Certifying and learning local quantum Hamiltonians

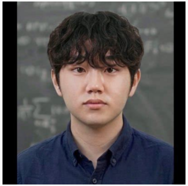
Andreas Bluhm
Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG

August 7, 2026, Foxconn Research Institute

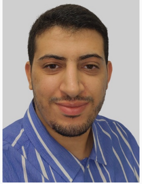
Collaborators

Matthias C. Caro
University of Warwick

Francisco Escudero Gutiérrez
Inria Saclay



Junseo Lee
Harvard University

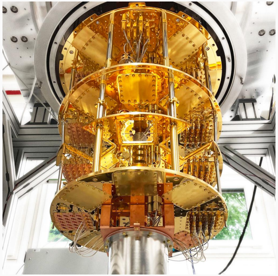


Aadil Oufkir
University Mohammed VI Polytechnic

Cambyse Rouzé
Inria Saclay, Télécom Paris, Institut Polytechnique

Myeongjin Shin
KAIST, Caltech

Why we want to learn and certify many-body systems



- Ever larger quantum devices are being built involving hundreds of qubits to serve as quantum computers and simulators
- Learning and testing of such systems needed to ensure that they function properly
- Learning of the dynamics needed, e.g., for error mitigation
- Certification needed to ensure that the device works as intended

Image taken from: https://funqs.de/

Access models and relevant figures of merit

Results typically depend on the access models. In this talk, we will consider two different ones:

- Access to dynamics
- Access to Gibbs (thermal equilibrium) states

Results often make assumptions about the quantum system at hand, such as locality, sparsity, low intersection, ...

Figures of merit:

- Number of queries (i.e., number of experiments)
- (expected) total evolution time
- classical post-processing time
- quantum complexity of input states and output measurements

Locality of Hamiltonians

A Hamiltonian is written as a sum of Pauli-string terms:

$$H = \sum_P h_P P, \qquad P \in \{I, X, Y, Z\}^{\otimes n}.$$

$Z_3 Z_7$**: 2-local**

A term is k-local if it touches at most k qubits.

$|P| = |\{i : P_i \neq I\}|$ counts touched qubits. For example, $|X_1 Y_4 Z_5| = 3$ and $|Z_3 Z_7| = 2$.

The Hamiltonian H is k-local if it has the form

$$H = \sum_{P \in \{I,X,Y,Z\}^{\otimes n} : |P| \leq k} h_P P,$$

i.e., only Pauli strings of weight at most k have non-zero coefficients.

Sparsity of Hamiltonians

Sparsity counts how many Pauli terms actually appear.

Definition

For an n-qubit Hamiltonian $H = \sum_{P \in \{I,X,Y,Z\}^{\otimes n}} h_P P$, its sparsity is the number s of coefficients with non-zero coefficients, i.e.,

$$s := |\{P : h_P \neq 0\}|$$

For example, $H = 2Z_1 - X_1 Z_2 + 3X_3$ has $s = 3$ active terms (but is 2-local).

Learning and certification in a nutshell

Learning

Dynamics:

$$U_H(t) = e^{-iHt} \quad \Longrightarrow \quad \widehat{H}.$$

Learns H by querying to e^{-iHt} for various t.

Gibbs:

$$\rho_H(\beta) \quad \Longrightarrow \quad \widehat{\rho}.$$

Learns directly the Gibbs state from access to copies.

Certification

Dynamics:

$$H_0 \text{ known}, \ U_H(t) \text{ queries } \Longrightarrow \text{CLOSE/FAR}.$$

Test whether H and H_0 are close or far.

Gibbs:

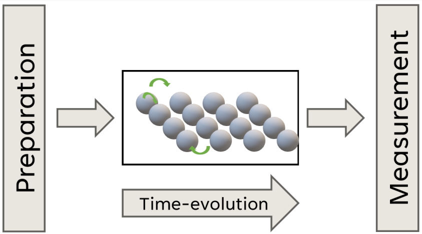
$$\rho_H(\beta), \ \rho_{H_0}(\beta) \Longrightarrow \text{CLOSE/FAR}.$$

Test whether thermal states are close or far.

Learning wants to figure out what an unknown device does, certification tests whether it behaves as expected.

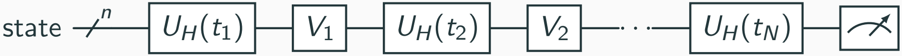
Certification from access to dynamics

Access to dynamics



- Algorithm: prepare suitable states, let them evolve under H for a time t, perform a measurement, classical post-processing of the results
- In other words, we can apply $U_H(t) := e^{-itH}$ for a time of our choosing

We assume that we have coherent access, i.e., we can run the time evolution for different times:



Total evolution time $T = \sum_{i=1}^{N} t_i$.

