gap $0 < \delta < 1/2$, with decision port x and postselection port y where $P_{C_w}(y = 0) > 2^{-6m-4}$. For some $0 < \delta' < \delta$, suppose we encode this circuit into an IQP circuit C using the fault-tolerance construction mentioned above with polynomial overhead, where $\epsilon = 2^{-6m-4}\frac{\delta'}{2+\delta'}$. By Lemma 13 and using the fact that $P_{C_w}(y = 0) > 2^{-6m-4}$, this means that $|P_{C,q}(x|y = 0, y' = 0) - P_{C_w}(x|y = 0)| < \delta'$. Clearly, post-selection on JW: y, y' registers for $P_{C,q}$ can also decide L, with confidence gap $\delta - \delta'$. There are standard complexity-theoretic reductions to show hardness of *exactly* sampling from probability distributions that decide PP-complete problems under postselection (see [62] for review). This concludes the outline of the proof from [38].

Now, note that $P_{C,q}(y' = 0) \geq (1 - q)^n$ because $y' = 0$ in $P_{C,q}$ when no bit-flip error occurs. This means that

$$\begin{aligned} P_{C,q}(y = 0, y' = 0) &= P_{C,q}(y = 0|y' = 0)P_{C,q}(y' = 0) && (17)\\ &\geq P_{C,q}(y = 0|y' = 0)(1 - q)^n && (18)\\ &\geq (P_{C_w}(y = 0) - \epsilon)(1 - q)^n && (19)\\ &> 2^{-6m-4}\left(1 - \frac{\delta'}{2 + \delta'}\right)(1 - q)^n && (20)\\ &> 2^{-6m-4}\frac{2}{2 + \delta'}(1 - q)^n && (21) \end{aligned}$$

For some $0 < \delta'' < \delta'$, suppose we consider some distribution Q such that $\|P_{C,q} - Q\|_1 \leq (1 - q)^n 2^{-6m-4}\frac{2}{2+\delta'}\frac{\delta''}{2+\delta''}$. Using the triangle inequality and two applications of Lemma 13, we have that

$$\begin{aligned} |Q(x|y = 0, y' = 0) - P_{C_w}(x|y = 0)| &\leq |Q(x|y = 0, y' = 0) - P_{C,q}(x|y = 0, y' = 0)| && (22)\\ &\quad + |P_{C,q}(x|y = 0, y' = 0) - P_{C_w}(x|y = 0)| && (23)\\ &< \delta'' + \delta' && (24) \end{aligned}$$

Thus, choosing δ'' to be close to $1/2$ and δ' to be close to 0, one can solve PP-complete problems by postselecting on Q whenever $\|P_{C,q} - Q\|_1 \leq (1 - q)^n 2^{-6n-4}/5$, which implies that sampling from Q is hard (assuming PH) □

B Equivalence of Gibbs States and Circuits with Input Noise

Here we provide a proof of Lemma 3 in the main text.

Lemma 16. *For any circuit C constructed from k-local gates of depth d, there exists a $O(kd)$-local parent Hamiltonian H_C, such that:*

$$\frac{1}{Z}e^{-\beta H_C} = C(\mathcal{D}_q^{\otimes n}(|0\rangle\langle 0|))C^\dagger$$

for $q = \frac{e^{-\beta}}{1+e^{-\beta}}$.

Proof. Let C denote an arbitrary quantum circuit. Define a corresponding parent Hamiltonian $H_C = CH_{\mathrm{NI}}C^\dagger$ with corresponding Gibbs state:

$$\begin{aligned} \frac{1}{Z}e^{-\beta H_C} &= \frac{1}{Z}e^{-\beta\sum_x \lambda_x C|x\rangle\langle x|C^\dagger} && (25)\\ &= \frac{1}{Z}\sum_{x\in\{0,1\}^n} e^{-\beta\,\mathrm{HW}(x)} C\,|x\rangle\langle x|\,C^\dagger && (26) \end{aligned}$$