Hamiltonian Certification

Theorem (Optimal $O(1)$-local Hamiltonian Certification)

Let H, H_0 be n-qubit $k = O(1)$-local traceless Hamiltonians. Assume that H_0 is known in advance. Then, there is an algorithm with access to a classical description of H_0 and to the time evolution of H that uses $\Theta(1/\varepsilon)$ total evolution time to determine whether $\|H - H_0\|_{nF} \leq \varepsilon/(8 \cdot 3^k)$ or $\|H - H_0\|_{nF} \geq \varepsilon$ with high probability.

- Note that the lower bound $\Omega(1/\varepsilon)$ comes from distinguishing $H = \varepsilon X$ and $H_0 = -\varepsilon X$.
- For $H = \sum_{P \in \{I,X,Y,Z\}^{\otimes n}} h_P P$, we have

$$\|H\|_{nF} = \sqrt{\sum_P |h_P|^2}.$$

Comparison to previous work

Figure of merit: total time evolution

Task	Previous best		This work
Certification	$\widetilde{O}(s^{3k/2}/\varepsilon)$[1] or $O(1/\varepsilon^2)$[2]	$\Longrightarrow$	$O(c^k/\varepsilon)$ for some constant c

Main message: For constant locality $k = O(1)$, certification is optimal $\Theta(1/\varepsilon)$.

[1] M. Gao, Z. Ji, Q. Wang, W. Yu, and Q. Zhao. Quantum Hamiltonian certification. *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2026.

[2] S. D. Sinha and Y. Tong. Improved Hamiltonian learning and sparsity testing through Bell sampling. *arXiv:2509.07937*.

Proof sketch: Intolerant certification

Let

$$\Delta H = H - H_0.$$

If $H = H_0$, then $\Delta H = 0$. If H is far from H_0, then ΔH has nonzero energy differences.

Thus, our aim is to distinguish

$$\boxed{\Delta H = 0} \quad \text{vs.} \quad \boxed{\|H - H_0\|_{nF} \geq \varepsilon}$$

The theorem is actually tolerant, but this simpler promise already explains the idea.

Proof sketch: Trotterization

Our experiments will use the evolution of the difference Hamiltonian:

$$e^{-it\Delta H} = e^{-it(H-H_0)}.$$

But the unknown device only gives e^{-itH}. How do we get access to the time evolution under ΔH?

Product-formula idea

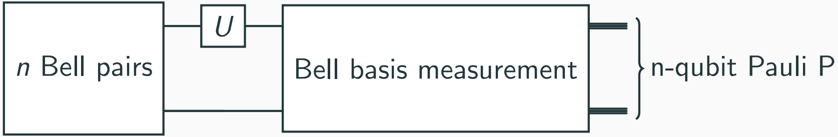
Use the known target H_0 as a control and interleave it with queries to the unknown evolution:

$$e^{-it(H-H_0)} \approx \left(e^{-iHt/(2r)} e^{+iH_0 t/r} e^{-iHt/(2r)}\right)^r.$$

Proof sketch: Bell sampling (1/2)

Conceptually, Bell sampling is Fourier sampling where the basis functions are Pauli operators.

$$\text{For } U = \sum_P u_P P \text{ want outcome } P \text{ with } \Pr[P] = |u_P|^2.$$



For certification, set $U = e^{-it\Delta H}$. The formal algorithm estimates the identity probability $|u_I|^2$ without quantum memory.

Proof sketch: Bell sampling (2/2)

If $H = H_0$, then $\Delta H = 0$ and

$$e^{-it\Delta H} = I.$$

Bell sampling always returns the identity Pauli string.

$H = H_0$	$\Longrightarrow$	identity outcome
H far from H_0	$\Longrightarrow$	non-identity appears

We need to estimate the probability of seeing a non-identity Pauli string

The certification algorithm

Repeat for $O(1)$ times:

1. Pick a random wait time $t \in [0, 2/\varepsilon]$.
2. Implement an approximation to $e^{-it(H-H_0)}$ using product formulas.
3. Estimate the identity probability $|(u_{\Delta H}(t))_I|^2$ of this difference evolution.
4. If the identity probability is noticeably below 1, output FAR and abort.

Output CLOSE.

Total evolution time:

$$T = O(9^{2k} \log(1/\delta)/\varepsilon);$$

for constant k, this is optimal $1/\varepsilon$ scaling.

Proof sketch: Analyzing Bell sampling probability (1/2)

We can query to the unitary $e^{-i\Delta H t}$. ($\Delta H = H - H_0$)

Lemma (Lemma[3])

The probability of Bell sampling I from $e^{-i\Delta H t}$ is represented as:

$$I(t) = \Pr[I] = \frac{\sum_{j,k} \cos[(\lambda_j - \lambda_k)t]}{4^n},$$

where $\lambda_1, \lambda_2, \ldots, \lambda_N$ are the eigenvalues of $\Delta H = H - H_0$.

[3] S. D. Sinha and Y. Tong. Improved Hamiltonian learning and sparsity testing through Bell sampling. *arXiv:2509.07937.*

Proof sketch: Analyzing Bell sampling probability (2/2)

Definition (The proportion of ε-far eigenvalue pairs)

For a Hermitian matrix $X \in \mathbb{C}^{N \times N}$, we have

$$\Lambda(X, \varepsilon) = \frac{1}{N^2} \sum_{|\lambda_j - \lambda_k| \geq \varepsilon} 1,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_N$ are the eigenvalues of X.

Lemma (Sufficient spectral condition)

Suppose there exists a constant $d > 0$ such that $\Lambda(H, \varepsilon) \geq d$. Then there is a randomized procedure that, using $O(\log(1/\delta))$ samples of $I(t)$ for uniformly random $t \in [0, 2/\varepsilon]$, finds a time t satisfying

$$I(t) \leq 1 - \frac{d}{4}$$

with probability at least $1 - \delta$.

Proof sketch: Remaining goal

In particular, suppose that for the case $\|\Delta H\|_{nF} \geq \varepsilon$, $\Lambda(\Delta H, \varepsilon)$ is always bounded below by a constant. Then, one can distinguish between $\|\Delta H\|_{nF} = 0$ and $\|\Delta H\|_{nF} \geq \varepsilon$ using total evolution time $O(1/\varepsilon)$.

Goal

Suppose that $\|\Delta H\|_{nF} \geq \varepsilon$. Let's prove that $\Lambda(\Delta H, \varepsilon) \geq d$ for some constant d.

Let $F(s,t) := \lambda_s - \lambda_t$. We realize that $\Lambda(\Delta H, \|\Delta H\|_{nF}) = \Pr_{t,s}[|F(s,t)| \geq \|\Delta H\|_{nF}]$. We therefore need a tail bound.

Paley-Zygmund inequality:

$$\Pr\left[F^2 \geq \tfrac{1}{2}\mathbb{E}F^2\right] \;\geq\; \frac{1}{4}\frac{(\mathbb{E}F^2)^2}{\mathbb{E}F^4}.$$

Proof sketch: Second and Fourth Moments

We can calculate

$$\mathbb{E}[F^2] = 2\|\Delta H\|_2^2.$$
$$\mathbb{E}[F^4] = 2\|\Delta H\|_4^4 + 6\|\Delta H\|_2^2.$$

Here, $\|\Delta H\|_p = (\frac{1}{2^n}\sum_s |\lambda_s|^p)^{1/p}$ (with this notation, $\|\Delta H\|_2 = \|\Delta H\|_{nF}$).

As ΔH is k-local, by hypercontractivity (quantum version of Bonami's lemma[4]):

$$\|\Delta H\|_4^4 \leq 9^k \|\Delta H\|_2^4.$$

Hence,

$$\|\Delta H\|_{nF} \geq \varepsilon \implies \Lambda(\Delta H, \varepsilon) = \Omega(9^{-k}).$$

Hence, $\Lambda(\Delta H, \varepsilon)$ is lower bounded by a constant for $k = O(1)$ and we are done.

[4] A. Montanaro and T. J. Osborne. Quantum Boolean functions. *Chicago Journal of Theoretical Computer Science*, pages 1–45, 2008.

Certification and learning from access to Gibbs states

Gibbs access model

In the Gibbs access model, we do not control time evolution. Instead, we receive many identical copies:

$$\rho_H(\beta) = \frac{e^{-\beta H}}{\mathrm{Tr}(e^{-\beta H})}.$$

$\boxed{\rho_H(\beta)}$ $\boxed{\rho_H(\beta)}$ $\boxed{\rho_H(\beta)}$ $\boxed{\rho_H(\beta)}$ $\boxed{\rho_H(\beta)}$

Each copy is measured once, usually with a simple Pauli measurement.

Here β is inverse temperature: larger β means lower temperature.