Note that because unitary transformations preserve the spectrum, the partition functions of H_{NI} and H_C are the same. Let $|x|$ be hamming weight of bitstring $x \in \{0,1\}^n$. We can then compare this to the output state of C with a single set of bit-flip noise of strength q occurring on the input.

$$C(\mathcal{D}_q^{\otimes n}(|0\rangle\langle 0|))C^\dagger = C\left(\sum_{x\in\{0,1\}^n} q^{\mathrm{HW}(x)}(1-q)^{n-\mathrm{HW}(x)}\,|x\rangle\langle x|^{\otimes n}\right)C^\dagger \tag{27}$$

$$= \sum_{x\in\{0,1\}^n}\left(\frac{q}{1-q}\right)^{\mathrm{HW}(x)}(1-q)^n C\,|x\rangle\langle x|^{\otimes n}\,C^\dagger \tag{28}$$

Comparing Eq. (26) to Eq. (28), and noting that $Z = (1+e^{-\beta})^n$, we can see that if we set $\frac{q}{1-q} = e^{-\beta}$ and $Z = (1+e^{-\beta})^n = (1+\frac{q}{1-q})^n = 1/(1-q)^n$, then these two states are the same. From here it is straightforward to see that the distribution $P(s) \coloneqq \langle s|\frac{e^{-\beta H_C}}{Z}|s\rangle$ satisfies:

$$P(s) = P_{C,q}(s)$$

for $q = \frac{e^{-\beta}}{1+e^{-\beta}}$. □

C CNOT Encoding

We use a result from Bremner et al. [39] which shows that IQP circuits can be made robust to bit-flip noise on the input distribution, at the cost of non-locality in the gate set. The fundamental idea is to encode a repetition code into a given hard-to-sample IQP circuit, and then decode with high probability.

We provide a brief explanation of the encoding as follows. Suppose we wish to encode some logical IQP circuit C into a physical circuit C_r. For each qubit i involved in C (and initialised in the $|0\rangle$ state), prepare a block of r qubits labelled $i_1,\ldots,i_r$ in C_r, which are initialised to the $|0\rangle$ state. Now, every 2-qubit diagonal gate between qubits i and j in C can be written in the form $D = e^{i\theta_1 Z_i + \theta_2 Z_j + \theta_3 Z_i Z_j}$ (modulo global phase), for some $\theta_1,\theta_2,\theta_3$. This notation is known as an 'X-program' [63, 64, 65, 39, 66]. The encoded circuit C_r then involves the following transformation for each 2-qubit diagonal gate

$$D = e^{i\theta_1 Z_i+\theta_2 Z_j+\theta_3 Z_i Z_j} \to D_{enc} = e^{i\theta_1 Z_{i_1}Z_{i_2}\ldots Z_{i_r}+\theta_2 Z_{j_1}Z_{j_2}\ldots Z_{j_r}+\theta_3 Z_{i_1}Z_{j_1}Z_{i_2}Z_{j_2}\ldots Z_{i_r}Z_{j_r}} \tag{29}$$

In Bremner et al., it is shown that this construction essentially encodes each logical output bit into r physical output bits, each with an independent probability q of being incorrect. Therefore, the logical output distribution can be recovered by taking a majority vote. This is captured in the following lemma.

Lemma 17. *(From [39]) Let C be an arbitrary IQP circuit constructed with 2-qubit gates of depth d on n qubits. Then, for any integer parameter $r \geq 1$, there is an encoded IQP circuit C_{enc} constructed with $2r$-local gates of depth d on nr qubits, and a decoding algorithm A such that,*

$$A(P_{C_{enc},q}) = P_{C,p_{fail}} \tag{30}$$

where $p_{fail} \leq (4q(1-q))^{r/2}$.