Gibbs access: cost

Sample cost

- $N =$ number of copies measured.
- Learning uses copies of $\rho_H(\beta)$.
- Certification uses copies of both $\rho_H(\beta)$ and $\rho_{H_0}(\beta)$.

Accuracy

For thermal states, the natural distance is trace norm:

$$\|\rho - \sigma\|_{\mathrm{tr}} = \mathrm{Tr}\,|\rho - \sigma|.$$

Comparison: Gibbs-state copies

Previous Gibbs-state methods usually tried the hard route:

$$\rho_H(\beta) \longrightarrow \widehat{H} \longrightarrow \rho_{\widehat{H}}(\beta),$$

and this route can be exponential in β.

Task	Previous route		This work
Learning	Exponential in β	$\Longrightarrow$	polynomial samples
Certification	Exponential in β	$\Longrightarrow$	polynomial samples and time

Learning: $\widetilde{O}(\text{poly}(n)\beta^2/\varepsilon^4)$ samples, certification: $\widetilde{O}(\text{poly}(n)\beta^2/\varepsilon^4)$ samples, poly time.

Hamiltonian learning from Gibbs states

Theorem

Let $\rho_H(\beta)$ be the Gibbs state of an unknown n-qubit local Hamiltonian H at temperature β with $|h_P| \leq 1$ for every P. Then, there is an algorithm that, with high probability, ε-learns $\rho_H(\beta)$ in trace norm using only $\widetilde{O}(\mathrm{poly}(n)\beta^2/\varepsilon^4)$ single copies of the state. However, our algorithm is not time efficient.

It is important to avoid learning the Hamiltonian, because learning has a $\Omega(e^{\beta})$ lower bound on the sample complexity[5].

[5] J. Haah, R. Kothari, and E. Tang. Optimal learning of quantum Hamiltonians from high-temperature Gibbs states. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 135–146. IEEE, 2022

Gibbs state learning strategy

Nearby local Hamiltonian coefficients give nearby Gibbs states:

$$\|\rho_H(\beta) - \rho_{H'}(\beta)\|_{\mathrm{tr}} \le \sqrt{2\beta\,\mathrm{Tr}[(\rho_H(\beta) - \rho_{H'}(\beta))(H - H')]} = O(\beta n^k \max_{P:|P|\le k} |h_P - h'_P|)$$

1. Build a finite candidate set of possible thermal states.
2. Measure many simple local Pauli measurements from copies of $\rho_H(\beta)$.
3. Use those measurement outcomes to choose or certify the matching thermal state.

$$\boxed{\text{copies}} \longrightarrow \boxed{\text{Pauli measurements}} \longrightarrow \boxed{\text{thermal state}}$$

Tool: classical shadows

Suppose we have known observables $O_1, \ldots, O_m$ and want all expectation values

$$\mu_i = \mathrm{Tr}(O_i \rho), \qquad i = 1, \ldots, m.$$

Classical shadows[6]

Randomized single-copy measurements let's one estimate many *simple* observables:

$$\widehat{\mu}_i \approx \mathrm{Tr}(O_i \rho) \quad \text{simultaneously for all } i.$$

by using $\mathcal{O}(\log m/\varepsilon^2)$ samples. *Simple* in this case includes k-local observables for constant k.

$$\text{copies of } \rho \quad \Longrightarrow \quad \text{shadow data} \quad \Longrightarrow \quad \text{all local Pauli statistics}$$

[6] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020.

Proof sketch (1/2)

Recall:

$$\|\rho_H(\beta) - \rho_{H'}(\beta)\|_{\mathrm{tr}} \leq \sqrt{2\beta\,\mathrm{Tr}[(\rho_H(\beta) - \rho_{H'}(\beta))(H - H')]} = O(\beta n^k \max_{P:|P|\leq k} |h_P - h'_P|)$$

Therefore, the set

$$\mathcal{S}_\eta = \{\rho_H(\beta) : \, H \in \mathcal{H}_\eta\}$$

of Gibbs states, where

$$\mathcal{H}_\eta = \{H : H \text{ } k\text{-local Hamiltonian with } h_P \in \eta\mathbb{Z} \cap [-1, 1] \;\forall P\},$$

is an ε-covering net of the set of k-local Gibbs states when taking η of the order $\varepsilon/(\beta n^k)$. The use of the ε-net is the reason why our algorithm is not efficient.

Proof sketch (2/2)

We note that for $H,\ H' \in \mathcal{H}_\eta$

$$H - H' = \sum_{P \in \{I,X,Y,Z\}^{\otimes n}:\, |P| \le k} \alpha_P P\,.$$

Hence, using classical shadows, we can simultaneously obtain accurate estimates

$$\Delta_{H,H'} \approx \mathrm{Tr}\big[\rho(H - H')\big]$$

in a sample-efficient manner, where ρ is the state to be learned. The rest of the proof is to show that

$$\hat{\rho} = \mathrm{argmin}_{\rho' \in \mathcal{S}_\eta} \max_{H,H' \in \mathcal{H}_\eta} \big|\Delta_{H,H'} - \mathrm{Tr}\big[\rho'(H - H')\big]\big|$$

satisfies $\|\rho - \hat{\rho}\|_{\mathrm{tr}} \le \varepsilon$ with high probability, where we make use of

$$\|\rho_H(\beta) - \rho_{H'}(\beta)\|_{\mathrm{tr}} \le \sqrt{2\beta\, \mathrm{Tr}[(\rho_H(\beta) - \rho_{H'}(\beta))(H - H')]}.$$

Certification of Gibbs states

Theorem

Let $\rho_H(\beta)$ *and* $\rho_{H_0}(\beta)$ *be the Gibbs states of an n-qubit local Hamiltonian* H *and* H_0 *at inverse temperature* β *with* $|h_P|, |(h_0)_P| \leq 1$ *for every* P*. Then, there is an algorithm that, with high probability, decides whether* $\rho_H(\beta) = \rho_{H_0}(\beta)$ *or* $\|\rho_H(\beta) - \rho_{H_0}(\beta)\|_{\mathrm{tr}} \geq \varepsilon$ *using only* $\widetilde{O}(\mathrm{poly}(n)\beta^2/\varepsilon^4)$ *single copies of the states.*

This algorithm is actually time efficient in every parameter. This answers an open question posed by Anshu (*Harvard Data Science Review*, 2022).

Proof sketch

Certification has a target Gibbs state $\rho_0 = \rho_{H_0}(\beta)$. We receive copies of $\rho = \rho_H(\beta)$ and ρ_0.

The key idea

Write Pauli coefficients as

$$\rho_P = \frac{1}{2^n} \operatorname{Tr}(P\rho).$$

If two bounded local Gibbs states are far, then some local Pauli coefficient must noticeably differ:

$$\|\rho - \rho_0\|_{\mathrm{tr}} \geq 2\varepsilon \quad \Longrightarrow \quad \exists P \text{ with } |P| \leq k \text{ such that } 2^n|\rho_P - (\rho_0)_P| \gtrsim \frac{\varepsilon^2}{\beta n^k}.$$

The latter follows from

$$\|\rho(\beta) - \rho'(\beta)\|_{\mathrm{tr}} \leq \sqrt{400\beta n^k \sup_{|P|\leq k} 2^n|\rho(\beta)_P - \rho'(\beta)_P|}.$$

Certification of Gibbs states: algorithm

1. Use classical shadows on copies of ρ.
2. Use the same kind of measurements on copies of ρ_0.
3. Estimate $2^n \rho_P$ and $2^n (\rho_0)_P$ for every $|P| \leq k$.
4. If some estimated coefficient gap exceeds $3\varepsilon^2/(400\beta n^k)$, output FAR.
5. Otherwise output CLOSE.

The certification theorem is tolerant. It distinguishes

$$\|\rho - \rho_0\|_{\mathrm{tr}} \leq \frac{\varepsilon^2}{400\beta n^k} \quad \text{versus} \quad \|\rho - \rho_0\|_{\mathrm{tr}} \geq 2\varepsilon.$$

Open questions

- Optimal certification from dynamics possible without locality of the Hamiltonian?
- Learning algorithm for Gibbs states that is both sample *and* time efficient?
- Is the algorithm for the certification of Gibbs states optimal?
- Optimal sparsity testing? Optimal locality testing?
- Generalization to dissipative dynamics?

Summary

- We have seen that certifying a local Hamiltonian in normalized Frobenius norm via access to its time-evolution can be achieved with only $O(1/\varepsilon)$ evolution time
- This is optimal, as it matches the Heisenberg-scaling lower bound of $\Omega(1/\varepsilon)$.
- We have designed an algorithm for learning Gibbs states of local Hamiltonians in trace norm that is sample-efficient in all relevant parameters.
- Previous approaches were exponential in the inverse temperature β
- We have given an algorithm for certifying Gibbs states of local Hamiltonians in trace norm that is both sample and time-efficient in all relevant parameters
- Certification can be much more efficient than Hamiltonian learning