It can be seen that the parent Hamiltonian of the above construction $H_{C_{enc}}$ will be $O(r)$-local. Now, we show that this encoding can be achieved without non-local gates by using a networks of $CNOT$s

Lemma 18. *Suppose IQP circuit C on n qubits is encoded into an encoded circuit C_{enc} on nr qubits as per Lemma 17. For each logical qubit i of C, let the block of r physical qubits be labelled*

$i_1, \ldots, i_r$. *Let* C_1 *denote the circuit* C *applied transversally*[3] *to the first qubits of each block* $(1_1, 2_1, \ldots, n_1)$. *Let the unitary* B *be composed of the following CNOT network,*

$$B = \prod_{i \in 1,\ldots,n} \prod_{j \in 2,\ldots,r} CNOT_{i_1,i_j} \tag{31}$$

where $CNOT_{x,y}$ *is a CNOT controlled on qubit* x *and targeting qubit* y. *Then,*

$$B^\dagger C_1 B = C_{enc} \tag{32}$$

Proof. Suppose C is given by the diagonal gates $D_1, \ldots, D_m$, so that $C = H^{\otimes nr} \prod_{l \in 1,\ldots,m} D_l H^{\otimes nr}$. Suppose C_{enc} is given by the diagonal gates $D'_1, \ldots, D'_m$, so that the unitary produced by C_r is $C_{enc} = H^{\otimes nr} \prod_{l \in 1,\ldots,m} D'_l H^{\otimes nr}$.

Define $B_H = H^{\otimes n} B H^{\otimes n}$. By a circuit identity, conjugating a CNOT with Hadamards on either qubit reverses the target and the control (i.e. $(H_x \otimes H_y) CNOT_{x,y} (H_x \otimes H_y) = CNOT_{y,x}$). Therefore,

$$B_H = \prod_{i \in 1,\ldots,n} \prod_{j \in 2,\ldots,r} CNOT_{i_j,i_1} \tag{33}$$

Due to another circuit identity that $CNOT_{x,y}(\mathbb{1}_x \otimes Z_y) CNOT_{x,y} = Z_x \otimes Z_y$, we can establish that B_H has the following property,

$$B_H^\dagger Z_{i_1} B_H = \prod_{j=1,\ldots,b} Z_{i_j} \tag{34}$$

Suppose we use $D_{l,1}$ to denote gate D_l applied transversely to the first qubits of each block $(1_1, 2_1, \ldots, nr_1)$. By examining Eq. (29) and comparing it the above, one can see that

$$B_H^\dagger D_{l,1} B_H = D'_1 \tag{35}$$

Therefore, we have,

$$\begin{aligned}
B^\dagger C_1 B &= B^\dagger \bigotimes_{i \in 1,\ldots,n} H_{i_1} \prod_{l \in 1,\ldots,m} D_{l,1} \bigotimes_{i \in 1,\ldots,n} H_{i_1} B && (36)\\
&= B^\dagger H^{\otimes nr} \prod_{l \in 1,\ldots,m} D_{l,1} H^{\otimes nr} B && (37)\\
&= H^{\otimes nr} B_H^\dagger \prod_{l \in 1,\ldots,m} D_{l,1} B_H H^{\otimes nr} && (38)\\
&= H^{\otimes nr} \left(\prod_{l \in 1,\ldots,m} B_H^\dagger D_{l,1} B_H \right) H^{\otimes nr} && (39)\\
&= H^{\otimes nr} D'_l H^{\otimes nr} && (40)\\
&= C_{enc} && (41)
\end{aligned}$$

where in the second step we have used the fact that $HH = I$ (so we can add a pair of Hadamards to every qubit without a Hadamard). □

[3] Transversally here means that for each gate between a pair of qubits (i, j), in the encoded circuit the same gate now acts between the pair of qubits (i_1, j_1).

Ref. [1] use a similar encoding scheme to ours. Specifically, they construct circuit $C' = B'C_1$, where B' is a different $CNOT$ network than B, but also acting amongst blocks of r qubits. By examining B', a simple adaptation of the above proof will show that $B'C_1B'^{\dagger} = C_{enc}$. This means that the encodings of Ref. [39], Ref.[1, Lemma 8.1], and Lemma 18 are all exactly equivalent up to CNOT networks on the output state. Since CNOTs commute with measurement, this means their output distributions are equivalent modulo post-processing. This suggests that further improvements to the noise-robustness of these constructions might require changes to the underlying error-correction code (the repetition code) rather than circuit-level optimisations. We also use the fact that CNOTs commute with measurement to establish the following lemma.

Lemma 19. *Let C be an arbitrary IQP circuit constructed with 2-qubit gates of depth d on n qubits. Then, for any integer parameter $r \geq 1$, there is an encoded IQP circuit C^* constructed with 2-qubit gates of depth $d + r$ on nr qubits, and a decoding algorithm A such that,*

$$A^*(P_{C^*,q}) = P_{C,p_{fail}} \tag{42}$$

where $p_{fail} \leq (4q(1-q))^{r/2}$. Furthermore, the parent Hamiltonian of C^* has locality $k \leq d+2$ and degree $\Delta \leq r(d+1)$

Proof. Using the notation of Lemma 18, define $C^* = C_1B$. $B^{\dagger}$ is composed of classical gates, which means it commutes with measurement. Therefore, we can apply it on the output distribution of C_1B, and then perform the decoding algorithm A of Lemma 17. That is,

$$A^*(P_{C^*,q}) = A(BP_{C^*,q}) = A(P_{BC^*,q}) = A(P_{C_r,q}) = P_{C,p_{\text{fail}}} \tag{43}$$

To calculate the locality, we use the following circuit identities.

$$CNOT_{x,y}(\mathbb{1}_x \otimes Z_y)CNOT_{x,y} = Z_x \otimes Z_y \tag{44}$$
$$CNOT_{x,y}(Z_x \otimes \mathbb{1}_y)CNOT_{x,y} = Z_x \otimes \mathbb{1}_y \tag{45}$$

We first conjugate H_{NI} with B

$$\begin{aligned}
&BH_{NI}B^{\dagger} \\
&= B \sum_{i\in[n],j\in[r]} \frac{\mathbb{1} - Z_{i_j}}{2} B^{\dagger} \\
&= \sum_{i\in[n]} B\frac{\mathbb{1} - Z_{i_1}}{2}B^{\dagger} + \sum_{i\in[n],j\in[2,r]} B\frac{\mathbb{1} - Z_{i_j}}{2}B^{\dagger} \\
&= \sum_{i\in[n]} \prod_{j\in[2,n]} CNOT_{i_1,i_j}\frac{\mathbb{1} - Z_{i_1}}{2}\prod_{j\in[2,n]} CNOT_{i_1,i_j} + \sum_{i\in[n],j\in[2,r]} CNOT_{i_1,i_j}\frac{\mathbb{1} - Z_{i_j}}{2}CNOT_{i_1,i_j} \\
&= \sum_{i\in[n]} \frac{\mathbb{1} - Z_{i_1}}{2} + \sum_{i\in[n],j\in[2,r]} \frac{\mathbb{1} - Z_{i_1}Z_{i_j}}{2}
\end{aligned}$$

Note that for each qubit i_1, there are at most d diagonal gates acting on qubit i_1 in C_1. Thus, when conjugating the $Z_{i_1}Z_{i_j}$ term with C_1, the resulting interaction term spans at most $d+2$ qubits, so $k \leq d+2$.

Each qubit i_1 appears in all $r-1$ terms of the form $C_1Z_{i_1}Z_{i_j}C_1^{\dagger}$. It also appears in all $r-1$ terms of the form $C_1Z_{i'_1}Z_{i'_j}C_1^{\dagger}$ for each of the d neighbours i'. If we count the Z_{i_1} terms as well, the total degree is $\Delta \leq (d+1)r$ □